Krasovskii Institute of Mathematics and Mechanics

as a manuscript

Yurii Averboukh

Probabilistic methods for analysis of game theoretical control problems

Summary of the thesis
for the purpose of obtaining academic degree
Doctor of Science in Mathematics

Moscow
2020

Introduction

The results included in the thesis are concerned with applications of probabilistic methods for game theoretical control problems. We examine both the case of finite player games and the limiting case of infinite player games. Notice that the control theory arises in various fields of science and technique. Among them are robotics, economics, finance and even biology. Additionally, the control theory is closely linked with the theory of partial differential equations due to the dynamical programming principle.

The thesis researches are simulated by the feedback approach to the control theory developed by Krasovskii and his followers. This methodology implies wide use of discontinuous strategies, the multivalued and nonsmooth analysis and the viability theory.

The summary is organized as follows. First, in Section 1 we give a short introduction to the control theory and the theories of differential and mean field games. We start with the classical finite horizon control theory on the finite dimensional state space. Then, we consider zero- and nonzero-sum differential games. Section 1 is completed with a very brief introduction to the mean field type control system and mean field games those are idealized models of differential games with many identical players. In Section 2 we present the thesis' results and the list of publications. Finally, Section 3 provides the detailed description of the results of the thesis.

1 Control theory and differential games. A brief survey

1.1 Finite dimensional control theory

The classic object of the control theory is the study of the dynamic system with the trajectories give by the ordinary differential equation

$$\frac{d}{dt}x(t) = f(t, x(t), u(t)) \tag{1}$$

on a finite time interval. Here $t \in [0, T]$ is the time, $x(t)$ is stands for the instantaneous state of the system, $u(t)$ is the instantaneous control parameter. Usually, it is assumed that $x(t) \in \mathbb{R}^d$, whereas the function $u(\cdot)$ is chosen from some set U by some decision maker acting with some purpose. There are many various problem statements in the control theory. We restrict our attention to two types:

- minimization (maximization) of some criterion;
- ensuring of the viability property.

Notice that other statements include minimization (maximization) of some criterion within various constraints. Additionally, there is a great interest to time optimal problems those imply that the decision maker tries to steer the system to the target as soon as possible.

The problem of minimization of some criterion means that the decision maker wishes to find a control $[0, T] \ni t \mapsto u(t) \in U$ minimizing the quantity

$$\gamma(x(T)) + \int_{t_0}^{T} g(t, x(t), u(t))dt \tag{2}$$

where $x(\cdot)$ satisfies (1) and initial condition $x(t_0) = x_0$. This optimal control problem is an extension of problems examined by the calculus of variations. The main feature of the optimal control theory is that it studies strong extremes. The necessary optimality condition for the optimal control problem is given by the Pontryagin maximum principle reducing the original (infinite dimensional) control problem to the boundary problem for the system of ODEs and the additional maximization condition that should be valid along the optimal trajectory.

The second method used for the computation of optimal control is called the dynamic programming principle. It is based on the observation that each part of the optimal trajectory should be optimal. To introduce it we are to assume that the decision maker can observe the current state of the system. Then, a control policy is a function of time t and state $x(t)$. In other worlds, now the decision maker uses the feedback strategies $u(t, x)$. (The control policies depending only on time are called open-loop strategies.) Notice that for the optimal control systems (1), (2) the open-loop and feedback strategies are equivalent.

If t_0 is an initial time, x_0 is an initial position, then denote

$$\begin{aligned}\mathrm{Val}(t_0, x_0) \triangleq \sup\Big\{\gamma(x(T)) + \int_{t_0}^{T} g(t, x(t), u(t))dt :\\ (x(\cdot), u(\cdot)) \text{ satisfying (1)}, \quad x(t_0) = x_0\Big\}.\end{aligned} \tag{3}$$

The dynamic programming principle implies that the given function φ is the value function if and only if it satisfies the following boundary value problem for following Bellman equation in the generalized (minimax/viscosity) sense [Sub95]:

$$\frac{\partial \varphi}{\partial t} + H(t, x, \nabla\varphi) = 0, \quad \varphi(T, x) = \gamma(x). \tag{4}$$

Here $H : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is the Hamiltonian given by the rule:

$$H(t, x, p) = \min_{u \in U}[\langle p, f(t, x, u)\rangle + g(t, x, u)].$$

Notice that for the considered control problem the Bellman equation is the first-order Hamilton-Jacobi PDE.

Additionally, if $\varphi : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ satisfies equation (4), then the optimal feedback strategy can be computed based on the derivatives of φ. In particular, in the smooth case we have that the optimal strategy is given by

$$u^*(t, x) = \underset{u \in U}{\mathrm{argmin}}[\langle \nabla\varphi(t, x), f(t, x, u)\rangle + g(t, x, u)].$$

The second problem is the viability problem. It is described as follows [Aub09], [CLSW98]. The set $K \subset \mathbb{R}^d$ is called viable with respect to system (1) if, for every time t_0 and every initial position $x_0 \in K$, there exists a control $u(\cdot)$ such that the corresponding trajectory starting at (t_0, x_0) lies at K. The sufficient and necessary conditions of the viability are given using the tangent and normal cones to the set [Aub09], [CLSW98]. Additionally, the optimal control problem can be reformulated as a viability problem. Indeed the value function satisfies the following conditions:

- its hypograph is viable with respect to the extended dynamical system with the dynamics:

$$\frac{d}{dt}x(t) = f(t, x(t), u(t)), \quad \frac{d}{dt}z(t) = -g(t, x(t), u(t));$$

- its epigraph is viable with respect to the dynamical systems:

$$\frac{d}{dt}x(t) = f(t, x(t), u), \quad \frac{d}{dt}z(t) = -g(t, x(t), u)$$

for any constant control u.

Above we consider only deterministic control problems. The stochastic control problem is its natural extension. The most general approach to the stochastic control theory assumes that the dynamics of the system is given by the generator of the Lévy–Khintchine type. Let $\mathcal{D}$ be a subspace of $C(\mathbb{R}^d)$ containing $C_c^2(\mathbb{R}^d)$ and let $L_t[u] : \mathcal{D} \to C(\mathbb{R}^d)$ be given by

$$L_t[u]\phi \triangleq \langle f(t, x, u), \nabla\rangle\phi(x) + \frac{1}{2}\langle G(t, x, u)\nabla, \nabla\rangle\phi(x) + \int_{\mathbb{R}^d} [\phi(x + y) - \phi(x) - \langle y, \nabla\phi(x)\rangle \mathbf{1}_{B_1}(y)]\nu(t, x, u, dy)$$

where G is a symmetric nonnegative matrix, B_a denotes the ball of radius $a > 0$ centered at the origin, ν is a Lévy measure be a generator of the Lévy–Khintchine type. Now it is assumed that payoff functional is

$$\mathbb{E}\left[\gamma(x(T)) + \int_{t_0}^{T} g(t, x(t), u(t))dt\right]. \tag{5}$$

Definition 1.1. We say (following [FS06]) that the 6-tuple $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[s,T]}, P, u, X)$ is a control process on $[s, T]$ admissible for the generator L, if

- $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[s,T]}, P)$ is a filtered probability space;
- u is a $\{\mathcal{F}_t\}_{t\in[s,T]}$-progressively measurable process with values in U; X is a $\{\mathcal{F}_t\}_{t\in[s,T]}$-adapted càdlàg process with values in $\mathbb{R}^d$;
- for any $\phi \in \mathcal{D}$, the process

$$\phi(X(t)) - \int_s^t L_\tau^n[u(\tau)]\phi(X(\tau))d\tau$$

is a $\{\mathcal{F}_t\}_{t\in[s,T]}$-martingale.

The Bellman equation for this type of dynamics takes the form

$$\frac{\partial\varphi}{\partial t} + \max_{u\in U}\Big[(L_t[u]\varphi)(x) + g(t, x, u)\Big] = 0, \quad \varphi(T, x) = \gamma(x). \tag{6}$$

Notice that this general approach is not well-studied. To the best of my knowledge, the dynamic programming principle is developed only under assumption that Bellman equation (6) has a smooth solution [FS06].

The exhaustive results are obtained in the less general cases, namely, in the stochastic control theory [YZ99] and Markov decision theory [GHL03]. Recall that the stochastic control theory examines systems with dynamics given by the stochastic differential equation

$$dX = f(t, X(t), u(t))dt + \sigma(t, X(t), u(t))dW_t.$$

This corresponds to the generator of the Lévy–Khintchine type

$$L_t[u]\phi \triangleq \langle f(t,x,u), \nabla\rangle\phi(x) + \frac{1}{2}\langle G(t,x,u)\nabla, \nabla\rangle\phi(x)$$

with $G(t,x,u) = \sigma(t,x,u)\sigma^T(t,x,u)$.

The Markov decision theory assumes that the dynamics of the system is given by a controlled Markov chain with Kolmogorov matrix equal to $(Q_{x,y}(t,u))_{x,y\in\mathcal{S}}$, where $\mathcal{S}$ is a finite state space. As above, the decision maker tries to maximize the outcome given by (5). The Markov decision problem can be reduced to the control problem for the system determined by Lévy-Khintchine type generator by letting

$$L_t[u]\phi(x) \triangleq \sum_{y\in\mathcal{S}, y\neq x} [\phi(y) - \phi(x)]Q_{x,y}(t,u).$$

Notice that for the stochastic control problem the Bellman equation is the second order PDE, whereas for the Markov decision problem the Bellman equation is the system of ODEs. Thus, the stochastic case is easier to examine from the viewpoint of dynamic programming principle.

1.2 Zero-sum differential games in the finite dimensional phase space

In this section we consider the zero-sum differential games. This problem can be regarded as the extension of control theory to the case when there are two decision makers (called players) with the opposite interests. We consider only finite-horizon differential games with the players' aims given by the payoff function[1]. This mean that the dynamics of the system is given by

$$\frac{d}{dt}x(t) = f(t, x(t), u(t), v(t)), \quad t \in [0,T], \quad x(t) \in \mathbb{R}^d, \quad u(t) \in U, \quad v(t) \in V. \tag{7}$$

We assume that the first player tries to minimize

$$\gamma(x(T)) + \int_{t_0}^{T} g(t, x(t), u(t), v(t))dt$$

whereas the other (second) player prevents him/her.

The crucial point in the differential game theory is the choice of the class of strategies.

First, one can assume that the players' strategies depend only on time. Such strategies are called open-loop. This approach implies that the players are blind and do not use the information about current position. Admitting sighted players we arrive to the concept of feedback strategies $u(t,x)$ and $v(t,x)$.

Given a feedback strategy $u(t,x)$ we have two approaches to construction of a control. First, one can plug directly $u(t,x)$ into dynamics (7). Second, one can use a stepwise scheme involving a finite partition $\Delta = \{t_i\}$ and use the control $u(t_i, x(t_i))$ on the time interval $[t_i, t_{i+1})$.

To realize the first approach we are either to assume continuity of u w.r.t. to x or to use differential inclusions. However, Barabanova (Subbotina) and Subbotin showed

[1] For other statements see [KS88].

that in the general case this approach does not provide optimal strategies comparing with the stepwise schemes [BS71].

The mathematical form of stepwise schemes was first proposed by Krasovskii and Subbotin [KS70] (see also [KS88]). Noticed that in fact this scheme holds on a short memory.

The first player's outcome at the initial position (t_0, x_0) is estimated as follows. For a feedback strategy u, a partition of $[t_0, T]$ $\Delta = \{t_i\}_{i=0}^n$, and some realization of the second player's control $v(\cdot)$, compute that quantity

$$J(t_0, x_0, u, \Delta, v) = \gamma(x(T)) + \int_{t_0}^{T} g(t, x(t), u(t), v(t))dt,$$

where $u(t)$ is a realization of the feedback strategy $u(t, x)$,

$$u(t) = u(t_i, x(t_i)), \text{ when } t \in [t_i, t_{i+1}),$$

$x(\cdot)$ denotes the corresponding trajectory. Recall that the first player wishes to correct his/her correct as frequent as possible and minimize J subject to any action of the second player. Thus he/she can expect the following outcome

$$\mathrm{Val}^+(t_0, x_0) = \inf_u \limsup_{\delta \downarrow 0} \sup\{J(t_0, x_0, u, \Delta, v) : d(\Delta) \leq \delta\}.$$

Here the supremum is taken over any realizations of the second player's control v.

Interchanging the players we can evaluate the expected outcomes of the second player by the function $\mathrm{Val}^-(t_0, x_0)$.

Obviously,

$$\mathrm{Val}^+(t_0, x_0) \geq \mathrm{Val}^-(t_0, x_0).$$

Furthermore, the value functions satisfies the Bellman equations

$$\frac{\partial \varphi}{\partial t} + H(t, x, \nabla \varphi) = 0, \quad \varphi(T, x) = \gamma(x)$$

with the Hamiltonian equal to

$$H^+(t, x, p) \triangleq \min_{u \in U} \max_{v \in V} [\langle p, f(t, x, u, v) \rangle + g(t, x, u, v)]$$

for the case of upper value function and equal to

$$H^-(t, x, p) \triangleq \max_{v \in V} \min_{u \in U} [\langle p, f(t, x, u, v) \rangle + g(t, x, u, v)]$$

for the case of lower value function.

When $H^+(t, x, p) = H^-(t, x, p)$ (this equality is called the Isaacs' condition), we have that (see [BD96], [KS88], [Sub95]) that the game has a value denoted below by Val and

$$\mathrm{Val}^+ = \mathrm{Val}^- = \mathrm{Val}.$$

Notice that the upper and lower value functions are characterized using the viability approach [KS88], [Sub95].

A function $\varphi : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ is called u-stable if, for any $s, r \in [0, T]$, $s < r$, $y \in \mathbb{R}^d$ and constant control of the second player $v \in V$, one can find a pair of functions $(x(\cdot), z(\cdot))$ satisfying

$$(\dot{x}(t), \dot{z}(t)) \in \mathrm{co}\{(f(t, x, u, v), g(t, x, u, v)) : u \in U\}, \quad x(s) = y, \quad z(s) = 0$$

and

$$\varphi(s,y) \geq \varphi(r,x(r)) + z(r).$$

Any u-stable function is an upper estimate of the upper value function. Moreover, the upper value function is itself u-stable.

The v-stability condition is introduced in the same way. A function $\varphi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is called v-stable if, for any $s, r \in [0,T]$, $s < r$, $y \in \mathbb{R}^d$ and constant control of the first player $u \in U$, one can find a solution of the initial value problem for the differential inclusion

$$(\dot{x}(t), \dot{z}(t)) \in \mathrm{co}\{(f(t,x,u,v), g(t,x,u,v)) : v \in V\}, \quad x(s) = y, \quad z(s) = 0$$

such that

$$\varphi(s,y) \leq \varphi(r,x(r)) + z(r).$$

As above, v-stable function provide lower estimates of the lower value function which in turn is v-stable.

Furthermore, given a u-stable function φ, a compact of initial positions $K \subset \mathbb{R}^d$ and a positive number ε one can construct the strategy of the first player $u^{\varepsilon}_{\varphi,K}(t,x)$ providing the outcome not greater than $\varphi(s,y)+\varepsilon$ at any initial position (s,y) from the set K. If we choose φ to be equal to the value function, we obtain ε-optimal strategy at any positions of K. The same can be performed for the second player using v-stable functions.

There several other ways to formalize the notion of differential game. Let us mention only nonanticipative strategies and control with guide strategies.

A nonanticipative strategy of the first players is a mapping α assigning to each masurable control of the second player a measurable control of the first player such that the equality $v_1(t) = v_2(t)$ on $[0,\tau]$, implies that $\alpha[v_1](t) = \alpha[v_2](t)$ $t \in [0,\tau]$. Nonanticpative strategies for the second player are introduced in the same way. Notice that feedback approach and formalization based on nonanticipative strategies are equivalent [SC81].

The control with guide strategies were introduced by Krasovski and Subbotion to provide stability with respect to informational disturbances. The control with guide strategies are the special case of full-memory strategies. The main idea of control with guide strategies is to form the control stepwise using on each step the additional information about the state of the auxiliary control system (guide). Notice that the value functions in the classes of feedback and control with guide strategies coincide [KS88], [SC81].

Above we have mentioned that the dynamic programming for deterministic differential games leads to viscosity/minimax solutions of first-order PDE those are generally non-smooth. To construct the approximately optimal feedback strategies one should use such constructions as quasidifferential or proximal sub- (super-) differentials [Sub95].

The required construction is simplified if we replaced the deterministic dynamics (7) with either dynamics given by the stochastic differential equation]

$$dX(t) = f(t, X(t), u(t), v(t))dt + \sigma(t, X(t), u(t), v(t))dW_t$$

or with the dynamics determined by the continuous-time Markov chain with the Kolmogorov matrix

$$(Q_{x,y}(t,u,v))_{x,y\in\mathcal{S}},$$

where $\mathcal{S}$ is a state space. In the first case, the Bellman equation is the parabolic PDE and the optimal strategy is determined by the derivatives of its solution [HL95]. In the second case the Bellman equation is a system of ODEs, whereas the optimal strategy is given by its solution [Zac64].

Krasovskii and Kotelnikova suggested to use the solutions of the stochastic differential game for construction full-memory strategy those are approximately optimal for the original game [KK10]. This approach can be extended to the general form of continuous-time stochastic game. The thesis includes the results on construction of the approximately optimal strategy for the continuous-time stochastic game based on the solution of the continuous-time stochastic game with various dynamics.

1.3 Nonzero-sum differential games

The general form of the finite-horizon, nonzero-sum differential games in the finite-dimensional space imply that the dynamics is given by

$$\begin{aligned}\frac{d}{dt}x(t) = f(t,x(t), u_1(t), \ldots, u_N(t)),\\ t \in [0,T], \quad x(t) \in \mathbb{R}^d, \quad u_i(t) \in U_i, \quad i = 1, \ldots, N,\end{aligned} \tag{8}$$

where the i-th player controls the variable $u_i(t)$ and wishes to maximize the outcome equal to

$$\gamma_i(x(T)) + \int_{t_0}^{T} g_i(t, x(t), u_1(t), \ldots, u_N(t))dt.$$

Notice that the two-player zero-sum differential game arises when we choose $N = 2$, $\gamma_1 = -\gamma_2$, $g_1 = -g_2$.

There several solution concepts for nonzero-sum games. We reduce our attention only to the Nash equilibrium, which means that the profile of strategies is equilibrium if any unilateral changing is not profitable.

The first way is to use the dynamic programming and reduce the game-theoretical problem to the system of Bellman equations. To apply this approach to the nonzero-sum differential game we are to assume that the following functions are well defined:

$$\begin{aligned}H_i(t, x, p_1, \ldots, p_N) \triangleq \\ \langle p_i, f(t, x, u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_N^*(t, x, p_1, \ldots, p_N))\rangle \\ + g_i(t, x, u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_N^*(t, x, p_1, \ldots, p_N))\end{aligned} \tag{9}$$

where the profile of strategies $(u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_N^*(t, x, p_1, \ldots, p_N))$ satisfy the property:

$$\begin{aligned}\langle p_i, f(t, x, u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_N^*(t, x, p_1, \ldots, p_N))\rangle \\ + g_i(t, x, u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_N^*(t, x, p_1, \ldots, p_N)) = \\ \max_{u_i \in U_i}[\langle p_i, f(t, x, u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_i, \ldots, u_N^*(t, x, p_1, \ldots, p_N))\rangle \\ + g_i(t, x, u_1^*(t, x, p_1, \ldots, p_N), \ldots, u_i, \ldots, u_N^*(t, x, p_1, \ldots, p_N))]\end{aligned} \tag{10}$$

In this case it is proved [Fri70] that if there exists a smooth solution of the system of Bellman equations

$$\frac{\partial \varphi_i}{\partial t} + H_i(t, x, \nabla\varphi_1, \dots, \varphi_N) = 0, \quad \varphi_i(T, x) = \gamma_i(x), \tag{11}$$

then the profile of strategies $(u_i^*(t, x, \nabla\varphi_1(t, x), \dots, \nabla\varphi_N(t, x)))_{i=1,\dots,N}$ provides the feedback Nash equilibrium at any initial position.

Since for the zero-sum games we have no smooth solutions, we are to develop some theory dealing with generalized solution of this system. This was realized only for several particular cases in [BS04a], [BS04b], [CP03].

However, up to now the general existence theorem for system (9), (10) was obtained. Moreover, there is the example demonstrating that, in the general case, this system admits only discontinuous solutions [Ave15].

The second approach to the nonzero-sum differential games is to use punishment strategies. Within this approach it is assumed that the players choose a trajectory that is feasible according to system (8). Any unilateral deviation from this trajectories are punished by other players. A trajectory provides a Nash equilibrium if the expected outcome of each player along it are greater or equal to the value function of the zero-sum differential game with dynamics given by (8) when this player wishes to maximize his/her payoff and other players prevent him/her. The punishment approach can be realized in the class of feedback strategies (in the sense of Krasovskii-Subbotin) for two-player games [Kle93]. When we consider games with N players ($N > 2$) the feedback Nash equilibrium within the punishment approach can be constructed when i-th player affects on his/her variable x_i. In the general case we are to assume that there exists an external center that informs the players about the deviating player [Kle93].

The main disadvantages of the punishment approach are the following.

1. The players are to negotiate about the desired trajectory.

2. There is a multiplicity of equilibria.

3. The construction of Nash equilibrium within the punishment approach relies on incredible threats.

4. The feedback Nash equilibria within punishment approach are not universal i.e. one can not construct a profile of feedback strategies providing the Nash equilibrium at any initial position.

Notice that when we consider continuous-time game with dynamics given either by stochastic differential equation or by Markov chain the system of Bellman equations (11) is replaced with either system of parabolic second-order PDEs for the case of stochastic differential games [Fri72] or by differential inclusion for the case of Markov games [Zac64], [Lev13, Theorems 2 and 5]. The existence theorems in these cases can be proved under rather mild assumptions [Man04], [Man14], [HM19], [Lev13, Theorem 6].

1.4 Controlled mean field dynamics and mean field games

The mean field games (originally proposed by Lasry and Lions [LL06a], [LL06b] and Caines, Malhamé, Huang [HMC05]) and mean field type control problems [AD01] refers to the systems constituted by infinitely many identical particles such that the dynamics of each particle is determined by his/her state, his/her control and the distribution of all other particles. This leads to the dynamic systems in the space of probabilities. Nowadays, it is generally accepted to endow this space with so called Kantorovich distance. Let us briefly describe this concept.

First, we need some additional notation. We denote the set of all Borel probabilities on X by $\mathcal{P}(X)$. If $(\Omega', \mathcal{F}')$ and $(\Omega'', \mathcal{F}'')$ are measurable spaces, m is a probability on $\mathcal{F}'$, $h : \Omega' \to \Omega''$ is measurable, then $h_{\#}m$ denotes the push-forward measure defined by the following rule, for $\Upsilon \in \mathcal{F}''$,

$$(h_{\#}m)(\Upsilon) \triangleq m(h^{-1}(\Upsilon)).$$

The projection operator plays the crucial role in the definition of the Kantorovich distance. If X_1, X_2 are sets, then we denote the natural projection of $X_1 \times X_2$ on X_i by p^i.

For $p \geq 1$, let $\mathcal{P}^p(X)$ denote the set of all probabilities m on X such that, for some (and, consequently, any) $x_0 \in X$

$$\int_X \rho_X^p(x, x_0) m(dx) < \infty.$$

The p-th Kantorovich metric[2] on $\mathcal{P}^p(X)$ is defined by the following rule (see [AGS05], [BK12]): if $m', m'' \in \mathcal{P}(X)$, then

$$W_p(m', m'') \triangleq \left[\inf \left\{ \int_{X \times X} \rho_X^p(x', x'') \pi(d(x', x'')) : \pi \in \Pi(m', m'') \right\} \right]^{1/p},$$

where $\Pi(m', m'')$ is the set of plans between m' and m'' i.e. $\pi \in \Pi(m', m'')$ iff $\pi \in \mathcal{P}(X)$ and $\mathrm{p}^1{}_{\#}\pi = m'$, $\mathrm{p}^2{}_{\#}\pi = m''$.

The space $\mathcal{P}^p(X)$ endowed with the metric W_p is a Polish space provided that X is itself a Polish space [AGS05]. If X is a compact, then $\mathcal{P}^p(X)$ is also compact [AGS05]. However, the σ-compactness property does not follow from the fact that X is σ-compact.

The mean field type control processes can be defined using the generators technique. In this case we assume that the dynamics of each particle is the stochastic process produced by the generator

$$\begin{aligned} L_t[m, u]\phi \triangleq \langle f(t, x, m, u), \nabla \rangle \phi(x) + \frac{1}{2} \langle G(t, x, m, u) \nabla, \nabla \rangle \phi(x) \\ + \int_{\mathbb{R}^d} [\phi(x + y) - \phi(x) - \langle y, \nabla \phi(x) \rangle \mathbf{1}_{B_1}(y)] \nu(t, x, m, u, dy). \end{aligned}$$

Here m is a probability corresponding to the distribution of particles. The control u can be chosen either as a open-loop control i.e. some stochastic process taking values in U or as a feedback control $u(t, x)$.

The first assumption leads to the following definition.

[2]It is also called the Wasserstein distance. The question about proper terminology is explained in [BK12, §1.1].

Definition 1.2. We say that a flow of probabilities $[s,T] \ni t \mapsto m(t) \in \mathcal{P}(\mathbb{R}^d)$ is generated by the generator L if there exists a control process $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\in[s,T]}, P, u, X)$ admissible for the generator $L_t[m(t), u]$, where

$$m(t) = X(t, \cdot)_\# P.$$

In this case the motion of each particle is determined by the generator $L_t[m(t), u]$.

Considering the feedback strategies $u(t, x)$ or $u(t, x, m)$ one can reduce them to the open-loop strategies letting $u(t, \omega) = u(t, X(t, \omega))$ or $u(t, \omega) = u(t, X(t, \omega), m(t))$.

The construction described above comprises the stochastic mean field control systems (controlled McKean-Vlasov dynamics) when the motion of each particle is given by

$$dX(t) = f(t, X(t), m(t), u)dt + \sigma(t, X(t), m(t), u)dW_t, \tag{12}$$

where $m(t)$ is the law of $X(t)$ and more general class of mean field type controlled system with jumps and diffusion. In the thesis we are primary concerned with the deterministic mean field type control systems those are obtained from (12) by letting $\sigma = 0$.

As it was mentioned above, the mean field game theory and mean field type control theory deal with the controlled dynamical in the space of probability measures constituted by identical particles. The difference between the mentioned theories is the solution concepts. The mean field game theory implies that each particle is an independent player who tries to maximize his/her own utility given by

$$\mathbb{E}\left[\gamma(X(T), m(T)) + \int_s^T g(t, X(t), m(t), u(t))dt\right], \tag{13}$$

whereas within the mean field type control theory it is assumed that the aim of all particles is to maximize the common utility given by

$$\gamma'(m(T)) + \mathbb{E}\int_s^T g(t, X(t), m(t), u(t))dt. \tag{14}$$

The mean field game theory examines symmetric Nash equilibria in infinite player continuous-time stochastic games with identical players. Thus, we arrive to the following formal definition.

Definition 1.3. Let t_0 be an initial time, and let $m_0 \in \mathcal{P}^2(\mathbb{R}^d)$ be a given initial distribution of players. We say that a 7-tuple $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\in[t_0,T]}, P, u, X, m(\cdot))$ provides a solution of the mean field game with the dynamics determined by the generator L and payoff of each player determined by (13), if

- $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\in[t_0,T]}, P, u, X)$ is admissible for the generator $L_t[m(t), u]$;
- $m(t) = X(t, \cdot)_\# P$;
- $m(t_0) = m_0$;
- for any $y \in \mathbb{R}^d$ and any $(\Omega', \mathcal{F}', \{\mathcal{F}'\}_{t\in[t_0,T]}, P', u', X')$ admissible for the generator $L_t[m(t), u]$ such that $X(t_0) = y$ P'-a.s., we have that

$$\mathbb{E}\Big[\gamma(X(T), m(T)) + \int_{t_0}^T g(t, X(t), m(t), u(t))dt\Big| X(t_0) = y\Big]$$
$$\geq \mathbb{E}'\Big[\gamma(X'(T), m(T)) + \int_{t_0}^T g(t, X'(t), m(t), u'(t))dt\Big]$$

Here $\mathbb{E}$ (respectively, $\mathbb{E}'$) stands for the expectation with respect to probability P (respectively P'.)

Definition 1.3 lies in the field of so called probabilistic approach to the mean field games [CD18]. A different approach (and more popular) is based on solutions of mean field game system. To introduce it, let us set, for $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]}, P, u, X)$ providing a solution of mean field games,

$$\varphi(s,y) \triangleq \mathbb{E}\Big[\gamma(X(T), m(T)) + \int_s^T g(t, X(t), m(t), u(t))dt \Big| X(s) = y].$$

The function φ is the expected outcome of the representative player who starts at the time t in the state y. Due to the dynamical programming principle, it satisfies (probably in a generalized sense) the following equation

$$\frac{\partial \varphi}{\partial t} + \max_{u\in U}\Big(L_t[m(t), u]\varphi(x) + g(t, x, m(t), u)\Big) = 0, \quad \varphi(T, x) = \gamma(x, m(T)). \qquad (15)$$

Simultaneously, the dynamics of the flow of probabilities $m(\cdot)$ is determined by the equality $m(t) = X(t, \cdot)_{\#}P$ with the stochastic process X determined by the optimal control. If we choose a feedback optimal control u satisfying

$$u^*(t,x) = \underset{u\in U}{\operatorname{argmax}}\Big(L_t[m(t), u]\varphi(x) + g(t, x, m(t), u)\Big), \qquad (16)$$

then the flow of probabilities $m(\cdot)$ obeys the kinetic equation

$$\frac{d}{dt}m(t) = L_t[m(t), u^*(t, \cdot)]m(t), \quad m(t_0) = m_0. \qquad (17)$$

The system of consisting of equation (15), (17) together with condition (16) is called the mean field game system. Its solution determines the solution of the mean field game in the sense of Definition 1.3. The existence result for the mean field game system is proved for the wide range of generators and running costs function g [KLY11], [KY13].

The mean field game can be regarded as a limit of the symmetric equilibria in the finite-player games. This insight becomes a strong statement for several cases including open-loop equilibria in stochastic differential games [Fis17], [Lac17]. The limit behaviour of feedback equilibria is studied with the help of the so called master equation [CDLL19]. This is a differential equation on the product of the finite dimensional space and the space of probability measures. The existence and uniqueness theorem on a given time interval for the master equation is proved only for the case of dynamics given by stochastic differential equation and the coercitive Hamiltonian [GS15]. Additionally, there are several short-time existence results. Approximate Nash equilibria in the finite player continuous-time games can be also constructed by solutions of the limiting mean field game [KLY11], [KTY14].

2 Results and publications

The thesis is concerned with the probabilistic methods for the differential games. First, we studied the two-person differential games (both zero-sum and nonzero-sum). The main object in this part of research is the construction of approximate solutions and estimates of the value function based on solutions of model games. To this end we consider a pair of continuous-time stochastic games. The first game is under our interest; the second one is model game and we are informed about its solution. We constructed an approximate solution for the first game. This result, in particular, provides various estimates of the value functions of the differential games. Second, we studied the games with many particle in the limit when the number of players tends to infinity. Assuming that the players are identical and interact via external media, we arrive to the mean field game theory first proposed by Lasry, Lions and (independently) by Huang, Malhamé, Caines. In this direction we studied the viability property of the system consisting of identical particles and properties of the first-order mean field games.

The main results of the thesis are the following.

1. We considered approximate solution for the zero-sum continuous-time stochastic game using on the solutions of model game. Generally, it is assumed that the dynamics of the original and the model games are different. A suboptimal strategy for the first (respectively, second) player is constructed based on a function satisfying so called u-stability (respectively, v-stability) condition for the model game. This provides the estimates of the value function for the original game by u- and v-stable functions for the model game.

2. The general result was applied to several particular cases. It gives approximations of the value function of the zero-sum differential game by either solutions of Cauchy problems for parabolic PDEs or solutions of the systems of ODEs. Furthermore, using the same methodology, we constructed approximately optimal strategies for the Markov games corresponding to large particle systems with mean field interaction.

3. For the two-person nonzero-sum differential game we constructed the approximate Nash equilibrium in the class of public-signal correlated strategies with memory. The proposed construction relies on a pair of function satisfying stability condition for a model continuous time differential game. It is proved that if the model game converge to the original one, then the limiting points of players' outcomes corresponding to the constructed approximate Nash equilibria lie in the convex hull of the set of Nash value in the class of punishment strategies.

4. We concretized the general construction of approximate Nash equilibrium based on a solution of a model game and designed approximate Nash equilibria using solutions of systems of parabolic PDEs and solutions of system of differential inclusions.

5. We studied the viability condition for the infinite system of identical players obeying deterministic evolution with the mean field interaction. We obtained the Nagumo-type viability theorem. To this end, we proposed an analog of a tangent cone.

6. We proposed the minimax approach to deterministic mean field games. In fact, this approach is the variant of the probabilistic approach that uses probabilities on the space of trajectories. We studied the stability of solutions of mean field

game under stochastic perturbations. Furthermore, we proved that each solution of mean field game within the probabilistic approach is a minimax solution. Additionally, we constructed an approximate Nash equilibrium for a finite player game based on a solution of mean field game.

7. We examined the dependence of the solution of mean field game on initial distribution of players. To this end we introduced the value multifunction that assigns to an initial time and an initial distribution of players a set of expected outcomes of the representative player. We defined a mean field game dynamics and proved that if a given multifunction is viable with respect to this dynamics, then it is a value mulifunction. Furthermore, using the methodology developed for the viability analysis of mean field type control systems, we derived the infinitesimal form of the viability condition. It can be regarded as a generalization of the master equation of mean field games.

List of author's publications submitted for the defense

[1] Yu. Averboukh. Approximate solutions of continuous-time stochastic games. *SIAM J. Control Optim.*, 54(5):2629–2649, 2016.

[2] Yu. Averboukh. Extremal shift rule for continuous-time zero-sum Markov games. *Dyn. Games Appl.*, 7(1):1–20, 2017.

[3] Yu. Averboukh. Approximate public-signal correlated equilibria for nonzero-sum differential games. *SIAM J. Control Optim.*, 57(1):743–772, 2019.

[4] Yu. Averboukh. Markov approximations of nonzero-sum differential games. *Vestnik Udmurtskogo Universiteta. Matematika. Mekhanika. Komp'yuternye Nauki*, 30:3–17, 2020.

[5] Yu. Averboukh. Viability theorem for deterministic mean field type control systems. *Set-Valued Var. Anal.*, 26:993–1008, 2018.

[6] Yu. Averboukh. A minimax approach to mean field games. *Sb. Math.*, 206(7):893–920, 2015.

[7] Yu. Averboukh. Deterministic limit of mean field games associated with nonlinear Markov processes. *Appl. Math. Opt.*, 81:711–738, 2020.

[8] Yu. Averboukh. Viability analysis of the first-order mean field games. *ESAIM Contr. Optim. Ca.*, 26:33, 35 pages, 2020.

3 Main results

3.1 Approximate solutions of continuous-time stochastic game

Here we briefly discuss the main results of papers [1], [2]. In those paper we assumed that we are given with two systems determined by the Lévy–Khintchine type generators $L^i_t[u,v]$, $i=1,2$:

$$L^i_t[u,v]\phi(x) = \frac{1}{2}\langle G^i(t,x,u,v)\nabla,\nabla\rangle\phi(x) + \langle f^i(t,x,u,v),\nabla\rangle\phi(x) \\ + \int_{\mathbb{R}^d}[\phi(x+y)-\phi(x)-\langle y,\nabla\phi(x)\rangle\mathbf{1}_{B_1}(y)]\nu^i(t,x,u,v,dy).$$

Here u (respectively, v) is the control of the first (respectively, second) player. Below we assume that each generator is defined on a space $\mathcal{D}^i\subset C(\mathbb{R}^d)$ that contains $C_b(\mathbb{R}^d)$ as well as the linear and quadratic functions. It is assumed that the first player tries to minimize

$$\mathbb{E}\gamma(X(T));$$

the aim of the second player is opposite. In the following L^1 provides the dynamics of the original system, whereas L^2 determines motions of the model system for which we know the solution or its estimates. The purpose of the mentioned papers is two construct approximate solution in the original system.

Now, let us introduce the definition of strategies. Here we assume that the players can use memory and additional information coming from the model that is stochastic. This drive us to the following concept of stochastic strategies with memory.

Definition 3.1 [1, Definition 1]**.** Let t_0 be an initial time. A strategy of the first player on $[t_0,T]$ is a 5-tuple $\mathfrak{u} = (\Omega^U,\mathcal{F}^U,\{\mathcal{F}^U_s\}_{s\in[t_0,T]},u_{x(\cdot)},P^U_{x(\cdot)})$ satisfying the following conditions:

1. $(\Omega^U,\mathcal{F}^U,\{\mathcal{F}^U_s\}_{s\in[t_0,T]})$ is a filtered space;
2. for each function $x(\cdot)\in\mathbb{D}_{t_0}$, $u_{x(\cdot)}$ is a $\{\mathcal{F}^U_s\}_{s\in[t_0,T]}$-progressive measurable stochastic process with values in U, whereas $P^U_{x(\cdot)}$ is a probability on $(\Omega^U,\mathcal{F}^U,\{\mathcal{F}^U_s\}_{s\in[t_0,T]})$;
3. if, for all $s\in[t_0,t]$, $y(s)=x(s)$, then
   - for any $A\in\mathcal{F}^U_t$, $P^U_{x(\cdot)}(A) = P^U_{y(\cdot)}(A)$;
   - for any $s\in[t_0,t]$, $u_{x(\cdot)}(s) = u_{y(\cdot)}(s)$ $P^U_{x(\cdot)}$-a.s.;
4. for any $t\in[t_0,T]$, the function $(x(\cdot),s,\omega)\mapsto u_{x(\cdot)}(s,\omega)$ is measurable with respect to $\mathbb{F}_{t_0,t}\otimes\mathcal{B}([t_0,t])\otimes\mathcal{F}^U_{t_0,t}$.

Here $x(\cdot)$ is a trajectory chosen from the Skorokhod space $\mathbb{D}_{t_0}\triangleq D([t_0,T],\mathbb{R}^d)$. The symbol $\mathbb{F}_{s,t}$ denotes the σ-algebra on $\mathbb{D}_{t_0}$ generated by events on $[s,t]$.

A strategy $\mathfrak{u} = (\Omega^U,\mathcal{F}^U,\{\mathcal{F}^U_s\}_{s\in[t_0,T]},u_{x(\cdot)},P^U_{x(\cdot)})$ is called *stepwise* if there exists a partition $\Delta=\{t_l\}_{l=1}^r$ of the interval $[t_0,T]$ such that equality $x(t_k)=y(t_k)$, $k=0,\dots,l-1$ implies that $P_{x(\cdot)}(A)=P_{y(\cdot)}(A)$ for any $A\in\mathcal{F}^U_{t_l-0}$ and $u_{x(\cdot)}(s)=u_{y(\cdot)}(s)$ for $s\in[0,t_l)$.

Note that the presented definition of strategy includes feedback strategies, and randomized feedback strategies.

A strategy of the second player is a 5-tuple $\mathfrak{v} = (\Omega^V,\mathcal{F}^V,\{\mathcal{F}^V_s\}_{s\in[t_0,T]},v_{x(\cdot)},P^V_{x(\cdot)})$ satisfying conditions similar to the conditions of Definition 3.1 with $v_{x(\cdot)}$ taking values in V.

Definition 3.2 [1, Definition 2]**.** Let (t_0, x_0) be an initial position, and let $\mathfrak{u} = (\Omega^U, \mathcal{F}^U, \{\mathcal{F}^U_s\}_{s\in[0,T]}, u_{x(\cdot)}, P^U_{x(\cdot)})$, $\mathfrak{v} = (\Omega^V, \mathcal{F}^V, \{\mathcal{F}^V_s\}_{s\in[t_0,T]}, v_{x(\cdot)}, P^V_{x(\cdot)})$ be strategies of the first and the second players respectively. A 5-tuple $(\Omega^X, \mathcal{F}^X, \{\mathcal{F}^X_s\}_{s\in[t_0,T]}, X(\cdot), P)$ is a *realization of the motion generated by the strategies* $\mathfrak{u}$, $\mathfrak{v}$ *and the initial position* (t_0, x_0) if the following conditions hold true:

1. $(\Omega^X, \mathcal{F}^X, \{\mathcal{F}^X_s\}_{s\in[t_0,T]})$ is a filtered space;
2. P is a probability on $(\Omega, \mathcal{F}, \{\mathcal{F}_s\}_{s\in[t_0,T]})$, where $\Omega \triangleq \Omega^X \times \Omega^U \times \Omega^V$, $\mathcal{F} \triangleq \mathcal{F}^X \otimes \mathcal{F}^U \otimes \mathcal{F}^V$, $\mathcal{F}_s \triangleq \mathcal{F}^X_s \otimes \mathcal{F}^U_s \otimes \mathcal{F}^V_s$;
3. $X(\cdot)$ is a $\{\mathcal{F}_s\}_{s\in[t_0,T]}$-adapted process on $(\Omega, \mathcal{F}, \{\mathcal{F}_s\}_{s\in[t_0,T]})$ with values in $\mathbb{R}^d$;
4. $X(t_0) = x_0$ P-a.s;
5. for any $\phi \in \mathcal{D}^1$, the process
$$\phi(X(t)) - \int_{t_0}^{t} L^1_\tau[u(\tau), v(\tau)]\phi(X(\tau))d\tau \tag{18}$$
is a $\{\mathcal{F}_s\}_{s\in[t_0,T]}$-martingale; here u and v are $\{\mathcal{F}_s\}_{s\in[t_0,T]}$-progressively measurable stochastic processes defined by the rules
$$u(\tau, \omega^X, \omega^U, \omega^V) \triangleq u_{X(\cdot,\omega^X,\omega^U,\omega^V)}(\tau, \omega^U),$$
$$v(\tau, \omega^X, \omega^U, \omega^V) \triangleq v_{X(\cdot,\omega^X,\omega^U,\omega^V)}(\tau, \omega^V),$$
where $(\omega^X, \omega^U, \omega^V) \in \Omega$;
6. for any $x(\cdot) \in \mathbb{D}_{t_0}$ and any random variable ϕ' on $(\Omega^U, \mathcal{F}^U)$,
$$\mathbb{E}^U_{x(\cdot)}\phi' = \mathbb{E}(\phi'|X(\cdot) = x(\cdot)),$$
where $\mathbb{E}^U_{x(\cdot)}$ denotes the expectation corresponding to the probability $P^U_{x(\cdot)}$;
7. for any $x(\cdot) \in \mathbb{D}_{t_0}$ and any random variable ϕ'' on $(\Omega^V, \mathcal{F}^V)$,
$$\mathbb{E}^V_{x(\cdot)}\phi'' = \mathbb{E}(\phi''|X(\cdot) = x(\cdot)),$$
where $\mathbb{E}^V_{x(\cdot)}$ denotes the expectation corresponding to the probability $P^V_{x(\cdot)}$.

Notice that if both strategies are stepwise, then one can prove the existence of at least one realization. However, given the strategies $\mathfrak{u}$, $\mathfrak{v}$, the outcome is not defined in the unique way. The values

$$J^*(t_0, x_0, \mathfrak{u}, \mathfrak{v}) \triangleq \sup\{\mathbb{E}g(X(T)) : (\Omega^X, \mathcal{F}^X, \{\mathcal{F}^X_s\}_{s\in[0,T]}, X(\cdot), P) \text{ realizing a motion generated by the strategies } \mathfrak{u} \text{ and } \mathfrak{v} \text{ and the initial position } (t_0, x_0)\},$$

$$J_*(t_0, x_0, \mathfrak{u}, \mathfrak{v}) \triangleq \inf\{\mathbb{E}g(X(T)) : (\Omega^X, \mathcal{F}^X, \{\mathcal{F}^X_s\}_{s\in[0,T]}, X(\cdot), P) \text{ realizing a motion generated by the strategies } \mathfrak{u} \text{ and } \mathfrak{v} \text{ and the initial position } (t_0, x_0)\}$$

are the upper and lower outcomes according to the strategies $\mathfrak{u}$ and $\mathfrak{v}$. The upper value of the game is

$$\mathrm{Val}_+(t_0, x_0) = \inf_{\mathfrak{u}} \sup_{\mathfrak{v}} J^*(t_0, x_0, \mathfrak{u}, \mathfrak{v}).$$

The lower value is equal to

$$\mathrm{Val}_-(t_0, x_0) = \sup_{\mathfrak{v}} \inf_{\mathfrak{u}} J_*(t_0, x_0, \mathfrak{u}, \mathfrak{v}).$$

Obviously,

$$\mathrm{Val}_-(t_0, x_0) \leq \mathrm{Val}_+(t_0, x_0).$$

The main result of paper [1] is the constructions of suboptimal strategies and the estimate upper and lower value functions for the original game using the model of the game given by the generator L^2.

We assume that the solution of the model game is given in the form of stable function.

Definition 3.3 [1, Definition 4]**.** A function $c_+ : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ is said to be u-stable with respect to the generator L^2 if

1. $c_+(T, x) = g(x)$;
2. for any $t, \theta \in [0, T]$, $t < \theta$, there exists a filtered space $(\widetilde{\Omega}^{t,\theta}, \widetilde{\mathcal{F}}^{t,\theta}, \{\widetilde{\mathcal{F}}_s^{t,\theta}\}_{s\in[t,\theta]})$ such that, for any $y \in \mathbb{R}^d$, $v \in V$, one can find a $\{\widetilde{\mathcal{F}}_s^{t,\theta}\}_{s\in[t,\theta]}$-progressively measurable generalized control of the first player on $[t, \theta]$ $\mu_{y,v}^{t,\theta}$, a $\{\widetilde{\mathcal{F}}_s^{t,\theta}\}_{s\in[t,\theta]}$-adapted process $Y_{y,v}^{t,\theta}$ with values in $\mathbb{R}^d$, and a probability $\widetilde{P}_{y,v}^{t,\theta}$ on $\widetilde{\Omega}^{t,\theta}$ such that $Y_{y,v}^{t,\theta}(t) = y$ $\widetilde{P}_{y,v}^{t,\theta}$-a.s., for any $\phi \in \mathcal{D}^2$,
$$\phi(Y_{y,v}^{t,\theta}(s)) - \int_t^s \int_U L_\tau^2[w, v]\phi(Y_{y,v}^{t,\theta}(\tau))\mu_{y,v}^{t,\theta}(\tau, dw)d\tau \qquad (19)$$
is a $\{\widetilde{\mathcal{F}}_s^{t,\theta}\}_{s\in[t,\theta]}$-martingale and
$$c_+(t, y) \geq \widetilde{\mathbb{E}}_{y,v}^{t,\theta} c_+(\theta, Y_{y,v}^{t,\theta}(\theta)); \qquad (20)$$
3. for any random variable ϕ on $\widetilde{\Omega}^{t,\theta}$, the dependence of $\widetilde{E}_{y,v}^{t,\theta}\phi$ on y and v is measurable;
4. for any $\phi \in \mathcal{D}^2$, the function $(y, v, s) \mapsto \widetilde{E}_{y,v}^{t,\theta}\phi(Y_{y,v}^{t,\theta}(s))$ is measurable.

Here $\widetilde{\mathbb{E}}_{y,v}^{t,\theta}$ denotes the expectation corresponding to the probability $\widetilde{P}_{y,v}^{t,\theta}$.

Further, put

$$\Sigma^i(t, x, u, v) \triangleq \sum_{j=1}^d G_{jj}^i(t, x, u, v) + \int_{\mathbb{R}^d} \|y\|^2 \nu^i(t, x, u, v, dy);$$

$$b^i(t, x, u, v) \triangleq f^i(t, x, u, v) + \int_{\mathbb{R}^d \setminus B_1} y\nu^i(t, x, u, v, dy).$$

The function Σ^i estimates the stochasticity of the generator L^i, whilst b^i plays the role of the effective drifts. We impose the continuity conditions (those include Lipschitz continuity of b^i), as well as the boundness of Σ^i and b^i. The most important assumption is the following analog of the Isaacs' condition which states that, for at least one $i = 1, 2$ and every $t \in [0, T]$, $x, p \in \mathbb{R}^d$, $u \in U$, $v \in V$,

$$\min_{u \in U} \max_{v \in V} \langle p, b^i(t, x, u, v) \rangle = \max_{v \in V} \min_{u \in U} \langle p, b^i(t, x, u, v) \rangle.$$

Further, let M_0^i be a constant such that

$$|\Sigma^i(t, x, u, v)| \leq M_0^i \quad t \in [0, T], \quad x \in \mathbb{R}^d, \quad u \in U, \quad v \in V,$$

K_i be a Lipschitz constant for the function $x \mapsto b^i(t, x, u, v)$,

$$\varkappa \triangleq \frac{1}{2} \sup_{t \in [0,T], x \in \mathbb{R}^d, u \in U, v \in V} \|b^1(t, x, u, v) - b^2(t, x, u, v)\|^2,$$

$$\epsilon \triangleq \varkappa + M_0^1 + M_0^2. \tag{21}$$

Now, let choose i such that for b^i the Isaacs' condition is fulfilled and set

$$\beta \triangleq 3 + 2K^i. \tag{22}$$

Finally, put $C \triangleq \sqrt{2Te^{\beta T}}$.

Theorem 3.4 [1, Theorem 8]. *If c_+ is u-stable with respect to L^2, then one can constructively build a strategy $\widehat{\mathfrak{u}}_\Delta$ such that, for any $(t_0, x_0) \in [0, T] \times \mathbb{R}^d$,*

$$\lim_{\delta \downarrow 0} \sup \{ J(t_0, x_0, \widehat{\mathfrak{u}}_\Delta, \mathfrak{v}) : d(\Delta) \leq \delta \} \leq c_+(t_0, x_0) + R \cdot C\sqrt{\epsilon}.$$

Here R is the Lipschitz constant for the payoff function γ. The proof of this theorem is based on the variant of Krasovski-Subbotin extremal shift rule for the continuous-time stochastic systems.

Using the definition of the value function we obtain the following.

Corollary 3.5 [1, Corollary 9]. *If c_+ is u-stable with respect to L^2, then, for $(t_0, x_0) \in [0, T] \times \mathbb{R}^d$,*

$$\mathrm{Val}_+(t_0, x_0) \leq c_+(t_0, x_0) + R \cdot C\sqrt{\epsilon}.$$

The notion of v-stability can be introduced in the same spirit as u-stability by interchanging of the players. Using it we obtain the following estimate

Corollary 3.6 [1, Corollary 10]. *If c_- is v-stable with respect to L^2, then, for $(t_0, x_0) \in [0, T] \times \mathbb{R}^d$,*

$$c_-(t_0, x_0) - R \cdot C\sqrt{\epsilon} \leq \mathrm{Val}_-(t_0, x_0).$$

3.2 Particular cases

Let us present estimates for some concrete class of continuous-time differential games.

3.2.1 Stochastic model for differential game

We assume here that the original game is given by the deterministic evolution

$$\frac{d}{dt}x(t) = f(t, x(t), u, v), \tag{23}$$

whereas the model system is described by the stochastic differential equation

$$dX(t) = f(t, X(t), u, t)dt + \sigma dW_t.$$

The generators now are equal to

$$L_t^1[u, v]\phi(x) = \langle f(t, x, u, v), \nabla\phi(x)\rangle,$$

$$L_\tau^2[u, v]\phi(x) = \langle\nabla\phi(x), f^1(t, x, u, v)\rangle + \frac{\sigma^2}{2}\cdot\triangle\phi(x).$$

We assume that the Isaacs' condition is fulfilled for the original game. Thus, the deterministic game has the value. As above, we denote it by Val. Applying the general theory we get the following.

Proposition 3.7 Proposition 16 and Corollary 18 of [1]. *If ψ_σ is a solution of*

$$\frac{\partial\psi}{\partial t} + \min_{u\in U}\max_{v\in V}\langle\nabla\psi, f^1(t, x, u, v)\rangle + \frac{\sigma^2}{2}\triangle\psi = 0, \quad \psi(T, x) = \gamma(x),$$

then ψ_σ is u- and v-stable with respect to the generator

$$L_\tau^2[u, v]\phi(x) = \langle\nabla\phi(x), f^1(t, x, u, v)\rangle + \frac{\sigma^2}{2}\cdot\triangle\phi(x). \tag{24}$$

Thus, there exists a constant C_1 such that

$$|\mathrm{Val}(t_0, x_0) - \psi_\sigma(t_0, x_0)| \leq RC_1\sigma.$$

3.2.2 Markov model for differential game

As above we consider the deterministic differential game with the dynamics given by (23). Now let us introduce a model system governed by a Markov chain. Let h be a positive number, $f^1(t, x, u, v) = (f_1^1(t, x, u, v), \ldots, f_d^1(t, x, u, v))$ and let e^i denote the i-th coordinate vector. Put

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

The quantity $\chi_i(t, x, u, v)$ indicates the direction of motion along the i-th axe according to the dynamics $f^1(t, x, u, v)$. For $A \subset \mathbb{R}^d$

$$\nu^2(t, x, u, v, A) \triangleq \frac{1}{h}\sum_{i=1}^{n}|f_i(t, x, u, v)|\delta_{h\chi_i(t,x,u,v)}(A).$$

Recall that δ_z denotes the Dirac measure concentrated at z.

Further, set

$$L_t^2[u,v]\phi(x) \triangleq \int_{\mathbb{R}^d} [\phi(x+y)-\phi(x)]\nu^2(t,x,u,v,dy) = \sum_{i=1}^{n} |f_i(t,x,u,v)| \frac{\phi(x+h\chi_i(t,x,u,v))-\phi(x)}{h}. \tag{25}$$

This generator corresponds to the continuous-time Markov chain on $h\mathbb{Z}^d$ with the Kolmogorov matrix

$$Q_{xy}^h(t,u,v) = \begin{cases} \frac{1}{h}|f_i(t,x,u,v)|, & y = x + h\chi_i(t,x,u,v), \\ -\frac{1}{h}\sum_{i=1}^d |f_i(t,x,u,v)|, & x = y, \\ 0, & y \neq x, \quad y \neq x + h\chi_i(t,x,u,v), \end{cases} \tag{26}$$

The following system of ODEs is the Isaacs–Bellman equation for the Markov game with the Kolmogorov matrix given by (26):

$$\frac{d}{dt}\zeta_h^+(t,x) + \min_{u\in U}\max_{v\in V} \sum_{i=1}^d |f_i(t,x,u,v)| \frac{\zeta_h^+(t,x+h\chi_i(t,x,u,v)) - \zeta_h^+(t,x)}{h} = 0, \qquad \zeta_h^+(T,x) = \gamma(x). \tag{27}$$

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

Theorem 3.8 [1, Proposition 20 and Theorem 21]**.** *Equation (27) has an unique solution. It provides the following estimate for the value function of the differential game:*

$$|\mathrm{Val}(t_0,x_0) - \zeta_h^+(t_0,x_0)| \leq RC_2\sqrt{h}.$$

Here C_2 is a constant determined by the function f.

3.2.3 Deterministic model for mean field interacting particle system

This result is concerned with the construction of the optimal strategies for the control of the system consisting of N interacting particles taking only finite number of states. We assume that

- the state space for each player is $\{1,\dots,d\}$, where d is a natural number;
- the dynamics of each particle obeys continuous-time Markov chain with the Kolmogorov matrox $Q(t,x,u,v) = (Q_{i,j}(t,x,u,v))_{i,j=\overline{1,d}}$; here $x = (x_1,\dots,x_d)$ is a vector describing the density of the states.

Notice that $x \in \mathbb{R}^d$, $x_i \geq 0$, $\sum_{i=1}^d x_i = 1$. Assuming that $h = 1/N$, we obtain that the examined mean field interacting particle system can be described by the genarator

$$L_t^1[u,v]\phi(x) = \sum_{i,j=1}^d \frac{1}{h} x_i Q_{ij}(t,x,u(t),v(t))[\phi(x-he^i+he^j)-\phi(x)].$$

As above, e^i stands for the i-th coordinate vector.

Letting $h \to 0$, we arrive at the following limiting system which serves as the model:

$$L_t^2[u,v]\phi(x) = \sum_{k=1}^{d} \sum_{i \neq k} [x_i Q_{ik}(t,x,u(t),v(t)) - x_k Q_{ki}(t,x,u(t),v(t))] \frac{\partial \phi}{\partial x_k}(x).$$

The characteristics of this system solve the ODE

$$\frac{d}{dt}x(t) = x(t)Q(t,x(t),u(t),v(t)). \tag{28}$$

In [2] (independently of [1]) the following result was obtained.

Theorem 3.9 [2, Theorem 1]**.** *Assume that c_+ is u-stable for the differential game with the dynamics (28), then one can define the strategy $\widehat{\mathfrak{u}}_\Delta$ such that, for any $(t_0,x_0) \in [0,T] \times \mathbb{R}^d$,*

$$\lim_{\delta \downarrow 0} \sup\{J(t_0,x_0,\widehat{\mathfrak{u}}_\Delta,\mathfrak{v}) : d(\Delta) \leq \delta\} \leq c_+(t_0,x_0) + R \cdot C_3\sqrt{h}.$$

Here C_3 is a constant dependent only on the Kolmogorov matrix Q and the number of states d.

An analogous property is fulfilled for the v-stable function. Finally, if we assume that the game with the mean field interacting particle dynamics for N particles ($N = 1/h$) has the value, then

$$|\mathrm{Val}^h(t_0,x_0) - \mathrm{Val}(t_0,x_0)| \leq RC_3\sqrt{h}.$$

Here we denote by Val^h the value of the game for N particle system ($N = 1/h$); when Val stands for the value of the limiting game with the dynamics given by (28).

3.3 Approximate equilibria in the nonzero-sum games

Here we extend the methodology developed in [1] to the nonzero-sum games. Recall that in many case the nonzero-sum stochastic differential games and the Markov games are easier to examine than the nonzero-sum differential. Thus, it is tempting to construct an approximate equilibrium for the differential game based on a solution of a continuous-time stochastic game.

We consider the differential game with the dynamics

$$\dot{x} = f_1(t,x,u) + f_2(t,x,v), \quad t \in [0,T], x \in \mathbb{R}^d, \quad u \in U, \quad v \in V. \tag{29}$$

Here u (respectively, v) denotes the control of the first (respectively, second) player. We assume that the purpose of the i-th player is to maximize the terminal payoff $\gamma_i(x(T))$. Below we assume that U and V are metric compacts. For the sake of shortness we use the notations: $f^1(t,x,u,v) \triangleq f_1(t,x,u) + f_2(t,x,v)$.

The approximate Nash equilibrium is built in the class of public-signal correlated strategies with memory. Informally, this class can be described as follows. We assume that both players at each time observe the random signal that is produced by an external device. Below this information will be a forecasting of a state of a game being a stochastic model of the original game. The players form their control using this shared information and the history of the game.

This idea can be formalized in the following way.

Definition 3.10 [3, Definition 2.1]**.** A 6-tuple $\mathfrak{w} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]}, u_{x(\cdot)}, v_{x(\cdot)}, P_{x(\cdot)})$ is called a profile of public-signal correlated strategies on $[t_0, T]$ if

(i) $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]})$ is a measurable space with a filtration;

(ii) for each $x(\cdot) \in C([t_0, T]; \mathbb{R}^d)$, $P_{x(\cdot)}$ is a probability on $\mathcal{F}$;

(iii) for each $x(\cdot) \in C([t_0, T]; \mathbb{R}^d)$, $u_{x(\cdot)}$ (respectively, $v_{x(\cdot)}$) is a $\{\mathcal{F}_t\}_{t\in[t_0,T]}$-progressively measurable process taking values in U (respectively, V);

(iv) if $x(t) = y(t)$ for all $t \in [t_0, r]$, then

- for any $A \in \mathcal{F}_r$, $P_{x(\cdot)}(A) = P_{y(\cdot)}(A)$,
- for any $t \in [t_0, r]$, $u_{x(\cdot)}(t) = u_{y(\cdot)}(t)$, $v_{x(\cdot)}(t) = v_{y(\cdot)}(t)$ $P_{x(\cdot)}$-a.s.

(v) for any r, the restrictions of functions $(x(\cdot), t, \omega) \mapsto u_{x(\cdot)}(t, \omega)$, $(x(\cdot), t, \omega) \mapsto v_{x(\cdot)}(t, \omega)$ on $C([t_0, T]; \mathbb{R}^d) \times [t_0, r] \times \Omega$ are measurable with respect to $\mathbb{F}_{t_0,r} \otimes \mathcal{B}([t_0, r]) \otimes \mathcal{F}_r$;

(vi) for any $A \in \mathcal{F}$, the function $x(\cdot) \mapsto P_{x(\cdot)}(A)$ is measurable with respect to $\mathbb{F}_{t_0,T}$.

The key ingredient of the Nash equilibrium is the unilateral deviation. Considering the public-signal correlated profile of strategies, we are to assume that the deviating player can use some additional external device. This lead to the following definition.

Definition 3.11 [3, Definition 2.2]**.** Given a profile of public-signal correlated strategies $\mathfrak{w} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]}, P_{x(\cdot)}, u_{x(\cdot)}, v_{x(\cdot)})$, we say that a profile of strategies $\mathfrak{w}^c = (\Omega^c, \mathcal{F}^c, \{\mathcal{F}^c_t\}_{t\in[t_0,T]}, P^c_{x(\cdot)}, u^c_{x(\cdot)}, v^c_{x(\cdot)})$ is an unilateral deviation by the first (respectively, the second) player if there exists a filtered measurable space $(\Omega', \mathcal{F}', \{\mathcal{F}'\}_{t\in[t_0,T]})$ such that

(i) $\Omega^c = \Omega \times \Omega'$;

(ii) $\mathcal{F}^c = \mathcal{F} \otimes \mathcal{F}'$;

(iii) $\mathcal{F}^c_t = \mathcal{F}_t \otimes \mathcal{F}'_t$ for $t \in [t_0, T]$;

(iv) for any $x(\cdot) \in C([t_0, T]; \mathbb{R}^d)$ and any $A \in \mathcal{F}$, $P^c_{x(\cdot)}(A \times \Omega') = P_{x(\cdot)}(A)$;

(v) for any $x(\cdot)$, $t \in [t_0, T]$, $\omega \in \Omega$, $\omega' \in \Omega'$, $v_{x(\cdot)}(t, \omega, \omega') = v_{x(\cdot)}(t, \omega)$ (respectively, $u_{x(\cdot)}(t, \omega, \omega') = u_{x(\cdot)}(t, \omega)$).

Now let us introduce the motion generated by the public-signal correlated profile of strategies.

Definition 3.12 [3, Definition 2.3]**.** Let $t_0 \in [0, T]$, $x_0 \in \mathbb{R}^d$, $\mathfrak{w} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[t_0,T]}, P_{x(\cdot)}, u_{x(\cdot)}, v_{x(\cdot)})$ be a profile of public-signal correlated strategies on $[t_0, T]$. We say that a pair $(X(\cdot), P)$ is a realization of the motion generated by $\mathfrak{w}$ and initial position (t_0, x_0) if

(i) P is a probability on $\mathcal{F}$;

(ii) $X(\cdot)$ is a $\{\mathcal{F}_t\}_{t\in[t_0,T]}$-adapted process taking values in $\mathbb{R}^d$;

(iii) $X(t_0) = x_0$ P-a.s.;

(iv) for P-a.e. $\omega \in \Omega$,

$$\frac{d}{dt}X(t,\omega) = f_1(t, X(t,\omega), u_{X(\cdot,\omega)}(t,\omega)) + f_2(t, X(t,\omega), v_{X(\cdot,\omega)}(t,\omega)).$$

(v) $P_{x(\cdot)} = P(\cdot|X(\cdot) = x(\cdot))$ i.e. given $A \in \mathcal{F}$,

$$P(A) = \int_{C([t_0,T];\mathbb{R}^d)} P_{x(\cdot)}(A)\chi(d(x(\cdot))),$$

where χ is a probability on $C([t_0,T];\mathbb{R}^d)$ defined by the rule: for any $\Upsilon \in \mathbb{F}_{t_0,T}$, $\chi(\Upsilon) \triangleq P\{\omega : X(\cdot,\omega) \in \Upsilon\}$.

Notice that if the public-signal correlated profile of strategies is stepwise (i.e. the strategies is determined only by history at the finite number of time instances), then there exist at least one realization.

Notice that dynamics (8) corresponds to the generator

$$L^1_t[u,v]\phi(x) \triangleq \langle f_1(t,x,u) + f_2(t,x,v), \nabla\phi(x)\rangle.$$

To construct an approximate equilibrium we will use the continuous-time stochastic game with the dynamics given by

$$(L^2_t[u,v]\phi)(x) \triangleq \frac{1}{2}\langle G^2(t,x,u,v)\nabla, \nabla\rangle\phi(x) + \langle f^2(t,x,u,v), \nabla\rangle\phi(x) \\ + \int_{\mathbb{R}^d}[\phi(x+y) - \phi(x) - \langle y, \nabla\phi(x)\rangle \mathbf{1}_{B_1}(y)]\nu^2(t,x,u,v,dy). \quad (30)$$

We assume that the objective function of the player i in the auxiliary stochastic game is equal to

$$\mathbb{E}\left[\gamma_i(X(T)) + \int_{t_0}^T g_i(t, X(t), u(t), v(t))dt\right]. \quad (31)$$

The following stability condition plays a key role in the construction of the approximate public-signal correlated equilibrium.

Definition 3.13 [3, Definition 3.2]. Let $c_1, c_2 : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ be continuous functions. We say that the pair (c_1, c_2) satisfies *Condition* ($\mathcal{C}$) if, for any $s, r \in [0,T]$, $s < r$, there exists a filtered measurable space $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}^{s,r}_t\}_{t\in[s,r]})$ satisfying the following properties:

(i) given $y \in \mathbb{R}^d$, one can find processes $\eta^{s,r}_y$, $\widehat{Y}^{s,r}_y$ and a probability $\widehat{P}^{s,r}_y$ such that the 6-tuple $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}^{s,r}_t\}_{t\in[s,r]}, \widehat{P}^{s,r}_y, \eta^{s,r}_y, \widehat{Y}^{s,r}_y)$ is a control system admissible for $L^2_t[u,v]$ and, for $i = 1, 2$,

$$\widehat{\mathbb{E}}^{s,r}_y\left[c_i(r, \widehat{Y}^{s,r}_y(r)) + \int_s^r \int_{U\times V} g_i(t, \widehat{Y}^{s,r}_y(t), u, v)\eta^{s,r}_y(t, d(u,v))dt\right] = c_i(s,y);$$

(ii) for any $y \in \mathbb{R}^d$ and $v \in V$, one can find a relaxed stochastic control of the first player $\mu^{s,r}_{y,v}$, a process $\overline{Y}^{1,s,r}_{y,v}$ taking values in $\mathbb{R}^d$ and a probability $\overline{P}^{1,s,r}_{y,v}$ such that the 6-tuple $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}^{s,r}_t\}_{t\in[s,r]}, \overline{P}^{1,s,r}_{y,v}, \mu^{s,r}_{y,v} \otimes \delta_v, \overline{Y}^{1,s,r}_{y,v})$ is a control system admissible for $L^2_t[u, v]$ and

$$\overline{\mathbb{E}}^{1,s,r}_{y,v}\left[c_2(r, \overline{Y}^{1,s,r}_{y,v}(r)) + \int_s^r \int_U g_2(t, \overline{Y}^{1,s,r}_{y,v}(t), u, v)\mu^{s,r}_{y,v}(t, du)dt\right] \leq c_2(s, y);$$

(iii) given $y \in \mathbb{R}^d$ and $u \in U$, one can find a second player's relaxed stochastic control $\nu^{s,r}_{y,u}$, a process $\overline{Y}^{2,s,r}_{y,u}$ and a probability $\overline{P}^{2,s,r}_{y,u}$ such that the 6-tuple $(\widehat{\Omega}^{s,r}, \widehat{\mathcal{F}}^{s,r}, \{\widehat{\mathcal{F}}^{s,r}_t\}_{t\in[s,r]}, \overline{P}^{2,s,r}_{y,u}, \delta_u \otimes \nu^{s,r}_{y,u}, \overline{Y}^{2,s,r}_{y,u})$ is a control system admissible for $L^2_t[u, v]$ and

$$\overline{\mathbb{E}}^{2,s,r}_{y,u}\left[c_1(r, \overline{Y}^{2,s,r}_{y,u}(r)) + \int_s^r \int_V g_1(t, \overline{Y}^{2,s,r}_{y,u}(t), u, v)\nu^{s,r}_{y,u}(t, dv)dt\right] \leq c_1(s, y).$$

Here $\widehat{\mathbb{E}}^{s,r}_y$ (respectively, $\overline{\mathbb{E}}^{1,s,r}_{y,u}$, $\overline{\mathbb{E}}^{2,s,r}_{y,u}$) denotes the expectation according to the probability $\widehat{P}^{s,r}_y$ (respectively, $\overline{P}^{1,s,r}_{y,u}$, $\overline{P}^{2,s,r}_{y,u}$).

Informally speaking, the meaning of Condition $(\mathcal{C})$ is as follows. The first part of this condition means that both players can maintain the value $(c_1(s, y), c_2(s, y))$ on the time interval $[s, r]$ choosing an appropriate controlled stochastic system. Parts (ii), (iii) mean that if player i uses a constant control on $[s, r]$, then the other player can find a control such that the outcome of the player i on $[s, r]$ is not greater than $c_i(s, y)$. Here we assume that the terminal part of the i-th player's reward on $[s, r]$ is given by $c_i(r, \cdot)$. Additionally, to avoid technical issues we assume that all mentioned controlled systems exploit the same filtered measurable space.

Now let us assume that the model system is close to the original one i.e., for any $t \in [0, T]$, $x \in \mathbb{R}^d$, $u \in U$, $v \in V$,

$$|\Sigma^2(t, x, u, v)| \leq \delta^2, \quad \|f^1(t, x, u, v) - b^2(t, x, u, v)\|^2 \leq 2\delta^2, \quad |g_i(t, x, u, v)| \leq \delta.$$

Here, as above,

$$\Sigma^2(t, x, u, v) \triangleq \sum_{j=1}^d G^2_{jj}(t, x, u, v) + \int_{\mathbb{R}^d} \|y\|^2 \nu^2(t, x, u, v, dy);$$

$$b^2(t, x, u, v) \triangleq f^2(t, x, u, v) + \int_{\mathbb{R}^d \setminus B_1} y\nu^2(t, x, u, v, dy).$$

Under this condition and some regularity assumption we prove the following.

Theorem 3.14 [3, Theorem 3.3]**.** *Let continuous functions $c_1, c_2 : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ be such that*

- $c_i(T, x) = \gamma_i(x)$;
- (c_1, c_2) *satisfies Condition* $(\mathcal{C})$.

Then, for any $(t_0, x_0) \in [0,T]\times\mathbb{R}^d$, and $\varepsilon > (RC+T)\delta$, there exists a profile of public-signal correlates strategies $\mathfrak{w}^*$ providing the ε-equilibrium at (t_0, x_0). Moreover, if X^* and P^* are generated by $\mathfrak{w}^*$ and (t_0, x_0), $\mathbb{E}^*$ denotes the expectation according to P^*, then

$$|\mathbb{E}^*\gamma_i(X^*(T)) - c_i(t_0, x_0)| \leq \varepsilon.$$

In Theorem 3.14 C is a constant determined by the dynamics of the original game.

The natural question here is the limiting behaviour when δ tends to zero. It was proved [3, Theorem 5.2]) that the limiting values of the expected outcomes corresponding to the public-signal correlated equilibria according to Theorem 3.1 of [3] lies in the convex hull of the Nash values in the class of deterministic punishment strategies.

3.4 Equilibria based on system of Bellman equations

3.4.1 Case of smooth solutions

The main result here is the following.

Theorem 3.15 [3, Theorem 4.1]. *Let functions $c_1, c_2 : [0,T]\times\mathbb{R}^d \to \mathbb{R}$, $u^0 : [0,T]\times\mathbb{R}^d \to U$, $v^0 : [0,T]\times\mathbb{R}^d \to V$ be of the class C^2 and satisfy the following conditions:*

$$\begin{aligned}&\frac{\partial c_i}{\partial t} + L_t^2[u^0(t,x), v^0(t,x)]c_i(t,x) + h_i(t,x,u^0(t,x),v^0(t,x)) = 0,\\ &c_i(T,x) = \gamma_i(x),\end{aligned} \tag{32}$$

where $u^0(t,x)$ and $v^0(t,x)$ provide the Nash equilibrium for the one-shot game with the payoff functions

$$L_t^2[u,v]c_i(t,x) + h_i(t,x,u,v), \quad i = 1,2.$$

Assume, additionally, that given $s, r \in [0,T]$, $s < r$, there exist solutions of the martingale problems on $[s,r]$ for the generators $L_t^2[u^0(t,\cdot), v^0(t,\cdot)]$, $L_t^2[u, v^0(t,\cdot)]$, $L_t^2[u^0(t,\cdot), v]$, $u \in U$, $v \in V$; moreover, one can find a common filtered measurable space with a filtration suitable for all mentioned problems.

Then, the pair (c_1, c_2) satisfies Condition $(\mathcal{C})$. In particular, for any $(t_0, x_0) \in [0,T]\times\mathbb{R}^d$, $\varepsilon > (RC+T)\delta$, there exists the public-signal correlated ε-equilibrium at (t_0, x_0).

3.4.2 Approximation by solutions of second-order PDEs

The assumption that the system of Hamilton-Jacobi equations admits a C^2-solution in many case is too restrictive. Let us weaken it for the case of model determined by the stochastic differential equation of the following form:

$$dX(t) = f^2(t, X(t), u(t), v(t))dt + \sigma(t, X(t))dW(t). \tag{33}$$

This system corresponds to the generator

$$L_t^2[u,v]\phi(x) = \langle f^2(t,x,u,v), \nabla\phi(x)\rangle + \frac{1}{2}\langle G^2(t,x)\nabla, \nabla\phi(x)\rangle, \tag{34}$$

where $G^2(t,x) = \sigma(t,x)\sigma^T(t,x)$. We assume that $W(t)$ is a d-dimensional Wiener process, σ is a nondegenerate and bounded $d\times d$-matrix. The crucial assumption is the following: there exist measurable functions $u^N(t,x,p_1,p_2)$, $v^N(t,x,p_1,p_2)$ taking values in U and V respectively such that, for any $t \in [0,T]$, $x, p_1, p_2 \in \mathbb{R}^d$, $u \in U$, $v \in V$,

$$\mathcal{H}_1(t,x,p_1,u^N(t,x,p_1,p_2),v^N(t,x,p_1,p_2)) \geq \mathcal{H}_1(t,x,p_1,u,v^N(t,x,p_1,p_2)),$$

$$\mathcal{H}_2(t,x,p_2,u^N(t,x,p_1,p_2),v^N(t,x,p_1,p_2)) \geq \mathcal{H}_2(t,x,p_2,u^N(t,x,p_1,p_2),v).$$

Here I stands for the identity matrix, whereas

$$\mathcal{H}_i(t,x,p,u,v) \triangleq \langle p, f^2(t,x,u,v)\rangle + g_i(t,x,u,v).$$

Theorem 3.16 [3, Theorem 4.5]. *Assume that (c_1,c_2) is a strong solution of*

$$\frac{\partial c_i}{\partial t} + \mathcal{H}_i(t,x,\nabla c_i, u^N(t,x,\nabla c_1,\nabla c_2), v^N(t,x,\nabla c_1,\nabla c_2)) \\ + \langle G(t,x)\nabla,\nabla\rangle c_i(t,x) = 0, \quad c_i(T,x) = \gamma_i(x).$$

Then (c_1,c_2) satisfies condition $(\mathcal{C})$ for the generator L^2 given by (34).

3.4.3 Approximation by solution of differential inclusions

Let us rewrite system (8) in the coordinate-wise form

$$\frac{d}{dt}x_j(t) = f_{1,j}(t,x_1(t),\ldots,x_d(t),u) + f_{2,j}(t,x_1(t),\ldots,x_d(t),v), \quad j = 1,\ldots,d.$$

Here $x_j(t)$ stands for the j-th coordinate of the vector $x(t)$. Let h be a positive number. As above e^j stands for the j-th coordinate vector.

Put

$$Q^1_{x,y}(t,u) \triangleq \begin{cases} \frac{1}{h}|f_{1,j}(t,x,u)|, & y = x + h\,\mathrm{sgn}(f_{1,j}(t,x,u))\cdot e^j, \\ -\frac{1}{h}\sum_{j=1}^d |f_{1,j}(t,x,u)|, & y = x, \\ 0, & \text{otherwise.} \end{cases}$$

Analogously, set

$$Q^2_{x,y}(t,v) \triangleq \begin{cases} \frac{1}{h}|f_{2,j}(t,x,v)|, & y = x + h\,\mathrm{sgn}(f_{2,j}(t,x,v))\cdot e^j, \\ -\frac{1}{h}\sum_{j=1}^d |f_{2,j}(t,x,v)|, & y = x, \\ 0, & \text{otherwise.} \end{cases}$$

We consider the Markov game with the state space equal to $h\mathbb{Z}^d$ and the Kolmogorov matrix $Q(t,u,v) = Q^1(t,u) + Q^2(t,v)$, where the first player's outcome is

$$\mathbb{E}\left[\gamma_1(X(T)) + \int_{t_0}^T g_1(t,X(t),u(t))dt\right],$$

while the second player maximizes

$$\mathbb{E}\left[\gamma_2(X(T)) + \int_{t_0}^T g_2(t,X(t),v(t))dt\right].$$

The Bellman equation for this controlled Markov chain is as follows. Given $t \in [0,T]$, $x \in h\mathbb{Z}^d$, $\zeta : h\mathbb{Z}^d \to \mathbb{R}$, $\mu \in \mathcal{P}(U)$, $\nu \in \mathcal{P}(V)$ set

$$\widehat{H}_1(t,x,\zeta,\mu) \triangleq \int_U \left[\sum_{y \in h\mathbb{Z}^d} Q^1_{x,y}(t,x,u)\zeta(y) + g_1(t,x,u) \right] \mu(du),$$

$$\widehat{H}_2(t,x,\zeta,\nu) \triangleq \int_V \left[\sum_{y \in h\mathbb{Z}^d} Q^2_{x,y}(t,x,v)\zeta(y) + g_2(t,x,v) \right] \nu(dv).$$

The functions $\widehat{H}_1$, $\widehat{H}_2$ play the role of pre-Hamiltonians. Further, put

$$\mathcal{O}_1(t,x,\zeta) \triangleq \underset{\mu \in \mathcal{P}(U)}{\mathrm{Argmax}} \widehat{H}_1(t,x,\zeta,\mu), \quad \mathcal{O}_2(t,x,\zeta) \triangleq \underset{\nu \in \mathcal{P}(V)}{\mathrm{Argmax}} \widehat{H}_2(t,x,\zeta,\nu).$$

Finally, if ζ_1, ζ_2 are real valued functions defined on $h\mathbb{Z}^d$, denote

$$\begin{aligned} \mathcal{H}_1(t,x,\zeta_1,\zeta_2) &\triangleq \max_{\mu \in \mathcal{P}(U)} \widehat{H}_1(t,x,\zeta_1,\mu) \\ &\quad + \left\{ \int_V \sum_{y \in h\mathbb{Z}^d} Q^2_{x,y}(t,x,v)\zeta_1(y)\nu(du) : \nu \in \mathcal{O}_2(t,x,\zeta_2) \right\}, \\ \mathcal{H}_2(t,x,\zeta_1,\zeta_2) &\triangleq \max_{\nu \in \mathcal{P}(V)} \widehat{H}_2(t,x,\zeta_2,\nu) \\ &\quad + \left\{ \int_U \sum_{y \in h\mathbb{Z}^d} Q^1_{x,y}(t,x,u)\zeta_2(y)\mu(du) : \mu \in \mathcal{O}_1(t,x,\zeta_1) \right\}. \end{aligned}$$

The multifunctions $\mathcal{H}_1$ and $\mathcal{H}_2$ are analogs of the Hamiltonians.

The following system of differential inclusions is the natural analog of the system of the Bellman equations

$$\frac{d}{dt}\zeta_i(t,x) \in -\mathcal{H}_i(t,x,\zeta_1,\zeta_2), \quad \zeta_i(T,x) = \gamma_i(x), \quad i = 1,2. \tag{35}$$

Theorem 3.17 [4, Theorem 2]**.** *Let $(\zeta_1(\cdot,\cdot), \zeta_2(\cdot,\cdot))$ solves system (35). Then it satisfies condition $(\mathcal{C})$. In particular, if $|g_1(t,x,u)|, |g_2(t,x,v)| \leq \sqrt{h}$, then one can construct an approximate public-signal correlated ε-equilibrium for every $\varepsilon > (RC+T)\sqrt{h}$.*

3.5 Viability for the mean field type control systems

The second part of the thesis is concerned with the study of mean field control systems.

Let us consider the first-order mean field type control system with the dynamics of each agent given by

$$\frac{d}{dt}x(t) = f(t,x(t),m(t),u(t)). \tag{36}$$

Here $t \in [0,T]$ is a time, $x(t)$ stands for a state of a representative agent (particle), $m(t)$ is a probability describing distribution of agents, $u(t)$ is a instantaneous control of a representative agent. For simplicity we assume that the phase space for each agent is the d-dimensional torus $\mathbb{T}^d \triangleq \mathbb{R}^d/\mathbb{Z}^d$. Integrating formally (36), we arrive at the following equation for the flow of probabilities $m(\cdot)$

$$\frac{d}{dt}m(t) = \langle f(t,\cdot,m(t),u(t)), \nabla\rangle m(t). \tag{37}$$

Certainly one may use general Definition 1.2 of the flow of probabilities for the generator $L_t[m,u]\phi(x) = \langle f(t,x,m,u), \nabla\phi(x)\rangle$. However, it is more convenient to specify the probability space as well as the class of controls.

To this end we need some additional designations. If X and Y are Polish spaces, m is a measure on X, then denote by $\Lambda(X,m,Y)$ the set of measures on $X \times Y$ with the first marginal equal to m. Using the disintegration theorem, we associate $\xi \in \Lambda(X,m,Y)$ with the measure valued function $X \ni x \mapsto \xi(\cdot|x) \in \mathcal{P}(Y)$ such that, for any $\phi \in C_b(X \times Y)$,

$$\int_{X\times Y} \phi(x,y)\xi(d(x,y)) = \int_X \int_Y \phi(x,y)\xi(dy|x)m(dx).$$

In the following, denote

$$\mathcal{U} \triangleq \Lambda([0,T],\lambda,U),$$

where λ stands for the Lebesgue measure. Given initial position $(s,y) \in [0,T] \times \mathbb{T}^d$, $\xi \in \mathcal{U}$ and a flow of probabilities $m(\cdot)$, we say that $x(\cdot)$ is generated by the initial position and control ξ if $x(\cdot)$ solves the initial value problem

$$\frac{d}{dt}x(t) = \int_U f(t,x(t),m(t),u)\xi(du|t), \quad x(s) = y.$$

We denote the operator assigning to y and ξ the corresponding motion by $\mathrm{traj}^s_{m(\cdot)}$.

Furthermore, if $t \in [0,T]$, then denote by e_t the evaluation operator form $C([0,T],\mathbb{T}^d)$ to $\mathbb{T}^d$ acting by the rule: for $x(\cdot) \in C([0,T],\mathbb{T}^d)$, we put

$$e_t(x(\cdot)) = x(t).$$

Now, given $m \in \mathcal{P}^p(\mathbb{T}^d)$, set

$$\mathcal{A}[m] \triangleq \Lambda(\mathbb{T}^d, m, \mathcal{U}).$$

The set $\mathcal{A}[m]$ is the set of distributions of pairs (y,ξ) with the marginal distribution on $\mathbb{T}^d$ equal to m. Below, if $\alpha \in \mathcal{A}[m]$, we assume that agents placed at the initial time in the state x choose their control according to the distribution $\alpha(d\xi|x)$.

Definition 3.18. Assume that $s \in [0,T]$ is an initial time, m_* is an initial distribution of agents, $\alpha \in \mathcal{A}(m_*)$. We say that the function $[0,t] \ni t \mapsto m(t) \in \mathcal{P}^p(\mathbb{T}^d)$ is a flow of probabilities generated by α, if there exists $\chi \in \mathcal{P}^p(C([0,T],\mathbb{T}^d))$ such that

1. $\chi = \mathrm{traj}^s_{m(\cdot)\#}\alpha$;
2. $m(t) = e_{t\#}\chi$;
3. $m(s) = m_*$.

Equivalently, we can use the approach based on differential inclusion. If we put

$$F(t,x,m) \triangleq \mathrm{co}\{f(t,x,m,u) : u \in U\},$$

control system (36) can be written as follows:

$$\frac{d}{dt}x(t) \in F(t,x(t),m(t)), \tag{38}$$

whereas equation (37) takes the form of the mean field type differential inclusion (MFDI)

$$\frac{d}{dt}m(t) \in \langle F(t,\cdot,m(t)), \nabla\rangle m(t). \tag{39}$$

As equation (37) above this inclusion is formal. We say that a flow of probabilities $[0,T] \ni t \mapsto m(t) \in \mathcal{P}^p(\mathbb{T}^d)$ is a solution to mean field type differential inclusion (39) if there exists $\chi \in \mathcal{P}^p(C([0,T];\mathbb{R}^d))$ such that

1. χ-a.e. $x(\cdot)$ satisfies (38);
2. $m(t) = e_{t\#}\chi$.

Notice that every flow of probabilities generated by s and α is a solution of (39). Conversely, given a flow of probabilities $m(\cdot)$ that solves (39), $s \in [0,T]$ and $m_* = m(s)$, one can find $\alpha \in \mathcal{A}[m_*]$ generating $m(\cdot)$.

Now let consider the viability property. It plays the crucial role in the analysis of first-order mean field games presented below. We restrict our attention to the case $p = 1$. Additionally, we assume time-homogeneous dynamics i.e. f does not depend on t.

Definition 3.19 [5, Definition 2]**.** We say that $K \subset \mathcal{P}^1(\mathbb{T}^d)$ is viable under MFDI (39) if, for any $m_0 \in K$, there exist $T > 0$ and a solution to MFDI (39) on $[0,T]$ $m(\cdot)$ such that $m(0) = m_0$, and $m(t) \in K$ for all $t \in [0,T]$.

To characterize the viability property let us introduce the notion of tangent distributions.

Definition 3.20 [5, Definition 2]**.** Let $a > 0$. We say that $\beta \in \mathcal{P}(\mathbb{T}^d \times \mathbb{R}^d)$ with margin on $\mathbb{T}^d$ equal to m is a tangent probability to K at $m \in \mathcal{P}^1(\mathbb{T}^d)$ with the radius a if there exist sequences $\{\tau_n\}_{n=1}^\infty \subset (0,+\infty)$, $\{\beta_n\}_{n=1}^\infty \subset \mathcal{P}(\mathbb{T}^d \times \mathbb{R}^d)$ such that $\mathrm{p}^1{}_\#\beta = m$, $\mathrm{supp}(\beta_n) \subset \mathbb{T}^d \times B_a$ and

$$\frac{1}{\tau_n}\mathrm{dist}(\Theta^{\tau_n}{}_\#\beta_n, K) \to 0, \quad W_1(\beta_n,\beta) \to 0, \quad \tau_n \to 0 \text{ as } n \to \infty.$$

Let us denote the set of tangent probabilities to K of the radius a by $\mathcal{T}_K^a(m)$.

Further, denote by $\mathcal{F}(m)$ the set of probabilities $\beta \in \mathcal{L}(m)$ such that

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} \mathrm{dist}(v, F(x,m))\beta(d(x,v)) = 0.$$

Theorem 3.21 [5, Theorem 1]**.** *A closed set $K \subset \mathcal{P}^1(\mathbb{T}^d)$ is viable under MFDI (39) if and only if, there exists a constant $a > 0$ such that, for any $m \in K$,*

$$\mathcal{T}_K^a(m) \cap \mathcal{F}(m) \neq \varnothing. \tag{40}$$

3.6 First-order mean field games

The following results are obtained for the case of the phase space of each player equal to $\mathbb{R}^d$. We consider the mean field game where each player tries to maximize the outcome given by

$$\gamma(x(T), m(T)) + \int_{t_0}^{T} g(t, x(T), m(t), u(t))dt \tag{41}$$

subject to conditions

$$\begin{gathered} \frac{d}{dt}x(t) = f(t, x(t), m(t), u(t)), \\ t \in [t_0, T], \quad x(t) \in \mathbb{R}^d, \quad m(t) \in \mathcal{P}^1(\mathbb{T}^d), \quad u(t) \in U, \end{gathered} \tag{42}$$

and $m(t)$ is a distribution of players at time t. It is more convenient to introduce the additional variable and consider the terminal payoff. In this case we introduce $z(\cdot)$ by the rule:

$$\frac{d}{dt}z(t) = g(t, x(t), m(t), u(t)), \quad z(t_0) = 0. \tag{43}$$

Thus, one can assume that each player tries to maximize the terminal payoff

$$\gamma(x(T), m(T)) + z(T). \tag{44}$$

Note that $(x(t), z(t))$ lies in the extended phase space $\mathbb{R}^d \times \mathbb{R}$.

We adapt the general definition of the solution to the case of first-order mean field games. With some abuse of notation we denote the evaluation operator from $C([0, T], \mathbb{R}^d \times \mathbb{R})$ to $\mathbb{R}^d$ by e_t: if $w(\cdot) = (x(\cdot), z(\cdot))$, then

$$e_t(w(\cdot)) = x(t).$$

The evaluation operator taking values in $\mathbb{R}^d \times \mathbb{R}$ is denoted by $\hat{e}_t$:

$$\hat{e}_t(w(\cdot)) = w(t).$$

Recall that the relaxation of the system with dynamics (42), (43) satisfy the differential inclusion

$$(\dot{x}(t), \dot{z}(t)) \in F(t, x(t), m(t)), \tag{45}$$

where

$$F(t, x, m) \triangleq \text{co}\{(f(t, x, m, u), g(t, x, m, u)) : u \in U\}. \tag{46}$$

Given a flow of probabilities $m(\cdot)$, $s, r \in [0, T]$, $s < r$, $y \in \mathbb{R}^d$, we denote the set of solutions of (45) on $[s, r]$ with the initial condition $x(s) = y$ by $\text{Sol}(r, s, y, m(\cdot))$. Furthermore,

$$\text{SOL}(r, s, m(\cdot)) \triangleq \bigcup_{y \in \mathbb{T}^d} \text{Sol}(r, s, y, m(\cdot)).$$

Definition 3.22. We say that a pair $(\varphi, m(\cdot))$, where $\varphi : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}$ is a continuous function and $[t_0, T] \ni t \mapsto m(t) \in \mathcal{P}^p(\mathbb{T}^d)$, is a (minimax) solution to mean field game (41), (42) if there exists a probability $\chi \in \mathcal{P}^p(C([t_0, T]; \mathbb{R}^d \times \mathbb{R}))$ such that

1. $m(t) = e_{t\#}\chi$;

2. $\varphi(s, y)$ is a value of the optimization problem

$$\text{maximize } [\gamma(x(T), m(T)) + z(T) - z(s)]$$
$$\text{subject to } (x(\cdot), z(\cdot)) \in \mathrm{Sol}(T, s, y, m(\cdot));$$

3. $\mathrm{supp}(\chi) \subset \mathrm{SOL}(T, t_0, m(\cdot))$;

4. for every $s, r \in [t_*, T]$, $s < r$, and each $(x(\cdot), z(\cdot)) \in \mathrm{supp}(\chi)$,

$$\varphi(s, x(s)) + z(s) = \varphi(r, x(r)) + z(r).$$

First, we have the existence theorem.

Theorem 3.23 [6, Theorem 1]**.** *Given $m_0 \in \mathcal{P}^1(\mathbb{T}^d)$, there exists a solution of mean field game (41), (42) such that $m(t_0) = m_0$.*

Remark 3.24. In [6] this theorem was proved under additional assumptions that $p = 1$ and m_0 is concentrated on some compact. However, it can be extended to the general case.

The presented concept of solution of the firts-order mean field games is stable with respect to stochastic perturbations. To describe the class of admissible perturbation, let us consider the sequence of stochastic mean field games with the dynamics of each player given by the Lévy-Khinchine generator:

$$\begin{aligned} L^n_t[m, u]\phi(x) = {} & \frac{1}{2}\langle G^n(t, x, m, u)\nabla, \nabla\rangle\phi(x) + \langle f^n(t, x, m, u), \nabla\rangle\phi(x) \\ & + \int_{\mathbb{R}^d} [\phi(x + y) - \varphi(x) - \langle y, \nabla\phi(x)\rangle \mathbf{1}_{B_1}(y)]\nu^n(t, x, m, u, dy). \end{aligned} \tag{47}$$

It is assumes that the payoff functional is determined by

$$\mathbb{E}\left[\gamma(X(T), m(T)) + \int_{t_0}^{T} g(t, X(t), m(t), u(t))dt\right],$$

where X stands for the stochastic process generated by L^n. Let us introduce the following notation:

$$\Sigma^n(t, x, m, u) \triangleq \sum_{i=1}^{d} G^n_{ii}(t, x, m, u) + \int_{\mathbb{R}^d} \|y\|^2 \nu^n(t, x, m, u, dy),$$

$$b^n(t, x, m, u) \triangleq f^n(t, x, m, u) + \int_{\mathbb{R}^d \setminus B_1} y\nu^n(t, x, m, u, dy).$$

We assume that the stochastic mean field games converge to the deterministic one in the following sense:

$$\sup_{t\in[0,T], x\in\mathbb{R}^d, m\in\mathcal{P}^2(\mathbb{R}^d)} \frac{\Sigma^n(t, x, m, u)}{1 + \|x\|^2 + \varsigma^2(m)} \to 0 \text{ as } n \to \infty;$$

$$\sup_{t\in[0,T], x\in\mathbb{R}^d, m\in\mathcal{P}^2(\mathbb{R}^d)} \frac{\|b^n(t, x, m, u) - f(t, x, m, u)\|}{(1 + \|x\| + \varsigma(m))} \to 0 \text{ as } n \to \infty.$$

Here

$$\varsigma(m) \triangleq \left[\int_{\mathbb{R}^d} \|x\|^2 m(dx)\right]^{1/2}.$$

Under certain regularity condition we have the following.

Theorem 3.25 [7, Theorem 1]**.** *Assume that for each n the pair $(\varphi^n, m^n(\cdot))$ solves the mean field game with the dynamics (47) and payoff functional (41).*

Then there exist a pair $(\varphi^*, m^*(\cdot))$ that is a solution to mean field game (41), (42) and a sequence $\{n_l\}_{l=1}^\infty$ such that

1. $\sup_{t\in[0,T]} W_2(m^{n_l}(t), \mu^*(t)) \to 0$ *as* $l \to \infty$;
2. $$\lim_{l\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\varphi^{n_l}(t,x) - \varphi^*(t,x)| = 0 \quad (48)$$

Given a solution of the mean field game for $p = 1$, one can construct a Nash equilibrium for the finite player game where each player controls a cloud of agents playing according to dynamics (42). It is assumed that each player tries to maximize the common outcome of his/her agents. If the agents affected by the player i start the position $x^i_{N,0}$, $\mathbf{x} = (x^1_{N,0}, \ldots, x^N_{N,0})$, then we denote

$$\delta^N_{\mathbf{x}} \triangleq \frac{1}{N}(\delta_{x^1_{N,0}} + \ldots + \delta_{x^N_{N,0}}).$$

Additionally, given a solution of the mean field game one can construct (by certain rule) the profile of strategies $\hat{\varrho}_N = (\hat{\varrho}^1_N, \ldots, \hat{\varrho}^N_N)$ where $\hat{\varrho}_N$ is a probability on $\mathcal{U}$.

Theorem 3.26 [6, Theorem 3]**.** *If*

$$W_1(m_0, \delta^N_{\mathbf{x}}) \to 0, \quad N \max_{i=\overline{1,N}} \int_{\mathbb{R}^d} \|x - x^i_{N,0}\| m^i_N(dx) \to 0, \text{ for } N \to \infty,$$

then, for any $\varepsilon > 0$, there exists a number N_0 such that, for all $N > N_0$, the profile of strategies $\hat{\varrho}_N$ is a ε-Nash equilibrium.

Here $m^1_N, \ldots, m^N_N$ are the measures such that $m_0 = m^1_N + \ldots + m^N_N$, $m^i_N(\mathbb{R}^d) = 1/N$ and

$$W_1(m_0, \delta^N_{\mathbf{x}}) = \sum_{i=1}^N \int_{\mathbb{R}^d} \|x - x^i_{N,0}\| m^i_N(dx).$$

3.7 Viability analysis of the mean field games

The last part of the thesis' results is concerned with the dependence of the solutions of the mean field game on the initial distribution of players. This dependence plays the crucial role in the master equation which is used to establish the convergence of feedback equilibria to the solution of the mean field game [CDLL19]. Now we assume that the phase space for each player is $\mathbb{T}^d \triangleq \mathbb{R}^d/\mathbb{Z}^d$. Moreover, we endow the space of probabilities over $\mathbb{T}^d$ by the 1-Kantorovich metric.

The main object of this part is the value multifunction defined as follows.

Definition 3.27 [8, Definition 3.1]**.** We say that an upper semicontinuous function multifunction $\mathcal{V} : [0,T] \times \mathcal{P}^1(\mathbb{T}^d) \rightrightarrows C(\mathbb{T}^d)$ is a value multifunction of mean field game (41), (42) if, for every $t_0 \in [0,T]$, $m_0 \in \mathcal{P}^1(\mathbb{T}^d)$, and $\phi \in \mathcal{V}(t_0, m_0)$, there exists a solution to mean field game (41), (42) $(\varphi, m(\cdot))$ such that

$$\varphi(t_0, \cdot) = \phi(\cdot), \quad m(t_0) = m_0. \quad (49)$$

This multifunction can be characterized via the viability property. To this end, let us introduce the mean field dynamics on $C(\mathbb{T}^d)\times\mathcal{P}^1(\mathbb{T}^d)$. It relies on several auxiliary definitions. First, for $s,r\in[0,T]$, $s\leq r$, $m(\cdot)\in C([s,r],\mathcal{P}^1(\mathbb{T}^d))$, define the operator $B^{s,r}_{m(\cdot)}:C(\mathbb{T}^d)\to C(\mathbb{T}^d)$ by the rule:

$$(B^{s,r}_{m(\cdot)}\psi)(y)\triangleq\sup\{\psi(x(r))+z(r)-z(s):(x(\cdot),z(\cdot))\in\mathrm{Sol}(r,s,y,m(\cdot))\}.$$

Additionally, if $\phi\in C(\mathbb{T}^d)$, $\nu\in\mathcal{P}^1(\mathbb{T}^d\times\mathbb{R})$, then denote by $[\phi,\nu]$ the averaging of the function $\phi(x)+z$ with respect to ν:

$$[\phi,\nu]\triangleq\int_{\mathbb{T}^d\times\mathbb{R}}(\phi(x)+z)\nu(d(x,z)).$$

Definition 3.28 [8, Definition 3.5]**.** For each $s,r\in[0,T]$, $s\leq r$, define the multifunction $\Psi^{r,s}:\mathcal{P}^1(\mathbb{T}^d)\times C(\mathbb{T}^d)\rightrightarrows\mathcal{P}^1(\mathbb{T}^d)\times C(\mathbb{T}^d)$ by the rule: $(\mu,\psi)\in\Psi^{r,s}(m,\phi)$ if and only if there exists a probability $\chi\in\mathcal{P}^1(C([s,r];\mathbb{T}^d\times\mathbb{R}))$, satisfying the following properties for $\nu(t)\triangleq\hat{e}_{t\#}\chi$, and $m(t)\triangleq e_{t\#}\chi$:

(Ψ1) $m(s)=m$, $m(r)=\mu$;

(Ψ2) $\phi=B^{s,r}_{m(\cdot)}\psi$;

(Ψ3) $[\psi,\nu(r)]\geq[\phi,\nu(s)]$.

The definition of the viability property is given in the classical way.

Definition 3.29 [8, Definition 3.7]**.** We say that an upper semicontinuous multifunction $\mathcal{V}:[0,T]\times\mathcal{P}^1(\mathbb{T}^d)\rightrightarrows C(\mathbb{T}^d)$ is viable with respect to the mean field game dynamics if, for any $s,r\in[0,T]$, $s\leq r$, $m\in\mathcal{P}^1(\mathbb{T}^d)$, $\phi\in\mathcal{V}(s,m)$, there exist $\mu\in\mathbb{T}^d$ and $\psi\in C(\mathbb{T}^d)$ such that

- $(\mu,\psi)\in\Psi^{r,s}(m,\phi)$;
- $\psi\in\mathcal{V}(r,\mu)$.

The following statement gives the link between the viability property and the mean field games.

Theorem 3.30 [8, Theorem 3.10]**.** *Assume that an upper semicontinuous multifunction $\mathcal{V}:[0,T]\times\mathbb{R}^d\rightrightarrows C(\mathbb{T}^d)$ is viable with respect to the mean field game dynamics and $\mathcal{V}(T,m)=\{\gamma(\cdot,m)\}$. Then $\mathcal{V}$ is a value multifunction.*

Further, we obtain the infinitesimal form of the viability condition for the mean field game dynamics. This relies on the set of tangent distribution to a multivalued function $\mathcal{V}$. First, assume that $m\in\mathcal{P}^1(\mathbb{T}^d)$, c is a positive number. With some abuse of notation, denote by $\mathcal{L}^c(m)$ the set of probabilities $\beta\in\mathcal{P}^1(\mathbb{T}^d\times\mathbb{R}^{d+1})$ concentrated on $\mathbb{T}^d\times B_c\times[-c,c]$ with the marginal distribution on $\mathbb{T}^d$ equal to m. If, additionally, $s,r\in[0,T]$, $s\leq r$, then denote by $A^{s,r}_m$ the operator on $C(\mathbb{T}^d)$ acting by the rule

$$(A^{s,r}_m\phi)(x)\triangleq\sup\{\phi(x+(r-s)a)+(r-s)b:(a,b)\in F(s,x,m)\}. \quad (50)$$

Here F is defined by (46). For $\tau\geq 0$, let the shift operator $\Theta^\tau:\mathbb{T}^d\times\mathbb{R}^{d+1}\to\mathbb{T}^d\times\mathbb{R}$ be given by

$$\Theta^\tau(x,a,b)\triangleq(x+\tau a,\tau b). \quad (51)$$

Definition 3.31 [8, Definition 4.1]**.** A probability $\beta \in \mathcal{L}^c(m)$ belongs to $\mathcal{D}^c_F\mathcal{V}(t,m,\phi)$ if there exist sequences $\{\tau_n\}_{n=1}^\infty \subset (0,+\infty)$, $\{\beta_n\}_{n=1}^\infty \subset \mathcal{L}^c(m)$ and $\{\phi_n\}_{n=1}^\infty \subset C(\mathbb{T}^d)$ satisfying the following properties for $\nu_n \triangleq \Theta^{\tau_n}{}_\#\beta_n$ and $m_n \triangleq \mathrm{p}_\#\nu_n$:

1. $\tau_n, W_1(\beta,\beta_n) \to 0$ as $n \to \infty$;
2. $\phi_n \in \mathcal{V}(t+\tau_n, m_n)$;
3. $$\lim_{n\to\infty} \frac{\|A_m^{t,t+\tau_n}\phi_n - \phi\|}{\tau_n} = 0;$$
4. $$\lim_{n\to\infty} \frac{[\phi_n,\nu_n] - [\phi,\widehat{m}]}{\tau_n} \geq 0;$$
5. $$\int_{\mathbb{T}^d\times\mathbb{R}^{d+1}} \mathrm{dist}(v; F(t,x,m))\beta(d(x,v)) = 0.$$

Here $\widehat{m}$ is a probability on $\mathbb{T}^d \times \mathbb{R}$ defined by the rule: for $\phi \in C(\mathbb{T}^d \times \mathbb{R})$

$$\int_{\mathbb{T}^d\times\mathbb{R}} \phi(x,z)\widehat{m}(d(x,x)) = \int_{\mathbb{T}^d} \phi(x,0)m(dx).$$

For $M, C > 0$, let $\mathrm{BL}_{M,C}$ denote the set of functions $\phi \in \mathrm{Lip}_C(\mathbb{T}^d)$ such that $\|\phi\| \leq M$.

Theorem 3.32 [8, Theorem 4.2]**.** *Assume that an upper semicontinuous multifunction $\mathcal{V} : [0,T] \times \mathcal{P}^1(\mathbb{T}^d) \rightrightarrows C(\mathbb{T}^d)$ has nonempty values and there exist constants M and C such that, for any $t \in [0,T]$, $m \in \mathcal{P}^1(\mathbb{T}^d)$,*

$$\mathcal{V}(t,m) \subset \mathrm{BL}_{M,C}(\mathbb{T}^d).$$

Then, $\mathcal{V}$ is viable with respect to the mean field game dynamics if and only if there exists a constant $c > 0$ such that, for any $t \in [0,T]$, $m \in \mathcal{P}^1(\mathbb{T}^d)$, $\phi \in \mathcal{V}(t,m)$,

$$\mathcal{D}^c_F\mathcal{V}(t,m,\phi) \neq \varnothing.$$

This theorem yield the sufficient condition in the infinitesimal form for a given multifunction to be a value multifunction.

Corollary 3.33 [8, Corollary 4.3]**.** *Let the upper semicontinuous multifunction $\mathcal{V}$: $[0,T] \times \mathcal{P}^1(\mathbb{T}^d) \rightrightarrows C(\mathbb{T}^d)$ have nonempty values. Assume that, for any $t \in [0,T]$, $m \in \mathcal{P}^1(\mathbb{T}^d)$, $\phi \in \mathcal{V}(t,m)$,*

- *$\mathcal{V}(t,m) \subset \mathrm{BL}_{M,C}(\mathbb{T}^d)$ where the constants M and C do not dependent on t and m;*
- *$\mathcal{V}(T,m) = \{\gamma(\cdot,m)\}$;*
- *$\mathcal{D}^c_F\mathcal{V}(t,m,\phi) \neq \varnothing$, where the constant c does not depend on t, m and ϕ.*

Then $\mathcal{V}$ is a value multifunction of mean field game (41), (42).

References

[AD01] N.U. Ahmed and X. Ding. Controlled McKean-Vlasov equation. *Commun. Appl. Anal.*, 5:183–206, 2001.

[AGS05] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows: in metric spaces and in the space of probability measures.* Lectures in Mathematics. ETH Zurich. Birkhäuser, Basel, 2005.

[Aub09] J.-P. Aubin. *Viability theory.* Birkhäuser, Boston, 2009.

[Ave15] Yu. Averboukh. Universal Nash equilibrium strategies for differential games. *J. Dyn. Control Syst.*, 21:329–350, 2015.

[BD96] M. Bardi and I. Capuzzo Dolcetta. *Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations.* Birkhäuser, Basel, 1996.

[BK12] V.I. Bogachev and A. V. Kolesnikov. The Monge-Kantorovich problem: achievements, connections, and perspectives. *Russ. Math. Surv+*, 67(5):785–890, 2012.

[BS71] N. Barabanova and A. Subbotin. On continuous evasion strategies in game problems on the encounter of motions. *J. Appl. Math. Mech.*, 34(5):765–772, 1971.

[BS04a] A. Bressan and W. Shen. Semi-cooperative strategies for differential games. *Internat. J. Game Theory*, 32:561–59, 2004.

[BS04b] A. Bressan and W. Shen. Small BV solutions of hyperbolic noncooperative differential games. *SIAM J. Control Optim.*, 43:194–215, 2004.

[CD18] Rene Carmona and Francois Delarue. *Probabilistic theory of mean field games with applications. Vol. I, II.* Springer, 2018.

[CDLL19] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. *The master equation and the convergence problem in mean field games.* Princeton University Press, Princeton, 2019.

[CLSW98] F. H. Clarke, Y. S. Ledyaev, R. Stern, and P. R. Wolenski. *Nonsmooth analysis and control theory.* Springer, Heidelberg, 1998.

[CP03] P. Cardaliaguet and S. Plaskacz. Existence and uniqueness of a Nash equilibrium feedback for a simple nonzero-sum differential game. *Internat. J. Game Theory*, 32:33–71, 2003.

[Fis17] M. Fischer. On the connection between symmetric N-player games and mean field games. *Ann. Appl. Probab.*, 27:757–810, 2017.

[Fri70] A. Friedman. Existence of value and of saddle points for differential games of pursuit and evasion. *J. Differential Equations*, 7(1):92–110, 1970.

[Fri72] A. Friedman. Stochastic differential games. *J. Differ. Equations*, 11(1):79–108, 1972.

[FS06] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*. Springer, New York, 2006.

[GHL03] X. Guo and O. Hernández-Lerma. Zero-sum games for continuous-time Markov chains with unbounded transitions and average payoff rates. *J. Appl. Probab.*, 40:327–345, 2003.

[GS15] W. Gangbo and A. Święch. Existence of a solution to an equation arising from the theory of mean field games. *J. Differential Equations*, 259:6573–6643, 2015.

[HL95] S. Hamadèène and J. P. Lepeltier. Backward equations, stochastic control and zero-sum stochastic differential games. *Stoch. Stoch. Rep.*, 54(3-4):221–231, 1995.

[HM19] S. Hamadène and P. Mannucci. Regularity of Nash payoffs of Markovian nonzero-sum stochastic differential games. *Stochastics*, 91(5):695–715, 2019.

[HMC05] M. Huang, R.P. Malhamé, and P.E. Caines. Nash equilibria for large population linear stochastic systems with weakly coupled agents. In E. K. Boukas and R.P. Malhamé, editors, *Analysis, control and optimization of complex dynamic systems*, pages 215–252. Springer, 2005.

[KK10] N. N. Krasovskii and A. N. Kotelnikova. An approach-evasion differential game: stochastic guide. *Proc. Steklov Inst. Math.*, 269(1 Supplement):S191–S213, 2010.

[Kle93] A.F. Kleimenov. *Nonzero-sum differential games.* Nauka, Ekaterinburg, 1993. in Russian.

[KLY11] V.N. Kolokoltsov, J.J. Li, and W. Yang. Mean field games and nonlinear Markov processes. Preprint at arXiv:1112.3744v2, 2011.

[KS70] N.N. Krasovskii and A.I. Subbotin. An alternative for the game problem of convergence. *J. Appl. Math. Mech.*, 34:948–965, 1970.

[KS88] N. N. Krasovskii and A. I. Subbotin. *Game-theoretical control problems.* Springer, New York, 1988.

[KTY14] V. N. Kolokoltsov, M. Troeva, and W. Yang. On the rate of convergence for the mean-field approximation of controlled diffusions with large number of players. *Dyn. Games Appl.*, 4:208–230, 2014.

[KY13] V. Kolokoltsov and W. Yang. Existence of solutions to path-dependent kinetic equations and related forward-backward systems. *Open Journal of Optimization*, 2:39–44, 2013.

[Lac17] D. Lacker. Limit theory for controlled McKean-Vlasov dynamics. *SIAM J. Control Optim.*, 55:1641–1672, 2017.

[Lev13] Y. Levy. Continuous-time stochastic games of fixed duration. *Dyn. Games Appl.*, 3:279–312, 2013.

[LL06a] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire (French) [Mean field games. I. The stationary case]. *C. R. Math. Acad. Sci. Paris*, 343:619–625, 2006.

[LL06b] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal (French) [Mean field games. II. Finite horizon and optimal control]. *C. R. Math. Acad. Sci. Paris*, 343:679–684, 2006.

[Man04] P. Mannucci. Nonzero-sum stochastic differential games with discontinuous feedback. *SIAM J. Control Optim.*, 43:1222–1233, 2004.

[Man14] P. Mannucci. Nash p for nonzero-sum stochastic differential games with separate Hamiltonians. *Dyn. Games Appl.*, 4:329–344, 2014.

[SC81] A. I. Subbotin and A. G. Chentsov. *Optimization of guarantee in control problems*. Nauka, Moscow, 1981. in Russian.

[Sub95] A. I. Subbotin. *Generalized solutions of first-order PDEs. The dynamical perspective*. Birkhauser, Boston, 1995.

[YZ99] J. Yong and X. Yu. Zhou. *Stochastic controls. Hamiltonian systems and HJB equations*. Springer, New York, 1999.

[Zac64] L. E. Zachrisson. Markov games. In M. Dresher, L. S. Shapley, and A. W. Tucker, editors, *Advances in game theory*, pages 211–253. Princeton University Press, Princeton, 1964.