S©MR

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ
Siberian Electronic Mathematical Reports
http://semr.math.nsc.ru
ISSN 1813-3304

Vol. 20, No. 2, pp. 144–145 (2023) УДК 517.977.1
https://doi.org/10.33048/semi.2023.20.??? MSC
93A14,93B11,93C10,93E10

APPROXIMATION OF DETERMINISTIC MEAN FIELD TYPE CONTROL SYSTEMS

YU. V. AVERBOUKH

Communicated by P.P. Petrov

Abstract: The paper is concerned with the approximation of the deterministic mean field type control system by a mean field Markov chain. It turns out that the dynamics of the distribution in the approximating system is described by a system of ordinary differential equations. Given a strategy for the Markov chain, we explicitly construct a control in the deterministic mean field type control system. Our method is a realization of the model predictive approach. The converse construction is also presented. These results lead to an estimate of the Hausdorff distance between the bundles of motions in the deterministic mean field type control system and the mean field Markov chain. Especially, we pay the attention to the case when one can approximate the bundle of motions in the mean field type system by solutions of a finite systems of ODEs. algebra.

Keywords: mean field type control, bundle of motions, mean field Markov chain, model predictive control

---

Averboukh Yu. V., Approximation of deterministic mean field type control systems.

The work was performed as part of research conducted in the Ural Mathematical Center with the financial support of the Ministry of Science and Higher Education of the Russian Federation (Agreement number 075-02-2023-913).

Received January, 1, 2023, Published December, 31, 2023.

1 Introduction

The paper studies mean field type control systems. These systems describe an evolution of many identical agents who play cooperatively and interact via some external field. The mean field type control systems appear within the modeling of swarm of robots, pedestrian flows, etc [10, 16, 18–20].

The concept of mean field models comes back to the model of plasma that was proposed by Vlasov in 1938 [36,37] and was formalized within the theory of McKean-Vlasov equation [31,35]. We will focus on the deterministic (first-order) mean field type systems. This means that the dynamics of each agent is described by an ordinary differential equation on the Euclidean space with the right-hand side depending on his/her state, control and the distribution of all agents.

Notice that the settings of the control theory include the study of optimal control problems as well as the examination of the qualitative properties of the bundle of trajectories.

The mean field type optimal control theory started with paper [1]. Nowadays, for the second order mean field type control system, i.e., when the dynamics of each agent is determined by a stochastic differential equation, the necessary and optimality conditions are derived (see [9, 15, 17, 22, 30, 32, 33]). The case of first-order mean field type optimal control problems was studied in papers [6,24,34], where the variants of dynamic programming and Pontryagin maximum principle were obtained.

The qualitative theory for mean field type control systems studies the general properties of bundle of motions as well as viability theory issues. For the deterministic mean field type control systems, the existence theorem was proved under the general assumption on the dynamics [8, 11, 24], while the Filippov and relaxation theorem are derived under additional assumption of continuity of the vector field [12, 14]. The viability theorem was obtained in the terms of tangent cones (see [4, 13]) and in the terms of proximal normal cones [7].

The main object of the paper is an approximation of the bundle of motions of the first order mean field type control system. Our approach is based on so called Markov approximations. They appear when one replace the ODE determining the dynamics of each agent by a continuous-time mean field Markov chain. The latter can be regarded as a system of infinitely many similar agents with the dynamics determined by transition rates depending, in particular, on a current distribution of agents. In this case, the dynamics of the whole distribution of agents obeys a nonlinear Kolmogorov equation. If, additionally, one can assure that the agents in the original system do not leave a compact space, the phase space for the approximating Markov chain can be chosen to be finite. This leads to an approximation of the first-order mean field type control system by a finite system of ODE.

The approximation of the deterministic control system by a Markov chain was proposed in [3]. In that paper, the approximation of the value function of the zero-sum differential game was studied based on stochastic control with guide approach. This concept comes back to research of Krasovskii and Kotelnikova [26–29]. The Markov approximation technique was extended to mean field type differential games in [5]. There, based on a modification of the extremal shift rule for the Kantorovich space and control of the guide strategies, the value function of the mean field type differential game was approximated by a solution of a finite dimensional differential game.

In the paper, we focus on the approximation of the mean field type control system and implement the model predictive control approach (see [23] and reference therein). Notice that, for the original deterministic mean field type control system, we assume the open-loop strategies. They are distribution of pairs consisting of an initial state and a control. Simultaneously, the approximating mean field Markov chain implies the feedback strategies. The latters can be regarded as a sequence of open loop strategies those work at the appropriate state.

As it was mentioned above, we use the methodology of the model predictive approach. For the considered deterministic mean field type control system, this means that, given a feedback strategy in the Markov chain, an agent uses on a small time interval a control borrowed from a state in the approximating Markov chain. The weights of the controls are determined by an optimal plan between the distribution in the original and approximating system. To solve the converse problem which implies the design of a feedback strategy in the Markov chain based on a given distribution of controls in the mean field type system, one can on each small time interval integrate the controls according to the optimal plan between distributions.

The main results of the paper includes also an approximation rates of the aforementioned constructions. They depend only on the original system, distance between the original and approximating systems, fineness of partition and the maximal transition rate in the mean field Markov chain. Notice also that, if the fineness of the partition tends to zero, the limiting approximation rates are determined only by the original system and the distance between the original and approximating systems. This, in particular, provides the distance between the bundles of motions in the deterministic mean field type control system and a finite dimensional control system. The earlier results gave only one-side estimate.

The rest of the paper is organized as follows. In Section 2, we introduce the general notation. The assumption on the deterministic mean field type control system as well as the class of admissible controls are presented in Section 3. The approximating mean field Markov chain is introduced in Section 4. Here, we define the general system and show the way to construct an approximating Markov chain for the original deterministic mean field type system. Section 5 deals with the model predictive control for the deterministic system. In this case, we assume that a feedback strategy for the mean field Markov chain is given and estimate the approximation rate of the constructed motion in the deterministic mean field type system. The model predictive control for the mean field Markov chain is discussed in Section 6. Finally, we estimate the Hausdorff distance between the bundle of motions in the deterministic mean field type control system and the Markov chain (see Section 7).

2 General notation

- If n is a natural number, $X_1, \ldots, X_n$ are sets, $i_1, \ldots, i_k$ are indices from $\{1, \ldots, n\}$, then $\mathrm{p}^{i_1,\ldots,i_k}$ is a projection from $X_1 \times \ldots X_n$ to $X_{i_1} \times \ldots X_{i_k}$, i.e,
$$p^{i_1,\ldots,i_k}(x_1, \ldots, x_n) \triangleq (x_{i_1}, \ldots, x_{i_k}).$$
- If (X, ρ_X), (Y, ρ_Y) are Polish spaces, then $C_b(X, Y)$ denotes the space of bounded continuous functions from X to Y. If Y is a normed vector space, then $C_b(X, Y)$ is also a normed vector space with usual sup-norm. Moreover, $C_b(X) \triangleq C_b(X, \mathbb{R})$.
- We always endow a Polish space (X, ρ_X) with the Borel σ-algebra denoted by $\mathcal{B}(X)$. Moreover, $\mathcal{M}(X)$ stands for the set of Borel nonnegative measures, whereas $\mathcal{P}(X)$ denotes the set of all Borel probabilities:
$$\mathcal{P}(X) \triangleq \{m \in \mathcal{M}(X) : m(X) = 1\}.$$
We consider on $\mathcal{M}(X)$ the topology of narrow convergence, i.e., a sequence of measures $\{m_n\}_{n=1}^{\infty}$ converges to a measure m if, for each $C_b(X)$,
$$\int_X \phi(x) m_n(dx) \to \int_X \phi(x) m(dx) \text{ as } n \to \infty.$$
Notice that $\mathcal{P}(X)$ is closed w.r.t. the narrow convergence.
- If $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$ are measurable spaces, m is a measure on $\mathcal{F}$, whilst $h : \Omega \to \Omega'$ is $\mathcal{F}/\mathcal{F}'$-measurable, then we denote by $h\sharp m$ the push-forward measure on $\mathcal{F}'$ defined by the rule: for each $\Upsilon \in \mathcal{F}'$,
$$(h\sharp m)(\Upsilon) \triangleq m(h^{-1}(\Upsilon)).$$
- If (X, ρ_X) and (Y, ρ_Y) are two Polish space, m is a measure on X, then we denote by $\Lambda(X, m; Y)$ the set of measures on $X \times Y$ with marginal distribution on X equal to m, i.e.,
$$\Lambda(X, m; Y) \triangleq \big\{\alpha \in \mathcal{M}(X \times Y) : \mathrm{p}^1 \sharp \alpha = m\big\}.$$
- A function $\beta : X \to \mathcal{P}(X)$ is called weakly measurable if, for each $\phi \in C_b(Y)$, the function
$$x \mapsto \int_Y \phi(y) \beta(x, dy)$$
is measurable. Furthermore, each weakly measurable function β generates a measure $m \star \beta \in \Lambda(X, m; Y)$ by the rule: for each

$\phi \in C_b(X \times Y)$,
$$\int_{X\times X} \phi(x,y)(m \star \beta)(d(x,y)) \triangleq \int_X \int_Y \phi(x,y)\beta(x,dy)m(dx).$$

- If $\alpha \in \Lambda(X, m; Y)$, then there exists a weakly measurable function β such that $\alpha = m \star \beta$ [21, III-70]. These function is called a disintegration of the measure α and denoted by $x \mapsto \alpha(\cdot|x)$. The disintegration exists and is unique a.e.
- If (X, ρ_X) is a Polish space, $p \geq 1$, then $\mathcal{P}^p(X)$ is the space of probabilities on X such that, for some (equivalently, every) $x_* \in X$,
$$\int_X \rho_X^p(x, x_*)m(dx) < \infty.$$
If X is a normed vector space, then we will always choose $x_* = 0$ and put
$$\varsigma_p(m) \triangleq \left[\int_X \|x\|^p m(dx)\right]^{1/p}.$$
- The space $\mathcal{P}^p(X)$ is endowed with the Kantorovich (also known as a Wasserstein) metric defined by the rule: if $m_1, m_2 \in \mathcal{P}^p(X)$, then
$$W_p(m_1,m_2) \triangleq \left[\inf\left\{\int_{X\times X} \rho_X^p(x_1,x_2)\pi(d(x_1,x_2)) : \pi \in \Pi(m_1,m_2)\right\}\right]^{1/p}.$$
Hereinafter, $\Pi(m_1, m_2)$ stands for the set of probabilities π on $X_1 \times X_2$ such that $\mathrm{p}^i \sharp\pi = m_i$, $i = 1, 2$. The convergence within W_p implies the narrow convergence, the converse holds true only if X is compact [2, Proposition 7.1.5]. Primary, we will consider the case $p = 2$.
- The set of all continuous curves in $\mathbb{R}^d$ on $[s, r]$ is denoted by $\Gamma_{s,r} \triangleq C([s, r], \mathbb{R}^d)$. If $t \in [s, r]$, then we denote by e_t the evaluation operator defined by the rule: for $x(\cdot) \in \Gamma_{s,r}$
$$e_t(x(\cdot)) \triangleq x(t).$$

3 First-order mean field type control system

In the paper, we consider a mean field type control system formally described by the continuity equation
$$\partial_t m(t) + \mathrm{div}(f(t, x, m(t), u(t, x))m(t)) = 0. \tag{1}$$
Here, $t \in [0, T]$, $x \in \mathbb{R}^d$ is a phase variable, $m(t)$ stands for a current distribution of agents, $u(t, x)$ is a control implemented by an agent at the time t and the state x. We will assume that the control is chosen from a set U. Notice that in (1) we indicate the dependence of the control u on t and x. However, in fact we will use distributions of open-loop strategies

(see Definition 1). The equivalence between the feedback controls and the distributions of open-loop contols is proved in [24, Theorem 1]. System (1) describes the behavior of the infinitely many identical agents such that the dynamics of each agent is determined by the ODE:

$$\frac{d}{dt}x(t) = f(t, x(t), m(t), u(t, x(t))). \tag{2}$$

Notice also that due to the equivalence of open-loop and feedback control for control systems, one can replace in (2) the feedback control $u(t, x(t))$ with $u(t)$. We impose the following conditions on U and f.

(C1) U is a metric compact;
(C2) f is a continuous function;
(C3) there exists a set $\mathcal{K} \subset \mathbb{R}^d$ such that, if $\operatorname{supp}(m) \subset \mathcal{K}$, then

$$f(t, x, m, u) = 0, \quad x \notin \mathcal{K}; \tag{3}$$

(C4) f is bounded on $[0, T] \times \mathcal{K} \times \mathcal{P}^2(\mathcal{K}) \times U$ by a constant R;
(C5) f is Lipschitz continuous w.r.t. the space and measure variables on $[0, T] \times \mathcal{K} \times \mathcal{P}^2(\mathcal{K}) \times U$; the Lipschitz constant is denoted by C_f.

It is convenient to use the relaxation of the controls. This means that we replace the set of instantaneous controls U with the set of probabilities on U. The time-dependent relaxed controls are defined as follows.

For $s, r \in [0, T]$, $s < r$, put

$$\mathcal{U}_{s,r} \triangleq \Lambda([s, r], \lambda; U),$$

where λ is the Lebesgue measure on $[s, r]$. An element of $\mathcal{U}_{s,r}$ is a control measure on $[s, r]$. Notice that $\mathcal{U}_{s,r}$ is a compact subset of $\mathcal{M}([s, r] \times U)$. For now, assume that we are given with a flow of probabilities $m(\cdot) : [s, r] \to \mathcal{P}^p(\mathbb{R}^d)$. Furthermore, let $y \in \mathbb{R}^d$ be an initial state, and let $\xi \in \mathcal{U}_{s,r}$ be a relaxed control. Then, the corresponding motion of an agent is a solution of the initial value problem

$$\frac{d}{dt}x(t) = \int_U f(t, x(t), m(t), u)\xi(du|t), \quad x(s) = y. \tag{4}$$

We denote this motion by $x(\cdot, s, y, m(\cdot), \xi)$. Let us denote by $\operatorname{traj}^{s,r}_{m(\cdot)}$ the operator assigning to $y \in \mathbb{R}^d$ and $\xi \in \mathcal{U}_{s,r}$ the trajectory $x(\cdot, s, y, m(\cdot), \xi) \in \Gamma_{s,r}$.

Further, for $m \in \mathcal{P}^p(\mathbb{R}^d)$, put

$$\mathcal{A}_{s,r}[m] \triangleq \Lambda(\mathbb{R}^d, m; \mathcal{U}_{s,r}).$$

The set $\mathcal{A}_{s,r}[m]$ is the set of distributions of pairs consisting of an initial state and a relaxed control compatible with the initial probability m.

Definition 1. *Let $s, r \in [0, T]$, $s < r$, $m_* \in \mathcal{P}^p(\mathbb{R}^d)$, $\alpha \in \mathcal{A}[m_*]$. We say that a function $m(\cdot) : [s, r] \to \mathcal{P}^p(\mathbb{R}^d)$ is a motion of deterministic mean field type system (1) produced by the initial time s and the distribution of controls $\alpha \in \mathcal{A}_{s,r}[m_*]$ if there exists a measure $\chi \in \mathcal{P}^2(\Gamma_{s,r})$ such that*

- $m(s) = m_*$;
- $\chi = \mathrm{traj}^{s,r}_{m(\cdot)} \sharp\alpha$;
- $m(t) = e_t\sharp\chi$ *on* $[s,r]$.

Below we denote the motion of the system (1) produced by the initial time s and the distribution of controls $\alpha \in \mathcal{A}_{s,r}[m_*]$ by $m(\cdot, s, \alpha)$.

Proposition 1. *For each* $s, r \in [0,T]$, $r > s$, $m_* \in \mathcal{P}^2(\mathcal{K})$, $\mathcal{A}_{s,r}[m_*]$, *there exists a unique motion* $m(\cdot, s, \alpha)$. *Moreover,* $m(t,s,\alpha) \in \mathcal{P}^2(\mathcal{K})$ *for all* $t \in [s,r]$.

This proposition is proved in [8].

Further, we introduce the concatenation of distributions of controls. First, if $s_0, s_1, s_2 \in [0,T]$, $s_0 < s_1 < s_2$, $\xi_0 \in \mathcal{U}_{s_0,s_1}$, $\xi_1 \in \mathcal{U}_{s_1,s_2}$, then the concatenation $\xi \triangleq \xi_0 \diamond_{s_1} \xi_1$ of these controls is defined by its disintegration w.r.t. the Lebesgue measure:

$$\xi(\cdot|t) \triangleq \begin{cases} \xi_0(\cdot|t), & t \in [s_0, s_1), \\ \xi_1(\cdot|t), & t \in [s_1, s_2]. \end{cases}$$

Definition 2. *Let*

- $s_0, s_1, s_2 \in [0,T]$, $s_0 < s_1 < s_2$;
- $m_0, m_1 \in \mathcal{P}^2(\mathbb{R}^d)$;
- $\alpha_0 \in \mathcal{A}_{s_0,s_1}[m_0]$, $\alpha_1 \in \mathcal{A}_{s_1,s_2}[m_1]$.

be such that

$$m_1 = m(s_1, s_0, \alpha_0).$$

A probability $\alpha \in \mathcal{A}_{s_0,s_2}[m_0]$ *defined by the rule: for every* $\phi \in C_b(\mathbb{R}^d \times \mathcal{U}_{s_0,s_2})$,

$$\int_{\mathbb{R}^d \times \mathcal{U}_{s_0,s_2}} \phi(y,\xi)\alpha(d(y,\xi)) \triangleq \int_{\mathbb{R}^d \times \mathcal{U}_{s_0,s_1}} \int_{\mathcal{U}_{s_1,s_2}} \phi(y, \xi_0 \diamond_{s_1} \xi_1)\alpha_1(d\xi_1|x^0(y,\xi_0))\alpha_0(d(y,\xi_0))$$

is called a concatination of distributions α_0 *and* α_1. *Here we use the designations*

$$x^0(y,\xi_0) \triangleq x(s_1, s_0, y, m^0(\cdot), \xi_0), \quad m^0(\cdot) \triangleq m(\cdot, s_0, y, m_0, \alpha_0).$$

With some abuse of notation, we denote the concatenation of distributions by $\alpha_0 \diamond_{s_1} \alpha_1$.

Proposition 2. *Assume that* $s_0 < s_1 < s_2$, $m_0, m_1 \in \mathcal{P}^2(\mathcal{K})$, $\alpha_0 \in \mathcal{A}_{s_0,s_1}[m_0]$, $\alpha_1 \in \mathcal{A}_{s_1,s_2}[m_1]$ *are such that*

$$m_1 = m(s_1, s_0, \alpha_0).$$

Then,

- $m(t, s_0, \alpha_0 \diamond_{s_1} \alpha_1) = m(t, s_0, \alpha_0)$ *when* $t \in [s_0, s_1]$;
- $m(t, s_0, \alpha_0 \diamond_{s_1} \alpha_1) = m(t, s_1, \alpha_1)$ *when* $t \in [s_1, s_2]$.

This proposition directly follows from the definition of concatenation.

Let us complete this section with the equivalent formalization of the motion in mean field type control system (1). The following result is proved in [24, Theorem 1].

Proposition 3. *Let $m(\cdot) : [s,r] \to \mathcal{P}^2(\mathcal{K})$ be equal to $m(\cdot, s, \alpha)$ for some distribution of controls α. Then, there exists a velocity field $v : [s,r] \times \mathbb{R}^d \to \mathbb{R}^d$ such that*

(V1) $v(t,x) \in \operatorname{co}\{f(t,x,m(t),u) : u \in U\}$ for a.e. $t \in [s,r]$, $m(t)$-a.e. $x \in \mathbb{R}^d$;

(V2) the continuity equation

$$\partial_t m(t) + \operatorname{div}(v(t,x)m(t)) = 0$$

holds in the sense of distributions, i.e., for every $\phi \in C_c^1([s,r] \times \mathbb{R}^d)$,

$$\int_s^r \int_{\mathbb{R}^d} [\partial_t \phi(t,x) + \langle \nabla \phi(t,x), v(t,x)\rangle] m(t,dx)dt = 0.$$

Conversely, if $m(\cdot)$ and $v(\cdot,\cdot)$ satisfy conditions (V1), (V2), then there exits a distribution of controls α such that

$$m(\cdot) = m(\cdot, s, \alpha).$$

4 Mean filed Markov chains

In this section, we introduce a controlled nonlinear Markov chain. Let $\mathcal{S} \subset \mathcal{K}$ be at most countable set. Distributions on $\mathcal{S}$ are described by sequences $\mu = (\mu_{\bar{x}})_{\bar{x} \in \mathcal{S}}$ such that

$$\mu_{\bar{x}} \geq 0, \quad \sum_{\bar{x} \in \mathcal{S}} \mu_{\bar{x}} = 1.$$

The set of such distributions is the simplex on $\mathcal{S}$ denoted below by Σ. Furthermore, let Σ^2 be a set of sequences $\mu = (\mu_{\bar{x}})_{\bar{x} \in \mathcal{S}}$ such that

$$\sum_{\bar{x} \in \mathcal{S}} \|\bar{x}\|^2 \mu_{\bar{x}} < \infty.$$

If $\mathcal{S}$ is finite, the sets Σ and Σ^2 coincide. There is a natural isomorphism between Σ and $\mathcal{P}(\mathcal{S})$:

$$\mu = (\mu_{\bar{x}})_{\bar{x} \in \mathcal{S}} \mapsto \mathscr{I}(\mu) \triangleq \sum_{\bar{x} \in \mathcal{S}} \delta_{\bar{x}} \mu_{\bar{x}}.$$

Hereinafter, δ_z stands for the Dirac measure concentrated at z.

For $t \in [0,T]$, $\mu \in \Sigma^2$, $u \in U$, let $Q(t,\mu,u) = (Q_{\bar{x},\bar{y}}(t,\mu,u))_{\bar{x},\bar{y} \in \mathcal{S}}$ be a Kolmogorov matrix, i.e., $Q_{\bar{x},\bar{y}}(t,\mu,u) \geq 0$ when $\bar{x} \neq \bar{y}$ and, for each $\bar{x} \in \mathcal{S}$,

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

Given $\mu \in \Sigma^2$, $u \in U$, the quantity $Q_{\bar{x},\bar{y}}$ is a probability rate of the transfer from the state $\bar{x}$ to the state $\bar{y}$ at the time t in the case when the distribution of all agent is μ, while the implemented control is u.

Now let us introduce a mean field Markov chain generated by this Kolmogorov matrix. To this end, we first consider a relaxation of the control space. As above, a relaxed control on $[s,r]$ is an element of $\mathcal{U}_{s,r}$. We will use the feedback approach. This means that we are given with a sequence of relaxed control $\zeta_{\mathcal{S}} = (\zeta_{\bar{x}})_{\bar{x}\in\mathcal{S}}$. Notice that the set of feedback controls is $\mathcal{U}_{s,r}^{\mathcal{S}}$. In this case, the instantaneous probability rate for transition from $\bar{x}$ to $\bar{y}$ is equal to

$$\mathcal{Q}_{\bar{x},\bar{y}}(t,\mu,\zeta_{\mathcal{S}}) \triangleq \int_U Q_{\bar{x},\bar{y}}(t,\mu,u)\zeta_{\bar{x}}(du|t).$$

Moreover, we denote

$$\mathcal{Q}(t,\mu,\zeta_{\mathcal{S}}) \triangleq (\mathcal{Q}_{\bar{x},\bar{y}}(t,\mu,\zeta_{\mathcal{S}}))_{\bar{x},\bar{y}\in\mathcal{S}}.$$

Definition 3. *Given a time interval $[s,r]$, an initial distribution of states $\mu_* \in \Sigma^2$ and a feedback control $\zeta_{\mathcal{S}} = (\zeta_{\bar{x}})_{\bar{x}\in\mathcal{S}}$, we say that $\mu(\cdot)$ is a motion in the mean field Markov chain if it satisfies the following initial value problem:*

$$\frac{d}{dt}\mu_{\bar{y}}(t) = \sum_{\bar{x}\in\mathcal{S}} \mu_{\bar{x}}(t)\mathcal{Q}_{\bar{x},\bar{y}}(t,\mu(t),\zeta_{\mathcal{S}}), \quad \mu_{\bar{y}}(s) = \mu_{*,\bar{y}}. \tag{5}$$

Notice that system (5) can be rewritten in the vector form

$$\frac{d}{dt}\mu(t) = \mu(t)\mathcal{Q}(t,\mu(t),\zeta_{\mathcal{S}}), \quad \mu(s) = \mu_*. \tag{6}$$

To guarantee the existence of the distribution $\mu(\cdot)$, it is sufficient to assume that Q has continuous entries, for each $\bar{x}$ only finite number of entries $Q_{\bar{x},\bar{y}}(t,\mu,u)$ are non-zero and the dependence of the matrix Q on μ is Lipschitz continuous.

Let us also give a probabilistic interpretation of Definition 3. Set

- $\Omega_{s,r} \triangleq D([s,r];\mathcal{S})$, where $D([s,r];\mathcal{S})$ stands for a Skorokhod space of càdlàg functions;
- $\mathcal{F}_{s,r} \triangleq \mathcal{B}(D([s,r];\mathcal{S}))$;
- $\mathbb{F}_{s,r} = \{\mathcal{F}_{s,r}^t\}_{t\in[s,r]}$, where $\mathcal{F}_{s,r}^t \subset \mathcal{F}_{s,r}$ is a σ-algebra such that projections of its elements on $[s,t]$ form the whole σ-algebra $\mathcal{B}(D([s,t];\mathcal{S}))$, while the projection on $[t,r]$ is a trivial σ-algebra;
- $X(t,\omega) \triangleq \omega(t)$.

Further, we define the generator $L_t[\mu,\zeta_{\mathcal{S}}]$ by the rule: for $\phi \in C_b(\mathcal{S})$,

$$L_t[\mu,\zeta_{\mathcal{S}}]\phi(x) \triangleq \sum_{\bar{y}\in\mathcal{S}} \mathcal{Q}_{\bar{x},\bar{y}}(t,\mu,\zeta_{\mathcal{S}})\phi(\bar{y}).$$

Definition 4. *Let $[s,r]$ be a time interval, μ_* be an initial distribution of states, $\zeta_{\mathcal{S}} \in \mathcal{U}_{s,r}^{\mathcal{S}}$ and let $\mu(\cdot) = \mu(\cdot,s,\mu_*,\zeta_{\mathcal{S}})$. We say that a probability $\mathbb{P}_{s,r}$ on $\mathcal{F}_{s,r}$ realizes $\mu(\cdot)$ if*

- $\mu_{\bar{x}}(t) = \mathbb{P}_{s,r}\{\omega(t) = \bar{x}\}$;

- *for each* $\phi \in C_b(\mathcal{S})$, *the process*

$$\phi(X(t)) - \int_s^t L_t[\mu(t), \zeta_{\mathcal{S}}]\phi(X(t'))dt'$$

is a $\mathbb{F}_{s,r}$*-martingale.*

Below, if $\mathbb{P}_{s,r}$ is a realization of $\mu(\cdot)$, $\mathbb{E}_{s,r}$ stands for the corresponding expectation.

Notice that there exists at least one representation of flow $\mu(\cdot)$ [25, Theorem 5.4.2].

The main result of the paper is proved under the following approximation assumptions.

(A1)

$$\max_{x\in\mathcal{K}} \min_{\bar{y}\in\mathcal{S}} \|x - \bar{y}\| \leq \varepsilon;$$

(A2) entries of the matrix Q are uniformly bounded by a number B_Q.

(A3) for each $t \in [0, T]$, $\bar{x} \in \mathcal{S}$, $\mu \in \Sigma^2$ and $u \in U$,

$$\Big\| f(t, \bar{x}, \mathscr{I}(\mu), u) - \sum_{\bar{y}\in\mathcal{S}} (\bar{y} - \bar{x}) Q_{\bar{x},\bar{y}}(t, \mu, u) \Big\| \leq \varepsilon;$$

(A4) for each $t \in [0, T]$, $\bar{x} \in \mathcal{S}$, $\mu \in \Sigma^2$ and $u \in U$,

$$\sum_{\bar{y}\in\mathcal{S}} \|\bar{y} - \bar{x}\|^2 Q_{\bar{x},\bar{y}}(t, \mu, u) \leq \varepsilon^2.$$

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

An example of the Markov chain that satisfies assumptions (A1)–(A4) can be constructed on a regular lattice as follows. First, we represent f in the coordinate-wise form

$$f(t, x, m, u) = (f_i(t, x, m, u))_{i=1}^d.$$

Let $h > 0$, $\mathcal{K}^h \triangleq \mathcal{K} + \mathbb{B}_h$, where $\mathbb{B}_h$ is a ball centered at the origin and radius h, and let e_i denote the i-th coordinate vector. We put

$$\mathcal{S} \triangleq \mathcal{K}^h \cap h\mathbb{Z}, \tag{7}$$

$$Q_{\bar{x},\bar{y}}(t, \mu, u) \triangleq \begin{cases} h^{-1}|f_i(t, \bar{x}, \mathscr{I}(\mu), u)|, & \bar{y} = \bar{x} + h\,\mathrm{sgn}(f_i(t, \bar{x}, \mathscr{I}(\mu), u))e_i, \\ -h^{-1}\sum_{j=1}^d |f_j(t, \bar{x}, \mathscr{I}(\mu), u)|, & \bar{y} = \bar{x}, \\ 0, & \text{otherwise}. \end{cases} \tag{8}$$

One can directly show that this Markov chain satisfies conditions (A1)–(A4) with $B_Q = dRh^{-1}$ and $\varepsilon \triangleq \max\{h, \sqrt{hdR}\}$.

Notice that if $\mathcal{K}$ is compact, the phase space for Markov chain introduced by rules (7), (8) is finite and equation (6) is a system of the finite number of ODEs.

5 Model predictive control of the first-order mean field type system

In this section, we show that a feedback control in the mean filed type Markov chain can be used directly to construct a motion in the first order mean filed type control system.

Let $\mu_0 \in \Sigma^2$, $\zeta_{\mathcal{S}} \in \mathcal{U}^{\mathcal{S}}_{0,T}$. Notice that there exists a unique motion in the mean field Markov chain on the time interval $[0, T]$ produced by the control $\zeta_{\mathcal{S}}$ and the initial distribution μ_0. For shortness, we denote it by $\mu(\cdot)$. Let $m_0 \in \mathcal{P}^2(\mathcal{K})$, $\Delta = \{s_i\}_{i=0}^n$ be a partition of $[0, T]$. Further, let $\hat{\zeta}_i$ assign to $x \in \mathcal{K}$ and $\bar{x} \in \mathcal{S}$ a measure $\zeta_{\bar{x}} \in \mathcal{U}_{s_i, s_{i+1}}$. A model predictive strategy for the first order mean field game is constructed as follows.

(D1) Let π_0 be an optimal plan between m_0 and $\mathscr{I}(\mu_0)$. We define

$$\alpha_0 \triangleq (\mathrm{p}^1, \hat{\zeta}_0)\sharp\pi_0.$$

(D2) Assume now that we already construct controls α_i, $i = 0, \ldots, k-1$, and a flow of probabilities $m(\cdot)$ on $[0, s_k]$ such that

$$\alpha_i \in \mathcal{A}_{s_i, s_{i+1}}[m(s_i)], \quad i = 0, \ldots, k-1,$$

$$m(t) = m(t, 0, \alpha_0 \diamond_{s_1} \ldots \diamond_{s_{k-1}} \alpha_{k-1}), \quad t \in [0, s_k].$$

Set $m_k \triangleq m(s_k, 0, \alpha_0 \diamond_{s_1} \ldots \diamond_{s_{k-1}} \alpha_{k-1})$ and choose π_k to be an optimal plan between m_k and $\mathscr{I}(\mu(s_k))$. As above $\pi_k(\cdot|x)$ is a disintegration of this plan w.r.t. m_k. We put

$$\alpha_k \triangleq (\mathrm{p}^1, \hat{\zeta}_k)\sharp\pi_k$$

and, for $t \in [s_k, s_{k+1}]$,

$$m(t) \triangleq m(s_k, 0, \alpha_0 \diamond_{s_1} \ldots \diamond_{s_{k-1}} \alpha_{k-1} \diamond_{s_k} \alpha_k).$$

Theorem 1. *If $\Delta = \{s_i\}_{i=0}^n$ is a partition of $[0, T]$, with $d(\Delta) \leq 1$, while $\alpha_0, \ldots, \alpha_{n-1}$ are constructed by the rules (D1), (D2) and $m(\cdot) \triangleq m(\cdot, 0, \alpha_0 \diamond_{s_1} \ldots \diamond_{s_{n-1}} \alpha_{n-1})$, then*

$$W_p(m(t), \mathscr{I}(\mu(t))) \leq C_0 W_2((\mu_0), m_0) + C_1\varepsilon + C_2 d^{1/2}(\Delta) + C_3 d(\Delta) + C_4 \varepsilon B_Q d(\Delta) + C_5 B_Q d^2(\Delta).$$

Here, the constants $C_0, \ldots, C_5$ depend only on f and T.

Lemma 1. *Let $\mu(\cdot)$ be a distribution of agents in the mean field Markov chain, $\mathbb{P}_{s,r}$ be its realization. Then,*

$$\mathbb{E}_{s,r}(\|X(t) - X(s)\|^2 | X(s) = \bar{z}) \leq \varepsilon^2 (t - s) + C_1'(t - s)^{3/2},$$

where

$$C_1' \triangleq 4(R + 1)e^{2(R+1)T}/3.$$

Proof. For fixed $\bar{z} \in \mathcal{S}$, let us denote

$$q_{\bar{z}}(\bar{x}) \triangleq \|\bar{x} - \bar{z}\|^2.$$

We have that

$$\begin{aligned} L_t[\mu, u]q_{\bar{z}}(\bar{x}) &= \sum_{\bar{y}\in\mathcal{S}} Q_{\bar{x},\bar{y}}(t,\mu,u)\|\bar{y}-\bar{z}\|^2 \\ &= \sum_{\bar{y}\in\mathcal{S}} Q_{\bar{x},\bar{y}}(t,\mu,u)(\|\bar{y}-\bar{x}\|^2 + \|\bar{x}-\bar{z}\|^2 + 2\langle \bar{y}-\bar{x}, \bar{x}-\bar{z}\rangle) \\ &= \sum_{\bar{y}\in\mathcal{S}} Q_{\bar{x},\bar{y}}(t,\mu,u)\|\bar{y}-\bar{x}\|^2 + \Big\langle \sum_{\bar{y}\in\mathcal{S}} Q_{\bar{x},\bar{y}}(t,\mu,u)(\bar{y}-\bar{x}), \bar{x}-\bar{z}\Big\rangle. \end{aligned}$$

Due to assumption (A4), we have that the first term is bounded by ε^2. Moreover, condition (A3) implies that

$$\left\| \sum_{\bar{y}\in\mathcal{S}} Q_{\bar{x},\bar{y}}(t,\mu,u)(\bar{y}-\bar{x}) \right\| \leq R + \varepsilon.$$

Using these estimates and Definition 4, we conclude that that

$$\begin{aligned} &\mathbb{E}_{s,rT}(\|X(t) - X(s)\|^2 | X(s) = \bar{z}) \\ &\quad \leq \varepsilon^2(t-s) + 2\int_s^t (R+\varepsilon)(\mathbb{E}_{0,T}\|X(t') - X(s)\| | X(s) = \bar{z})dt'. \end{aligned} \tag{9}$$

Since we assume that $\varepsilon \leq 1$, we have that

$$\mathbb{E}_{0,T}(\|X(t) - X(s)\|^2 | X(s) = \bar{z}) \leq C_1''(t-s),$$

where

$$C_1'' \triangleq e^{2(R+1)T}.$$

Plugging this estimate to (9), we obtain the statement of the lemma. □

Lemma 2. *Assume that $s, r \in [0,T]$, $s < r$, $\nu_* = (\nu_{*,\bar{x}})_{\bar{x}\in\mathcal{S}} \in \Sigma^2$, $\zeta_{\mathcal{S}} \in \mathcal{U}_{s,r}^{\mathcal{S}}$, $\mu(\cdot) : [0,T] \to \Sigma^2$, while $\nu(\cdot)$ satisfies*

$$\frac{d}{dt}\nu(t) = \nu(t)Q(t,\mu(t),\zeta_{\mathcal{S}}(t)), \nu(s) = \nu_*. \tag{10}$$

Then, for each $\bar{x} \in \mathcal{S}$,

$$|\nu_{\bar{x}}(t) - \nu_{*,\bar{x}}| \leq B_Q(t-s).$$

Proof. We have that

$$|\nu_{\bar{x}}(t) - \nu_{*,\bar{x}}| \leq \int_s^t \int_U \sum_{\bar{y}\in\mathcal{S}} \nu_{\bar{y}}(t')\big|Q_{\bar{y},\bar{x}}(t',\mu(t),u)\big|\zeta_{\bar{y}}(du|t'))dt' \leq B_Q(t-s).$$

□

Proof of Theorem 1. Let us estimate the squared Kantorovich distance $W_2^2(\mu(t), m(t))$ for $t \in [s_k, s_{k+1}]$, $k = 0, \dots, n-1$. Let $x_k(\cdot, y, \bar{z})$ be a solution on $[s_k, s_{k+1}]$ of the differential equation

$$\frac{d}{dt}x(t) = \int_U f(t, x(t), m(t), u)\zeta_{\bar{z}}(du|t), \quad x(s_k) = y.$$

Notice that $m(t) = x_k(t, \cdot, \cdot)\sharp\pi_k$, while by construction $\mathscr{I}(\mu(t)) \triangleq X(t)\sharp\mathbb{P}_{s_k,s_{k+1}}$. Recall that $\mathbb{P}_{s_k,s_{k+1}}$ is a probability that realizes a flow of distributions $\mu(\cdot)$ on $[s_k, s_{k+1}]$, whereas $\mathbb{E}_{s_k,s_{k+1}}$ is the corresponding expectation. Thus, we have that

$$\begin{aligned} &W_2^2(\mathscr{I}(\mu(t)), m(t)) \\ &\qquad = \int_{\mathcal{K}\times\mathcal{S}} \mathbb{E}_{s_k,s_{k+1}}\big(\|X(t) - x_k(t, y, \bar{z})\|^2 | X(s_k) = \bar{z}\big)\pi_k(d(y, \bar{z})). \end{aligned} \tag{11}$$

Now, let us evaluate the quantity $\mathbb{E}_{s_k,s_{k+1}}(\|X(t) - x_k(t, y, \bar{z})\|^2 | X(s_k) = \bar{z})$. First, notice that

$$\begin{aligned} \|X(t) - x_k(t, y, \bar{z})\|^2 &\leq \|X(s_k) - x_k(s_k, y, \bar{z})\|^2 \\ &+ 2\|X(t) - X(s_k)\|^2 + 2\|x_k(t, y, \bar{z}) - x_k(s_k, y, \bar{z})\|^2 \\ &\quad + 2\langle X(t) - X(s_k), X(s_k) - x_k(s_k, y, \bar{z})\rangle \\ &\quad - 2\langle x_k(t, y, \bar{z}) - x_k(s_k, y, \bar{z}), X(s_k) - x_k(s_k, y, \bar{z})\rangle. \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{E}_{s_k,s_{k+1}}\big(\|X(t) - x_k(t, y, \bar{z})\|^2 | X(s_k) = \bar{z}\big) &\leq \|y - \bar{z}\|^2 \\ + 2\mathbb{E}_{s_k,s_{k+1}}\big(\|X(t) - \bar{z}\|^2 | X(s_k) = \bar{z}\big) &+ 2\|x_k(t, y, \bar{z}) - y\|^2 \\ + \big\langle \mathbb{E}_{s_k,s_{k+1}}(X(t) - \bar{z} | X(s_k) = \bar{z}), \bar{z} - y\big\rangle \\ - \langle x_k(t, y, \bar{z}) - y, \bar{z} - y\rangle. \end{aligned} \tag{12}$$

Due to Lemma 1, and the boundness of f, we have that

$$\begin{aligned} 2\mathbb{E}_{s_k,s_{k+1}}\big(\|X(t) - \bar{z}\|^2 | X(s_k) = \bar{z}\big) + 2\|x_k(t, y, \bar{z}) - y\|^2 \\ \leq 2\varepsilon^2(t - s) + 2C_1'(t - s)^{3/2} + R^2(t - s)^2. \end{aligned} \tag{13}$$

Further,

$$\begin{aligned} &\mathbb{E}_{s_k,s_{k+1}}(X(t) - \bar{z} | X(s_k) = \bar{z}) \\ &\qquad = \mathbb{E}_{s_k,s_{k+1}}\left(\int_{s_k}^t \sum_{\bar{y}\in\mathcal{S}} Q_{X(t'),y}(y - X(t')) | X(s) = \bar{z}\right)dt'. \end{aligned}$$

This and condition (A3) yield that

$$\Bigg\| \mathbb{E}_{s_k,s_{k+1}}(X(t)-\bar{z}|X(s_k)=\bar{z}) \\ - \int_{s_k}^{t} \mathbb{E}_{s_k,s_{k+1}}\left(\int_U f(t',X(t'),\mathscr{I}(\mu(t)),u)\zeta_{X(t')}(du|t')|X(s)=\bar{z}\right)dt' \Bigg\| \\ \leq \varepsilon(t-s_k).$$

Recall that

$$\mathbb{E}_{s_k,s_{k+1}}\left(\int_U f(t',X(t'),\mathscr{I}(\mu(t)),u)\zeta_{X(t')}(du|t')|X(s)=\bar{z}\right)dt' \\ = \sum_{\bar{x}\in\mathcal{S}} \int_U f(t',\bar{x},\mathscr{I}(\mu(t)),u)\zeta_{X(t')}(du|t')\nu_{*,\bar{x}}.$$

Thus, due to Lemma 2, we have that

$$\Bigg\| \mathbb{E}_{s_k,s_{k+1}}(X(t)-\bar{z}|X(s_k)=\bar{z}) - \int_{s_k}^{t}\int_U f(t',\bar{z},\mathscr{I}(\mu(t)),u)\zeta_{\bar{z}}(du|t')dt' \Bigg\| \\ \leq \varepsilon(t-s_k)+RB_Q(t-s_k)^2.$$

This and definition of the motion $x(\cdot,y,\bar{z})$ imply that

$$\Big| \big\langle \mathbb{E}_{s_k,s_{k+1}}(X(t)-\bar{z}|X(s_k)=\bar{z}),\bar{z}-y\big\rangle - \langle x_k(t,y,\bar{z})-y,\bar{z}-y\rangle \Big| \\ \leq \int_{s_k}^{t}\int_U \|f(t',\bar{z},\mathscr{I}(\mu(t)),u) \\ - f(t',x(t',y,\bar{z}),m(t),u)\|\zeta_{\bar{z}}(du|t')dt'\cdot\|y-\bar{z}\| \\ + (\varepsilon(t-s_k)+RB_Q(t-s_k)^2)\|y-\bar{z}\|.$$

Using the Lipschitz continuity of the function f, we obtain the following:

$$\Big| \big\langle \mathbb{E}_{s_k,s_{k+1}}(X(t)-\bar{z}|X(s_k)=\bar{z}),\bar{z}-y\big\rangle - \langle x_k(t,y,\bar{z})-y,\bar{z}-y\rangle \Big| \\ \leq C_f R(t-s_k)^2\|y-\bar{z}\| \\ + C_f\int_{s_k}^{t} W_2(\mathscr{I}(\mu(t')),m(t'))\|y-\bar{z}\| \\ + (\varepsilon(t-s_k)+RB_Q(t-s_k)^2)\|y-\bar{z}\|.$$

Plugging this estimate into (12) and taking into account (13), we derive the estimate:

$$\begin{aligned}\mathbb{E}_{s_k,s_{k+1}}\big(\|X(t)-x_k(t,y,\bar z)\|^2|X(s_k)=\bar z\big) &\leq \|y-\bar z\|^2\\ &+2\varepsilon^2(t-s_k)+2C_1'(t-s)^{3/2}+R(t-s_k)^2\\ &+C_fR(t-s_k)^2+C_fR(t-s_k)\|y-\bar z\|^2\\ &+C_f(t-s_k)W_2^2(\mathscr{I}(\mu(s_k)),m(s_k))+3C_f(t-s_k)\|y-\bar z\|^2\\ &+C_f\varepsilon^2(t-s_k)^2+C_1'C_f(t-s_k)^{5/2}+RC_f(t-s_k)^2\\ &+(\varepsilon+RB_Q(t-s_k))^2(t-s_k)+(t-s_k)\|y-\bar z\|^2\\ \leq \|y-\bar z\|^2+C_2'(t-s_k)\|y-\bar z\|^2 &\\ &+C_3'(t-t_k)W_2^2(\mathscr{I}(\mu(s_k)),m(s_k))\\ &+3\varepsilon(t-s_k)+2C_1'(t-s_k)^{3/2}+C_4'(t-s_k)^2\\ &+C_5'\varepsilon B_Q(t-s_k)^2+R^2B_Q^2(t-s_k)^3.\end{aligned}$$

Due to (11), we arrive at the inequality

$$\begin{aligned}W_2^2(\mathscr{I}(\mu(t)),m(t))\leq (1+C_6'(t-s_k))&W_2^2(\mathscr{I}(\mu(s_k)),m(s_k))\\ &+3\varepsilon(t-s_k)+2C_1'(t-s_k)^{3/2}+C_4'(t-s_k)^2\\ &+C_5'\varepsilon B_Q(t-s_k)^2+R^2B_Q^2(t-s_k)^3.\end{aligned}\tag{14}$$

Applying this inequality sequentially, we deduce the statement of the theorem. □

6 Model predictive control for Markov chains

In this section, given an initial distribution for the deterministic mean field type control system $m_0\in\mathcal{P}^2(\mathcal{K})$, a distribution of controls $\alpha\in\mathcal{U}_{0,T}$ such that $\alpha\in\mathcal{A}_{0,T}[m_0]$, and an initial system for mean field Markov chain $\mu_0\in\Sigma^2$, we construct a feedback strategy $\zeta_{\mathcal{S}}$ such that the corresponding motion of the Markov chain starting at μ_0 approximates the motion $m(\cdot,0,\alpha)$. Within this section, we denote

$$m(\cdot)\triangleq m(\cdot,0,\alpha).$$

Further, for $(y,\xi)\in\mathbb{R}^d\times\mathcal{U}_{0,T}$ and $s\in[0,T]$, we put

$$\mathscr{T}^s(y,\xi)\triangleq x(s,0,y,m(\cdot),\xi).$$

Notice that, if $(y',\xi')=\mathscr{T}^s(y,\xi)$, then

$$x(\cdot,0,y,m(\cdot),\xi)=x(\cdot,s,y',m(\cdot),\xi').$$

Informally, the operator $\mathscr{T}^s$ transfers the initial condition and the control from $t=0$ to the time s.

Below, if $\zeta_{\mathcal{S}},\zeta'_{\mathcal{S}}$ are feedback controls for the Markov chain on $[s,r]$ and $[r,\theta]$ respectively, then we denote by $\zeta_{\mathcal{S}}\diamond_r\zeta'_{\mathcal{S}}$ the feedback control such that $\zeta_{\mathcal{S}}\diamond_r\zeta'_{\mathcal{S}}=(\zeta_{\bar x}\diamond_r\zeta'_{\bar x})_{\bar x\in\mathcal{S}}$.

Finally, let $\Delta = \{s_i\}_{i=0}^n$ be a partition of $[0,T]$.

The construction is stepwise.

(M1) Let π_0 be an optimal plan between m_0 and $\mathscr{I}(\mu_0)$ and let $\pi_0(\cdot|\bar{x})$ be its disintegration w.r.t. $\mathscr{I}(\mu_0)$. We define a probability $\zeta_{\mathcal{S},0} = (\zeta_{\bar{x},0})_{\bar{x}} \in \mathcal{U}^{\mathcal{S}}_{s_0,s_1}$ by the rule: for $\phi \in C_b([s_0,s_1]\times U)$

$$\int_{[s_0,s_1]\times U} \phi(t,u)\zeta_{\bar{x},0}(d(t,u)) \triangleq \int_{\mathcal{K}}\int_{\mathcal{U}_{s_0,s_1}}\int_{[s_0,s_1]\times U} \phi(t,u)\xi(d(t,u))\alpha_0(\xi|y)\pi_0(dy|\bar{x}).$$

Hereinafter, α_0 is a restriction of α on $[s_0,s_1]$.

(M2) Assume now that we already constructed controls $\zeta_{\mathcal{S},0},\ldots,\zeta_{\mathcal{S},k-1}$ and put, for $t\in[0,s_k]$,

$$\mu(t) \triangleq \mu(t,0,\mu_0,\zeta_{\mathcal{S},0}\diamond_{s_1}\ldots\diamond_{s_{k-1}}\zeta_{\mathcal{S},k-1}).$$

To extend the control to the next time step, we first set α_k to be a restriction of the distribution of controls $\mathscr{T}^{s_k}\sharp\alpha$ to the time interval $[s_k,s_{k+1}]$. Further, let π_k be an optimal plan between $\mathscr{I}(\mu(s_k))$ and $m(s_k)$. A feedback control $\zeta_{\mathcal{S},k} = (\zeta_{\bar{x},k})_{\bar{x}\in\mathcal{S}} \in \mathcal{U}^{\mathcal{S}}_{[s_k,s_{k+1}]}$ is defined by the rule: if $\phi\in C([s_k,s_{k+1}]\times U)$, then

$$\int_{[s_0,s_1]\times U}\phi(t,u)\zeta_{\bar{x},k}(d(t,u)) \triangleq \int_{\mathcal{K}}\int\mathcal{U}_{s_k,s_{k+1}}\int_{[s_k,s_{k+1}]\times U}\phi(t,u)\xi(d(t,u))\alpha_k(d\xi|y)\pi_k(dy|\bar{x}).$$

Theorem 2. *Let conditions (A1)–(A4) hold true and let* $\mu(\cdot) \triangleq \mu(\cdot,0,\mu_0,\zeta_{\mathcal{S},0}\diamond_{s_1}\ldots\diamond_{s_{n-1}}\zeta_{\mathcal{S},n-1})$*, then*

$$W_p(m(t),\mathscr{I}(\mu(t))) \le C_0W_2(\mathscr{I}(\mu_0),m_0) + C_1\varepsilon + C_2d^{1/2}(\Delta) + C_3d(\Delta) + C_4\varepsilon B_Qd(\Delta) + C_5B_Qd^2(\Delta).$$

Here $C_0,\ldots,C_5$ *are the same constant as in Theorem 1.*

Proof. The proof mimics the proof of Theorem 1. We consider a time interval $[s_k,s_{k+1}]$ and choose $\mathbb{P}_{s_k,s_{k+1}}$ to be a realization of the flow $\mu(\cdot)$ on $[s_k,s_{k+1}]$. If $t\in[s_{k-1},s_k]$, then

$$W_2^2(m(t),\mathscr{I}(\mu(t))) \le \int_{\mathcal{K}\times\mathcal{S}}\int_{\mathcal{U}_{s_k,s_{k+1}}}\mathbb{E}_{s_k,s_{k+1}}(\|X(t)-x_k(t,y,\xi)\|^2|X(s_k)=\bar{z})\,\pi_k(d(y,\bar{z})). \tag{15}$$

Here, as above, we denote

$$x_k(t,y,\xi)\triangleq x(t,s_k,y,m(\cdot),\xi).$$

Further, we have that

$$\begin{aligned}\mathbb{E}_{s_k,s_{k+1}}(\|X(t)-x_k(t,y,\xi)\|^2|X(s_k)=\bar{z}) \leq \|\bar{z}-y\|^2 \\ + 2\mathbb{E}_{s_k,s_{k+1}}(\|X(t)-X(s_k)\|^2|X(s_k)=\bar{z}) \\ + 2\|x_k(t,y,\xi)-y\|^2 \\ + 2\mathbb{E}_{s_k,s_{k+1}}(\langle X(t)-X(s_k),\bar{z}-y\rangle|X(s_k)=\bar{z}) \\ - 2\langle x_k(t,y,\xi)-y,\bar{z}-y\rangle.\end{aligned}$$

Using Lemma 1, we conclude that

$$\begin{aligned}2\mathbb{E}_{s_k,s_{k+1}}\big(\|X(t)-X(s_k)\|^2|X(t_k)=\bar{z}\big) + 2\|x_k(t,y,\bar{z})-y\|^2 \\ \leq 2\varepsilon^2(t-s) + 2C_1'(t-s)^{3/2} + R^2(t-s)^2.\end{aligned} \tag{16}$$

Further, as in the proof of Theorem 1, we have that

$$\begin{aligned}\left\|\mathbb{E}_{s_k,s_{k+1}}(X(t)-\bar{z}|X(s_k)=\bar{z}) - \int_{s_k}^t\int_U f(t',\bar{z},\mathscr{I}(\mu(t')),u)\zeta_{\bar{z},k}(du|t')dt'\right\| \\ \leq \varepsilon(t-s_k) + RB_Q(t-s_k)^2.\end{aligned}$$

Simultaneously,

$$x_k(t,y,\xi)-y = \int_{s_k}^t\int_U f(t',x_k(t',y,\xi),m(t'),u)\xi(du|t')dt'$$

Plugging this into (15) and taking into account the definition of $\zeta_{\mathcal{S},k}$, we conclude that

$$\begin{aligned}W_2^2(m(t),\mathscr{I}(\mu(t))) \leq W_2^2(m(s_k),\mathscr{I}(\mu(s_k))) \\ + 2\varepsilon^2(t-s) + 2C_1'(t-s)^{3/2} + R(t-s)^2 + \varepsilon(t-s_k) + B_Q(t-s_k)^2 \\ + \int_{\mathcal{K}\times\mathcal{S}}\int_{\mathcal{U}_{s_k,s_{k+1}}}\Bigg\|\int_U f(t',\bar{z},\mathscr{I}(\mu(t')),u)\xi(du|t') \\ - \int_U f(t',x_k(t',y,\xi),m(t'),u)\xi(du|t')\Bigg|\cdot\|\bar{z}-y\|.\end{aligned}$$

Estimating the last term as in the proof of Theorem 1, we obtain the inequality:

$$\begin{aligned}W_2^2(\mathscr{I}(\mu(t)),m(t)) \leq (1+C_6'(t-s_k))W_2^2(\mathscr{I}(\mu(s_k)),m(s_k)) \\ + 3\varepsilon(t-s_k) + 2C_1'(t-s_k)^{3/2} + C_4'(t-s_k)^2 \\ + C_5'\varepsilon B_Q(t-s_k)^2 + R^2B_Q^2(t-s_k)^3.\end{aligned} \tag{17}$$

Here the constants are the same as in (14). Applying (17) sequentially, we arrive at the statement of the theorem. □

7 Hausdorff distance between bundles of flows

In this short section, we consider the bundles of flows of probabilities generated by the original first order mean field type control system and the mean field Markov chain. To define them, let

$$\mathcal{X}(m_0) \triangleq \{m(\cdot,0,\alpha) : \alpha \in \mathcal{A}_{0,T}[m_0]\},$$

$$\mathcal{X}_Q(\mu_0) \triangleq \{\mu(\cdot,0,\mu,\zeta_{\mathcal{S}}) : \zeta_{\mathcal{S}} \in \mathcal{U}^{\mathcal{S}}_{0,T}\}.$$

Notice that $\mathcal{X}(m_0) \subset C([0,T];\mathcal{P}^2(\mathbb{R}^d))$, while $\mathcal{X}_Q \subset C([0,T];\Sigma^2)$

If $m(\cdot) \in C([0,T];\mathcal{P}^2(\mathbb{R}^d))$, $\mu \in C([0,T];\Sigma^2)$, we denote

$$\mathbb{d}(m(\cdot),\mu(\cdot)) \triangleq \sup_{t\in[0,T]} W_2(m(t),\mathscr{I}(\mu(t))).$$

If $\Upsilon_1 \subset C([0,T];\mathcal{P}^2(\mathbb{R}^d))$, $\Upsilon_2 \subset C([0,T],\Sigma^2)$ are closed sets, then we introduce the Hausdorff distance in the standard way:

$$\mathbb{H}(\Upsilon_1,\Upsilon_2) \triangleq \max\left\{\sup_{m(\cdot)\in\Upsilon_1}\inf_{\mu(\cdot)\in\Upsilon_2}\mathbb{d}(m(\cdot),\mu(\cdot)),\ \sup_{\mu(\cdot)\in\Upsilon_2}\inf_{m(\cdot)\in\Upsilon_1}\mathbb{d}(m(\cdot),\mu(\cdot))\right\}.$$

The main result of this short section is the following.

Proposition 4. *Assume that the approximation conditions (A1)–(A4) are in force. Then,*

$$\mathbb{H}(\mathcal{X}(m_0),\mathcal{X}_Q(\mu_0)) \leq C_0 W_2(\mathscr{I}(\mu_0),m_0) + C_1\varepsilon,$$

where C_0 and C_1 are constants dependent only on f and T.

Proof. From Theorem 2, it follows that, given $m(\cdot) \in \mathcal{X}(m_0)$ and arbitrary $\delta > 0$, one can find a partition of $[0,T]$ $\Delta = \{s_i\}_{i=0}^n$ and controls $\zeta_{\mathcal{S}},\ldots,\zeta_{\mathcal{S},n-1}$ such that

$$\mathbb{d}(m(\cdot),\mu(\cdot)) \leq C_0 W_2((\mu_0),m_0) + C_1\varepsilon + \delta,$$

where $\mu(\cdot) = \mu(\cdot,0,\mu_0,\zeta_{\mathcal{S},0} \diamond_{s_1} \ldots \diamond_{s_{n-1}} \zeta_{\S,n-1})$. Passing to the limit while $\delta \to 0$, we conclude that

$$\sup_{m(\cdot)\in\mathcal{X}(m_0)}\inf_{\mu(\cdot)\in\mathcal{X}_Q(\mu_0)}\mathbb{d}(m(\cdot),\mu(\cdot)) \leq C_0 W_2(\mathscr{I}(\mu_0),m_0) + C_1\varepsilon.$$

The inequality

$$\sup_{\mu(\cdot)\in\mathcal{X}_Q(\mu_0)}\inf_{m(\cdot)\in\mathcal{X}(m_0)}\mathbb{d}(m(\cdot),\mu(\cdot)) \leq C_0 W_2(\mathscr{I}(\mu_0),m_0) + C_1\varepsilon$$

is proved in the similar way using Theorem 1. □

References

[1] N. Ahmed and X. Ding. *Controlled McKean-Vlasov equation*, Commun. Appl. Anal., **5** (2001), 183–206, .

[2] 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.

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

[4] Y. Averboukh. *Viability theorem for deterministic mean field type control systems*, Set-Valued Var. Anal., **26**:4 (2018), 993–1008.

[5] Y. Averboukh. *Lattice approximations of the first-order mean field type differential games*, Nonlinear Differ. Equ. Appl., **28**:6 (2021), 42. Id/No 65.

[6] Y. Averboukh and D. Khlopin. *Pontryagin maximum principle for the deterministic mean field type optimal control problem via the lagrangian approach.* Preprint at arXiv:2207.01892, 2022.

[7] Y. Averboukh, A. Marigonda, and M. Quincampoix. *Extremal shift rule and viability property for mean field-type control systems*, J. Optim. Theory Appl., **189**:1 (2021), 244–270.

[8] Y. V. Averboukh. *A mean field type differential inclusion with upper semicontinuous right-hand side*, Vestn. Udmurt. Univ.: Mat. Mekhanika Komp'yuternye Nauki, **32**:4 (2022), 489–501.

[9] 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** (2018), 2115–2160.

[10] N. Bellomo, B. Piccoli, and A. Tosin. *Modeling crowd dynamics from a complex system viewpoint*, Math. Models Methods Appl. Sci., **22** (2012), 1230004, 29.

[11] M. Bivas and M. Quincampoix. *Nonsmooth feedback control for multi-agent dynamics*, Set-Valued Var. Anal., **29**:2 (2021), 501–518.

[12] B. Bonnet and H. Frankowska. *Differential inclusions in Kantorovich spaces: the Cauchy-Lipschitz framework*, J. Differ Equ., **271** (2021), 594–637.

[13] B. Bonnet and H. Frankowska. *Viability and exponentially stable trajectories for differential inclusions in Wasserstein spaces.* In 2022 IEEE 61st Conference on Decision and Control (CDC), pages 5086–5091. IEEE, 2022.

[14] B. Bonnet-Weil and H. Frankowska. *Carathéodory theory and a priori estimates for continuity inclusions in the space of probability measures.* Preprint at arXiv:2302.00963, 2023.

[15] R. Buckdahn, B. Djehiche, and J. Li. *A general stochastic maximum principle for SDEs of mean-field type*, Appl. Math. Optim., **64**:2 (2011), 197–216.

[16] F. Bullo, J. Cortés, and S. Martínez. *Distributed control of robotic networks: a mathematical approach to motion coordination algorithms.* Princeton Ser. Appl. Math. Princeton, NJ: Princeton University Press, 2009.

[17] R. Carmona and F. Delarue. *Forward-backward stochastic differential equations and controlled McKean–Vlasov dynamics*, Ann. Probab., **43**:5 (2015), 2647–2700.

[18] R. M. Colombo, M. Garavello, and M. Lécureux-Mercier. *Non-local crowd dynamics*, C. R. Acad. Sci. Paris Sér. I Math., 349(13-14):769–772, 2011.

[19] R. M. Colombo and M. D. Rosini. *Pedestrian flows and non-classical shocks*, Math. Methods Appl. Sci., **28**:13 (2005), 1553–1567.

[20] E. Cristiani, B. Piccoli, and A. Tosin. *Multiscale modeling of pedestrian dynamics*, volume 12 of *MS&A, Model. Simul. Appl.* Springer, 2014.

[21] C. Dellacherie and P. Meyer. *Probabilities and potential.* North-Holland Publishing Co., Amsterdam, 1978.

[22] B. Djehiche and S. Hamadène. *Optimal control and zero-sum stochastic differential game problems of mean-field type*, Appl. Math. Optim., **81** (2020), 933–960.

[23] L. Grüne and J. Pannek. *Nonlinear Model Predictive Control. Theory and Algorithms.* Springer, 2011.

[24] C. Jimenez, A. Marigonda, and M. Quincampoix. *Optimal control of multiagent systems in the Wasserstein space*, Calc. Var. Partial Differ. Equ., **59** (2020), Article number: 58.

[25] V. N. Kolokoltsov. *Markov Processes, Semigroups and Generators.* De Gruyter, Berlin, 2011.

[26] N. N. Krasovskii and A. N. Kotelnikova. *Unification of differential games, generalized solutions of the hamilton-jacobi equations, and a stochastic guide*, Differ. Equ., **45**:11 (2009), 1653–1668.

[27] N. N. Krasovskii and A. N. Kotelnikova. *An approach-evasion differential game: stochastic guide*, Proc. Steklov Inst. Math., **269**: Supplement 1 (2010), 191–213.

[28] N. N. Krasovskii and A. N. Kotelnikova. *On a differential interception game*, Proc. Steklov Inst. Math., **268**:1 (2010), 161–206.

[29] N. N. Krasovskii and A. N. Kotelnikova. *Stochastic guide for a time-delay object in a positional differential game*, Proc. Steklov Inst. Math., **277**:Supplement 1 (2012), 145–151.

[30] M. Laurière and O. Pironneau. *Dynamic programming for mean-field type control*, C. R. Math. Acad. Sci. Paris, **352**:9 (2014), 707–713.

[31] H. P. McKean. *A class of Markov processes associated with nonlinear parabolic equations*, Proc. Natl. Acad. Sci. U.S.A., **56** (1966), 1907–1911.

[32] H. Pham and X. Wei. *Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics*, SIAM J. Control Optim., **55** (2017), 1069–1101.

[33] H. Pham and X. Wei. *Bellman equation and viscosity solutions for mean-field stochastic control problem*, ESAIM Control Optim. Calc. Var., **24**:1 (2018), 437–461.

[34] N. Pogodaev. *Optimal control of continuity equations*, Nonlinear Differ. Equ. Appl., **23** (2016), Art21, 24 pp..

[35] A.-S. Sznitman. *Topics in propagation of chaos.* In Ecole dété de probabilités de Saint-Flour XIX – 1989, volume 1464 of *Lecture Notes in Math.* Springer, Berlin, 1991.

[36] A. Vlasov. *On vibration properties of electron gas*, J. Exp. Theor. Phys., **8**:3 (1938), 291–318. (in Russian).

[37] A. Vlasov. *Many-particle theory and its application to plasma.* Gordon and Breach, New York, 1961.

Yurii Vladimirovich Averboukh
Krasovskii Institute of Mathematics and Mechanics,
16 S.Kovalevskaya Str.,
620108, Yekaterinburg, Russia
Email address: ayv@imm.uran.ru