Quantitative Finance

ISSN: 1469-7688 (Print) 1469-7696 (Online) Journal homepage: https://www.tandfonline.com/loi/rquf20

Numerical methods for dynamic Bertrand oligopoly and American options under regime switching

Swathi Amarala & Justin W. L. Wan

To cite this article: Swathi Amarala & Justin W. L. Wan (2016) Numerical methods for dynamic Bertrand oligopoly and American options under regime switching, Quantitative Finance, 16:11, 1741-1762, DOI: 10.1080/14697688.2016.1167281

To link to this article: https://doi.org/10.1080/14697688.2016.1167281

Published online: 28 Apr 2016.

Quantitative Finance, 2016
Vol. 16, No. 11, 1741–1762, http://dx.doi.org/10.1080/14697688.2016.1167281

Numerical methods for dynamic Bertrand oligopoly and American options under regime switching

SWATHI AMARALA* and JUSTIN W. L. WAN

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

(*Received 5 August 2015; accepted 3 March 2016; published online 28 April 2016*)

Bertrand oligopolies are competitive markets in which a small number of firms producing similar goods use price as their strategic variable. In particular, each firm wants to determine the optimal price that maximizes its expected discounted lifetime profit. The oligopoly problem can be modeled as nonzero-sum games which can be formulated as systems of Hamilton–Jacobi–Bellman (HJB) partial differential equations (PDEs). In this paper, we propose fully implicit, positive coefficient finite difference schemes that converge to the viscosity solution for the HJB PDE from dynamic Bertrand monopoly and the two-dimensional HJB system from dynamic Bertrand duopoly. Furthermore, we develop fast multigrid methods for solving these systems of discrete nonlinear HJB PDEs. The new multigrid methods are general and can be applied to other systems of HJB and HJB-Isaacs PDEs arising from American options under regime switching and American options with unequal lending/borrowing rates and stock borrowing fees under regime switching, respectively. We provide a theoretical analysis for the smoother, restriction and interpolation operators of the multigrid methods. Finally, we demonstrate the effectiveness of our method by numerical examples from the dynamic Bertrand problem and pricing American options under regime switching.

Keywords: Bertrand oligopoly; American options; Regime switching; Systems of HJB and HJBI PDEs; Monotone scheme; Multigrid methods; Jump in control

JEL Classification: JEL C

1. Introduction

There are different models describing the operation of an oligopolistic market. Cournot oligopolies are markets in which the firms compete by using quantity as their strategic variable and price is determined by the market through an inverse demand function (Cournot 1838). Bertrand oligopolies are competitive markets in which a small number of firms producing similar goods use price as their strategic variable under randomly fluctuating demands (Bertrand 1883). In reality, some markets can be better modeled as Cournot and some as Bertrand. In this paper, we consider continuous time Bertrand models. The firms in this market sell differentiated but substitutable goods. Many products that are sold in markets fit this structure. For example, in the energy market, oil, coal and natural gas are commodities that can be substituted for one another. However, they have different prices per unit of energy produced. The price strategies of the firms in Bertrand oligopoly are characterized using the solution to a system of N_p-dimensional nonlinear Hamilton–Jacobi–Bellman (HJB) partial differential equations (PDEs), where N_p is the number of firms.

Bertrand oligopolies under linear demand functions were analyzed by Ledvina and Sircar (2011) using an asymptotic approximation in the limit of small competition. Numerical solutions were further used to analyze cases with a high degree of substitutability. In particular, they analyze the effects of substitutability and relative firm size on prices, demands and profits. The main finding is that customers benefit the most when a market is composed of many firms of the same relative size producing highly substitutable goods.

The systems of nonlinear HJB PDEs resulting from dynamic oligopoly problems do not admit an explicit solution, except possibly in the monopoly case. Therefore, one needs numerical methods to compute the value functions and the equilibrium strategies. Naive numerical methods quickly become computationally infeasible as the number of players goes beyond three. Moreover, even in the two player case, these PDEs are highly coupled when the competition is strong and are hard to handle. In this paper, we develop an efficient discretization scheme and

*Corresponding author. Email: samarala@uwaterloo.ca

fast numerical solver for nonlinear HJB systems, which has not been addressed in the literature so far.

Seemingly reasonable discretizations of nonlinear PDEs may not converge to the viscosity solution (Pooley *et al*. 2003), which is the financially relevant solution. Therefore, it is important to construct discretization schemes that converge to the viscosity solution. For the systems of HJB PDEs from the dynamic Bertrand duopoly problem, there is no coupling of derivative terms among the individual PDEs and hence the extended definition of the viscosity solution from Ishii and Koike (1991) can be applied here. For nonlinear second-order PDEs, any monotone, consistent, l_∞ stable discretization scheme converges to the viscosity solution provided that the strong comparison property holds (Barles and Souganidis 1991, Ishii and Koike 1991). Among these properties, monotonicity is, in general, hard to achieve. Positive coefficient discretization typically results in monotone schemes.

For the duopoly problem, the Bertrand model results in a two-dimensional system of HJB PDEs. The presence of the cross derivative terms makes the construction of positive coefficient schemes non-trivial. A skewed co-ordinate system can be used to transform the PDEs so that it results in a zero diffusion correlation. One can also rotate the grid by an appropriate angle so that it eliminates the cross derivative term (Zvan *et al*. 2001). The latter approach has the advantage that it preserves the orthogonality of the co-ordinate system. Another alternative is to enforce a spacing restriction on the original finite difference grid such that a positive coefficient condition results (Clift and Forsyth 2008). We adopt this approach as it is computationally inexpensive compared to the other approaches. We theoretically prove that our fully implicit finite difference discretization converges to the unique viscosity solution of the two-dimensional nonlinear HJB system.

We also generalize the fast numerical solver to other systems of HJB and HJB-Isaacs (HJBI) PDEs arising from computational finance applications. The Black–Scholes model has been used with great success to value options. However, it is valid only for processes with constant volatility. Although empirical tests have shown that the Black–Scholes price is fairly close to observed prices, it has been observed that financial models based on stochastic processes having constant volatility are not consistent with market prices. Recent research shows that models based on stochastic volatility, jump diffusion and regime switching processes produce results that better fit market data. Regime switching models result in a system of coupled HJB/HJBI PDEs, which are intuitively appealing and computationally inexpensive compared to the stochastic volatility and jump diffusion models. A two state regime switching model with constant parameters can reproduce a volatility smile (Yao *et al*. 2006). Therefore, we are interested in developing fast numerical solvers for models based on regime switching.

Discretization schemes that converge to the viscosity solution for HJB/HJBI equations associated with American options under regime switching applications have been studied by Forsyth and Labahn (2007) and Huang *et al*. (2011). Among the solvers, Policy iteration (Bellman 1971, Howard 1960, Lions and Mercier 1980) is an efficient and convergent solver for discrete HJB equations. However, convergence is not guaranteed for HJBI equations (Bokanowski *et al*. 2009, Wal 1978). An alternative is to use relaxation-type iterative methods (Barles and Jakobsen 2005), which are convergent for both HJB and HJBI equations. However, a major drawback for relaxation methods is their slow convergence. Therefore, we propose to develop multigrid methods based on a full approximation scheme (FAS). The rate of convergence of multigrid methods is often independent of the grid size (Trottenberg *et al*. 2001). The components of the multigrid method are very specific for the underlying PDE. To the best of our knowledge, multigrid methods have not been developed for systems of HJB and HJBI PDEs previously.

The control and solution of the HJB/HJBI equations are highly nonlinearly coupled. Standard multigrid techniques do not work well for problems with jumps in control or when the control is unbounded (Han and Wan 2013), which often happens in practice. Hoppe (1986) proposed multigrid methods for HJB equations which can be directly applied to the nonlinear problem. The convergence of these methods is slow because jumps in control are ignored. Han and Wan (2013) proposed a multigrid method using a damped relaxation scheme as the smoother and grid transfer techniques which address the issue of jumps in the control. This method is computationally inefficient as the control set size increases. Multigrid methods for linear complementarity problems (LCP) with an application to American options were proposed by Oosterlee (2003). They use a projected pointwise Gauss-Siedel smoother and standard grid transfer operators. The LCP formulation treats the unknown boundary explicitly in a post processing step. Therefore, there is no issue of jumps in control in this formulation. Moreover, not all HJB and HJBI equations can be formulated as complementarity problems. Therefore, techniques for one formulation are not generally applicable to the other. Furthermore, none of these methods are designed for systems of nonlinear HJB and HJBI PDEs.

The main results of this paper can be summarized as follows:

- We first construct a fully implicit, consistent, unconditionally l_∞ stable and monotone discretization that converges to the viscosity solution for the two-dimensional system of HJB PDEs resulting from the dynamic Bertrand duopoly problem.
- We develop a multigrid method based on the FAS to solve systems of discrete HJB and HJBI PDEs.
- We show by a two grid Fourier analysis that the multigrid method gives efficient convergence. We also prove that the multigrid method is monotone, which ensures smooth convergence.
- We demonstrate the effectiveness of our multigrid solver by numerical examples including dynamic Bertrand duopoly and American option pricing under regime switching.

2. Model problems

In this section, we present the model problems and its discretization in detail starting with dynamic Bertrand duopoly, followed by American options under regime switching and American options with unequal lending/borrowing rates with stock borrowing fees under regime switching.

2.1. Dynamic Bertrand oligopoly

We consider a market with two firms which make their decisions dynamically through time. Each firm has a fixed lifetime capacity of production at time $t = 0$ denoted by $x_l(0)$, $l = 1, 2$. At any later time t, the remaining capacity is given by $x_l(t)$. When $x_l(t) = 0$, the firm has exhausted its capacity and is out of business. The cost of production in the dynamic game is assumed to be zero. However it is noticed that shadow costs associated with the scarcity of goods as they run down are introduced in the system.

The price of the good for each firm in general depends on the capacity of all firms, i.e. $p_l = p_l(\mathbf{x}(t))$, where $\mathbf{x}(t) = (x_1(t), x_2(t))$ and is chosen by a Markovian dynamic strategy. Given these prices, each firm expects the market to demand at a rate $D_l(p_1, p_2)$, which is affine in prices (Ledvina and Sircar 2011):

$$D_l(p_1, p_2) = a_1 - a_2\ p_l + a_3\ p_m, \quad l, m = 1, 2,\ m \neq l,$$

where a_1, a_2, a_3 are positive parameters such that $a_2 > a_3$. The intercept parameter a_1 is a measure of general level of demand due to business cycles and recessions and a_3/a_2 is the measure of substitutability. The actual demand from the market $d_l(t)$, however, undergoes short term unpredictable fluctuations, which is modeled by

$$d_l(t) = D_l(p_1, p_2) - \sigma_l\ \upsilon_l(t), \quad l = 1, 2,$$

where $\upsilon_l(t)$ are correlated Gaussian white noise sequences and σ_l is the volatility of the demand of firm l. The lifetime capacity of each firm depletes over time according to the market demand for its good. As a result, the dynamics of the lifetime capacity of the firms are given by $\mathrm{d}x_l(t) = -d_l(t)\mathrm{d}t$. Consequently, the stochastic differential equation for the lifetime capacity is given by

$$\mathrm{d}x_l(t) = -D_l\,(p_1, p_2)\,\mathrm{d}t + \sigma_l\ \mathrm{d}W_l(t), \quad \text{if } x_l > 0,\ l = 1, 2,$$

where $W_l(t)$ are correlated Brownian motions. If $x_l(t) = 0$, then for all $s \geq t$, $x_l(s) = 0$.

Given the initial lifetime capacity $x_l(0) > 0$, the players seek to maximize their expected discounted lifetime profit, also known as the value function $V_l(x_1, x_2)$, in the Nash equilibrium sense. Each player l maximizes its value function by assuming that the other player is using its equilibrium pricing strategy p_m^*:

$$\begin{aligned} &V_l(x_1, x_2) \\ &= \sup_{p_l \geq 0} \mathbb{E}\left\{\int_0^\infty e^{-rt}\, p_l(\mathbf{x}(t)) D_l\left(p_l(\mathbf{x}(t)),\, p_m^*(\mathbf{x}(t))\right)\right. \\ &\quad \left.\times\ \mathbb{1}_{\{x_l(t)>0\}}\mathrm{d}t\right\},\ l, m = 1, 2,\ m \neq l, \end{aligned} \tag{1}$$

where $r > 0$ is the discount rate. Using a dynamic programming argument for nonzero sum differential games (Dockner *et al*. 2000), the equations for the value functions (1) can be reformulated as a system of coupled backward HJB PDEs:

$$(V_l)_\tau = \sup_{p_l \geq 0}\left\{\mathcal{L}^{p_l} V_l\right\}, \quad l = 1, 2, \tag{2}$$

where the differential operators $\mathcal{L}^{p_l}$ are given by

$$\begin{aligned} \mathcal{L}^{p_1} &= \frac{1}{2}\sigma_1^2\frac{\partial^2 V_1}{\partial x_1^2} + \rho\sigma_1\sigma_2\frac{\partial^2 V_1}{\partial x_1 \partial x_2} + \frac{1}{2}\sigma_2^2\frac{\partial^2 V_1}{\partial x_2^2} \\ &\quad - \left(a_1 - a_2 p_1 + a_3 p_2^*\right)\frac{\partial V_1}{\partial x_1} \\ &\quad + \left[\frac{\gamma}{\eta}\left(a_1 - a_2 p_1 + a_3 p_2^*\right) - \frac{\kappa - p_2^*}{\eta}\right]\frac{\partial V_1}{\partial x_2} - rV_1 \\ &\quad + p_1\left(a_1 - a_2 p_1 + a_3 p_2^*\right). \\ \mathcal{L}^{p_2} &= \frac{1}{2}\sigma_1^2\frac{\partial^2 V_2}{\partial x_1^2} + \rho\sigma_1\sigma_2\frac{\partial^2 V_2}{\partial x_1 \partial x_2} + \frac{1}{2}\sigma_2^2\frac{\partial^2 V_2}{\partial x_2^2} \\ &\quad + \left[\frac{\gamma}{\eta}\left(a_1 - a_2 p_2 + a_3 p_1^*\right) - \frac{\kappa - p_1^*}{\eta}\right]\frac{\partial V_2}{\partial x_1} \\ &\quad - \left(a_1 - a_2 p_2 + a_3 p_1^*\right)\frac{\partial V_2}{\partial x_2} - rV_2 \\ &\quad + p_2\left(a_1 - a_2 p_2 + a_3 p_1^*\right), \end{aligned}$$

where ρ is the correlation coefficient of the Brownian motions. The parameters κ, η and γ are positive and are defined as (Ledvina and Sircar 2011)

$$\gamma = \frac{a_3}{(a_2 - a_3)(a_2 + a_3)}, \quad \kappa = \gamma\ a_1\left(\frac{a_2}{a_3} + 1\right), \quad \eta = \gamma\ \frac{a_2}{a_3}.$$

The parameter γ gives a measure of degree of substitutability. We note that coupling in the system of HJB PDEs (2) is due to the equilibrium pricing strategies (p_1^*, p_2^*).

The domain of the PDE is $x_1 > 0,\ x_2 > 0,\ \tau > 0$. For computational purposes, the domain is truncated to $\Omega^D = (x_1, x_2, \tau) \in [0, x_{1,\max}] \times [0, x_{2,\max}] \times [0, T]$. When one firm runs out of capacity, the other has a monopoly. If $v(x)$ is the value function of a monopolist, then on $x_1 = 0,\ x_2 > 0$, we have $V_1(0, x_2) = 0$, $V_2(0, x_2) = v(x_2)$. Similarly when $x_1 > 0,\ x_2 = 0$, we have $V_1(x_1, 0) = v(x_1)$ and $V_2(x_1, 0) = 0$. We look for solutions in which $\lim_{x_l \to \infty} \partial V_l/\partial x_l = 0,\ l = 1, 2$ and $\lim_{x_m \to \infty} \partial V_l/\partial x_m = 0,\ l, m = 1, 2,\ m \neq l$. As a result, we use Neumann conditions on $x_1 = x_{1,\max}$ and $x_2 = x_{2,\max}$. When $\gamma = 0$, both firms are monopolists in disjoint markets of their own goods and hence $V_l(x_1, x_2) = v(x_l)$. The value function $v(x)$ of the monopoly firm is given by the following HJB equation (Ledvina and Sircar 2011)

$$v_\tau = \sup_{p \geq 0}\left\{\mathcal{L}^{M,p} v\right\}, \tag{3}$$

where

$$\mathcal{L}^{M,p} v = \frac{1}{2}\sigma^2\frac{\partial^2 v}{\partial x^2} - \frac{1}{\eta}(\kappa - p)\frac{\partial v}{\partial x} - rv + \frac{p}{\eta}(\kappa - p), \tag{4}$$

where x denotes the firm's remaining life time capacity and p is the price of the firm's good. The domain of the PDE is $x > 0,\ \tau > 0$. The boundary conditions are $v(0) = 0$ and $\lim_{x \to \infty} \partial v/\partial x = 0$. For computational purposes, the domain is localized to $\Omega^M = (x, \tau) \in [0, x_{\max}] \times [0, T]$.

Ledvina and Sircar (2011) mainly consider the analysis of the HJB PDEs (2) and (3). In this paper, we develop a fully implicit, positive coefficient, finite difference discretization scheme for (2) and (3). We prove that the discretization scheme for (2) converges to the viscosity solution in section 3. It can also be shown that the discretization scheme for (3) converges to the viscosity solution, but we omit the details here due to space limitations.

2.1.1. Discretization. We first briefly provide the discretization details for the scalar monopoly problem (3). The spatial domain is discretized into a set of nodes $\{x_0, x_1, \ldots, x_{N-1}\}$ with a uniform grid spacing Δx. Let v_i^n be the approximate solution of (3) at (x_i, τ^n). We assume a mesh and control discretization parameter h such that

$$\Delta x = C_1 h, \quad \Delta\tau = C_2 h, \quad \Delta p = C_3 h, \tag{5}$$

where C_1, C_2 and C_3 are constants independent of h.

The differential term $\mathcal{L}^{M,p}v$ in (3) is discretized using a fully implicit finite difference discretization, which results in

$$\mathcal{L}_h^{M,p} v_i^n = \alpha_i(p) v_{i-1}^n + \beta_i(p) v_{i+1}^n - \left(\alpha_i(p) + \beta_i(p) + r\right) v_i^n + \frac{p}{\eta}(\kappa - p), \tag{6}$$

where $\alpha_i(p)$ and $\beta_i(p)$ are given in algorithm 2. A combination of central and upstream differencing is used for the drift term such that the coefficients $\alpha_i(p)$ and $\beta_i(p)$ are positive, while ensuring that central differencing is used as much as possible (Wang and Forsyth 2008). Using a fully implicit time stepping and (6), the discrete form of (3) is then given by

$$\frac{v_i^{n+1} - v_i^n}{\Delta\tau} = \sup_{p\geq 0}\left\{\mathcal{L}_h^{M,p} v_i^{n+1}\right\}. \tag{7}$$

We next consider the two-dimensional HJB system (2). The spatial domain is discretized into a set of nodes $\{(x_1)_0, (x_1)_1, \ldots, (x_1)_{N_1-1}\} \times \{(x_2)_0, (x_2)_1, \ldots, (x_2)_{N_2-1}\}$ with a uniform grid spacing of size Δx_1 and Δx_2 in the x_1 and x_2 directions, respectively. Let $(V_l)_{i,j}^n$ be the approximate solution of (2) at $((x_1)_i, (x_2)_j, \tau^n)$ for $l = 1, 2$. We assume a mesh and control discretization parameter h such that

$$\begin{aligned} &\Delta x_1 = C_4\, h, \quad \Delta x_2 = C_5\, h, \quad \Delta\tau = C_6\, h, \\ &\Delta p_1 = \Delta p_2 = C_7\, h, \end{aligned} \tag{8}$$

where C_4, C_5, C_6 and C_7 are constants independent of h.

A seven point stencil (Clift and Forsyth 2008) is used to discretize the cross derivative term. For $\rho \geq 0$, the stencil in figure 1(a) is used and the finite difference formula is given by

$$\frac{\partial^2 V_l}{\partial x_1 \partial x_2} \approx \frac{2\,(V_l)_{i,j}^n + (V_l)_{i+1,j+1}^n + (V_l)_{i-1,j-1}^n - (V_l)_{i+1,j}^n - (V_l)_{i-1,j}^n - (V_l)_{i,j+1}^n - (V_l)_{i,j-1}^n}{2\Delta x_1 \Delta x_2}. \tag{9}$$

For $\rho < 0$, the stencil in figure 1(b) is used and the corresponding formula is given by

$$\frac{\partial^2 V_l}{\partial x_1 \partial x_2} \approx \frac{-2\,(V_l)_{i,j}^n - (V_l)_{i+1,j-1}^n - (V_l)_{i-1,j+1}^n + (V_l)_{i+1,j}^n + (V_l)_{i-1,j}^n + (V_l)_{i,j+1}^n + (V_l)_{i,j-1}^n}{2\Delta x_1 \Delta x_2}. \tag{10}$$

Standard three point central differencing is used for $\frac{\partial^2 V_l}{\partial x_1^2}$ and $\frac{\partial^2 V_l}{\partial x_2^2}$. The first order derivatives are discretized using central differencing as much as possible and forward or backward differencing when central differencing fails to satisfy positive coefficient discretization. The discrete form of the objective function in (2) is then given by

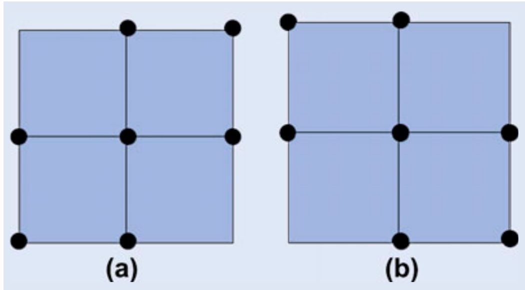

Figure 1. Seven point stencil for finite differencing. (a) $\rho \geq 0$. (b) $\rho < 0$.

$$\begin{aligned} &\mathcal{L}_h^{p_l} V_{i,j}^{n+1} \\ &= \left((\alpha_l)_{i,j}^{x_1} - \xi_{i,j}\right)(V_l)_{i-1,j}^{n+1} + \left((\beta_l)_{i,j}^{x_1} - \xi_{i,j}\right)(V_l)_{i+1,j}^{n+1} \\ &\quad + \left((\alpha_l)_{i,j}^{x_2} - \xi_{i,j}\right)(V_l)_{i,j-1}^{n+1} + \left((\beta_l)_{i,j}^{x_2} - \xi_{i,j}\right)(V_l)_{i,j+1}^{n+1} \\ &\quad + \mathbb{1}_{\rho\geq 0}\, \xi_{i,j}\left((V_l)_{i+1,j+1}^{n+1} + (V_l)_{i-1,j-1}^{n+1}\right) \\ &\quad + \mathbb{1}_{\rho< 0}\, \xi_{i,j}\left((V_l)_{i+1,j-1}^{n+1} + (V_l)_{i-1,j+1}^{n+1}\right) \\ &\quad - \left((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1} + (\alpha_l)_{i,j}^{x_2} + (\beta_l)_{i,j}^{x_2} - 2\xi_{i,j} + r\right)(V_l)_{i,j}^{n+1} \\ &\quad + p_l\left(a_1 - a_2 p_l + a_3 (p_m^*)_{i,j}\right), \qquad l, m = 1, 2, \quad m \neq l, \end{aligned} \tag{11}$$

where $(\alpha_l)_{i,j}^{x_1}$, $(\beta_l)_{i,j}^{x_1}$, $(\alpha_l)_{i,j}^{x_2}$, $(\beta_l)_{i,j}^{x_2}$ and $\xi_{i,j}$ are given in algorithm 3. Using a fully implicit time stepping and (11), the discrete form of (2) is given by

$$\frac{(V_l)_{i,j}^{n+1} - (V_l)_{i,j}^n}{\Delta\tau} = \sup_{p_l\geq 0}\left\{\mathcal{L}_h^{p_l}(V_l)_{i,j}^{n+1}\right\}, \qquad l = 1, 2. \tag{12}$$

We note that the presence of the cross derivative term poses a challenge to construct a positive coefficient discretization. The following theorem illustrates how the constraint on grid spacing results in a positive coefficient discretization.

Theorem 2.1 *If $\rho \neq 0$ and with the use of a seven point stencil, if the grid spacing is chosen such that*

$$\left|2\rho\frac{\sigma_2}{\sigma_1}\right| \leq \frac{\Delta x_2}{\Delta x_1} \leq \left|\frac{1}{2\rho}\frac{\sigma_2}{\sigma_1}\right|, \tag{13}$$

then the positive coefficient conditions are satisfied, i.e.

$$\begin{aligned} &(\alpha_l)_{i,j}^{x_1} - \xi_{i,j} \geq 0, \quad (\beta_l)_{i,j}^{x_1} - \xi_{i,j} \geq 0, && (14)\\ &(\alpha_l)_{i,j}^{x_2} - \xi_{i,j} \geq 0, \quad (\beta_l)_{i,j}^{x_2} - \xi_{i,j} \geq 0, && (15)\\ &\xi_{i,j} \geq 0, && (16)\\ &l = 1, 2, \quad 1 \leq i \leq N_1 - 1, \quad 1 \leq j \leq N_2 - 1. \end{aligned}$$

Proof The use of the seven point stencil ensures that $\xi_{i,j} \geq 0$ irrespective of the sign of the correlation ρ. This is obvious from (A9). From algorithm 3, conditions (14) and (15) can be rewritten as

$$(\alpha_l)_{i,j}^{x_1,ups} - \xi_{i,j} \geq 0, \quad (\beta_l)_{i,j}^{x_1,ups} - \xi_{i,j} \geq 0, \tag{17}$$

$$(\alpha_l)_{i,j}^{x_2,ups} - \xi_{i,j} \geq 0, \quad (\beta_l)_{i,j}^{x_2,ups} - \xi_{i,j} \geq 0. \tag{18}$$

From (A1), (A2), (A5) and (A6), it is clear that the conditions (17) are satisfied when

$$\frac{\sigma_1^2}{2\Delta x_1^2} - \frac{\rho\sigma_1\sigma_2}{\Delta x_1 \Delta x_2} \geq 0. \tag{19}$$

Similarly for the x_2 dimension, it is clear from (A3), (A4), (A7) and (A8), that the conditions (18) hold when

$$\frac{\sigma_2^2}{2\Delta x_2^2} - \frac{\rho\sigma_1\sigma_2}{\Delta x_1 \Delta x_2} \geq 0. \tag{20}$$

Conditions (19) and (20) then result in the restriction on the grid spacing as given in (13). □

2.2. *American options under regime switching*

Consider a regime switching model with N_m regimes and a finite set of discrete volatilities σ_j, $j = 1, 2, \ldots, N_m$. A continuous Markov chain process controls the shifts between these regimes. The stochastic process for the underlying asset S under the real world measure is

$$\mathrm{d}S = \mu_j^{\mathbb{P}}\, S\, \mathrm{d}t + \sigma_j\, S\, \mathrm{d}Z + \sum_{m=1}^{N_m} (\xi_{jm} - 1)\, S\, \mathrm{d}X_{jm}, \quad j = 1, \ldots, N_m, \tag{21}$$

where $\mathrm{d}Z$ is the increment of a Wiener process and $\mu_j^{\mathbb{P}}$ is the drift in regime j. The superscript $\mathbb{P}$ denotes the objective probability measure. The term $\mathrm{d}X_{jm}$ is given by

$$\mathrm{d}X_{jm} = \begin{cases} 1, & \text{with probability } \lambda_{jm}^{\mathbb{P}}\mathrm{d}t + \delta_{jm}, \\ 0, & \text{with probability } 1 - \lambda_{jm}^{\mathbb{P}}\mathrm{d}t - \delta_{jm}, \end{cases}$$

where the transition probability $\lambda_{jm}^{\mathbb{P}} \geq 0$, $j \neq m$ and $\lambda_{jj}^{\mathbb{P}} = -\sum_{m\neq j} \lambda_{jm}^{\mathbb{P}}$. The asset price jumps from S to $\xi_{jm} S$ when a transition from j to m occurs and $\xi_{jj} = 1$. The jump amplitudes ξ_{jm} are assumed to be deterministic functions of (S, t). In practice, the quantities ξ_{jm} and λ_{jm} are determined by calibration to market prices (Ayache 2010).

Let $V_j(S, \tau)$ be the no-arbitrage value of the contingent claim in regime j, where $\tau = T - t$ with T being the expiry time of the contingent claim and t the time variable. Consider a hedging portfolio P such that

$$P = -V_j + \Delta_s S + \sum_{m=1}^{N_m - 1} \Delta_m F_m,$$

where Δ_s is the number of units of underlying asset with price S and Δ_m is the number of units of additional hedging instruments with price F_m. It is possible to set up a perfect hedge under the assumption that the set of assets $\{S, F_1, \ldots, F_{N_m-1}\}$ forms a non-redundant set (Kennedy 2007). The existence of a perfect hedge allows us to define the risk neutral transition probabilities λ_{jm} and the quantities ρ_j and λ_j as (Huang *et al.* 2011)

$$\rho_j = \sum_{m\neq j} \lambda_{jm}(\xi_{jm} - 1), \quad \lambda_j = \sum_{m\neq j} \lambda_{jm}, \quad \lambda_{jj} = -\lambda_j. \tag{22}$$

Let $V = [V_1, V_2, \ldots, V_{N_m}]^T$. The differential operators $\mathcal{L}_j V_j$ and $\mathcal{J}_j V$ are defined as

$$\mathcal{L}_j V_j = \frac{\sigma_j^2 S^2}{2} \frac{\partial^2 V_j}{\partial S^2} + (r - \rho_j) S \frac{\partial V_j}{\partial S} - (r + \lambda_j) V_j, \tag{23}$$

$$\mathcal{J}_j V = \sum_{m\neq j} \frac{\lambda_{jm}}{\lambda_j} V_m(\xi_{jm} S, \tau), \tag{24}$$

where r is the risk free interest rate. The no-arbitrage price of the American option $V_j(S, \tau)$ is then given by (Kennedy 2007)

$$\min\left[V_{j,\tau} - \mathcal{L}_j V_j - \lambda_j \mathcal{J}_j V,\ V_j - V^*\right] = 0, \quad j = 1, \ldots, N_m, \tag{25}$$

where $V^*(S)$ is the payoff function. We consider the truncated domain $(S, \tau) \in [0, S_{\max}] \times [0, T]$ for computational purposes. No boundary condition is required at $S = 0$, whereas at $S = S_{\max}$, we follow the standard approach and use a Dirichlet condition with $V(S_{\max}, \tau) = V^*(S_{\max})$ (Huang *et al.* 2011). The initial condition is given by the payoff function at $\tau = 0$, which is denoted by

$$V(S, 0) = V^*(S).$$

The minimization problem (25) can be solved in different ways. A straight forward approach is to enforce the constraint explicitly. But, the resulting solution is inconsistent and the option delta is not continuous across the early exercise boundary (Forsyth 2012). Alternatively, (25) can be reformulated using different optimal control formulations (Huang *et al.* 2011) to overcome the issues. In this paper, we adopt the penalty method. The penalized form of (25) (Forsyth and Vetzal 2002) is

$$V_{j,\tau} = \mathcal{L}_j V_j + \lambda_j \mathcal{J}_j V + \max_{\varphi\in\{0,1\}} \left[\varphi \frac{(V^* - V_j)}{\epsilon}\right], \tag{26}$$

where ϵ is the penalty parameter and $\varphi \in \{0, 1\}$ is the control parameter. For efficient convergence of the multigrid method, it is important that the consistency of control from the fine to the coarse grids is maintained during restriction (Han and Wan 2013). This can be easily enforced when the penalty formulation (26) is used. We discuss this in detail in section 4.3.

Equation (26) can be written in the general form as

$$V_{j,\tau} = \max_{\varphi\in\{0,1\}} \left\{ a_j(S, \tau, \varphi) \frac{\partial^2 V_j}{\partial S^2} + b_j(S, \tau, \varphi) \frac{\partial V_j}{\partial S} - c_j(S, \tau, \varphi) V_j + d_j(S, \tau, \varphi) + \lambda_j \mathcal{J}_j V \right\}, \tag{27}$$

where

$$a_j(S, \tau, \varphi) = \frac{\sigma_j^2 S^2}{2}, \quad b_j(S, \tau, \varphi) = S(r - \rho_j),$$

$$c_j(S, \tau, \varphi) = \left(r + \lambda_j + \frac{\varphi}{\epsilon}\right), \quad d_j(S, \tau, \varphi) = \varphi \frac{V_j^*}{\epsilon}. \tag{28}$$

Fully implicit finite differencing is used to discretize (27) as described in the following section.

2.2.1. Discretization. We discretize (27) using a fully implicit, positive coefficient, finite difference discretization, which ensures convergence to the viscosity solution (Forsyth and Labahn 2007). The spatial domain is discretized into a set of nodes $\{S_0, S_1, \ldots, S_{N-1}\}$ and the nth time step is denoted by $\tau^n = n\Delta\tau$, where $\Delta\tau$ is the time step size. Let $V_{i,j}^n$ be the discrete approximation to $V_j(S_i, \tau^n)$. The discretized differential terms in (27) are represented as

$$\left(a_j(S,\tau,\varphi)\frac{\partial^2 V_j}{\partial S^2} + b_j(S,\tau,\varphi)\frac{\partial V_j}{\partial S} - c_j(S,\tau,\varphi)V_j\right)_i^{n+1}$$
$$= \alpha_{i,j}(\varphi)\, V_{i-1,j}^{n+1} + \beta_{i,j}(\varphi)\, V_{i+1,j}^{n+1}$$
$$- \left(\alpha_{i,j}(\varphi) + \beta_{i,j}(\varphi) + c_{i,j}(\varphi)\right) V_{i,j}^{n+1}. \tag{29}$$

A weighted average of central and upstream differencing (Huang *et al*. 2011) is used such that the positive coefficient condition ($\alpha_{i,j} \geq 0$, $\beta_{i,j} \geq 0$) is satisfied and central differencing is used as much as possible. The details are given in algorithm 4.

We append all the discrete vectors of the approximation V_j^n to form a long vector V^n of size NM,

$$V^n = \left[V_{0,1}^n, \ldots, V_{N-1,1}^n, \ldots, V_{0,N_m}^n, \ldots, V_{N-1,N_m}^n\right]^T.$$

Let $\mathcal{J}_j^h$ denote the discrete form of the operator $\mathcal{J}_j$, the discretization for the regime switching term $\mathcal{J}_j V$ is then given by

$$[\mathcal{J}_j^h V^{n+1}]_{i,j} = \sum_{m\neq j} \frac{\lambda_{jm}}{\lambda_j} V_m(\min(S_{\max}, \xi_{jm}S_i), \tau^{n+1}), \tag{30}$$

where $V_m(\min(S_{\max}, \xi_{jm}S_i), \tau^{n+1})$ is approximated by linear interpolation, which is given by

$$V_m(\min(S_{\max}, \xi_{jm}S_i), \tau^{n+1}) = w_m V_{i_m,m}^{n+1}$$
$$+ (1 - w_m)V_{i_m+1,m}^{n+1}, \quad w_m \in [0,1]. \tag{31}$$

Note that we truncate any jumps that require data outside the computational domain in (30). The error due to this approximation is small when $S_{\max}$ is sufficiently large (Kennedy 2007). The HJB system (27) is then discretized using (29), (30) and a fully implicit time stepping as

$$\frac{V_{i,j}^{n+1} - V_{i,j}^n}{\Delta\tau}$$
$$= \max_{\varphi\in\{0,1\}}\Big\{\alpha_{i,j}(\varphi)V_{i-1,j}^{n+1} + \beta_{i,j}(\varphi)V_{i+1,j}^{n+1}$$
$$- \left(\alpha_{i,j}(\varphi) + \beta_{i,j}(\varphi) + c_{i,j}(\varphi)\right) V_{i,j}^{n+1}$$
$$+ d_{i,j}(\varphi) + [\mathcal{J}_j^h V^{n+1}]_{i,j}\Big\}, \quad i < N-1;\ j = 1, \ldots, N_m, \tag{32}$$

$$V_{i,j}^{n+1} = V_{i,j}^*, \quad i = N-1;\ j = 1, \ldots, N_m.$$

2.3. *American options with unequal lending/borrowing rates under regime switching*

Consider the model where the cash borrowing rate, r_b, and the lending rate, r_l, are not necessarily equal (with $r_b \geq r_l$) along with the stock borrowing fees, r_f. Such models result in nonlinear HJB PDEs (Bergman 1995, Forsyth and Labahn 2007). We consider these models under a N_m-state regime switching process (21) combined with the American early exercise, which results in a system of HJBI PDEs. Let $V_j(S,\tau)$ be the no-arbitrage value of the contingent claim in regime j and $V = [V_1, V_2, \ldots, V_{N_m}]^T$. We define the following differential operators for a long position in the contingent claim:

$$\mathcal{L}_j^Q V_j = \frac{\sigma_j^2 S^2}{2}\frac{\partial^2 V_j}{\partial S^2} + (q_3 q_1 + (1-q_3)(r_l - r_f) - \rho_j)$$
$$\times S\frac{\partial V_j}{\partial S} - (q_3 q_1 + q_2(1-q_3) + \lambda_j)V_j$$
$$\mathcal{J}_j V = \sum_{m\neq j}\frac{\lambda_{jm}}{\lambda_j} V_m\left(\xi_{jm}S, \tau\right),$$

where ρ_j and λ_j are given by (22). When combined with American early exercise, we have,

$$\min\left[V_{j,\tau} - \inf_{Q\in\hat{Q}}\{\mathcal{L}_j^Q V_j\} - \lambda_j \mathcal{J}_j V,\ V_j - V^*\right] = 0, \tag{33}$$

with $Q = (q_1, q_2, q_3)$ and $\hat{Q} = (\{r_l, r_b\}, \{r_l, r_b\}, 0, 1)$. As mentioned in section 2.2, directly solving the minimization problem (33) leads to an inconsistent solution. Therefore, we reformulate (33) to a penalty form, which results in the following system of HJBI PDEs,

$$V_{j,\tau} = \sup_{\varphi\in\{0,1\}}\inf_{Q\in\hat{Q}}\left\{\mathcal{L}_j^Q V_j + \lambda_j\mathcal{J}_j V + \varphi\frac{(V^* - V_j)}{\epsilon}\right\} \tag{34}$$

We rewrite (34) in the general form as

$$V_{j,\tau} = \sup_{\varphi\in\{0,1\}}\inf_{Q\in\hat{Q}}\Big\{a_j(S,\tau,Q,\varphi)\frac{\partial^2 V_j}{\partial S^2} + b_j(S,\tau,Q,\varphi)$$
$$\times \frac{\partial V_j}{\partial S} - c_j(S,\tau,Q,\varphi)V_j$$
$$+ d_j(S,\tau,Q,\varphi) + \lambda_j\mathcal{J}_j V\Big\}, \tag{35}$$

where

$$a_j(S,\tau,Q,\varphi) = \frac{\sigma_j^2 S^2}{2},\quad b_j(S,\tau,Q,\varphi)$$
$$= S(q_3 q_1 + (1-q_3)(r_l - r_f) - \rho_j),$$
$$c_j(S,\tau,Q,\varphi) = (q_3 q_1 + q_2(1-q_3) + \lambda_j + \frac{\varphi}{\epsilon}),$$
$$d_j(S,\tau,Q,\varphi) = \varphi\frac{V_j^*}{\epsilon}.$$

2.3.1. Discretization. Following the discretization procedure in section 2.2.1, i.e. a fully implicit positive coefficient discretization and linear interpolation (30) for $\mathcal{J}_j V$, we obtain the following for (35),

$$V_{i,j}^{n+1} = V_{i,j}^n + \Delta\tau \sup_{\varphi\in\{0,1\}}\inf_{Q\in\hat{Q}}\Big\{\alpha_{i,j}(Q,\varphi)V_{i-1,j}^{n+1}$$
$$+ \beta_{i,j}(Q,\varphi)V_{i+1,j}^{n+1}$$
$$- \left(\alpha_{i,j}(Q,\varphi) + \beta_{i,j}(Q,\varphi) + c_{i,j}(Q,\varphi)\right) V_{i,j}^{n+1}$$
$$+ d_{i,j}(Q,\varphi) + \lambda_j[\mathcal{J}_j^h V^{n+1}]_{i,j}\Big\}, \quad i < N-1,$$
$$V_{i,j}^{n+1} = V_{i,j}^*, \quad i = N-1. \tag{36}$$

The coefficients $\alpha_{i,j}$ and $\beta_{i,j}$ are defined using a weighted average of central and upstream differencing as described in algorithm 4.

Remark 1 Equations (25) and (33) are special cases of the more general systems of variational inequalities (VIs) considered by Crepey (2010), where it is shown that such VIs have unique viscosity solutions. The definition of a viscosity solution must be generalized for systems of PDEs (Crepey 2010, Ishii and Koike 1991). The discretization schemes (32) and (36) for American options under regime switching can be shown to be unconditionally l_∞ stable, consistent and monotone and hence converge to the viscosity solution in a straightforward way by using methods in Forsyth and Labahn (2007).

In the next section, we prove that the discretization scheme (12), for dynamic Bertrand duopoly, converges to the viscosity solution.

3. Discretization analysis

For nonlinear second-order PDEs, any monotone, consistent and l_∞ stable discretization scheme converges to the viscosity solution provided that the strong comparison property holds (Barles and Souganidis 1991, Ishii and Koike 1991). For the system of HJB PDEs (2), we note that there is no coupling of derivative terms among the individual PDEs and hence the extended definition of the viscosity solution from Ishii and Koike (1991) can be applied here. In this Section, we verify that our discretization (12) converges to the viscosity solution of the system (2).

For compactness of analysis, let

$$\mathbf{x} = (x_1, x_2, \tau), \qquad DV_l(\mathbf{x}) = \left(\frac{\partial V_l}{\partial x_1}, \frac{\partial V_l}{\partial x_2}, \frac{\partial V_l}{\partial \tau}\right),$$

$$D^2 V_l(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 V_l}{\partial x_1^2} & \frac{\partial^2 V_l}{\partial x_1 \partial x_2} \\ \frac{\partial^2 V_l}{\partial x_1 \partial x_2} & \frac{\partial^2 V_l}{\partial x_2^2} \end{pmatrix}.$$

Definition 1 The domain Ω^D is partitioned into

$$\begin{aligned}
\Omega^D_{in} &= (x_1, x_2, \tau) \in (0, (x_1)_{\max}] \times (0, (x_2)_{\max}] \times (0, T], \\
\Omega^D_{(x_1)_0} &= (x_1, x_2, \tau) \in \{0\} \times (0, (x_2)_{\max}] \times (0, T], \\
\Omega^D_{(x_2)_0} &= (x_1, x_2, \tau) \in (0, (x_1)_{\max}] \times \{0\} \times (0, T], \\
\Omega^D_{\tau_0} &= (x_1, x_2, \tau) \in [0, (x_1)_{\max}] \times [0, (x_2)_{\max}] \times \{0\}.
\end{aligned}$$

The system of HJB equations (2) is then written in compact form as

$$F^l V_l = F^l\left(\mathbf{x}, V_l(\mathbf{x}), DV_l(\mathbf{x}), D^2 V_l(\mathbf{x}), \{V_m(\mathbf{x})\}_{m \neq l}\right) = 0, \quad \times\, l, m = 1, 2, \tag{37}$$

where $F^l V_l$ is defined as

$$F^l V_l = \begin{cases}
F^l_{in} V_l \equiv F^l_{in}\left(\mathbf{x}, V_l(\mathbf{x}), DV_l(\mathbf{x}), D^2 V_l(\mathbf{x}), \{V_m(\mathbf{x})\}_{m\neq l}\right) \\
\quad = (V_l)_\tau - \sup_{p_l \geq 0}\left\{\mathcal{L}^{p_l} V_l\right\}, \quad \mathbf{x} \in \Omega^D_{in}, \\
F^l_{(x_1)_0} V_l \equiv F^l_{(x_1)_0}\left(\mathbf{x}, V_l(\mathbf{x}), DV_l(\mathbf{x}), D^2 V_l(\mathbf{x}), \{V_m(\mathbf{x})\}_{m\neq l}\right) \\
\quad = \begin{cases} V_1 \\ V_2 - \sup_{p_2 \geq 0}\left\{\mathcal{L}^{M,p_2} V_2\right\} \end{cases}, \quad \mathbf{x} \in \Omega^D_{(x_1)_0}, \\
F^l_{(x_2)_0} V_l \equiv F^l_{(x_2)_0}\left(\mathbf{x}, V_l(\mathbf{x}), DV_l(\mathbf{x}), D^2 V_l(\mathbf{x}), \{V_m(\mathbf{x})\}_{m\neq l}\right) \\
\quad = \begin{cases} V_1 - \sup_{p_1 \geq 0}\left\{\mathcal{L}^{M,p_1} V_1\right\} \\ V_2 \end{cases}, \quad \mathbf{x} \in \Omega^D_{(x_2)_0}, \\
F^l_{\tau_0} U_l \equiv F^l_{\tau_0}(\mathbf{x}, V_l(\mathbf{x})) \\
\quad = V_l - (V_l)_0(x_1, x_2), \quad \mathbf{x} \in \Omega^D_{\tau_0},
\end{cases}$$

where $(V_l)_0(x_1, x_2)$ is the initial condition.

Definition 2 (Viscosity solution of the system of PDEs (37)) A $\mathbf{R}^2$-valued function $V = (V_1, V_2)$, where each $V_l : \Omega^D \to \mathbf{R}$ is locally bounded, is called a viscosity subsolution (respectively supersolution) of the system of PDEs (37) if and only if for all smooth test functions $\phi_l \in C^\infty(\Omega^D)$, and for all maximum (respectively minimum) points $\mathbf{x}$ of $V_l^* - \phi_l$ (respectively $V_{l*} - \phi_l$), one has

$$F^l_*\left(\mathbf{x}, V_l^*(\mathbf{x}), D\phi_l(\mathbf{x}), D^2\phi_l(\mathbf{x}), \{V_m^*(\mathbf{x})\}_{m\neq l}\right) \leq 0$$
$$\left(\text{respectively } F^{l*}\left(\mathbf{x}, V_{l*}(\mathbf{x}), D\phi_l(\mathbf{x}), D^2\phi_l(\mathbf{x}), \{V_{m*}(\mathbf{x})\}_{m\neq l}\right) \geq 0\right).$$

A locally bounded function V is a viscosity solution if it is both a viscosity subsolution and a viscosity supersolution.

There is no coupling of derivative terms among the individual PDEs of the system (37), hence the test function for the l-th equation is scalar valued and replaces only the l-th component of the solution V, as in the above definition of the viscosity solution. It is in this sense that we extend the convergence result of Barles and Souganidis (1991) to system of PDEs that arise in dynamic Bertrand duopoly. Related work that also generalize the result of Barles and Souganidis (1991) to systems of PDEs are Ishii and Koike (1991), Reisinger and Forsyth (2015), Tse (2012).

ASSUMPTION 3.1 (Strong comparison property) *If U is an upper semi-continuous sub-solution of* (37) *and if V is a lower semi-continuous super-solution of* (37), *then*

$$U \leq V.$$

The strong comparison property has been proven for first order equations for all kinds of classical equations and boundary conditions. It has also been proven for second order equations with Neumann boundary conditions and classical Dirichlet boundary conditions (Barles 1997, Barles and Rouy 1998, Chaumont 2004). *Only fully degenerate equations are not well understood. As such, it is clear that the strong comparison property holds for all PDEs in this paper.*

3.1. Consistency

In this section, we prove that the discretization scheme (12) is a consistent approximation to the system of PDEs (2) in the viscosity sense. Let $\mathcal{G}^l(.)$ be the discrete approximation to

F^l_{in} for $\mathbf{x} \in \Omega^D_{in}$ and $\mathbf{x}^{n+1}_{i,j} = ((x_1)_i, (x_2)_j, \tau^{n+1})$. Then for $\mathbf{x}^{n+1}_{i,j} \in \Omega^D_{in}$, we rewrite (12) as

$$\begin{aligned}&\mathcal{G}^l\Bigg(h, \mathbf{x}^{n+1}_{i,j}, (V_l)^{n+1}_{i,j}, \{(V_l)^{n+1}_{i',j'}\}_{\substack{i'\neq i\\ or\ j'\neq j}},\\ &\quad\times\{(V_l)^n_{i',j'}\}, \{(V_m)^{n+1}_{i',j'}\}_{m\neq l}\Bigg)\\ &= \frac{(V_l)^{n+1}_{i,j} - (V_l)^n_{i,j}}{\Delta\tau} - \sup_{p_l\geq 0}\left\{\mathcal{L}^{p_l}_h (V_l)^{n+1}_{i,j}\right\} = 0,\\ &\quad l, m = 1, 2.\end{aligned} \tag{38}$$

For $\mathbf{x}^{n+1}_{i,j} \in \Omega^D_{(x_1)_0}$, we have

$$\begin{aligned}&\mathcal{G}^l\Bigg(h, \mathbf{x}^{n+1}_{i,j}, (V_l)^{n+1}_{i,j}, \{(V_l)^{n+1}_{i',j'}\}_{\substack{i'\neq i\\ or\ j'\neq j}},\\ &\times\{(V_l)^n_{i',j'}\}, \{(V_m)^{n+1}_{i',j'}\}_{m\neq l}\Bigg)\\ &= \begin{cases} V^{n+1}_l = 0, & l = 1,\\ \frac{(V_l)^{n+1}_{i,j} - (V_l)^n_{i,j}}{\Delta\tau} - \sup_{p_l\geq 0}\left\{\mathcal{L}^{M,p_l}_h (V_l)^{n+1}_{i,j}\right\} = 0, & l = 2.\end{cases}\end{aligned} \tag{39}$$

Similarly, for $\mathbf{x}^{n+1}_{i,j} \in \Omega^D_{(x_2)_0}$, we have

$$\begin{aligned}&\mathcal{G}^l\Bigg(h, \mathbf{x}^{n+1}_{i,j}, (V_l)^{n+1}_{i,j}, \{(V_l)^{n+1}_{i',j'}\}_{\substack{i'\neq i\\ or\ j'\neq j}},\\ &\times\{(V_l)^n_{i',j'}\}, \{(V_m)^{n+1}_{i',j'}\}_{m\neq l}\Bigg)\\ &= \begin{cases} \frac{(V_l)^{n+1}_{i,j} - (V_l)^n_{i,j}}{\Delta\tau} - \sup_{p_l\geq 0}\left\{\mathcal{L}^{M,p_l}_h (V_l)^{n+1}_{i,j}\right\} = 0, & l = 1,\\ V^{n+1}_l = 0, & l = 2\end{cases}\end{aligned} \tag{40}$$

Finally, for $\mathbf{x}^{n+1}_{i,j} \in \Omega^D_{\tau_0}$, we have

$$\begin{aligned}&\mathcal{G}^l\Bigg(h, \mathbf{x}^{n+1}_{i,j}, (V_l)^{n+1}_{i,j}, \{(V_l)^{n+1}_{i',j'}\}_{\substack{i'\neq i\\ or\ j'\neq j}},\\ &\times\{(V_l)^n_{i',j'}\}, \{(V_m)^{n+1}_{i',j'}\}_{m\neq l}\Bigg)\\ &= (V_l)((x_1)_i, (x_2)_j, 0) - (V_l)_0((x_1)_i, (x_2)_j) = 0, \quad l, m = 1, 2.\end{aligned} \tag{41}$$

Definition 3 Let $\{d_m\}_{m\neq l}$ be a set of real values d_m. We use the notation

$$\mathcal{G}^l\left(h, \mathbf{x}^{n+1}_{i,j}, (V_l)^{n+1}_{i,j}, \left\{(V_l)^{n+1}_{i',j'}\right\}_{\substack{i'\neq i\\ or\ j'\neq j}}, \left\{(V_l)^n_{i',j'}\right\}, \{d_m\}_{m\neq l}\right)$$

to mean

$$\begin{aligned}&\mathcal{G}^l\Bigg(h, \mathbf{x}^{n+1}_{i,j}, (V_l)^{n+1}_{i,j}, \left\{(V_l)^{n+1}_{i',j'}\right\}_{\substack{i'\neq i\\ or\ j'\neq j}},\\ &\times\left\{(V_l)^n_{i',j'}\right\}, \left\{(V_m)^{n+1}_{i',j'} = d_m\right\}_{m\neq l}\Bigg),\end{aligned}$$

which implies that for a fixed m, $(V_m)^{n+1}_{i',j'}$ all have the same value d_m.

Definition 4 (Consistency) For any C^∞ function $\phi_l(x_1, x_2, \tau)$ in Ω^D with $(\phi_l)^{n+1}_{i,j} = \phi_l\left(\mathbf{x}^{n+1}_{i,j}\right) = \phi_l\left((x_1)_i, (x_2)_j, \tau^{n+1}\right)$, the discretization scheme $\mathcal{G}^l(.)$ is consistent in the viscosity sense if $\forall\hat{\mathbf{x}} = \left(\hat{x}_1, \hat{x}_2, \hat{\tau}\right)$ with $\mathbf{x}^{n+1}_{i,j} = \left((x_1)_i, (x_2)_j, \tau^{n+1}\right)$ and $l, m = 1, 2$ and for a small constant ψ, the following holds

$$\begin{aligned}&\limsup_{\substack{h\to 0\\ \psi\to 0\\ \mathbf{x}^{n+1}_{i,j}\to\hat{\mathbf{x}}}} \mathcal{G}^l\Bigg(h, \mathbf{x}^{n+1}_{i,j}, (\phi_l)^{n+1}_{i,j} + \psi, \left\{(\phi_l)^{n+1}_{i',j'} + \psi\right\}_{\substack{i'\neq i\\ or\ j'\neq j}},\\ &\quad\times\left\{(\phi_l)^n_{i',j'} + \psi\right\}, \{d_m\}_{m\neq l}\Bigg)\\ &\leq F^{l*}\left(\hat{\mathbf{x}}, \phi_l(\hat{\mathbf{x}}), D\phi_l(\hat{\mathbf{x}}), D^2\phi_l(\hat{\mathbf{x}}), \{d_m\}_{m\neq l}\right),\end{aligned}$$

and

$$\begin{aligned}&\liminf_{\substack{h\to 0\\ \psi\to 0\\ \mathbf{x}^{n+1}_{i,j}\to\hat{\mathbf{x}}}} \mathcal{G}^l\Bigg(h, \mathbf{x}^{n+1}_{i,j}, (\phi_l)^{n+1}_{i,j} + \psi, \left\{(\phi_l)^{n+1}_{i',j'} + \psi\right\}_{\substack{i'\neq i\\ or\ j'\neq j}},\\ &\quad\times\left\{(\phi_l)^n_{i',j'} + \psi\right\}, \{d_m\}_{m\neq l}\Bigg)\\ &\geq F^l_*\left(\hat{\mathbf{x}}, \phi_l(\hat{\mathbf{x}}), D\phi_l(\hat{\mathbf{x}}), D^2\phi_l(\hat{\mathbf{x}}), \{d_m\}_{m\neq l}\right).\end{aligned}$$

Lemma 3.2 (Local consistency) *Suppose the mesh and control discretization parameter h satisfies (8), then for any C^∞ function $\phi_l(x_1, x_2, \tau)$ in Ω^D, with $(\phi_l)^{n+1}_{i,j} = \phi_l\Big((x_1)_i, (x_2)_j, \tau^{n+1}\Big) = \phi_l\left(\mathbf{x}^{n+1}_{i,j}\right)$, and for h and for a sufficiently small constant ψ, we have that*

$$\begin{aligned}&\mathcal{G}^l\Bigg(h, \mathbf{x}^{n+1}_{i,j}, (\phi_l)^{n+1}_{i,j} + \psi, \left\{(\phi_l)^{n+1}_{i',j'} + \psi\right\}_{\substack{i'\neq i\\ or\ j'\neq j}}\\ &\quad\times\left\{(\phi_l)^n_{i',j'} + \psi\right\}, \{d_m\}_{m\neq l}\Bigg)\\ &= \begin{cases} F^l_{in}\,(\phi_l)^{n+1}_{i,j} + O(h) + O(\psi), & \mathbf{x}^{n+1}_{i,j} \in \Omega^D_{in},\\ F^l_{(x_1)_0}\,(\phi_l)^{n+1}_{i,j} + O(h) + O(\psi), & \mathbf{x}^{n+1}_{i,j} \in \Omega^D_{(x_1)_0},\\ F^l_{(x_1)_0}\,(\phi_l)^{n+1}_{i,j} + O(h) + O(\psi), & \mathbf{x}^{n+1}_{i,j} \in \Omega^D_{(x_2)_0},\\ F^l_{\tau_0}\,(\phi_l)^{n+1}_{i,j} + O(\psi), & \mathbf{x}^{n+1}_{i,j} \in \Omega^D_{\tau_0}.\end{cases}\end{aligned} \tag{42}$$

Proof Let

$$\begin{aligned}\mathcal{L}^{p_l}\,(\phi_l)^{n+1}_{i,j} &\equiv \mathcal{L}^{p_l}\phi_l\left((x_1)_i, (x_2)_j, \tau^{n+1}\right),\\ ((\phi_l)_\tau)^{n+1}_{i,j} &= (\phi_l)_\tau\left((x_1)_i, (x_2)_j, \tau^{n+1}\right).\end{aligned}$$

For $\mathbf{x}^{n+1}_{i,j} \in \Omega^D_{in}$, $\mathcal{L}^{p_l}_h(\phi_l)^{n+1}_{i,j}$ given by (11) is a locally consistent discretization of the linear operator $\mathcal{L}^{p_l}$, i.e. by use of Taylor series, we get

$$\mathcal{L}^{p_l}_h(\phi_l)^{n+1}_{i,j} = \mathcal{L}^{p_l}(\phi_l)^{n+1}_{i,j} + O(h).$$

Further, we have

$$\begin{aligned}\mathcal{L}^{p_l}_h\left((\phi_l)^{n+1}_{i,j} + \psi\right) &= \mathcal{L}^{p_l}_h(\phi_l)^{n+1}_{i,j} - r\psi,\\ \frac{(\phi_l)^{n+1}_{i,j} - (\phi_l)^n_{i,j}}{\Delta\tau} &= ((\phi_l)_\tau)^{n+1}_{i,j} + O(h).\end{aligned}$$

From (38), we then have

$$\mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (\phi_l)_{i,j}^{n+1} + \psi, \{(\phi_l)_{i',j'}^{n+1} + \psi\}_{\substack{i'\neq i \\ or\ j'\neq j}},\right.$$
$$\left.\times \{(\phi_l)_{i',j'}^{n} + \psi\}, \{d_m\}_{m\neq l}\right)$$
$$= \frac{(\phi_l)_{i,j}^{n+1} - (\phi_l)_{i,j}^{n}}{\Delta\tau} - \sup_{p_l\geq 0}\left\{\mathcal{L}_h^{p_l}(\phi_l)_{i,j}^{n+1}\right\} + O(\psi)$$
$$= ((\phi_l)_\tau)_{i,j}^{n+1} - \sup_{p_l\geq 0}\left\{\mathcal{L}^{p_l}(\phi_l)_{i,j}^{n+1}\right\} + O(h) + O(\psi)$$
$$= F_{in}^l(\phi_l)_{i,j}^{n+1} + O(h) + O(\psi), \quad \mathbf{x}_{i,j}^{n+1} \in \Omega_{in}^D.$$

The remaining results in (42) are proved similarly using Taylor series and (39)–(41). □

LEMMA 3.3 (Consistency) *Provided all the conditions in Lemma 3.2 are satisfied, the scheme (38)–(41) is consistent according to the Definition 4.*

Proof The proof follows in a straightforward fashion from Lemma 3.2 and following the analysis in Huang and Forsyth (2012) □

3.2. Stability

Definition 5 (Stability) Discretization (12) is stable if

$$||V_l^n||_\infty \leq C_{10},$$

for $0 \leq n \leq N$ and $h \to 0$, where C_{10} is independent of h.

LEMMA 3.4 (Stability) *Given the positive coefficient conditions (14)–(16) are satisfied, the discretization (12) is unconditionally l_∞ stable, as the mesh parameter (8) $h \to 0$ satisfying*

$$||V_l^n||_\infty \leq ||V_l^0||_\infty + C_{11}, \tag{43}$$

where $C_{11} = T\frac{(a_1+a_3(p_m^*)_{\max})^2}{4a_2}$ *and* $(p_m^*)_{\max} = \max_{i,j}(p_m^*)_{i,j}$.

Proof The discrete equations given by (12) are

$$\begin{aligned}(V_l)_{i,j}^{n+1} = (V_l)_{i,j}^{n} - \Delta\tau\Big((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1}\\
+ (\alpha_l)_{i,j}^{x_2} + (\beta_l)_{i,j}^{x_2} - 2\xi_{i,j} + r\Big)(V_l)_{i,j}^{n+1}\\
+ \Delta\tau\left((\alpha_l)_{i,j}^{x_1} - \xi_{i,j}\right)(V_l)_{i-1,j}^{n+1}\\
+ \Delta\tau\left((\beta_l)_{i,j}^{x_1} - \xi_{i,j}\right)(V_l)_{i+1,j}^{n+1}\\
+ \Delta\tau\left((\alpha_l)_{i,j}^{x_2} - \xi_{i,j}\right)(V_l)_{i,j-1}^{n+1}\\
+ +\Delta\tau\left((\beta_l)_{i,j}^{x_2} - \xi_{i,j}\right)(V_l)_{i,j+1}^{n+1}\\
+ \mathbb{1}_{\rho\geq 0}\,\Delta\tau\xi_{i,j}\left((V_l)_{i+1,j+1}^{n+1} + (V_l)_{i-1,j-1}^{n+1}\right)\\
+ \mathbb{1}_{\rho< 0}\,\Delta\tau\xi_{i,j}\left((V_l)_{i+1,j-1}^{n+1} + (V_l)_{i-1,j+1}^{n+1}\right)\\
+ \Delta\tau(p_l)_{i,j}^{n+1}\left(a_1 - a_2(p_l)_{i,j}^{n+1} + a_3(p_m^*)_{i,j}\right).\end{aligned} \tag{44}$$

To avoid notational clutter, we suppressed the p_l dependence of the discrete coefficients in (44). It is implied that the coefficients are the limiting values at the optimal p_l. The following inequality is obtained from (44)

$$\begin{aligned}\left|(V_l)_{i,j}^{n+1}\right|\Big(1 + \Delta\tau\Big((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1} + (\alpha_l)_{i,j}^{x_2}\\
+ (\beta_l)_{i,j}^{x_2} - 2\xi_{i,j} + r\Big)\Big)\\
\leq ||V_l^n||_\infty + \Delta\tau\Big((\alpha_l)_{i,j}^{x_1} + (\beta_l)_{i,j}^{x_1} + (\alpha_l)_{i,j}^{x_2}\\
+ (\beta_l)_{i,j}^{x_2} - 2\xi_{i,j} + r\Big)||V_l^{n+1}||_\infty\\
+ \Delta\tau\left|(p_l)_{i,j}^{n+1}\left(a_1 - a_2(p_l)_{i,j}^{n+1} + a_3(p_m^*)_{i,j}\right)\right|.\end{aligned} \tag{45}$$

Let $||V_l^{n+1}||_\infty = |(V_l)_{i',j'}^{n+1}|$, then (45) becomes

$$||(V_l)_{i,j}^{n+1}||_\infty(1 + r\Delta\tau) \leq ||U_l^n||_\infty$$
$$+ \Delta\tau\left|(p_l)_{i',j'}^{n+1}\left(a_1 - a_2(p_l)_{i',j'}^{n+1} + a_3(p_m^*)_{i',j'}\right)\right|.$$

Letting $(p_m^*)_{\max} = \max_{i,j}(p_m^*)_{i,j}$ and $\hat{d}_{\max}^{n+1} = \max_{i,j}\left((p_l)_{i,j}^{n+1}\left(a_1 - a_2(p_l)_{i,j}^{n+1} + a_3(p_m^*)_{\max}\right)\right) = \frac{(a_1+a_3(p_m^*)_{\max})^2}{4a_2}$, we get

$$||V_l^{n+1}||_\infty \leq ||V_l^n||_\infty + \Delta\tau\hat{d}_{\max}^{n+1},$$

which then results in (43). □

3.3. Monotonicity

The notion of monotonicity in Barles and Souganidis (1991) needs to be extended for systems of PDEs. Quasi-monotone property is an important assumption in the theory of viscosity solution for systems of PDEs (Ishii and Koike 1991). We first show that the system (37) satisfies the quasi-monotone property.

PROPOSITION 3.5 *Let $w_1, w_2 \in \mathbf{R}^2$ and $l \in \{1, 2\}$. We use the notation $w_1 \geq_l w_2$, which means that $w_1 \geq w_2$ component-wise and $(w_1)_l = (w_2)_l$. The system of PDEs (37) is called quasi-monotone Ishii and Koike (1991), if for all $\mathbf{x} \in \Omega^D$ and $\phi_l \in C^\infty$, whenever $w_1 \geq_l w_2$, then*

$$F^l\left(\mathbf{x}, (w_1)_l(\mathbf{x}), D\phi_l(\mathbf{x}), D^2\phi_l(\mathbf{x}), \{(w_1)_m(\mathbf{x})\}_{m\neq l}\right)$$
$$\leq F^l\left(\mathbf{x}, (w_2)_l(\mathbf{x}), D\phi_l(\mathbf{x}), D^2\phi_l(\mathbf{x}), \{(w_2)_m(\mathbf{x})\}_{m\neq l}\right).$$

Proof This follows from a straightforward calculation by noting that the coupling among the individual PDEs in (37) is only due to the control and not the solution. □

We now prove that our discretization (38)–(41) is monotone. Note that the definition of monotonicity with respect to the last argument of $\mathcal{G}^l(.)$ in (46) below is a discrete version of the quasi-monotone property given in proposition 3.5.

Definition 6 (Monotonicity) The discretization scheme (38)–(41) is monotone if for any two $\mathbf{R}^2$-valued discrete functions W_h and U_h defined on Ω^D such that $W_h \geq U_h$ and $(W_l)_{i,j}^{n+1} = (U_l)_{i,j}^{n+1}$,

$$\mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (W_l)_{i,j}^{n+1}, \left\{(W_l)_{i',j'}^{n+1}\right\}_{\substack{i'\neq i \\ or\ j'\neq j}},\right.$$
$$\left.\times\left\{(W_l)_{i',j'}^{n}\right\}, \left\{(W_m)_{i',j'}^{n+1}\right\}_{m\neq l}\right)$$

$$\leq \mathcal{G}^l\left(h, \mathbf{x}_{i,j}^{n+1}, (U_l)_{i,j}^{n+1}, \left\{(U_l)_{i',j'}^{n+1}\right\}_{\substack{i'\neq i \\ or\ j'\neq j}},\right.$$
$$\left.\times\left\{(U_l)_{i',j'}^{n}\right\}, \left\{(U_m)_{i',j'}^{n+1}\right\}_{m\neq l}\right). \tag{46}$$

LEMMA 3.6 (Monotonicity) *If the scheme (38)–(41) satisfies the conditions required for Lemma 3.4, then the discretization is monotone according to Definition 6.*

Proof Our discretization (38)–(41) satisfies positive coefficient conditions $\forall p_l \geq 0$, therefore monotonicity is proved using the same steps from Forsyth and Labahn (2007). □

3.4. Convergence

THEOREM 3.7 (Convergence) *Assuming that the discretization (38)–(41) satisfies all the conditions required by Lemma 3.3, 3.4 and 3.6 and that Assumption 3.1 holds for (37), then the numerical scheme converges to the unique viscosity solution of the system (37).*

Proof Since the scheme is monotone, consistent and l_∞ stable, the convergence follows from the results in Barles and Souganidis (1991), Ishii and Koike (1991). □

4. Multigrid method for HJB and HJBI systems

We develop a multigrid method based on the FAS for solving (32), (12) and (36). FAS is a multigrid method which directly handles the nonlinearity of the underlying PDE (Brandt, 1977). The idea of multigrid methods is to accelerate the convergence of a relaxation scheme by removing the low frequency error efficiently. First, an iterative method, such as Gauss–Seidel or Jacobi relaxation, is applied to the fine grid problem. The resulting error is smooth and hence can be accurately represented on a coarser grid. Since the coarse grid is much smaller than the fine grid, it is much less expensive to work on the coarse grid. In addition, resolving the error on a coarser grid is effective for low frequency error reduction. The fine grid solution is then updated with the interpolated coarse grid error followed by post-smoothing iterations.

We define the problem on the fine grid Ω_h as

$$N^h(.) = F^h. \tag{47}$$

For the HJB system (12) due to dynamic Bertrand oligopoly,

$$N^h\left(V_l^{n+1,h}\right) \equiv V_l^{n+1,h} - \Delta\tau \sup_{p_l\geq 0}\left\{A^{n+1}\left(p_l, p_m^*\right)V_l^{n+1,h} + B^{n+1}\left(p_l, p_m^*\right)\right\},$$
$$F^h \equiv V_l^{n,h}, \quad l, m = 1, 2; \quad m \neq l, \tag{48}$$

where $A^{n+1}\left(p_l, p_m^*\right)V_l^{n+1,h}$ is the matrix form of the differential operator $\mathcal{L}_h^{p_l}$ dependent on $V_l^{n+1,h}$ and $B^{n+1}\left(p_l, p_m^*\right)$ is the vector form of $p_l\left(a_1 - a_2 p_l + a_3 (p_m^*)_{i,j}\right)$ in (12).

For HJB system (32) the operators are defined by

$$N^h\left(V^{n+1,h}\right) \equiv V^{n+1,h} - \Delta\tau \max_{\varphi\in\{0,1\}}\left\{A^{n+1}(\varphi)V^{n+1,h} + B^{n+1}(\varphi)\right\}, \quad F^h \equiv V^{n,h}, \tag{49}$$

where $A^{n+1}(\varphi)V^{n+1,h}$ is the matrix form of the objective function dependent on $V^{n+1,h}$ and $B^{n+1}(\varphi)$ is the vector form of $d_{i,j}(\varphi)$ in (32).

For HJBI system (36), we have

$$N^h\left(V^{n+1,h}\right) \equiv V^{n+1,h} - \Delta\tau \sup_{\varphi\in\{0,1\}}\inf_{Q\in\hat{Q}}\left\{A^{n+1}(Q,\varphi)V^{n+1,h} + B^{n+1}(Q,\varphi)\right\}, \quad F^h \equiv V^{n,h} \tag{50}$$

where $A^{n+1}(Q,\varphi)V^{n+1,h}$ is the matrix form of the objective function dependent on $V^{n+1,h}$ and $B^{n+1}(Q,\varphi)$ is the vector form of $d_{i,j}(\varphi)$ in (36).

Given $V^{n,h}$, the two-grid FAS V-cycle to compute $V^{n+1,h}$ is given in algorithm 1. Recursively applying the two grid method gives the multigrid method. The core components of the multigrid algorithm are the smoothing procedure $\mathcal{S}(.)$, the coarsening strategy, the restriction $\{\mathcal{R}_v, \mathcal{R}_r\}$ and interpolation operators $\mathcal{P}$ and the coarse grid operator $N_H(.)$, where Ω_H denotes the coarse grid.

Algorithm 1 Two-grid FAS V-cycle.

(1) Pre-smoothing
Compute $\tilde{V}^{n,h}$ by applying ν_1 smoothing iterations $\mathcal{S}(.)$ to $V^{n,h}$:
$\tilde{V}^{n,h} = \mathcal{S}^{\nu_1}\left(V^{n,h}, N^h, F^h\right)$

(2) Coarse Grid Correction
Compute the residual: $\tilde{R}^{n,h} = F^h - N^h(\tilde{V}^{n,h})$
Restrict the residual using $\mathcal{R}_r$: $\tilde{R}^{n,H} = \mathcal{R}_r\tilde{R}^{n,h}$
Restrict the solution using $\mathcal{R}_v$: $V^{n,H} = \mathcal{R}_v\tilde{V}^{n,h}$
Compute the right hand side: $F^H = \tilde{R}^{n,H} + N_H(V^{n,H})$
Solve $N^H(\tilde{V}^{n,H}) = F^H$ for $\tilde{V}^{n,H}$
Compute correction: $\tilde{E}^{n,H} = \tilde{V}^{n,H} - V^{n,H}$
Interpolate the correction using $\mathcal{P}$: $\tilde{E}^{n,h} = \mathcal{P}\tilde{E}^{n,H}$
Correct the approximation: $\hat{V}_h^n = \tilde{V}^{n,h} + \tilde{E}^{n,h}$

(3) Post-smoothing
Compute $V^{n+1,h}$ by applying ν_2 smoothing iterations $\mathcal{S}(.)$ to $\hat{V}^{n,h}$:
$$V^{n+1,h} = \mathcal{S}^{\nu_2}\left(\hat{V}^{n,h}, N^h, F^h\right)$$

For HJB and HJBI systems, in addition to the solution and residual, the control should also be carefully considered during restriction and interpolation. Standard FAS techniques using fully weighted restriction and linear interpolation, in general, work well when the control is continuous and bounded. However, when the control is discrete with large jumps, the convergence of the standard FAS deteriorates or it may not converge at all in certain situations (Han and Wan 2013). For American options formulated in HJB/HJBI form, there is typically a large jump in control. For efficient convergence of the multigrid method, it is important that the consistency of control between the fine and the coarse grid is preserved during restriction and interpolation. Also, the optimal control at the jump locations must be accurately captured during interpolation. We address these issues and develop efficient multigrid methods for the HJB and HJBI systems. A weighted relaxation scheme, described in section 4.1, is used as the smoother. We propose novel interpolation techniques which are presented in section 4.2. The restriction operator is chosen such that it preserves consistency of the control, the details are presented in section 4.3.

4.1. Weighted relaxation smoother

Relaxation type iterative methods are efficient in damping the high frequency error components (Trottenberg *et al*. 2001, Han and Wan 2013). We use weighted relaxation scheme as the smoother. Consider the HJB system resulting from American options under regime switching. Rearranging (32), we have

$$\max_{\varphi\in\{0,1\}}\Big\{\Delta\tau\alpha_{i,j}^{n+1}(\varphi)V_{i-1,j}^{n+1}+\Delta\tau\beta_{i,j}^{n+1}(\varphi)V_{i+1,j}^{n+1} -\left(1+\Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi)+\beta_{i,j}^{n+1}(\varphi)+c_{i,j}^{n+1}(\varphi)\right)\right)V_{i,j}^{n+1} +\Delta\tau d_{i,j}^{n+1}(\varphi)+\Delta\tau[\mathcal{J}_j^h V^{n+1}]_{i,j}+V_{i,j}^n\Big\}=0. \qquad (51)$$

We note that the coefficient of $V_{i,j}^{n+1}$ in (51) is non-negative and hence the equation can be rewritten as

$$\max_{\varphi\in\{0,1\}}\left\{-V_{i,j}^{n+1}+\frac{\Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi)V_{i-1,j}^{n+1}+\beta_{i,j}^{n+1}(\varphi)V_{i+1,j}^{n+1}+d_{i,j}^{n+1}(\varphi)+[\mathcal{J}_j^h V^{n+1}]_{i,j}\right)+V_{i,j}^n}{1+\Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi)+\beta_{i,j}^{n+1}(\varphi)+c_{i,j}^{n+1}(\varphi)\right)}\right\}=0 \qquad (52)$$

Let $\widetilde{V}^k$ be the kth estimate for V^{n+1}. A relaxation scheme can then be derived from (52) as

$$\widetilde{V}_{i,j}^{k+1}=\max_{\varphi\in\{0,1\}}\left\{\frac{\Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi)\widetilde{V}_{i-1,j}^{k}+\beta_{i,j}^{n+1}(\varphi)\widetilde{V}_{i+1,j}^{k}+d_{i,j}^{n+1}(\varphi)+[\mathcal{J}_j^h \widetilde{V}^{k}]_{i,j}\right)+V_{i,j}^n}{1+\Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi)+\beta_{i,j}^{n+1}(\varphi)+c_{i,j}^{n+1}(\varphi)\right)}\right\}. \qquad (53)$$

This relaxation scheme is not efficient in reducing the high frequency components. Therefore, we introduce a damping factor ω to obtain a weighted relaxation scheme, which is given by

$$\widetilde{V}_{i,j}^{k+1}=(1-\omega)\widetilde{V}_{i,j}^{k}+\omega\max_{\varphi\in\{0,1\}}\left\{\frac{\Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi)\widetilde{V}_{i-1,j}^{k}+\beta_{i,j}^{n+1}(\varphi)\widetilde{V}_{i+1,j}^{k}+d_{i,j}^{n+1}(\varphi)+[\mathcal{J}_j^h \widetilde{V}^{k}]_{i,j}\right)+V_{i,j}^n}{1+\Delta\tau\left(\alpha_{i,j}^{n+1}(\varphi)+\beta_{i,j}^{n+1}(\varphi)+c_{i,j}^{n+1}(\varphi)\right)}\right\} \qquad (54)$$

Following similar derivation, the weighted relaxation scheme for the HJB system (12) resulting from dynamic Bertrand duopoly is obtained as

$$\begin{aligned}(\widetilde{V}_l)_{i,j}^{k+1} &= (1-\omega)\,(\widetilde{V}_l)_{i,j}^{k}\\ &+\omega\sup_{p_l\ge 0}\left\{\frac{\Delta\tau\left(\left((\alpha_l)_{i,j}^{x_1}-\xi_{i,j}\right)(V_l)_{i-1,j}^k+\left((\beta_l)_{i,j}^{x_1}-\xi_{i,j}\right)(V_l)_{i+1,j}^k\right)}{1+\Delta\tau\left((\alpha_l)_{i,j}^{x_1}+(\beta_l)_{i,j}^{x_1}+(\alpha_l)_{i,j}^{x_2}+(\beta_l)_{i,j}^{x_2}-2\xi_{i,j}+r\right)}\right.\\ &+\frac{\Delta\tau\left(\left((\alpha_l)_{i,j}^{x_2}-\xi_{i,j}\right)(V_l)_{i,j-1}^k+\left((\beta_l)_{i,j}^{x_2}-\xi_{i,j}\right)(V_l)_{i,j+1}^k\right)}{1+\Delta\tau\left((\alpha_l)_{i,j}^{x_1}+(\beta_l)_{i,j}^{x_1}+(\alpha_l)_{i,j}^{x_2}+(\beta_l)_{i,j}^{x_2}-2\xi_{i,j}+r\right)}\\ &+\frac{\Delta\tau\left(\mathbb{1}_{\rho\ge 0}\,\xi_{i,j}\left((V_l)_{i+1,j+1}^k+(V_l)_{i-1,j-1}^k\right)+\mathbb{1}_{\rho<0}\,\xi_{i,j}\left((V_l)_{i+1,j-1}^k+(V_l)_{i-1,j+1}^k\right)\right)}{1+\Delta\tau\left((\alpha_l)_{i,j}^{x_1}+(\beta_l)_{i,j}^{x_1}+(\alpha_l)_{i,j}^{x_2}+(\beta_l)_{i,j}^{x_2}-2\xi_{i,j}+r\right)}\\ &\left.+\frac{\Delta\tau\left(p_l\left(a_1-a_2p_l+a_3\left(p_m^*\right)_{i,j}^k\right)\right)}{1+\Delta\tau\left((\alpha_l)_{i,j}^{x_1}+(\beta_l)_{i,j}^{x_1}+(\alpha_l)_{i,j}^{x_2}+(\beta_l)_{i,j}^{x_2}-2\xi_{i,j}+r\right)}\right\},\quad l,m=1,2;\ m\ne l. \end{aligned} \qquad (55)$$

Similarly, for HJBI systems, the weighted relaxation scheme for (36) is given by

$$\widetilde{V}_{i,j}^{k+1}=(1-\omega)\widetilde{V}_{i,j}^{k}+\omega\sup_{\varphi\in\{0,1\}}\inf_{Q\in\hat{Q}} \times\left\{\frac{\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi)\widetilde{V}_{i-1,j}^{k}+\beta_{i,j}^{n+1}(Q,\varphi)\widetilde{V}_{i+1,j}^{k}+d_{i,j}^{n+1}(Q,\varphi)+[\mathcal{J}_j^h \widetilde{V}^{k}]_{i,j}\right)+V_{i,j}^n}{1+\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi)+\beta_{i,j}^{n+1}(Q,\varphi)+c_{i,j}^{n+1}(Q,\varphi)\right)}\right\}. \qquad (56)$$

LFA shows that $\omega=0.67$ results in efficient reduction of the high frequency error components for both HJB and HJBI systems. We present the details of the analysis in section 5.1. We now prove that the weighted relaxation scheme is globally convergent.

THEOREM 4.1 (HJBI Systems) *Suppose the discretization* (36) *satisfies a positive coefficient condition* (Forsyth and Labahn 2007)*, then the weighted relaxation scheme* (56) *is globally convergent for any initial condition given that* $0 < \omega < 2/(1+\gamma)$. *Furthermore, we have*

$$||\widetilde{V}^{k+1} - \widetilde{V}^{k}||_\infty \le (|1-\omega| + \omega\gamma)||\widetilde{V}^{k} - \widetilde{V}^{k-1}||_\infty$$

where

$$\gamma = \max_{i,j} \sup_{\phi\in\{0,1\}} \sup_{Q\in\hat{Q}} \left\{ \frac{\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi) + \beta_{i,j}^{n+1}(Q,\varphi) + \lambda_j\right)}{1+\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi) + \beta_{i,j}^{n+1}(Q,\varphi) + c_{i,j}^{n+1}(Q,\varphi)\right)} \right\}.$$

Proof Using (56), we have $|\widetilde{V}_{i,j}^{k+1} - \widetilde{V}_{i,j}^{k}|$ as

$$\begin{aligned}
&|\widetilde{V}_{i,j}^{k+1} - \widetilde{V}_{i,j}^{k}| \\
&\le |1-\omega||\widetilde{V}_{i,j}^{k} - \widetilde{V}_{i,j}^{k-1}| \\
&+\omega\Bigg| \sup_{\varphi\in\{0,1\}} \inf_{Q\in\hat{Q}} \left\{ \frac{\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi)\widetilde{V}_{i-1,j}^{k} + \beta_{i,j}^{n+1}(Q,\varphi)\widetilde{V}_{i+1,j}^{k} + d_{i,j}^{n+1}(Q,\varphi) + \lambda_j[\mathcal{J}_j^h\widetilde{V}^k]_{i,j}\right) + V_{i,j}^n}{1+\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi) + \beta_{i,j}^{n+1}(Q,\varphi) + c_{i,j}^{n+1}(Q,\varphi)\right)} \right\} \\
&- \sup_{\varphi\in\{0,1\}} \inf_{Q\in\hat{Q}} \left\{ \frac{\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi)\widetilde{V}_{i-1,j}^{k-1} + \beta_{i,j}^{n+1}(Q,\varphi)\widetilde{V}_{i+1,j}^{k-1} + d_{i,j}^{n+1}(Q,\varphi) + \lambda_j[\mathcal{J}_j^h\widetilde{V}^{k-1}]_{i,j}\right) + V_{i,j}^n}{1+\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi) + \beta_{i,j}^{n+1}(Q,\varphi) + c_{i,j}^{n+1}(Q,\varphi)\right)} \right\}\Bigg|
\end{aligned}$$

Using the properties of sup-inf operators (Forsyth and Labahn 2007) and replacing the regime switching term $[\mathcal{J}_j^h\widetilde{V}^k]_{i,j}$ by (30) and (31), we have

$$\begin{aligned}
&|\widetilde{V}_{i,j}^{k+1} - \widetilde{V}_{i,j}^{k}| \\
&\le |1-\omega||\widetilde{V}_{i,j}^{k} - \widetilde{V}_{i,j}^{k-1}| \\
&\quad+\omega \sup_{\varphi\in\{0,1\}} \sup_{Q\in\hat{Q}} \Bigg\{\Bigg| \frac{\Delta\tau\left[\alpha_{i,j}^{n+1}(Q,\varphi)\left(\widetilde{V}_{i-1,j}^{k} - \widetilde{V}_{i-1,j}^{k-1}\right) + \beta_{i,j}^{n+1}(Q,\varphi)\left(\widetilde{V}_{i+1,j}^{k} - \widetilde{V}_{i+1,j}^{k-1}\right)\right]}{1+\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi) + \beta_{i,j}^{n+1}(Q,\varphi) + c_{i,j}^{n+1}(Q,\varphi)\right)} \\
&\quad+ \frac{\lambda_j\Delta\tau\left(\sum_{m\neq j}\frac{\lambda_{jm}}{\lambda_j}\left[w_m\left(\widetilde{V}_{i_m,m}^{k} - \widetilde{V}_{i_m,m}^{k-1}\right) + (1-w_m)\left(\widetilde{V}_{i_m+1,m}^{k} - \widetilde{V}_{i_m+1,m}^{k-1}\right)\right]\right)}{1+\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi) + \beta_{i,j}^{n+1}(Q,\varphi) + c_{i,j}^{n+1}(Q,\varphi)\right)}\Bigg|\Bigg\} \\
&\le |1-\omega||\widetilde{V}_{i,j}^{k} - \widetilde{V}_{i,j}^{k-1}| \\
&\quad+\omega \sup_{\varphi\in\{0,1\}} \sup_{Q\in\hat{Q}} \left\{ \frac{\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi) + \beta_{i,j}^{n+1}(Q,\varphi) + \lambda_j\right)}{1+\Delta\tau\left(\alpha_{i,j}^{n+1}(Q,\varphi) + \beta_{i,j}^{n+1}(Q,\varphi) + c_{i,j}^{n+1}(Q,\varphi)\right)} \right\} |\widetilde{V}_{i,j}^{k} - \widetilde{V}_{i,j}^{k-1}|.
\end{aligned}$$

Therefore,

$$||\widetilde{V}^{k+1} - \widetilde{V}^{k}||_\infty \le (|1-\omega| + \omega\gamma)\,||\widetilde{V}^{k} - \widetilde{V}^{k-1}||_\infty.$$

Note that $\gamma < 1$ as $\alpha_{i,j}^{n+1}(\varphi)$, $\beta_{i,j}^{n+1}(\varphi)$, $c_{i,j}^{n+1}(\varphi)$ and λ_j are all non-negative. Therefore, the weighted relaxation scheme converges if $0 < \omega < 2/(1+\gamma)$. □

By a similar argument, it can be shown that the relaxation schemes (55) and (54) are globally convergent. We omit the proof here.

4.2. Interpolation

We use the HJB system (49) resulting from American options under regime switching as an example to explain the interpolation and restriction operators, but they work well for the Bertrand duopoly (48) and the unequal lending/borrowing (50) cases as demonstrated in the numerical results.

Consider a two grid method. Given $\{S_i^h\}$, $i = 0, 1, \ldots, N-1$, the grid points with even indices are selected as coarse grid points, i.e. $\{S_i^H\}$, $i = 0, 2, \ldots, N-1$. Given the coarse grid solution $\widetilde{V}_{i,j}^{n,H}$, $i = 0, 2, \ldots, N-1$, we want to interpolate the solution on S_i^h, $i = 0, 1, \ldots, N-1$ for all $j = 1, 2, \ldots, N_m$. Let the interpolated solution be denoted by $\hat{V}^{n,h}$.

For the example problem (49), the solution after interpolation must satisfy the American constraint $\hat{V}^{n,h} \ge V^*$. Consider figure 2, where $\widetilde{V}^{n,H}_{i-1,j}$ and $\widetilde{V}^{n,H}_{i+1,j}$ denoted by black circles are solutions from the coarse grid at S^H_{i-1} and S^H_{i+1}, respectively. Suppose the payoff function is a hat function as shown in figure 2. If linear interpolation is used for the noncoarse grid point S^h_i, then the solution $\hat{V}^{n,h}_{i,j}$, denoted by the white circle, lies below the payoff function. It is clear from this example that standard linear interpolation fails to capture the optimal control and hence the correct solution. In the following sections, we present two novel interpolation techniques which address this issue.

4.2.1. Direct interpolation of the solution. We propose a new interpolation technique which is derived from the relaxation scheme and hence we can ensure that the optimal control is accurately captured. We first copy the coarse grid solution to the fine grid points which coincide with the coarse grid points, i.e.

$$\hat{V}^{n,h}_{i,j} = \widetilde{V}^{n,H}_{i,j}, \quad i=0,2,\ldots,N-1, \quad j=1,2,\ldots,N_m \tag{57}$$

For the noncoarse grid points, $i = 1,3,\ldots,N-2$, $j = 1,2,\ldots,N_m$ we interpolate $\hat{V}^{n,h}_{i,j}$ using one iteration of the relaxation scheme (53). We rewrite it here for convenience:

$$\hat{V}^{n,h}_{i,j} = \max_{\varphi\in\{0,1\}} \left\{ \frac{\Delta\tau\left(\alpha^{n+1}_{i,j}(\varphi)\hat{V}^{n,h}_{i-1,j} + \beta^{n+1}_{i,j}(\varphi)\hat{V}^{n,h}_{i+1,j} + d^{n+1}_{i,j}(\varphi) + [\mathcal{J}^h_j\hat{V}^{n,h}]_{i,j}\right) + V^{n,h}_{i,j}}{1+\Delta\tau\left(\alpha^{n+1}_{i,j}(\varphi)+\beta^{n+1}_{i,j}(\varphi)+c^{n+1}_{i,j}(\varphi)\right)} \right\} \tag{58}$$

where $\hat{V}^{n,h}_{i-1,j}$, $\hat{V}^{n,h}_{i+1,j}$ are given by (57). Whenever there is a switch from regime j to regime m, the term $[\mathcal{J}^h_j\hat{V}^{n,h}]_{i,j}$ is approximated with the values $\hat{V}^{n,h}_{i_m,m}$ and $\hat{V}^{n,h}_{i_m+1,m}$ (See (31)). If i_m is odd and $m > j$, then $\hat{V}^{n,h}_{i_m,m}$ is an unknown. In such cases, we use standard linear interpolation to approximate $\hat{V}^{n,h}_{i_m,m}$, i.e.

$$\hat{V}^{n,h}_{i_m,m} = \widetilde{V}^{n,h}_{i_m,m} + 0.5\left(e^{n,H}_{i_m-1,m} + e^{n,H}_{i_m+1,m}\right),$$

where $\widetilde{V}^{n,h}_{i_m,m}$ is the solution after presmoothing and

$$e^{n,H}_{i_m-1,m} = \widetilde{V}^{n,H}_{i_m-1,m} - V^{n,H}_{i_m-1,m}.$$

Similar approximation is used when for $\hat{V}^{n,h}_{i_m+1,m}$ when i_m is even and $m > j$. This approximation does not hamper the convergence of our multigrid method. Since we use the relaxation iteration (58), it is guaranteed that the optimal control is accurately captured. We theoretically prove that the constraint $\hat{V}^{n,h} \geq V^*$ is satisfied when (57)–(58) is used and the resulting multigrid method is monotone in section 5.3. Similarly, for the HJB system (48) and HJBI system (50), we use their respective relaxation iterations for interpolation.

Note that this interpolation is different from the standard interpolation, where the fine grid solution is corrected with the interpolated coarse grid error. Instead, we directly interpolate the solution and the interpolation formula depends on the underlying PDE. We also develop another interpolation technique which is based on the traditional idea of interpolating the coarse grid error. This approach again depends on the underlying PDE as detailed in the following section.

4.2.2. Interpolation of the error. Let the exact solution for time step $n+1$ be $\bar{V}$ and the approximate solution after the kth iteration be $\hat{V}^k = \bar{V} + \varepsilon^k$, where ε^k is the error after the kth approximation. Using the relaxation scheme (53), we obtain

$$\begin{aligned}
&\bar{V}^{k+1}_{i,j} + \varepsilon^{k+1}_{i,j} \\
&= \max_{\varphi\in\{0,1\}} \left\{ \frac{\Delta\tau\left[\alpha^{n+1}_{i,j}(\varphi)(\bar{V}^k_{i-1,j}+\varepsilon^k_{i-1,j}) + \beta^{n+1}_{i,j}(\varphi)(\bar{V}^k_{i+1,j}+\varepsilon^k_{i+1,j})\right]}{1+\Delta\tau\left(\alpha^{n+1}_{i,j}(\varphi)+\beta^{n+1}_{i,j}(\varphi)+c^{n+1}_{i,j}(\varphi)\right)} \right. \\
&\quad \left. + \frac{\Delta\tau\left[d^{n+1}_{i,j}(\varphi) + \lambda_j[\mathcal{J}^h_j(\bar{V}^k+\varepsilon^k)]_{i,j}\right] + V^n_{i,j}}{1+\Delta\tau\left(\alpha^{n+1}_{i,j}(\varphi)+\beta^{n+1}_{i,j}(\varphi)+c^{n+1}_{i,j}(\varphi)\right)} \right\}.
\end{aligned} \tag{59}$$

Let $\varphi^k_{i,j} = \bar{\varphi}_{i,j}$ for all i, j, where $\bar{\varphi}_{i,j}$ is the optimal control of the exact solution $\bar{V}_{i,j}$. Then (59) reduces to

$$\begin{aligned}
&\varepsilon^{k+1}_{i,j} \\
&= \frac{\Delta\tau\left[\alpha^{n+1}_{i,j}(\bar{\varphi}_{i,j})\varepsilon^k_{i-1,j} + \beta^{n+1}_{i,j}(\bar{\varphi}_{i,j})\varepsilon^k_{i+1,j} + \lambda_j[\mathcal{J}^h_j\varepsilon^k]_{i,j}\right]}{1+\Delta\tau\left(\alpha^{n+1}_{i,j}(\bar{\varphi}_{i,j})+\beta^{n+1}_{i,j}(\bar{\varphi}_{i,j})+c^{n+1}_{i,j}(\bar{\varphi}_{i,j})\right)}.
\end{aligned} \tag{60}$$

Given the coarse grid error $e^{n,H} = \widetilde{V}^{n,H} - V^{n,H}$, we first copy the error to the fine grid points that coincide with the coarse grid points, i.e.

$$e^{n,h}_{i,j} = e^{n,H}_{i,j}, \quad i=0,2,\ldots,N-1, \quad j=1,2,\ldots,N_m.$$

For interpolation at noncoarse grid points, $S^h_i, i = 1,3,\ldots,N-2$, $j = 1,2,\ldots,N_m$, we use (60), i.e.

$$\begin{aligned}
&e^{n,h}_{i,j} \\
&= \frac{\Delta\tau\left[\alpha^{n+1}_{i,j}(\bar{\varphi}_{i,j})e^{n,h}_{i-1,j} + \beta^{n+1}_{i,j}(\bar{\varphi}_{i,j})e^{n,h}_{i+1,j} + \lambda_j[\mathcal{J}^h_j e^{n,h}]_{i,j}\right]}{1+\Delta\tau\left(\alpha^{n+1}_{i,j}(\bar{\varphi}_{i,j})+\beta^{n+1}_{i,j}(\bar{\varphi}_{i,j})+c^{n+1}_{i,j}(\bar{\varphi}_{i,j})\right)}.
\end{aligned} \tag{61}$$

A major challenge in using (61) is that the exact solution is unknown and hence its optimal control $\bar{\varphi}$ is also unknown. We address this issue by approximating $\bar{\varphi}$ using the relaxation iteration, i.e. for every noncoarse grid point S^h_i, $i = 1,3,\ldots,N-2$, we set $\bar{\varphi}_{i,j}$ to be the optimal control which maximizes

$$\max_{\varphi\in\{0,1\}} \left\{ \frac{\Delta\tau\left[\alpha^{n+1}_{i,j}(\varphi)\widetilde{V}^{n,H}_{i-1,j} + \beta^{n+1}_{i,j}(\varphi)\widetilde{V}^{n,H}_{i+1,j} + d^{n+1}_{i,j}(\varphi) + \lambda_j[\mathcal{J}^h_j\widetilde{V}^{n,H}]_{i,j}\right] + V^{n,h}_{i,j}}{1+\Delta\tau\left(\alpha^{n+1}_{i,j}(\varphi_{i,j})+\beta^{n+1}_{i,j}(\varphi_{i,j})+c^{n+1}_{i,j}(\varphi_{i,j})\right)} \right\} \tag{62}$$

The regime switching terms $[\mathcal{J}^h_j\widetilde{V}^{n,H}]_{i,j}$ and $[\mathcal{J}^h_j e^{n,h}]_{i,j}$ are handled as described in section 4.2.1. The fine grid solution is then updated using the standard procedure as

$$\hat{V}^{n,h}_{i,j} = \widetilde{V}^{n,h}_{i,j} + e^{n,h}_{i,j}, \quad i=0,1,\ldots,N-1, \quad j=1,2,\ldots,N_m, \tag{63}$$

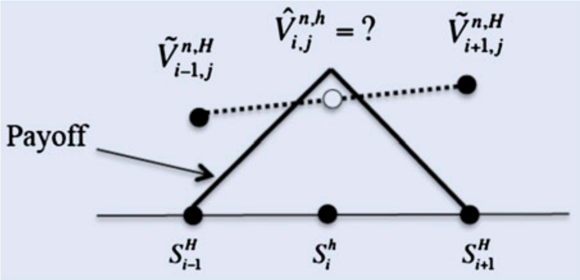

Figure 2. Illustration of a scenario where standard linear interpolation fails to capture the accurate optimal control.

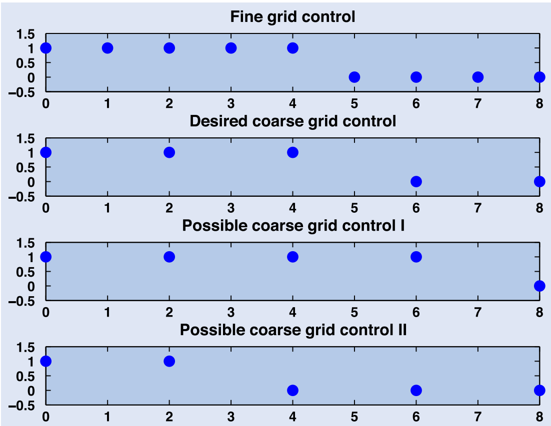

Figure 3. The different possibilities for restriction of control.

Similar derivation applies for interpolation of the error for (12) and (36) as well. We skip the details here.

4.3. *Restriction*

In the FAS scheme, we restrict the solution and the residual from the fine to the coarse grid. For HJB and HJBI problems, with jumps in the control, the control also needs to be carefully considered while performing restriction. Given the solution and its optimal control on the fine grid, we have to perform restriction such that the optimal control on the coarse grid is consistent with that on the fine grid. Let the fine grid control be given by the first plot of figure 3, with a jump in control between grid points 4 and 5. The coarse grid control may assume one of the three possible values shown in the last three plots of figure 3 based on the choice of the restriction operator. The coarse grid controls in the third and fourth plots of figure 3 are off by one grid point from that of the fine grid. The desired coarse grid control is the one which is consistent with the fine grid control as given in the second plot of figure 3. This can be justified by the FAS coarse grid right hand side F^H, which from algorithm 1 is given by

$$F^H = \mathcal{R}_r \widetilde{R}^{n,h} + N^H(\mathcal{R}_v \widetilde{V}^{n,h}). \tag{64}$$

Suppose $N^H(.)$ is a direct discretization on the coarse grid. Then $N^H(\mathcal{R}_v \widetilde{V}^{n,h})$ depends on the optimal control on the coarse grid, $\varphi^{*,H}$, which is computed from the restricted fine grid solution, $\mathcal{R}_v \widetilde{V}^{n,h}$. Whereas, the fine grid residual $\widetilde{R}^{n,h}$ depends on the optimal control on the fine grid, $\varphi^{*,h}$. When $\varphi^{*,h} \neq \varphi^{*,H}$ as shown in the third and fourth plots of figure 3, then the two components of F^H are inconsistent with each other, leading to slow convergence of the multigrid method (Han and Wan 2013). This inconsistency can occur due to improper choice of the restriction operator and is most visible near the jumps in control. Hence preserving consistency of control from the fine to the coarse grid is important.

There are two possible ways to achieve the consistency in control. One is to choose the restriction operator of the solution $\mathcal{R}_v$ such that the optimal control remains consistent. The other is to fix the controls on the coarse grid to be consistent with that on the fine grid and solve a local linear problem to update the coarse grid solution accordingly. The latter approach was used by Han and Wan (2013). We choose $\mathcal{R}_v$ to be the injection operator, which in a certain sense is constraint preserving as given in theorem 4.2.

THEOREM 4.2 *Let $\varphi^{*,h}$ and $\varphi^{*,H}$ be the optimal control on the fine and coarse grids respectively for problems discretized using penalty method. Let $\mathcal{R}_v$ be an injection operator to restrict the solution V^h, then $\varphi^{*,H} = \varphi^{*,h}$ for all the coarse grid points.*

Proof The optimal control for the HJB and HJBI systems in the penalty form (26) and (34) is determined by

$$\max_{\varphi\in\{0,1\}} \left[\frac{\varphi}{\epsilon}(V_i^* - V_{i,j}^n)\right]. \tag{65}$$

We note that the optimal control at the grid point (S_i, τ^n) depends only on $V_{i,j}^n$ for any regime j. By choosing $\mathcal{R}_v$ to be an injection operator, we have

$$V_{i,j}^{n,H} = V_{i,j}^{n,h}, \quad i = 0, 2, \ldots, N-1; \; j = 1, 2, \ldots, N_m. \tag{66}$$

Using (66), the solution at the coarse grid points is unchanged from that of the fine grid points. Therefore, the optimal control given by (65) on the coarse grid is consistent with that of the fine grid, i.e. $\varphi^{*,H} = \varphi^{*,h}$. □

The residual restriction operator $\mathcal{R}_r$ can either be a fully weighted restriction or an injection operator.

5. Theoretical analysis

In this section, we perform LFA to analyse the smoothing property of the weighted relaxation scheme and a two grid Fourier analysis to analyse the convergence behavior of the multigrid method. We also prove that the multigrid method using direct interpolation of the solution is monotone.

5.1. *Smoothing analysis*

We perform LFA for the weighted relaxation scheme to determine its efficiency as a smoother. LFA is a popular tool for the quantitative analysis of multigrid methods (Trottenberg *et al*., 2001). It is applied on linear discrete operators with constant coefficients, which are obtained by locally linearizing the nonlinear discrete operators with non-constant coefficients. LFA is based on grid functions of the form $\psi(\theta, x) = e^{i\theta x/h}$, where $i = \sqrt{-1}$.

We present a detailed smoothing analysis of the weighted relaxation scheme (54) for the HJB system resulting from American options under regime switching. We transform (26) into the log domain for simplicity of the analysis, which is a common practice in option pricing literature. Using $X = \log S$, we rewrite (26) in the log domain as

$$V_{j,\tau} = \max_{\varphi \in \{0,1\}} \left\{ \frac{\sigma_j^2}{2} V_{j,XX} + \left(r - \rho_j - \frac{\sigma_j^2}{2} \right) \times V_{j,X} - (r + \lambda_j + \frac{\varphi}{\epsilon}) V_j + \varphi \frac{V^*}{\epsilon} + \lambda_j \mathcal{J}_j^X V \right\},$$

where

$$\mathcal{J}_j^X V = \sum_{m \neq j} \frac{\lambda_{jm}}{\lambda_j} V_m(\xi_{jm} e^X, \tau).$$

The coefficients for (27) in the log domain are given by

$$a_j(\tau, \varphi) = \frac{\sigma_j^2}{2}, \quad b_j(\tau, \varphi) = r - \rho_j - \frac{\sigma_j^2}{2}, \\ c_j(\tau, \varphi) = r + \lambda_j + \frac{\varphi}{\epsilon}, \quad d_j(\tau, \varphi) = \frac{V^*}{\epsilon}. \tag{67}$$

We note that the coefficients on the log grid are independent of X, which is a desirable property for LFA.

Let $\bar{V}$ be the exact solution for time step $n+1$ and let the approximate solution after the kth smoothing iteration be $\widetilde{V}^k = \bar{V} + \varepsilon^k$, where ε^k is the error after the kth iteration. Since LFA is applied on linear operators, we will assume the optimal control for every grid point will not change from iteration to iteration. Let $\bar{\varphi}_l$ be the optimal control corresponding to the exact solution $\bar{V}$. Let $\varphi_l^k = \bar{\varphi}_l$ be the optimal control for $\widetilde{V}_l^k$ and let $\alpha_l^* = \alpha_l^{n+1}(\varphi_l^k)$, $\beta_l^* = \beta_l^{n+1}(\varphi_l^k)$, $c_l^* = c_l^{n+1}(\varphi_l^k)$ and $\kappa_l^* = \alpha_l^* + \beta_l^* + c_l^*$. Then using (53), we obtain the error equation as

$$\varepsilon^{k+1} = S^k \cdot \varepsilon^k,$$

where S^k is an $NN_m \times NN_m$ matrix. For $l = (j-1)N + i,\ i = 0, 1, \ldots N-1,\ j = 1, 2, \ldots, N_m$ and $l_m = (m-1)N + i_m,\ m = 1, 2, \ldots, N_m,\ m \neq j$, the elements of S^k are given by

$$S_{l,l-1}^k = \frac{\Delta\tau\alpha_l^*}{1 + \Delta\tau\kappa_l^*}, \quad S_{l,l+1}^k = \frac{\Delta\tau\beta_l^*}{1 + \Delta\tau\kappa_l^*}, \\ S_{l,l_m}^k = \frac{\Delta\tau w_m \lambda_{jm}^*}{1 + \Delta\tau\kappa_l^*}, \quad S_{l,l_m+1}^k = \frac{\Delta\tau(1 - w_m)\lambda_{jm}^*}{1 + \Delta\tau\kappa_l^*}, \tag{68} \\ S_{l,l'}^k = 0, \quad l' \neq l-1, l+1, l_m, l_m+1.$$

Using (68), we obtain the symbol of the smoothing operator for the relaxation scheme (53) as

$$\widetilde{S}_l^k(\theta) = \frac{\Delta\tau \left[\alpha_l^* e^{-i\theta} + \beta_l^* e^{i\theta} + \sum_{m \neq j} \lambda_{jm}^* \left(w_m e^{i(l_m - l)\theta} + (1 - w_m) e^{i(l_m - l + 1)\theta} \right) \right]}{1 + \Delta\tau\kappa_l^*}.$$

Since we use the weighted relaxation scheme (54) for smoothing, the symbol of its smoothing operator is given by

$$\widetilde{S}_l^k(\theta, \omega) = (1 - \omega) \\ + \omega \frac{\Delta\tau \left[\alpha_l^* e^{-i\theta} + \beta_l^* e^{i\theta} + \sum_{m \neq j} \lambda_{jm}^* \left(w_m e^{i(l_m - l)\theta} + (1 - w_m) e^{i(l_m - l + 1)\theta} \right) \right]}{1 + \Delta\tau\kappa_l^*} \tag{69}$$

We are interested in the smoothing effect, i.e. the reduction of high frequency error components. Hence, the smoothing factor $\mu(S_l^k)$ is defined as

$$\mu(S_l^k) = \sup \left\{ |\widetilde{S}_l^k(\theta, \omega)| : \theta \in [-\pi, \pi) \setminus \left[-\frac{\pi}{2}, \frac{\pi}{2} \right) \right\}.$$

Since generating a useful analytical expression for $\mu(S_l^k)$ is very complicated, we consider specific frequency values and analyse the behavior of the smoother. In addition, we demonstrate the efficiency of the smoother for the entire frequency range by plotting $|\widetilde{S}_l^k(\theta, \omega)|$ for $\theta \in [-\pi, \pi)$ and different ω.

Consider a high frequency point $\theta = -\pi$. We note that $l_m - l$ is an integer. We then have $\cos(z\theta) = -1$ if z is odd and $\cos(z\theta) = 1$ if z is even and $\sin(z\theta) = 0$ for all z. Using these values in (69), we have

$$\widetilde{S}_l^k(\theta, \omega) \\ = \frac{(1 - \omega)[1 + \Delta\tau\kappa_l^*] + \omega\Delta\tau \left[-(\alpha_l^* + \beta_l^*) + \sum_{m \neq j} \lambda_{jm}^* w_m \right]}{1 + \Delta\tau\kappa_l^*}, \tag{70}$$

where

$$w_m^* = \begin{cases} 2w_m - 1 & \text{if } l_m - l is even, \\ 1 - 2w_m & \text{if } l_m - l is odd, \end{cases}$$

with $|w_m^*| \leq 1$ since $w_m \in [0, 1]$. Adding and subtracting $\omega\Delta\tau c_l^*$ from the numerator of (70), we get

$$\widetilde{S}_l^k(\theta, \omega) \\ = \frac{(1 - \omega)[1 + \Delta\tau\kappa_l^*] + \omega\Delta\tau \left[-\left(\alpha_l^* + \beta_l^* + c_l^* \right) + c_l^* + \sum_{m \neq j} \lambda_{jm} w_m^* \right]}{1 + \Delta\tau\kappa_l^*} \\ = \frac{(1 - \omega)[1 + \Delta\tau\kappa_l^*] + \omega\Delta\tau \left[-\kappa_l^* + c_l^* + \sum_{m \neq j} \lambda_{jm} w_m^* \right]}{1 + \Delta\tau\kappa_l^*} \\ = \frac{(1 - 2\omega)[1 + \Delta\tau\kappa_l^*] + \omega + \omega\Delta\tau \left[c_l^* + \sum_{m \neq j} \lambda_{jm} w_m^* \right]}{1 + \Delta\tau\kappa_l^*} \\ = (1 - 2\omega) + \frac{\omega \left[1 + \Delta\tau \left(c_l^* + \sum_{m \neq j} \lambda_{jm} w_m^* \right) \right]}{1 + \Delta\tau\kappa_l^*} \tag{71}$$

Using (B10), (B11) and (67) in (71), we have

$$\widetilde{S}_l^k(\theta,\omega) = (1-2\omega) + \frac{\omega\left[1+\Delta\tau\left(r+\lambda_j+\frac{\varphi_l^k}{\epsilon}+\sum_{m\neq j}\lambda_{jm}w_m^*\right)\right]}{1+\Delta\tau\left(\frac{2\sigma_j^2}{h^2}+r+\lambda_j+\frac{\varphi_l^k}{\epsilon}\right)}. \quad (72)$$

We consider a three regime model with the following parameters:

$$r = 0.02,\ \ \Delta\tau = 10^{-3},\ \ \epsilon = 10^{-4}\Delta\tau,$$
$$\sigma = \begin{bmatrix} 0.2 \\ 0.15 \\ 0.3 \end{bmatrix},\ \lambda = \begin{bmatrix} -3.2 & 0.2 & 3.0 \\ 1.0 & -1.08 & 0.08 \\ 3.0 & 0.2 & -3.2 \end{bmatrix},$$
$$\xi = \begin{bmatrix} 1.0 & 0.9 & 1.1 \\ 1.2 & 1.0 & 1.3 \\ 0.95 & 0.8 & 1.0 \end{bmatrix}. \quad (73)$$

As $h \to 0$ and for the parameters given in (73), equation (72) reduces to $\widetilde{S}_l^k(\theta,\omega) = 1-2\omega$. For convergence, we should have $|\widetilde{S}_l^k(\theta,\omega)| < 1$, which is satisfied when $\omega \in (0,1)$. The smoothing factor is minimized when $\omega = 0.5$. Similarly, for low and medium frequency components, $|\widetilde{S}_l^k(\theta,\omega)| < 1$ when $\omega \in (0,2)$. Optimal smoothing is obtained with the ω which satisfies

$$\min_{\omega}\sup_{\theta}\left\{|\widetilde{S}_l^k(\theta,\omega)| : \theta \in [-\pi,\pi)\backslash\left[-\frac{\pi}{2},\frac{\pi}{2}\right)\right\}.$$

Since analytically evaluating this expression is very complicated, we numerically determine the optimal ω by plotting $|\widetilde{S}_l^k(\theta,\omega)|$ against θ for different $\omega \in (0,1]$ and for different grid sizes as $h \to 0$. Figure 4(a) shows the smoothing factor when $\varphi = 0$ and $h = 0.0125$ and for different $\omega \in (0,1]$. The relaxation scheme is convergent for the entire frequency range. Furthermore, $\omega = 0.67$ and 0.8 have small smoothing factors for the high frequency range. We now analyse the smoothing property for $\omega = 0.5, 0.67$ and 0.8 as $h \to 0$. We are interested in $\omega = 0.5$ as it minimizes the smoothing factor at $\theta = \pm\pi$ as $h \to 0$. Figure 4(b)–(d) shows the smoothing factors for $\omega = 0.5, 0.67$ and 0.8, respectively. Among the different choices, $\omega = 0.67$ has small smoothing factor, i.e. for all $\theta \in [-\pi,\pi)\backslash[-\frac{\pi}{2},\frac{\pi}{2})$ and different h, $|\widetilde{S}_l^k(\theta,0.67)| \le 1/3$. The smoothing factor has the same upper bound with $\omega = 0.67$ for $\varphi = 1$ as well. Therefore, we use $\omega = 0.67$ as the damping parameter.

Similar results are obtained for the weighted relaxation schemes for the dynamic Bertrand oligopoly and the HJBI systems. Therefore, we use $\omega = 0.67$ for all model problems.

5.2. Two-grid analysis

We apply LFA to the two grid operator to study the convergence properties of the multigrid method. An analysis of the two grid method in general provides sufficient insight into the behavior of multigrid methods. Let ε^k be the error after the kth two grid iteration, then

$$\varepsilon^{k+1} = M_h^H\varepsilon^k,$$

where M_h^H is the two grid operator,

$$M_h^H = S^{h,\nu_1}K_h^H S^{h,\nu_2},$$

where S^h is the smoothing operator and ν_1 and ν_2 are the number of pre and post smoothing iterations, respectively. K_h^H is the coarse grid correction operator given by

$$K_h^H = I^h - \mathcal{P}(L^H)^{-1}\mathcal{R}L^h,$$

where I^h is an identity matrix, L^h is the fine grid discrete operator, L^H is the coarse grid operator, $\mathcal{P}$ and $\mathcal{R}$ are the interpolation and restriction operators. The spectral radius of M_h^H gives an indication of the asymptotic rate of multigrid convergence. The convergence factors of M_h^H are computed by analyzing how the operators S^h, L^h, $\mathcal{R}$, L^H and $\mathcal{P}$ act on the Fourier components $\psi(\theta,x) = e^{\mathrm{i}\theta x/h}$. Let $x^h = \{ih, i \in \mathbb{Z}\}$ be the infinite fine grid and $x^H = \{iH, i \in \mathbb{Z}\}$ be the corresponding coarse grid with $H = 2h$.

The Fourier space $E^h = \mathrm{span}\{e^{\mathrm{i}\theta x/h} : \theta \in (-\pi,\pi]\}$ contains any infinite grid function on x^h (Trottenberg *et al*. 2001). The current approximation $V^{k,h}$ and the error $\varepsilon^{k,h}$ can be represented as linear combinations of the basis functions $e^{\mathrm{i}\theta x/h} \in E^h$. We note that E^h can be divided into two dimensional subspaces, also called the harmonics:

$$E^{\theta,h} = \mathrm{span}\left[\psi(\theta^\gamma,x) : \gamma \in \{0,1\}\right], x \in x^h,$$

where

$$\theta^0 \in (-\pi/2,\pi/2],\ \ \theta^1 = \theta^0 - \mathrm{sign}\left(\theta^0\right)\pi.$$

For an arbitrary $\theta \in (-\pi/2,\pi/2]\backslash\{\theta : \hat{L}^H(2\theta^0) = 0\}$, $E^{\theta,h}$ is invariant under the coarse grid correction operator K_h^H and the smoother,

$$K_h^H : E^{\theta,h} \to E^{\theta,h},\ \ S^h : E^{\theta,h} \to E^{\theta,h}.$$

Hence M_h^H is orthogonally equivalent to a 2×2 block matrix given by

$$\hat{M}_h^H(\theta,\omega) = \hat{S}^{h,\nu_2}(\theta,\omega)\,\hat{K}_h^H(\theta)\,\hat{S}^{h,\nu_1}(\theta,\omega), \quad (74)$$

where

$$\hat{K}_h^H(\theta) = I^h - \hat{\mathcal{P}}(\theta)\left(\hat{L}^H(2\theta)\right)^{-1}\hat{\mathcal{R}}(\theta)\,\hat{L}^h(\theta),$$

where $\hat{L}^h$, $\hat{\mathcal{R}}$, $\hat{L}^H$, $\hat{\mathcal{P}}$ and $\hat{\mathcal{S}}^h$ are matrices built with the Fourier symbols of their respective multigrid operators. We present these matrices for the case of American options under regime switching (32).

5.2.1. Discrete fine grid operator $\hat{L}^h(\theta)$.

$\hat{L}^h(\theta)$ is a 2×2 matrix given by

$$\hat{L}^h(\theta) = \begin{pmatrix} \tilde{L}^h(\theta^0) & \\ & \tilde{L}^h(\theta^1) \end{pmatrix},$$

where

$$\tilde{L}^h(\theta^\gamma) = -\Delta\tau\alpha_l^{*,h}e^{-\mathrm{i}\theta^\gamma} + \left(1+\Delta\tau\kappa_l^{*,h}\right) - \Delta\tau\beta_l^{*,h}e^{\mathrm{i}\theta^\gamma} - \Delta\tau\sum_{m\neq j}\lambda_{jm}\left(w_m e^{\mathrm{i}(l_m-l)\theta^\gamma} + (1-w_m)e^{\mathrm{i}(l_m-l+1)\theta^\gamma}\right) \quad (75)$$

5.2.2. Restriction operator $\hat{\mathcal{R}}(\theta)$.

The restriction operator is denoted by the following 1×2 matrix

$$\hat{\mathcal{R}}(\theta) = \left[\widetilde{\mathcal{R}}(\theta^0)\ \widetilde{\mathcal{R}}(\theta^1)\right]. \quad (76)$$

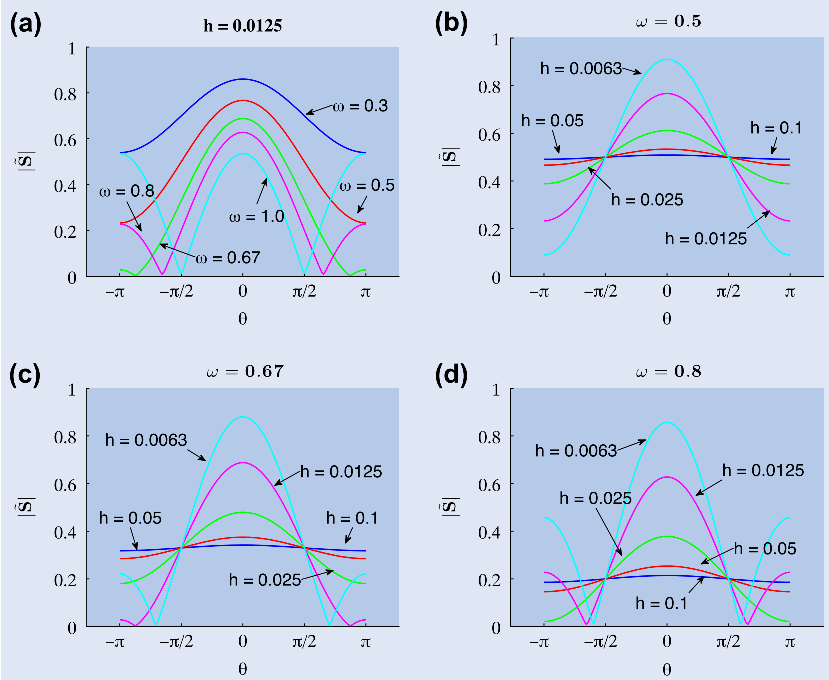

Figure 4. Smoothing factor for (54) with $\varphi = 0$. (a) $h = 0.0125$ and $\omega \in (0, 1]$, (b) $\omega = 0.5$, (c) $\omega = 0.67$, (d) $\omega = 0.8$.

For injection the Fourier symbols $\widetilde{\mathcal{R}}(\theta^{\gamma})$ for all frequencies are 1.

5.2.3. Interpolation operator $\hat{\mathcal{P}}$. The interpolation matrix is in general given by

$$\hat{\mathcal{P}}(\theta) = \begin{pmatrix} \widetilde{\mathcal{P}}(\theta^0) \\ \widetilde{\mathcal{P}}(\theta^1) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+\delta \\ 1-\delta \end{pmatrix}$$

where for (32), we have

$$\delta = \frac{\Delta\tau\left[\alpha_l^{*,h} e^{-i\theta^0} + \beta_l^{*,h} e^{i\theta^0} + \sum_{m\neq j}\left(w_m e^{\mathrm{i}(l_m-l)\theta^0} + (1-w_m)e^{\mathrm{i}(l_m-l+1)\theta^0}\right)\right]}{1+\Delta\tau\kappa_l^{*,h}}. \tag{77}$$

5.2.4. Smoothing operator $\hat{S}^h$. The smoothing operator is a 2×2 matrix,

$$\hat{S}^h(\theta,\omega) = \begin{pmatrix} \widetilde{S}^h(\theta^0,\omega) & \\ & \widetilde{S}^h(\theta^1,\omega) \end{pmatrix},$$

where $\widetilde{S}^h$ is given by (69).

5.2.5. Coarse grid discrete operator $\hat{L}^H(2\theta)$. $\hat{L}^H(2\theta)$ is a 1×1 matrix whose symbol is given by

$$\begin{aligned}
&\tilde{L}^H(2\theta) \\
&= -\Delta\tau\alpha_l^{*,H} e^{-i2\theta} + \left(1+\Delta\tau\kappa_l^{*,H}\right) - \Delta\tau\beta_l^{*,H} e^{i2\theta} \\
&\quad - \Delta\tau\sum_{m\neq j}\lambda_{jm}\left(w_m e^{\mathrm{i}(l_m-l)2\theta} + (1-w_m)e^{\mathrm{i}(l_m-l+1)2\theta}\right).
\end{aligned} \tag{78}$$

We then construct $\hat{M}_h^H(\theta)$ using (69), (75), (76), (77) and (78). We can now determine the spectral radius of M_h^H by calculating the spectral radius of $\hat{M}_h^H(\theta)$:

$$\rho\left(M_h^H\right) = \max_{\theta\in(-\pi/2,\pi/2]} \rho\left(\hat{M}_h^H(\theta)\right).$$

We recall that Fourier analysis is exact only for linear operators with constant (or frozen) coefficients (Trottenberg *et al.* 2001). Therefore, we fix the parameters as given in (73) and $\varphi = 0$. Figure 5(a) shows the plot of $\rho(\hat{M}_h^H(\theta))$ against θ for different h for the HJB system (32). As $h \to 0$, $\rho\left(M_h^H\right) \to 0.12$, which is a very satisfactory convergence rate. Figure 5(b) shows the convergence rates for HJBI systems. Similar results are observed for $\varphi = 1$, the details are omitted here.

5.3. *Monotonicity*

Monotonicity properties often result in smooth and fast convergence of the multigrid solution (Amarala and Wan 2013). In this section, we present detailed analysis of the monotonicity property for the HJB system resulting from American options under regime switching (25). We note that (25) can also be formally stated in a linear complementarity form:

$$\begin{aligned}
V_{j,\tau} - \mathcal{L}_j V_j - \lambda_j \mathcal{J}_j V &\geq 0, \\
V_j - V^* &\geq 0, \\
\left(V_{j,\tau} - \mathcal{L}_j V_j - \lambda_j \mathcal{J}_j V\right)\left(V_j - V^*\right) &= 0.
\end{aligned} \tag{79}$$

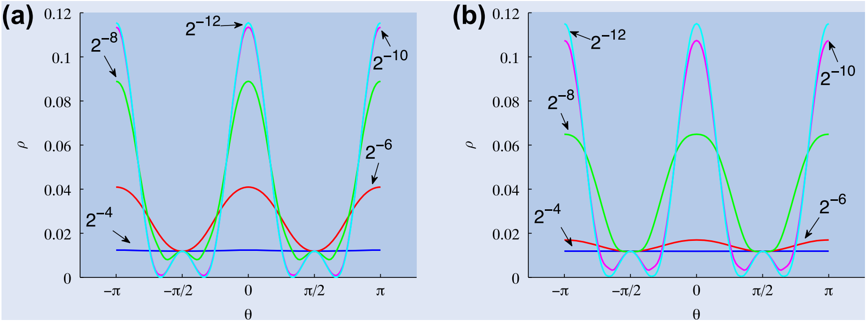

Figure 5. Convergence rate of the two-grid algorithm for different grid sizes and $\varphi = 0$. (a) HJB system and (b) HJBI system.

Table 1. Multigrid convergence for the HJB system (12).

$N_1 \times N_2$	65×65	129×129	257×257	513×513	1025×1025
$\gamma = 0.1$	3	3	4	5	9
$\gamma = 0.5$	3	3	4	6	9

We use the linear complementarity formulation to define the monotonicity property.

Definition 7 (Holtz and Kunoth 2007) A multigrid method for the linear complimentary problem (79) is monotone if, as $\epsilon \to 0$, the interpolated fine grid solution $\hat{V}^{n,h}$ satisfies the constraint

$$\hat{V}^{n,h} - V^{*,h} \geq 0. \tag{80}$$

THEOREM 5.1 *The multigrid method using direct interpolation for the solution for the HJB system (32) is monotone as $\epsilon \to 0$.*

Proof Given $V^{n,h}$, let $\hat{V}^{n,h}$ be the solution after interpolation, which is given by (57) and (58). The updated coarse grid solution $\widetilde{V}^{n,H}_{i,j}$ is obtained by solving the following coarse grid problem:

$$\begin{aligned}\widetilde{V}^{n,H}_{i,j} - \Delta\tau \Big[& \alpha^{n,H}_{i,j} \widetilde{V}^{n,H}_{i-1,j} + \beta^{n,H}_{i,j} \widetilde{V}^{n,H}_{i+1,j} \\ & - \left(\alpha^{n,H}_{i,j} + \beta^{n,H}_{i,j} + c^{n,H}_{i,j}\right) + \lambda_j \left[\mathcal{J}^H_j \widetilde{V}^{n,H}\right]_{i,j} \\ & + \max_{\varphi\in\{0,1\}} \left(\frac{\varphi^{n,H}_{i,j}}{\epsilon} \left(V^{*,H}_{i,j} - \widetilde{V}^{n,H}_{i,j} \right) \right) \Big] \\ = F^{n,H}_{i,j}, \qquad & i = 0, 2, \ldots, N-3,\end{aligned} \tag{81}$$

where $F^{n,H}_{i,j}$ is the coarse grid right hand side (See algorithm 1) and $V^{*,H}_{i,j}$ is the payoff function on the coarse grid, which is given by

$$V^{*,H}_{i,j} = V^{*,h}_{i,j}, \quad i = 0, 2, \ldots, N-3. \tag{82}$$

Suppose we use relaxation scheme to solve the coarse grid problem. Let $(\widetilde{V}^{n,H})^0 = V^{n,H}$, then the relaxation iteration for (81) is given by

$$\left(\widetilde{V}^{n,H}_{i,j}\right)^{k+1} = \max_{\varphi\in\{0,1\}} \left[\frac{\Delta\tau \left(\alpha^{n,H}_{i,j} \left(\widetilde{V}^{n,H}_{i-1,j}\right)^k + \beta^{n,H}_{i,j} \left(\widetilde{V}^{n,H}_{i+1,j}\right)^k + d^{n,H}_{i,j} + \left[\mathcal{J}^H_j \left(\widetilde{V}^{n,H}\right)^k\right]_{i,j} \right) + F^{n,H}_{i,j}}{1 + \Delta\tau \left(\alpha^{n,H}_{i,j} + \beta^{n,H}_{i,j} + c^{n,H}_{i,j} \right)} \right]$$
$$i = 0, 2, \ldots, N-3; \quad j = 1, 2, \ldots, N_m; \quad k = 0, 1, \ldots, \text{until convergence}. \tag{83}$$

Using $c^{n,H}_{i,j}$ and $d^{n,H}_{i,j}$ from (28) in (83), we obtain

$$\begin{aligned}\left(\widetilde{V}^{n,H}_{i,j}\right)^{k+1} = \max \Bigg[& \frac{\Delta\tau \left(\alpha^{n,H}_{i,j} \left(\widetilde{V}^{n,H}_{i-1,j}\right)^k + \beta^{n,H}_{i,j} \left(\widetilde{V}^{n,H}_{i+1,j}\right)^k + \left[\mathcal{J}^H_j \left(\widetilde{V}^{n,H}\right)^k\right]_{i,j} \right) + F^{n,H}_{i,j}}{1 + \Delta\tau \left(\alpha^{n,H}_{i} + \beta^{n,H}_{i} + r + \lambda_j \right)}, \\ & \frac{\Delta\tau \left(\alpha^{n,H}_{i,j} \left(\widetilde{V}^{n,H}_{i-1,j}\right)^k + \beta^{n,H}_{i,j} \left(\widetilde{V}^{n,H}_{i+1,j}\right)^k + \frac{1}{\epsilon} V^{*,H}_{i,j} + \left[\mathcal{J}^H_j \left(\widetilde{V}^{n,H}\right)^k\right]_{i,j} \right) + F^{n,H}_{i,j}}{1 + \Delta\tau \left(\alpha^{n,H}_{i,j} + \beta^{n,H}_{i,j} + r + \lambda_j + \frac{1}{\epsilon} \right)} \Bigg]\end{aligned}$$

As $\epsilon \to 0$, we have

$$\left(\widetilde{V}_{i,j}^{n,H}\right)^{k+1} = \max\left[\frac{\Delta\tau\left(\alpha_{i,j}^{n,H}\left(\widetilde{V}_{i-1,j}^{n,H}\right)^k + \beta_{i,j}^{n,H}\left(\widetilde{V}_{i+1,j}^{n,H}\right)^k + \left[\mathcal{J}_j^H\left(\widetilde{V}^{n,H}\right)^k\right]_{i,j}\right) + F_{i,j}^{n,H}}{1+\Delta\tau\left(\alpha_i^{n,H}+\beta_i^{n,H}+r+\lambda_j\right)},\ V_{i,j}^{*,H}\right] \quad (84)$$

From (84), it is clear that

$$\left(\widetilde{V}_{i,j}^{n,H}\right)^{k+1} \geq V_{i,j}^{*,H}, \quad i = 0, 2, \dots, N-3 \quad (85)$$

for any k. From (57), (82) and (85), we have

$$\hat{V}_{i,j}^{n,h} \geq V_{i,j}^{*,h}, \quad i = 0, 2, \dots, N-3. \quad (86)$$

For the noncoarse grid points, the interpolated fine grid solution is given by (58), which as $\epsilon \to 0$ becomes:

$$\hat{V}_{i,j}^{n,h} = \max\left[\frac{\Delta\tau\left(\alpha_{i,j}^{n,h}\hat{V}_{i-1,j}^{n,h} + \beta_{i,j}^{n,h}\hat{V}_{i+1,j}^{n,h} + \left[\mathcal{J}_j^h\hat{V}^{n,h}\right]_{i,j}\right) + V_{i,j}^{n,h}}{1+\Delta\tau\left(\alpha_{i,j}^{n,h}+\beta_{i,j}^{n,h}+r+\lambda_j\right)},\ V_{i,j}^{*,h}\right], \quad i = 1, 3, \dots, N-2,$$

which results in

$$\hat{V}_{i,j}^{n,h} \geq V_{i,j}^{*,h}, \qquad i = 1, 3, \dots, N-2. \quad (87)$$

From (86), (87) and definition 7, the two grid method using direct interpolation for the solution is monotone. Using induction, we can prove that a L-grid method ($L \geq 2$) is monotone. □

6. Numerical results

We test our multigrid method with two pre and post-smoothing steps on the model problems presented in section 2. We present the results using direct interpolation of the solution (58). The convergence using interpolation of the error (63) is very similar and hence we omit the details here.

Example 6.1 2D HJB System: Dynamic Bertrand Duopoly (12).

We use the parameters $T = 0.25$, $r = 1$, $\sigma_1 = \sigma_2 = 0.6$, $\rho = 0.1$, $\gamma = 0.1$, $\kappa = 6$, $\eta = 1$, $\Delta\tau = 0.025$ and a convergence tolerance of 10^{-6}. The two dimensional grid is coarsened using the multiple coarsening strategy (Amarala and Wan 2013). We use multiple grids such that the coarsest grid has 17×17 grid points. Since the convergence is similar in each time step, we only show the convergence results for the very first time step in table 1. The relaxation scheme alone takes 877 iterations for the grid size of 1025×1025, whereas our multigrid method converges in only 9 iterations.

Example 6.2 HJB System: American Options under Regime Switching (32).

We consider a three regime model for evaluating the multigrid method. The transition probabilities λ, jump amplitudes ξ and the volatilities are given in (88). The other parameters are given in table 2. We consider American options with three different payoffs: put, straddle and butterfly. Numerical tests are performed on a uniform log grid. Multiple grids are used for different grid sizes such that the coarsest grid had only 9 grid points. The convergence results for the very first timestep for different grid sizes are given in table 3. The multigrid method converges in a very small number of iterations irrespective of the grid size.

$$\lambda = \begin{bmatrix} -3.2 & 0.2 & 3.0 \\ 1.0 & -1.08 & 0.08 \\ 3.0 & 0.2 & -3.2 \end{bmatrix}; \xi = \begin{bmatrix} 1.0 & 0.9 & 1.1 \\ 1.2 & 1.0 & 1.3 \\ 0.95 & 0.8 & 1.0 \end{bmatrix}; \sigma = \begin{bmatrix} 0.2 \\ 0.15 \\ 0.3 \end{bmatrix} \quad (88)$$

Example 6.3 HJBI System: American Option and Stock Borrowing Fees (36).

We use the parameters given in table 2 and (88) for the HJBI systems under a three regime model. The borrowing rate $r_b = 0.05$, lending rate $r_l = 0.03$ and the stock borrowing fee $r_f = 0.004$. The convergence results for different initial conditions are given in table 4. Similar to the case of HJB systems, the multigrid method for HJBI system also converges in a very small number of iterations independent of the grid size.

Table 2. Parameters used for American options.

Expiry time, T	0.5
Exercise	American
Strike, K	100
Butterfly parameters, K_1, K_2	90, 110
Risk free interest rate, r	0.02
Time step, $\Delta\tau$	10^{-3}
Penalty parameter, ϵ	$10^{-4}\Delta\tau$
Convergence tolerance	10^{-6}

Table 3. Multigrid convergence for the HJB system (32).

N	65	129	257	513	1025	2049
Put	2	3	3	3	3	4
Straddle	3	3	3	3	3	4
Butterfly	3	3	3	3	3	5

Table 4. Multigrid convergence for the HJBI system (36).

N	65	129	257	513	1025	2049
Put	2	3	3	3	3	4
Straddle	3	3	3	3	3	4
Butterfly	3	3	3	3	4	5

7. Conclusion

We constructed fully implicit, consistent, unconditionally l_∞ stable and monotone discretization schemes that converge to the viscosity solution for the HJB PDE resulting from dynamic Bertrand monopoly and the two-dimensional systems of nonlinear HJB PDEs from duopoly. We developed multigrid methods for discrete systems of nonlinear HJB and HJBI PDEs associated with dynamic Bertrand duopoly and regime switching applications. A weighted relaxation scheme is used as the smoother. We show that the smoother is convergent for both HJB and HJBI systems, in contrast to policy iteration which may not converge for HJBI problems. A smoothing analysis shows that the weighted relaxation scheme with $\omega = 0.67$ effectively damps the high frequency error components. We choose injection for restriction, which preserves the consistency of the control from the fine to the coarse grids. We introduce new interpolation techniques which efficiently capture the optimal control in the presence of jumps. We analyze the convergence behavior of the multigrid method through a two grid Fourier analysis, which gives a convergence factor as low as 0.12 as $h \to 0$. Numerical tests on practical examples show that the multigrid method converges in a very small number of iterations irrespective of the grid size.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Amarala, S. and Wan, J.W., Multigrid methods for systems of hyperbolic conservation laws. *SIAM J. Multiscale Model. Simul.*, 2013, **11**, 586–614.

Ayache, E., Equity-credit problem. In *Encyclopedia of Quantitative Finance*, edited by R. Cont, pp. 571–575, 2010 (Wiley: New York).

Barles, G., Convergence of numerical schemes for degenerate parabolic equations arising in finance. In *Numerical Methods in Finance*, edited by L.C.G Rogers and D. Talay, pp. 1–21, 1997 (Cambridge University Press: Cambridge).

Barles, G. and Rouy, E., A strong comparison result for the Bellman equation arising in stochastic exit time control problems and applications. *Commun. Partial Differ. Equ.*, 1998, **23**, 1945–2033.

Barles, G. and Souganidis, P., Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 1991, **4**, 271–283.

Barles, G. and Jakobsen, E.R., Error bounds for monotone approximation schemes for Hamilton–Jacobi–Bellman equations. *SIAM J. Numer. Anal.*, 2005, **43**, 540–558.

Bellman, R., *Introduction to the Mathematical Theory of Control Processes: Nonlinear processes*. Mathematics in Science and Engineering, 1971 (Academic Press: New York).

Bergman, Y., Option pricing with differential interest rates. *Rev. Financ. Stud.*, 1995, **8**, 475–500.

Bertrand, J., Théorie mathématique de la richesse sociale [Mathematical theory of social wealth]. *J. Savants*, 1883, **67**, 499–508.

Bokanowski, O., Maroso, S. and Zidani, H., Some convergence results for Howard's algorithm. *SIAM J. Numer. Anal.*, 2009, **47**, 3001–3026.

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

Chaumont, S., A strong comparison result for viscosity solutions to Hamilton–Jacobi–Bellman equations with Dirichlet conditions on a non-smooth boundary. Preprint, Institute Elie Cartan, Universitè Nancy I, 2004.

Clift, S.S. and Forsyth, P.A., Numerical solution of two asset jump diffusion models for option valuation. *Appl. Numer. Math.*, 2008, **58**, 743–782.

Cournot, A., *Recherches sur les Principes Mathématique de la Théorie des Richesses* [Researches into the Mathematical Principles of the Theory of Wealth], 1838 (Hachette: Paris). English translation by N.T. Bacon, published in Economic Classics, Macmillan, 1897, and reprinted in 1960 by Augustus M. Kelley, 1838.

Crepey, S., About pricing equations in finance. In *Proceedings of the Paris-Princeton Lectures on Mathematical Finance*, edited by A. Carmona, pp. 63–203, 2010 (Berlin: Springer).

Dockner, E., Jorgensen, S., Long, N.V. and Sorger, G., *Differential Games in Economics and Management Science*, 2000 (Cambridge University Press: New York).

Forsyth, P.A., Numerical computation for financial modelling. Course notes, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada, 2012.

Forsyth, P.A. and Labahn, G., Numerical methods for controlled Hamilton–Jacobi–Bellman PDEs in finance. *J. Comput. Finance*, 2007, **11**, 1–44.

Forsyth, P.A. and Vetzal, K.R., Quadratic convergence of a penalty method for valuing American options. *SIAM J. Sci. Comput.*, 2002, **23**, 2095–2122.

Han, D. and Wan, J.W., Multigrid methods for second order Hamilton–Jacobi–Bellman and Hamilton–Jacobi–Bellman-Isaacs equations. *SIAM J. Sci. Comput.*, 2013, **35**, S323–S344.

Holtz, M. and Kunoth, A., B-Spline based monotone multigrid methods. *SIAM J. Numer. Anal.*, 2007, **45**, 1175–1199.

Hoppe, R.H., Multi-grid methods for Hamilton–Jacobi–Bellman equations. *Numer. Math.*, 1986, **49**, 239–254.

Howard, R.A., *Dynamic Programming and Markov Process*, 1960 (MIT Press: Cambridge, MA).

Huang, Y., Forsyth, P. and Labahn, G., Methods for pricing American options under regime switching. *SIAM J. Sci. Comput.*, 2011, **33**, 2144–2168.

Huang, Y. and Forsyth, P.A., Analysis of a penalty method for pricing a guaranteed minimum withdrawal benefit (GMWB). *IMA J. Numer. Anal.*, 2012, **32**, 320–351.

Ishii, H. and Koike, S., Viscosity solutions for monotone systems of second order elliptic PDEs. *Commun. Partial Differ. Equ.*, 1991, **16**, 1095–1128.

Kennedy, J., Hedging contingent claims in markets with jumps. PhD Thesis, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada, 2007.

Ledvina, A. and Sircar, R., Dynamic Bertrand oligopoly. *Appl. Math. Optim.*, 2011, **63**, 11–44.

Lions, P. and Mercier, B., Approximation numerique des equations Hamilton–Jacobi–Bellman [Numerical approximation of the Hamilton-Jacobi-Bellman equations]. *RAIRO – Anal. Numer.*, 1980, **14**, 369–393.

Oosterlee, C.W., On multigrid for linear complimentary problems with application to American-style options. *Electron. Trans. Numer. Anal.*, 2003, **15**, 165–185.

Pooley, D.M., Forsyth, P.A. and Vetzal, K.R., Numerical convergence properties of option pricing PDEs with uncertain volatility. *IMA J. Numer. Anal.*, 2003, **23**, 241–267.

Reisinger, C. and Forsyth, P.A., Piecewise constant policy approximations to Hamilton–Jacobi–Bellman equations. Techinical Report, 2015.

Trottenberg, U., Oosterlee, C.W. and Schuller, A., *Multigrid*, 2001 (Academic Press: San Diego, CA).

Tse, S.T., Numerical methods for optimal trade execution. PhD Thesis, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada, 2012.

Wal, J., Discounted Markov games: Generalized policy iteration method. *J. Optim. Theory Appl.*, 1978, **25**, 125–138.

Wang, J. and Forsyth, P.A., Maximal use of central differencing for Hamilton–Jacobi–Bellman PDEs in finance. *SIAM J. Numer. Anal.*, 2008, **46**, 1580–1601.

Yao, D., Zhang, Q. and Zhou, X.Y., A regime switching model for European options. In *Stochastic Processes, Optimization and Control Theory: Applications in Financial Engineering, Queueing Networks and Manufacturing Systems*, edited by H. Yan, G. Yin and Q. Zhang, Vol. 94, pp. 281–300, 2006 (Springer International Series in Operations Research and Management Science: New York).

Zvan, R., Forsyth, P.A. and Vetzal, K.R., A finite volume approach for contingent claims valuation. *IMA J. Numer. Anal.*, 2001, **21**, 703–731.

Appendix 1. Dynamic Bertrand oligopoly

A.1. Monopoly problem: discrete equation coefficients

The discrete coefficients α_i and β_i for (6) are presented here. Standard central differencing for all derivatives of (4) results in

$$\alpha_i^{cent} = \frac{\sigma^2}{2\Delta x^2} - \frac{\kappa - p}{2\eta\Delta x},$$
$$\beta_i^{cent} = \frac{\sigma^2}{2\Delta x^2} + \frac{\kappa - p}{2\eta\Delta x}.$$

Using upstream differencing for first order terms and central for second order terms, we have

$$\alpha_i^{ups} = \frac{\sigma^2}{2\Delta x^2} + \max\left\{0, -\frac{\kappa - p}{2\eta\Delta x}\right\},$$
$$\beta_i^{ups} = \frac{\sigma^2}{2\Delta x^2} + \max\left\{0, \frac{\kappa - p}{2\eta\Delta x}\right\}.$$

A combination of central and upstream differencing is then chosen according to algorithm 2 on a node by node basis.

Algorithm 2 Differencing method for monopoly problem

if $\alpha_i^{cent} \geq 0$ **and** $\beta_i^{cent} \geq 0$ **then**

$\alpha_i \leftarrow \alpha_i^{cent}$
$\beta_i \leftarrow \beta_i^{cent}$

else

$\alpha_i \leftarrow \alpha_i^{ups}$
$\beta_i \leftarrow \beta_i^{ups}$

end if

A.2. Duopoly problem: discrete equation coefficients

The coefficients of the discrete equations (12) are given here. Using standard central differencing for the first and second order derivatives, we have for player 1,

$$(\alpha_1)_{i,j}^{x_1,cent} = \frac{\sigma_1^2}{2\Delta x_1^2} + \frac{a_1 - a_2 p_1 + a_3(p_2^*)_{i,j}}{2\Delta x_1}$$
$$(\beta_1)_{i,j}^{x_1,cent} = \frac{\sigma_1^2}{2\Delta x_1^2} - \frac{a_1 - a_2 p_1 + a_3(p_2^*)_{i,j}}{2\Delta x_1}$$
$$(\alpha_1)_{i,j}^{x_2,cent} = \frac{\sigma_2^2}{2\Delta x_2^2} - \frac{1}{2\Delta x_2}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_1 + a_3(p_2^*)_{i,j}\right) - \frac{\kappa - (p_2^*)_{i,j}}{\eta}\right)$$
$$(\beta_1)_{i,j}^{x_2,cent} = \frac{\sigma_2^2}{2\Delta x_2^2} + \frac{1}{2\Delta x_2}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_1 + a_3(p_2^*)_{i,j}\right) - \frac{\kappa - (p_2^*)_{i,j}}{\eta}\right)$$

Similarly for player 2, we have

$$(\alpha_2)_{i,j}^{x_1,cent} = \frac{\sigma_1^2}{2\Delta x_1^2} - \frac{1}{2\Delta x_1}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_2 + a_3(p_1^*)_{i,j}\right) - \frac{\kappa - (p_1^*)_{i,j}}{\eta}\right)$$
$$(\beta_2)_{i,j}^{x_1,cent} = \frac{\sigma_1^2}{2\Delta x_1^2} + \frac{1}{2\Delta x_1}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_2 + a_3(p_1^*)_{i,j}\right) - \frac{\kappa - (p_1^*)_{i,j}}{\eta}\right)$$
$$(\alpha_2)_{i,j}^{x_2,cent} = \frac{\sigma_2^2}{2\Delta x_2^2} + \frac{a_1 - a_2 p_2 + a_3(p_1^*)_{i,j}}{2\Delta x_2}$$
$$(\beta_2)_{i,j}^{x_2,cent} = \frac{\sigma_2^2}{2\Delta x_2^2} - \frac{a_1 - a_2 p_2 + a_3(p_1^*)_{i,j}}{2\Delta x_2}$$

Using upstream for first order derivatives and central for second order derivatives results in

$$(\alpha_1)_{i,j}^{x_1,ups} = \frac{\sigma_1^2}{2\Delta x_1^2} + \max\left\{0, \frac{\left(a_1 - a_2 p_1 + a_3(p_2^*)_{i,j}\right)}{\Delta x_1}\right\} \tag{A1}$$
$$(\beta_1)_{i,j}^{x_1,ups} = \frac{\sigma_1^2}{2\Delta x_1^2} + \max\left\{0, -\frac{\left(a_1 - a_2 p_1 + a_3(p_2^*)_{i,j}\right)}{\Delta x_1}\right\} \tag{A2}$$
$$(\alpha_1)_{i,j}^{x_2,ups} = \frac{\sigma_2^2}{2\Delta x_2^2} + \max\left\{0, -\frac{1}{\Delta x_2}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_1 + a_3(p_2^*)_{i,j}\right) - \frac{\kappa - (p_2^*)_{i,j}}{\eta}\right)\right\} \tag{A3}$$
$$(\beta_1)_{i,j}^{x_2,ups} = \frac{\sigma_2^2}{2\Delta x_2^2} + \max\left\{0, \frac{1}{\Delta x_2}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_1 + a_3(p_2^*)_{i,j}\right) - \frac{\kappa - (p_2^*)_{i,j}}{\eta}\right)\right\} \tag{A4}$$
$$(\alpha_2)_{i,j}^{x_1,ups} = \frac{\sigma_1^2}{2\Delta x_1^2} + \max\left\{0, -\frac{1}{\Delta x_1}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_2 + a_3(p_1^*)_{i,j}\right) - \frac{\kappa - (p_1^*)_{i,j}}{\eta}\right)\right\} \tag{A5}$$
$$(\beta_2)_{i,j}^{x_1,ups} = \frac{\sigma_1^2}{2\Delta x_1^2} + \max\left\{0, \frac{1}{\Delta x_1}\left(\frac{\gamma}{\eta}\left(a_1 - a_2 p_2 + a_3(p_1^*)_{i,j}\right) - \frac{\kappa - (p_1^*)_{i,j}}{\eta}\right)\right\} \tag{A6}$$
$$(\alpha_2)_{i,j}^{x_2,ups} = \frac{\sigma_2^2}{2\Delta x_2^2} + \max\left\{0, \frac{\left(a_1 - a_2 p_2 + a_3(p_1^*)_{i,j}\right)}{\Delta x_2}\right\} \tag{A7}$$
$$(\beta_2)_{i,j}^{x_2,ups} = \frac{\sigma_2^2}{2\Delta x_2^2} + \max\left\{0, -\frac{\left(a_1 - a_2 p_2 + a_3(p_1^*)_{i,j}\right)}{\Delta x_2}\right\}. \tag{A8}$$

The coefficient due to the cross derivative term for both the players is given by

$$\xi_{i,j} = \begin{cases} \frac{\rho\sigma_1\sigma_2}{\Delta x_1 \Delta x_2}, & \text{if } \rho \geq 0 \\ \frac{-\rho\sigma_1\sigma_2}{\Delta x_1 \Delta x_2}, & \text{if } \rho < 0 \end{cases} \tag{A9}$$

A combination of central and upstream differencing is then chosen according to algorithm 3.

Algorithm 3 Differencing in the x_k, $k = 1, 2$ and for each player $l = 1, 2$.

if $(\alpha_l)_{i,j}^{x_k,cent} - \xi_{i,j} \geq 0$ **then**

$(\alpha_l)_{i,j}^{x_k} \leftarrow (\alpha_l)_{i,j}^{x_k,cent}$
$(\beta_l)_{i,j}^{x_k} \leftarrow (\beta_l)_{i,j}^{x_k,cent}$

else

$(\alpha_l)_{i,j}^{x_k} \leftarrow (\alpha_l)_{i,j}^{x_k,ups}$
$(\beta_l)_{i,j}^{x_k} \leftarrow (\beta_l)_{i,j}^{x_k,ups}$

end if

Appendix 2. American option under regime switching: discrete equation coefficients

The coefficients of the discrete equations (29) and (36) are given here. Using standard three point stencil and central differencing for the first and second order derivatives, we have

$$\alpha_{i,j}^{cent} = \frac{2a_{i,j}(S,\tau,\varphi)}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} - \frac{b_{i,j}(S,\tau,\varphi)}{(S_{i+1} - S_{i-1})}, \quad \text{(B10)}$$

$$\beta_{i,j}^{cent} = \frac{2a_{i,j}(S,\tau,\varphi)}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} + \frac{b_{i,j}(S,\tau,\varphi)}{(S_{i+1} - S_{i-1})}. \quad \text{(B11)}$$

Using upstream (forward/backward) differencing for the first order derivative and central differencing for the second order derivative, we have

$$\alpha_{i,j}^{ups} = \frac{2a_{i,j}(S,\tau,\varphi)}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} + \max\left\{0, -\frac{b_{i,j}(S,\tau,\varphi)}{(S_{i+1} - S_{i-1})}\right\}, \quad \text{(B12)}$$

$$\beta_{i,j}^{ups} = \frac{2a_{i,j}(S,\tau,\varphi)}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} + \max\left\{0, \frac{b_{i,j}(S,\tau,\varphi)}{(S_{i+1} - S_{i-1})}\right\}. \quad \text{(B13)}$$

A weighted average of central and upstream differencing is used on a node by node basis such that a positive coefficient discretization is obtained as detailed in algorithm 4.

Algorithm 4 Differencing method.

for $i = 1, 2, \ldots$ **do**
  $\omega = 1$
  if $\alpha_{i,j}^{cent} < 0$ **then**
    $\omega = \frac{\alpha_{i,j}^{ups}}{\alpha_{i,j}^{ups} - \alpha_{i,j}^{cent}}$
  else
    if $\beta_{i,j}^{cent} < 0$ **then**
      $\omega = \frac{\beta_{i,j}^{ups}}{\beta_{i,j}^{ups} - \beta_{i,j}^{cent}}$
    end if
  end if
  $\alpha_{i,j} = \omega\alpha_{i,j}^{cent} + (1-\omega)\alpha_{i,j}^{ups}$
  $\beta_{i,j} = \omega\beta_{i,j}^{cent} + (1-\omega)\beta_{i,j}^{ups}$
end for