

Lattice approximations of the first-order mean field type differential games

Yurii Averboukh^{ab}

Abstract

The theory of first-order mean field type differential games examines the systems of infinitely many identical agents interacting via some external media under assumption that each agent is controlled by two players. We study the approximations of the value function of the first-order mean field type differential game using solutions of model finite-dimensional differential games. The model game appears as a mean field type continuous time Markov game, i.e., the game theoretical problem with the infinitely many agents and dynamics of each agent determined by a controlled finite state nonlinear Markov chain. Given a supersolution (resp. subsolution) of the Hamilton-Jacobi equation for the model game, we construct a suboptimal strategy of the first (resp. second) player and evaluate the approximation accuracy using the modulus of continuity of the reward function and the distance between the original and model games. This gives the approximations of the value function of the mean field type differential game by values of the finite-dimensional differential games. Furthermore, we present the way to build a finite-dimensional differential game that approximates the original game with a given accuracy.

MSC Classification (2020): 49L25, 49J45, 49J53, 49N70.

Keywords: mean field type differential games, approximate solutions, suboptimal strategies, viscosity solutions, extremal shift rule.

1 Introduction

The mean field type differential games are mathematical models of systems consisting of many agents those are influenced by two player. The main feature of the mean field approach is the assumption that the agents interact via some external media, i.e., the dynamics of each agent depends only on time, his/her state, controls produced by the players and the distribution of all agents. In the paper, we study the zero-sum first-order mean field type differential games. This means that the agents' dynamics is given by the ordinary differential equation, when the players' purposes are opposite.

Notice that the theory of mean field type differential games comprises the mean field type control problems. In fact, one can regard the mean field type control problem as

^aKrasovskii Institute of Mathematics and Mechanics e-mail:ayv@imm.uran.ru

^bUral Federal University

the mean field type differential games with the fictions second player. It is worth to mention that the mean field type control theory is a counterpart of the mean field game theory introduced by Lasry and Lions [24], [25] and (independently) by Huang, Malhamé, Caines [18]. However, the mean field game theory implies that each agent is a player and tries to optimize his/her own payoff.

The mean field type control theory started with paper [1]. Nowadays, for the mean field type control system such methods as dynamic programming and analogs of Pontryagin maxim principle are developed [4], [8], [9], [10], [11], [26], [29], [30]. Notice that the first-order mean field type control problems are more challenging due to the possible crowding of the agents. They were examined from the dynamic programming point of view in [12], [13], [19].

First, the mean field type differential games was analyzed in [15]. In that the paper it was assumed that the dynamics of each agent is given by a non-degenerate stochastic differential equation. The main result is the existence of the value function in the class of progressively measurable strategies. Papers [14], [28] examine the mean field type differential games using nonanticipative strategies. There is proved the existence theorem for the value function. The feedback approach to the mean field type differential games is developed in [5]. That paper extends the Krasovskii-Subbotin extremal shift rule to the mean-field type control system. Furthermore, the existence of the value function is obtained using so called programmed iteration method. The infinitesimal characterization of the value function is derived in [6] and [14]. Paper [14] shows that the value function of the mean field type differential game is the viscosity solution of the corresponding Hamilton-Jacobi equation in the space of probability measures. The notion of viscosity solution used there is based on embedding of the set of probability measures into the space of square integrable random variables. The characterization of the value function of the mean field type differential game via minimax inequalities on direction derivatives was obtained in [6].

The aforementioned results are counterparts of the finite-dimensional differential games theory. However, the mean field type control theory raises its own questions. One of these questions is a finite dimensional approximations. Recall that the mean field type control theory implies that the phase space of the whole system is the Wasserstein space that is only metric, when the corresponding Hamilton-Jacobi equation involves the derivatives with respect to probability (various approaches to this problem can be found in [14], [27], [29], [30]).

At the same time, the finite-dimensional differential games are simpler and well-studied [7],[22], [32]. In particular, they are reduced to the finite dimensional Hamilton-Jacobi equations. Thus, it is tempting to approximate a mean field type control system by finite dimensional systems and construct finite-dimensional approximations of the value function. Since the mean field type control system implies the infinitely many agents, the most natural approximation arises if we consider finite agent system assuming that the probability in the right-hand side of the dynamics is determined by the empirical measure. This approach was realized in [17], [23]. There, it is shown that solutions of the control problems with finite number of agents converge to the solution of the limiting mean field control problem when the number of agents tends to infinity. This result,

in particular, shows that the mean field type control theory is an appropriate model for systems with large number of agents. However, the aforementioned papers do not provide the approximation rate.

We develop a quite different approach. We examine a mean field type differential game assuming, for simplicity, that the phase space of each agent is the flat torus and approximate it by the finite-state mean field type Markov game, i.e., the approximative system consists now of infinitely many identical agents with dynamics of each agent given by the controlled continuous-time Markov chain defined on some lattice. It is well known that the dynamics of the distribution of agents in the Markov chain obeys the Kolmogorov equation which, for the mean field type case, is a system of nonlinear ODEs. Thus, the approximative problem turns to be a finite-dimensional differential game. Notice that, if the agents' state space is Euclidean, such approximation can be performed using a lattice contained in a sufficiently large ball.

The main result of the paper (see Theorem 4.1) states that, if the dynamics of each agent according to the Markov chain approximates the original deterministic evolution, then the value function of the finite-state mean field type Markov game provides the approximation of the value function of the original mean field type differential game. Moreover, the approximation accuracy can be estimated by the modulus of continuity of the reward function and distance between the original and approximative systems. The proof is constructive, i.e., given a supersolution (respectively, subsolution) of the Hamilton-Jacobi equation corresponding to the mean field type Markov game, we construct a suboptimal strategy of the first (respectively, second) player in the original mean field type differential game providing the desired approximation accuracy. Furthermore, one can always build at least one approximative system defined on a regular lattice (see Section 5).

The paper is organized as follows. In Section 2, we introduce the general notation used throughout the paper (see Subsection 2.1) and recall the feedback approach to the mean field type differential games first developed in [5] (see Subsection 2.2). Section 3 is concerned with the finite-state mean field type Markov games. Here, we introduce this object and, utilizing the relaxed feedback strategies, reduce such game theoretical problems to a finite-dimensional differential games defined on a simplex. The main result of the paper is formulated in Section 4. Recall that it states that, given a mean field type Markov game that approximates the original one and the supersolution (respectively, subsolution) of corresponding finite-dimensional Hamilton-Jacobi equation, one can find a suboptimal strategy of the first (respectively, second) player in the original game. A construction of the mean-field type Markov game that approximates the original game with given accuracy is presented in Section 5. The rest of the paper is devoted to the proof of the main result. In Section 6, we give the properties of the original and model systems those are utilized in the proof of the main theorem. We introduce the desired suboptimal strategy with memory in Section 7. Its construction relies on the extremal shift rule first proposed by Krasovskii and Subbotin for finite-dimensional differential games [22]. Finally, proof of the main the theorem is given in Section 8.

2 Preliminaries

2.1 General notation

- If X_1, \dots, X_n are the set, i_1, \dots, i_k are distinct indexes, then $p^{i_1, \dots, i_k} : X_1 \times \dots \times X_n \rightarrow X_{i_1} \times \dots \times X_{i_k}$ acts by the rule

$$p^{i_1, \dots, i_k}(x_1, \dots, x_n) \triangleq (x_{i_1}, \dots, x_{i_k}).$$

- Let (Ω', \mathcal{F}') and $(\Omega'', \mathcal{F}'')$ be measurable spaces, m be a measure on \mathcal{F}' , $h : \Omega' \rightarrow \Omega''$ be a measurable function. Denote by $h\#m$ the push-forward measure defined by the rule: for $\Upsilon \in \mathcal{F}''$,

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

- If (X, ρ_X) , (Y, ρ_Y) are metric spaces, then $C(X, Y)$ denotes the set of continuous functions from X to Y . We write $C(X)$ for $C(X, \mathbb{R})$.
- $C_b(X)$ stands for the set of bounded continuous functions from X to \mathbb{R} .
- Denote by $\mathcal{M}^+(X)$ the set of (nonnegative) measures on X . We endow $\mathcal{M}^+(X)$ with the topology of narrow convergence: a sequence $\{m_n\}_{n=1}^\infty \subset \mathcal{M}^+(X)$ narrowly converges to $m \in \mathcal{M}^+(X)$ if, for any $\phi \in C_b(X)$,

$$\int_X \phi(x)m_n(dx) \rightarrow \int_X \phi(x)m(dx) \text{ as } n \rightarrow \infty.$$

- $\mathcal{P}(X)$ is the set of probability measures, i.e.,

$$\mathcal{P}(X) \triangleq \{m \in \mathcal{M}^+(X) : m(X) = 1\}.$$

- We denote the Dirac measure concentrated at $z \in X$ by δ_z , i.e., for a Borel set Υ ,

$$\delta_z(\Upsilon) \triangleq \begin{cases} 1, & z \in \Upsilon, \\ 0, & z \notin \Upsilon. \end{cases}$$

- For $p \geq 1$, let $\mathcal{P}^p(X)$ stand for the set of probability measures with finite p -th moment i.e. $m \in \mathcal{P}(X)$ lies in $\mathcal{P}^p(X)$, if, for some (equivalently, every) $x_0 \in X$,

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

- We consider on $\mathcal{P}^p(X)$ the p -Wasserstein distance defined as follows:

$$W_p(m_1, m_2) \triangleq \left[\inf_{\pi \in \Pi(m_1, m_2)} \int_{X \times X} (\rho_X(x_1, x_2))^p \pi(d(x_1, x_2)) \right]^{1/p}.$$

Here $\Pi(m_1, m_2)$ is the set of probabilities on $X \times X$ such that $p^i \# \pi = m_i$, $i = 1, 2$. Notice that $\mathcal{P}^p(X)$ endowed with the metric W_p is Polish [3]. If, additionally, X is compact, then $\mathcal{P}^p(X)$ is also compact [3]. In the paper we primary consider the quadratic case, i.e., $p = 2$.

- Given $m \in \mathcal{M}^+(X)$, let $\mathcal{M}_{\text{product}}^+(X, m, Y)$ be a set of measures on $X \times Y$ compatible with m , i.e., $\alpha \in \mathcal{M}_{\text{product}}^+(X, m, Y)$ if $\alpha \in \mathcal{M}^+(X \times Y)$ and $p^1 \# \alpha = m$. The set $\mathcal{M}_{\text{product}}^+(X, m, Y)$ inherits the topology of narrow convergence from $\mathcal{M}^+(X \times Y)$. Notice that $\mathcal{M}_{\text{product}}^+(X, m, Y)$ is compact provided that both X and Y are compact.
- The function $\gamma : X \rightarrow \mathcal{P}(Y)$ is called weakly measurable if, for every $\phi \in C_b(X \times Y)$, the function $x \mapsto \int_Y \phi(x, y) \gamma(x, dy)$ is measurable. If m is a measure on X , $\gamma : X \rightarrow \mathcal{P}(Y)$ is weakly measurable, then define the measure $m \star \gamma$ by the rule: for $\varphi \in C_b(X \times Y)$,

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

Obviously, $m \star \gamma \in \mathcal{M}_{\text{product}}^+(X, m, Y)$.

- The disintegration theorem [3] states that, given $\alpha \in \mathcal{M}_{\text{product}}^+(X, m, Y)$, one can construct a weakly measurable function $\gamma : X \rightarrow \mathcal{P}(Y)$ such that $\alpha = m \star \gamma$. If γ is a disintegration of α , we denote

$$\alpha(dy|x) \triangleq \gamma(x, dy).$$

The disintegration is unique m -a.e.

2.2 Mean field type differential game

The main object of the paper is the infinite system of identical agents with the dynamics governed by ordinary differential equations influenced by two players with opposite purposes. We assume the mean field interaction between the agents. For simplicity, let the phase space of each agent be $\mathbb{T}^d \triangleq \mathbb{R}^d / \mathbb{Z}^d$. An element of \mathbb{T}^d is a set $x = \{x' + n : n \in \mathbb{Z}^d\}$. We denote the metric on \mathbb{T}^d by $\|x - y\|$:

$$\|x - y\| = \min\{\|x' - y'\| : x' \in x, y' \in y\}.$$

Let $\mathcal{C}_{s,r} \triangleq C([s, r], \mathbb{T}^d)$ be the set of all trajectories on $[s, r]$. If $x(\cdot), y(\cdot) \in \mathcal{C}_{s,r}$, then distance between $x(\cdot), y(\cdot)$ is denoted (with some abuse of notation) by $\|x(\cdot) - y(\cdot)\|$

$$\|x(\cdot) - y(\cdot)\| \triangleq \sup_{t \in [s,r]} \|x(t) - y(t)\|.$$

Furthermore, if $t \in [s, r]$, then denote by e_t the evaluation operator from $\mathcal{C}_{s,r}$ to \mathbb{T}^d acting by the rule

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

Let the dynamics of each agent be given by

$$\begin{aligned} \frac{d}{dt} x(t) &= f(t, x(t), m(t), u(t), v(t)), \\ t \in [0, T], \quad x(t) &\in \mathbb{T}^d, \quad m(t) \in \mathcal{P}^2(\mathbb{T}^d), \quad u(t) \in U, \quad v(t) \in V. \end{aligned} \tag{1}$$

It is assumed that $u(t)$ and $v(t)$ are control of the first and second players respectively. The concept of mean field type control processes implies that agents are influenced by the players independently.

We assume that the first (respectively, second) player tries to minimize (respectively, maximize) the value

$$g(m(T)).$$

The following conditions are imposed:

- the sets U and V are metric compacts;
- the functions f and g are continuous;
- the function f is Lipschitz continuous with respect to x and m , i.e., there exists a constant $L > 0$ such that, for any $t \in [0, T]$, $x_1, x_2 \in \mathbb{T}^d$, $m_1, m_2 \in \mathcal{P}^2(\mathbb{T}^d)$, $u \in U$, $v \in V$,

$$\|f(t, x_1, m_1, u, v) - f(t, x_2, m_2, u, v)\| \leq L(\|x_1 - x_2\| + W_2(m_1, m_2)); \quad (2)$$

- (Isaacs' condition) for every $t \in [0, T]$, $x \in \mathbb{T}^d$, $m \in \mathcal{P}^2(\mathbb{T}^d)$, and $w \in \mathbb{R}^d$,

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

These conditions implies that

$$R = \max_{x \in \mathbb{T}^d} \|f(t, x, m, u, v)\| < \infty. \quad (3)$$

Let us recall the definitions of open-loop and feedback strategies and corresponding motions of system (1) proposed in [5]. First, we consider the motion of each agent assuming that the flow of probabilities $m(\cdot)$ is given. Then, we introduce the flow of probabilities generated by the distribution of open-loop strategies. Finally, we introduce the feedback strategies and strategies with memory of the players and the corresponding flows of probabilities.

As it was mentioned above, we start with the motions of each player. First, recall the classes of strategies.

- A constant control of the first (respectively, second) player is an element $u \in U$ (respectively, $v \in V$).
- A measurable control of the first player is a measurable function $u : [0, T] \rightarrow U$. Analogously, a second player's measurable control is a measurable function $v : [0, T] \rightarrow V$. Denote the set of measurable controls of the first (respectively, second) player by \mathcal{U}_0 (respectively, \mathcal{V}_0).
- A relaxed control of the first (respectively, second) player is a measure on $[0, T] \times U$ (respectively, on $[0, T] \times V$) compatible with the Lebesgue measure. This means that the set of relaxed controls of the first player is

$$\mathcal{U} \triangleq \mathcal{M}_{\text{product}}^+([0, T], \lambda, U),$$

where λ stands for the Lebesgue measure on $[0, T]$. Analogously, the set of relaxed control of the second player is

$$\mathcal{V} \triangleq \mathcal{M}_{\text{product}}^+([0, T], \lambda, V).$$

This means that, given $\xi \in \mathcal{U}$, $\zeta \in \mathcal{V}$, we regard their disintegrations with respect to the Lebesgue measure $\xi(\cdot|t)$ and $\zeta(\cdot|t)$ as distributions of instantaneous players' controls applied at time t .

Notice that the sets of constant control are embedded into the sets of measurable controls. Furthermore, we identify a measurable control $u(\cdot) : [0, T] \rightarrow U$ with a relaxed control $t \mapsto \delta_{u(t)}$. Recall that δ_z denotes the Dirac measure concentrated at z . One can also identify a measurable function $v(\cdot)$ with a weakly measurable mapping $t \mapsto \delta_{v(t)}$. This provides the embedding of the sets of measurable controls of the players to the set of relaxed controls. Below, with some abuse of notation we assume that

$$U \subset \mathcal{U}_0 \subset \mathcal{U}, \tag{4}$$

$$V \subset \mathcal{V}_0 \subset \mathcal{V}. \tag{5}$$

Let s be an initial time, $r \in (s, T]$, y be an initial position of an agent, $\xi \in \mathcal{U}$, $\zeta \in \mathcal{V}$ be players' controls acting upon an agent, $[s, r] \ni t \mapsto m(t) \in \mathcal{P}^2(\mathbb{T}^d)$ be a flow of probabilities. Denote by $x(\cdot, s, y, m(\cdot), \xi, \zeta)$ the unique solution of the initial value problem

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

Let $\text{traj}_{m(\cdot)}^{s,r} : \mathbb{T}^d \times \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{C}_{s,r}$ denote the operator assigning to y, ξ and ζ the trajectory $x(\cdot, s, y, m(\cdot), \xi, \zeta)$.

Now we turn to the definition of evolution of the whole mean field type system driven by open-loop strategies. Recall that the state space of the whole system is $\mathcal{P}^2(\mathbb{T}^d)$. Since the agents placed in one point can be influenced by various controls, we use distribution of controls. They are probabilities on the product of phase space and space of controls. Using disintegration, one can regard them as weakly measurable functions of position taking values in the space of controls. Let $m \in \mathcal{P}^2(\mathbb{T}^d)$ be a distribution of agents. First, set

$$\mathcal{A}^c[m] \triangleq \mathcal{M}_{\text{product}}^+(\mathbb{T}^d, m, U), \quad \mathcal{B}^c[m] \triangleq \mathcal{M}_{\text{product}}^+(\mathbb{T}^d, m, V).$$

Elements $\mathcal{A}^c[m]$ and $\mathcal{B}^c[m]$ are distributions of the constant controls of the first and second players respectively compatible with the measure m .

In the same way, one can introduce the distribution of relaxed controls of the first and second players by the rules. Put

$$\mathcal{A}[m] \triangleq \mathcal{M}_{\text{product}}^+(\mathbb{T}^d, m, \mathcal{U}), \quad \mathcal{B}[m] \triangleq \mathcal{M}_{\text{product}}^+(\mathbb{T}^d, m, \mathcal{V}).$$

Due to conventions (4) and (5), we assume that

$$\mathcal{A}^c[m] \subset \mathcal{A}[m], \quad \mathcal{B}^c[m] \subset \mathcal{B}[m].$$

The definition of feedback strategies proposed in [5] requires the notion of joint distributions of players' controls. A joint distribution of players' controls is a probability on $\mathbb{T}^d \times \mathcal{U} \times \mathcal{V}$ compatible with the distribution of agent m , i.e., the set of joint distribution of players controls is

$$\mathcal{D}[m] \triangleq \mathcal{M}_{\text{product}}^+(\mathbb{T}^d, m, \mathcal{U} \times \mathcal{V}).$$

If $\alpha \in \mathcal{A}[m]$, then denote by $\mathcal{D}_1[\alpha]$ the set of joint distribution of players' control produced by the reaction of the second players to α , i.e.,

$$\mathcal{D}_1[\alpha] \triangleq \{\varkappa \in \mathcal{D}[m] : p^{1,2} \# \varkappa = \alpha\}.$$

In the same way we define the set of joint distributions of players' controls compatible with the distribution of first player's controls $\beta \in \mathcal{B}[m]$:

$$\mathcal{D}_2[\beta] \triangleq \{\varkappa \in \mathcal{D}[m] : p^{1,3} \# \varkappa = \beta\}.$$

Definition 2.1. Let $s, r \in [0, T]$, $s < r$, $m_* \in \mathcal{P}^2(\mathbb{T}^d)$ be an initial distribution of players, $\varkappa \in \mathcal{D}[m_*]$ be a joint distribution of players' controls consistent with m_* , we say that $[s, r] \ni t \mapsto m(t) \in \mathcal{P}^2(\mathbb{T}^d)$ is a flow of probabilities generated by s , m_* and \varkappa and write $m(\cdot) = m(\cdot, s, m_*, \varkappa)$ if there exists $\chi \in \mathcal{P}^2(\mathcal{C}_{s,r})$ such that

- $m(s) = m_*$;
- $m(t) = e_t \# \chi$;
- $\chi = \text{traj}_{m(\cdot)}^{s,r} \# \varkappa$.

It can be proved in the same spirit as in [33] that there exists the unique flow of probabilities $m(\cdot, s, m_*, \varkappa)$.

It is natural to assume that the players can use information about the state of the whole system to adjust their control. Recall that the state of the system in the mean field setting is provided by the distribution of agents. We use the Krasovskii-Subbotin approach [22] which assumes that the controls is formed using the information about the state of the system in the finite number of time instants.

Definition 2.2. Let t_0 be an initial time, $\Delta = \{t_i\}_{i=0}^N$ be a partition of $[t_0, T]$. A stepwise strategy with memory of the first player is a collection $\mathfrak{u}^\Delta = \{\mathfrak{u}_k^\Delta\}_{k=0}^{N-1}$ where $\mathfrak{u}_k^\Delta : (\mathcal{P}^2(\mathbb{T}^d))^{k+1} \rightarrow \mathcal{P}^2(\mathbb{T}^d \times U)$ are such that

$$p^1 \# \mathfrak{u}_k^\Delta [m_0, \dots, m_k] = m_k.$$

The value $\mathfrak{u}_k^\Delta [m_0, \dots, m_k]$ determines the distribution of controls used by the first player on the time interval $[t_k, t_{k+1})$.

Definition 2.3. Let t_0 be an initial time, m_0 be an initial position, Δ be a partition of $[t_0, T]$, $\mathfrak{u}^\Delta = \{\mathfrak{u}_k^\Delta\}_{k=0}^{N-1}$ be a stepwise strategy with memory of the first player. We say that a flow of probability $m(\cdot)$ is generated by t_0 , m_0 and \mathfrak{u}^Δ if there exists a sequence of probabilities $\alpha_k \in \mathcal{P}^2(\mathbb{T}^d \times U)$, $\varkappa_k \in \mathcal{P}^2(\mathbb{T}^d \times U \times \mathcal{V})$, $k = 0, \dots, N-1$, such that

- $m(t_0) = m_0$;
- $\alpha_k \triangleq \mathfrak{u}_k^\Delta[m(t_0), m(t_1), \dots, m(t_k)], \mathfrak{x}_k \in \mathcal{D}_1[\alpha_k]$;
- $m(t) = m(t, t_k, m(t_k), \mathfrak{x}_k), t \in [t_k, t_{k+1}], k = 0, \dots, N-1$.

We denote the set of flows of probabilities generated by t_0, m_0, Δ , and \mathfrak{u}^Δ by $\mathcal{X}_1(t_0, m_0, \Delta, \mathfrak{u}^\Delta)$.

The second player's strategies are defined in the similar way.

Definition 2.4. Given $t_0 \in [0, T]$ and a partition of $[t_0, T]$ $\Delta = \{t_k\}_{k=0}^N$, a collection $\mathfrak{u}^\Delta = \{\mathfrak{u}_k^\Delta\}_{k=0}^{N-1}$ is called a stepwise strategy of the second player provided that $\mathfrak{v}_k^\Delta : (\mathcal{P}^2(\mathbb{T}^d))^{k+1} \rightarrow \mathcal{P}^2(\mathbb{T}^d \times V)$ are such that

$$\mathfrak{p}^1 \# \mathfrak{v}_k^\Delta[m_0, \dots, m_k] = m_k.$$

Definition 2.5. We say that a flow of probabilities $m(\cdot)$ is generated by the initial time t_0 , the initial distribution of agents m_0 , the partition of $[t_0, T]$ $\Delta = \{t_k\}_{k=0}^N$ and the stepwise strategy with memory of the second player $\mathfrak{v}^\Delta = \{\mathfrak{v}_k^\Delta\}_{k=0}^{N-1}$ if one can find joint distributions of players' controls $\beta_k \in \mathcal{P}^2(\mathbb{T}^d \times V), \mathfrak{x}_k \in \mathcal{P}^2(\mathbb{T}^d \times \mathcal{U} \times V), k = 0, \dots, N-1$, such that

- $m(t_0) = m_0$;
- $\beta_k = \mathfrak{v}_k^\Delta[m(t_0), \dots, m(t_k)], \mathfrak{x}_k \in \mathcal{D}_2[\beta_k]$;
- $m(t) = m(t, t_k, m(t_k), \mathfrak{x}_k), t \in [t_k, t_{k+1}], k = 0, \dots, N-1$.

Denote the set of flow of probabilities generated by t_0, m_0, Δ and \mathfrak{v}^Δ by $\mathcal{X}_2(t_0, m_0, \Delta, \mathfrak{v}^\Delta)$.

Now let us briefly describe the feedback approach to the mean field type differential games [5]. It implies that a feedback strategy of the first player is a mapping $\mathfrak{u} : [0, T] \times \mathcal{P}^2(\mathbb{T}^d) \rightarrow \mathcal{P}^2(\mathbb{T}^d \times U)$ such that

$$\mathfrak{p}^1 \# \mathfrak{u}[t, m] = m.$$

In this case, given a partition $\Delta = \{t_k\}_{k=0}^N$ of the time interval $[t_0, T]$, the stepwise strategy $\mathfrak{u}^\Delta = \{\mathfrak{u}_k^\Delta\}_{k=0}^{N-1}$ realizing the feedback strategy \mathfrak{u} is defined as

$$\mathfrak{u}_k^\Delta[m_0, \dots, m_k] \triangleq \mathfrak{u}[t_k, m_k].$$

Analogously, a feedback strategy of the second player is a mapping $\mathfrak{v} : [0, T] \times \mathcal{P}^2(\mathbb{T}^d) \rightarrow \mathcal{P}^2(\mathbb{T}^d \times V)$ such that

$$\mathfrak{p}^1 \# \mathfrak{v}[t, m] = m.$$

Given a partition of the interval $[t_0, T]$ $\Delta = \{t_k\}_{k=0}^N$, one can reduce the feedback strategy to the strategies with memory letting

$$\mathfrak{v}_k^\Delta[m_0, \dots, m_k] \triangleq \mathfrak{v}[t_k, m_k].$$

The upper and lower value functions by the rules

$$\text{Val}^+(t_0, m_0) = \inf_{\Delta, \mathfrak{u}^\Delta} \sup_{m(\cdot) \in \mathcal{X}_1(t_0, m_0, \Delta, \mathfrak{u}^\Delta)} g(m(T)), \quad (6)$$

$$\text{Val}^-(t_0, m_0) = \sup_{\Delta, \mathfrak{v}^\Delta} \inf_{m(\cdot) \in \mathcal{X}_2(t_0, m_0, \Delta, \mathfrak{v}^\Delta)} g(m(T)). \quad (7)$$

Allowing in (6) only the stepwise realizations of feedback strategies, we obtain the upper value function in the class of feedback strategies. Denote it by $\text{Val}_f^+(t_0, m_0)$. Analogously, reducing the class of strategies in (7) to the stepwise realizations of second player's feedback strategies, we arrive at the notion of lower value function in the class of feedback strategies $\text{Val}_f^-(t_0, m_0)$.

Obviously,

$$\text{Val}^-(t_0, m_0) \leq \text{Val}_f^-(t_0, m_0) \leq \text{Val}_f^+(t_0, m_0) \leq \text{Val}^+(t_0, m_0).$$

It is proved (see [5, Theorem 2]) that, under imposed conditions, the game has the value in the class of feedback strategies, i.e.,

$$\text{Val}_f^-(t_0, m_0) = \text{Val}_f^+(t_0, m_0) = \text{Val}^+(t_0, m_0) = \text{Val}^-(t_0, m_0) = \text{Val}(t_0, m_0).$$

Hereinafter, Val denotes the value of the game. Below we estimate it using sub- and supersolutions of the Hamilton-Jacobi equations corresponding to the finite-dimensional differential games.

3 Finite state mean field type Markov games

In this section, we consider the finite state mean field type continuous-time game. The main feature of the finite state mean field systems is that a phase variable is an element of a finite-dimensional simplex. Thus, the finite state mean field type Markov games can be reduced to the finite-dimensional differential game theory.

Let \mathcal{S} be a finite set. We will examine the approximation of original mean field type differential game by the mean field type Markov game with the state space of each agent equal to \mathcal{S} . Thus, without loss of generality, we assume that $\mathcal{S} \subset \mathbb{T}^d$. Denote by $d(\mathcal{S})$ the fineness of \mathcal{S} :

$$d(\mathcal{S}) \triangleq \min_{\bar{x}, \bar{y} \in \mathcal{S}, \bar{x} \neq \bar{y}} \|\bar{x} - \bar{y}\|. \quad (8)$$

Furthermore, let Σ be a simplex on $\{1, \dots, |\mathcal{S}|\}$:

$$\Sigma \triangleq \left\{ \mu = (\mu_{\bar{x}})_{\bar{x} \in \mathcal{S}} : \mu_{\bar{x}} \geq 0, \sum_{x \in \mathcal{S}} \mu_{\bar{x}} = 1 \right\}.$$

Notice that

$$\Sigma \subset \mathbb{R}^{|\mathcal{S}|}.$$

Denote by $\mathbb{1}_{\bar{y}}$ the distribution concentrated at \bar{y} , i.e., $\mathbb{1}_{\bar{y}} = (\mathbb{1}_{\bar{y},\bar{x}})_{\bar{x} \in \mathcal{S}}$ where

$$\mathbb{1}_{\bar{y},\bar{x}} = \begin{cases} 1, & \bar{x} = \bar{y}, \\ 0, & \bar{x} \neq \bar{y}. \end{cases} \quad (9)$$

The distributions $\mathbb{1}_{\bar{y}}$ are extremal points of Σ .

Given $p \geq 1$, the p -th metric on $\mathbb{R}^{|\mathcal{S}|}$ (and, thus, on Σ) is defined by the rule: for $\mu^1 = (\mu_{\bar{x}}^1)_{\bar{x} \in \mathcal{S}}, \mu^2 = (\mu_{\bar{x}}^2)_{\bar{x} \in \mathcal{S}} \in \mathbb{R}^{|\mathcal{S}|}$,

$$\|\mu^1 - \mu^2\|_p \triangleq \left[\sum_{x \in \mathcal{S}} |\mu_x^1 - \mu_x^2|^p \right]^{1/p}.$$

If $\mu = (\mu_{\bar{x}})_{\bar{x} \in \mathcal{S}} \in \Sigma$, then let $\tilde{\mu}$ stand for the measure on \mathcal{S} defined by

$$\tilde{\mu} \triangleq \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} \delta_{\bar{x}} \in \mathcal{P}(\mathcal{S}). \quad (10)$$

The mapping

$$\Sigma \ni \mu \mapsto \tilde{\mu} \in \mathcal{P}(\mathcal{S})$$

is an isomorphism between Σ and $\mathcal{P}(\mathcal{S})$. However, the metrics on Σ and $\mathcal{P}(\mathcal{S})$ are not equivalent. The relation between them is given in the following.

Proposition 3.1. *There exist positive constants C_1 and C_2 (depending on \mathcal{S}) such that, if $\mu^1 = (\mu_{\bar{x}}^1)_{\bar{x} \in \mathcal{S}}, \mu^2 = (\mu_{\bar{x}}^2)_{\bar{x} \in \mathcal{S}} \in \Sigma, \tilde{\mu}^1 = \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}}^1 \delta_{\bar{x}}, \tilde{\mu}^2 = \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}}^2 \delta_{\bar{x}}$, then*

$$\|\mu^1 - \mu^2\|_p \leq C_1 W_p(\tilde{\mu}^1, \tilde{\mu}^2), \quad (11)$$

$$W_p(\tilde{\mu}^1, \tilde{\mu}^2) \leq C_2 (\|\mu^1 - \mu^2\|_p)^{1/p}. \quad (12)$$

This proposition is proved in the Appendix.

We regard the set \mathcal{S} to be the state space for each agent in the mean field type Markov game, whereas Σ is the set of distributions on \mathcal{S} . The dynamics of the finite state mean field type game is given by the controlled Markov chain with the transition rates determined by the Kolmogorov matrix

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

where $t \in [0, T], \mu \in \Sigma, u \in U, v \in V$. As above, the variable u is controlled by the first player, whereas v is chosen by the second player. The mean field methodology implies that $\mu(t) = \text{Law}(X(t))$, where $X(t)$ is the stochastic process produced by the nonlinear Markov chain with the Kolmogorov matrix $Q(t, \mu(t), u(t), v(t))$. We assume that the first (second) player tries to minimize (maximize) the quantity

$$g(\widetilde{\mu(T)}).$$

We impose the following conditions on the matrix Q :

(M1) for every $(t, \mu, u, v) \in [0, T] \times \Sigma \times U \times V$, $Q_{\bar{x}, \bar{y}}(t, \mu, u, v) \geq 0$ when $\bar{x} \neq \bar{y}$ and

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

(M2) the functions $[0, T] \times \Sigma \times U \times V \ni (t, \mu, u, v) \mapsto Q_{\bar{x}, \bar{y}}(t, \mu, u, v) \in \mathbb{R}$ are continuous;

(M3) there exists a constant L' such that for any $t \in [0, T]$, $\bar{x}, \bar{y} \in \mathcal{S}$, $\mu^1, \mu^2 \in \Sigma$, $u \in U$, $v \in V$,

$$|Q_{\bar{x}, \bar{y}}(t, \mu^1, u, v) - Q_{\bar{x}, \bar{y}}(t, \mu^2, u, v)| \leq L' \|\mu^1 - \mu^2\|_2.$$

As above, considering the finite state mean field type differential games, we first define the motion of the representative agent, then turn to the behavior of the whole system. On the first stage, we assume that the flow of probabilities $[s, r] \ni t \mapsto \mu(t) \in \Sigma$ is given, while the players use relaxed feedback strategies. For a Markov game, it is convenient to introduce relaxed feedback controls of the players using weakly measurable functions.

Let

$$\mathcal{U}_{\text{instant}} \triangleq \mathcal{P}(U), \quad \mathcal{V}_{\text{instant}} \triangleq \mathcal{P}(V)$$

be sets of players' controls applied at each time instant and each state from \mathcal{S} . Therefore, the instantaneous controls of the players lie in the sets

$$\begin{aligned} \mathcal{U}_{\text{instant}}^{\mathcal{S}} &= \{\gamma_{\mathcal{S}} = (\gamma_{\bar{x}})_{\bar{x} \in \mathcal{S}} : \gamma_{\bar{x}} \in \mathcal{U}_{\text{instant}}\}, \\ \mathcal{V}_{\text{instant}}^{\mathcal{S}} &= \{\vartheta_{\mathcal{S}} = (\vartheta_{\bar{x}})_{\bar{x} \in \mathcal{S}} : \vartheta_{\bar{x}} \in \mathcal{V}_{\text{instant}}\}. \end{aligned}$$

We say that a function $[s, r] \ni t \mapsto \gamma_{\mathcal{S}}(t) = (\gamma_{\bar{x}})_{\bar{x} \in \mathcal{S}} \in \mathcal{U}_{\text{instant}}^{\mathcal{S}}$ is a measurable control of the first player in the finite state mean field type continuous-times game if each coordinate function $\gamma_{\bar{x}}(\cdot)$ is weakly measurable. The measurable controls of the second player are introduced in the same way. Given measurable controls $\gamma_{\bar{x}} : [0, T] \rightarrow \mathcal{U}_{\text{instant}}$, $\vartheta_{\bar{x}} : [0, T] \rightarrow \mathcal{V}_{\text{instant}}$ of the first and second players respectively, we regard $\gamma_{\bar{x}}(t)$, $\vartheta_{\bar{x}}(t)$ as the instantaneous controls of the players acting upon the agent who occupy the state \bar{x} at time t .

Considering the relaxed feedback strategy, we arrive at the following Kolmogorov matrix:

$$\mathcal{Q}(t, \mu, \gamma_{\mathcal{S}}, \vartheta_{\mathcal{S}}) = (\mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, \gamma_{\bar{x}}, \vartheta_{\bar{x}}))_{\bar{x}, \bar{y} \in \mathcal{S}}$$

with

$$\mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, \gamma_{\bar{x}}, \vartheta_{\bar{x}}) \triangleq \int_U \int_V Q_{\bar{x}, \bar{y}}(t, \mu, u, v) \gamma_{\bar{x}}(du) \vartheta_{\bar{x}}(dv).$$

The probabilities of the states $\nu(t) = (\nu_{\bar{y}}(t))_{\bar{y} \in \mathcal{S}}$, $\nu_{\bar{y}} = P(X(t) = \bar{y})$ obeys the backward Kolmogorov equation

$$\frac{d}{dt} \nu_{\bar{y}}(t) = \sum_{\bar{x} \in \mathcal{S}} \nu_{\bar{x}}(t) \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu(t), \gamma_{\bar{x}}(t), \vartheta_{\bar{x}}(t)), \quad \bar{y} \in \mathcal{S}.$$

Assuming that $\nu = (\nu_{\bar{x}})_{\bar{x} \in \mathcal{S}}$ is a row-vector with some ordering of \mathcal{S} , we rewrite this equations in the vector form

$$\frac{d}{dt} \nu(t) = \nu(t) \mathcal{Q}(t, \mu(t), \gamma_{\mathcal{S}}(t), \vartheta_{\mathcal{S}}(t)), \quad \nu(s) = \nu_*. \quad (13)$$

The mean field game methodology implies that the probability of the event that an agent occupies the state \bar{x} at time t is equal to $\mu_{\bar{x}}(t)$. This leads to the following equations:

$$\frac{d}{dt}\mu_{\bar{y}}(t) = \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}}(t) \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu(t), \gamma_{\bar{x}}(t), \vartheta_{\bar{x}}(t)), \quad \bar{y} \in \mathcal{S}$$

or in the vector form

$$\frac{d}{dt}\mu(t) = \mu(t) \mathcal{Q}(t, \mu(t), \gamma_{\mathcal{S}}(t), \vartheta_{\mathcal{S}}(t)). \quad (14)$$

Equation (14) describes the dynamics of finite-dimensional differential game with the controls of the players $\gamma_{\mathcal{S}}(t)$, $\vartheta_{\mathcal{S}}(t)$. Recall that the first players tries to minimize

$$\hat{g}(\mu(T)), \quad (15)$$

where

$$\hat{g}(\mu) \triangleq g(\tilde{\mu}) = g\left(\sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} \delta_{\bar{x}}\right). \quad (16)$$

The purpose of the second player is assumed to be opposite.

The solution of this differential game can be described using the notion of the viscosity/minimax solution of the corresponding Hamilton-Jacobi equation. Let $t \in [0, T]$, $\mu = (\mu_{\bar{x}})_{\bar{x} \in \mathcal{S}} \in \Sigma$, $w = (w_{\bar{x}})_{\bar{x} \in \mathcal{S}} \in \mathbb{R}^{\mathcal{S}}$. Put

$$\begin{aligned} \mathcal{H}^{\mathcal{Q}}(t, \mu, w) &\triangleq \min_{\gamma_{\mathcal{S}} \in \mathcal{U}_{\text{instant}}^{\mathcal{S}}} \max_{\vartheta_{\mathcal{S}} \in \mathcal{V}_{\text{instant}}^{\mathcal{S}}} \sum_{\bar{y} \in \mathcal{S}} \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, \gamma_{\bar{x}}, \vartheta_{\bar{x}}) w_{\bar{y}} \\ &= \min_{\gamma_{\mathcal{S}} \in \mathcal{U}_{\text{instant}}^{\mathcal{S}}} \max_{\vartheta_{\mathcal{S}} \in \mathcal{V}_{\text{instant}}^{\mathcal{S}}} \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} \sum_{\bar{y} \in \mathcal{S}} \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, \gamma_{\bar{x}}, \vartheta_{\bar{x}}) w_{\bar{y}}. \end{aligned}$$

Since we admit the feedback controls $\gamma_{\mathcal{S}} = (\gamma_{\bar{x}})_{\bar{x} \in \mathcal{S}}$, $\vartheta_{\mathcal{S}} = (\vartheta_{\bar{x}})_{\bar{x} \in \mathcal{S}}$, we have that

$$\mathcal{H}^{\mathcal{Q}}(t, \mu, w) = \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} \min_{\gamma_{\bar{x}} \in \mathcal{U}_{\text{instant}}} \max_{\vartheta_{\bar{x}} \in \mathcal{V}_{\text{instant}}} \sum_{\bar{y} \in \mathcal{S}} \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, \gamma_{\bar{x}}, \vartheta_{\bar{x}}) w_{\bar{y}}.$$

Furthermore, by construction of \mathcal{Q} , we have

$$\sum_{\bar{y} \in \mathcal{S}} \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, \gamma_{\bar{x}}, \vartheta_{\bar{x}}) w_{\bar{y}} = \int_U \int_V \sum_{\bar{y} \in \mathcal{S}} \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, u, v) w_{\bar{y}} \gamma_{\bar{x}}(du) \vartheta_{\bar{x}}(dv).$$

Thus, the function

$$\sum_{\bar{y} \in \mathcal{S}} \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, \gamma_{\bar{x}}, \vartheta_{\bar{x}}) w_{\bar{y}}$$

is convex w.r.t. $\gamma_{\bar{x}}$ and concave w.r.t. $\vartheta_{\bar{x}}$. From the minimax theorem [31], it follows that

$$\begin{aligned} \mathcal{H}^{\mathcal{Q}}(t, \mu, w) &= \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} \max_{\vartheta_{\bar{x}} \in \mathcal{V}_{\text{instant}}} \min_{\gamma_{\bar{x}} \in \mathcal{U}_{\text{instant}}} \sum_{\bar{y} \in \mathcal{S}} \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, \gamma_{\bar{x}}, \vartheta_{\bar{x}}) w_{\bar{y}} \\ &= \max_{\vartheta_{\mathcal{S}} \in \mathcal{V}_{\text{instant}}^{\mathcal{S}}} \min_{\gamma_{\mathcal{S}} \in \mathcal{U}_{\text{instant}}^{\mathcal{S}}} \sum_{\bar{y} \in \mathcal{S}} \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, \gamma_{\bar{x}}, \vartheta_{\bar{x}}) w_{\bar{y}} \end{aligned}$$

In particular, we showed that differential game (14), (15) satisfies the Isaacs' condition. For $\varphi : [0, T] \times \Sigma \rightarrow \mathbb{R}$, we denote by $\nabla\varphi(t, \mu)$ the vector of partial derivatives

$$\nabla\varphi(t, \mu) \triangleq \left(\frac{\partial\varphi}{\partial\mu_{\bar{x}}}(t, \mu) \right)_{\bar{x} \in \mathcal{S}}.$$

Consider the Hamilton-Jacobi equations

$$\frac{\partial\varphi}{\partial t} + \mathcal{H}^{\mathcal{Q}}(t, \mu, \nabla\varphi) = 0. \quad \varphi(T, \mu) = \hat{g}(\mu),$$

where \hat{g} is introduced by (16).

Notice that a supersolutions (respectively, subsolution) of equation (3) is an upper (lower) bound for the value function for the differential game (14), (15) [32]. The solution of (3) is the value function of differential game (14), (15) (see [32]).

4 Main result

The purpose of this Section is to formulate Theorem 4.1 which states that, if the set \mathcal{S} and the Markov chain with the Kolmogorov matrix $Q(t, \mu, u, v)$ approximate the phase space for the original mean field type differential game \mathbb{T}^d and the dynamics f , then supersolutions (respectively, subsolutions) of (3) provide upper (respectively, lower) estimates of the players' outcomes in the original mean field type differential game. This also yields the estimates of the value function of the mean field type differential game by the value function of finite-dimensional differential game (see Corollary 4.2).

First, let $\ell : \mathbb{T}^d \times \mathbb{T}^d \rightarrow \mathbb{R}^d$ be a measurable function assigning to a pair of elements $x, y \in \mathbb{T}^d$ a vector $z' \in x - y$ of the minimal norm. The existence of this function follows from the measurable maximum theorem [2, Theorem 18.19]. Notice that, for $x, y, z \in \mathbb{T}^d$,

$$\|x - y\|^2 - \|y - z\|^2 = \|x - z\|^2 - 2\langle \ell(x, z), \ell(y, z) \rangle. \quad (17)$$

Recall that we assume that $\mathcal{S} \subset \mathbb{T}^d$. Let $\varepsilon > 0$ be such that

$$\max_{x \in \mathbb{T}^d} \min_{\bar{y} \in \mathcal{S}} \|x - \bar{y}\| \leq \varepsilon, \quad (18)$$

$$\max_{t \in [0, T], \bar{x} \in \mathcal{S}, \mu \in \Sigma, u \in U, v \in V} \left\| f(t, \bar{x}, \tilde{\mu}, u, v) - \sum_{\bar{y} \in \mathcal{S}, \bar{y} \neq \bar{x}} \ell(\bar{y}, \bar{x}) Q_{\bar{x}, \bar{y}}(t, \mu, u, v) \right\| \leq \varepsilon, \quad (19)$$

and

$$\max_{t \in [0, T], \bar{x} \in \mathcal{S}, \mu \in \Sigma, u \in U, v \in V} \sum_{\bar{y} \in \mathcal{S}} \|\bar{y} - \bar{x}\|^2 Q_{\bar{x}, \bar{y}}(t, \mu, u, v) \leq \varepsilon^2. \quad (20)$$

Below, without loss of generality, we assume that $\varepsilon \leq 1$.

Further, given $m \in \mathcal{P}^2(\mathbb{T}^d)$, denote by $\text{pr}_{\mathcal{S}}(m)$ an element of Σ such that $\widetilde{\text{pr}_{\mathcal{S}}(m)}$ is a proximal to m element of $\mathcal{P}^2(\mathcal{S})$. This means that $\text{pr}_{\mathcal{S}}(m)$ minimizes the function

$$\Sigma \ni \mu = (\mu_{\bar{x}})_{\bar{x} \in \mathcal{S}} \mapsto W_2(\tilde{\mu}, m) = W_2 \left(\sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} \delta_{\bar{x}}, m \right).$$

As above, due to the measurable maximum theorem (see [2, Theorem 18.19]) one may choose the mapping $m \mapsto \text{pr}_S(m)$ to be measurable.

Let ς_g be a modulus of continuity for the function g , i.e.,

$$\varsigma_g(\theta) \triangleq \sup\{|g(m') - g(m'')| : m', m'' \in \mathcal{P}^2(\mathbb{T}^d), W_2(m', m'') \leq \theta\}.$$

Set

$$C^* \triangleq \sqrt{1 + 2T}e^{(2L+1/2)T}.$$

Here L is a Lipschitz constant of the function f w.r.t. x and m (see (2)).

Theorem 4.1. *(i) Given a supersolution of (3) φ^+ , $t_0 \in [0, T]$, and a partition of $[t_0, T]$ Δ , there exists a first player's strategy with memory \mathfrak{u}^Δ such that, for every $m_0 \in \mathcal{P}^2(\mathbb{T}^d)$ and $m(\cdot) \in \mathcal{X}_1(t_0, m_0, \Delta, \mathfrak{u}^\Delta)$,*

$$g(m(T)) \leq \varphi^+(t_0, \text{pr}_S(m_0)) + \varsigma_g(C^*\varepsilon + \varsigma^*(d(\Delta))).$$

(ii) For a function φ^- that is a subsolution of (3), an initial time t_0 and a partition of $[t_0, T]$ Δ , one can construct a strategy with memory of the second player \mathfrak{v}^Δ , satisfying the following: for every $m_0 \in \mathcal{P}^2(\mathbb{T}^d)$ and $m(\cdot) \in \mathcal{X}_2(t_0, m_0, \Delta, \mathfrak{v}^\Delta)$,

$$g(m(T)) \geq \varphi^-(t_0, \text{pr}_S(m_0)) - \varsigma_g(C^*\varepsilon + \varsigma^*(d(\Delta))).$$

Here $\varsigma^*(\varepsilon)$ is a function taking values in $[0, +\infty)$ satisfying $\varsigma^*(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$.

This theorem is proved in Sections 7–8.

Furthermore, Theorem 4.1 implies the following estimates on the value function of the mean field type differential game.

Corollary 4.2. *Let φ^+ and φ^- be a supersolution and a subsolution of (3) respectively. If $t_0 \in [0, T]$, $m_0 \in \mathcal{P}^2(\mathbb{T}^d)$, then*

$$\varphi^-(t_0, \text{pr}_S(m_0)) - \varsigma_g(C^*\varepsilon) \leq \text{Val}(t_0, m_0) \leq \varphi^+(t_0, \text{pr}_S(m_0)) + \varsigma_g(C^*\varepsilon).$$

If φ is the viscosity/minimax solution of (3), then

$$|\text{Val}(t_0, m_0) - \varphi(t_0, \text{pr}_S(m_0))| \leq \varsigma_g(C^*\varepsilon).$$

5 Design of the approximating Markov chain

In this section, given a system (1), we construct an approximating Markov chain. Here we assume additionally that the function f is Lipschitz continuous with respect to m in 1-Wasserstein metric, i.e., for every $t \in [0, T]$, $x \in \mathbb{T}^d$, $m_1, m_2 \in \mathcal{P}^1(\mathbb{T}^d) = \mathcal{P}^2(\mathbb{T}^d)$, $u \in U$, $v \in V$,

$$\|f(t, x, m_1, u, v) - f(t, x, m_2, u, v)\| \leq L''W_1(m_1, m_2)$$

for some constant L'' .

Let $h > 0$ be such that $1/h \in \mathbb{N}$. Put

$$\mathcal{S}_h \triangleq h\mathbb{Z}^d \cap \mathbb{T}^d.$$

To define the approximating Markov chain, we write the dynamics in the coordinatewise form

$$f(t, x, m, u, v) = (f_1(t, x, m, u, v), \dots, f_d(t, x, m, u, v)).$$

Let e^i stand for the i -th coordinate vector. The approximating Markov chain is defined as follows. For $t \in [0, T]$, $\bar{x}, \bar{y} \in \mathcal{S}_h$, $\mu \in \Sigma$, $u \in U$, $v \in V$, set

$$Q_{\bar{x}, \bar{y}}^h(t, \mu, u, v) \triangleq \begin{cases} \frac{1}{h} |f_i(t, x, \tilde{\mu}, u, v)|, & \bar{y} = \bar{x} + h \operatorname{sgn}(f_i(t, x, \tilde{\mu}, u, v)) e^i, \\ -\frac{1}{h} \sum_{j=1}^d |f_j(t, x, \tilde{\mu}, u, v)|, & \bar{x} = \bar{y}, \\ 0, & \text{otherwise.} \end{cases} \quad (21)$$

Proposition 5.1. *The matrix Q defined by (21) satisfies condition (M1)–(M3) and (18)–(20) with*

$$\varepsilon = \sqrt{h} \cdot \max \left\{ \sqrt{R} \sqrt[4]{d}, \frac{\sqrt[4]{d}}{\sqrt{2}} \right\}. \quad (22)$$

Proof. First, let us check whether the matrix Q^h satisfies conditions (M1)–(M3). Indeed, direct calculations yields that

$$\sum_{\bar{y} \in \mathcal{S}_h, \bar{y} \neq \bar{x}} Q_{\bar{x}, \bar{y}}^h(t, \mu, u, v) = \sum_{i=1}^d \frac{1}{h} |f_i(t, x, \tilde{\mu}, u, v)| - \frac{1}{h} \sum_{j=1}^d |f_j(t, x, \tilde{\mu}, u, v)| = 0.$$

This gives condition (M1). Condition (M2) follows from very definition of the matrix Q^h . Now, recall that we additionally assumed that the function f is Lipschitz continuous with respect to the 1-Wasserstein metric. This implies that

$$|Q_{\bar{x}, \bar{y}}^h(t, \mu^1, u, v) - Q_{\bar{x}, \bar{y}}^h(t, \mu^2, u, v)| \leq \frac{L''}{h} W_1(\tilde{\mu}^1, \tilde{\mu}^2).$$

From Proposition 3.1 it follows that

$$|Q_{\bar{x}, \bar{y}}^h(t, \mu^1, u, v) - Q_{\bar{x}, \bar{y}}^h(t, \mu^2, u, v)| \leq \frac{C_2 L''}{h} \|\mu^1 - \mu^2\|_1.$$

Using Holder's inequality we deduce condition (M3).

Now let us show that the constructed Markov chain approximated the original mean field type control system. We have that

$$\max_{x \in \mathbb{T}^d} \min_{\bar{y} \in \mathcal{S}} \|x - \bar{y}\| \leq \sqrt{d}h/2.$$

Further, let $t \in [0, T]$ $\bar{x} \in \mathcal{S}$, $\mu \in \Sigma$, $u \in U$, $v \in V$. We have that

$$\begin{aligned} \sum_{\bar{y} \in \mathcal{S}_h, \bar{y} \neq \bar{x}} (\bar{y} - \bar{x}) Q_{\bar{x}, \bar{y}}^h(t, \mu, u, v) &= \sum_{i=1}^d h \operatorname{sgn}(f_i(t, x, \tilde{\mu}, u, v)) \frac{1}{h} |f_i(t, x, \tilde{\mu}, u, v)| e^i \\ &= \sum_{i=1}^d f_i(t, x, \tilde{\mu}, u, v) e^i = f(t, x, \tilde{\mu}, u, v). \end{aligned}$$

Finally, we have that

$$\begin{aligned} \sum_{\bar{y} \in \mathcal{S}_h} \|\bar{y} - \bar{x}\|^2 Q_{\bar{x}, \bar{y}}^h(t, \mu, u, v) &= \sum_{i=1}^d \frac{h^2}{h} |f_i(t, x, \tilde{\mu}, u, v)| \\ &\leq h\sqrt{d} \left[\sum_{i=1}^d |f_i(t, x, \tilde{\mu}, u, v)|^2 \right]^{1/2} \leq R\sqrt{d}h, \end{aligned}$$

where the constant R estimates the norm of f (see (3)). Thus, conditions (18)–(20) hold with ε determined by (22). \square

For the Markov chain with Kolmogorov matrix given by (21), we have that

$$\begin{aligned} \mathcal{H}^{\mathcal{Q}}(t, \mu, w) &= \frac{1}{h} \sum_{\bar{x} \in \mathcal{S}_h} \min_{\gamma_{\bar{x}} \in \mathcal{U}_{\text{instant}}} \max_{\vartheta_{\bar{x}} \in \mathcal{V}_{\text{instant}}} \int_U \int_V \sum_{i=1}^d |f_i(t, \bar{x}, \mu, u, v)| \\ &\quad (w_{\bar{x} + h e^i \text{sgn}(f_i(t, \bar{x}, \mu, u, v))} - w_{\bar{x}}) \gamma_{\bar{x}}(du) \vartheta_{\bar{x}}(dv). \end{aligned}$$

The computations of the Hamiltonian are simplified when we assume that

$$f(t, x, m, u, v) = f^1(t, x, m, u) + f^2(t, x, m, v).$$

Here, we also assume that f^1 and f^2 are Lipschitz continuous with respect to m in W_1 . As above we put $\mathcal{S}_h \triangleq \mathbb{T}^d \cap h\mathbb{Z}^d$. Representing the functions f^1 and f^2 in the coordinate-wise form

$$\begin{aligned} f^1(t, x, m, u) &= (f_1^1(t, x, m, u), \dots, f_d^1(t, x, m, u)), \\ f^2(t, x, m, v) &= (f_1^2(t, x, m, v), \dots, f_d^2(t, x, m, v)), \end{aligned}$$

we introduce two Kolmogorov matrices for the dynamics f^1 and f^2 . Set

$$\begin{aligned} Q^{1,h} h_{\bar{x}, \bar{y}}(t, \mu, u) &\triangleq \begin{cases} \frac{1}{h} |f_i^1(t, x, \tilde{\mu}, u)|, & \bar{y} = \bar{x} + h \text{sgn}(f_i^1(t, x, \tilde{\mu}, u)) e^i, \\ -\frac{1}{h} \sum_{j=1}^d |f_j^1(t, x, \tilde{\mu}, u)|, & \bar{x} = \bar{y}, \\ 0, & \text{otherwise,} \end{cases} \\ Q^{2,h} h_{\bar{x}, \bar{y}}(t, \mu, v) &\triangleq \begin{cases} \frac{1}{h} |f_i^2(t, x, \tilde{\mu}, v)|, & \bar{y} = \bar{x} + h \text{sgn}(f_i^2(t, x, \tilde{\mu}, v)) e^i, \\ -\frac{1}{h} \sum_{j=1}^d |f_j^2(t, x, \tilde{\mu}, v)|, & \bar{x} = \bar{y}, \\ 0, & \text{otherwise,} \end{cases} \\ Q_{\bar{x}, \bar{y}}^h(t, \mu, u, v) &\triangleq Q_{\bar{x}, \bar{y}}^{1,h}(t, \mu, u) + Q_{\bar{x}, \bar{y}}^{2,h}(t, \mu, v). \end{aligned} \tag{23}$$

Proposition 5.2. *The Kolmogorov matrix defined by (23) satisfies conditions (M1)–(M3) and (18)–(20) with ε given by (22).*

The proof of this proposition is similar to the proof of Proposition 5.1. In this case, we have that

$$\begin{aligned} \mathcal{H}^{\mathcal{Q}}(t, \mu, w) &= \frac{1}{h} \sum_{\bar{x} \in \mathcal{S}_h} \min_{\gamma_{\bar{x}} \in \mathcal{U}_{\text{instant}}} \int_U \sum_{i=1}^d |f_i^1(t, \bar{x}, \mu, u)| (w_{\bar{x} + h e^i \text{sgn}(f_i^1(t, \bar{x}, \mu, u))} - w_{\bar{x}}) \gamma_{\bar{x}}(du) \\ &\quad + \frac{1}{h} \sum_{\bar{x} \in \mathcal{S}_h} \max_{\vartheta_{\bar{x}} \in \mathcal{V}_{\text{instant}}} \int_V \sum_{i=1}^d |f_i^2(t, \bar{x}, \mu, v)| (w_{\bar{x} + h e^i \text{sgn}(f_i^2(t, \bar{x}, \mu, v))} - w_{\bar{x}}) \vartheta_{\bar{x}}(dv). \end{aligned}$$

6 Properties of mean field type controlled processes

In this section, we study the properties of the original and model systems. The results proved below are used in the proof of Theorem 4.1.

First, let us mention the following property of the flow of probabilities generated by the original mean field type dynamics.

Lemma 6.1. *If $m(\cdot) = m(\cdot, s, m_*, \varkappa)$ for some $s, m_* \in \mathcal{P}^2(\mathbb{T}^d)$, and distribution of controls $\varkappa \in \mathcal{D}[m_*]$, then*

$$W_2^2(m(t), m(s)) \leq R^2(t - s)^2.$$

Proof. The conclusion of the lemma follows directly from definition (2.1) and the estimate $\|f(t, x, u, v)\| \leq R$. \square

We obtain the properties of the approximating system using the nonlinear Markov processes. To define this object, it is convenient to introduce the generators (see [16], [20], [21]). The Markov chain with Kolmogorov matrix $\mathcal{Q}(t, \mu, \gamma_S, \vartheta_S)$ corresponds to the generator $\Lambda_t^{\mathcal{Q}}[\mu, \gamma_S, \vartheta_S]$ acting on $C(\mathcal{S})$ by the rule: if $\phi \in C(\mathcal{S})$, $\bar{x} \in \mathcal{S}$, then

$$\Lambda_t^{\mathcal{Q}}[\mu, \gamma_S, \vartheta_S]\phi(\bar{x}) \triangleq \sum_{\bar{y} \in \mathcal{S}} [\phi(\bar{y}) - \phi(\bar{x})] \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, \gamma_{\bar{x}}, \vartheta_{\bar{x}}). \quad (24)$$

Definition 6.2. Given an initial distribution $\nu_* = (\nu_{*, \bar{x}})_{\bar{x} \in \mathcal{S}}$, a function $[s, r] \ni t \mapsto \mu(t) \in \Sigma$, players' strategies $[s, r] \ni t \mapsto \gamma_S(t) \in \mathcal{U}_{\text{instant}}^{\mathcal{S}}$ and $[s, r] \ni t \mapsto \vartheta_S(t) \in \mathcal{V}_{\text{instant}}^{\mathcal{S}}$, we say that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [s, r]}, P, X)$ is a Markov chain generated by the Kolmogorov matrix $\mathcal{Q}(t, \mu(t), \gamma_S(t), \vartheta_S(t))$ if

- $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [s, r]}, P)$ is a filtered probability space;
- $X(\cdot)$ is a $\{\mathcal{F}_t\}_{t \in [s, r]}$ -adapted stochastic taking values in \mathcal{S} ;
- for every function $\phi \in C(\mathcal{S})$,

$$\phi(X(t)) - \int_s^t \Lambda_{\tau}^{\mathcal{Q}}[\mu(\tau), \gamma_S(\tau), \vartheta_S(\tau)]\phi(X(\tau))d\tau \quad (25)$$

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

- $P(X(s) = \bar{x}) = \nu_{*, \bar{x}}$.

Hereinafter, we denote by \mathbb{E} the expectation according to the probability P .

Remark 6.3. From [21] it follows that there exists at least one Markov chain generated by the Kolmogorov matrix $\mathcal{Q}(t, \mu(t), \gamma_S(t), \vartheta_S(t))$. Furthermore, one can assume that $\Omega = D([s, r], \mathcal{S})$, \mathcal{F} is equal to the family of Borel sets on $D([s, r], \mathcal{S})$, whilst \mathcal{F}_t is the family of sets such that their natural projections on $D([s, t], \mathcal{S})$ are Borel, whereas their natural projections on $[t, s]$ are equal to $D([t, r], \mathcal{S})$. Here $D([s, r], \mathcal{S})$ stands the Skorokhod space of càdlàg functions on $[s, r]$ taking values in \mathcal{S} .

Notice that, if we denote

$$\nu_{\bar{x}}(t) \triangleq P(X(t) = \bar{x}),$$

then $\nu(\cdot)$ satisfies the forward Kolmogorov differential equation (13). In particular, if $\nu_* = \mu(s)$, then $\nu(\cdot) = \mu(\cdot)$.

Set

$$\begin{aligned} R_1 &\triangleq 1 + 2(1 + Rd), \\ \varsigma_1(\epsilon) &\triangleq \frac{4}{3}(1 + R)R_1\epsilon^{1/2}. \end{aligned}$$

Notice that $\varsigma_1(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$.

Lemma 6.4. *Let $s, r \in [0, T]$, $s < r$, $\gamma_{\mathcal{S}}(\cdot)$ and $\vartheta_{\mathcal{S}}(\cdot)$ be measurable functions taking values in $\mathcal{U}_{instant}^{\mathcal{S}}$ and $\mathcal{V}_{instant}^{\mathcal{S}}$ respectively, and let $\bar{z} \in \mathcal{S}$. If $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [s, r]}, P, X)$ is a Markov chain generated by the Kolmogorov matrix $\mathcal{Q}(t, \mu(t), \gamma_{\mathcal{S}}(t), \vartheta_{\mathcal{S}}(t))$ and initial distribution at time s equal to $\mathbb{1}_{\bar{z}}$, then*

$$\mathbb{E}\|X(t) - \bar{z}\|^2 \leq R_1^2(t - s),$$

and

$$\mathbb{E}\|X(t) - \bar{z}\|^2 \leq \epsilon^2(t - s) + \varsigma_1(t - s) \cdot (t - s).$$

Proof. Consider the function $q_{\bar{z}}(x) \triangleq \|x - \bar{z}\|^2$. Due to Definition 6.2, we have that

$$\mathbb{E}q_{\bar{z}}(X(t)) = \int_s^t \mathbb{E}\Lambda_{\tau}^{\mathcal{Q}}q_{\bar{z}}(X(\tau))d\tau \quad (26)$$

Further, using definition of the generator $\Lambda^{\mathcal{Q}}$ (see (24)) and equality (17), we obtain the equality:

$$\begin{aligned} &\mathbb{E}\Lambda_{\tau}^{\mathcal{Q}}q_{\bar{z}}(X(\tau)) \\ &= \sum_{\bar{x} \in \mathcal{S}} \nu_{*, \bar{x}}(\tau) \sum_{\bar{y} \in \mathcal{S}, \bar{y} \neq \bar{x}} (\|\bar{y} - \bar{z}\|^2 - \|\bar{x} - \bar{z}\|^2) \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), \gamma_{\bar{x}}(\tau), \vartheta_{\bar{x}}(\tau)) \\ &= \sum_{\bar{x}} \nu_{\bar{x}}(\tau) \sum_{\bar{y} \in \mathcal{S}} (\|\bar{y} - \bar{x}\|^2 - 2\langle \ell(\bar{y}, \bar{x}), \ell(\bar{z}, \bar{x}) \rangle) \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), \gamma_{\bar{x}}(\tau), \vartheta_{\bar{x}}(\tau)). \end{aligned} \quad (27)$$

Using (20), we conclude that

$$\sum_{\bar{y} \in \mathcal{S}} \|\bar{y} - \bar{x}\|^2 \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), \gamma_{\bar{x}}(\tau), \vartheta_{\bar{x}}) \leq \epsilon^2. \quad (28)$$

To evaluate the last term, recall that

$$\mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), \gamma_{\bar{x}}(\tau), \vartheta_{\bar{x}}(\tau)) = \int_U \int_V \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), u, v) \gamma_{\bar{x}}(\tau, du) \vartheta_{\bar{x}}(\tau, dv). \quad (29)$$

Further, notice that

$$\begin{aligned} &-\sum_{\bar{y} \in \mathcal{S}} \langle \ell(\bar{y}, \bar{x}), \ell(\bar{z}, \bar{x}) \rangle \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), u, v) \\ &= -\left\langle \sum_{\bar{y} \in \mathcal{S}} \ell(\bar{y}, \bar{x}) \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), u, v) - f(\tau, \bar{x}, u, v), \ell(\bar{z}, \bar{x}) \right\rangle - \langle f(\tau, \bar{x}, u, v), \ell(\bar{z}, \bar{x}) \rangle. \end{aligned}$$

Thanks to (19), we obtain that

$$\left| \left\langle \sum_{\bar{y} \in \mathcal{S}} \ell(\bar{y}, \bar{x}) Q_{\bar{x}, \bar{y}}(\tau, \mu(\tau), u, v) - f(\tau, \bar{x}, u, v), \ell(\bar{z}, \bar{x}) \right\rangle \right| \leq \varepsilon |\ell(\bar{z}, \bar{x})|,$$

while (see (3))

$$|\langle f(\tau, \bar{x}, u, v), \ell(\bar{z}, \bar{x}) \rangle| \leq R |\ell(\bar{z}, \bar{x})|.$$

This, (27), (28), (29) and the inequality $|\ell(\bar{z}, \bar{x})| \leq \sqrt{q_{\bar{z}}(\bar{x})}$ give the following estimate:

$$\mathbb{E} \Lambda_{\tau}^{\mathcal{Q}} q_{\bar{z}}(X(\tau)) \leq \varepsilon^2 + 2(\varepsilon + R) [\mathbb{E} q_{\bar{z}}(X(\tau))]^{1/2}. \quad (30)$$

Taking into account the assumption that $\varepsilon \leq 1$, the inequality $0 \leq q_{\bar{z}}(x) \leq d$ for every $x \in \mathbb{T}^d$ and plugging (30) into (26), we conclude that

$$\mathbb{E} q_{\bar{z}}(X(t)) \leq R_1^2 (t - s).$$

Further, estimating the right-hand side of (30) using this inequality, we deduce that

$$\mathbb{E} \Lambda_{\tau}^{\mathcal{Q}} q_{\bar{z}}(X(\tau)) \leq \varepsilon^2 + 2(\varepsilon + R) R_1 \sqrt{\tau - s}.$$

Plugging this into (26), we obtain that

$$\mathbb{E} q_{\bar{z}}(X(t)) \leq \varepsilon^2 (t - s) + \frac{4}{3} (1 + R) R_1 (t - s)^{3/2}.$$

This gives the conclusion of the Lemma. □

Corollary 6.5. *Let $\mu(\cdot)$ satisfy (14). Then*

$$W_2^2(\widetilde{\mu}(t), \widetilde{\mu}(s)) \leq R_1^2 (t - s).$$

Proof. First notice [21] that one can find a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [s, r]}, P)$ and stochastic processes $X^{\bar{z}}, \bar{z} \in \mathcal{S}$, X those are defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [s, r]}, P)$ such that

- $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [s, r]}, P, X^{\bar{z}})$, $\bar{z} \in \mathcal{S}$, and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [s, r]}, P, X)$ are generated by the Kolmogorov matrix $\mathcal{Q}(t, \mu(t), \gamma_{\mathcal{S}}(t), \vartheta_{\mathcal{S}}(t))$;
- $X^{\bar{z}}(s) = \bar{z}$, P -a.s.,
- $P(X(s) = \bar{x}) = \mu_{\bar{x}}(s)$.

This, in particular, means that $\text{Law}(X(t)) = \mu(t)$. Furthermore, for every $\phi \in C(\mathcal{S})$,

$$\mathbb{E} \phi(X(t)) = \sum_{\bar{z} \in \mathcal{S}} \mu_{\bar{z}}(s) \phi(X^{\bar{z}}(t)).$$

Thus, we have that

$$W_2^2(\widetilde{\mu}(t), \widetilde{\mu}(s)) \leq \sum_{\bar{z} \in \mathcal{S}} \mathbb{E} \|X^{\bar{z}}(t) - \bar{z}\|^2 \mu_{\bar{z}}(s).$$

From Lemma 6.4 we obtain that

$$W_2^2(\widetilde{\mu}(t), \widetilde{\mu}(s)) \leq R_1^2 (t - s).$$

□

7 Construction of suboptimal strategy

In this section, we define the strategy \mathfrak{u}^Δ desired in Theorem 4.1. The construction relies on the adaptation of Krasovskii-Subbotin extremal shift rule.

We will consider only the first statement of Theorem 4.1. The proof of the second statement can be obtained by interchanging of the players.

Let φ^+ be a supersolution of (3). There are several equivalent definitions of the supersolution. We will use one involving the control theory. Recall [32, §12] that a lower semicontinuous function φ^+ is a supersolution of (3) provided that

$$(S1) \quad \varphi^+(T, \mu) \geq \hat{g}(\mu) = g(\tilde{\mu});$$

$$(S2) \quad \text{for every } s, r \in [0, T], s < r, \mu_* \in \Sigma, \vartheta_S \in \mathcal{V}_{\text{instant}}^S, \text{ there exists a control } [s, r] \ni t \mapsto \gamma_S(t) \in \mathcal{U}_{\text{instant}}^S \text{ such that, if } \mu(\cdot) \text{ solves}$$

$$\frac{d}{dt}\mu(t) = \mu(t)\mathcal{Q}(t, \mu(t), \gamma_S(t), \vartheta_S), \quad \mu(s) = \mu_*, \quad (31)$$

then the following inequality holds:

$$\varphi^+(r, \mu(r)) \leq \varphi^+(s, \mu(s)). \quad (32)$$

Remark 7.1. Notice that the original definition [32, §12] requires the convexification of the right-hand side of (31) with respect to the control of the first player. However, we avoid it due to the fact that the function

$$\gamma_S \mapsto \mu\mathcal{Q}(t, \mu, \gamma_S, \vartheta_S) = \left(\int_U \int_V \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} \mathcal{Q}_{\bar{x}, \bar{y}}(t, \mu, u, v) \gamma_{\bar{x}}(du) \vartheta_{\bar{x}}(dv) \right)_{\bar{y} \in \mathcal{S}}$$

is convex w.r.t. γ_S .

Furthermore, let \hat{u} and \hat{v} be measurable functions defined on $[0, T] \times \mathbb{T}^d \times \mathbb{T}^d \times \mathcal{P}^2(\mathbb{T}^d)$ and taking values in U and V respectively satisfying the following condition: for $t \in [0, T]$, $m \in \mathcal{P}^2(\mathbb{T}^d)$, $x \in \mathbb{T}^d$, $z \in \mathbb{R}^d$,

$$\hat{u}(t, x, y, m) \in \underset{u \in U}{\text{Argmin}} \max_{v \in V} \langle \ell(x, y), f(t, x, m, u, v) \rangle, \quad (33)$$

$$\hat{v}(t, x, y, m) \in \underset{v \in V}{\text{Argmax}} \min_{u \in U} \langle \ell(x, y), f(t, x, m, u, v) \rangle. \quad (34)$$

Existence of these functions directly follows from the measurable maximum theorem [2, 18.19].

Let t_0 be an initial time, $\Delta = \{t_k\}_{k=0}^N$ be a partition of $[t_0, T]$. The strategy $\mathfrak{u}^\Delta = \{\mathfrak{u}_k^\Delta\}_{k=0}^{N-1}$ is defined based on the auxiliary motions $\eta(\cdot)$, $\mu(\cdot)$ those serve as guides for the original system governed by (1). We define the controls \mathfrak{u}_k^Δ , $k = 0, \dots, N-1$ and the motions $\eta(\cdot)$, $\mu(\cdot)$ on intervals $[t_{k-1}, t_k]$ inductively.

$$(u1) \quad \text{Put } \eta(t_0) = \mu(t_0) \triangleq \text{pr}_S(m_0).$$

(u2) If $\eta(t)$, $\mu(t)$, $t \in [t_0, t_k]$ are already defined, $m_0, \dots, m_k \in \mathcal{P}^2(\mathbb{T}^d)$, then let π_k be an optimal plan between m_k and $\widetilde{\eta}(t_k)$. Denote

$$\alpha_k \triangleq (\mathfrak{p}^1, \hat{u}(t_k, \cdot, \cdot, m_k))\# \pi_k, \quad (35)$$

$$\beta_k \triangleq (\mathfrak{p}^2, \hat{v}(t_k, \cdot, \cdot, m_k))\# \pi_k. \quad (36)$$

Put

$$\mathfrak{u}_k^\Delta[m_0, \dots, m_k] \triangleq \alpha_k.$$

Further, notice that β_k is supported on $\mathcal{S} \times V$ and $\mathfrak{p}^1 \# \beta_k = \widetilde{\eta}(t_k)$. Let $\beta_k(dv|\bar{x})$ be disintegration of β_k w.r.t. $\widetilde{\eta}(t_k)$. We regard each probability $\beta_k(dv|\bar{x})$ as a constant control of the second player in the finite state mean field type game applied at the state \bar{x} . Set

$$\vartheta_{k,\mathcal{S}} \triangleq (\vartheta_{k,\bar{x}})_{\bar{x} \in \mathcal{S}}, \text{ with } \vartheta_{k,\bar{x}}(dv) \triangleq \beta_k(dv|\bar{x}). \quad (37)$$

Now let $\mu(t)$ $t \in [t_k, t_{k+1}]$ be such that (31) and (32) holds for $s = t_k$, $r = t_{k+1}$, $\mu_* = \mu(t_k)$ and $\vartheta_{\mathcal{S}}(t) \equiv \vartheta_{k,\mathcal{S}}$, $t \in [t_k, t_{k+1}]$, and some control of the first player $[t_k, t_{k+1}] \ni t \mapsto \gamma_{k,\mathcal{S}}(t) \in \mathcal{U}_{\text{instant}}^{\mathcal{S}}$.

For $x_* \in \mathbb{T}^d$, $\bar{y}_* \in \mathcal{S}$, denote by $\vartheta_{k,\mathcal{S}}^{x_*,\bar{y}_*} = (\vartheta_{k,\bar{x}}^{x_*,\bar{y}_*})_{\bar{x} \in \mathcal{S}}$ the constant control of the second player with

$$\vartheta_{k,\bar{x}}^{x_*,\bar{y}_*} = \begin{cases} \vartheta_{k,\bar{x}}, & \bar{x} \neq \bar{y}_* \\ \delta_{\hat{v}(t_k, x_*, \bar{y}_*, m_k)}, & \bar{x} = \bar{y}_*. \end{cases} \quad (38)$$

Let $\eta_k^{\Delta, x_*, \bar{y}_*}$ solve the differential equation

$$\frac{d}{dt} \eta_k^{x_*, \bar{y}_*}(t) = \eta_k^{x_*, \bar{y}_*}(t) \mathcal{Q}(t, \mu(t), \gamma_{k,\mathcal{S}}(t), \vartheta_{k,\mathcal{S}}^{x_*, \bar{y}_*}), \quad \eta_k^{x_*, \bar{y}_*}(t_k) = \mathbb{1}_{\bar{y}_*}, \quad (39)$$

where $\mathbb{1}_{\bar{y}_*}$ stands for the distribution on $\{1, \dots, |\mathcal{S}|\}$ concentrated at \bar{y}_* (see (9)) Finally, put, for $t \in [t_k, t_{k+1}]$,

$$\bar{\eta}_k^{\bar{y}_*}(t) \triangleq \int_{\mathbb{T}^d} \eta_k^{x_*, \bar{y}_*}(t) \pi(dx_* | \bar{y}_*), \quad (40)$$

$$\eta(t) \triangleq \sum_{\bar{y}_* \in \mathcal{S}} \bar{\eta}_k^{\bar{y}_*}(t_k) \eta^{\bar{y}_*}(t). \quad (41)$$

In the next section, we show that the strategy \mathfrak{u}^Δ defined in this manner provides the estimates desired in Theorem 4.1. This is performed using a modification of the extremal shift rule.

8 Extremal shift rule

Let \mathfrak{u}^Δ be introduced by rules (u1), (u2) and let $m(\cdot) \in \mathcal{X}_1(t_0, m_0, \Delta, \mathfrak{u}^\Delta)$. This means that on each interval there exists a distribution of players' controls $\varkappa_k \in \mathcal{D}_1[\alpha_k]$ (here α_k is given by (35)) for $m_k = m(t_k)$ such that

$$m(t) = m(t, t_k, m(t_k), \varkappa_k), \quad t \in [t_k, t_{k+1}].$$

Recall that the design of strategy \mathfrak{u}^Δ involves auxiliary motions $\mu(\cdot)$ and $\eta(\cdot)$ of the Markov chain those satisfies the initial condition $\mu(t_0) = \eta(t_0) = \text{pr}_{\mathcal{S}}(m_0)$ and are defined on the intervals $[t_k, t_{k+1}]$ by (37)–(41) (see (u1), (u2)).

To simplify designation, put

$$u_k^\natural(x, y) \triangleq \hat{u}(t_k, x, y, m(t_k)), \quad (42)$$

$$v_k^\natural(x, y) \triangleq \hat{v}(t_k, x, y, m(t_k)). \quad (43)$$

Furthermore, for $\zeta \in \mathcal{V}$, set

$$X_k^{x_*, \bar{y}^*}(t, \zeta) \triangleq x(t, t_k, x_*, m(t), u_k^\natural(x_*, \bar{y}^*), \zeta).$$

Recall that the probability $\eta(t)$ for $t \in [t_k, t_{k+1}]$ is the averaging of the probabilities $\eta_k^{x_*, \bar{y}^*}(t)$ those are defined by (39). The squared 2-Wasserstein distance between $\delta_{X_k^{x_*, \bar{y}^*}(t_{k+1}, \zeta)}$ and $\widetilde{\eta_k^{x_*, \bar{y}^*}(t_{k+1})}$ is equal to

$$W_2^2(\delta_{X_k^{x_*, \bar{y}^*}(t_{k+1}, \zeta)}, \widetilde{\eta_k^{x_*, \bar{y}^*}(t_{k+1})}) = \sum_{\bar{x} \in \mathcal{S}} \|X_k^{x_*, \bar{y}^*}(t_k, \zeta) - \bar{x}\|^2 \eta_{k, \bar{x}}^{x_*, \bar{y}^*}(t_{k+1}),$$

when the squared distance between $m(t_{k+1})$ and $\widetilde{\eta(t_{k+1})}$ is estimated as follows:

$$\begin{aligned} & W_2(m(t_{k+1}), \widetilde{\eta(t_{k+1})}) \\ & \leq \int_{\mathbb{T}^d \times \mathcal{S}} \int_{\mathcal{V}} W_2^2(\delta_{X_k^{x_*, \bar{y}^*}(t_{k+1}, \zeta)}, \widetilde{\eta_k^{x_*, \bar{y}^*}(t_{k+1})}) \varkappa_k(d\zeta | x_*, u_k^\natural(x_*, y_*)) \pi_k(d(x_*, \bar{y}_*)) \\ & = \int_{\mathbb{T}^d \times \mathcal{S}} \int_{\mathcal{V}} \|X_k^{x_*, \bar{y}^*}(t_{k+1}, \zeta) - \bar{x}\|^2 \eta_{k, \bar{x}}^{x_*, \bar{y}^*}(t_{k+1}) \varkappa_k(d\zeta | x_*, u_k^\natural(x_*, y_*)) \pi_k(d(x_*, \bar{y}_*)), \end{aligned} \quad (44)$$

where π_k is the optimal plan between $m(t_k)$ and $\widetilde{\eta(t_k)}$

Below, we estimate $W_2^2(\delta_{X_k^{x_*, \bar{y}^*}(t_{k+1}, \zeta)}, \widetilde{\eta_k^{x_*, \bar{y}^*}(t_{k+1})})$ by some function of $\|x_* - \bar{y}_*\|^2$, $W_2^2(m(t_k), \widetilde{\eta(t_k)})$ and $W_2^2(m(t_k), \mu(t_k))$ (Lemma 8.1). Then we evaluate the distance between $\mu(t_{k+1})$ and $\eta(t_{k+1})$. Based on this we derive the estimate of squared Wasserstein distance between $m(t_{k+1})$ and $\widetilde{\eta(t_{k+1})}$ by $W_2^2(m(t_k), \widetilde{\eta(t_k)})$. This will give the proof of the main theorem.

To state the Lemma 8.1, denote

$$\begin{aligned} \varsigma_f(\theta) & \triangleq \sup\{\|f(t', x, m, u, v) - f(t'', x, m, u, v)\| : \\ & t', t'' \in [0, T], x \in \mathbb{T}^d, m \in \mathcal{P}^2(\mathbb{T}^d), u \in U, v \in V, |t' - t''| \leq \theta\}. \end{aligned}$$

Since f is continuous on the compact $[0, T] \times \mathbb{T}^d \times \mathcal{P}^2(\mathbb{T}^d) \times U \times V$, the function ς_f is such that $\varsigma_f(\theta) \rightarrow 0$ as $\theta \rightarrow 0$.

Lemma 8.1. *For any $\zeta \in \mathcal{V}$, the following estimate holds:*

$$\begin{aligned} W_2^2(\delta_{X_k^{x_*, \bar{y}^*}(t_{k+1})}, \widetilde{\eta_k^{x_*, \bar{y}^*}(t_{k+1})}) & \leq \|x_* - \bar{y}_*\|^2 (1 + (3L + 1)(t_{k+1} - t_k)) \\ & + LW_2^2(m(t_k), \widetilde{\mu(t_k)})(t_{k+1} - t_k) + 2\varepsilon^2(t_{k+1} - t_k) \\ & + \varsigma_2(t_{k+1} - t_k) \cdot (t_{k+1} - t_k), \end{aligned} \quad (45)$$

where $\varsigma_2(\cdot)$ is determined only by f , Q , \mathcal{S} and satisfies $\varsigma_2(\theta) \rightarrow 0$ as $\theta \rightarrow 0$.

Proof. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [t_k, t_{k+1}]}, Y, P)$ be a Markov chain generated by the Kolmogorov matrix $\mathcal{Q}(t, \mu(t), \gamma_{k,S}(t), \vartheta_{k,S}^{x_*, \bar{y}_*})$ and the initial distribution at the time s equal to $\mathbb{1}_{\bar{y}_*}$. Notice that $P(Y(t) = \bar{x}) = \eta_{k, \bar{x}}^{x_*, \bar{y}_*}(t)$.

We shall prove that

$$\begin{aligned} \mathbb{E} \|X_k^{x_*, \bar{y}_*}(t_{k+1}) - Y(t_{k+1})\|^2 &\leq \|x_* - \bar{y}_*\|^2 (1 + (3L + 1)(t_{k+1} - t_k)) \\ &\quad + LW_2^2(m(t_k), \widetilde{\mu}(t_k))(t_{k+1} - t_k) + 2\varepsilon^2(t_{k+1} - t_k) \\ &\quad + \varsigma_2(t_{k+1} - t_k) \cdot (t_{k+1} - t_k) \end{aligned} \quad (46)$$

for some function ς_2 such that $\varsigma_2(\theta) \rightarrow 0$ as $\theta \rightarrow 0$. This will imply (45).

For shortness, set

$$u_* \triangleq u_k^\natural(x_*, \bar{y}_*) = \hat{u}(t_k, x_*, \bar{y}_*, m(t_k)), \quad (47)$$

$$v_* \triangleq v_k^\natural(x_*, \bar{y}_*) = \hat{v}(t_k, x_*, \bar{y}_*, m(t_k)). \quad (48)$$

Recall that $X_k^{x_*, \bar{y}_*}(\cdot, \zeta)$ satisfies

$$\frac{d}{dt} X_k^{x_*, \bar{y}_*}(t) = \int_V f(t, X_k^{x_*, \bar{y}_*}(t), m(t), u_*, v) \zeta(dv|t), \quad x(t_k) = x_*.$$

If $x' \in \mathbb{R}^d$, then, with some abuse of notation, we write $f(t, x', m, u, v)$ instead $f(t, \{x'\} + \mathbb{Z}^d, m, u, v)$. Now, let $x'_* \in x_*$. Consider the solution of the following differential equation in \mathbb{R}^d :

$$\frac{d}{dt} x'(t) = \int_V f(t, x'(t), m(t), u_*, v) \zeta(dv|t), \quad x'(t_k) = x'_*. \quad (49)$$

Notice that $x'(t) \in X_k^{x_*, \bar{y}_*}(t)$.

For any $Y'_{t_k} \in Y(t_k)$, $Y'_{t_{k+1}} \in Y(t_{k+1})$, we have that

$$\begin{aligned} \|x(t_{k+1}) - Y(t_{k+1})\|^2 &\leq \|x'(t_{k+1}) - Y'_{t_{k+1}}\|^2 \\ &= \|(x'(t_{k+1}) - x'(t_k)) + (x'(t_k) - Y'_{t_k}) - (Y'_{t_{k+1}} - Y'_{t_k})\|^2 \\ &= \|x'(t_k) - Y'_{t_k}\|^2 + \|Y'_{t_{k+1}} - Y'_{t_k}\|^2 + \|x'(t_{k+1}) - x'(t_k)\|^2 \\ &\quad - 2\langle x'(t_{k+1}) - x'(t_k), Y'_{t_{k+1}} - Y'_{t_k} \rangle \\ &\quad + 2\langle x'(t_{k+1}) - x'(t_k), x'(t_k) - Y'_{t_k} \rangle - 2\langle Y'_{t_{k+1}} - Y'_{t_k}, x'(t_k) - Y'_{t_k} \rangle \\ &\leq \|x'(t_k) - Y'_{t_k}\|^2 + 2\|Y'_{t_{k+1}} - Y'_{t_k}\|^2 + 2\|x'(t_{k+1}) - x'(t_k)\|^2 \\ &\quad + 2\langle x'(t_{k+1}) - x'(t_k), x'(t_k) - Y'_{t_k} \rangle - 2\langle Y'_{t_{k+1}} - Y'_{t_k}, x'(t_k) - Y'_{t_k} \rangle. \end{aligned}$$

Now choose $Y'_{t_k} \in Y(t_k)$, $Y'_{t_{k+1}} \in Y(t_{k+1})$ such that

$$x'(t_k) - Y'_{t_k} = \ell(x_*, Y(t_k)), \quad Y'_{t_{k+1}} - Y'_{t_k} = \ell(Y(t_{k+1}), Y(t_k)).$$

Therefore,

$$\begin{aligned} \|x(t_{k+1}) - Y(t_{k+1})\|^2 &\leq \|x(t_k) - Y(t_k)\|^2 + 2\|Y(t_{k+1}) - Y(t_k)\|^2 + 2\|x'(t_{k+1}) - x'(t_k)\|^2 \\ &\quad + 2\langle x'(t_{k+1}) - x'(t_k), \ell(x(t_k), Y(t_k)) \rangle \\ &\quad - 2\langle \ell(Y(t_{k+1}), Y(t_k)), \ell(x(t_k), Y(t_k)) \rangle. \end{aligned} \quad (50)$$

We have that

$$\|x'(t_{k+1}) - x'(t_k)\|^2 \leq R(t-s)^2. \quad (51)$$

Lemma 6.4 implies that

$$\mathbb{E}\|Y(t_{k+1}) - Y(t_k)\|^2 \leq \varepsilon^2(t-s) + \varsigma_1(t-s) \cdot (t-s). \quad (52)$$

Now, let us consider the term $\langle x'(t_{k+1}) - x'(t_k), \ell(x(t_k), Y(t_k)) \rangle$. To evaluate it, recall that $x'(\cdot)$ satisfies differential equation (49). We have that

$$\begin{aligned} & \langle x'(t_{k+1}) - x'(t_k), \ell(x_*, Y(t_k)) \rangle \\ &= \left\langle \int_s^t \int_V f(\tau, x'(\tau), m(\tau), u_*, v) \zeta(dv|\tau) d\tau, \ell(x_*, Y(t_k)) \right\rangle. \end{aligned} \quad (53)$$

Further, the continuity of f and the Lipschitz continuity with respect to x and m yields

$$\begin{aligned} & \|f(\tau, x'(\tau), m(\tau), u_*, v) - f(t_k, x_*, m(t_k), u_*, v)\| \\ & \leq \varsigma_f(t_{k+1} - t_k) + L\|x'(\tau) - x'(t_k)\| + LW_2(m(\tau), m(t_k)), \end{aligned}$$

when $t \in [t_k, t_{k+1}]$. This and inequalities $\|x'(\tau) - x'(t_k)\| \leq R|\tau - t_k|$ (see (3)), $W_2(m(\tau), m(t_k)) \leq R|\tau - t_k|$ (see Lemma 6.1) imply

$$\|f(\tau, x'(\tau), m(\tau), u_*, v) - f(t_k, x_*, m(t_k), u_*, v)\| \leq \varsigma_f(\tau - t_k) + 2LR(\tau - t_k).$$

Plugging this estimate into (53), we get the following:

$$\begin{aligned} & \left| \langle x'(t_{k+1}) - x'(t_k), \ell(x_*, Y(t_k)) \rangle \right. \\ & \quad \left. - \left\langle \int_s^t \int_V f(t_k, x_*, m(t_k), u_*, v) \zeta(dv|\tau) d\tau, \ell(x_*, Y(t_k)) \right\rangle \right| \\ & \leq \sqrt{d}[\varsigma_f(t_{k+1} - t_k) + LR(t_k - t_{k+1})] \cdot (t_{k+1} - t_k). \end{aligned} \quad (54)$$

Next, we evaluate the term $\mathbb{E}\langle \ell(Y(t_{k+1}), Y(t_k)), \ell(x_*, Y(t_k)) \rangle$. Recall that $Y(t_k) = \bar{y}_*$ P -a.s. Furthermore, the function $t \mapsto \eta_{k,S}^{x_*, \bar{y}_*}(t) = (\eta_{k, \bar{x}}^{x_*, \bar{y}_*}(t))_{\bar{x} \in \mathcal{S}}$ is such that $\eta_{k, \bar{x}}^{x_*, \bar{y}_*}(t) \triangleq P(Y(\tau) = \bar{x})$ and satisfies ODE (13) with initial condition $\nu_* = \mathbb{1}_{\bar{y}_*}$. Thus, since $Q(t, \mu, u, v)$ is defined on the compact space, we have that

$$\|\eta_{k,S}^{x_*, \bar{y}_*}(t) - \mathbb{1}_{\bar{y}_*}\|_2 \leq C_3(t - t_k), \quad (55)$$

where C_3 is a constant dependent only on Q .

Using the fact that the process (25) is the $\{\mathcal{F}_t\}_{t \in [s, r]}$ -martingale for the function $\phi(x) \triangleq \langle \ell(x, \bar{y}_*), \ell(x_*, \bar{y}_*) \rangle$ when the controls $\gamma_{k,S}$ and $\vartheta_{k,S}^{x_*, \bar{y}_*}$ are used, we conclude that

$$\begin{aligned} & \mathbb{E}\langle \ell(Y(t_{k+1}), Y(t_k)), \ell(x_*, Y(t_k)) \rangle \\ &= \mathbb{E} \int_{t_k}^{t_{k+1}} \Lambda_\tau^Q[\mu(\tau), \gamma_{k,S}(\tau), \vartheta_{k,S}^{x_*, \bar{y}_*}] \left\langle \ell(\cdot, \bar{y}_*), \ell(x_*, \bar{y}_*) \right\rangle (Y(\tau)) d\tau. \end{aligned}$$

The definition of the generator $\Lambda_t^{\mathcal{Q}}$ (see (24)) implies that

$$\begin{aligned} & \Lambda_{\tau}^{\mathcal{Q}}[\mu(\tau), \gamma_{k,\mathcal{S}}(\tau), \vartheta_{k,\mathcal{S}}^{x_*, \bar{y}_*}] \langle \ell(\cdot, \bar{y}_*), \ell(x_*, \bar{y}_*) \rangle(\bar{x}) \\ &= \left\langle \sum_{\bar{y} \in \mathcal{S}} (\ell(\bar{y}, \bar{y}_*) - \ell(\bar{x}, \bar{y}_*)) \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), \gamma_{k,\mathcal{S}}(\tau), \vartheta_{k,\mathcal{S}}^{x_*, \bar{y}_*}), \ell(x_*, \bar{y}_*) \right\rangle \end{aligned}$$

Hence,

$$\begin{aligned} & \mathbb{E} \langle \ell(Y(t_{k+1}), Y(t_k)), \ell(x_*, Y(t_k)) \rangle \\ &= \int_{t_k}^{t_{k+1}} \left\langle \sum_{\bar{x} \in \mathcal{S}} \eta_{k,\bar{x}}^{x_*, \bar{y}_*}(\tau) \sum_{\bar{y} \in \mathcal{S}} (\ell(\bar{y}, \bar{y}_*) - \ell(\bar{x}, \bar{y}_*)) \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), \gamma_{k,\mathcal{S}}(\tau), \vartheta_{k,\mathcal{S}}^{x_*, \bar{y}_*}), \ell(x_*, \bar{y}_*) \right\rangle. \end{aligned} \quad (56)$$

Notice that, from (19) and the fact that $\vartheta_{k,\bar{y}_*}^{x_*, \bar{y}_*} = \delta_{v_*}$, it follows, that, if $\bar{x} = \bar{y}_*$,

$$\begin{aligned} & \left| \sum_{\bar{y} \in \mathcal{S}} (\ell(\bar{y}, \bar{y}_*) - \ell(\bar{x}, \bar{y}_*)) \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), \gamma_{k,\mathcal{S}}(\tau), \vartheta_{k,\mathcal{S}}^{x_*, \bar{y}_*}) \right. \\ & \quad \left. - \int_U f(\tau, \bar{y}_*, \widetilde{\mu}(\tau), u, v_*) \gamma_{k,\bar{y}_*}(\tau, du) \right| \leq \varepsilon. \end{aligned} \quad (57)$$

Furthermore, (55) gives that $|\eta_{k,\bar{y}_*}^{x_*, \bar{y}_*}(\tau) - 1| \leq C_3(\tau - s)$.

If $\bar{x} \neq \bar{y}_*$, then, as above, $\eta_{k,\bar{x}}^{x_*, \bar{y}_*}(\tau) \leq C_3(\tau - s)$. Additionally, $\|\ell(x, y)\| \leq \sqrt{d}$ on $\mathbb{T}^d \times \mathbb{T}^d$, when the functions $\mathcal{Q}_{\bar{x}, \bar{y}}$ are bounded. Thus,

$$\begin{aligned} & \left| \sum_{\bar{x} \in \mathcal{S}} \eta_{k,\bar{x}}^{x_*, \bar{y}_*}(\tau) \sum_{\bar{y} \in \mathcal{S}} (\ell(\bar{y}, \bar{y}_*) - \ell(\bar{x}, \bar{y}_*)) \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), \gamma_{k,\bar{x}}(\tau), \vartheta_{k,\bar{x}}^{x_*, \bar{y}_*}) \right. \\ & \quad \left. - \sum_{\bar{y} \in \mathcal{S}} \ell(\bar{y}, \bar{y}_*) \mathcal{Q}_{\bar{y}_*, \bar{y}}(\tau, \mu(\tau), \gamma_{k,\bar{y}_*}(\tau), \vartheta_{k,\bar{y}_*}^{x_*, \bar{y}_*}) \right| \\ & \leq |\eta_{k,\bar{y}_*}^{x_*, \bar{y}_*}(\tau) - 1| \sum_{\bar{y} \in \mathcal{S}} |\ell(\bar{y}, \bar{y}_*)| \cdot |\mathcal{Q}_{\bar{y}_*, \bar{y}}(\tau, \mu(\tau), \gamma_{k,\bar{y}_*}(\tau), \vartheta_{k,\bar{y}_*}^{x_*, \bar{y}_*})| \\ & \quad + \sum_{\bar{x} \neq \bar{y}_*} |\eta_{k,\bar{x}}^{x_*, \bar{y}_*}(\tau)| \sum_{\bar{y} \in \mathcal{S}} |\ell(\bar{y}, \bar{y}_*) - \ell(\bar{x}, \bar{y}_*)| \cdot |\mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), \gamma_{k,\bar{x}}(\tau), \vartheta_{k,\bar{x}}^{x_*, \bar{y}_*})|. \end{aligned}$$

The right-hand side of this inequality can be estimated by $C_4|\tau - s|$, where C_4 is a constant determined by the matrix Q . This and (57) yield that

$$\begin{aligned} & \left\langle \sum_{\bar{x} \in \mathcal{S}} \eta_{k,\bar{x}}^{x_*, \bar{y}_*}(\tau) \sum_{\bar{y} \in \mathcal{S}} (\ell(\bar{y}, \bar{y}_*) - \ell(\bar{x}, \bar{y}_*)) \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), \gamma_{k,\bar{x}}(\tau), \vartheta_{k,\bar{x}}^{x_*, \bar{y}_*}), \ell(x_*, \bar{y}_*) \right\rangle \\ & \quad - \left\langle \int_U f(\tau, \bar{y}_*, \widetilde{\mu}(\tau), u, v_*) \gamma_{k,\bar{y}_*}(\tau, du), \ell(x_*, \bar{y}_*) \right\rangle \leq \varepsilon |\ell(x_*, \bar{y}_*)| + C_4 \sqrt{d} (\tau - t_k). \end{aligned}$$

Recall that we denote the modulus of continuity of f w.r.t. time variable by ς_f . Furthermore, $W_2(\widetilde{\mu}(t_k), \widetilde{\mu}(\tau)) \leq R_1(\tau - t_k)^{1/2}$. Hence,

$$\begin{aligned} & \left\langle \sum_{\bar{x} \in \mathcal{S}} \eta_{k,\bar{x}}^{x_*, \bar{y}_*}(\tau) \sum_{\bar{y} \in \mathcal{S}} (\ell(\bar{y}, \bar{y}_*) - \ell(\bar{x}, \bar{y}_*)) \mathcal{Q}_{\bar{x}, \bar{y}}(\tau, \mu(\tau), \gamma_{k,\bar{x}}(\tau), \vartheta_{k,\bar{x}}^{x_*, \bar{y}_*}), \ell(x_*, \bar{y}_*) \right\rangle \\ & \quad - \left\langle \int_U f(t_k, \bar{y}_*, \widetilde{\mu}(t_k), u, v_*) \gamma_{k,\bar{y}_*}(\tau, du), \ell(x_*, \bar{y}_*) \right\rangle \leq \varepsilon |\ell(x_*, \bar{y}_*)| + \varsigma_2'(\tau - t_k). \end{aligned}$$

Here $\zeta'_2(\theta) \triangleq C_4\sqrt{d}\theta + \zeta_f(\theta) + LR_1\theta^{1/2}$ is such that $\zeta'_2(\theta) \rightarrow 0$ as $\theta \rightarrow 0$.

Combining the last estimate with (56), we arrive at the following estimate:

$$\begin{aligned} & \left| \mathbb{E} \langle \ell(Y(t_{k+1}), Y(t_k)), \ell(x_*, Y(t_k)) \rangle \right. \\ & \quad \left. - \int_{t_k}^{t_{k+1}} \left\langle \int_U f(t_k, \bar{y}_*, \widetilde{\mu}(t_k), u, v_*) \gamma_{k, \bar{y}_*}(\tau, du) d\tau, \ell(x_*, \bar{y}_*) \right\rangle \right| \\ & \leq \frac{\varepsilon^2}{2} (t_{k+1} - t_k) + \frac{\|x_* - \bar{y}_*\|^2}{2} (t_{k+1} - t_k) + \zeta'_2(t_{k+1} - t_k) (t_{k+1} - t_k). \end{aligned} \quad (58)$$

Now, turn back to inequality (50). Expectation of its right-hand side is evaluated in (51), (52), (54), (58). They imply the following estimate:

$$\begin{aligned} \mathbb{E} \|x(t_{k+1}) - Y(t_{k+1})\|^2 & \leq \|x_* - y_*\|^2 + 2\varepsilon^2(t_{k+1} - t_k) + \|x_* - \bar{y}_*\|^2(t_{k+1} - t_k) \\ & \quad + 2 \int_{t_k}^{t_{k+1}} \int_V \langle f(t_k, x_*, m(t_k), u_*, v), \ell(x_*, \bar{y}_*) \rangle \zeta(dv|\tau) d\tau \\ & \quad - 2 \int_{t_k}^{t_{k+1}} \int_U \langle f(t_k, \bar{y}_*, \widetilde{\mu}(t_k), u, v_*) \gamma_{k, \bar{y}_*}(\tau, du) d\tau \\ & \quad + \varsigma_2(t_{k+1} - t_k) \cdot (t_{k+1} - t_k), \end{aligned} \quad (59)$$

where

$$\varsigma_2(\theta) \triangleq R\theta + \varsigma_1(\theta) + 2[\zeta_f(\theta) + LR\theta] + 2\zeta'_2(\theta). \quad (60)$$

Since f is Lipschitz continuous w.r.t. to x and m we have that

$$\begin{aligned} & |\langle f(t_k, \bar{y}_*, \widetilde{\mu}(t_k), u, v_*), \ell(x_*, \bar{y}_*) \rangle - \langle f(t_k, x_*, m(t_k), u, v_*), \ell(x_*, \bar{y}_*) \rangle| \\ & \leq L\|x_* - \bar{y}_*\|^2 + LW_2(m(t_k), \widetilde{\mu}(t_k))\|x_* - \bar{y}_*\| \\ & \leq \frac{3}{2}L\|x_* - \bar{y}_*\|^2 + \frac{1}{2}LW_2^2(m(t_k), \widetilde{\mu}(t_k)). \end{aligned}$$

Using this, we can estimate the right-hand side of (59) as follows:

$$\begin{aligned} & \mathbb{E} \|x(t_{k+1}) - Y(t_{k+1})\|^2 \\ & \leq \|x_* - \bar{y}_*\|^2 + 2\varepsilon^2(t_{k+1} - t_k) + (1 + 3L)\|x_* - \bar{y}_*\|^2(t_{k+1} - t_k) \\ & \quad + LW_2^2(m(t_k), \widetilde{\mu}(t_k))(t_{k+1} - t_k) \\ & \quad + 2 \int_{t_k}^{t_{k+1}} \int_V \langle f(t_k, x_*, m(t_k), u_*, v), \ell(x_*, \bar{y}_*) \rangle \zeta(dv|\tau) d\tau \\ & \quad - 2 \int_{t_k}^{t_{k+1}} \int_U \langle f(t_k, x_*, m(t_k), u, v_*) \gamma_{k, \bar{y}_*}(\tau, du) d\tau \\ & \quad + \varsigma_2(t_{k+1} - t_k) \cdot (t_{k+1} - t_k), \end{aligned} \quad (61)$$

The choice of the controls u_* and v_* (see (33), (34), (47), (48)) gives that, for any $u \in U$, $v \in V$,

$$\langle f(t_k, x_*, m(t_k), u_*, v), \ell(x_*, \bar{y}_*) \rangle - \langle f(t_k, x_*, m(t_k), u, v_*), \ell(x_*, \bar{y}_*) \rangle \leq 0.$$

This and (61) imply (46). Simultaneously, very definition of the function ς_2 implies that $\varsigma_2(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$. □

The following lemma provides the estimate between $\mu(\cdot)$ and $\eta(\cdot)$ defined in (u2).

Lemma 8.2. *Let $\mu(\cdot)$ solve (31) on each interval $[t_k, t_{k+1}]$ with $\vartheta_{k,S}$ given by (37) and let $\eta(\cdot)$ be defined by (41). Then there exists a constant C_5 determined only by Q such that*

$$\|\mu(t) - \eta(t)\|_2 \leq C_5 d(\Delta).$$

Proof. First, consider $\nu(\cdot)$ satisfying ODE (13) on $[s, r]$. We have that

$$\nu(t) = \nu_* \exp \left\{ \int_s^t \mathcal{Q}(\tau, \mu(\tau), \gamma_S(\tau), \vartheta_S(\tau)) d\tau \right\}$$

Expanding the exponent and using the fact that the matrices $Q(t, \mu, u, v)$ (and, thus, $\mathcal{Q}(t, \mu, \gamma_S, \vartheta_S)$) are uniformly bounded, we conclude that

$$\left\| \nu(r) - \nu(s) - \nu(s) \int_s^r \mathcal{Q}(\tau, \mu(\tau), \gamma_S(\tau), \vartheta_S(\tau)) d\tau \right\|_2 \leq C'_5 (r - s)^2. \quad (62)$$

Here C'_5 is a constant (certainly, dependent on Q).

Now let $x_* \in \mathbb{T}^d$, whereas $\bar{y}_* \in \mathcal{S}$. Recall that $\eta_k^{x_*, \bar{y}_*}$ solves (39). Estimate (62) implies that

$$\left\| \eta_k^{x_*, \bar{y}_*}(t_{k+1}) - \mathbb{1}_{\bar{y}_*} - \mathbb{1}_{\bar{y}_*} \int_{t_k}^{t_{k+1}} \mathcal{Q}(\tau, \mu(\tau), \gamma_{k,S}(\tau), \vartheta_{k,S}^{x_*, \bar{y}_*}) d\tau \right\|_2 \leq C'_5 (t_{k+1} - t_k)^2.$$

Notice that the very definitions of $\vartheta_{k,S}$ and $\vartheta_{k,S}^{x_*, \bar{y}_*}$ (see (37) and (38)) yield that, for every $t \in [0, T]$, $\mu' \in \Sigma$, $\gamma_S \in \mathcal{U}_{\text{instant}}^S$, $\bar{y}_* \in \mathcal{S}$,

$$\mathcal{Q}(t, \mu', \gamma_S, \vartheta_{k,S}) = \int_{\mathbb{T}^d} \mathcal{Q}(t, \mu', \gamma_S, \vartheta_{k,S}^{x_*, \bar{y}_*}) \pi(dx_* | \bar{y}_*).$$

Hence,

$$\left\| \eta_k^{\bar{y}_*}(t_{k+1}) - \mathbb{1}_{\bar{y}_*} - \mathbb{1}_{\bar{y}_*} \int_{t_k}^{t_{k+1}} \mathcal{Q}(\tau, \mu(\tau), \gamma_{k,S}(\tau), \vartheta_{k,S}) d\tau \right\|_2 \leq C'_5 (t_{k+1} - t_k)^2.$$

Here $\eta_k^{\bar{y}_*}(\cdot)$ is defined by (40). Now, let $\mu_k^{\bar{y}_*}(\cdot)$ solve (13) on $[t_k, t_{k+1}]$ with $\mu_k^{\bar{y}_*}(t_k) = \mathbb{1}_{\bar{y}_*}$ and controls $\gamma_{k,S}$ and $\vartheta_{k,S}$. Due to (62), we have that

$$\left\| \mu_k^{\bar{y}_*}(t_{k+1}) - \mathbb{1}_{\bar{y}_*} - \mathbb{1}_{\bar{y}_*} \int_{t_k}^{t_{k+1}} \mathcal{Q}(\tau, \mu(\tau), \gamma_{k,S}(\tau), \vartheta_{k,S}) d\tau \right\|_2 \leq C'_5 (t_{k+1} - t_k)^2.$$

Therefore,

$$\|\eta_k^{\bar{y}_*}(t_{k+1}) - \mu_k^{\bar{y}_*}(t_{k+1})\|_2 \leq 2C'_5 (t_{k+1} - t_k)^2. \quad (63)$$

Now recall that

$$\eta(t) = \sum_{\bar{y}_* \in \mathcal{S}} \eta_{k, \bar{y}_*}(t_k) \eta_k^{\bar{y}_*}(t), \quad \mu(t) = \sum_{\bar{y}_* \in \mathcal{S}} \mu_{\bar{y}_*}(t_k) \mu_k^{\bar{y}_*}(t).$$

Thus,

$$\begin{aligned} & \|\eta(t_{k+1}) - \mu(t_{k+1})\|_2 \\ & \leq \left\| \sum_{\bar{y}_* \in \mathcal{S}} \eta_{k, \bar{y}_*}(t_k) \eta_k^{\bar{y}_*}(t_{k+1}) - \sum_{\bar{y}_* \in \mathcal{S}} \eta_{k, \bar{y}_*}(t_k) \mu_k^{\bar{y}_*}(t_{k+1}) \right\|_2 \\ & \quad + \left\| \sum_{\bar{y}_* \in \mathcal{S}} \eta_{k, \bar{y}_*}(t_k) \mu_k^{\bar{y}_*}(t_{k+1}) - \sum_{\bar{y}_* \in \mathcal{S}} \mu_{\bar{y}_*}(t_k) \mu_k^{\bar{y}_*}(t_{k+1}) \right\|_2 \end{aligned} \quad (64)$$

The first term in the right-hand side of this inequality can be estimated as follows:

$$\begin{aligned} & \left\| \sum_{\bar{y}_* \in \mathcal{S}} \eta_{k, \bar{y}_*}(t_k) \eta_k^{\bar{y}_*}(t_{k+1}) - \sum_{\bar{y}_* \in \mathcal{S}} \eta_{k, \bar{y}_*}(t_k) \mu_k^{\bar{y}_*}(t_{k+1}) \right\|_2 \\ & \leq \sum_{\bar{y}_* \in \mathcal{S}} \eta_{k, \bar{y}_*}(t_k) \|\eta_k^{\bar{y}_*}(t_{k+1}) - \mu_k^{\bar{y}_*}(t_{k+1})\|_2 \\ & \leq \sum_{\bar{y}_* \in \mathcal{S}} \eta_{k, \bar{y}_*}(t_k) \cdot 2C'_5(t_{k+1} - t_k)^2 = 2C''_5(t_{k+1} - t_k)^2. \end{aligned}$$

In the last inequality above we used estimate (63).

To evaluate the second term, notice that the functions

$$t \mapsto \sum_{\bar{y}_* \in \mathcal{S}} \eta_{k, \bar{y}_*}(t_k) \mu_k^{\bar{y}_*}(t) \quad \text{and} \quad t \mapsto \sum_{\bar{y}_* \in \mathcal{S}} \mu_{\bar{y}_*}(t_k) \mu_k^{\bar{y}_*}(t)$$

satisfy (13) with the initial conditions at the time t_k $\nu_* = \eta(t_k)$ and $\nu_* = \mu(t_k)$ respectively. Thus, there exists a constant C''_5 such that

$$\left\| \sum_{\bar{y}_* \in \mathcal{S}} \eta_{k, \bar{y}_*}(t_k) \mu_k^{\bar{y}_*}(t) - \sum_{\bar{y}_* \in \mathcal{S}} \mu_{\bar{y}_*}(t_k) \mu_k^{\bar{y}_*}(t) \right\|_2 \leq \|\eta(t_k) - \mu(t_k)\|_2 e^{C''_5(t_{k+1} - t_k)}.$$

As above, C''_5 is a constant determined only by Q . Using (64), we conclude that

$$\|\eta(t_{k+1}) - \mu(t_{k+1})\|_2 \leq 2C''_5(t_{k+1} - t_k)^2 + \|\eta(t_k) - \mu(t_k)\|_2 e^{C''_5(t - t_k)}.$$

From this and the fact that $\eta(t_0) = \mu(t_0)$ one can derive the estimate

$$\|\eta(t_k) - \mu(t_k)\|_2 \leq 2C''_5 e^{C''_5(t_k - t_0)} (t_k - t_0) d(\Delta).$$

This gives the conclusion of the lemma with $C_5 \triangleq 2C''_5 e^{C''_5 T} T$. \square

Lemma 8.3. *The following inequality holds for every $k = 0, \dots, N - 1$:*

$$\begin{aligned} W_2^2(m(t_{k+1}), \widetilde{\eta(t_{k+1})}) & \leq W_2^2(m(t_k), \widetilde{\eta(t_k)}) (1 + (4L + 1)(t_{k+1} - t_k)) \\ & \quad + 2\varepsilon^2(t_{k+1} - t_k) + \varsigma_3(d(\Delta)) \cdot (t_{k+1} - t_k). \end{aligned}$$

Here $\varsigma_3(\cdot)$ is determined only by f , Q , T and has the zero limit at 0.

Proof. Plugging the estimate proved in Lemma 8.1 into (44), we conclude that

$$\begin{aligned}
& W_2^2(m(t_{k+1}), \widetilde{\eta(t_{k+1})}) \\
& \leq (1 + (3L + 1)(t_{k+1} - t_k)) \int_{\mathbb{T}^d \times \mathcal{S}} \|x_* - \bar{y}_*\|^2 \pi_k(d(x_*, \bar{y}_*)) \\
& \quad + LW_2^2(m(t_k), \widetilde{\mu(t_k)})(t_{k+1} - t_k) + 2\varepsilon^2(t_{k+1} - t_k) + \varsigma_2(t_{k+1} - t_k) \cdot (t_{k+1} - t_k) \\
& = (1 + (3L + 1)(t_{k+1} - t_k)) W_2^2(m(t_k), \widetilde{\eta(t_k)}) \\
& \quad + LW_2^2(m(t_k), \widetilde{\mu(t_k)})(t_{k+1} - t_k) + 2\varepsilon^2(t_{k+1} - t_k) + \varsigma_2(t_{k+1} - t_k) \cdot (t_{k+1} - t_k).
\end{aligned}$$

Here we used the fact that π_k is an optimal plan between $m(t_k)$ and $\widetilde{\eta(t_k)}$ (see 60). Then, notice that $\varsigma_2(t_{k+1} - t_k) \leq \varsigma_2(d(\Delta))$. Further, due to Proposition 3.1 and Lemma 8.2 we have that

$$\begin{aligned}
& W_2^2(m(t_k), \widetilde{\mu(t_k)}) \\
& \leq W_2^2(m(t_k), \widetilde{\eta(t_k)}) + W_2^2(\widetilde{\mu(t_k)}, \widetilde{\eta(t_k)}) + 2W_2(m(t_k), \widetilde{\eta(t_k)}) \cdot W_2(\widetilde{\mu(t_k)}, \widetilde{\eta(t_k)}) \\
& \leq W_2^2(m(t_k), \widetilde{\eta(t_k)}) + C_2 C_5 d(\Delta) + 2\sqrt{d C_2 C_5} \cdot \sqrt{d(\Delta)}.
\end{aligned}$$

Hence we obtain the conclusion of the Lemma with

$$\varsigma_3(\theta) \triangleq \varsigma_2(\theta) + L \left(C_2 C_5 \cdot \theta + 2\sqrt{d C_2 C_5} \cdot \theta \right).$$

□

Proof of Theorem 4.1. Given $m(\cdot) \in \mathcal{X}_1(t_0, m_0, \Delta, \mathbf{u}^\Delta)$, we obtain the auxiliary motions of the Markov chain $\mu(\cdot)$ and $\eta(\cdot)$. Notice that $\mu(\cdot)$ satisfies

$$\varphi^+(t_{k+1}, \mu(t_{k+1})) \leq \varphi^+(t_k, \mu(t_k)).$$

Thus,

$$\varphi^+(T, \mu(T)) \leq \varphi^+(t_0, \mu(t_0)). \quad (65)$$

Applying Lemma 8.3 sequentially, we obtain that

$$W_2^2(m(T), \widetilde{\eta(T)}) \leq e^{(4L+1)(T-t_0)} [W_2^2(m(t_0), \widetilde{\eta(t_0)}) + 2\varepsilon^2(T-t_0) + \varsigma_3(d(\Delta)) \cdot (T-t_0)].$$

Recall that $\widetilde{\mu(t_0)} = \eta(t_0) = \text{pr}_{\mathcal{S}}(m_0)$ (see (u1)). Therefore taking into account the fact that $\text{pr}_{\mathcal{S}}(m_0)$ is the nearest to m_0 element of $\mathcal{P}^2(\mathcal{S})$ and assumption (18), we have that

$$W_2^2(m(t_0), \widetilde{\eta(t_0)}) \leq \varepsilon^2.$$

Therefore,

$$\begin{aligned}
W_2(m(T), \widetilde{\eta(T)}) & \leq e^{(2L+1/2)T} [(1 + 2T)\varepsilon^2 + T\varsigma_3(d(\Delta))]^{1/2} \\
& \leq \sqrt{1 + 2T} e^{(2L+1/2)T} \varepsilon + \sqrt{T\varsigma_3(d(\Delta))}.
\end{aligned} \quad (66)$$

Lemma 8.2 and Proposition 3.1 give that

$$W_2(\widetilde{\eta}(T), \widetilde{\mu}(T)) \leq \sqrt{C_2 C_5 d(\Delta)}.$$

Combining this with (66), we arrive at the estimate

$$W_2(m(T), \widetilde{\mu}(T)) \leq C^* \varepsilon + \varsigma^*(d(\Delta)). \quad (67)$$

Here, we use the designation introduced in Section 4:

$$C^* \triangleq \sqrt{1 + 2T} e^{(2L+1/2)T},$$

and additionally put

$$\varsigma^*(\theta) \triangleq \sqrt{T \varsigma_3(\theta)} + \sqrt{C_2 C_5 \theta}.$$

Notice that $\varsigma^*(\theta) \rightarrow 0$ as $\theta \rightarrow 0$. Recall that ς_g is a modulus of continuity of the function g . Estimate (67) implies that

$$g(m(T)) \leq g(\widetilde{\mu}(T)) + \varsigma_g(C^* \varepsilon + \varsigma^*(d(\Delta))).$$

To complete the proof of the first part of the theorem it suffices to apply inequality (65) and condition (S1) that states that

$$g(\widetilde{\mu}(T)) \leq \varphi^+(T, \mu(T)).$$

As it was mentioned above, the second part of Theorem 4.1 can be proved by interchanging of the players. □

Appendix. A property of Wasserstein metric on a simplex

The purpose of the Appendix is to prove the inequalities between the Wasserstein metric on $\mathcal{P}(\mathcal{S})$ and the p -th metric on Σ formulated in Proposition 3.1.

Proof of Proposition 3.1. Since the set \mathcal{S} is finite we have that, for $\pi \in \Pi(\widetilde{\mu}^1, \widetilde{\mu}^2)$, there exist nonnegative numbers $b_{\bar{x}, \bar{y}}[\pi]$, $\bar{x}, \bar{y} \in \mathcal{S}$, such that

$$\pi = \sum_{\bar{x}, \bar{y} \in \mathcal{S}} b_{\bar{x}, \bar{y}}[\pi] \delta_{\bar{x}, \bar{y}}. \quad (68)$$

Now let us prove inequality (11). Let $\pi_0 \in \Pi(\widetilde{\mu}^1, \widetilde{\mu}^2)$ be such that

$$W_p^p(\widetilde{\mu}^1, \widetilde{\mu}^2) = \int_{\mathcal{S} \times \mathcal{S}} \|x - y\|^p \pi_0(d(x, y)).$$

From (68) it follows that

$$W_p^p(\tilde{\mu}^1, \tilde{\mu}^2) = \sum_{\bar{x}, \bar{y} \in \mathcal{S}} b_{\bar{x}, \bar{y}}[\pi_0] \|\bar{x} - \bar{y}\|^p.$$

Notice that

$$\sum_{\bar{y} \in \mathcal{S}, \bar{y} \neq \bar{x}} b_{\bar{x}, \bar{y}}[\pi_0] \geq (\mu_{\bar{x}}^2 - \mu_{\bar{x}}^1)^+,$$

where a^+ is equal to a provided that a is positive, and 0 otherwise. Using this estimate and the definition of the fineness of the lattice \mathcal{S} (see (8)), we conclude that

$$W_p^p(\tilde{\mu}^1, \tilde{\mu}^2) \geq \sum_{\bar{x} \in \mathcal{S}} \sum_{\bar{y} \in \mathcal{S}} b_{\bar{x}, \bar{y}}[\pi_0] (d(\mathcal{S}))^p \geq (d(\mathcal{S}))^p \sum_{\bar{x} \in \mathcal{S}} (\mu_{\bar{x}}^2 - \mu_{\bar{x}}^1)^+.$$

Since $\mu^1, \mu^2 \in \Sigma$, we have that

$$\sum_{\bar{x} \in \mathcal{S}} (\mu_{\bar{x}}^2 - \mu_{\bar{x}}^1)^+ = \frac{1}{2} \sum_{\bar{x} \in \mathcal{S}} |\mu_{\bar{x}}^2 - \mu_{\bar{x}}^1|. \quad (69)$$

Furthermore, for every $\bar{x} \in \mathcal{S}$, $|\mu_{\bar{x}}^2 - \mu_{\bar{x}}^1| \leq 1$. Thus, we have that

$$W_p^p(\tilde{\mu}^1, \tilde{\mu}^2) \geq \frac{(d(\mathcal{S}))^p}{2} \sum_{\bar{x} \in \mathcal{S}} |\mu_{\bar{x}}^2 - \mu_{\bar{x}}^1| \geq \frac{(d(\mathcal{S}))^p}{2} \sum_{\bar{x} \in \mathcal{S}} |\mu_{\bar{x}}^2 - \mu_{\bar{x}}^1|^p.$$

This proves (11).

To prove inequality (12), choose a probability $\pi' \in \Pi(\tilde{\mu}^1, \tilde{\mu}^2)$ such that

$$b_{\bar{x}, \bar{x}}[\pi'] \triangleq \mu_{\bar{x}}^1 \wedge \mu_{\bar{x}}^2.$$

Here $b_{\bar{x}, \bar{y}}[\pi']$ are given by representation (68). Therefore,

$$\sum_{\bar{y} \in \mathcal{S}, \bar{y} \neq \bar{x}} b_{\bar{x}, \bar{y}}[\pi'] = (\mu_{\bar{x}}^2 - \mu_{\bar{x}}^1)^+. \quad (70)$$

The inclusion $\mathcal{S} \subset \mathbb{T}^d$ yields

$$W_p^p(\tilde{\mu}^1, \tilde{\mu}^2) \leq \sum_{\bar{x}, \bar{y} \in \mathcal{S}} \|\bar{x} - \bar{y}\|^p b_{\bar{x}, \bar{y}}[\pi'] \leq d^{p/2} \sum_{\bar{x} \in \mathcal{S}} \sum_{\bar{y} \in \mathcal{S}, \bar{y} \neq \bar{x}} b_{\bar{x}, \bar{y}}[\pi'].$$

This and equalities (69), (70) imply that

$$W_p^p(\tilde{\mu}^1, \tilde{\mu}^2) \leq \frac{d^{p/2}}{2} \sum_{\bar{x} \in \mathcal{S}} |\mu_{\bar{x}}^2 - \mu_{\bar{x}}^1| = \frac{d^{p/2}}{2} \|\mu^2 - \mu^1\|_1.$$

The Holder's inequality gives that

$$W_p^p(\tilde{\mu}^1, \tilde{\mu}^2) \leq \frac{d^{p/2+p'}}{2} \|\mu^2 - \mu^1\|_p,$$

where p' is such that $1/p' + 1/p = 1$. This proves (12) \square

References

- [1] N. Ahmed and X. Ding. Controlled McKean-Vlasov equation. *Commun. Appl. Anal.*, 5:183–206, 2001.
- [2] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis: A Hitchhiker’s Guide*. Springer, Berlin, Heidelberg, 2006.
- [3] 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.
- [4] D. Andersson and B. Djehiche. A maximum principle for SDEs of mean-field type. *Appl. Math. Optim.*, 63(3):341–356, 2011.
- [5] Y. Averboukh. Krasovskii–Subbotin approach to mean field type differential games. *Dyn. Games Appl.*, 9:573–593, 2019.
- [6] Y. Averboukh. A stability property in mean field type differential games. Preprint at ArXiv:1903.11152, 2019.
- [7] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations*. Birkhäuser, Basel, 1996.
- [8] E. Bayraktar, A. Cosso, and H. Pham. Randomized dynamic programming principle and Feynman-Kac representation for optimal control of McKean-Vlasov dynamics. *Trans. Amer. Math. Soc.*, 370:2115–2160, 2018.
- [9] A. Bensoussan, J. Frehse, and P. Yam. *Mean field games and mean field type control theory*. Springer, New York, 2013.
- [10] R. Buckdahn, B. Djehiche, and J. Li. A general stochastic maximum principle for SDEs of mean-field type. *Appl. Math. Optim.*, 64(2):197–216, 2011.
- [11] R. Carmona and F. Delarue. Forward-backward stochastic differential equations and controlled McKean–Vlasov dynamics. *Ann. Probab.*, 43(5):2647–2700, 2015.
- [12] G. Cavagnari and A. Marigonda. Time-optimal control problem in the space of probability measures. In *Large-Scale Scientific Computing*, volume 9374 of *Lecture Notes in Computer Science*, pages 109–116, 2015.
- [13] G. Cavagnari, A. Marigonda, K. Nguyen, and F. Priuli. Generalized control systems in the space of probability measures. *Set-Valued Var. Anal.*, 26(3):663–691, 2018.
- [14] A. Cosso and P. H. Zero-sum stochastic differential games of generalized McKean-Vlasov type. *J. Math. Pures Appl.*, 129:180–212, 2019.
- [15] B. Djehiche and S. Hamadène. Optimal control and zero-sum stochastic differential game problems of mean-field type. *Appl. Math. Optim.*, 81:933–960, 2020.

- [16] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*. Springer, New York, 2006.
- [17] W. Gangbo, S. Mayorga, and A. Świąch. Finite dimensional approximations of Hamilton-Jacobi-bellman equations in spaces of probability measures. Preprint, 2020.
- [18] M. Huang, R. Malhamé, and P. Caines. Nash equilibria for large population linear stochastic systems with weakly coupled agents. In E. Boukas and M. R.P., editors, *Analysis, Control and Optimization of Complex Dynamic Systems*, pages 215–252. Springer, 2005.
- [19] C. Jimenez, A. Marigonda, and Q. M. Optimal control of multiagent systems in the wasserstein space. *Calc. Var. Partial Differ. Equ.*, 59:Article number: 58, 2020.
- [20] V. N. Kolokoltsov. *Nonlinear Markov process and kinetic equations*. Cambridge University Press, Cambridge, 2010.
- [21] V. N. Kolokoltsov. *Markov processes, semigroups and generators*, volume 38 of *De Gruyter Studies in Mathematics*. De Gruyter, 2011.
- [22] N. N. Krasovskii and A. I. Subbotin. *Game-theoretical control problems*. Springer, New York, 1988.
- [23] D. Lacker. Limit theory for controlled McKean–Vlasov dynamics. *SIAM J. Control Optim.*, 55(3):1641–1672, 2017.
- [24] 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.
- [25] 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.
- [26] M. Laurière and O. Pironneau. Dynamic programming for mean-field type control. *C. R. Math. Acad. Sci. Paris*, 352(9):707–713, 2014.
- [27] A. Marigonda and M. Quincampoix. Mayer control problem with probabilistic uncertainty on initial positions. *J. Differential Equations*, 264(5):3212–3252, 2018.
- [28] J. Moon and T. Başar. Zero-sum differential games on the Wasserstein space. Preprint at ArXiv:1912.06084, 2019.
- [29] H. Pham and X. Wei. Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics. *SIAM J. Control Optim.*, 55:1069–1101, 2017.
- [30] H. Pham and X. Wei. Bellman equation and viscosity solutions for mean-field stochastic control problem. *ESAIM Control Optim. Calc. Var.*, 24(1):437–461, 2018.
- [31] M. Sion. On general minimax theorems. *Pac. J. Math.*, 8:171–176, 1958.

- [32] A. I. Subbotin. *Generalized solutions of first-order PDEs. The dynamical perspective.* Birkhäuser, Boston, 1995.
- [33] A. Sznitman. *Topics in propagation of chaos*, volume 1464 of *Lecture Notes in Mathematics*, pages 165–251. Springer, Berlin/Heidelberg, 1991.