

Markov Approximation of Zero-sum Differential Games

Yurii Averboukh

Krasovskii Institute of Mathematics and Mechanics

S. Kovalevskoy str, 16,

620990 Yekaterinburg, Russia.

ayv@imm.uran.ru

Abstract

The paper is concerned with approximations of a value function of a differential game. To this end we introduce a approximation of the original differential game by a continuous-time Markov game with infinite state space. The value function of this game in a given region is approximated by a solution of a finite state Markov game. This yields the approximation of the value function of the original differential game by the finite system of ODEs.

1 Introduction

The paper deals with the analysis of the deterministic differential games. These problems arise in many applications and primary are concerned with the control problem for mechanical or economical systems. The key concept of the differential game theory is a concept of a value function. Given the value function of the differential game one can construct a near optimal feedback strategies of the players [Krasovskii & Subbotin, 1988], [Subbotin, 1995]. The stable under information disturbances strategies are constructed within the framework of control with guide strategies [Krasovskii & Subbotin, 1973], [Krasovskii & Subbotin, 1988]. Equivalent result occurs when the players are informed about the control of the opponent realized up to the current time [Varaya & Lin, 1969], [Elliott & Kalton, 1972], [Subbotin & Chentsov, 1981]. The theory of zero-sum differential games is closely related to the theory of first-order PDEs. It is proved that the value function is the minimax/viscosity solution of the corresponding Hamilton-Jacobi PDE.

The direct study of the value function is extremely complicated. The fully discrete approximations leads to the numerical approximations (see [Botkin et al.,] and references therein). In the paper we introduce the approximation based on continuous-time stochastic game. This implies the approximations of the value function by solutions of systems of ODEs or by solutions of second-order PDEs.

The approximation of the differential game by the stochastic differential game is studied in [Buslaeva, 1977], [Krasovskii & Kotelnikova, 2010a], [Krasovskii & Kotelnikova, 2010b], [Krasovskii & Kotelnikova, 2009], [Camilli & Marchi, 2012]. It was shown that one can construct a near-optimal strategy based on a solution of the stochastic differential game. The close problem was considered in [Kolokoltsov, 2012] where the near optimal strategies for the continuous-time Markov game was designed based on a value function of a differential game. The approximation of the stochastic differential game with the reflection term by the continuous-time Markov game was proposed in [Kushner, 2007]. The approximation of the differential game by the discrete time Markov chain was studied in [Li & Song, 2007] (see, also, survey [Jorgensen & Zaccour, 2006]). Note that these

Copyright © by the paper's authors. Copying permitted for private and academic purposes.

In: Yu. G. Evtushenko, M. Yu. Khachay, O. V. Khamisov, Yu. A. Kochetov, V.U. Malkova, M.A. Posypkin (eds.): Proceedings of the OPTIMA-2017 Conference, Petrovac, Montenegro, 02-Oct-2017, published at <http://ceur-ws.org>

results are closed to the problems of the approximation of the control system by Markov chains considered in [Boue & Dupuis, 1999], [Kushner & Dupuis, 2001].

Paper [Averboukh, 2016] is concerned with the general continuous-time stochastic game. A near optimal strategy for the continuous-time stochastic game is constructed using the value function of a game with a different dynamics. The general theory implies the approximation of the value function of the differential game by the infinite system of ODEs. This system is obtained using dynamic programming arguments for the approximating continuous-time Markov game.

In the paper we introduce the approximation of the value function of the differential game by the finite system of ODEs. This approach can be used for numerical schemes. The paper is organized as follows. In Section 2 we recall the general theory of differential games. Section 3 deals with the approximation of the original differential game by a Markov game with infinite state space. This leads to the approximation of the value function by infinite system of ODEs. The approximation of the differential game by a Markov game with the finite state space is considered in Section 4. In this section we obtain an approximation of the value function of the differential game by the finite system of ODEs.

2 Differential Games

We consider the zero-sum differential game with the dynamics given by

$$\frac{d}{dt}x(t) = f(t, x(t), u(t), v(t)), \quad t \in [0, T], \quad x(t) \in \mathbb{R}^d, \quad u(t) \in U, \quad v(t) \in V. \quad (1)$$

Here $u(t)$ (respectively, $v(t)$) is a instantaneous control of the first (respectively, second) player. The aim of the first (respectively, second) player is to minimize (respectively, maximize) the terminal payoff $\sigma(x(T))$.

We use feedback approach proposed by Krasovskii and Subbotin. To introduce this formalization let us denote by $\mathcal{U}[t_0]$ (respectively, $\mathcal{V}[t_0]$) the set of all measurable functions from $[t_0, T]$ to U (receptively, V).

We assume that

1. the sets U and V are metric compacts;
2. the functions f and σ are continuous;
3. the functions f and σ are bounded by the constant M ;
4. there exists a constant K such that, for any $t \in [0, T]$, $x', x'' \in \mathbb{R}^d$, $u \in U$, $v \in V$,

$$\|f(t, x', u, v) - f(t, x'', u, v)\| \leq K\|x' - x''\|;$$

5. there exists a function $\alpha : \mathbb{R} \rightarrow \mathbb{R}$ such that $\alpha(\delta) \rightarrow 0$ as $\delta \rightarrow 0$ and, for any $t', t'' \in [0, T]$, $x \in \mathbb{R}^d$, $u \in U$, $v \in V$,

$$\|f(t', x, u, v) - f(t'', x, u, v)\| \leq \alpha(t' - t'');$$

6. the function σ is Lipschitz continuous with the constant R ;
7. (Isaacs' condition) for any $t \in [0, T]$, $x, \xi \in \mathbb{R}^d$, $u \in U$, $v \in V$,

$$\min_{u \in U} \max_{v \in V} \langle \xi, f(t, x, u, v) \rangle = \max_{v \in V} \min_{u \in U} \langle \xi, f(t, x, u, v) \rangle.$$

Below, let p be an arbitrary function from $[0, T] \times \mathbb{R}^d$ to U . We will construct a stepwise control that is constant between the times of control correction. Let t_0 be an initial time, $\Delta = \{t_i\}_{i=0}^N$ be a partition of $[t_0, T]$. The times t_i are times of control correction. If $x_0 \in \mathbb{R}^d$, $v \in \mathcal{V}[t_0]$, then let $x^1[\cdot, t_0, x_0, p, \Delta, v]$ be a solution of the problem

$$\frac{d}{dt}x(t) = f(t, x(t), p(t_{i-1}, x(t_{i-1})), v(t)), \quad t \in [t_{i-1}, t_i], \quad i = 1, \dots, N, \quad x(t_0) = x_0.$$

Analogously, the strategy of the second player is determined by the function $q : [0, T] \times \mathbb{R}^d \rightarrow V$. If (t_0, x_0) is an initial position, $\Delta = \{t_i\}_{i=0}^N$ is a partition of $[t_0, T]$, $u \in \mathcal{U}[t_0]$, then denote by $x^2[\cdot, t_0, x_0, q, \Delta, u]$ a solution of the problem

$$\frac{d}{dt}x(t) = f(t, x(t), u(t), q(t_{i-1}, x(t_{i-1}))), \quad t \in [t_{i-1}, t_i], \quad i = 1, \dots, N, \quad x(t_0) = x_0.$$

Krasovskii and Subbotin proved that there exist functions $p^* : [0, T] \times \mathbb{R}^d \rightarrow U$, $q^* : [0, T] \times \mathbb{R}^d \rightarrow V$ such that

$$\begin{aligned} & \limsup_{\delta \downarrow 0} \{\sigma(x^1[T, t_0, x_0, p^*, \Delta, v]) : d(\Delta) \leq \delta, v \in \mathcal{V}[t_0]\} \\ &= \inf_{p \in U^{[0, T] \times \mathbb{R}^d}} \limsup_{\delta \downarrow 0} \{\sigma(x^1[T, t_0, x_0, p, \Delta, v]) : d(\Delta) \leq \delta, v \in \mathcal{V}[t_0]\} \\ &= \sup_{q \in V^{[0, T] \times \mathbb{R}^d}} \liminf_{\delta \downarrow 0} \{\sigma(x^2[T, t_0, x_0, q, \Delta, u]) : d(\Delta) \leq \delta, u \in \mathcal{U}[t_0]\} \\ &= \liminf_{\delta \downarrow 0} \{\sigma(x^2[T, t_0, x_0, q^*, \Delta, u]) : d(\Delta) \leq \delta, u \in \mathcal{U}[t_0]\} = \text{Val}(t_0, x_0). \end{aligned}$$

Here B^A stands for the set of functions from A to B .

The function Val is value function of the differential game. It is a viscosity/minimax solution of the Hamilton-Jacobi PDE

$$\frac{\partial \phi}{\partial t} + H(t, x, \nabla \phi), \quad \phi(T, x) = \sigma(x).$$

Here

$$H(t, x, \xi) \triangleq \min_{u \in U} \max_{v \in V} \langle \xi, f(t, x, u, v) \rangle. \quad (2)$$

Note that given a supersolution of (2) one can construct a suboptimal strategy of the first player. Analogous result holds true for the second player. She is to use a subsolution of (2).

3 General Theory of Markov Games

To evaluate the value function we will use zero-sum Markov game. In the general case a Markov game is defined in the following way. Let S be a set. Assume that S is at most countable. Below S stands for the state space. For any $t \in [0, T]$, $u \in U$, $v \in V$, let $Q(t, u, v) = (Q_{x,y}(t, u, v))_{x,y \in S}$ be a Kolmogorov matrix i.e., for any $t \in [0, T]$, $u \in U$, $v \in V$, $x \in S$,

$$\sum_{y \in S} Q_{x,y}(t, u, v) = 0$$

and for any $y \in S$, $y \neq x$

$$Q_{x,y}(t, u, v) \geq 0.$$

If $u \in \mathcal{U}[t_0]$, $v \in \mathcal{V}[t_0]$ be open-loop controls of the players, the state of the Markov chain at $t \in [0, T]$ is x , then under some continuity conditions the probability of the state $y \neq x$ at time $t^+ > t$ is

$$\int_t^{t^+} Q_{x,y}(\tau, u(\tau), v(\tau)) d\tau + o(\tau);$$

whereas the probability of keeping of x at $[t, t^+]$ is

$$1 + \int_t^{t^+} Q_{x,x}(\tau, u(\tau), v(\tau)) d\tau + o(\tau).$$

Recall that $Q_{x,x}(t, u, v)$ is always nonpositive. If we denote the probability of the state x at time t by $P_x(t)$, we obtain the vector $P(t) = (P_x(t))_{x \in S}$. Below we assume that $P(t)$ is a row-vector. The dynamics of the vector of probabilities $P(t)$ is given by the Kolmogorov equation

$$\frac{d}{dt} P(t) = P(t)Q(t, u(t), v(t)), \quad P(0) = P^0. \quad (3)$$

Here $P^0 = (P_x^0)_{x \in S}$ is an initial distribution.

Assume that the first (receptively, second) player wishes to minimize (respectively, maximize) the terminal payoff given by $E\sigma(X(T))$, where $X(t)$ is a state of the Markov chain corresponding to the Kolmogorov matrix $Q(t, u(t), v(t))$ at time t .

We do not put any analog of Isaacs condition on the Markov games. Thus, we are to suppose that one player is informed about the current player of her partner or use mixed strategies. Within the paper we assume that the first player has an information only on a current position, whereas the second player is informed about a

current position and a current control of the first player. In this case the strategy of the first player is $u(t, x)$, the strategy of the second player $v(t, x, u)$. These strategies produce a Markov chain with the Kolmogorov matrix

$$\widehat{Q}_{x,y}(t) \triangleq Q_{x,y}(t, u(t, x), v(t, x, u(t, x))).$$

Denote the corresponding Markov chain starting at (t_*, x_*) by $X^{t_*, x_*, u, v}(\cdot)$. The corresponding probability is denoted by $P^{t_*, x_*, u, v}$. One may introduce the upper value function by the rule:

$$\eta^+(t_*, x_*) \triangleq \min_u \max_v E^{t_*, x_*, u, v} \sigma(X^{t_*, x_*, u, v}(T)).$$

Here $E^{t_*, x_*, u, v}$ stands for the expectation according to the probability $P^{t_*, x_*, u, v}$.

Under some regularity conditions the value function satisfies Zachrisson equation that is an analog of Isaacs-Bellman equation:

$$\frac{d}{dt} \eta^+(t, x) + \min_{u \in U} \max_{v \in V} \sum_{y \in S} \eta^+(t, y) Q_{x,y}(t, u, v) = 0, \quad \eta^+(T, x) = \sigma(x), \quad x \in S. \quad (4)$$

Note that (4) is a system of ODE. The existence result was established for the case of finite S in [Zachrisson, 1964].

If a solution of the (4) exist one can define the strategies of the players as follows:

$$u^*(t, x) \triangleq \operatorname{argmin}_{u \in U} \max_{v \in V} \left[\sum_{y \in S} \eta^+(t, y) Q_{x,y}(t, u, v) \right];$$

$$v^*(t, x, u) \triangleq \operatorname{argmax}_{v \in V} \left[\sum_{y \in S} \eta^+(t, y) Q_{x,y}(t, u, v) \right].$$

Using the standard verification arguments, one can prove that the strategies u^* and v^* are optimal, i.e.

$$\begin{aligned} \eta^+(t_*, x_*) &= E^{t_*, x_*, u^*, v^*} \sigma(X(T, t_*, x_*, u^*, v^*)) = \min_u E^{t_*, x_*, u, v^*} \sigma(X(T, t_*, x_*, u, v^*)) \\ &= \max_v E^{t_*, x_*, u^*, v} \sigma(X(T, t_*, x_*, u^*, v)) = \min_u \max_v E^{t_*, x_*, u, v} \sigma(X(T, t_*, x_*, u, v)). \end{aligned} \quad (5)$$

4 Approximating Infinite State Markov Chain

Given a differential game with dynamics (1) define the approximating Markov game in the following way.

Let h be a positive number, $f(t, x, u, v) = (f_1(t, x, u, v), \dots, f_d(t, x, u, v))$ and let e^i denote the i -th coordinate vector. Put

$$\chi_i(t, x, u, v) = \begin{cases} e^i, & f_i^1(t, x, u, v) > 0, \\ -e^i, & f_i^1(t, x, u, v) < 0, \\ 0, & f_i^1(t, x, u, v) = 0. \end{cases}$$

Let the state space be $S \triangleq h\mathbb{Z}^d$. Put the Kolmogorov matrix be equal to

$$Q_{xy}^h(t, u, v) = \begin{cases} \frac{1}{h} |f_i(t, x, u, v)|, & y = x + h\chi_i(t, x, u, v), \\ -\frac{1}{h} \sum_{i=1}^d |f_i(t, x, u, v)|, & x = y, \\ 0, & y \neq x, \quad y \neq x + h\chi_i(t, x, u, v), \end{cases} \quad (6)$$

In this case the Zachrisson equation (4) takes the form:

$$\frac{d}{dt} \eta_h^+(t, x) + \min_{u \in U} \max_{v \in V} \sum_{i=1}^d |f_i(t, x, u, v)| \frac{\eta_h^+(t, x + h\chi_i(t, x, u, v)) - \eta_h^+(t, x)}{h} = 0, \quad \eta_h^+(T, x) = \sigma(x). \quad (7)$$

Here $x \in h\mathbb{Z}^d$ is a parameter.

The following statement is proved in [Averboukh, 2016].

Proposition 1. *There exists an unique solution of (7).*

The following theorem is also proved in [Averboukh, 2016].

Theorem 1. *There exists a constant C_1 determined by the function f such that if η_h^+ is a solution of (7), then, for $t_0 \in [0, T]$, $x_0 \in h\mathbb{Z}^d$,*

$$|\text{Val}(t_0, x_0) - \eta_h^+(t_0, x_0)| \leq RC_1\sqrt{h}. \quad (8)$$

Note that in [Averboukh, 2016] it is proved that $C_1 = \sqrt{2dMT}e^{(3+2K)T}$.

Further, the optimal strategies of the players take the form:

$$u^*(t, x) \triangleq \underset{u \in U}{\operatorname{argmin}} \max_{v \in V} \left[\sum_{i=1}^d |f_i(t, x, u, v)| \frac{\eta_h^+(t, x + h\chi_i(t, x, u, v)) - \eta_h^+(t, x)}{h} \right];$$

$$v^*(t, x, u) \triangleq \underset{v \in V}{\operatorname{argmax}} \left[\sum_{i=1}^d |f_i(t, x, u, v)| \frac{\eta_h^+(t, x + h\chi_i(t, x, u, v)) - \eta_h^+(t, x)}{h} \right].$$

5 Approximating finite state Markov chain

Assume that we are interesting in the approximation of the value function in the set

$$\Lambda_r \triangleq \left\{ x \in h\mathbb{Z}^d : \max_{i=1, \dots, d} |x_i| \leq r \right\}.$$

To this end we consider the Markov chain with the state space

$$\Lambda_\rho \triangleq \left\{ x \in h\mathbb{Z}^d : \max_{i=1, \dots, d} |x_i| \leq \rho \right\}.$$

Let

$$\Gamma_\rho \triangleq \left\{ x \in h\mathbb{Z}^d : \max_{i=1, \dots, d} |x_i| = \rho \right\}.$$

Define the Kolmogorov matrix $Q^{\natural, h}(t, u, v) = (Q_{x, y}^{\natural, h}(t, u, v))_{x, y \in \Lambda_\rho}$ in the following way:

$$Q_{x, y}^{\natural, h}(t, u, v) \triangleq \begin{cases} Q_{x, y}^h(t, u, v), & x \in \Lambda_{\rho-h}, \\ 0, & x \in \Gamma_\rho. \end{cases}$$

We assume that the purposes of the player in this game is the same as above i.e. the first (respectively, second) player wishes to minimize (respectively, maximize) the terminal payoff $E\sigma(Y(T))$, where $Y(t)$ denotes the state of the corresponding Markov chain.

For the Kolmogorov matrix $Q^{\natural, h}(t, u, v)$ equation (4) takes the form:

$$\begin{aligned} \frac{d}{dt} \eta_h^+(t, x) + \min_{u \in U} \max_{v \in V} \sum_{i=1}^d |f_i(t, x, u, v)| \frac{\eta_h^+(t, x + h\chi_i(t, x, u, v)) - \eta_h^+(t, x)}{h} &= 0, \quad x \in \Lambda_{\rho-h}, \\ \frac{d}{dt} \eta_h^+(t, x) &= 0, \quad x \in \Gamma_\rho, \\ \eta_h^+(T, x) &= \sigma(x). \end{aligned} \quad (9)$$

Note that system (9) is a finite system of ODEs. Thus, it can be solved using a numerical method.

Let us estimate the difference between $\eta^h(t, x)$ and $\eta^{\natural, h}(t, x)$ for $t \in [0, T]$, $x \in \Lambda_r$. This estimate and Theorem 1 will lead the approximation of the value function by the solution of the finite system of ODEs.

If $u(t, x)$ is a strategy of the first player, $v(t, x, u)$ is a strategy of the second player, then denote the corresponding Markov chain determined by the Kolmogorov matrix $(Q_{x, y}^{\natural, h}(t, u(t, x), v(t, x, u(t, x))))$ and starting at (t_*, x_*) by $Y^{T_*, x_*, u, v}$. Further, denote by $\mathcal{P}^{t_*, x_*, u, v}$ and $\mathcal{E}^{t_*, x_*, u, v}$ the corresponding probability and the expectation respectively.

Without loss of generality we can assume that $\mathcal{P}^{t_*, x_*, u, v}$ and $\mathcal{E}^{t_*, x_*, u, v}$ are defined on the same measurable space.

Let

$$A_\rho^{t_*, x_*, u, v} = \{X^{t_*, x_*, u, v}(t) \notin \Lambda_{\rho-h} \text{ for some } t \in [t_*, T]\}.$$

Analogously, let

$$B_\rho^{t_*, x_*, u, v} = \{Y^{t_*, x_*, u, v}(t) \notin \Lambda_{\rho-h} \text{ for some } t \in [t_*, T]\}.$$

We have that, for any event C such that $C \cap A_\rho^{t_*, x_*, u, v} = C \cap B_\rho^{t_*, x_*, u, v} = \emptyset$,

$$P^{t_*, x_*, u, v}(C) = \mathcal{P}^{t_*, x_*, u, v}(C). \quad (10)$$

Lemma 1. For any $t_* \in [0, T]$, $x_* \in \Lambda_r$, and any strategies of the players u and v ,

$$P^{t_*, x_*, u, v}(A_\rho^{t_*, x_*, u, v}) \leq d \frac{(r^2 + (h + M^2)T^2)e^T}{\rho} \quad (11)$$

$$P^{t_*, x_*, u, v}(B_\rho^{t_*, x_*, u, v}) \leq d \frac{(r^2 + (h + M^2)T^2)e^T}{\rho} \quad (12)$$

Proof. We shall prove estimate (11). To this end given a controlled Markov chain with the Kolmogorov matrix $Q(t, u, v)$ let us introduce the generator acting on the continuous functions by the following rule:

$$L_t[u, v]\varphi(x) = \sum_{y \in S} Q_{x, y}(t, u, v)\varphi(y).$$

It follows from [Kolokoltsov, 2011] that if $X(t)$ is a Markov chain corresponding to the Kolmogorov matrix Q and E is the corresponding probability, then

$$\frac{d}{dt} E\varphi(X(t)) = EL_t[u(t), v(t)]\varphi(X(t)). \quad (13)$$

For the Markov chain with the Kolmogorov matrix $Q^h(t, u, v)$ the generator takes the form:

$$L_t^h[u, v]\varphi = \frac{1}{h} |f_i(t, x, u, v)| \sum_{i=1}^d [\varphi(x + h\chi_i(t, x, u, v)) - \varphi(x)]. \quad (14)$$

Further, for $x = (x_1, \dots, x_d)$, put $\psi_i(x) \triangleq \|x_i\|^2$. Equality (14) implies that for any u, v

$$L_t[u, v]\psi(x) = h|\chi_i(t, x, u, v)|^2 + 2x_i\chi_i(t, x, u, v)|f_i(t, x, u, v)|.$$

Therefore, (13) yields the following equation:

$$\begin{aligned} \frac{d}{dt} E^{t_*, x_*, u, v}(X_i^{t_*, x_*, u, v}(t))^2 &= hE^{t_*, x_*, u, v}|\chi_i(t, X_i^{t_*, x_*, u, v}(t), u, v)|^2 \\ &+ E^{t_*, x_*, u, v}2X_i^{t_*, x_*, u, v}(t)\chi_i(t, X_i^{t_*, x_*, u, v}(t), u, v)|f_i(t, X_i^{t_*, x_*, u, v}(t), u, v)|. \end{aligned}$$

Here $X_i^{t_*, x_*, u, v}(t)$ stands for the i -th coordinate of $X^{t_*, x_*, u, v}(t)$.

Since, $|\chi_i(t, x, u, v)| = 1$, we have that

$$\frac{d}{dt} E^{t_*, x_*, u, v}(X_i^{t_*, x_*, u, v}(t))^2 \leq h + E^{t_*, x_*, u, v}2|f_i(t, X_i^{t_*, x_*, u, v}(t), u, v)||X_i^{t_*, x_*, u, v}(t)|.$$

Since f_i is bounded by M we have that

$$\frac{d}{dt} E^{t_*, x_*, u, v}(X_i^{t_*, x_*, u, v}(t))^2 \leq h + 2ME^{t_*, x_*, u, v}|X_i^{t_*, x_*, u, v}(t)|.$$

Thus,

$$\frac{d}{dt} E^{t_*, x_*, u, v}(X_i^{t_*, x_*, u, v}(t))^2 \leq h + M^2 + E^{t_*, x_*, u, v}(X_i^{t_*, x_*, u, v}(t))^2.$$

Hence,

$$E^{t_*, x_*, u, v} (X_i^{t_*, x_*, u, v}(t))^2 \leq \|x_{*,i}\|^2 + (h + M^2)(t - t_*) + \int_{t_*}^t E^{t_*, x_*, u, v} (X_i^{t_*, x_*, u, v}(\tau))^2 d\tau.$$

Using Gronwall's inequality, we get that

$$E^{t_*, x_*, u, v} (X_i^{t_*, x_*, u, v}(t))^2 \leq (\|x_{*,i}\|^2 + (h + M^2)(t - t_*))e^{t-t_*}.$$

Therefore,

$$E^{t_*, x_*, u, v} \sup_{t' \in [t_*, T]} (X_i^{t_*, x_*, u, v}(t'))^2 \leq \sup_{t' \in [t_*, T]} E^{t_*, x_*, u, v} (X_i^{t_*, x_*, u, v}(t'))^2 \leq (r^2 + (h + M^2)T^2)e^T.$$

Using Markov inequality, we get

$$P \left(\sup_{t \in [t_*, T]} |X_i^{t_*, x_*, u, v}(t)| \geq \rho \right) \leq \frac{(r^2 + (h + M^2)T^2)e^T}{\rho}.$$

Therefore,

$$P(A_\rho^{t_*, x_*, u, v}) \leq \sum_i P \left(\sup_{t \in [t_*, T]} |X_i^{t_*, x_*, u, v}(t)| \geq \rho \right) \leq d \frac{(r^2 + (h + M^2)T^2)e^T}{\rho}.$$

Inequality (12) is proved in the same way. \square

Theorem 2. *There exists a constant C_2 such that for any h one can choose ρ satisfying*

$$|\text{Val}(t_0, x_0) - \eta^{\natural, h}(t_0, x_0)| \leq C_2 h.$$

Proof. We shall estimate $|\eta^h(t_*, x_*) - \eta^{\natural, h}(t_*, x_*)|$.

Further, let $\bar{A}_\rho^{t_*, x_*, u, v}$ (respectively, $\bar{B}_\rho^{t_*, x_*, u, v}$) be a complement of $A^{t_*, x_*, u, v}$ (respectively, $B_\rho^{t_*, x_*, u, v}$).

Using Lemma 1, we get

$$\begin{aligned} \eta^h(t_*, x_*) &= \min_u \max_v E^{t_*, x_*, u, v} \sigma(X^{t_*, x_*, u, v}) \\ &\leq \min_u \max_v E^{t_*, x_*, u, v} \sigma(X^{t_*, x_*, u, v}) \mathbf{1}_{\bar{A}^{t_*, x_*, u, v}} + \max_u \max_v E^{t_*, x_*, u, v} \sigma(X^{t_*, x_*, u, v}) \mathbf{1}_{A^{t_*, x_*, u, v}} \\ &\leq \min_u \max_v \mathcal{E}^{t_*, x_*, u, v} \sigma(Y^{t_*, x_*, u, v}) \mathbf{1}_{\bar{B}^{t_*, x_*, u, v}} + Md \frac{(r^2 + (h + M^2)T^2)e^T}{\rho} \\ &\leq \min_u \max_v \mathcal{E}^{t_*, x_*, u, v} \sigma(Y^{t_*, x_*, u, v}) + 2Md \frac{(r^2 + (h + M^2)T^2)e^T}{\rho} = \eta^{\natural, h}(t_*, x_*) + 2Md \frac{(r^2 + (h + M^2)T^2)e^T}{\rho}. \end{aligned}$$

The opposite inequality is proved in the same way. Thus,

$$|\eta^h(t_*, x_*) - \eta^{\natural, h}(t_*, x_*)| \leq 2Md \frac{(r^2 + (h + M^2)T^2)e^T}{\rho}.$$

Once can choose ρ such that $2Md(r^2 + (h + M^2)T^2)e^T/\rho \leq h$. Using Theorem 1, we get the conclusion of the Theorem. \square

Acknowledgements

This work was supported by the Russian Science Foundation (project no. 17-11-01093).

References

- [Averboukh, 2016] Averboukh, Y. (2016). Approximate solutions of continuous-time stochastic games. *SIAM J. Control Optim.*, 54(5):2629–2649.
- [Botkin et al.,] Botkin, N., Hoffmann, K.-H., and Turova, V. L. Stable numerical schemes for solving hamilton-jacobibellmanisaacs equations. *SIAM J. Control Optim.*, 33(2):9921007.

- [Boue & Dupuis, 1999] Boue, M. and Dupuis, P. (1999). Markov chain approximations for deterministic control problems with affine dynamics and quadratic cost in the control. *Siam J. Numer. Anal.*, 36:667–695.
- [Buslaeva, 1977] Buslaeva, L. T. (1977). Stochastic control in a differential game. *J. Appl. Math. Mech.*, 42(4):609–623.
- [Camilli & Marchi, 2012] Camilli, F. and Marchi, C. (2012). Continuous dependence estimates and homogenization of quasi-monotone systems of fully nonlinear second order parabolic equations. *Nonlinear Anal.*, 75(13):5103–5118.
- [Elliott & Kalton, 1972] Elliott, R. J. and Kalton, N. J. (1972). Values in differential games. *Bull. Amer. Math. Soc.*, 78(3):427–431.
- [Jorgensen & Zaccour, 2006] Jorgensen, S. and Zaccour, G. (2006). Developments in differential game theory and numerical methods: economic and management applications. *Computational Management Science*, 4:159181.
- [Kolokoltsov, 2011] Kolokoltsov, V. N. (2011). *Markov processes, semigroups and generators*, volume 38 of *De Gruyter Studies in Mathematics*. De Gruyter.
- [Kolokoltsov, 2012] Kolokoltsov, V. N. (2012). Nonlinear Markov games on a finite state space (mean-field and binary interactions). *Int. J. Stat. Probab.*, 1(1):77–91.
- [Krasovskii & Kotelnikova, 2009] Krasovskii, N. N. and Kotelnikova, A. N. (2009). Unification of differential games, generalized solutions of the hamilton-jacobi equations, and a stochastic guide. *Differ. Equ.*, 45(11):1653–1668.
- [Krasovskii & Kotelnikova, 2010a] Krasovskii, N. N. and Kotelnikova, A. N. (2010a). An approach-evasion differential game: stochastic guide. *Proc. Steklov Inst. Math.*, 269 Supplement(1 Supplement):191–213.
- [Krasovskii & Kotelnikova, 2010b] Krasovskii, N. N. and Kotelnikova, A. N. (2010b). On a differential interception game. *Proc. Steklov Inst. Math.*, 268(1):161–206.
- [Krasovskii & Subbotin, 1973] Krasovskii, N. N. and Subbotin, A. I. (1973). Approximation in a differential game. *J. Appl. Math. Mech.*, 37(2):185–192.
- [Krasovskii & Subbotin, 1988] Krasovskii, N. N. and Subbotin, A. I. (1988). *Game-theoretical control problems*. Springer, New York.
- [Kushner, 2007] Kushner, H. (2007). Numerical approximations for stochastic differential games: The ergodic cost criterion. In Jorgensen, S., Quincampoix, M., and Vincent, T. L., editors, *Advances in Dynamic Game Theory Numerical Methods, Algorithms, and Applications to Ecology and Economics*, pages 617–638. Springer.
- [Kushner & Dupuis, 2001] Kushner, H. and Dupuis, P. G. (2001). *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, New York.
- [Li & Song, 2007] Li, X. and Song, Q. (2007). Markov chain approximation methods on generalized hjb equation. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 4069–4074, New Orleans, LA, USA.
- [Subbotin, 1995] Subbotin, A. I. (1995). *Generalized solutions of first-order PDEs. The dynamical perspective*. Birkhauser, Boston.
- [Subbotin & Chentsov, 1981] Subbotin, A. I. and Chentsov, A. G. (1981). *Optimization of guarantee in control problems*. Nauka, Moscow. in Russian.
- [Varaya & Lin, 1969] Varaya, P. and Lin, J. (1969). Existence of saddle points in differential games. *Siam J. Control*, 7(1):142–157.
- [Zachrisson, 1964] Zachrisson, L. E. (1964). Markov games. In Dresher, M., Shapley, L. S., and Tucker, A. W., editors, *Advances in game theory*, pages 211–253. Princeton University Press, Princeton.