Volume 44 • Number 3 • March 2008 ISSN 0005-1098

ELSEVIER

automatica

A Journal of IFAC The International Federation of Automatic Control

IFAC

www.elsevier.com/locate/automatica

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Automatica 44 (2008) 761–766
automatica
www.elsevier.com/locate/automatica

Brief paper

Convergence of Markov chain approximation on generalized HJB equation and its applications ☆

Q.S. Song

Department of Mathematics, University of Southern California, 3620 South Vermont Ave. KAP 108 Los Angeles, Los Angeles, CA 90089-2532, USA

Received 13 March 2007; received in revised form 27 May 2007; accepted 2 July 2007
Available online 11 December 2007

Abstract

This work is concerned with numerical methods for a class of stochastic control optimizations and stochastic differential games. Numerical procedures based on Markov chain approximation techniques are developed in a framework of generalized Hamilton–Jacobi–Bellman equations. Convergence of the algorithms is derived by means of viscosity solution methods.

Keywords: Stochastic control; Differential games; Convergence of numerical methods; Dynamic programming

1. Introduction

Stochastic control has its wide applications in manufacturing, communication theory, signal processing, and wireless networks; see for example Kushner and Dupuis (2001), Fleming and Soner (2006) and references therein. On the other hand, zero sum stochastic differential games, as the theory of two-controller, extends the control theory into more realistic problems. Many problems arising in, for example, pursuit evasion games, queueing systems in heavy traffic, risk sensitive control, and constrained optimization problems, can be formulated as two-player stochastic differential games (Basar & Bernhard, 1991; Elliott & Kalton, 1972).

It is well known that the value functions of stochastic optimal controls of such systems lead to systems of Hamilton–Jacobi–Bellman (HJB) equations, and the value functions of stochastic differential games satisfy Hamilton–Jacobi–Isaac (HJI) equations. Such a HJB or HJI equations are usually nonlinear and difficult to solve in closed form. Thus numerical methods become viable alternative. One of the most effective methods is the Markov chain approximation approach. The proof of convergence using probability methods is referred to Kushner (1990), Kushner and Dupuis (2001), and Song, Yin, and Zhang (2006) for stochastic controls and Kushner (2002), and Song and Yin (2006) for stochastic differential games. Viscosity solution methods provides another way to prove the convergence, see Barles and Souganidis (1991) for stochastic controls and Souganidis (1999) for stochastic differential games.

The idea of the generalized operator for associated HJB equations in this work is motivated from Littman and Szepesvari (1996) on Q-learning problem, in which many applications of such a generalization are introduced in the framework of Markov decision process (MDP). In this paper, we aim to introduce generalized HJB equations using the same type of generalized operator used in Littman and Szepesvari (1996). Upwind finite difference schemes and its interpretations of Markov chain approximation are developed on generalized HJB equations. The convergence of numerical solution is provided using viscosity solution technique. Different from previous literature aforementioned, such a development can be applied to a class of applications. In this paper, we focus on two of these applications: stochastic controls and zero-sum stochastic differential games in finite time horizon with stopping time associated with generalized HJB equations. Nevertheless, many other applications can fit for our framework as long as it satisfies conditions (C1)–(C3).

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor George Yin under the direction of Editor Ian Petersen.

E-mail address: qingshus@usc.edu.

0005-1098/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2007.07.014

The rest of paper is arranged as follows. Section 2 begins with description of generalized HJB equation. Associated stochastic control and stochastic differential games are formulated as its applications. Section 3 presents an effective upwind finite difference scheme with its probability interpretations. Section 4 proves the convergence of numerical scheme. Section 5 concludes the paper with further remarks.

2. Generalized HJB equations

Throughout the paper, we use following notations. K is generic constant. $\mathcal{O}$ is bounded open set in $\mathbb{R}^d$. The real-valued function $V(\cdot,\cdot)$ is defined on its domain $[0,T]\times\bar{\mathcal{O}}$. U is a compact subset of Euclidian space $\mathbb{R}^d$. $\sigma(\cdot,\cdot):\bar{\mathcal{O}}\times U\to\mathbb{R}^{d\times d}$ is a matrix-valued function, and $a(x,v)=\sigma(x,v)\sigma^{\mathrm{T}}(x,v)$, where $\sigma^{\mathrm{T}}(x,v)$ is the transpose of $\sigma(x,v)$. Let $L:\bar{\mathcal{O}}\times U\to\mathbb{R}$ be the running cost, $\psi(\cdot):\bar{\mathcal{O}}\to\mathbb{R}$ be the terminal cost. xy for $x,y\in\mathbb{R}^d$ is abbreviation of inner product $x^{\mathrm{T}}y=\sum_{i=1}^n x_iy_i$.

Consider generalized HJB equation: for $\forall(t,x)\in Q$

$$V_t+\otimes_v^x[f(x,v)D_xV+\tfrac{1}{2}\mathrm{tr}(a(x,v)D_x^2V)+L(x,v)]=0 \quad (1)$$

with boundary condition $V(t,x)=\psi(x)$ for $x\in[0,T)\times\partial\mathcal{O}\cup\{T\}\times\bar{\mathcal{O}}$. Here $\otimes_v^x$ is an operator that summarizes values over actions as a function of the state, such that, for any real-valued function ϕ_1,ϕ_2 and constant c, there exists some constant K

(C1) $\otimes_v^x[c\phi_1(x,v)+\phi_2(x)]=c\otimes_v^x[\phi_1(x,v)]+\phi_2(x)$.
(C2) $\otimes_v^x\phi_1(x,v)\leqslant\otimes_v^x\phi_2(x,v)$, whenever $\phi_1\leqslant\phi_2$.
(C3) $|\otimes_v^x\phi_1(x,v)-\otimes_v^x\phi_2(x,v)|\leqslant K\max_v|\phi_1(x,v)-\phi_2(x,v)|$.

Many natural operators satisfy the above conditions. For instance, $\max_v\phi(x,v)$ and $\min_v\phi(x,v)$ are related to classical stochastic control problems; $\min_{v_1}\max_{v_2}\phi(x,v_1,v_2)$ and $\min_{v_1}\max_{v_2}\max_{v_3}\phi(x,v_1,v_2,v_3)$ for $v=(v_1,v_2,v_3)$ are related to stochastic differential games either of zero-sum or non-zero-sum type. For a non-min–max type of operator satisfying above conditions, $\int_U\phi(x,v)m(\mathrm{d}v)$, where $m(\cdot)$ is a measure on a Borel σ-algebra of U, is associated to HJB equations with relaxed controls used.

To proceed, we need following regular assumption.

(A1) a,f,L,ψ are continuous and bounded. For $\phi=a,f,L,\psi$, function ϕ and its partial derivatives $\phi_{x_i},\phi_{x_ix_j}$ are continuous and bounded on $\mathbb{R}^d\times U$ for $i,j=1,2,\ldots,n$.

$(\Omega,\mathcal{F},\mathcal{F}_t,P,W)$ is a given probability space driven by Wiener Process W_t with filtration $\mathcal{F}_t$. In the following, we present two applications of generalized HJB equation in the form of (1): *stochastic control problem* and *stochastic differential games.*

2.1. Classical stochastic control problem

Suppose X_s satisfies controlled stochastic differential equation (SDE)

$$\mathrm{d}X_s=f(X_s,u_s)\,\mathrm{d}s+\sigma(X_s,u_s)\,\mathrm{d}W_s,\quad \forall s\in[t,T], \quad (2)$$

with initial condition $X_t=x$.

Definition 1. An admissible control process u on $[t,T]$ is an $\mathcal{F}_t$-progressively measurable process taking values in U. The set of all admissible controls is denoted by $\mathcal{U}(t)$.

Define stopping time $\tau=\inf\{s:X_s\notin\mathcal{O}\}$. Cost function for a given admissible control $u(\cdot)\in\mathcal{U}$ is defined as

$$J(t,x,u)=E\left[\int_t^{\tau\wedge T}L(X_s,u_s)\,\mathrm{d}s+\psi(X_{\tau\wedge T})\right], \quad (3)$$

and value function is defined as

$$V(t,x)=\inf_{u\in\mathcal{U}(t)}J(t,x,u). \quad (4)$$

It is well known that $V(t,x)$ is unique viscosity solution of HJB equation (1) with $\otimes_v^x=\min_v$. Similarly, if take sup over all admissible controls in (4), then $V(t,x)$ is unique viscosity solution of (1) with $\otimes_v^x=\max_v$. See more detail in Fleming and Rishel (1975), Crandall, Ishii, and Lions (1992) and Fleming and Soner (2006).

2.2. Stochastic differential games

Let $U=U_1\times U_2$ and $u=(u_1,u_2)\in\mathcal{U}(t)$ is an admissible control. X_s satisfies SDE (2). u_1 and u_2 are controls offered by player 1 and player 2, respectively. The collection of admissible controls on $[t,T]$ of player 1 and player 2 are denoted by $\mathcal{U}_1(t)$ and $\mathcal{U}_2(t)$. Player 1 (resp. player 2) wants to minimize (resp. maximize) the cost (3). In the following, we define Elliott–Kalton type upper and lower values of differential games.

Definition 2. An admissible strategy α (resp. β) for player 2 (resp. player 1) on $[t,T]$ is a mapping $\alpha:\mathcal{U}_1(t)\to\mathcal{U}_2(t)$ (resp. $\beta:\mathcal{U}_2(t)\to\mathcal{U}_1(t)$), such that, for $t<r<T$, $u_1(s)=\tilde{u}_1(s)$ (resp. $u_2(s)=\tilde{u}_2(s)$) for almost all $s\in[t,r]$ implies $\alpha(u_1(s))=\alpha(\tilde{u}_1(s))$ (resp. $\beta(u_2(s))=\beta(\tilde{u}_2(s))$) almost everywhere. Let $\mathcal{L}_1(t)$ (resp. $\mathcal{L}_2(t)$) denote the class of all admissible strategies of player 2 and player 1 on $[t,T]$.

The upper value $V^+(t,x)$ and lower value $V^-(t,x)$ are defined as

$$V^+(t,x)=\sup_{\alpha\in\mathcal{S}_1(t)}\inf_{u_1\in\mathcal{U}_1(t)}J(t,x,u_1,\alpha(u_1)) \quad (5)$$

and

$$V^-(t,x)=\inf_{\beta\in\mathcal{S}_2(t)}\sup_{u_2\in\mathcal{U}_2(t)}J(t,x,\beta(u_2),u_2). \quad (6)$$

It is well known that $V^+(t,x)$ (resp. $V^-(t,x)$) is unique viscosity solution of HJB equation (1) with $\otimes_v^x=\min_{v_1}\max_{v_2}$ (resp. $\max_{v_2}\min_{v_1}$). If $V^+(t,x)=V^-(t,x)$ holds for all $(t,x)\in Q$, then the differential game is said to have a saddle point, and its value is denoted by $V(t,x)$. See more detail in and Elliott and Kalton (1972) and Fleming and Souganidis (1989).

3. Numerical solutions

Let e_i be ith unit basis of $\mathbb{R}^d$ for $i=1,2,\ldots,d$. For a given positive discretized parameter δ, h, define discrete spaces in state and time by

$$\Sigma^\delta = \left\{ x \in \mathcal{O} : x = \sum_{i=1}^n k_i \delta e_i, k_i \in \mathbb{Z} \right\},$$
$$[t,T]^h = [t,T] \cap \{t = kh + T : k \in \mathbb{Z}\}. \tag{7}$$

To proceed, the following assumptions will be given.

(A2) $a(x,v)$ satisfies $|a_{ii}(x,v)| - \sum_{j\neq i} |a_{ij}(x,v)| \geqslant 0$.
(A3) Discrete parameter $\delta = \delta(h)$ is a function of h, s.t.

$$h \sum_{i=1}^n \left[a_{ii}(x,v) - \frac{1}{2} \sum_{j\neq i} |a_{ij}(x,v)| + \delta |f_i(x,v)| \right] \leqslant \delta^2.$$

Assumption (A2) requires that the diffusion matrix be diagonally dominated. If the given dynamic system does not satisfy (A2), then we can adjust the coordinate system to satisfy Assumption (A2); see Kushner and Dupuis (2001, p. 110) and Fleming and Soner (2006, p. 329). Assumption (A3) gives the relation between two parameters δ and h, which are used in discretization.

By $V^h(\cdot,\cdot)$ on $\Sigma^\delta \times [0,T]^h$, we denote numerical solution of (1) with parameters δ, h of (A3) used. Note that, for simplicity, we use V^h instead of $V^{\delta,h}$.

Numerical solution V^h can be obtained by upwind finite difference numerical scheme, that is, for $\forall \phi(t,x)$

$$\begin{aligned}
\Delta_{x_i}^{\delta,\pm}\phi &= \delta^{-1}[\phi(t,x\pm\delta e_i) - \phi(t,x)],\\
\Delta_{x_i}^{2,\delta}\phi &= \delta^{-2}[\phi(t,x+\delta e_i) + \phi(t,x-\delta e_i) - 2\phi(t,x)]\\
&\triangleq \Delta_{x_i x_i}^{\delta,+}\phi \triangleq \Delta_{x_i x_i}^{\delta,-}\phi,\\
\Delta_{x_i x_j}^{\delta,+}\phi &= \tfrac{1}{2}\delta^{-2}[2\phi(t,x) + \phi(t,x+\delta e_i+\delta e_j)\\
&\quad + \phi(t,x-\delta e_i-\delta e_j)]\\
&\quad - \tfrac{1}{2}\delta^{-2}[\phi(t,x+\delta e_i) + \phi(t,x-\delta e_i)\\
&\quad + \phi(t,x+\delta e_j) + \phi(t,x-\delta e_j)],\\
\Delta_{x_i x_j}^{\delta,-}\phi &= -\tfrac{1}{2}\delta^{-2}[2\phi(t,x) + \phi(t,x+\delta e_i-\delta e_j)\\
&\quad + \phi(t,x-\delta e_i+\delta e_j)]\\
&\quad + \tfrac{1}{2}\delta^{-2}[\phi(t,x+\delta e_i) + \phi(t,x-\delta e_i)\\
&\quad + \phi(t,x+\delta e_j) + \phi(t,x-\delta e_j)],\\
\Delta_t^{h,-}\phi &= (\phi(t,x) - \phi(t-h,x))/h,\\
\Delta_t^{h,+}\phi &= (\phi(t+h,x) - \phi(t,x))/h.
\end{aligned}$$

Applying above upwind finite difference scheme to (1), one can write explicit numerical scheme as

$$\begin{aligned}
&\Delta_t^{h,-}V^h + \otimes_v^x [f^+(x,v)\Delta_x^{\delta,+}V^h\\
&\quad + \tfrac{1}{2}\mathrm{tr}(a^+(x,v)\Delta_x^{2,\delta,+}V^h) - f^-(x,v)\Delta_x^{\delta,-}V^h\\
&\quad - \tfrac{1}{2}\mathrm{tr}(a^-(x,v)\Delta_x^{2,\delta,-}V^h) + L(x,v)] = 0,
\end{aligned} \tag{8}$$

where

$$\begin{aligned}
&a^\pm = \max\{\pm a, 0\},\\
&\Delta_x^{\delta,\pm}\phi = (\Delta_{x_1}^{\delta,\pm}\phi, \Delta_{x_2}^{\delta,\pm}\phi, \ldots, \Delta_{x_d}^{\delta,\pm}\phi)^{\mathrm{T}},\\
&\Delta_x^{2,\delta,\pm}\phi = (\Delta_{x_i x_j}^{\delta,\pm}\phi)_{i,j=1,\ldots,d}.
\end{aligned} \tag{9}$$

Note that $\Delta_x^{2,\delta,\pm}$ is symmetric matrix. In the following, we give equivalent Markov chain approximation interpretation of above upwind finite difference scheme. One can rewrite (8) with boundary conditions by

$$\begin{aligned}
V^h(t-h,x) = \otimes_v^x \Bigg[&\sum_{y\in\Sigma^\delta} p^h(x,y,v)V^h(t,y)\\
&+ hL(x,v) \Bigg], \quad t \in [h,T]^h, \quad x \in \Sigma^\delta,
\end{aligned}$$
$$\begin{aligned}
&V^h(T,x) = \psi(x), \quad x \in \Sigma^\delta,\\
&V^h(t,x) = \psi(x), \quad (t,x) \in [0,T) \times \partial\mathcal{O},
\end{aligned} \tag{10}$$

where

$$\begin{aligned}
&p^h(x, x\pm\delta e_i, v)\\
&\quad = \frac{h}{2\delta^2}\left[a_{ii}(x,v) - \sum_{j\neq i}|a_{ij}(x,v)| + 2\delta f_i^\pm(x,v) \right],\\
&p^h(x, x+\delta e_i \pm \delta e_j, v) = \frac{h}{2\delta^2} a_{ij}^\pm(x,v), \quad i \neq j,\\
&p^h(x, x-\delta e_i \pm \delta e_j, v) = \frac{h}{2\delta^2} a_{ij}^\mp(x,v), \quad i \neq j,\\
&p^h(x,x,v)\\
&\quad = 1 - \frac{h}{\delta^2}\sum_{i=1}^n \left[a_{ii}(x,v) - \frac{1}{2}\sum_{j\neq i}|a_{ij}(x,v)| + \delta|f_i(x,v)| \right],\\
&p^h(x,y,v) = 0 \quad \text{otherwise.}
\end{aligned}$$

Note that, under Assumptions (A2) and (A3), we have

$$\sum_{y\in\Sigma^\delta} p^h(x,y,v) = 1, \quad p^h(x,y,v) \geqslant 0. \tag{11}$$

It can be seen from (11), we can consider $p^h(\cdot)$ as a one step controlled transition probability of a Markov chain $\{x_n^h : n = 0,1,2,\ldots\}$ in state space Σ^δ, on which appropriate cost function is defined.

Remark 3. The numerical scheme obtained in (10) is explicit backward in time, since each $V^h(t-h,\cdot)$ can be explicitly solved in terms of $V^h(t,\cdot)$ from terminal condition $V^h(T,\cdot)$. An implicit numerical scheme also can be obtained by replacing $\Delta_t^{h,-}\phi$ by $\Delta_t^{h,+}\phi$ in (8), i.e.

$$\begin{aligned}
&\Delta_t^{h,+}V^h + \otimes_v^x [f^+(x,v)\Delta_x^{\delta,+}V^h\\
&\quad + \tfrac{1}{2}\mathrm{tr}(a^+(x,v)\Delta_x^{2,\delta,+}V^h) - f^-(x,v)\Delta_x^{\delta,-}V^h\\
&\quad - \tfrac{1}{2}\mathrm{tr}(a^-(x,v)\Delta_x^{2,\delta,-}V^h) + L(x,v)] = 0.
\end{aligned} \tag{12}$$

The above implicit numerical scheme also have probability interpretations when we deal discrete time as another state variable, see Kushner and Dupuis (2001, Chapter 12).

The next section is to find sufficient conditions such that V^h of explicit scheme (8) converge to unique viscosity solution V of (1). For the convergence of implicit scheme (12), we can follow analogous method.

4. Convergence

To show the convergence of V^h of explicit scheme (8) with boundary conditions, one can rewrite (8) as

$$\begin{aligned} &V^h(t-h,x) = F_h[V^h(t,\cdot)](x), \quad t \in [h,T]^h, \quad x \in \Sigma^\delta, \\ &V^h(T,x) = \psi(x), \quad x \in \Sigma^\delta, \\ &V^h(t,x) = \psi(x), \quad (t,x) \in [0,T) \times \partial\mathcal{O}, \end{aligned} \tag{13}$$

where $F_h[\phi](x)$ is an operator for any function $\phi : \mathbb{R}^d \to \mathbb{R}$, such that

$$\begin{aligned} F_h[\phi](x) = \phi(x) + h\otimes_v^x [&f^+(x,v)\Delta_x^{\delta,+}\phi(x) \\ &- f^-(x,v)\Delta_x^{\delta,-}\phi(x) + \tfrac{1}{2}\mathrm{tr}(a^+(x,v)\Delta_x^{2,\delta,+}\phi(x)) \\ &- \tfrac{1}{2}\mathrm{tr}(a^-(x,v)\Delta_x^{\delta,-}\phi(x)) + L(x,v)]. \end{aligned} \tag{14}$$

Note that, by condition (C1), one can rewrite (14) as

$$F_h[\phi](x) = \otimes_v^x \left[\sum_{y\in\Sigma^\delta} p^h(x,y,v)\phi(y) + hL(x,v) \right]. \tag{15}$$

Lemma 4. *Assume* (A1)–(A3). *Following properties hold*:

$$F_h[\phi_1] \leqslant F_h[\phi_2] \quad \textit{for } \forall \phi_1 \leqslant \phi_2, \tag{16}$$

$$F_h(\phi + c) = F_h(\phi) + c, \quad \forall c \in \mathbb{R}, \tag{17}$$

$$\|V^h\|_\infty \leqslant K, \quad \forall 0 < h < 1, \tag{18}$$

and for all $\phi \in C^{1,2}((0,T)\times\mathcal{O})$

$$\begin{aligned} &\lim_{\substack{(s,y)\to(t,x)\\ h\to 0}} \frac{F_h[\phi(s,\cdot)](y) - \phi(s-h,y)}{h} \\ &\quad = \phi_t + \otimes_v^x [f(x,v)D_x\phi(t,x) \\ &\qquad + \tfrac{1}{2}\mathrm{tr}(a(x,v)D_x^2\phi(t,x)) + L(x,v)]. \end{aligned} \tag{19}$$

Proof. Note that under (A2), (A3) and (11), one can have (16) and (17). Rewrite (13) as

$$V^h(t-h,x) = \otimes_v^h \left[\sum_{y\in\Sigma^\delta} p^h(x,y,v)V^h(t,y) + hL(x,v) \right].$$

Then, for any $t \in [h,T]$

$$\begin{aligned} V^h(t-h,x) &\leqslant \otimes_v^h \left[\max_y V^h(t,y) + hL(x,v) \right] \\ &\leqslant \max_y V^h(t,y) + Kh\|L\|_1. \end{aligned}$$

It leads to stability of F_h, that is, for any $0 \leqslant m \leqslant T/h$,

$$\max_x V^h(T-mh,x) \leqslant \max_x V^h(T,x) + Kmh\|L\|_1 \leqslant K.$$

Hence, (18) holds.

For any test function $\phi \in C^{1,2}([0,T]\times\bar{\mathcal{O}})$, one can verify the consistency (19) as following,

$$\begin{aligned} &\lim_{\substack{(s,y)\to(t,x)\\ h\to 0}} \frac{F_h[\phi(s,\cdot)](y) - \phi(s-h,y)}{h} \\ &= \lim_{\substack{(s,y)\to(t,x)\\ h\to 0}} \frac{\phi(s,y) - \phi(s-h,y)}{h} \\ &\quad + \lim_{\substack{(s,y)\to(t,x)\\ h\to 0}} \otimes_v^y \Big[L(y,v) + f^+(y,v)\Delta_x^{\delta,+}\phi(s,y) \\ &\quad - f^-(y,v)\Delta_x^{\delta,-}\phi(s,y) + \frac{1}{2}\mathrm{tr}(a^+(y,v)\Delta_x^{2,\delta,+}\phi(s,y)) \\ &\quad - \frac{1}{2}\mathrm{tr}(a^-(y,v)\Delta_x^{2,\delta,-}\phi(s,y)) \Big] \\ &= \phi_t + \otimes_v^x [f(x,v)D_x\phi(t,x) \\ &\quad + \tfrac{1}{2}\mathrm{tr}(a(x,v)D_x^2\phi(t,x)) + L(x,v)]. \end{aligned}$$

This completes the proof. □

Definition 5. We say that V is a viscosity solution of equation (1) if (a) $V(t,x)$ is upper semicontinuous function on $[0,T]\times\bar{\mathcal{O}}$ and for each smooth function $\phi \in C^{1,2}((0,T)\times\mathcal{O})$,

$$\begin{aligned} &\phi_t(\bar{t},\bar{x}) + \otimes_v^x [f(\bar{x},v)D_x\phi(\bar{t},\bar{x}) \\ &\quad + \tfrac{1}{2}\mathrm{tr}(a(\bar{x},v)D_x^2\phi(\bar{t},\bar{x})) + L(\bar{x},v)] \geqslant 0, \end{aligned} \tag{20}$$

at every $(\bar{t},\bar{x}) \in (0,T)\times\mathcal{O}$ which is strict maximizer of $V-\phi$ on $[0,T]\times\bar{\mathcal{O}}$.

(b) $V(t,x)$ is lower semicontinuous function on $[0,T]\times\bar{\mathcal{O}}$ and for each $\phi \in C^{1,2}((0,T)\times\mathcal{O})$,

$$\begin{aligned} &\phi_t(\bar{t},\bar{x}) + \otimes_v^x [f(\bar{x},v)D_x\phi(\bar{t},\bar{x}) \\ &\quad + \tfrac{1}{2}\mathrm{tr}(a(\bar{x},v)D_x^2\phi(\bar{t},\bar{x})) + L(\bar{x},v)] \leqslant 0, \end{aligned} \tag{21}$$

at every $(\bar{t},\bar{x}) \in (0,T)\times\mathcal{O}$ which is strict minimizer of $V-\phi$ on $[0,T]\times\bar{\mathcal{O}}$.

If (a) (respectively (b)) holds, then V is said to be subsolution (respectively supersolution) of (1).

For $(t,x) \in (0,T)\times\mathcal{O}$, define upper and lower semicontinuous envelope of solution V^h as

$$\begin{aligned} V^*(t,x) &= \limsup_{\substack{(s,y)\to(t,x)\\ h\to 0}} V^h(s,y), \\ V_*(t,x) &= \liminf_{\substack{(s,y)\to(t,x)\\ h\to 0}} V^h(s,y). \end{aligned} \tag{22}$$

Lemma 6. V^* (*resp.* V_*) *defined in* (22) *is a viscosity subsolution* (*resp. supersolution*) *of Eq.* (1) *under Assumptions* (A1), (A2), *and* (A3).

Proof. Suppose that $\phi \in C^{1,2}((0,T)\times\mathcal{O})$ is a test function such that $V^*-\phi$ has strict maximum at $(\bar{t},\bar{x}) \in (0,T)\times\mathcal{O}$

by upper semicontinuity. Then there is a sequence converging to zero denoted by h, such that $V^h - \phi$ has a maximum on $[0,T]^h \times \Sigma^\delta$ at (s_h, y_h) such that $(s_h, y_h) \to (\bar{t}, \bar{x})$ as $h \to 0$. For all $y \in \Sigma^h$, therefore

$$\phi(s_h + h, y) - \phi(s_h, y_h) \geqslant V^h(s_h + h, y) - V^h(s_h, y_h).$$

By virtue of (16) and (17),

$$F_h[\phi(s_h + h, \cdot)](y_h) - \phi(s_h, y_h) \geqslant F_h[V^h(s_h + h, \cdot)](y_h) - V^h(s_h, y_h). \tag{23}$$

By (13), the right-hand side of (23) is zero. By dividing h and forcing $h \to 0$, left-hand side of (23) goes to (19). Thus, (20) holds. One can prove V_* is supersolution in similar fashion. □

In the following lemma, by $A \geqslant B$ for symmetric matrices, we mean $A - B$ is symmetric positive definite.

Lemma 7. *Suppose* (A1)–(A3) *hold. Let* ϕ *and* $\bar{\phi}$ *be bounded viscosity subsolution and supersolution of* (1). *Then*

$$\sup_{[0,T]\times\bar{\mathcal{O}}} (\phi - \bar{\phi}) = \sup_{(t,y)\in[0,T)\times\partial\mathcal{O}\cup\{T\}\times\bar{\mathcal{O}}} (\phi(t,y) - \bar{\phi}(t,y)).$$

Proof. By virtue (Fleming & Soner, 2006, Theorem V.8.1) it is enough to show that there exists a constant K, such that

$$\begin{aligned}&\otimes_v^y[\alpha f(y,v)(x-y) + \tfrac{1}{2}\mathrm{tr}(a(y,v)B) + L(y,v)]\\&\quad - \otimes_v^x[\alpha f(x,v)(x-y) + \tfrac{1}{2}\mathrm{tr}(a(x,v)B) + L(x,v)]\\&\quad\quad \leqslant K(\alpha|x-y|^2 + |x-y|),\end{aligned}$$

for every $(t,x), (t,y) \in Q$, $\alpha > 0$, and symmetric matrices A, B satisfying

$$-3\alpha\begin{bmatrix} I & 0 \\ 0 & I\end{bmatrix} \leqslant \begin{bmatrix} A & 0 \\ 0 & -B\end{bmatrix} \leqslant 3\alpha\begin{bmatrix} I & -I \\ -I & I\end{bmatrix}.$$

By condition (C3) of the operator $\otimes_v^x$, one can write

$$\begin{aligned}&\otimes_v^y[\alpha f(y,v)(x-y) + \tfrac{1}{2}\mathrm{tr}(a(y,v)B) + L(y,v)]\\&\quad - \otimes_v^x[\alpha f(x,v)(x-y) + \tfrac{1}{2}\mathrm{tr}(a(x,v)B) + L(x,v)]\\&\quad\quad \leqslant K\max_v|\alpha(f(y,v) - f(x,v))(x-y)|\\&\quad\quad\quad + K\max_v|L(y,v) - L(x,v)|\\&\quad\quad\quad + K\max_v|\mathrm{tr}(a(y,v)B - a(x,v)A)|.\end{aligned}$$

Note that Assumption (A1) implies Lipschitz continuity of function f and L. Hence

$$\begin{aligned}&\max_v|\alpha(f(y,v) - f(x,v))(x-y)|\\&\quad + \max_v|L(y,v) - L(x,v)| \leqslant K(\alpha|x-y|^2 + |x-y|).\end{aligned}$$

The last term is

$$\begin{aligned}&\mathrm{tr}(a(y,v)B - a(x,v)A)\\&\quad = \mathrm{tr}\left(\begin{bmatrix}\sigma^{\mathrm{T}}(y,v)\sigma(y,v) & \sigma(y,v)\sigma^{\mathrm{T}}(x,v)\\ \sigma(x,v)\sigma^{\mathrm{T}}(y,v) & \sigma(x,v)\sigma^{\mathrm{T}}(x,v)\end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & -B\end{bmatrix}\right)\\&\quad \leqslant 3\alpha\,\mathrm{tr}\left(\begin{bmatrix}\sigma(y,v)\sigma^{\mathrm{T}}(y,v) & \sigma(y,v)\sigma^{\mathrm{T}}(x,v)\\ \sigma(x,v)\sigma^{\mathrm{T}}(y,v) & \sigma(x,v)\sigma^{\mathrm{T}}(x,v)\end{bmatrix}\begin{bmatrix} I & -I \\ -I & I\end{bmatrix}\right)\\&\quad = 3\alpha\,\mathrm{tr}(\sigma(y,v)\sigma^{\mathrm{T}}(y,v) - \sigma(y,v)\sigma^{\mathrm{T}}(x,v)\\&\quad\quad - \sigma(x,v)\sigma^{\mathrm{T}}(y,v) + \sigma(x,v)\sigma^{\mathrm{T}}(x,v))\\&\quad = 3\alpha\,\mathrm{tr}((\sigma(y,v) - \sigma(x,v))(\sigma(y,v) - \sigma(x,v))^{\mathrm{T}})\\&\quad = 3\alpha\|\sigma(y,v) - \sigma(x,v)\|^2 \leqslant K\alpha|x-y|^2.\end{aligned}$$

The above inequalities imply the result. □

Theorem 8. *Suppose* (A1)–(A3) *hold. Then* $V^* = V_* \triangleq V$ *is the unique viscosity solution of* (1) *on* $[0,T] \times \bar{\mathcal{O}}$.

Proof. By (22), $V^* \geqslant V_*$. Moreover, (13) and (22) implies V^* and V_* has the same boundary condition on $[0,T) \times \partial\mathcal{O} \cup \{T\} \times \bar{\mathcal{O}}$. Applying Lemmas 6 and 7, one can write $V^* \leqslant V_*$. Thus, $V = V^* = V_*$ is viscosity solution of (1). Uniqueness also follows from Lemma 7. □

Following corollaries are direct results from Theorem 8.

Corollary 9. *Suppose* (A1)–(A3) *hold.* $V(t,x)$ *is the value function defined by* (3), *and* $V^h(t,x)$ *is approximated value function of* (8) *with* $\otimes_v^x$ *replaced by* $\min_v$. *Then* $V^h(t,x)$ *converge to* $V(t,x)$ *as* $h \to 0$.

Corollary 10. *Suppose* (A1), (A2), *and* (A3) *hold.* $V^+(t,x)$ (*resp.* $V^-(t,x)$) *is the value defined by* (5) (*resp.* (6)), *and* $V^{h+}(t,x)$ (*resp.* $V^{h-}(t,x)$) *is approximated value function of* (8) *with* $\otimes_v^x$ *replaced by* $\min_{v_1}\max_{v_2}$ (*resp.* $\max_{v_2}\min_{v_1}$). *Then* $V^{h+}(t,x)$ (*resp.* $V^-(t,x)$) *converge to* $V^+(t,x)$ (*resp.* $V^-(t,x)$) *as* $h \to 0$.

5. Further remarks

In this work, the generalized HJB equation is proposed, which is associated with stochastic control and stochastic differential games. The proof of convergence is given by the viscosity solution method. Probability methods analogous to Kushner (1990) and Song and Yin (2006) can also be used to prove the convergence. Numerical examples on stochastic control and stochastic differential games are referred to Song, Yin, and Zhang (2006) and Song, Yin, and Zhang (2007) due to the limit of space.

Another approachable result is controlled stochastic hybrid system. Such a formulation has extensive recent applications in risk theory, financial engineering, and insurance modeling, see Rolski, Schmidli, Schmidt, and Teugels (1999), Yin and Zhang (1998), and Yin and Zhang (2005). It might be interesting if one could find more applications in the manner of generalized HJB equations, such as, risk-sensitive models, exploration-sensitive models, and non zero sum differential games.

Acknowledgements

Research of this author is supported in part by the U.S. Army Research Office MURI Grant W911NF-06-1-0094 at the University of Southern California. The author gratefully acknowledges many useful comments and suggestions from the referees and editors.

References

Barles, G., & Souganidis, P. E. (1991). Convergence of approximation schemes for fully nonlinear second order equations. *Journal of Asymptotic Analysis*, *4*, 271–283.

Basar, T., & Bernhard, P. (1991). *H_∞-optimal control and related minimax problems*. Boston: Birkhäuser.

Crandall, M. G., Ishii, H., & Lions, P. L. (1992). User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, *27*(1), 1–67.

Elliott, R. J., & Kalton, N. J. (1972). *Existence of value in differential games* (Vol. 126). Providence, RI: Memoires of American Mathematical Society.

Fleming, W. H., & Rishel, R. W. (1975). *Deterministic and stochastic optimal control*. New York: Springer.

Fleming, W. H., & Soner, H. M. (2006). *Controlled Markov processes and viscosity solutions*. (2nd ed.), New York: Springer.

Fleming, W. H., & Souganidis, P. E. (1989). On the existence of value functions of two-player, zero-sum stochastic differential games. *Indiana University Mathematics Journal*, *38*(2), 293–314.

Kushner, H. J. (1990). Numiercal methods for stochastic control problems in continuous time. *SIAM Journal of Control and Optimization*, *28*, 999–1048.

Kushner, H. J. (2002). Numerical approximations for stochastic differential games. *SIAM Journal of Control and Optimization*, *41*(2), 457–486.

Kushner, H. J., & Dupuis, P. (2001). *Numerical methods for stochastic control problems in continuous time*. (2nd ed.), New York: Springer.

Littman, M. L., & Szepesvari, C. (1996). A generalized reinforcement-learning model: Convergence and applications. *Proceedings of the 13th international conference on machine learning* (ICML-Vol. 96, pp. 310–318). Bari, Italy.

Rolski, T., Schmidli, H., Schmidt, V., & Teugels, J. (1999). *Stochastic processes for insurance and finance*. New York: Wiley.

Song, Q. S., & Yin, G. (2006). Existence of saddle points in discrete Markov games and its application in numerical methods for stochastic differential games. *Proceedings of the 45th IEEE conference on decision & control CDC06* (pp. 6325–6330).

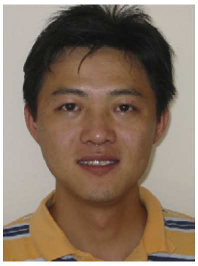
Song, Q. S., Yin, G., & Zhang, Z. (2006). Numerical method for controlled regime-switching diffusions and regime-switching jump diffusions. *Automatica*, *42*, 1147–1158.

Song, Q. S., Yin, G., & Zhang, Z. (2007). Numerical solutions for stochastic differential games with regime switching. *In revision*, 28pp.

Souganidis, P. E. (1999). Two-player zero-sum differential games and viscosity solutions. Stochastic and differential games. *Annals of the International Society of Dynamical Games*, *4*, 69–104.

Yin, G., & Zhang, Q. (1998). *Continuous-time Markov chains and applications: A singular perturbation approach*. New York: Springer.

Yin, G., & Zhang, Q. (2005). *Discrete-time Markov chains: Two time-scale methods and applications*. New York: Springer.



Qingshuo Song received his B.S. in Automatic Control and Systems from Nankai University in 1996, M.A. in Computer Science from Nankai University in 1999, and Ph.D. in Mathematics from Wayne State University in 2006. He is currently an NTT assistant professor in Department of Mathematics at University of Southern California. His research interests include stochastic control and related PDE, mathematical finance, numerical analysis, stochastic differential games.