

# Mean Field Control Hierarchy

Giacomo Albi<sup>1</sup> · Young-Pil Choi<sup>2</sup> · Massimo Fornasier<sup>3</sup> ·  
Dante Kalise<sup>4</sup>

Published online: 6 July 2017  
© Springer Science+Business Media, LLC 2017

**Abstract** In this paper we model the role of a government of a large population as a mean field optimal control problem. Such control problems are constrained by a PDE of continuity-type, governing the dynamics of the probability distribution of the agent population. We show the existence of mean field optimal controls both in the stochastic and deterministic setting. We derive rigorously the first order optimality conditions useful for numerical computation of mean field optimal controls. We introduce a novel approximating hierarchy of sub-optimal controls based on a Boltzmann approach, whose computation requires a very moderate numerical complexity with respect to the one of the optimal control. We provide numerical experiments for models in opinion formation comparing the behavior of the control hierarchy.

**Keywords** PDE constrained optimization multi-agent systems · Boltzmann equations · Kinetic equations

---

✉ Massimo Fornasier  
massimo.fornasier@ma.tum.de

Giacomo Albi  
giacomo.albi@univr.it

Young-Pil Choi  
ypchoi@inha.ac.kr

Dante Kalise  
dante.kalise@oeaw.ac.at

<sup>1</sup> Department of Computer Science, University of Verona, Str. Le Grazie,15, 37134 Verona, Italy

<sup>2</sup> Department of Mathematics, Inha University, Incheon 402-751, Republic of Korea

<sup>3</sup> Department of Mathematics, TU München, Boltzmannstr. 3, 85748 Garching bei München, Germany

<sup>4</sup> RICAM, Austrian Academy of Sciences, Altenbergerstr. 69, 4040 Linz, Austria

## 1 Introduction

Self-organization in social interactions is a fascinating mechanism, which inspired the mathematical modeling of multi-agent interactions towards formation of coherent global behaviors, with applications in the study of biological, social, and economical phenomena. Recently there has been a vigorous development of literature in applied mathematics and physics describing collective behavior of multiagent systems [41–43, 52, 56, 58, 82], towards modeling phenomena in biology, such as cell aggregation and motility [21, 59, 60, 73], coordinated animal motion [12, 27, 34, 37–39, 43, 66, 70, 71, 74, 79, 85], coordinated human [40, 44, 76] and synthetic agent behavior and interactions, such as cooperative robots [35, 63, 72, 77]. As it is very hard to be exhaustive in accounting all the developments of this very fast growing field, we refer to [28–30, 33, 81] for recent surveys.

Two main mechanisms are considered in such models to drive the dynamics. The first, which takes inspiration, e.g., from physics laws of motion, is based on binary forces encoding observed “first principles” of biological, social, or economical interactions. Most of these models start from particle-like systems, borrowing a leaf from Newtonian physics, by including fundamental “social interaction” forces within classical systems of 1st or 2nd order equations. In this paper we mix general principles with concrete modeling instances to encounter the need of both a certain level of generality and to provide immediately concrete applications. Accordingly, we consider here mainly large particle/agent systems of form:

$$dx_i = \left( \frac{1}{N} \sum_{j=1}^N P(x_i, x_j)(x_j - x_i) \right) dt + \sqrt{2\sigma} dB_i^t, \quad i = 1, \dots, N, \quad t > 0, \quad (1.1)$$

where  $P(\cdot, \cdot)$  represents the communication function between agents  $x_i \in \mathbb{R}^d$  and  $B_i^t$  is a  $d$ -dimensional Brownian motion.

The second mechanism, which we do not address in detail here, is based on evolutive games, where the dynamics is driven by the simultaneous optimization of costs by the players, perhaps subjected to selection, from game theoretic models of evolution [54] to mean field games, introduced in [62] and independently under the name Nash Certainty Equivalence (NCE) in [55], later greatly popularized, e.g., within consensus problems, for instance in [67, 68].

The common viewpoint of these branches of mathematical modeling of multi-agent systems is that the dynamics are based on the *free* interaction of the agents or *decentralized* control. The wished phenomenon to be described is their self-organization in terms of the formation of complex macroscopic patterns.

One fundamental goal of these studies is in fact to reveal the possible relationship between the simple binary forces acting at individual level, being the “first principles” of social interaction or the game rules, and the potential emergence of a global behavior in the form of specific patterns.

For instance one can use the model in (1.1), for  $d = 1$  and  $x_i \in I = [-1, 1]$ , a bounded interval, to formulate classical opinion models, where  $x_i$  represents an opin-

ion in the continuous set between two opposite opinions  $\{-1, 1\}$ . According to the choice of the communication function  $P(\cdot, \cdot)$ , consensus can emerge or not, and different studies have been made in order to enforce the emergence of a global consensus, [5–7, 45, 80]. The mathematical property for a system to form patterns is actually its persistent *compactness*. There are actually several mechanisms of promotion of compactness to yield eventually self-organization. In the recent paper [65], for instance, the authors name the *heterophilia*, i.e., the tendency to bond more with those who are “different” rather than those who are similar, as a positive mechanism in consensus models to reach accord. However also in *homophilious* societies influenced by more local interactions, global self-organization towards consensus can be expected as soon as enough initial coherence is given. At this point, and perhaps reminiscently of biblical stories from the Genesis, one could enthusiastically argue “*Let us give them good rules and they will find their way!*” Unfortunately, this is not true, at all. In fact, in homophilious regimes there are plenty of situations where patterns will not spontaneously form. In Sect. 5 below we mathematically demonstrate with a few simple numerical examples the incompleteness of the self-organization paradigm, and we refer to [18] for its systematic discussion. Consequently, we propose to amend it by allowing possible external interventions in form of *centralized* controls. The human society calls them *government*.

The general idea consists in considering dynamics of the form

$$dx_i = \left( \frac{1}{N} \sum_{j=1}^N P(x_i, x_j)(x_j - x_i) \right) dt + f_i dt + \sqrt{2\sigma} dB_i^t, \quad i = 1, \dots, N, \quad t > 0, \tag{1.2}$$

where the control  $f = (f_1, \dots, f_N)$  minimizes a given functional  $J(x, f)$ . As an example we can consider the following variational formulation

$$f = \arg \min_{g \in \mathcal{U}} J(x, g) := \mathbb{E} \left[ \int_0^T \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{2} |x_i - \bar{x}|^2 + \gamma \Psi(g_i) \right) dt \right], \tag{1.3}$$

where  $\bar{x}$  represents a target point,  $\gamma$  is the penalization parameter of the control  $g$ , which is chosen among the admissible controls in  $\mathcal{U}$ , and  $\Psi : \mathbb{R}^d \rightarrow \mathbb{R}_+ \cup \{0\}$  is a convex function. The choice of this particular cost function, and especially of the term  $\int_0^T \frac{1}{2} \int |x - \bar{x}|^2 \mu(x, t) dx$  is absolutely arbitrary. It is consistent with our wish of mixing general statements with instances of applications, and the cost function is so given to provide immediately a specific instance of application oriented to opinion consensus problems. Similar models as (1.3) have been studied recently also for the flocking dynamics in [4, 17, 24, 51] and one can of course consider many more instances, as soon as one ensures enough continuity of the cost, see, e.g., [51].

As the number of particles  $N \rightarrow \infty$ , the finite dimensional optimal control problem with ODE constraints (1.2)–(1.3) converges to the following mean field optimal control problem [1, 15, 51, 61]:

$$\partial_t \mu + \nabla \cdot ((\mathcal{P}[\mu] + f) \mu) = \sigma \Delta \mu, \quad (1.4)$$

where the interaction force  $\mathcal{P}$  is given by

$$\mathcal{P}[\mu](x) = \int P(x, y)(y - x)\mu(y, t) dy \quad (1.5)$$

and the solution  $\mu$  is controlled by the minimizer of the cost functional

$$J(\mu, f) = \int_0^T \left( \frac{1}{2} \int |x - \bar{x}|^2 \mu(x, t) dx + \gamma \int \Psi(f)\mu(x, t) dx \right) dt. \quad (1.6)$$

To a certain extent, the mean field optimal control problem (1.4)–(1.6) can be viewed as a generalization of optimal transport problems [14] for which the term  $P \equiv 0$ , the term  $\int_0^T \frac{1}{2} \int |x - \bar{x}|^2 \mu(x, t) dx$  does not appear in the cost, and final conditions are given. Differently from mean field games [62] the goal here is not to derive the equilibria of a multi-player game, rather to compute mean field optimal government strategies for a population so large that the curse of dimensionality would otherwise prohibit numerical solutions. The mean field optimal control problem (1.4)–(1.6) provides an artificial confinement vector field  $f$ , inducing the right amount of compactness to have global convergence to steady states (pattern formation). *Local* convergence towards, e.g., to global Maxwellians, is provided for certain second order mean field-type of equations in [32, 46]. Hence, our results can be also interpreted as an external model perturbation to induce *global* stability.

In this paper we provide a friendly introduction to mean field optimal controls of the type (1.4)–(1.6), showing their main analytical properties and furnish a simple route to their numerical solutions, which we call “the control hierarchy”. Although some of the results contained in this paper are certainly also derived elsewhere, see, e.g., [15, 51], we made an effort to present them in a simplified form as well as providing rigorous derivations.

In particular, in Sect. 2, we show existence of mean field optimal controls for first order models in case of both stochastic and deterministic control problems. We also derive rigorously in Sect. 3 the corresponding first order optimality conditions, resulting in a coupled system of forward/backward time-dependent PDEs. The forward equation is given by (1.4), while the backward one is a nonlocal integro-differential advection-reaction-diffusion equation. The presence of nonlocal interaction terms in form of integral functions is another feature, which distinguishes mean field optimal control problems from classical mean field games [62] and optimal transport problems [14], where usually  $P \equiv 0$ . The nonlocal terms pose additional challenges in the numerical solution, which are subject of recent studies [22].

Although mean field optimal controls are designed to be independent of the number  $N$  of agents to provide a way to circumvent the curse of dimensionality of  $N \rightarrow \infty$ , still their numerical computation needs to be realized by solving the first-order optimality conditions. The complexity of their solution depends on the intrinsic dimensionality  $d$  of the agents, which is affordable only at moderate dimensions (e.g.,  $d \leq 3$ ). For this reason, in Sect. 4 we approach the solution of the mean field optimal control,

by means of a novel hierarchy of suboptimal controls, computed by a Boltzmann approach: first one derives a control for a system of two representative particles, then one plugs it into a collisional operator considering the statistics of the interactions of a distribution of agents, and finally one performs a *quasi-invariant limit* to approximate the PDE of continuity-type, governing the dynamics of the probability distribution of the agent population. For the two particle system considered in the first step of the Boltzmann approach above, we propose two suboptimal controls stemming from the binary Boltzmann approach: the first level is given by an instantaneous *model predictive control* on two interacting agents—we shall call this control *instantaneous control* (IC), while the second stems from the solution of the binary optimal control problem by means of the Bellman dynamical programming principle—we shall call this control *finite horizon control* (FH). These two controls have the advantage that the complexity of their computation is dramatically reduced with respect to the mean field optimal control (OC) in its full glory, still retaining their ability to induce government of the population. We describe in detail how they can be efficiently numerically computed. In Sect. 5 we provide simple numerical approaches, easily implementable, for solving one-dimensional mean field optimal control problems of the type (1.4)–(1.6). We eventually numerically compare the control hierarchy with the mean field optimal control in a model of opinion formation and we show the quasi-optimality of the Boltzmann–Bellman (FH) control. To facilitate the reproducibility of our results and to allow other scientists to easily access this very exciting field, we provide at the link <https://www-m15.ma.tum.de/Allgemeines/SoftwareSite> the Matlab code used to produce our numerical experiments.

## 2 Existence of Mean Field Optimal Controls

### 2.1 Deterministic Case

In this section, we study global existence and uniqueness of weak solutions for Eq. (1.4) in  $\mathbb{R}^d$  without the diffusion, i.e.,  $\sigma = 0$ , namely

$$\partial_t \mu + \nabla \cdot ((\mathcal{P}[\mu] + f) \mu) = 0, \quad x \in \mathbb{R}^d, \quad t > 0. \tag{2.1}$$

We also investigate the mean field limit of the ODE constrained control problem (1.2)–(1.3) in the deterministic setting. Let us denote by  $\mathcal{M}(\mathbb{R}^d)$  and  $\mathcal{M}_p(\mathbb{R}^d)$  the sets of all probability measures and the ones with finite moments of order  $p \in [1, \infty)$  on  $\mathbb{R}^d$ , respectively. We first define a notion of weak solutions to the equation to (2.1).

**Definition 2.1** For a given  $T > 0$ , we call  $\mu \in \mathcal{C}([0, T]; \mathcal{M}_1(\mathbb{R}^d))$  a weak solution of (2.1) on the time-interval  $[0, T]$  if for all compactly supported test functions  $\varphi \in C_c^\infty(\mathbb{R}^d \times [0, T])$ ,

$$\begin{aligned} & \int_{\mathbb{R}^d} \varphi(x, T) \mu_T(dx) - \int_0^T \int_{\mathbb{R}^d} (\partial_t \varphi + (\mathcal{P}[\mu_t] + f) \cdot \nabla \varphi) \mu_t(dx) dt \\ & = \int_{\mathbb{R}^d} \varphi_0(x) \mu_0(dx). \end{aligned}$$

We also introduce a set of admissible controls  $\mathcal{F}_\ell([0, T])$  in the definition below.

**Definition 2.2** For a given  $T$  and  $q \in [1, \infty)$ , we fix a control bound function  $\ell \in L^q(0, T)$ . Then  $f \in \mathcal{F}_\ell([0, T])$  if and only if

- (i)  $f : \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}^d$  is a Carathéodory function.
- (ii)  $f(\cdot, t) \in W_{loc}^{1,\infty}(\mathbb{R}^d)$  for almost every  $t \in [0, T]$ .
- (iii)  $|f(0, t)| + \|f(\cdot, t)\|_{Lip} \leq \ell(t)$  for almost every  $t \in [0, T]$ .

For the existence and mean field limit, we use the topology on probability measures induced by the Wasserstein distance, which is defined by

$$\mathcal{W}_p(\mu, \nu) := \inf_{\pi \in \Gamma(\mu, \nu)} \left( \int_{\mathbb{R}^{2d}} |x - y|^p \pi(dx, dy) \right)^{1/p} \quad \text{for } p \geq 1 \text{ and } \mu, \nu \in \mathcal{M}(\mathbb{R}^d),$$

where  $\Gamma(\mu, \nu)$  is the set of all probability measures on  $\mathbb{R}^{2d}$  with first and second marginals  $\mu$  and  $\nu$ , respectively. Note that  $\mathcal{M}_1(\mathbb{R}^d)$  is a complete metric space endowed with the  $\mathcal{W}_1$  distance, and  $\mathcal{W}_1$  is equivalently characterized in duality with Lipschitz continuous functions [84].

The following result is a rather straightforward adaptation from [51] and we shall prove it rather concisely. For more details we address the interested reader to [51], which has been written in a more scholastic and perhaps accessible form.

**Theorem 2.1** *Let the initial data  $\mu_0 \in \mathcal{M}(\mathbb{R}^d)$  and assume that  $\mu_0$  is compactly supported, i.e., there exists  $R > 0$  such that*

$$\text{supp } \mu_0 \subset B(0, R),$$

where  $B(0, R) := \{x \in \mathbb{R}^d : |x| < R\}$ . Furthermore, we assume that  $P \in W^{1,\infty}(\mathbb{R}^{2d})$ . Then, for a given  $f \in \mathcal{F}_\ell([0, T])$ , there exists a unique weak solution  $\mu \in \mathcal{C}([0, T]; \mathcal{M}_1(\mathbb{R}^d))$  to Eq. (1.4) with  $\sigma = 0$ . Furthermore,  $\mu$  is determined as the push-forward of the initial measure  $\mu_0$  through the flow map generated by the locally Lipschitz velocity field  $\mathcal{P}[\mu] + f$ . Moreover, if  $\mu^i, i = 1, 2$  are two such with initial data  $\mu_0^i$  satisfying the above assumption, we have

$$\mathcal{W}_1(\mu_t^1, \mu_t^2) \leq C \mathcal{W}_1(\mu_0^1, \mu_0^2) \quad \text{for } t \in [0, T],$$

where  $C > 0$  depends only on  $\|P\|_{W^{1,\infty}}, R, T$ , and  $\|\ell\|_{L^q}$ .

*Proof • Existence and Uniqueness* Let  $\mu \in \mathcal{C}([0, T]; \mathcal{M}_1(\mathbb{R}^d))$  with compact support in  $B(0, R)$  for some positive constant  $R > 0$ . Then we can easily show that the interaction force  $\mathcal{P}$  is locally bounded and Lipschitz:

$$|\mathcal{P}[\mu](x)| \leq C(\|P\|_{L^\infty}, R)(1 + |x|),$$

and

$$|\mathcal{P}[\mu](x) - \mathcal{P}[\mu](y)| \leq C(\|P\|_{W^{1,\infty}}, R)(1 + |x|)|x - y|.$$

On the other hand, since  $f \in \mathcal{F}_\ell([0, T])$ , we obtain that the vector field  $\mathcal{P}[\mu] + f$  is also locally bounded and Lipschitz. Then this together with employing the argument in [23, Theorem 3.10] and existence theory for Carathéodory differential equation in [50], we can get the local-in-time existence and uniqueness of weak solutions to the system (1.4) with  $\sigma = 0$  in the sense of Definition 2.1. Note that those solutions exist as long as that solutions are compactly supported. Set

$$R(t) := \max_{x, y \in \text{Supp}(\mu_t)} |x - y| \quad \text{for } t \in [0, T].$$

Let us consider the following characteristic  $X(t) := X(t; s, x) : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ :

$$\frac{dX(t; s, x)}{dt} = \mathcal{P}[\mu_t](X(t; s, x), t) + f(X(t; s, x), t) \quad \text{for all } t, s \in [0, T], \quad (2.2)$$

with the initial data  $X_0 = x \in \mathbb{R}^d$ . We notice that characteristic is well-defined on the time interval  $[0, T]$  due to the regularity of the velocity field. A straightforward computation yields that for  $x, y \in \text{supp}(\mu_0)$

$$\begin{aligned} & \frac{d|X(t) - Y(t)|^2}{dt} \\ &= (X(t) - Y(t)) \cdot \frac{d(X(t) - Y(t))}{dt} \\ &\leq |X(t) - Y(t)| |\mathcal{P}[\mu_t](X(t), t) - \mathcal{P}[\mu_t](Y(t), t)| \\ &\quad + |X(t) - Y(t)| |f(X(t), t) - f(Y(t), t)| \\ &\leq 2\|P\|_{L^\infty} |X(t) - Y(t)| \int_{\mathbb{R}^d} |z - X(t)| \mu(z, t) dz + \|P\|_{L^\infty} |X(t) - Y(t)|^2 \\ &\quad + \|f(\cdot, t)\|_{\text{Lip}} |X(t) - Y(t)|^2. \end{aligned}$$

This deduces

$$\frac{dR(t)}{dt} \leq (3\|P\|_{L^\infty} + \|f(\cdot, t)\|_{\text{Lip}}) R(t) \leq (3\|P\|_{L^\infty} + \ell(t)) R(t),$$

and

$$R(t) \leq CR_0 \quad \text{for } t \in [0, T],$$

where  $C$  depends only on  $T, \|P\|_{L^\infty}$ , and  $\|\ell\|_{L^q}$ . Thus, by continuity arguments, we have the global existence of weak solutions. We can also find that for  $h \in \mathcal{C}_c^\infty(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} \mu(x, t) h(x) dx = \int_{\mathbb{R}^d} \mu_0(x) h(X(0; t, x)) dx \quad \text{for } t \in [0, T].$$

This implies that  $\mu$  is determined as the push-forward of the initial density through the flow map (2.2).

• *Stability Estimate* Let  $T > 0$  and  $\mu^i, i = 1, 2$  be the weak solutions to Eq. (1.4) with  $\sigma = 0$  obtained in the above. Let  $X_i$  be the characteristic flows defined in (2.2) generated by the velocity fields  $\mathcal{P}[\mu^i] + f$ , respectively. For a fixed  $t_0 \in [0, T]$ , we choose an optimal transport map for  $\mathcal{W}_1$  denoted by  $\mathcal{T}^0(x)$  between  $\mu_{t_0}^1$  and  $\mu_{t_0}^2$ , i.e.,  $\mu_{t_0}^2 = \mathcal{T}^0 \# \mu_{t_0}^1$ . It also follows from the above that  $\mu_t^i = X_i(t; t_0, \cdot) \# \mu_{t_0}^i$  for  $t \geq t_0$ . Furthermore, we get  $\mathcal{T}^t \# \mu_t^1 = \mu_t^2$  with  $\mathcal{T}^t = X_2(t; t_0, \cdot) \circ \mathcal{T}^0 \circ X_1(t_0; t, \cdot)$  for  $t \in [t_0, T]$ . Then we obtain

$$\begin{aligned} & \left. \frac{d^+ \mathcal{W}_1(\mu_t^1, \mu_t^2)}{dt} \right|_{t=t_0+} \\ & \leq \int_{\mathbb{R}^d} \left| \mathcal{P}[\mu_{t_0}^1](X_1(t; t_0, x), t) - \mathcal{P}[\mu_{t_0}^2](X_2(t; t_0, \mathcal{T}^0(x)), t) \right| \mu_{t_0}^1(dx) \Big|_{t=t_0+} \\ & \quad + \int_{\mathbb{R}^d} \left| f(X_1(t; t_0, x), t) - f(X_2(t; t_0, \mathcal{T}^0(x)), t) \right| \mu_{t_0}^1(dx) \Big|_{t=t_0+} \\ & = I_1 + I_2, \end{aligned}$$

where  $I_i, i = 1, 2$  are estimated as follows.

$$\begin{aligned} I_1 & \leq \int_{\mathbb{R}^{2d}} \left| P(x, y)(y - x) - P(\mathcal{T}^0(x), \mathcal{T}^0(y))(\mathcal{T}^0(y) - \mathcal{T}^0(x)) \right| \mu_{t_0}^1(dx) \mu_{t_0}^1(dy) \\ & \leq \int_{\mathbb{R}^{2d}} |P(x, y) - P(\mathcal{T}^0(x), \mathcal{T}^0(y))| |y - x| \mu_{t_0}^1(dx) \mu_{t_0}^1(dy) \\ & \quad + \int_{\mathbb{R}^{2d}} |P(\mathcal{T}^0(x), \mathcal{T}^0(y))| (|y - \mathcal{T}^0(y)| + |x - \mathcal{T}^0(x)|) \mu_{t_0}^1(dx) \mu_{t_0}^1(dy) \\ & \leq C \|P\|_{W^{1,\infty}} \mathcal{W}_1(\mu_{t_0}^1, \mu_{t_0}^2), \\ I_2 & = \int_{\mathbb{R}^d} \left| f(x, t) - f(\mathcal{T}^0(x), t) \right| \mu_{t_0}^1(dx) \leq \|f(\cdot, t)\|_{\text{Lip}} \mathcal{W}_1(\mu_{t_0}^1, \mu_{t_0}^2) \\ & \leq \ell(t) \mathcal{W}_1(\mu_{t_0}^1, \mu_{t_0}^2), \end{aligned}$$

where we used the fact that  $\mu$  has the compact support for the estimate of  $I_1$ . We now combine the above estimates together with being  $t_0$  arbitrary in  $[0, T]$  to conclude

$$\frac{d^+ \mathcal{W}_1(\mu_t^1, \mu_t^2)}{dt} \leq C (\|P\|_{W^{1,\infty}} + \ell(t)) \mathcal{W}_1(\mu_t^1, \mu_t^2), \quad \text{for } t \in [0, T].$$

This completes the proof. □

In Theorem 2.1, we show the global existence and uniqueness of weak solutions  $\mu$  to Eq. (1.4) with  $\sigma = 0$  for a given control  $f \in \mathcal{F}_\ell([0, T])$ . In the rest of this part, we show the rigorous derivation of the infinite dimensional optimal control problem from the finite dimensional one as  $N \rightarrow \infty$ . Let us recall the finite/infinite dimensional optimal control problems:



- *Finite dimensional optimal control problem:*

$$\min_{f \in \mathcal{F}_\ell} J(x, f) := \min_{f \in \mathcal{F}_\ell} \int_0^T \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{2} |x_i - \bar{x}|^2 + \gamma \Psi(f(x_i, t)) \right) dt, \quad (2.3)$$

where  $x_i$  is a unique solution of

$$\dot{x}_i = \frac{1}{N} \sum_{j=1}^N P(x_i, x_j)(x_j - x_i) + f(x_i, t), \quad i = 1, \dots, N, \quad t > 0, \quad (2.4)$$

- *Infinite dimensional optimal control problem:*

$$\min_{f \in \mathcal{F}_\ell} J(\mu_t, f) := \min_{f \in \mathcal{F}_\ell} \int_0^T \left( \frac{1}{2} \int_{\mathbb{R}^d} |x - \bar{x}|^2 \mu_t(dx) + \gamma \int_{\mathbb{R}^d} \Psi(f) \mu_t(dx) \right) dt, \quad (2.5)$$

where  $\mu \in \mathcal{C}([0, T]; \mathcal{M}_1(\mathbb{R}^d))$  is a unique weak solution of

$$\begin{aligned} \partial_t \mu_t &= \nabla \cdot ((\mathcal{P}[\mu_t] + f) \mu_t), \quad (x, t) \in \mathbb{R}^d \times [0, T], \\ \mathcal{P}[\mu_t](x) &= \int_{\mathbb{R}^d} P(x, y)(y - x) \mu_t(dy). \end{aligned} \quad (2.6)$$

For the convergence from (2.3)–(2.4) to (2.5)–(2.6), we need a weak compactness result in  $\mathcal{F}_\ell$  whose proof can be found in [51, Corollary 2.7].

**Lemma 2.2** *Let  $p \in (1, \infty)$ . Suppose that  $(f_j)_{j \in \mathbb{N}} \in \mathcal{F}_\ell$  with  $\ell \in L^q(0, T)$  for  $1 \leq q < \infty$ . Then there exists a subsequence  $(f_{j_k})_{k \in \mathbb{N}}$  and a function  $f \in \mathcal{F}_\ell$  such that*

$$f_{j_k} \rightharpoonup f \text{ weakly* in } L^q(0, T; W^{1,p}(\mathbb{R}^d)) \text{ as } k \rightarrow \infty, \quad (2.7)$$

i.e.,

$$\begin{aligned} \lim_{k \rightarrow \infty} \int_0^T \int_{\mathbb{R}^d} \phi(x, t) (f_{j_k}(x, t) - f(x, t)) dx dt &= 0 \\ \text{for all } \phi &\in L^{q'}(0, T; W^{-1,p'}(\mathbb{R}^d)). \end{aligned}$$

Define the empirical measure  $\mu^N$  associated to the particle system (2.4) as

$$\mu_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{x_i(t)} \quad \text{for } t \geq 0.$$

Then we are now in a position to state our theorem on the mean field limit of the optimal control problem.

**Theorem 2.3** *Let  $T > 0$ . Suppose that  $P \in W^{1,\infty}(\mathbb{R}^{2d})$  and  $\Psi$  satisfies that there exist  $C \geq 0$  and  $1 \leq q < \infty$*

$$\text{Lip}(\Psi, B(0, R)) \leq CR^{q-1} \text{ for all } R > 0.$$

*Let  $\ell(t)$  be a fixed function in  $L^q(0, T)$ . Furthermore we assume that  $\{x_i^0\}_{i=1}^N \subset B(0, R_0)$  for  $R_0 > 0$  independent of  $N$ . For all  $N \in \mathbb{N}$ , let us denote the control function  $f_N \in \mathcal{F}_\ell$  as a solution of the finite dimensional optimal control problem (2.3)–(2.4). If there exists a compactly supported initial data  $\mu_0 \in \mathcal{M}_1(\mathbb{R}^d)$  such that  $\lim_{N \rightarrow \infty} \mathcal{W}_1(\mu_0^N, \mu_0) = 0$ , then there exists a subsequence  $(f_t^{N_k})_{k \in \mathbb{N}}$  and a function  $f_t^\infty$  such that  $f_t^{N_k} \rightarrow f_t^\infty$  in the sense of (2.7). Moreover,  $f_t^\infty$  and the corresponding  $\mu_t^\infty$  are solutions of the infinite dimensional optimal control problem (2.5)–(2.6).*

*Proof* We first notice that the existence of an optimal control  $f_t^N$  on the time interval  $[0, T]$  for the finite dimensional optimal problem (2.3)–(2.4) can be obtained by using the weak compactness estimate in Lemma 2.2 together with the strong regularity of velocity field  $\mathcal{P} + f$ , see [51, Theorem 3.3]. For any  $f \in \mathcal{F}_\ell([0, T])$ , let us denote  $(\mu_f)_t^N$  by the solution to the equation (2.4) with the initial data  $(\mu_f)_0^N$  satisfying  $\lim_{N \rightarrow \infty} \mathcal{W}_1((\mu_f)_0^N, \mu_0) = 0$ . Let denote also by  $\mu_t^{f_t}$  is a solution associated to (2.6) with the control  $f_t$  and that initial data  $\mu_0$ , which is ensured by Theorem 2.1. Moreover, by Theorem 2.1,  $\lim_{N \rightarrow \infty} \mathcal{W}_1((\mu_f)_t^N, \mu_t^{f_t}) = 0$ . On the other hand, it follows from Lemma 2.2 that there exists a subsequence  $f_t^{N_k}$  such that  $f_t^{N_k} \rightharpoonup f_t^\infty$  weakly\* in  $L^q(0, T; W^{1,p}(\mathbb{R}^d))$  as  $k \rightarrow \infty$  for some  $f_t^\infty \in \mathcal{F}_\ell$ . Let  $\mu_t^\infty$  is the solution to (2.6) with the control function  $f_t^\infty$ . Then, by the lower-semicontinuity of the onset functional, we get

$$\liminf_{k \rightarrow \infty} J(\mu_t^{N_k}, f_t^{N_k}) \geq J(\mu_t^\infty, f_t^\infty),$$

where  $\mu_t^{N_k}$  is a solution to the particle equation (2.4) with the optimal control  $f_t^{N_k}$ . Then, due to the minimality of  $f_t^{N_k}$ , it is clear that

$$J((\mu_f)_t^{N_k}, f_t) \geq J(\mu_t^{N_k}, f_t^{N_k}) \text{ for each } k \in \mathbb{N}.$$

We finally use the convergence of  $\lim_{k \rightarrow \infty} \mathcal{W}_1((\mu_f)_t^{N_k}, \mu_t^f) = 0$  together with the compactly supported solution  $\mu_t$  to have

$$J(\mu_t^f, f_t) = \lim_{k \rightarrow \infty} J((\mu_f)_t^{N_k}, f_t) \geq \liminf_{k \rightarrow \infty} J(\mu_t^{N_k}, f_t^{N_k}) \geq J(\mu_t^\infty, f_t^\infty).$$

Since  $f_t$  is arbitrarily chosen in  $\mathcal{F}_\ell([0, T])$ , this concludes

$$\min_{f_t \in \mathcal{F}_\ell} J(\mu_t, f_t) = J(\mu_t^\infty, f_t^\infty),$$

i.e.,  $f_t^\infty$  is the optimal control for the problem (2.5)–(2.6). □

### 2.2 Stochastic Case

In this section, we study the parabolic optimal control problem in a bounded domain. In this section we are to a certain extent inspired by the work [20]. As we are deviating from that in certain estimates, we take the burden somehow of presenting the results in more details than in the previous section.

Let  $\Omega$  denote an open, bounded, smooth subset of  $\mathbb{R}^d$ . We first introduce function spaces:

$$V := L^2(0, T; H^1(\Omega)) \cap \dot{H}^1(0, T; H_*^{-1}(\Omega)), \quad \text{and} \quad H_*^{-1}(\Omega) = H^1(\Omega)',$$

and the set of admissible controls

$$Q_M := \left\{ \|f\|_{L^2(0,T;L^\infty(\Omega))} \leq M : f \in L^2(0, T; L^\infty(\Omega)) \right\},$$

for a given  $M > 0$ . Then our optimization problem is to show the existence of

$$\min_{f \in Q_M} J(\mu, f) := \min_{f \in Q_M} \int_0^T \left( \frac{1}{2} \int_\Omega |x - \bar{x}|^2 \mu(x, t) dx + \gamma \int_\Omega \Psi(f)\mu(x, t) dx \right) dt, \tag{2.8}$$

where  $\mu$  is a weak solution to the following parabolic equation:

$$\partial_t \mu + \nabla \cdot (\mathcal{P}[\mu]\mu + f\mu) = \sigma \Delta \mu, \quad (x, t) \in \Omega_T := \Omega \times [0, T], \tag{2.9}$$

with the initial data

$$\mu(\cdot, 0) = \mu_0(x) \quad x \in \Omega,$$

and the zero-flux boundary condition

$$\langle \sigma \nabla \mu - (\mathcal{P}[\mu] + f)\mu, n(x) \rangle = 0, \quad (x, t) \in \partial\Omega \times [0, T],$$

where  $n(x)$  is the outward normal to  $\partial\Omega$  at the point  $x \in \partial\Omega$ . Here the interaction term is given by

$$\mathcal{P}[\mu](x, t) = \int_\Omega P(x, y)(y - x)\mu(y, t) dy.$$

We next provide a notion of weak solution to Eq. (2.9).

**Definition 2.3** For a given  $T > 0$ , a function  $\mu : \Omega_T \rightarrow [0, \infty)$  is a weak solution of Eq. (2.9) on the time-interval  $[0, T]$  if and only if

1.  $\mu \in L^2(0, T; H^1(\Omega))$  and  $\partial_t \mu \in L^2(0, T; H_*^{-1}(\Omega))$ .

2. For any  $\varphi \in C^1(\overline{\Omega_T})$  with  $\varphi(\cdot, 0) = \varphi(\cdot, T) = 0$ ,

$$\int_0^T \int_{\Omega} \mu \partial_t \varphi + (\mathcal{P}[\mu]\mu + f\mu - \sigma \nabla \mu) \cdot \nabla \varphi \, dx dt = 0.$$

**Theorem 2.4** For a given  $T, M > 0$ , let  $f \in Q_M$  and  $\mu_0 \in L^2(\Omega)$ . Furthermore, we assume  $P \in L^\infty(\Omega^2)$ . Then there exists a unique weak solution  $\mu$  to Eq. (2.9) in the sense of Definition 2.3.

*Proof Existence.*- We first employ the following iteration scheme: Let  $\mu^1(x, t) := \mu_0(x)$  for  $(x, t) \in \Omega_T$ . For  $n \geq 1$ , let  $\mu^{n+1}$  be the solution of

$$\partial_t \mu^{n+1} + \nabla \cdot (\mathcal{P}[\mu^n]\mu^{n+1} + f\mu^{n+1}) = \sigma \Delta \mu^{n+1} \tag{2.10}$$

with the initial data  $\mu^n(x)|_{t=0} = \mu_0(x)$  for all  $n \geq 1, x \in \Omega$  and the zero-flux boundary conditions. It is clear that  $\int_{\Omega} \mu^n(x, t) \, dx = \int_{\Omega} \mu_0(x) \, dx$ . Note that for given  $\mu^n \in V$  we can have a unique weak solution to Eq. (2.10) since  $\mathcal{P}[\mu^n] \in L^\infty(\Omega)$  and  $f \in L^\infty(\Omega)$ . We next show that  $\mu^{n+1} \in V$ . A straightforward computation yields

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \int_{\Omega} (\mu^{n+1})^2 \, dx + \sigma \int_{\Omega} |\nabla \mu^{n+1}|^2 \, dx &= \int_{\Omega} \nabla \mu^{n+1} \cdot (\mathcal{P}[\mu^n]\mu^{n+1} + f\mu^{n+1}) \, dx \\ &=: I_1 + I_2, \end{aligned}$$

where  $I_2$  can be easily estimated as

$$I_2 \leq \int_{\Omega} |\nabla \mu^{n+1}| |f| \mu^{n+1} \, dx \leq \frac{\epsilon}{2} \int_{\Omega} |\nabla \mu^{n+1}|^2 \, dx + C_\epsilon \|f\|_{L^\infty}^2 \int_{\Omega} (\mu^{n+1})^2 \, dx.$$

For the estimate of  $I_1$ , we use the fact that

$$\|\mathcal{P}[\mu^n]\|_{L^\infty} \leq \text{diam}(\Omega) \|P\|_{L^\infty} \|\mu_0\|_{L^1} < \infty, \tag{2.11}$$

to obtain

$$|I_1| \leq \int_{\Omega} |\nabla \mu^{n+1}| |\mathcal{P}[\mu^n]| \mu^{n+1} \, dx \leq \frac{\epsilon}{2} \int_{\Omega} |\nabla \mu^{n+1}|^2 \, dx + C_\epsilon \int_{\Omega} (\mu^{n+1})^2 \, dx.$$

Combining the above estimates and choosing  $\epsilon < \sigma$ , we find

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} (\mu^{n+1})^2 \, dx + (\sigma - \epsilon) \int_{\Omega} |\nabla \mu^{n+1}|^2 \, dx \leq C_\epsilon (1 + \|f\|_{L^\infty}^2) \int_{\Omega} (\mu^{n+1})^2 \, dx.$$

Applying Gronwall’s inequality to the above differential inequality deduces

$$\int_{\Omega} (\mu^{n+1})^2 \, dx + \int_0^t \int_{\Omega} |\nabla \mu^{n+1}|^2 \, dx ds \leq C(T, \sigma, \|\mu_0\|_{L^2}, M). \tag{2.12}$$

We also get that for all  $\psi \in H^1(\Omega)$

$$\begin{aligned} \|\partial_t \mu^{n+1}\|_{H_*^{-1}} &= \sup_{\|\psi\|_{H^1} \leq 1} |\langle \partial_t \mu^{n+1}, \psi \rangle| \\ &\leq \sup_{\|\psi\|_{H^1} \leq 1} \left| \langle \mathcal{P}[\mu^n] \mu^{n+1} + f \mu^{n+1} + \sigma \nabla \mu^{n+1}, \nabla \psi \rangle \right| \\ &\leq (\|\mathcal{P}[\mu^n]\|_{L^\infty} + \|f\|_{L^\infty}) \|\mu^{n+1}\|_{L^2} + \sigma \|\nabla \mu^{n+1}\|_{L^2}. \end{aligned}$$

Thus we obtain  $\partial_t \mu^{n+1} \in L^2(0, T; H_*^{-1}(\Omega))$  due to (2.11) and (2.12). This concludes  $\mu^n \in V$  for all  $n \geq 2$ . Note that this also implies  $\mu^n \in C([0, T]; L^2(\Omega))$  for all  $n \geq 2$ . Indeed, we have

$$\max_{0 \leq t \leq T} \|\mu^n(t)\|_{L^2} \leq C \left( \|\mu^n\|_{L^2(0,T;H^1)} + \|\partial_t \mu^n\|_{L^2(0,T;H_*^{-1})} \right) \quad \text{for all } n \geq 2,$$

where  $C$  only depends on  $T$ . Then, by Aubin–Lions lemma, there exist a subsequence  $\mu^{n_k}$  and a function  $\mu \in L^2(\Omega_T)$  such that

$$\mu^{n_k} \rightarrow \mu \quad \text{in } L^2(\Omega_T) \quad \text{as } k \rightarrow \infty. \tag{2.13}$$

We next show that the above limiting function  $\mu$  solves Eq. (2.9) in the sense of Definition 2.3. For this, it suffices to take into account the interaction term  $\mathcal{P}[\mu]\mu$  since the other terms are linear with respect to  $\mu$ . Using the linearity of the functional  $\mathcal{P}$  together with (2.11) and the following fact

$$\|\mathcal{P}[f]\|_{L^\infty} \leq \text{diam}(\Omega) \|P\|_{L^\infty} \sqrt{|\Omega|} \|f\|_{L^2},$$

we get

$$\begin{aligned} &\int_0^T \int_\Omega |\mu^{n_{k+1}} \mathcal{P}[\mu^{n_k}] - \mu \mathcal{P}[\mu]|^2 dx dt \\ &\leq 2 \int_0^T \int_\Omega |\mu^{n_{k+1}} - \mu|^2 |\mathcal{P}[\mu^{n_k}]|^2 dx dt + 2 \int_0^T \int_\Omega \mu^2 |\mathcal{P}[\mu^{n_k} - \mu]|^2 dx dt \\ &\leq C_0 \int_0^T \int_\Omega |\mu^{n_{k+1}} - \mu|^2 + |\mu^{n_k} - \mu|^2 dx dt \rightarrow 0 \quad \text{as } k \rightarrow \infty, \end{aligned} \tag{2.14}$$

where  $C_0 > 0$  is given by

$$C_0 := 2 \text{diam}(\Omega)^2 \|P\|_{L^\infty}^2 \left( \|\mu_0\|_{L^1}^2 + |\Omega| \|\mu\|_{L^\infty(0,T;L^2)}^2 \right).$$

Hence we have that the limiting function  $\mu$  satisfies

$$\int_0^T \int_\Omega \mu \partial_t \varphi + (\mathcal{P}[\mu]\mu + f\mu - \sigma \nabla \mu) \cdot \nabla \varphi dx dt = 0.$$

*Uniqueness.*- Let  $\mu_i, i = 1, 2$  be two solutions to Eq. (2.9) with initial data  $\mu_i(0) \in L^2(\Omega)$ . Then, by using the similar estimate as in (2.14), we find

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \int_{\Omega} |\mu_1 - \mu_2|^2 dx + \sigma \int_{\Omega} |\nabla(\mu_1 - \mu_2)|^2 dx \\ &= \int_{\Omega} \nabla(\mu_1 - \mu_2) \cdot (\mathcal{P}[\mu_1 - \mu_2]\mu_1 + \mathcal{P}[\mu_2](\mu_1 - \mu_2) + f(\mu_1 - \mu_2)) dx \\ &\leq \epsilon \int_{\Omega} |\nabla(\mu_1 - \mu_2)|^2 dx + C_{\epsilon} \left(1 + \|f\|_{L^{\infty}}^2\right) \int_{\Omega} |\mu_1 - \mu_2|^2 dx, \end{aligned}$$

where  $C_{\epsilon}$  depends only on  $\Omega, \epsilon, \|\mu_1\|_{L^{\infty}(0,T;L^2)}$ , and  $\|\mu_2(0)\|_{L^1}$ . Finally, we apply the Gronwall’s inequality to the above differential inequality to get

$$\|\mu_1 - \mu_2\|_{L^{\infty}(0,T;L^2)}^2 + \|\nabla(\mu_1 - \mu_2)\|_{L^2(0,T;L^2)}^2 \leq C_1 \|\mu_1(0) - \mu_2(0)\|_{L^2}^2$$

where  $C_1$  depends only on  $T, \sigma, \|\mu_2(0)\|_{L^2}, M, \Omega$ , and  $\|\mu_1\|_{L^{\infty}(0,T;L^2)}$ . This completes the proof. □

**Theorem 2.5** *For a given  $T, M > 0$ , let us assume  $\mu_0 \in L^2(\Omega)$ . Furthermore, we assume that  $P \in L^{\infty}(\Omega^2)$  and  $\Psi$  satisfies that for all  $R > 0$*

$$W^{1,\infty}(\Psi, B(0, R)) \leq CR,$$

for some  $C > 0$ . Then there exist  $f^{\infty} \in Q_M$  and the corresponding density  $\mu^{\infty}$  solving the optimal control problem (2.8)–(2.9).

*Proof* For  $f \in Q_M$ , by Theorem 2.4, there exists a weak solution  $\mu$  in the sense of Definition 2.3. Note that  $0 \in Q_M$  and

$$J(\mu^0, 0) = \frac{1}{2} \int_0^T \int_{\Omega} |x - \bar{x}|^2 \mu(x, t) dx dt \leq C(T, \Omega) \|\mu_0\|_{L^1(\Omega)} \leq C,$$

where  $\mu^0$  is a weak solution of Eq. (2.9) with  $f = 0$ . Since  $J(\mu, f) \geq 0$  for all  $(\mu, f) \in V \times Q_M$ , there exist a sequence  $(f^j)_{j \in \mathbb{N}} \in Q_M$  and the corresponding density  $(\mu^j)_{j \in \mathbb{N}} \in V$  solving (2.9) such that

$$\lim_{j \rightarrow \infty} J(\mu^j, f^j) = \inf_{f \in Q_M} J(\mu, f).$$

On the other hand, since  $(\mu^j, f^j)_{j \in \mathbb{N}} \in V \times Q_M$ , by Banach–Alaoglu theorem, there exist a subsequence  $(\mu^{j_k}, f^{j_k}) \in V \times Q_M$  and  $(\mu^{\infty}, f^{\infty}) \in V \times Q_M$  such that

$$\mu^{j_k} \rightharpoonup \mu^{\infty} \text{ in } L^2(\Omega_T) \text{ and } f^{j_k} \overset{*}{\rightharpoonup} f^{\infty} \text{ in } L^2(0, T; L^{\infty}(\Omega)). \tag{2.15}$$

We next show that  $(\mu^{\infty}, f^{\infty})$  is a solution to (2.9). Since the term involving  $\mathcal{P}[\mu]$  can be easily handled by using the similar estimate to (2.14) and the above strong convergence (2.15), it is enough to show that

$$I_k := \int_0^T \int_{\Omega} (f^{j_k} \mu^{j_k} - f^{\infty} \mu^{\infty}) \phi \, dx dt \rightarrow 0 \text{ as } k \rightarrow \infty,$$

for  $\phi \in L^2(0, T; H^1(\Omega))$ . For this, we decompose  $I_k$  into two parts as

$$I_k = \int_0^T \int_{\Omega} (f^{j_k} - f^{\infty}) \mu^{j_k} \phi \, dx dt + \int_0^T \int_{\Omega} (\mu^{j_k} - \mu^{\infty}) f^{\infty} \phi \, dx dt =: I_k^1 + I_k^2.$$

Since

$$L^2(0, T; L^{\infty}(\Omega)) = \left( L^2(0, T; L^1(\Omega)) \right)' \text{ and } \mu^{j_k} \phi \in L^2(0, T; L^1(\Omega)),$$

it is clear from (2.15) that  $I_k^1 \rightarrow 0$  as  $k \rightarrow \infty$ . For the convergence of  $I_k^2$ , we get

$$\begin{aligned} I_k^2 &\leq \int_0^T \|f^{\infty}\|_{L^{\infty}} \|\mu^{j_k} - \mu^{\infty}\|_{L^2} \|\phi\|_{L^2} \, dt \\ &\leq \|\phi\|_{L^{\infty}(0, T; L^2)} \|f^{\infty}\|_{L^2(0, T; L^{\infty})} \|\mu^{j_k} - \mu^{\infty}\|_{L^2(0, T; L^2)} \rightarrow 0 \text{ as } k \rightarrow \infty. \end{aligned}$$

Thus we conclude that  $(\mu^{\infty}, f^{\infty})$  is a solution to (2.9). Furthermore, we obtain

$$\int_0^T \int_{\Omega} |x - \bar{x}|^2 \mu^{j_k} \, dx dt \rightarrow \int_0^T \int_{\Omega} |x - \bar{x}|^2 \mu^{\infty} \, dx dt \text{ as } k \rightarrow \infty,$$

due to  $|\Omega| < \infty$ . We also find

$$\lim_{k \rightarrow \infty} \int_0^T \int_{\Omega} \Psi(f^{j_k}) \mu^{j_k} \, dx dt \geq \int_0^T \int_{\Omega} \Psi(f^{\infty}) \mu^{\infty} \, dx dt. \tag{2.16}$$

More precisely, we estimate

$$\begin{aligned} &\int_0^T \int_{\Omega} (\Psi(f^{j_k}) \mu^{j_k} - \Psi(f^{\infty}) \mu^{\infty}) \, dx dt \\ &= \int_0^T \int_{\Omega} (\Psi(f^{j_k}) - \Psi(f^{\infty})) \mu^{j_k} + \Psi(f^{\infty}) (\mu^{j_k} - \mu^{\infty}) \, dx dt \\ &\geq \int_0^T \int_{\Omega} \nabla \Psi(f^{\infty}) \cdot (f^{j_k} - f^{\infty}) \mu^{j_k} \, dx dt + \int_0^T \int_{\Omega} \Psi(f^{\infty}) (\mu^{j_k} - \mu^{\infty}) \, dx dt \\ &=: J_k^1 + J_k^2, \end{aligned}$$

where we used the convexity of  $\Psi$  and the positivity of  $\mu^{j_k}$ . We then claim that  $\lim_{k \rightarrow \infty} J_k^1 \geq 0$  and  $\lim_{k \rightarrow \infty} J_k^2 \geq 0$ , and for this, we show that

$$\begin{aligned} &\nabla \Psi(f^{\infty}) \cdot f^{j_k} \mu^{j_k} \xrightarrow{*} \nabla \Psi(f^{\infty}) \cdot f^{\infty} \mu^{j_k} \\ &\text{and } \Psi(f^{\infty}) \mu^{j_k} \xrightarrow{*} \Psi(f^{\infty}) \mu^{\infty} \text{ in } \mathcal{M}(\Omega_T) \text{ as } k \rightarrow \infty. \end{aligned}$$

For  $\phi \in C_c(\Omega_T)$ , we get

$$\begin{aligned} \int_0^T \int_{\Omega} \Psi(f^\infty)(\mu^{j_k} - \mu^\infty)\phi \, dx dt &\leq C \|\phi\|_{L^\infty(\Omega_T)} \int_0^T \|f^\infty\|_{L^\infty} \|\mu^{j_k} - \mu^\infty\|_{L^2} \, dt \\ &\leq M \|\phi\|_{L^\infty(\Omega_T)} \|\mu^{j_k} - \mu^\infty\|_{L^2(\Omega_T)}, \end{aligned}$$

thus  $J_k^2 \rightarrow 0$  as  $k \rightarrow \infty$ . Similarly as before, for the estimate of  $J_k^1$ , we use the fact that  $\nabla \Psi(f^\infty) \mu^{j_k} \phi \in L^2(0, T; L^1(\Omega))$  uniformly in  $k$  and  $(L^2(0, T; L^1))' = L^2(0, T; L^\infty)$  to obtain  $J_k^1 \rightarrow 0$  as  $k \rightarrow \infty$ . Then, this and together with de la Vallée–Poussin’s theorem, provides the semicontinuity (2.16). This yields

$$\liminf_{k \rightarrow \infty} J(\mu^{j_k}, f^{j_k}) \geq J(\mu^\infty, f^\infty).$$

Hence we conclude

$$\inf_{f \in Q_M} J(\mu, f) = \lim_{j \rightarrow \infty} J(\mu^j, f^j) = \liminf_{k \rightarrow \infty} J(\mu^{j_k}, f^{j_k}) \geq J(\mu^\infty, f^\infty). \quad \square$$

### 3 First Order Optimality Conditions

In this section, we derive first order optimality conditions for the mean field optimal control problem studied in Sect. 2:

$$\partial_t \mu + \nabla \cdot ((\mathcal{P}[\mu] + f)\mu) = \sigma \Delta \mu, \quad x \in \Omega, \quad t > 0, \tag{3.1}$$

where the control  $f$  is the solution of the minimization of the following cost functional:

$$J(\mu, f) = \int_0^T \left( \frac{1}{2} \int_{\Omega} |x - \bar{x}|^2 \mu(x, t) \, dx + \gamma \int_{\Omega} \Psi(f)\mu(x, t) \, dx \right) dt. \tag{3.2}$$

#### 3.1 Formal Derivation of the Optimality Conditions

Let us first write the Lagrangian of the mean field optimal control defined by (3.1) and (3.2), as follows

$$\begin{aligned} \mathcal{L}(\mu, \psi, f) &= \int_0^T \left( \frac{1}{2} \int_{\Omega} |x - \bar{x}|^2 \mu(x, t) \, dx + \gamma \int_{\Omega} \Psi(f)\mu(x, t) \, dx \right) dt \\ &\quad - \int_0^T \left[ \int_{\Omega} \psi(x, t) (\partial_t \mu(x, t) + \nabla \cdot (\mathcal{P}[\mu(x, t)] \right. \\ &\quad \left. + f(x, t))\mu(x, t)) - \sigma \Delta \mu(x, t) \right] dx dt. \end{aligned} \tag{3.3}$$



Integrating by parts and taking the terminal data  $\psi(x, T) = 0$ , we get

$$\begin{aligned} \mathcal{L}(\mu, \psi, f) &= \int_0^T \left( \frac{1}{2} \int_{\Omega} |x - \bar{x}|^2 \mu \, dx + \gamma \int_{\Omega} \Psi(f) \mu \, dx \right) dt \\ &\quad + \int_{\Omega} \psi(x, 0) \mu(x, 0) \, dx + \int_0^T \int_{\Omega} \partial_t \psi \, \mu \, dx \, dt \\ &\quad + \int_0^T \int_{\Omega} \nabla \psi \cdot (\mathcal{P}[\mu] \mu) \, dx \, dt \\ &\quad + \int_0^T \int_{\Omega} \nabla \psi \cdot (f \mu) \, dx \, dt + \sigma \int_0^T \int_{\Omega} \mu \Delta \psi \, dx \, dt, \end{aligned} \tag{3.4}$$

where we omit the dependency on  $(x, t)$  where not necessary. We compute the functional derivatives of the Lagrangian with respect to the state function  $\mu$  and the control  $f$ ,

$$\frac{\delta \mathcal{L}}{\delta f} = \gamma \nabla \Psi(f) \mu + \nabla \psi \, \mu = (\gamma \nabla \Psi(f) + \nabla \psi) \mu, \tag{3.5}$$

$$\begin{aligned} \frac{\delta \mathcal{L}}{\delta \mu} &= \frac{1}{2} |x - \bar{x}|^2 + \gamma \Psi(f) + \partial_t \psi + \nabla \psi \cdot f + \sigma \Delta \psi \\ &\quad + \int_{\Omega} (P(x, y) \nabla \psi(x, t) - P(y, x) \nabla \psi(y, t)) \cdot (y - x) \mu(y, t) \, dy. \end{aligned} \tag{3.6}$$

Let  $(\mu^*, \psi^*, f^*)$  be the solution to the optimal control problem. Then we have

$$\left. \frac{\delta \mathcal{L}}{\delta f} \right|_{(\mu, \psi, f) = (\mu^*, \psi^*, f^*)} = 0 \quad \text{and} \quad \left. \frac{\delta \mathcal{L}}{\delta \mu} \right|_{(\mu, \psi, f) = (\mu^*, \psi^*, f^*)} = 0.$$

This yields from (3.5) that

$$\gamma \nabla \Psi(f^*) = -\nabla \psi^* \quad \text{on the support of } \mu^*. \tag{3.7}$$

We also find from (3.6) that  $\psi^*$  satisfies

$$\begin{aligned} \partial_t \psi^* + \frac{1}{2} |x - \bar{x}|^2 + \gamma \Psi(f^*) + \nabla \psi^* \cdot f^* + \sigma \Delta \psi^* \\ + \int_{\Omega} (P(x, y) \nabla \psi^*(x, t) - P(y, x) \nabla \psi^*(y, t)) \cdot (y - x) \mu^*(y, t) \, dy = 0, \end{aligned}$$

or equivalently

$$\begin{aligned} \partial_t \psi^* + \frac{1}{2} |x - \bar{x}|^2 + \gamma (\Psi(f^*) - \nabla \Psi(f^*) \cdot f^*) + \sigma \Delta \psi^* \\ + \int_{\Omega} (P(x, y) \nabla \psi^*(x, t) - P(y, x) \nabla \psi^*(y, t)) \cdot (y - x) \mu^*(y, t) \, dy = 0, \end{aligned} \tag{3.8}$$

due to (3.7), where  $\mu^*$  satisfies

$$\partial_t \mu^* + \nabla \cdot ((\mathcal{P}[\mu^*] + f^*)\mu^*) = \sigma \Delta \mu^* \quad \text{with} \quad \nabla \Psi(f^*) = -\frac{1}{\gamma} \nabla \psi^*.$$

### 3.2 Rigorous Derivation of the Optimality Conditions

The first order optimality conditions (3.10) are of utmost relevance as they are often used for the numerical computation of mean field optimal controls and we show how to proceed for that in Sect. 5. Although they are very often *formally* derived, as we do above, and used in several contributions, see, e.g. [15], as a relatively straightforward consequence of the Lagrange multiplier theorem, we feel that presenting their *rigorous* derivation can be useful for a reader not familiar with such derivations. Moreover, by doing so, we highlight more precisely certain technical difficulties and aspects, which one may in fact encounter along the process, and are often left to a certain extent as for granted. Let us recall then the Lagrange multiplier theorem in Banach spaces.

Let  $X$  and  $Y$  be Banach spaces, and let a functional  $J : U(x^*) \subseteq X \rightarrow \mathbb{R}$  and a mapping  $G : U(x^*) \subseteq X \rightarrow Y$  be continuously differentiable on an open neighbourhood of  $x^*$ . Consider the following optimal problem:

$$J(x) \rightarrow \inf, \quad G(x) = 0. \tag{3.9}$$

Then we recall the following first order optimality condition whose proof can be found in [86, Sect. 4.14].

**Theorem 3.1** *Let  $x^*$  be a solution to the problem (3.9), and let the range of the operator  $G'(x^*) : X \rightarrow Y$  be closed. Then there exists a nonzero pair  $(\lambda, p) \in \mathbb{R} \times Y'$  such that*

$$\mathcal{L}'_x(x^*, \lambda, p)(x) = 0 \quad \text{for all } x \in X,$$

where

$$\mathcal{L}(x, \lambda, p) = \lambda J(x) + G(x)(p).$$

Moreover; if  $\text{Im } G'(x^*) = Y$ , then  $\lambda \neq 0$  in the above, thus we can assume that  $\lambda = 1$ .

In order to apply the above theorem, we set

$$X = V \times L^2(\Omega_T), \quad Y = L^2(0, T; H^{-1}(\Omega)),$$

$$J(\mu, f) = \int_0^T \left( \frac{1}{2} \int_{\Omega} |x - \bar{x}|^2 \mu(x, t) dx + \gamma \int_{\Omega} \Psi(f)\mu(x, t) dx \right) dt,$$

and

$$G(\mu, f)(\psi) = - \int_0^T \int_{\Omega} \partial_t \psi \mu dxdt + \int_0^T \int_{\Omega} \nabla \psi \cdot (\mathcal{P}[\mu]\mu) dxdt$$

$$+ \int_0^T \int_{\Omega} \nabla \psi \cdot (f\mu) \, dxdt - \sigma \int_0^T \int_{\Omega} \nabla \mu \cdot \nabla \psi \, dxdt,$$

for  $\psi \in Y' = L^2(0, T; H_0^1(\Omega))$ . Then straightforward computations yield

$$\begin{aligned} G'_\mu(\mu, f)(v, \psi) &= - \int_0^T \int_{\Omega} \partial_t \psi \, v \, dxdt \\ &+ \int_0^T \int_{\Omega} \nabla \psi \cdot (\mathcal{P}[v]\mu + \mathcal{P}[\mu]v + fv) \, dxdt \\ &- \sigma \int_0^T \int_{\Omega} \nabla v \cdot \nabla \psi \, dxdt, \end{aligned}$$

for  $(v, \psi) \in V \times Y'$ , and

$$G'_f(\mu, f)(g, \psi) = \int_0^T \int_{\Omega} \nabla \psi \cdot (g\mu) \, dxdt \quad \text{for } (g, \psi) \in Q_M \times V'.$$

Note that the interaction terms on the right hand side of the equality for  $G'_\mu(\mu, f)(v, \psi)$  can be rewritten as

$$\begin{aligned} &\int_0^T \int_{\Omega} \nabla \psi \cdot (\mathcal{P}[v]\mu + \mathcal{P}[\mu]v) \, dxdt \\ &= \int_0^T \int_{\Omega^2} \nabla \psi(x) P(x, y) \cdot (y - x)(v(y)\mu(x) + \mu(y)v(x)) \, dxdydt \\ &= - \int_0^T \int_{\Omega^2} \nabla \psi(y) P(y, x) \cdot (y - x)(v(y)\mu(x) + \mu(y)v(x)) \, dxdydt \\ &= \frac{1}{2} \int_0^T \int_{\Omega^2} (P(x, y)\nabla \psi(x) - P(y, x)\nabla \psi(y)) \\ &\quad \cdot (y - x)(v(x)\mu(y) + \mu(x)v(y)) \, dxdydt. \end{aligned}$$

We now present our main result on the first order optimality condition in the theorem below.

**Theorem 3.2** *Let  $(\mu^*, f^*) \in V \times Q_M$  be a solution to the problem (3.1)–(3.2). Suppose that there exists a  $\mu_\ell > 0$  such that  $\mu^* \geq \mu_\ell$  for all  $(x, t) \in \Omega_T$ . Then there exists  $\psi^* \in Y'$  such that*

$$\begin{aligned} G'_\mu(\mu^*, f^*)(v, \psi^*) &= J'_\mu(\mu^*, f^*)(v), \quad \text{for all } v \in V, \\ G'_f(\mu^*, f^*)(g, \psi^*) &= J'_f(\mu^*, f^*)(g), \quad \text{for all } g \in L^2(\Omega_T). \end{aligned} \tag{3.10}$$

Before presenting the proof of the first order optimality conditions (3.10), let us comment the positivity principle on the existence of  $\mu_\ell > 0$  such that  $\mu^* \geq \mu_\ell$  for all  $(x, t) \in \Omega_T$ . If we assume that  $\mu_0, f, P \in C^2$  and  $\mu_0$  is bounded from below by

a positive constant, then by Feynman–Kac formula, we can show that  $\mu$  is bounded from below by some positive constant until the fixed time  $T$ . However, we a priori assume it to avoid any further stronger regularity assumption for the control  $f$ . Later, we will verify this property numerically in Sect. 5.

*Proof* For the proof, we show that linear operators  $G'_\mu(\mu^*, f^*) : V \rightarrow Y$  and  $G'_f(\mu^*, f^*) : L^2(\Omega_T) (\supseteq Q_M) \rightarrow Y$  are surjective. Then, by Theorem 3.1, we conclude our desired results.

**Surjectivity of  $G'_\mu(\mu^*, f^*)$ .** Let  $(\mu^*, f^*) \in V \times Q_M$  be a solution to (3.1)–(3.2). We want to show that for any  $\eta \in Y$  there exists a  $v \in V$  such that

$$G'_\mu(\mu^*, f^*)(v) = \eta, \quad \text{i.e.,} \quad G'_\mu(\mu^*, f^*)(v, \psi) = \eta(\psi) \quad \text{for all } \psi \in Y'.$$

Note that finding the above equality is equivalent to show that for given  $(\mu^*, f^*, \eta) \in V \times Q_M \times Y$ , there exists a solution  $v \in V$  to the Cauchy problem:

$$\partial_t v + \nabla \cdot (\mathcal{P}[v]\mu^* + \mathcal{P}[\mu^*]v + f^*v) = \sigma \Delta v - \eta, \quad x \in \Omega, \quad t > 0, \quad (3.11)$$

with the initial data  $v_0 \in L^2(\Omega)$  and the boundary condition:

$$\langle \sigma \nabla v - \mathcal{P}[v]\mu^* - (\mathcal{P}[\mu^*] + f^*)v, n(x) \rangle = 0, \quad (x, t) \in \partial\Omega \times \mathbb{R}_+.$$

We notice that (3.11) is linear parabolic equation of  $v$ . Thus the existence of  $v \in V$  is enough to show the following a priori estimates which are very similar to that in the proof of Theorem 2.4:

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|v\|_{L^2}^2 + \sigma \|\nabla v\|_{L^2}^2 \\ & \leq \|\nabla v\|_{L^2} (\|\mathcal{P}[v]\mu^*\|_{L^2} + \|\mathcal{P}[\mu^*]v\|_{L^2} + \|f^*v\|_{L^2}) + \|\eta\|_{H^{-1}} \|v\|_{H^1} \\ & \leq \frac{\sigma}{2} \|\nabla v\|_{L^2}^2 + C \left( \|\mathcal{P}[v]\|_{L^\infty}^2 \|\mu^*\|_{L^2}^2 + \left( \|\mathcal{P}[\mu^*]\|_{L^\infty}^2 + \|f^*\|_{L^\infty}^2 \right) \|v\|_{L^2}^2 \right) \\ & \quad + \|\eta\|_{H^{-1}}^2 + \|v\|_{L^2}^2 \\ & \leq \frac{\sigma}{2} \|\nabla v\|_{L^2}^2 + C \left( \|\mu^*\|_{L^2}^2 + \|f^*\|_{L^\infty}^2 + 1 \right) \|v\|_{L^2}^2 + \|\eta\|_{H^{-1}}^2, \\ & \quad \|\partial_t v\|_{H^{-1}} \leq \|\mathcal{P}[v]\|_{L^\infty} \|\mu^*\|_{L^2} + (\|\mathcal{P}[\mu^*]\|_{L^\infty} + \|f^*\|_{L^\infty}) \|v\|_{L^2} \\ & \quad + \sigma \|\nabla v\|_{L^2} + \|\eta\|_{H^{-1}} \\ & \lesssim (\|\mu^*\|_{L^2} + \|f^*\|_{L^\infty}) \|v\|_{L^2} + \sigma \|\nabla v\|_{L^2} + \|\eta\|_{H^{-1}}. \end{aligned}$$

Here we used

$$\|\mathcal{P}[v]\|_{L^\infty} \leq \text{diam}(\Omega) \sqrt{|\Omega|} \|P\|_{L^\infty} \|v\|_{L^2},$$

and similarly

$$\|\mathcal{P}[\mu^*]\|_{L^\infty} \leq \text{diam}(\Omega) \sqrt{|\Omega|} \|P\|_{L^\infty} \|\mu^*\|_{L^2}.$$

This yields

$$\begin{aligned} & \|v(\cdot, t)\|_{L^2}^2 + \int_0^t \|\nabla v(\cdot, s)\|_{L^2}^2 ds \\ & \leq \left( \|v_0\|_{L^2}^2 + \|\eta\|_{L^2(0,T;H^{-1})}^2 \right) \exp \left( C \int_0^T \left( \|\mu^*(\cdot, s)\|_{L^2}^2 + \|f^*(\cdot, s)\|_{L^\infty}^2 + 1 \right) ds \right) \end{aligned}$$

and

$$\begin{aligned} \|\partial_t v\|_{L^2(0,T;H^{-1})} & \lesssim \|v\|_{L^\infty(0,T;L^2)} \left( \|\mu^*\|_{L^2(\Omega_T)} + \|f^*\|_{L^2(0,T;L^\infty)} \right) \\ & \quad + \sigma \|\nabla v\|_{L^2(\Omega_T)} + \|\eta\|_{L^2(0,T;H^{-1})}. \end{aligned}$$

**Surjectivity of  $G'_f(\mu^*, f^*)$ .** For  $\xi \in Y$ , we first consider the following weak formulation of Poisson equation:

$$\int_0^t \int_\Omega \nabla \psi \cdot \nabla u \, dx ds = \int_0^t \int_\Omega \xi \psi \, dx ds, \quad \text{for any } \psi \in H_0^1(\Omega), \tag{3.12}$$

where we already took account the space–time decomposition of the test function. Note that solving Eq. (3.12) is equivalent to finding  $u \in L^2(0, T; H_0^1(\Omega))$  such that

$$a(u, v) = (f, v) \quad \text{for all } L^2(0, T; H_0^1(\Omega)),$$

with

$$a(u, v) := (\nabla u, \nabla v) = \int_0^T \int_\Omega \nabla u \cdot \nabla v \, dx dt,$$

and  $(\cdot, \cdot)$  is the inner product in  $L^2(\Omega_T)$ . Due to Poincaré inequality, we find that  $a(\cdot, \cdot)$  is an inner product on  $L^2(0, T; H_0^1(\Omega))$  with the induced norm:

$$\|v\|_{L^2(0,T;H_0^1)}^2 = a(v, v).$$

Define

$$F(v) := \int_0^T \int_\Omega f v \, dx dt \quad \text{for } v \in L^2(0, T; H_0^1(\Omega)).$$

Then this functional is continuous on  $L^2(0, T; H_0^1(\Omega))$  since  $|F(v)| \leq \|f\|_{L^2(0,T;H^{-1})} \|v\|_{L^2(0,T;H_0^1)}$ . Thus, by Riesz representation theorem, there exists a unique  $u \in L^2(0, T; H_0^1(\Omega))$  solving Eq. (3.12).

We now get back to our original problem. Our goal was to show that for given  $\mu^* \in V$  and  $\xi \in Y$ , there exists a function  $g \in L^2(\Omega_T)$  such that

$$\int_0^T \int_\Omega \nabla \psi \cdot (g \mu^*) \, dx dt = \int_0^T \int_\Omega \xi \psi \, dx dt \quad \text{for any } \psi \in Y'.$$

Then we now construct the solution  $g$  to the above equation by

$$g\mu^* = \nabla u, \quad \text{i.e., } g = \frac{\nabla u}{\mu^*} \quad \text{on the support of } \mu^*,$$

where the existence of  $u \in L^2(0, T; H_0^1(\Omega))$  was guaranteed in the beginning of the proof. Moreover, by the assumption  $\mu^*(x, t) > \mu_\ell > 0$  in  $\Omega \times [0, T]$ , we have

$$\begin{aligned} \int_0^T \int_\Omega |g(x, t)|^2 dx dt &= \int_0^T \int_\Omega \left| \frac{\nabla u(x, t)}{\mu^*(x, t)} \right|^2 dx dt \\ &\leq \frac{1}{\mu_\ell^2} \int_0^T \int_\Omega |\nabla u(x, t)|^2 dx dt < \infty, \end{aligned}$$

due to  $u \in L^2(0, T; H_0^1(\Omega))$ . This completes the proof.  $\square$

#### 4 Hierarchy of Controls via Boltzmann Equation

For large values of  $N$ , the solution of finite horizon control problems of the type (1.2)–(1.3) through standard methods stumble upon prohibitive computational costs, due to the nonlinear constraints and the lack of convexity in the cost. Although mean field optimal controls (1.4)–(1.6) are designed to be independent of the number  $N$  of agents to provide a way to circumvent the curse of dimensionality of  $N \rightarrow \infty$ , still their numerical computation needs to be realized by solving the first-order optimality conditions. The complexity of their solution depends on the intrinsic dimensionality  $d$  of the agents, which is affordable only at moderate dimensions (e.g.,  $d \leq 3$ ). In order to tackle these difficulties, we introduce a novel reduced setting, by introducing a binary dynamics whose evolution can be described by means of a Boltzmann-type equation, [3, 69]. Hence we will show that this description, under a proper scaling [80, 83], converges to the mean field equation (1.4), [6, 36, 80]. This type of approach allows to embed the control dynamics into two different ways:

- (i) we can assume the control  $f$  to be a given function, possibly obtained from the solution of the optimal control problem (1.2)–(1.3);
- (ii) alternatively, the control is obtained as a solution of the reduced optimal control problem associated to the dynamics of two single agents. We refer to this approach as *binary control*.

Similar ideas have been used in a control context in [5–7, 45, 49]. We devote the forthcoming sections to show different strategies to derive such binary controls. Thus we want to approach the mean field optimal control problem (1.2)–(1.3) as the last step of a control hierarchy, starting from an instantaneous control strategy and going towards a binary Hamilton–Jacobi–Bellman control.

### 4.1 Binary Controlled Dynamics

We consider the discrete controlled system (1.2)–(1.3) in the simplified case of only two interacting agents  $(x_i(t), x_j(t))$  and in absence of noise, i.e.  $\sigma = 0$ . Hence, by defining the sample time  $\Delta t$  such that  $t_m = m \Delta t$ , so that  $0 = t_0 < \dots < t_m < \dots < t_M = T$  and introducing a forward Euler discretization, we write (1.2) as follows

$$\begin{aligned} x_i^{m+1} &= x_i^m + \frac{\Delta t}{2} P(x_i^m, x_j^m)(x_j^m - x_i^m) + \Delta t u_i^m, \\ x_j^{m+1} &= x_j^m + \frac{\Delta t}{2} P(x_j^m, x_i^m)(x_i^m - x_j^m) + \Delta t u_j^m, \end{aligned} \tag{4.1}$$

where from now on we denote the control pair  $u := (u_i, u_j)$  associated to the state variable  $x := (x_i, x_j)$ , and having used the compact notation for  $x_i^m = x_i(t_m), u_i^m = u_i(t_m)$ .

The discretized form for the functional (1.3) for the binary dynamics (4.1) reads

$$J_M(x, u) := \sum_{m=0}^{M-1} \int_{t_m}^{t_{m+1}} L(x(t), u(t)) dt, \tag{4.2}$$

where the stage cost is given by

$$L(x, u) = \frac{1}{2} \left( |x_i - \bar{x}|^2 + |x_j - \bar{x}|^2 \right) + \gamma \left( \Psi(u_i) + \Psi(u_j) \right). \tag{4.3}$$

In the following we propose two alternative methods in order to characterize  $u_i, u_j$  as (sub-)optimal feedback controller. In both cases, we will consider the controlled dynamics in the deterministic case. Nonetheless, we will show in Sect. 5.3 that such controls are robust with respect to the presence of noise, ( $\sigma > 0$ ) and they shall be employed in the corresponding stochastic setting as well.

#### 4.1.1 Instantaneous Control

A first approach towards obtaining a low complexity computational realization of the solution of the optimal control problem (4.1)–(4.2) is the so-called *model predictive control* (MPC). This strategy furnishes a suboptimal control by an iterative solution over a sequence of finite time steps, representing the *predictive horizon* [6, 8, 64]. Since we are only interested in instantaneous control strategies, we limit the MPC method to a single time prediction horizon, therefore we reduce the original optimization into the minimization on every time interval  $[t_m, t_{m+1}]$  of the following functional

$$\begin{aligned} J_{\Delta t}(x^m, u^m) &= \Delta t L(x(t_{m+1}), u(t_m)) = \Delta t \left( \frac{1}{2} \left( |x_i^{m+1} - \bar{x}|^2 + |x_j^{m+1} - \bar{x}|^2 \right) \right. \\ &\quad \left. + \gamma \left( \Psi(u_i^m) + \Psi(u_j^m) \right) \right). \end{aligned} \tag{4.4}$$

Note that from (4.1) we have that  $x^{m+1}$  depends linearly on  $u^m$ , thus

$$U_{ij}^m := U(x_i, x_j, t_m) = \arg \min_{u^m} J_{\Delta t}(x^m, u^m)$$

can be directly computed from the following system

$$\begin{aligned} \Delta t^2 U_{ij}^m + 2\gamma \nabla_{u_i} \Psi(U_{ij}^m) + \Delta t(x_i^m - \bar{x}) + \frac{\Delta t^2}{2} P(x_i^m, x_j^m)(x_j^m - x_i^m) &= 0, \\ \Delta t^2 U_{ji}^m + 2\gamma \nabla_{u_j} \Psi(U_{ji}^m) + \Delta t(x_j^m - \bar{x}) + \frac{\Delta t^2}{2} P(x_j^m, x_i^m)(x_i^m - x_j^m) &= 0. \end{aligned} \quad (4.5)$$

In the case of a quadratic penalization of the control, i.e.  $\Psi(c) := |c|^2/2$ , we can furnish the following explicit expression for the minimizers

$$\begin{aligned} U_{ij}^m &= \frac{\Delta t}{2\gamma + \Delta t^2} \left( (\bar{x} - x_i^m) - \frac{\Delta t}{2} P(x_i^m, x_j^m)(x_j^m - x_i^m) \right), \\ U_{ji}^m &= \frac{\Delta t}{2\gamma + \Delta t^2} \left( (\bar{x} - x_j^m) - \frac{\Delta t}{2} P(x_j^m, x_i^m)(x_i^m - x_j^m) \right), \end{aligned} \quad (4.6)$$

hence (4.5) gives a feedback control for the full binary dynamics, which can be plugged as an *instantaneous control* into (4.1).

*Remark 4.1* Note that the instantaneous control (4.6) embedded into the discretized dynamics (4.1), is of order  $o(\Delta t)$ . To obtain an effective contribution of the control in the dynamics we will assume that the penalization parameter  $\gamma$  scales with the time discretization, in this way the leading order is recovered, [6, 8], e.g. for  $\gamma = \Delta t \bar{\gamma}$  we have

$$U_{ij}^m = \frac{1}{2\bar{\gamma} + \Delta t} \left( (\bar{x} - x_i^m) - \frac{\Delta t}{2} P(x_i^m, x_j^m)(x_j^m - x_i^m) \right). \quad (4.7)$$

#### 4.1.2 Finite Horizon Optimal Control

The instantaneous feedback control derived in the previous section is the optimal control action for the binary system with a single step prediction horizon. An improved, yet more complex optimal feedback synthesis can be performed by considering an extended finite horizon control problem. Let us define the value function associated to the finite horizon discrete cost (4.2) as

$$V(x_i, x_j, t_m) := \inf_{u \in \mathcal{U}} \sum_{k=m}^{M-1} \Delta t L(x_i(t_k), x_j(t_k), u(t_k)), \quad \text{for } m = 0, \dots, M-1, \quad (4.8)$$



with terminal condition  $V(x_i, x_j, t_M) = 0$ . It is well-known that the application of the Dynamic Programming Principle [13] with the discrete time dynamics (4.1) characterizes the value function as the solution of the following recursive Bellman equation

$$\begin{aligned} V(x_i, x_j, t_M) &= 0, \\ V(x_i, x_j, t_m) &= \inf_{u \in \mathcal{U}} \{ \Delta t L(x_i, x_j, u) + V(x + \Delta t(F(x_i, x_j) + u), t_{m+1}) \}, \\ m &= M - 1, \dots, 0, \end{aligned} \tag{4.9}$$

where  $x = (x_i, x_j)$ ,  $u = (u_i, u_j)$ , and  $F(x_i, x_j) := (P(x_i, x_j)(x_j - x_i), P(x_i, x_j)(x_j - x_i))$ . Once this functional relation has been solved, for every time step the optimal control is recovered from the optimality condition as follows

$$U(x_i, x_j, t_m) = \arg \min_{u \in \mathcal{U}} \{ \Delta t L(x_i, x_j, u) + V(x + \Delta t(F(x_i, x_j) + u), t_{m+1}) \}. \tag{4.10}$$

As in the expression (4.5), this optimal control is also in feedback form, depending not only on the current states of binary system  $(x_i, x_j)$ , but also on the discrete time variable  $t_m$ .

*Remark 4.2* The system (4.9) is a first-order approximation of the Hamilton–Jacobi–Bellman equation

$$\partial_t V(x, t) + \inf_{u \in \mathcal{U}} \{ L(x, u) + \nabla V(x, t) \cdot [F(x) + u] \} = 0, \tag{4.11}$$

related to the continuous time optimal control problem. In fact, this latter equation corresponds to the adjoint (3.6) when the nonlocal integral terms are neglected, and therefore this approach although optimal for the binary system, cannot be expected to satisfy the optimality system (3.5)–(3.6) related to the mean field optimal control problem.

### 4.2 Boltzmann Description

We introduce now a Boltzmann framework in order to describe the statistical evolution of a system of agents ruled by binary interactions, [8, 69].

Let  $\mu(x, t)$  denote the kinetic density of agents in position  $x \in \Omega$  at time  $t \geq 0$ , such that the total mass is normalized

$$\rho(t) = \int_{\Omega} \mu(x, t) dx = 1,$$

and the time evolution of the density  $\mu$  is given as a balance between the bilinear gain and loss of the agents position due to the binary interaction. In a general formulation, we assume that two agents have positions  $x, y \in \Omega$  and modify their positions according to the following rule

$$x^* = x + \alpha P(x, y)(y - x) + \alpha U_{\alpha}(x, y, t) + \sqrt{2\alpha}\xi,$$

$$y^* = y + \alpha P(y, x)(x - y) + \alpha U_\alpha(y, x, t) + \sqrt{2\alpha}\zeta, \tag{4.12}$$

where  $(x^*, y^*)$  are the post-interaction positions, the parameter  $\alpha$  measures the influence strength of the different terms,  $(\xi, \zeta)$  is a vector of i.i.d. random variables with a symmetric distribution  $\Theta(\cdot)$  with zero mean and variance  $\sigma$ , and  $U_\alpha(x, y, t)$  indicates the forcing term due to the control dynamics.

We consider now a kinetic model for the evolution of the density  $\mu = \mu(x, t)$  of agents with  $x \in \mathbb{R}^d$  at time  $t \geq 0$  and ruled by the following Boltzmann-type equation

$$\partial_t \mu(x, t) = Q_\alpha(\mu, \mu)(x, t), \tag{4.13}$$

where the interaction operator  $Q_\alpha(\mu, \mu)$  in (4.13), accounts the loss and gain of agents in position  $x$  at time  $t$ , as follows

$$Q_\alpha(\mu, \mu)(x, t) = \mathbb{E} \left[ \int_\Omega \left( \mathcal{B}_* \frac{1}{\mathcal{J}_\alpha} \mu(x_*, t) \mu(y_*, t) - \mathcal{B} \mu(x, t) \mu(y, t) \right) dy \right], \tag{4.14}$$

where  $(x_*, y_*)$  are the pre-interaction positions that generate arrivals  $(x, y)$ . The bilinear operator  $Q_\alpha(\cdot, \cdot)$  includes the expectation value with respect to  $\xi^x$  and  $\xi^y$ , while  $\mathcal{J}_\alpha$  represents the Jacobian of the transformation  $(x, y) \rightarrow (x^*, y^*)$ , described by (4.12). Here  $\mathcal{B}_* = \mathcal{B}_{(x_*, y_*) \rightarrow (x, y)}$  and  $\mathcal{B} = \mathcal{B}_{(x, y) \rightarrow (x^*, y^*)}$  are the transition rate functions. More into the details we take into account

$$\mathcal{B}_{(x, y) \rightarrow (x^*, y^*)} = \eta \chi_\Omega(x^*) \chi_\Omega(y^*),$$

as the functions with an interaction rate  $\eta > 0$ , and where  $\chi_\Omega$  is the characteristic function of the domain  $\Omega$ . Note that in this case the transition functions depends on the relative position, similarly to [80], as we introduced a bounded domain  $\Omega$  into the dynamics. A major simplification occurs in the case the bounded domain is preserved by the binary interactions itself, therefore the transition is constant and the interaction operator (4.14) reads

$$Q_\alpha(\mu, \mu)(x, t) = \eta \mathbb{E} \left[ \int_\Omega \left( \frac{1}{\mathcal{J}_\alpha} \mu(x_*, t) \mu(y_*, t) - \mu(x, t) \mu(y, t) \right) dy \right]. \tag{4.15}$$

In [6,80] authors showed that in opinion dynamics binary interactions are able to preserve the boundary, according to the choice of a small support of the symmetric random variable  $\xi$  and introducing a suitable function  $D(x)$  acting as a local weight on the noise in (4.12).

In the next section we will perform the analysis of this model in the simplified case of  $\Omega = \mathbb{R}^d$  and constant rate of interaction  $\eta$ .

*Remark 4.3* Note that the binary dynamics (4.12) is equivalent to the Euler–Maruyama discretization for Eq. (1.2) in the two agents case

$$\begin{aligned} x_i^{m+1} &= x_i^m + \frac{\Delta t}{2} P(x_i^m, x_j^m)(x_j^m - x_i^m) + \Delta t U_{ij}^m + \sqrt{2\sigma} \Delta B_i^m, \\ x_j^{m+1} &= x_j^m + \frac{\Delta t}{2} P(x_j^m, x_i^m)(x_i^m - x_j^m) + \Delta t U_{ji}^m + \sqrt{2\sigma} \Delta B_j^m, \end{aligned} \tag{4.16}$$

where we impose that  $\alpha = \Delta t/2$ ,  $\alpha U_\alpha(x_i, x_j) = \Delta t U_{ij}^m$ , and  $\sqrt{2\alpha\xi} = \sqrt{2\sigma} \Delta B_i^m$  is a random variable normally distributed with zero mean value and variance  $\Delta t$ , for  $\Delta B_i^m$  defined as the  $\Delta B_i^m = B_i(t_{m+1}) - B_i(t_m)$ .

### 4.2.1 The Quasi-Invariant Limit

We consider now the Boltzmann operator (4.15) in the case  $\Omega = \mathbb{R}^d$ , and in order to obtain a more regular description we introduce the so-called *quasi-invariant interaction limit*, whose basic idea is considering a regime where interactions strength is low and frequency is high. This technique, analogous to the grazing collision limit in plasma physics, has been thoroughly studied in [83] and specifically for first order models in [36, 80], and allows to pass from Boltzmann equation (4.13) to a mean field equation of the Fokker–Planck-type, [5, 6]. In order to state the main result we start fixing some notation and terminology.

**Definition 4.1** (*Multi-index*) For any  $a \in \mathbb{N}^d$  we set  $|a| = \sum_{i=1}^d a_i$ , and for any function  $h \in C^q(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R})$ , with  $q \geq 0$  and any  $a \in \mathbb{N}^d$  such that  $|a| \leq q$ , we define for every  $(x, v) \in \mathbb{R}^d \times \mathbb{R}^d$

$$\partial_x^a h(x) := \frac{\partial^{|a|} h}{\partial^{a_1} x_1 \dots \partial^{a_d} x_d}(x),$$

with the convention that if  $a = (0, \dots, 0)$  then  $\partial_x^a h(x) := h(x)$ .

**Definition 4.2** (*Test functions*) We denote by  $\mathcal{T}_\delta$  the set of compactly supported functions  $\varphi$  from  $\mathbb{R}^d$  to  $\mathbb{R}$  such that for any multi-index  $a \in \mathbb{N}^d$  we have,

1. if  $|a| < 2$ , then  $\partial_x^a \varphi(\cdot)$  is continuous for every  $x \in \mathbb{R}^d$ ;
2. if  $|a| = 2$ , then there exists  $C > 0$  such that,  $\partial_x^a \varphi(\cdot)$  is uniformly Hölder continuous of order  $\delta$  for every  $x \in \mathbb{R}^d$  with Hölder bound  $C$ , that is for every  $x, y \in \mathbb{R}^d$

$$\|\partial_x^a \varphi(x) - \partial_x^a \varphi(y)\| \leq C \|x - y\|^\delta,$$

and  $\|\partial_x^a \varphi(x)\| \leq C$  for every  $x \in \mathbb{R}^d$ .

**Definition 4.3** ( $\delta$ -weak solution) Let  $T > 0$ ,  $\delta > 0$ , we call a  $\delta$ -weak solution of the initial value problem for Eq. (4.13), with initial datum  $\mu^0 = \mu(x, 0) \in \mathcal{M}_0(\mathbb{R}^d)$  in the interval  $[0, T]$ , if  $\mu \in L^2([0, T], \mathcal{M}_0(\mathbb{R}^d))$  such that,  $\mu(x, 0) = \mu^0(x)$  for every  $x \in \mathbb{R}^d$ , and there exists  $R_T > 0$  such that  $\text{supp}(\mu(t)) \subset B_{R_T}(0)$  for every  $t \in [0, T]$  and  $\mu$  satisfies the weak form of Eq. (4.13), i.e.,

$$\frac{d}{dt} \langle \mu, \varphi \rangle = \langle \mathcal{Q}_\alpha(\mu, \mu), \varphi \rangle, \tag{4.17}$$

for all  $t \in (0, T]$  and all  $\varphi \in \mathcal{T}_\delta$ , where

$$\langle Q_\alpha(\mu, \mu), \varphi \rangle = \mathbb{E} \left[ \int_{\mathbb{R}^{2d}} \eta (\varphi(x^*) - \varphi(x)) \mu(x)\mu(y) dx dy \right]. \tag{4.18}$$

Moreover, we assume that

- (a) the system (4.12) constitutes invertible changes of variables from  $(x, y)$  to  $(x^*, y^*)$ ;
- (b) there exists an integrable function  $K(x, y, t)$  such that the following limit is well defined

$$\lim_{\alpha \rightarrow 0} U_\alpha(x, y, t) = K(x, y, t). \tag{4.19}$$

In the case of instantaneous control of type (4.6), we can explicitly give an expression to the limit as  $K(x, y, t) = (\bar{x} - x)/\gamma$ .

We state the following theorem.

**Theorem 4.4** *Let us fix a control  $U_\alpha \in \mathcal{U}$  and  $\alpha \geq 0$ , and  $T > 0, \delta > 0, \varepsilon > 0$ , and assume that density  $\Theta \in \mathcal{M}_{2+\delta}(\mathbb{R}^d)$  and the function  $P(\cdot, \cdot) \in L^q_{loc}$  for  $q = 2, 2 + \delta$  and for every  $t \geq 0$ . We consider a  $\delta$ -weak solution  $\mu$  of Eq. (4.13) with initial datum  $\mu_0(x)$ . Thus introducing the following scaling*

$$\alpha = \varepsilon, \quad \eta = 1/\varepsilon, \tag{4.20}$$

for the binary interaction (4.12) and defining by  $\mu^\varepsilon(x, t)$  a solution for the scaled equation (4.13), for  $\varepsilon \rightarrow 0$   $\mu^\varepsilon(x, t)$  converges pointwise, up to a subsequence, to  $\mu(x, t)$  where  $\mu$  satisfies the following Fokker–Planck-type equation,

$$\partial_t \mu + \nabla \cdot ((\mathcal{P}[\mu] + \mathcal{K}[\mu])\mu) = \sigma \Delta \mu, \tag{4.21}$$

with initial data  $\mu_0(x) = \mu(x, 0)$  and where  $\mathcal{P}$  represents the interaction kernel (1.5) and  $\mathcal{K}$  is the control.

$$\mathcal{K}[\mu](x, t) = \int_{\mathbb{R}^d} K(x, y, t)\mu(y, t) dy. \tag{4.22}$$

with  $K(x, y, t)$  defined as in (4.19).

*Proof* • *Taylor approximation* We consider the weak formulation of the Boltzmann equation (4.17) and we expand  $\varphi(x^*)$  inside the operator (4.18) in Taylor series of  $x^* - x$  up to the second order, obtaining

$$\langle Q_\alpha(\mu, \mu), \varphi \rangle = T_1^\varphi + T_2^\varphi + R_1^\varphi, \tag{4.23}$$

where the first and second order terms are

$$T_1^\varphi := \eta \mathbb{E} \left[ \int_{\mathbb{R}^{2d}} \nabla \varphi(x) \cdot (x^* - x) \mu(x) \mu(y) dx dy \right], \tag{4.24}$$

$$T_2^\varphi := \frac{\eta}{2} \mathbb{E} \left[ \int_{\mathbb{R}^{2d}} \left( \sum_{i,j=1}^d \partial_x^{(i,j)} \varphi(x) (x^* - x)_i (x^* - x)_j \right) \mu(x) \mu(y) dx dy \right], \tag{4.25}$$

and  $R_1^\varphi(\varepsilon)$  is the reminder of the Taylor expansion, with a form

$$R_1^\varphi := \frac{\eta}{2} \mathbb{E} \left[ \int_{\mathbb{R}^{2d}} \left( \sum_{i,j=1}^d \left( \partial_x^{(i,j)} \varphi(x) - \partial_x^{(i,j)} \varphi(\bar{x}) \right) (x^* - x)_i (x^* - x)_j \right) \mu(x) \mu(y) dx dy \right],$$

with  $\bar{x} := (1 - \theta)x^* + \theta x$ , for some  $\theta \in [0, 1]$ . By using the relation given by the scaled interaction rule (4.12), i.e.

$$x^* - x = \alpha F_\alpha(x, y) + \sqrt{2\alpha} \xi$$

where for the sake of brevity we denoted  $F_\alpha(x, y) := P(x, y)(y - x) + U_\alpha(x, y)$ . Note that from the hypothesis it follows that  $F_\alpha \in L_{loc}^q$ . Thus we obtain

$$\begin{aligned} T_1^\varphi &= \eta \mathbb{E} \left[ \alpha \int_{\mathbb{R}^{2d}} \nabla \varphi(x) \cdot \left( F_\alpha(x, y) + \sqrt{2/\alpha} \xi \right) \mu(x) \mu(y) dx dy \right] \\ &= \eta \alpha \int_{\mathbb{R}^{2d}} \nabla \varphi(x) \cdot F_\alpha(x, y) \mu(x) \mu(y) dx dy \end{aligned}$$

where the noise term,  $\xi$  is canceled out since it has zero mean. For the same reason in the second order term  $T_2^\varphi$  all the mixed products between  $F_\alpha$  and  $\xi$  vanish, the same hold for all the crossing terms  $\xi_i \xi_j$  since  $\xi_i$  are supposed to be independent variables. Hence the only contribution we have reads

$$\begin{aligned} T_2^\varphi &= \frac{\eta}{2} \mathbb{E} \left[ \int_{\mathbb{R}^{2d}} \alpha^2 \left( \sum_{j=1}^d \partial_x^{(j,j)} \varphi(x) (F_\alpha(x, y)_j)^2 \right) \right. \\ &\quad \left. + \left( \sum_{j=1}^d \partial_x^{(j,j)} \varphi(x) (2\alpha \xi_j^2) \right) \mu(x) \mu(y) dx dy \right] \\ &= \eta \alpha \int_{\mathbb{R}^{2d}} \sigma \Delta \varphi(x) \mu(x) \mu(y) dx dy \end{aligned}$$

$$\begin{aligned}
 & + \frac{\eta\alpha^2}{2} \int_{\mathbb{R}^{2d}} \left( \sum_{j=1}^d \partial_x^{(j,j)} \varphi(x) (F_\alpha(x, y)_j)^2 \right) \mu(x)\mu(y) dx dy, \\
 & =: T_{22}^\varphi + R_2^\varphi.
 \end{aligned}$$

• *Quasi-invariant limit* We now introduce the scaling (4.20), for which we can substitute in the previous equations,  $\eta\alpha = 1$  and  $\eta\alpha^2 = \varepsilon$ , thus we have that terms  $T_1^\varphi$  and  $T_{22}^\varphi$  represent the leading order and  $R^\varphi(\varepsilon) := R_1^\varphi + R_2^\varphi$  a reminder, so we can recast the scaled expression (4.23) as follows

$$\int_{\mathbb{R}^{2d}} (\nabla\varphi \cdot F_\varepsilon(x, y) + \sigma \Delta\varphi(x)) \mu(x)\mu(y) dx dy + R^\varphi(\varepsilon). \tag{4.26}$$

Let us now consider the limit  $\varepsilon \rightarrow 0$ , assuming that for every  $\varphi \in \mathcal{T}_\delta$

$$\lim_{\varepsilon \rightarrow 0} R^\varphi(\varepsilon) = 0 \tag{4.27}$$

holds true, we have thanks to (4.19) and (4.26) that the weak scaled Boltzman equation (4.17) converges pointwise to the Fokker–Planck-type equation (4.21) as follows

$$\frac{d}{dt} \langle \mu, \varphi \rangle = \langle \mu, \nabla\varphi \cdot (\mathcal{P}[\mu] + \mathcal{K}[\mu]) + \sigma \Delta\varphi \rangle, \tag{4.28}$$

where the operators  $\mathcal{P}[\mu]$  and  $\mathcal{K}[\mu]$  are defined in (1.5) and (4.22). Since  $\varphi$  has compact support, Eq. (4.28) can be revert in strong form by means of integration by parts, we eventually obtain system (4.21).

• *Estimates for the reminder* In order to conclude the proof it is sufficient to show that the limit (4.27) for  $R^\varphi(\varepsilon)$  vanishes. From the definition of  $\bar{x}$  it follows that  $\|\bar{x} - x\| \leq \|x^* - x\|$ , then for every  $\varphi \in \mathcal{T}_\delta$  we have

$$\left\| \partial_x^{(i,j)} \varphi(x) - \partial_x^{(i,j)} \varphi(\bar{x}) \right\| \leq C \|\bar{x} - x\|^\delta \leq C \|x^* - x\|^\delta.$$

Hence for  $R_1^\varphi$  we get

$$\begin{aligned}
 \|R_1^\varphi\| & \leq \frac{C}{2\varepsilon} \mathbb{E} \left[ \int_{\mathbb{R}^{2d}} \|x^* - x\|^{2+\delta} \mu(x)\mu(y) dx dy \right] \\
 & = \frac{C}{2} \varepsilon^{1+\delta} \mathbb{E} \left[ \int_{\mathbb{R}^{2d}} \|F_\varepsilon(x, y) + \sqrt{2/\varepsilon} \xi\|^{2+\delta} \mu(x)\mu(y) dx dy \right]
 \end{aligned}$$

from the inequality  $|a + b|^{2+\delta} \leq 2^{2+2\delta}(|a|^{2+\delta} + |b|^{2+\delta})$  for some  $a, b$  we obtain

$$\begin{aligned}
 \|R_1^\varphi\| & \leq 2^{1+2\delta} C \\
 & \left( \varepsilon^{1+\delta} \int_{\mathbb{R}^{2d}} \|F_\varepsilon(x, y)\|^{2+\delta} \mu(x)\mu(y) dx dy + 2^{1+\delta/2} \varepsilon^{\delta/2} \mathbb{E} \left[ \|\xi\|^{2+\delta} \right] \right).
 \end{aligned}$$

Analogous computation can be yield for  $R_2^\varphi$  for which we have the following inequality

$$\|R_2^\varphi\| \leq \frac{\varepsilon C}{2} \int_{\mathbb{R}^{2d}} \|F_\varepsilon(x, y)\|^2 \mu(x)\mu(y) dx dy.$$

Since  $F_\varepsilon \in L^q_{loc}$  for  $q = 2, 2 + \delta$  and  $\Theta \in \mathcal{M}_{2+\delta}(\mathbb{R}^d)$  we can conclude that for  $\varepsilon \rightarrow 0$  the limit (4.27) holds true. □

*Remark 4.4* Note that in the case  $U_\alpha(x, y, t) = U_\alpha(x, t)$ , namely if the feedback control depends only by the position  $x$  of the agents at time  $t$ , then the kernel  $\mathcal{K}[\mu](x, t)$  reduces to  $K(x, t)$ . This observation holds also if we consider a sampling from the optimal control, i.e.  $U_\alpha(x, y, t) = f(x, t)$ , thus Eq. (4.21) becomes exactly the original equation (1.2).

### 5 Numerical Methods

In this section we are concerned with the development of numerical methods for the mean field optimal control problem (1.2)–(1.3). First, we present direct simulation Monte Carlo methods for the constrained Boltzmann-type model (4.13), and discuss the implementation of the binary feedback controllers introduced in Sect. 4.1. Next, we describe a sweeping algorithm based on the iterative solution of the optimality system, (3.1)–(3.8).

#### 5.1 Asymptotic Constrained Binary Algorithms

One of the most common approaches to solve Boltzmann-type equations is based on Monte Carlo methods. For this, we consider the initial value problem given by Eq. (4.13), in the grazing interaction regime (4.20), with initial data  $\mu(x, t = 0) = \mu_0(x)$ , as follows

$$\begin{cases} \frac{d}{dt} \mu(x, t) = \frac{1}{\varepsilon} [Q_\varepsilon^+(\mu, \mu)(x, t) - \mu(x, t)], \\ \mu(x, 0) = \mu_0(x). \end{cases} \tag{5.1}$$

Here we have made explicit the dependence of the interaction operator  $Q_\varepsilon(\cdot, \cdot)$  on the frequency of interactions  $1/\varepsilon$ , and decomposing it into its gain and loss parts according to (4.15). With  $Q_\varepsilon^+(\cdot, \cdot)$  we denote the gain part, which accounts the density of agents gained at position  $x$  after the binary interaction (4.12).

We tackle the Boltzmann-type equation (5.1) by means of a binary interaction algorithm [3, 69], where the basic idea is to solve the binary exchange of information described by (4.12), under the grazing interaction scaling (4.20), in order to obtain in the limit an approximate solution of the mean field equation (4.21). Note that the consistency of this procedure is given by Theorem 4.4.

Let us now consider a time interval  $[0, T]$  discretized in  $M_{tot}$  intervals of size  $\Delta t$ . We denote by  $\mu^m$  the approximation of  $\mu(x, m \Delta t)$ , thus the first order forward scheme of the scaled Boltzmann-type equation (5.1) reads

$$\mu^{m+1} = \left(1 - \frac{\Delta t}{\varepsilon}\right) \mu^m + \frac{\Delta t}{\varepsilon} Q_\varepsilon^+(\mu^m, \mu^m), \tag{5.2}$$

where, since  $\mu^m$  is a probability density, thanks to mass conservation, and also  $Q_\varepsilon^+(\mu^m, \mu^m)$  is a probability density. Under the restriction  $\Delta t \leq \varepsilon$ ,  $\mu^{m+1}$  is a probability density, since it is a convex combination of probability densities.

From a Monte Carlo point of view, Eq. (5.2) can be interpreted as follows: an individual with position  $x$  will not interact with other individuals with probability  $1 - \Delta t/\varepsilon$  and it will interact with others with probability  $\Delta t/\varepsilon$  according to the interaction law stated by  $Q_\varepsilon^+(\mu^m, \mu^m)$ . Note that, since we aim at small values of  $\varepsilon$  and we have to fulfill the condition  $\Delta t \leq \varepsilon$ , the natural choice is to take  $\Delta t = \varepsilon$ . At every time step, this choice maximizes the number of interactions among the agents.

For the numerical treatment of the operator  $Q_\varepsilon^+(\mu^m, \mu^m)$ , we have to account that, every interaction includes action of the feedback control. In the case of *instantaneous control* this can be evaluated directly, for example in the case of quadratic functional defining the scaling version of (4.7) as

$$U_\varepsilon(x, y, t) = \frac{1}{\gamma + \varepsilon} ((\bar{x} - x) + \alpha P(x, y)(y - x)).$$

On the other hand, the realization of the optimal feedback controller in the *finite horizon setting* requires the numerical approximation of the Bellman equation (4.9). This approximation is performed offline and only once, previous to the simulation of the mean field model. For a state space of moderate dimension, such as in our binary model, several numerical schemes for the approximation of Hamilton–Jacobi–Bellman equations are available, and we refer the reader to [47, Chap. 8] for a comprehensive description of the different available techniques. Since the binary model is already introduced in discrete time, a natural choice is to solve eq. (4.9) by means of a sequential semi-Lagrangian scheme, following the same guidelines as in the recent works [11, 48, 57]. Once the value function has been approximated, online feedback controllers can be implemented through the evaluation of the optimality condition (4.10).

We report in Algorithm 1 a stochastic procedure to solve (5.2), based on Nanbu’s method for plasma physics, [3, 16].

Where function  $\text{IROUND}(\cdot)$  denotes the integer stochastic rounding defined as

$$\text{IROUND}(x) = \begin{cases} [x] + 1, & \zeta < x - [x], \\ [x], & \text{elsewhere} \end{cases}$$

with  $\zeta$  a uniform  $[0, 1]$  random number and  $[\cdot]$  the integer part.

*Remark 5.1 (Efficiency)* In general, computing the interactions among a multi-agent system is a procedure of quadratic cost with respect to the number of agents, since every agent needs to evaluate its influence with every other. Note that with the proposed algorithm this cost becomes linear with respect to the number of samples introduced  $O(N_s)$ , since only binary interactions are accounted. A major difference compared to standard algorithms for Boltzmann equations is the way in which particles are sampled from  $Q_\varepsilon^+(\mu^m, \mu^m)$  which does not require the introduction of a space grid [16].



**Algorithm 1** *Asymptotic constrained binary algorithm*

0. Pre-compute the feedback control  $U_\varepsilon(x, y, t)$  on an appropriate discretized grid of the domain  $\Omega \times [0, T]$ .
1. Given  $N_s$  samples  $\{x_k^0\}_{k=1}^{N_s}$ , from the initial distribution  $\mu_0(x)$ ;
  - for  $m = 0$  to  $M_{tot} - 1$ 
    - a. set  $N_c = \text{ROUND}(N_s/2)$ ;
    - b. select  $N_c$  random pairs  $(i, j)$  uniformly without repetition among all possible pairs of individuals at time level  $t_m$ ;
    - c. evaluate  $P(x_i, x_j)$ ,  $P(x_j, x_i)$  and  $U_\varepsilon(x_i, x_j, t_m)$ ,  $U_\varepsilon(x_j, x_i, t_m)$ ;
    - d. compute the post-interaction position  $x_i^*$ ,  $x_j^*$  for each pair  $(i, j)$  using relations (4.12) and  $\xi_i, \xi_j$  sampled from a normal distribution  $\mathcal{N}(0, \sigma)$ ;
    - e. set  $x_i^{n+1} = x_i^*$ ,  $x_j^{n+1} = x_j^*$ .
  - end for

*Remark 5.2 (Accuracy)* The choice  $\Delta t = \varepsilon$  is optimal if  $\varepsilon$  is of the order of  $O(N_s^{-1/2})$ . Indeed, the accuracy of the method will not increase for smaller values of  $\Delta t$ , because the numerical error is dominated by the fluctuations of the Monte Carlo method. For further details we refer to [3, 69].

**5.2 Numerical Approximation of the Optimality Conditions**

As shown in Sect. 3, the solution of the mean field optimal control problem (3.1)–(3.2) satisfies the optimality system

$$\partial_t \mu = -\nabla \cdot ((\mathcal{P}[\mu] + f)\mu) + \sigma \Delta \mu, \tag{5.3}$$

$$\begin{aligned}
 -\partial_t \psi &= \frac{1}{2}|x - \bar{x}|^2 + \gamma \Psi(f) + \nabla \psi \cdot f + \sigma \Delta \psi \\
 &+ \int_{\Omega} (P(x, y)\nabla \psi(x, t) - P(y, x)\nabla \psi(y, t)) \cdot (y - x)\mu(y, t) dy, \tag{5.4}
 \end{aligned}$$

$$\nabla \Psi(f) = -\frac{1}{\gamma} \nabla \psi, \quad \mu(x, 0) = \mu_0(x), \quad \psi(x, T) = 0. \tag{5.5}$$

**5.2.1 Forward Equation**

In order to solve Eq. (5.3), we consider a first order forward scheme the time evolution and the Chang–Cooper scheme for the space discretization, [31]. The formulation is based on the finite volume approximation of the density  $\mu$  and  $f$ . Defining the operator  $\mathcal{G}[\mu, f] := \mathcal{F}[\mu, f] + \sigma \nabla \mu$ , with  $\mathcal{F}[\mu, f] = \mathcal{P}[\mu] + f$ , then we can write in the one-dimensional domain  $[-L, L]$  the (semi)-discretized equation (5.3) as

$$\frac{d}{dt} \mu_i(t) = \frac{\mathcal{G}_{i+1/2}[\mu, f] - \mathcal{G}_{i-1/2}[\mu, f]}{\delta x}, \quad \text{with} \quad \mu_i(t) = \frac{1}{\delta x} \int_{x-1/2}^{x+1/2} \mu(x, t) dx, \tag{5.6}$$

where we have introduced the uniform grid  $x_i = -L + i\delta x$ ,  $i = 0, \dots, N$ , with  $\delta x = 2L/N$ , and denoted by  $x_{i\pm 1/2} = x_i \pm \delta x/2$ . Thus, the operator  $\mathcal{G}_{i+1/2}[\mu, f]$  in

the case of constant diffusion  $\sigma$  reads

$$\mathcal{G}_{i+1/2}[\mu, f] = ((1 - \theta_{i+1/2})\mu_{i+1} + \theta_{i+1/2}\mu_i) \mathcal{F}[\mu_{i+1/2}, f_{i+1/2}] + \frac{\sigma(\mu_{i+1} - \mu_i)}{\delta x}, \tag{5.7}$$

where the weights  $\theta_{i+1/2}$  are in general depending on the solution and the parameters of Eq. (5.3). Hence the flux functions are defined as a combination of upwind and centered discretizations, and such that for  $\sigma = 0$  the scheme reduces to an upwind scheme, i.e.  $\theta_{i+1/2} = 0$ . The choice of the weights is the key point of the scheme (5.6), which allows to preserve steady state solutions and the non-negativity of the numerical density. We refer to [9, 19, 31] for the details on the properties and analysis of the Chang–Cooper scheme for similar Fokker–Planck models and to [75], and references therein, for applications to control problems.

Alternatively, scheme (5.2) furnishes a consistent method to solve the forward equation (5.3), which we expect to be more efficient for problems with high dimensionality, since it relies on a stochastic evaluation of the nonlocal operator  $\mathcal{P}[f]$ .

### 5.2.2 Backward Equation

The main difficulty of the integro-differential advection-reaction-diffusion equation (5.4) resides on the efficient approximation of the integral term. We follow a finite difference approach, which we describe in the following. First, with time parameter  $\delta t$  as in the forward problem, we consider the first-order temporal approximation

$$\begin{aligned} -\frac{\psi^m - \psi^{m+1}}{\delta t} &= \frac{1}{2}|x - \bar{x}|^2 + \gamma \Psi(f^{m+1}) \\ &+ \left( f^{m+1} + \int_{\Omega} P(x, y) \cdot (y - x) \mu^{m+1} dy \right) \cdot \nabla \psi^{m+1} \\ &+ \sigma \Delta \psi^{m+1} - \int_{\Omega} \left( P(y, x) \nabla_y \psi^{m+1} \right) \\ &\cdot (y - x) \mu^{m+1} dy, \quad m = 0, \dots, M \end{aligned}$$

where  $\psi^M = 0$ . At this level,  $f$ ,  $\mu$ , and  $\nabla \psi$  are treated as external data available at every discrete instance. In particular  $\nabla_y$  (inside the integral) is reconstructed by numerical differentiation. Then, the integral terms are evaluated with a Monte Carlo method generating  $M_s$  samples according to the distribution  $\mu$ , and values of  $\nabla_y \psi$  are obtained by interpolation of the reconstructed variable. The advection term is approximated with a space-dependent upwind scheme, and diffusion is approximated with centered differences.

### 5.2.3 Optimality Condition and Sweeping Iteration

Once the forward–backward system has been discretized, what remains is to establish a coupling procedure in order to find the solution of the optimality system matching both initial and terminal conditions. For this, a first possibility is to consider the full space–

time discretization of the forward–backward system, together with the optimality condition  $\nabla\Psi(f) = -\frac{1}{\gamma}\nabla\psi$ , and cast it as a large-scale set of nonlinear equations, which can be solved via a Newton method. This idea has been already successfully applied in the context of mean field games in [2]. We pursue a different approach that has proven to be equally effective, developed in [26], where the authors apply a sweeping algorithm, which in our setting reads as follows.

---

**Algorithm 2** *Sweeping algorithm*

---

- 0. Given initial guess  $f_0$ , tolerance  $tol$ , and  $i = 0$
  - 1. while  $\|f_i - f_{i-1}\| \leq tol$ 
    - a. Perform a forward solve (5.3) with data  $f_i$  for  $\mu_i$ ;
    - b. Perform a backward solve (5.4) with data  $f_i, \mu_i$ , for  $\psi_i$ ;
    - c. Update through  $\nabla\Psi(f_{i+1}) = -\frac{1}{\gamma}\nabla\psi_i$ ;
    - d. set  $i = i + 1$ .
  - end while
- 

Our numerical experience is consistent with what has been already reported in [26], in the sense that solutions satisfying the optimality system can be found after few sweeps. Convergence of a similar sweeping iteration in the context of mean-field games has been recently proven in [25]. An alternative approach is to follow a gradient-type method, as in [20].

**5.3 Numerical Experiments**

In order to validate our previous analysis we focus on models for opinion dynamics, [53,69,78,80], thus in the unidimensional case the state variable  $x \in [-L, L]$  represents the agent opinion with respect to two opposite opinions  $\{-L, +L\}$ , and the control  $f(x, t)$  can be interpreted as the strategy of a policy maker, [5,6].

Therefore we consider the following initial value problem

$$\partial_t\mu + \partial_x \left( \left( \int_{-L}^{+L} P(x, y)(y - x)\mu(y)dy + f \right) \mu \right) = \sigma\partial_x^2\mu, \quad \mu(x, 0) = \mu^0(x) \tag{5.8}$$

with no-flux boundary conditions, and where  $f$  denotes the control term, solution of

$$f = \arg \min_{g \in \mathcal{U}} \frac{1}{2} \int_0^T \int_{-L}^{+L} (|x - \bar{x}|^2 + \gamma g^2) \mu(x, t) dx dt, \tag{5.9}$$

where we consider a quadratic penalization of the control, i.e.  $\Psi(c) = |c|^2/2$ .

For different interaction kernels  $P(\cdot, \cdot)$ , we will study the performance of the proposed controllers  $f = f(x, t)$ , obtained through the following synthesis procedures: instantaneous control (IC), finite horizon (FH), and the sweeping algorithm (OC).

We report in Table 1 the choice of the algorithms and parameters, indicating for which method they have been used to compute (5.8)–(5.9).

**Table 1** Parameters choice for the various algorithms and optimization methods

	Algorithm	$N_s$	$\varepsilon = \delta t$	$\delta x$	Tol
IC/FH	Alg 1	$5 \times 10^5$	$2.5 \times 10^{-3}$	$2.5 \times 10^{-2}$	–
Uncontrolled/OC	Alg 2	–	$2.5 \times 10^{-3}$	$2.5 \times 10^{-2}$	$10^{-5}$

### 5.3.1 Test 1: Sznajd Model

We consider the Sznajd model, [10,78] for which the interaction operator  $P(\cdot, \cdot)$  in (5.8) is defined as follows

$$P(x, y) = \beta(1 - x^2), \quad (5.10)$$

for  $\beta$  a constant. Note that in this case the interaction kernel  $P(\cdot, \cdot)$  models the propensity of voters to change their opinions within the domain  $\Omega = [-1, 1]$ , and for values close to the extremal opinions  $\{-1, 1\}$  the influence is low, conversely for opinions close to zero the influence is high. The dynamics is such that for  $\beta > 0$  concentration of the density profile appears, whereas for  $\beta < 0$  separation occurs, namely concentration around  $x = 1$  and  $x = -1$ , see [10].

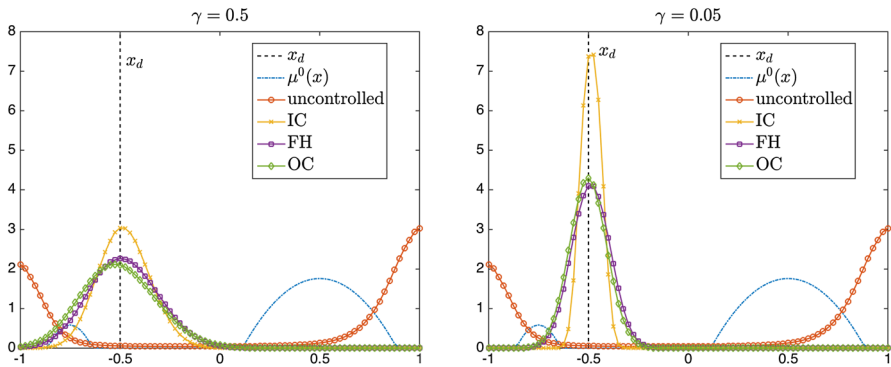
For our first test we fix  $\beta = -1$  and we define in the time interval  $[0, T]$ ,  $T = 8$ . We solve the control problem (5.8)–(5.9), with a bivariate initial data  $\mu^0(x) := \varrho_+(x + 0.75; 0.05, 0.5) + \varrho_+(x - 0.5; 0.15, 1)$ , where  $\varrho_+(y; a, b) := \max\{-(y/b)^2 + a, 0\}$ , with diffusion coefficient  $\sigma = 0.01$ , and desired state  $\bar{x} = -0.5$ .

In Fig. 1 we depict the final state of (5.10) at time  $T = 8$  for the uncontrolled and controlled dynamics. The simulations show the concentration of the profiles around the reference position  $\bar{x}$  in presence of the control, instead in the uncontrolled case the density tends to concentrate around the boundary. The left-hand side figure refers to a penalization of the control  $\gamma = 0.5$ , the right-hand side figure with  $\gamma = 0.05$ . As expected, with smaller control penalizations, the final state is driven closer to the desired reference.

Figure 2 illustrates the transient behavior of the density  $\mu(x, t)$  and the control  $f(x, t)$  in the  $[-1, +1] \times [0, T]$  frame, respectively for  $\gamma = 0.5$  and  $\gamma = 0.05$ , and we report the values of the cost function  $J(\mu, f)$  corresponding to the different methods. Note that that the action of the instantaneous control is almost constant in time steering the system toward  $\bar{x}$  but with the higher cost  $J(\mu, f)$ , on the other hand the optimal finite horizon for the binary dynamics (FH) produces a similar control with respect to the optimal control obtained by the sweeping algorithm (OC), with a small difference between the values of the cost functional.

### 5.3.2 Test 2: Hegselmann–Krause Model

In this second test we consider the mean field Hegselmann–Krause model [53], also known as bounded confidence model, whose interaction kernel reads



**Fig. 1** Test #1: Final states at time  $T = 8$  of the Sznajd model (5.10) for  $\beta = -1$  with initial data  $\mu^0(x)$ . Concentration around the desired state  $\bar{x}$  is observed in presence of the controls: instantaneous control (IC), finite horizon approach (FH), optimal control (OC), separation is observed in the uncontrolled setting. *Left figure  $\gamma = 0.5$ , right figure  $\gamma = 0.05$*

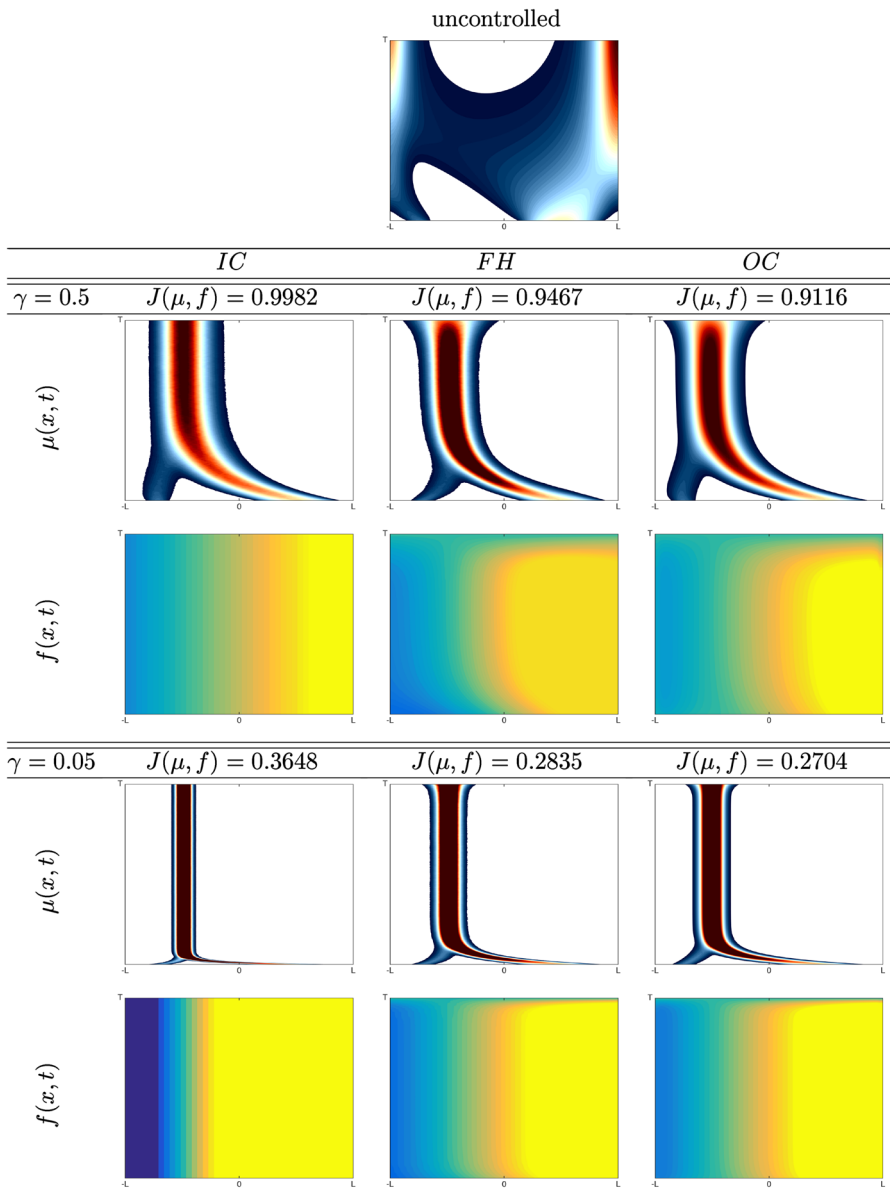
$$P(x, y) = \chi_{\{|x-y| \leq \kappa\}}(y). \tag{5.11}$$

This type of model describes the propensity of agents to interact only within a confidence range  $K = [x - \kappa, x + \kappa]$  of their opinion  $x$ , in the present experiment we fix  $\kappa = 0.15$ . Thus we study the evolution of the control problem (5.8)–(5.9) up to time  $T = 20$  with initial data defined as  $\mu^0(x) = C_0(0.5 + \epsilon(1 - x^2))$ , for  $\epsilon = 0.01$  and  $C_0$  such that the total density is a probability distribution. The diffusion coefficient is  $\sigma = 10^{-5}$ , the penalization parameter  $\gamma = 2.5$ , and the desired state  $\bar{x} = 0$ .

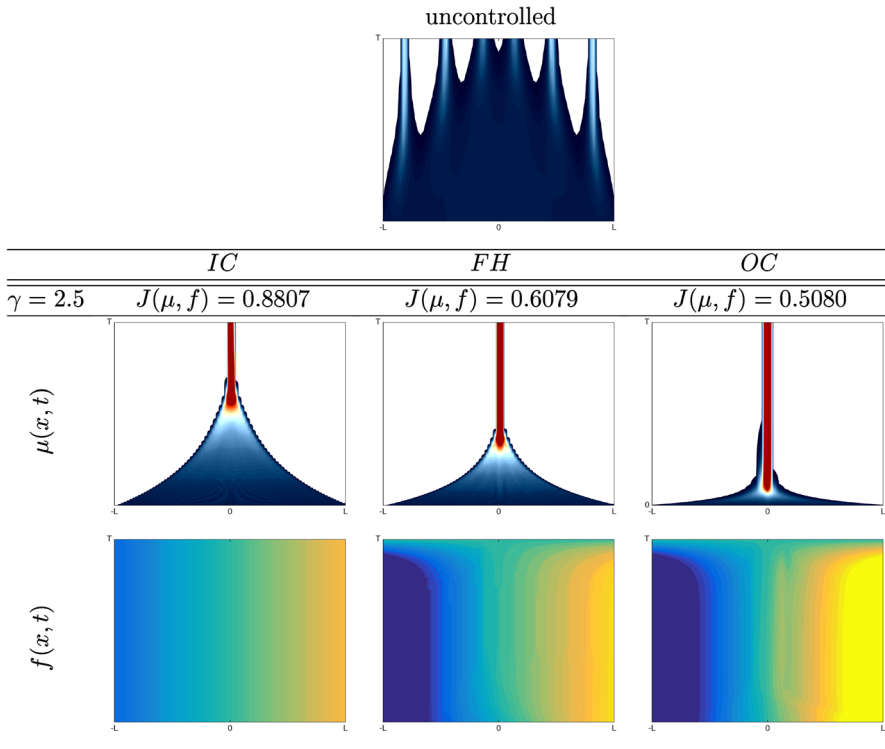
The uncontrolled evolution of this model shows the emergence of multiple clusters, as it is shown in the top picture of Fig. 3, due to the small value of  $\kappa$  and small diffusion. Figure 3 depicts the transient behavior of the density  $\mu(x, t)$  and the control signal  $f(x, t)$  in the frame  $\Omega \times [0, T]$ .

We observe in Fig. 3 that for the instantaneous control (IC), consensus is slowly reached with a cost functional value of  $J_{IC}(\mu, f) = 0.8807$ ; the finite horizon control (FH) and the solution of the optimality conditions (OC) are able to steer faster the system towards  $\bar{x}$ , respectively with cost  $J_{FH}(\mu, f) = 0.6079$ , and  $J_{OC} = 0.5080$ .

These experiments are showing very clearly the hierarchy of the controls (IC)  $\rightarrow$  (FH)  $\rightarrow$  (OC). In particular, it is evident the quasi-optimality of (FH), to the extent that we can claim (FH)  $\approx$  (OC). The intuition is that (FH) is an optimal control on the binary dynamics of two particles, and, through the Boltzmann collisional operator, its binary optimality is “smeared” over the entire population. However, we have no quantitative method yet to assess such an approximation. In fact, as commented in Remark 4.2, although the (FH) fulfills a Hamilton–Jacobi–Bellman equation, its synthesis by means of (4.22) to control (4.21) unfortunately does not fulfill the backward equation (5.4) of the optimality conditions, even not approximately: by testing (4.22) within (5.4), there a few useful cancellations, but, because of lack of symmetry, certain terms remains, whose magnitude is still hard to estimate. We expect that those terms are actually not so large and this would somehow justify the quasi-optimality of (FH). This issue remains an interesting open problem.



**Fig. 2** Test #1: Transient behavior of the density  $\mu(x, t)$  and the control  $f(x, t)$  in  $[-L, +L] \times [0, T]$ , with  $L = 1, T = 8$ , for the Sdnajz's model, (5.8)–(5.10). The top picture depicts the transient density of the unconstrained dynamics. Value of the cost functional are reported in correspondence of the choice of the method and the penalization parameter  $\gamma$



**Fig. 3** Test #2: Transient behavior of the density  $\mu(x, t)$  and the control  $f(x, t)$  in  $[-L, +L] \times [0, T]$ , with  $L = 1, T = 20$ , for the Hegelmann–Krause’s model, (5.8)–(5.9). The top picture shows the emergence of opinion clustering in the unconstrained dynamics. Value of the cost functional are reported in correspondence of the choice of the method with penalization parameter  $\gamma = 2.5$

### 6 Concluding Remarks

In this paper, we have presented a hierarchy of control designs for mean field dynamics. At the bottom of the hierarchy, we have introduced optimal feedback controls which are derived for two-agent models, and which are subsequently realized at the mean field level through a Boltzmann approach. At the top of the hierarchy, one finds the mean field optimal control problem and its correspondent optimality conditions. In both cases, we presented a theoretical and numerical analysis of the proposed designs, as well as computational implementations. From the numerical experiments presented in the last section, we observe that although the numerical realization of the mean field optimality system yields the best controller in terms of the cost functional value, feedback controllers obtained for the binary system perform reasonably well, and provide a much simpler control synthesis. We expect to further proceed along this direction of research, in particular in relation to the computation of feedback controllers via Dynamic Programming and Hamilton–Jacobi–Bellman equations for the binary system, as it provides a versatile framework to address different control problems.

**Acknowledgements** G.A., Y.P.C., and M.F. acknowledge the support of the ERC-Starting Grant HDSP-CONTR “High-Dimensional Sparse Optimal Control”. Y.P.C. is also supported by the Alexander Humboldt Foundation through the Humboldt Research Fellowship for Postdoctoral Researchers. D.K. acknowledges the support of the ERC-Advanced Grant OCLOC “From Open-Loop to Closed-Loop Optimal Control of PDEs”.

## References

- Achdou, Y., Laurière, M.: Mean field type control with congestion. *Appl. Math. Optim.* **73**(3), 393–418 (2016)
- Achdou, Y., Camilli, F., Capuzzo-Dolcetta, I.: Mean field games: numerical methods for the planning problem. *SIAM J. Control Optim.* **50**(1), 77–109 (2012)
- Albi, G., Pareschi, L.: Binary interaction algorithms for the simulation of flocking and swarming dynamics. *Multiscale Model. Simul.* **11**, 1–29 (2013)
- Albi, G., Pareschi, L.: Modeling of self-organized systems interacting with a few individuals: from microscopic to macroscopic dynamics. *Appl. Math. Lett.* **26**, 397–401 (2013)
- Albi, G., Pareschi, L., Zanella, M.: Boltzmann-type control of opinion consensus through leaders. *Philos. Trans. R. Soc. A* **372**, 20140138/1–20140138/18 (2014)
- Albi, G., Herty, M., Pareschi, L.: Kinetic description of optimal control problems and applications to opinion consensus. *Commun. Math. Sci.* **13**(6), 1407–1429 (2015)
- Albi, G., Bongini, M., Cristiani, E., Kalise, D.: Invisible control of self-organizing agents leaving unknown environments. *SIAM J. Appl. Math.* **76**, 1683–1710 (2016)
- Albi, G., Pareschi, L., Toscani, G., Zanella, M.: Recent advances in opinion modeling: control and social influence. In: Bellomo, N., Degond, P., Tadmor, E. (eds.) *Active Particles, Vol. 1: Theory, Methods, and Applications*. Birkhauser-Springer, Boston (2016)
- Albi, G., Pareschi, L., Zanella, M.: Opinion dynamics over complex networks: kinetic modeling and numerical methods. *Kinet. Relat. Mod.* **10**(1), 1–32 (2017)
- Aletti, G., Naldi, G., Toscani, G.: First-order continuous models of opinion formation. *SIAM J. Appl. Math.* **67**(3), 837–853 (2007)
- Alla, A., Falcone, M., Kalise, D.: An efficient policy iteration algorithm for dynamic programming equations. *SIAM J. Sci. Comput.* **37**(1), A181–A200 (2015)
- Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, L., Lecomte, L., Orlandi, A., Parisi, G., Procaccini, A., Viale, M., Zdravkovic, V.: Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. *PNAS* **105**(4), 1232–1237 (2008)
- Bellman, R., Kalaba, R.E.: *Dynamic Programming and Modern Control Theory*, vol. 81. Citeseer (1965)
- Benamou, J.-D., Brenier, Y.: A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.* **84**(3), 375–393 (2000)
- Bensoussan, A., Frehse, J., Yam, P.: *Mean Field Games and Mean Field Type Control Theory*. Springer, New York (2013)
- Bobylev, A., Nanbu, K.: Theory of collision algorithms for gases and plasmas based on the Boltzmann equation and the Landau-Fokker-Planck equation. *Phys. Rev. E* **61**(4), 4576 (2000)
- Bongini, M., Fornasier, M.: Sparse stabilization of dynamical systems driven by attraction and avoidance forces. *Netw. Heterog. Media* **9**(1), 1–31 (2014)
- Bongini, M., Fornasier, M.: Sparse control of multiagent systems. In: Bellomo, N., Degond, P., Tadmor, E. (eds.) *Active Particles, vol. 1: Theory, Methods, and Applications*. Birkhauser-Springer, Boston (2016)
- Buet, C., Dellacherie, S.: On the Chang and Cooper scheme applied to a linear Fokker-Planck equation. *Commun. Math. Sci.* **8**(4), 1079–1090 (2010)
- Burger, M., Francesco, M.D., Markowich, P.A., Wolfram, M.-T.: Mean field games with nonlinear mobilities in pedestrian dynamics. *Discret. Contin. Dyn. Syst. Ser. B* **19**(5), 1311–1333 (2014)
- Camazine, S., Deneubourg, J., Franks, N., Sneyd, J., Theraulaz, G., Bonabeau, E.: *Self-organization in Biological Systems*. Princeton University Press, Princeton (2003)
- Camilli, F., Jakobsen, E.R.: A finite element like scheme for integro-partial differential Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.* **47**(4), 2407–2431 (2009)



23. Cañizo, J.A., Carrillo, J.A., Rosado, J.: A well-posedness theory in measures for some kinetic models of collective motion. *Math. Models Methods Appl. Sci.* **21**(3), 515–539 (2011)
24. Caponigro, M., Fornasier, M., Piccoli, B., Trélat, E.: Sparse stabilization and optimal control of the Cucker–Smale model. *Math. Control Relat. Fields* **3**, 447–466 (2013)
25. Cardaliaguet, P., Hadikhhanloo, S.: Learning in mean field games: the fictitious play. *ESAIM: COCV* **23**(2), 569–591 (2017)
26. Carlini, E., Silva, F.J.: A fully discrete semi-Lagrangian scheme for a first order mean field game problem. *SIAM J. Numer. Anal.* **52**(1), 45–67 (2014)
27. Carrillo, J.A., D’Orsogna, M.R., Panferov, V.: Double milling in self-propelled swarms from kinetic theory. *Kinet. Relat. Models* **2**(2), 363–378 (2009)
28. Carrillo, J.A., Fornasier, M., Toscani, G., Vecil, F.: Particle, kinetic, and hydrodynamic models of swarming. In: Naldi, G., Pareschi, L., Toscani, G., Bellomo, N. (eds.) *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences. Modeling and Simulation in Science, Engineering and Technology*, pp. 297–336. Birkhäuser Boston, Boston (2010)
29. Carrillo, J.A., Choi, Y.-P., Hauray, M.: The derivation of swarming models: mean-field limit and Wasserstein distances. In: Muntean, A., Toschi, F. (eds.) *Collective Dynamics from Bacteria to Crowds. CISM International Centre for Mechanical Sciences*, pp. 1–46. Springer, New York (2014)
30. Carrillo, J.A., Choi, Y.-P., Pérez, S.: A review on attractive–repulsive hydrodynamics for consensus in collective behavior. In: Bellomo, N., Degond, P., Tadmor, E. (eds.) *Active Particles. Modeling and Simulation in Science, Engineering and Technology*, vol. 1, pp. 259–298. Birkhäuser, Cham (2017)
31. Chang, J., Cooper, G.: A practical difference scheme for Fokker–Planck equations. *J. Comput. Phys.* **6**(1), 1–16 (1970)
32. Choi, Y.-P.: Global classical solutions of the Vlasov–Fokker–Planck equation with local alignment forces. *Nonlinearity* **29**(7), 1887–1916 (2016)
33. Choi, Y.-P., Ha, S.-Y., Li, Z.: Emergent dynamics of the Cucker–Smale flocking model and its variants. In: Bellomo, N., Degond, P., Tadmor, E. (eds.) *Active Particles. Modeling and Simulation in Science, Engineering and Technology*, vol 1, pp. 299–331. Birkhäuser, Cham (2017)
34. Chuang, Y., D’Orsogna, M., Marthaler, D., Bertozzi, A., Chayes, L.: State transition and the continuum limit for the 2D interacting, self-propelled particle system. *Physica D* **232**, 33–47 (2007)
35. Chuang, Y., Huang, Y., D’Orsogna, M., Bertozzi, A.: Multi-vehicle flocking: scalability of cooperative control algorithms using pairwise potentials. In: *IEEE International Conference on Robotics and Automation*, pp. 2292–2299 (2007)
36. Cordier, S., Pareschi, L., Toscani, G.: On a kinetic model for a simple market economy. *J. Stat. Phys.* **120**(1–2), 253–277 (2005)
37. Couzin, I., Franks, N.: Self-organized lane formation and optimized traffic flow in army ants. *Proc. R. Soc. Lond. B* **270**, 139–146 (2002)
38. Couzin, I., Krause, J., Franks, N., Levin, S.: Effective leadership and decision making in animal groups on the move. *Nature* **433**, 513–516 (2005)
39. Cristiani, E., Piccoli, B., Tosin, A.: Modeling self-organization in pedestrians and animal groups from macroscopic and microscopic viewpoints. In: Naldi, G., Pareschi, L., Toscani, G., Bellomo, N. (eds.) *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences, Modeling and Simulation in Science. Engineering and Technology*. Birkhäuser Boston, Boston (2010)
40. Cristiani, E., Piccoli, B., Tosin, A.: Multiscale modeling of granular flows with application to crowd dynamics. *Multiscale Model. Simul.* **9**(1), 155–182 (2011)
41. Cucker, F., Dong, J.-G.: A general collision-avoiding flocking framework. *IEEE Trans. Autom. Control* **56**(5), 1124–1129 (2011)
42. Cucker, F., Mordecki, E.: Flocking in noisy environments. *J. Math. Pures Appl.* (9) **89**(3), 278–296 (2008)
43. Cucker, F., Smale, S.: Emergent behavior in flocks. *IEEE Trans. Autom. Control* **52**(5), 852–862 (2007)
44. Cucker, F., Smale, S., Zhou, D.: Modeling language evolution. *Found. Comput. Math.* **4**(5), 315–343 (2004)
45. Degond, P., Herty, M., Liu, J.-G.: Meanfield games and model predictive control. *Comm. Math. Sci.* **15**(5), 1403–1422 (2017)
46. Duan, R., Fornasier, M., Toscani, G.: A kinetic flocking model with diffusion. *Commun. Math. Phys.* **300**, 95–145 (2010)
47. Falcone, M., Ferretti, R.: *Semi-Lagrangian Approximation Schemes for Linear and Hamilton–Jacobi Equations*. Society for Industrial and Applied Mathematics, Philadelphia (2013)

48. Festa, A.: Reconstruction of independent sub-domains for a class of Hamilton–Jacobi equations and application to parallel computing. *ESAIM: M2AN* **50**(4), 1223–1240 (2016)
49. Festa, A., Wolfram, M.-T.: Collision avoidance in pedestrian dynamics. In: 2015 54th IEEE Conference on Decision and Control (CDC), pp. 3187–3192 (2015)
50. Filippov, A.F.: *Differential Equations with Discontinuous Righthand Sides*. Mathematics and Its Applications. Kluwer, Dordrecht (1988)
51. Fornasier, M., Solombrino, F.: Mean-field optimal control. *ESAIM Control Optim. Calc. Var.* **20**(4), 1123–1152 (2014)
52. Grégoire, G., Chaté, H.: Onset of collective and cohesive motion. *Phys. Rev. Lett.* **92**(2), 025702 (2004)
53. Hegselmann, R., Krause, U.: Opinion dynamics and bounded confidence: models, analysis and simulation. *J. Artif. Soc. Soc. Simul.* **5**(3), 1–33 (2002)
54. Hofbauer, J., Sigmund, K.: *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge (1998)
55. Huang, M., Caines, P., Malhamé, R.: Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions. In: Proceedings of the 42nd IEEE Conference on Decision and Control Maui, Hawaii, USA, December 2003, pp. 98–103 (2003)
56. Jadbabaie, A., Lin, J., Morse, A.S.: Correction to: “Coordination of groups of mobile autonomous agents using nearest neighbor rules”. *IEEE Trans. Autom. Control* **48**(9), 1675 (2003)
57. Kalise, D., Kröner, A., Kunisch, K.: Local minimization algorithms for dynamic programming equations. *SIAM J. Sci. Comput.* **38**(3), A1587–A1615 (2016)
58. Ke, J., Minett, J., Au, C.-P., Wang, W.-Y.: Self-organization and selection in the emergence of vocabulary. *Complexity* **7**, 41–54 (2002)
59. Keller, E.F., Segel, L.A.: Initiation of slime mold aggregation viewed as an instability. *J. Theor. Biol.* **26**(3), 399–415 (1970)
60. Koch, A., White, D.: The social lifestyle of myxobacteria. *Bioessays* **20**, 1030–1038 (1998)
61. Lacker, D.: Limit theory for controlled McKean-Vlasov dynamics. *SIAM J. Control Optim.* **55**(3), 1641–1672 (2017)
62. Lasry, J.-M., Lions, P.-L.: Mean field games. *Jpn. J. Math.* (3) **2**(1), 229–260 (2007)
63. Leonard, N., Fiorelli, E.: Virtual leaders, artificial potentials and coordinated control of groups. In: Proceeding of 40th IEEE Conference on Decision and Control, pp. 2968–2973 (2001)
64. Mayne, D.Q., Rawlings, J.B., Rao, C.V., Scokaert, P.O.M.: Constrained model predictive control: stability and optimality. *Autom. J. IFAC* **36**(6), 789–814 (2000)
65. Motsch, S., Tadmor, E.: Heterophilous dynamics enhances consensus. *SIAM Rev.* **56**(4), 577–621 (2014)
66. Niwa, H.: Self-organizing dynamic model of fish schooling. *J. Theor. Biol.* **171**, 123–136 (1994)
67. Nuorian, M., Caines, P., Malhamé, R.: Synthesis of Cucker–Smale type flocking via mean field stochastic control theory: Nash equilibria. In: Proceedings of the 48th Allerton Conference on Communication, Control, and Computing, Monticello, Illinois, pp. 814–819, September 2010, pp. 814–815 (2010)
68. Nuorian, M., Caines, P., Malhamé, R.: Mean field analysis of controlled Cucker–Smale type flocking: linear analysis and perturbation equations. In: Proceedings of 18th IFAC World Congress Milano (Italy), 28 August–2 September 2011, pp. 4471–4476 (2011)
69. Pareschi, L., Toscani, G.: *Interacting Multi-agent Systems*. Kinetic Equations & Monte Carlo Methods. Oxford University Press, Oxford (2013)
70. Parrish, J., Edelstein-Keshet, L.: Complexity, pattern, and evolutionary trade-offs in animal aggregation. *Science* **294**, 99–101 (1999)
71. Parrish, J., Viscido, S., Gruenbaum, D.: Self-organized fish schools: an examination of emergent properties. *Biol. Bull.* **202**, 296–305 (2002)
72. Perea, L., Gómez, G., Elosegui, P.: Extension of the Cucker–Smale control law to space flight formations. *AIAA J. Guid. Control Dyn.* **32**, 527–537 (2009)
73. Perthame, B.: *Transport Equations in Biology*. Birkhäuser, Basel (2007)
74. Romey, W.: Individual differences make a difference in the trajectories of simulated schools of fish. *Ecol. Model.* **92**, 65–77 (1996)
75. Roy, S., Annunziato, M., Borzi, A.: A Fokker-Planck feedback control-constrained approach for modeling crowd motion. *J. Comput. Theor. Transp.* **45**(6), 442–458 (2016)

76. Short, M.B., D'Orsogna, M.R., Pasour, V.B., Tita, G.E., Brantingham, P.J., Bertozzi, A.L., Chayes, L.B.: A statistical model of criminal behavior. *Math. Models Methods Appl. Sci.* **18**(suppl.), 1249–1267 (2008)
77. Sugawara, K., Sano, M.: Cooperative acceleration of task performance: foraging behavior of interacting multi-robots system. *Physica D* **100**, 343–354 (1997)
78. Sznajd-Weron, K., Sznajd, J.: Opinion evolution in closed community. *Int. J. Mod. Phys. C* **11**(06), 1157–1165 (2000)
79. Toner, J., Tu, Y.: Long-range order in a two-dimensional dynamical xy model: how birds fly together. *Phys. Rev. Lett.* **75**, 4326–4329 (1995)
80. Toscani, G.: Kinetic models of opinion formation. *Commun. Math. Sci.* **4**(3), 481–496 (2006)
81. Vicsek, T., Zafeiris, A.: Collective motion. *Phys. Rep.* **517**, 71–140 (2012)
82. Vicsek, T., Czirok, A., Ben-Jacob, E., Cohen, I., Shochet, O.: Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.* **75**, 1226–1229 (1995)
83. Villani, C.: On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. *Arch. Ration. Mech. Anal.* **143**(3), 273–307 (1998)
84. Villani, C.: Optimal Transport, vol. 338 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Berlin (2009)
85. Yates, C., Erban, R., Escudero, C., Couzin, L., Buhl, J., Kevrekidis, L., Maini, P., Sumpter, D.: Inherent noise can facilitate coherence in collective swarm motion. *Proc. Natl Acad. Sci. U.S.A.* **106**, 5464–5469 (2009)
86. Zeidler, E.: Applied Functional Analysis. Applied Mathematical Sciences. Springer, New York (1995)