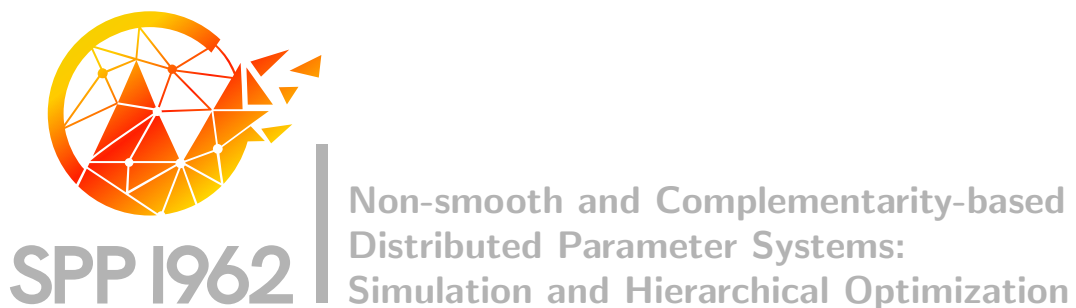


**DFG** Deutsche  
Forschungsgemeinschaft  
Priority Programme 1962

# *Multiscale Control of Stackelberg Games*

Michael Herty, Sonja Steffensen, Anna Thünen



Preprint Number SPP1962-150

received on November 6, 2020

Edited by  
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)  
Leibniz Institute in the Forschungsverbund Berlin e.V.  
Mohrenstraße 39, 10117 Berlin, Germany  
E-Mail: [spp1962@wias-berlin.de](mailto:spp1962@wias-berlin.de)

World Wide Web: <http://spp1962.wias-berlin.de/>

# Multiscale Control of Stackelberg Games

Michael Herty\*, Sonja Steffensen† and Anna Thünen‡

November 6, 2020

## Abstract

We present a linear–quadratic Stackelberg game with a large number of followers and we also derive the mean field limit of infinitely many followers. The relation between optimization and mean-field limit is studied and conditions for consistency are established. Finally, we propose a numerical method based on the derived models and present numerical results.

**Keywords:** Multi-level Game, Multiscale Control, Stackelberg Game, Nash Equilibrium, Mean-Field Game

**MSC(2020):** 82B40 91A65 49N80 91A16

## 1 Introduction

Game theory extends classical optimization by allowing for competing goals of possibly many actors. Early considerations and economic applications are described in [33]. A theoretical breakthrough has been made by Nash by formalizing the concept of equilibrium [30]. Stackelberg extended the models by putting one player in an special position, called the leader [34], establishing the class of Stackelberg games. In the last decades such multilevel games served as a tool for the analysis of systems of multiple competing interests and hierarchies. A prominent application is the analysis of electricity markets [6, 18, 21] using multi-leader follower games. Other applications include traffic and tolling [16, 22] as well as telecommunication [31, 35].

These applications usually involve modeling large populations of followers. For example, the demand of all customers for electricity is represented by one single independent system operator (ISO), which currently also provides a precise model for the current practice in energy markets, e.g. [8, 9, 10]. Many commuters in tolling models are modeled as one unit seeking a Wardrop equilibrium, whereby the interaction of these units does not play a role in road traffic, e.g. [16]. Similarly, in [35] the internet providers are modeled as individual leaders but data traffic is not further addressed.

We are interested in the study of Stackelberg games under possibly infinitely many followers. Models of interacting agents or followers have been studied e.g. in [12, 15, 32]. In particular, opinion formation and consensus as social models are discussed in [17, 28]. Other applications include economic and financial market models [29] as well as traffic models [19]. Game theoretic foundations of the analysis of these interacting agent systems discussed e.g. [23]. In [1] the control of a two-population model is investigated, where a population with a leading role is modeled through the dynamics. The agents of this leading population are however not leaders in the sense of game theory.

---

\*herty@igpm.rwth-aachen.de, RWTH Aachen University, Germany

†steffensen@igpm.rwth-aachen.de, RWTH Aachen University, Germany

‡thuenen@igpm.rwth-aachen.de, RWTH Aachen University, Germany

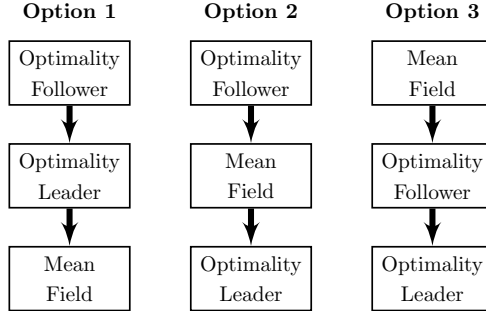


Figure 1: Schematic overview of the various order of operations for the Stackelberg game.

In this paper we consider a Stackelberg game with one leader and a possibly infinite number of followers. This population of followers is modeled as a dynamic system. We are interested in an equilibrium of the game which we characterize by first-order optimality conditions. We propose the following approach: The follower level is optimized first with the leader’s control as parameter, then the leader’s problem is solved provided certain regularity assumptions hold true, see e.g. [11]. The limit to infinitely many followers can be derived at different stages of the optimization yielding mean-field descriptions of the model. The focus of this article is the analysis of the interchangeability of optimization and derivation of the mean-field, see Figure 1. We establish consistency conditions for the Lagrange multipliers which link the different options. The novelty of our work lies not only in the two level problem but also in open loop controls compared, different to prior work studying feedback control techniques, c.f. [3, 4, 2, 5]. Also, compared to [20], we derive consistent optimality conditions for a Stackelberg game.

Other related work includes a linear quadratic Stackelberg game of a large follower population governed by stochastic differential equations in [27]. There, a local optimal control problem of the followers is solved where the control of the leader is considered as an exogenous stochastic process. This leads to  $\varepsilon$ -Nash equilibria and it is shown that  $\varepsilon \rightarrow 0$  as the number of followers grows to infinity. A related model is studied in [25] where it is distinguished between one major and a large number of minor players. In contrast to the work in [25, 27], we study a partial differential equation (PDE) on the probability density of the players’ states.

Here, we limit ourselves to formal computations in this article. Other approaches, rigorous derivations, and analytical results on derivatives with respect to measures may be found in [13, 14, 23].

This article is structured as follows: We begin with the derivation of consistent optimality conditions for two optimal control problems in Section 2. A model including a single control and a second model where each agent has an individual control, are studied there. We apply these results to a Stackelberg game in Section 3. In Section 4 a numerical scheme to solve the optimality system of Option 3 is derived. We conclude with some numerical results in Section 5.

## 2 Single Level Problems

In this section, we study two optimal control problems of interacting agents systems that differ in the nature of the application of the control. The problem in Section 2.1 is controlled by one control. In contrast, the problem discussed in Section 2.2 captures one control for each agent.

The optimality conditions to each optimal control problem can be derived prior to the derivation of the mean-field limit or after—resulting in two different optimality systems. We compare both systems and

establish consistency conditions to build a link between them, c.f. Lemma 2.1 and Lemma 2.2.

Note, the superscript MO indicates that the mean-field limit is derived prior to optimization. The superscript OM indicates the opposite order, see also Figure 1.

## 2.1 Single Control System

We consider the optimal control problem of a system of  $N$  interacting agents as follows:

$$\begin{aligned} \min_u \int_0^T \left[ J(u, m(\vec{x})) + \frac{\alpha}{2} u^2 \right] dt \\ \text{s.t. } \dot{x}_i = \frac{1}{N} \sum_{j=1}^N G(x_i, x_j, u), \quad x_i(0) = x_{i,0}, \quad i = 1, \dots, N, \end{aligned} \quad (1)$$

where the states  $x_i = x_i(t)$  are considered to be in  $\mathbb{R}^n$  for the agents  $i = 1, \dots, N$  and the according initial states are given by  $x_{i,0}$ . The concatenation of all agents' states is denoted by  $\vec{x} = (x_i)_{i=1}^N \in \mathbb{R}^{Nn}$ . The (common) control is  $u = u(t) \in \mathbb{R}^{n_u}$ . The explicit dependence on time  $t$  is omitted whenever the intention is clear. The agents' dynamics is described by  $G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}^n$  and is assumed to be at least differentiable in all arguments.

The common control  $u$  is to be chosen such that an objective functional is minimized over the time horizon  $[0, T]$ . The function  $J : \mathbb{R}^{n_u} \times \mathbb{R}^n \rightarrow \mathbb{R}$  takes the control  $u$  and  $m$  as arguments and is assumed to be differentiable. The value  $m = m(\vec{x})$  is considered to be a vector of a moment of the states, i.e.  $m : \mathbb{R}^{Nn} \rightarrow \mathbb{R}^n$  with

$$m(\vec{x}) = \frac{1}{N} \sum_{i=1}^N \tilde{m}(x_i)$$

and  $\tilde{m} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . The objective is regularized by a quadratic term of the control  $u$  with a scalar weighting parameter  $\alpha > 0$ .

Assuming the agents are identically independent and the interaction  $G$  is symmetric, we compute admissible variations with respect to the mean-field density  $h^{\text{MO}} = h^{\text{MO}}(t, x)$ . With e.g. [20, Proposition 2.1], we have the mean-field evolution equation of the state variables and the objective functional form the following optimal control problem in the strong form:

$$\begin{aligned} \min_u \int_0^T \left[ J(u, m_{h^{\text{MO}}}(t)) + \frac{\alpha}{2} u^2 \right] dt \\ \text{s.t. } 0 = \partial_t h^{\text{MO}} + \text{div}_x \left( h^{\text{MO}} \int G(x, \hat{x}, u) h^{\text{MO}}(t, \hat{x}) d\hat{x} \right) \\ h^{\text{MO}}(0, x) = h_0^{\text{MO}}(x), \end{aligned} \quad (2)$$

where  $m_{h^{\text{MO}}}(t) = \int \tilde{m}(x) h^{\text{MO}} dx$ . Formally, the dynamics and the cost of (1) and (2) are recovered for the mean-field probability density  $h^{\text{MO}}$  chosen as empirical distribution:

$$\mu_N(t, x) = \frac{1}{N} \sum_{i=1}^N \delta(x - x_i(t)), \quad (3)$$

where  $\delta$  denotes the Dirac delta. Similarly, the initial distribution  $h_0^{\text{MO}}$  is obtained as limit for  $N \rightarrow \infty$  of the empirical distribution centered at the initial data  $\vec{x}_0$ .

**Lemma 2.1** (Single Control System). *Consider the optimal control problem in (1) of  $N$  agents and the optimal control problem (2) for the density  $h^{\text{MO}} : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$  of agents. Let  $h^{\text{OM}} : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  be the distribution function, which satisfies the mean-field limit of the first-order the optimality conditions of problem (1). The multiplier to the optimality system of (2) is  $\Lambda = \Lambda(t, x)$ .*

*Then, the solution of the mean-field of the optimality conditions of (1) and the solution of the optimality conditions of (2) are formally identified by:*

$$\nabla_x \Lambda(t, x) = - \int \lambda h_2^{\text{OM}}(t, x, \lambda) \, d\lambda, \quad (4a)$$

for all  $t \geq 0$  and all  $x$  in the support of  $h^{\text{MO}}$ . The function  $h_2^{\text{OM}}$  is the marginal of  $h^{\text{OM}}$ :

$$h^{\text{OM}}(t, x, \lambda) = h^{\text{MO}}(t, x) h_2^{\text{OM}}(t, x, \lambda). \quad (4b)$$

Note, the consistency conditions in Lemma 2.1 state a decomposition of the probability density  $h^{\text{OM}}(t, x, \lambda)$  to the probability density  $h^{\text{MO}}(t, x)$  in (4b) and the Lagrange multipliers  $\lambda$  and  $\Lambda(t, x)$  in (4a).

The proof of Lemma 2.1 is omitted since it is analogous to the proof of Lemma 2.2 presented below.

## 2.2 Individual Control System

We consider the interacting agent system of  $N$  agents which reads as follows:

$$\begin{aligned} \min_u \quad & \frac{1}{N} \sum_{i=1}^N \int_0^T \left[ J(u_i, \tilde{m}(x_i)) + \frac{\alpha}{2} u_i^2 \right] dt \\ \text{s.t.} \quad & \dot{x}_i = \frac{1}{N} \sum_{j=1}^N G(x_i, x_j, u_i), \quad x_i(0) = x_{i,0}, \quad i = 1, \dots, N \end{aligned} \quad (5)$$

Contrary to Section 2.1, each agent  $i$  influences the model by its control  $u_i = u_i(t) \in \mathbb{R}^{n_u}$ . Formally, we obtain a mean-field optimal control problem as:

$$\begin{aligned} \min_u \quad & \int_0^T \int \left[ J(u, \tilde{m}(x)) + \frac{\alpha}{2} u^2 \right] f^{\text{MO}} \, dx \, dt \\ \text{s.t.} \quad & 0 = \partial_t f^{\text{MO}} + \text{div}_x \left( f^{\text{MO}} \int G(x, \hat{x}, u) f^{\text{MO}}(t, \hat{x}) \, d\hat{x} \right) \\ & f^{\text{MO}}(0, x) = f_0^{\text{MO}}(x) \end{aligned} \quad (6)$$

A difference in (6) compared to the problem in (2) is that the mean-field control  $u$  is additionally dependent on the state space, i.e.  $u = u(t, x)$ . If the mean-field density  $f^{\text{MO}}$  is chosen to be the empirical measure (3) then dynamics and cost of the problems (5) and (6) coincide if we define  $u_i(t) = u(t, x_i)$ .

As for the problem of a single control, we can derive consistency conditions which connect the optimality conditions of (5) and (6)

**Lemma 2.2** (Individual Control System). *Consider the optimal control problem in (5) of  $N$  agents and the optimal control problem (6) for the density  $f^{\text{MO}} : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$  of agents. Let  $f^{\text{OM}} : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  be the distribution function, which satisfies the mean-field limit of the first-order the optimality conditions to (5). The multiplier to the the optimality system of (6) is  $\Lambda = \Lambda(t, x)$ .*

*Then, the solution of the mean-field of the optimality conditions of (5) and the solution of the optimality conditions of (6) can be formally identified by:*

$$\nabla_x \Lambda(t, x) = - \int \lambda f_2^{\text{OM}}(t, x, \lambda) \, d\lambda, \quad (7a)$$

for all  $t \geq 0$  and all  $x$  in the support of  $f^{\text{MO}}$ . The function  $f_2^{\text{OM}}$  is the marginal of  $f^{\text{OM}}$ :

$$f^{\text{OM}}(t, x, \lambda) = f^{\text{MO}}(t, x) f_2^{\text{OM}}(t, x, \lambda). \quad (7b)$$

The proof of Lemma 2.2 is carried out in Section 2.3. In particular, the optimality conditions of (5) may be found in (9a-9c) and its mean-field limit in (10a-10b). The optimality conditions of (6) are given in (11a-11c).

**Corollary 2.3** (Parameterized Problems). *Consider the optimal control problem in (5) with the parameterized objective  $J = J(u_i, \tilde{m}(x_i); p)$ . Then Lemma 2.2 holds also for the parameterized objective.*

Before proving Lemma 2.2, we address an aspect related to the usage of  $L^2$  calculus for the formal computations of the optimality conditions of (6). In the proof, Gateaux derivatives of the Lagrangian are computed, see (11). In particular, the derivative with respect to the probability density  $f^{\text{MO}}$  is computed. Probability densities are nonnegative and their integral is one. A consistent derivative with respect to such a function conserves these properties also with the variation, e.g. in Wasserstein calculus. This means a suitable variation  $\eta$  of the probability density  $f^{\text{MO}}$  satisfies:

$$f^{\text{MO}}(t, x) + \eta(x) \geq 0 \quad \text{and} \quad \int_{\mathbb{R}^n} (f^{\text{MO}}(t, x) + \eta(x)) \, dx = 1, \quad (8)$$

which is not the case in  $L^2$  calculus. However, this relation is recovered by (7). Assume in the following paragraph that  $n = 1$ . The Lagrangian of the problem (6) contains the scalar product of the evolution equation of  $f^{\text{MO}}$  with the multiplier  $\Lambda$ :

$$\begin{aligned} \mathcal{L}^{\text{MO}}(f^{\text{MO}}, u, \Lambda) &= \int_0^T \int [J(u, \tilde{m}(x)) + \frac{\alpha}{2} u^2] f^{\text{MO}} \, dx \, dt \\ &\quad + \left\langle \partial_t f^{\text{MO}} + \text{div}_x \left( f^{\text{MO}} \int G(x, \hat{x}, u) f^{\text{MO}}(t, \hat{x}) \, d\hat{x} \right), \Lambda \right\rangle. \end{aligned}$$

If one uses now instead of the standard  $L^2$  scalar product the following scalar product:

$$\langle f^{\text{MO}}, \Lambda \rangle := \int f^{\text{MO}}(t, x) \partial_x \Lambda(t, x) \, dx,$$

we have that  $\partial_x \Lambda$  is a consistent variation of  $f^{\text{MO}}$  for compactly supported  $\Lambda$  since:

$$\int_{\mathbb{R}} \partial_x \Lambda(t, x) \, dx = 0$$

Hence the suitable test function in (8) is  $\eta(x) = \partial_x \Lambda(t, x)$ .

## 2.3 Proof of Lemma 2.2

We refer to [20] for a detailed discussion.

Notation. A function with an hat is evaluated in space or multiplier of the hat variable, e.g.  $\hat{f}^{\text{OM}} = f^{\text{OM}}(t, \hat{x}, \hat{\lambda})$  and  $\hat{f}^{\text{OM}} = f^{\text{OM}}(t, \hat{x})$ .

**First Optimize then Mean-Field Limit** Under regularity assumptions on the cost and the dynamics  $G$ , Pontryagin's maximum principle provides optimality conditions to (5).

The first-order optimality conditions are composed of the state dynamics and the dynamics of the Lagrange multipliers  $(\lambda_i)_{i=1}^N \in \mathbb{R}^{Nn}$ :

$$\dot{x}_i = \frac{1}{N} \sum_{j=1}^N G(x_i, x_j, u_i), \quad (9a)$$

$$\begin{aligned} \dot{\lambda}_i &= -D_x \tilde{m}(x_i)^\top \nabla_m J(u_i, \tilde{m}(x_i)) \\ &\quad - \frac{1}{N} \sum_{j=1}^N \left[ D_1 G(x_i, x_j, u_i)^\top \lambda_i + D_2 G(x_j, x_i, u_j)^\top \lambda_j \right], \end{aligned} \quad (9b)$$

with  $x_i(0) = x_{i,0}$  and  $\lambda_i(T) = 0$  for  $i = 1, \dots, N$  and in addition to that, the control is determined by:

$$0 = \nabla_u J(u_i, \tilde{m}(x_i)) + \alpha u_i + \frac{1}{N} \sum_{j=1}^N D_u G(x_i, x_j, u_i)^\top \lambda_i. \quad (9c)$$

In this paragraph, we derive evolution equation of the probability density  $f^{\text{OM}} = f^{\text{OM}}(t, x, \lambda)$ . To derive the mean field limit, we assume there exists an  $u : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n_u}$  such that:

$$u(t, x_i(t)) = u_i(t),$$

for  $t \geq 0$  and all  $i = 1, \dots, N$ . The mean-field equation related to the many particle limit of the dynamical system (9a-9b) is:

$$\begin{aligned} 0 = & \partial_t f^{\text{OM}} + \text{div}_x \left( f^{\text{OM}} \int G(x, \hat{x}, u) \hat{f}^{\text{OM}} d\hat{x} d\hat{\lambda} \right) \\ & - \text{div}_\lambda \left( f^{\text{OM}} \left[ \int \left[ D_1 G(x, \hat{x}, u)^\top \lambda + D_2 G(\hat{x}, x, \hat{u})^\top \hat{\lambda} \right] \hat{f}^{\text{OM}} d\hat{x} d\hat{\lambda} \right. \right. \\ & \left. \left. + D_x \tilde{m}(x)^\top \nabla_m J(u, \tilde{m}(x)) \right] \right), \end{aligned} \quad (10a)$$

with the initial condition  $f^{\text{OM}}(0, x, \lambda) = f_0^{\text{OM}}(x, \lambda)$  for all  $(x, \lambda)$ . The mean-field limit to (9c) is:

$$0 = \nabla_u J(u, m_{f^{\text{OM}}}(t)) + \alpha u + \int D_u G(x, \hat{x}, u)^\top \lambda \hat{f}^{\text{OM}} d\hat{x} d\hat{\lambda}, \quad (10b)$$

where  $m_{f^{\text{OM}}}(t) = \int \tilde{m}(x) f^{\text{OM}} dx d\lambda$ .

**First Mean-Field Limit then Optimize** The formal first-order optimality conditions of (6) in the  $L^2$ -sense are given by:

$$0 = \partial_t f^{\text{MO}} + \text{div}_x \left( f^{\text{MO}} \int G(x, \hat{x}, u) \hat{f}^{\text{MO}} d\hat{x} \right), \quad (11a)$$

$$\begin{aligned} 0 = & J(u, \tilde{m}(x)) + \frac{\alpha}{2} u^2 - \partial_t \Lambda \\ & - \int \left( G(x, \hat{x}, u)^\top \nabla_x \Lambda + G(\hat{x}, x, \hat{u})^\top \nabla_{\hat{x}} \hat{\Lambda} \right) \hat{f}^{\text{MO}} d\hat{x}, \end{aligned} \quad (11b)$$

$$0 = \nabla_u J(u, \tilde{m}(x)) + \alpha u - \int D_u G(x, \hat{x}, u)^\top \nabla_x \Lambda \hat{f}^{\text{MO}} d\hat{x}, \quad (11c)$$

with the initial value  $f^{\text{MO}}(0, x) = f_0^{\text{MO}}(x)$  and the terminal condition  $\Lambda(T, x) = 0$  for all  $x$ .



**The Relation between the Approaches** In this paragraph, we connect the two approaches and derive the consistency conditions (7). We may assume that there exists a decomposition such that:

$$f^{\text{OM}}(t, x, \lambda) = f_1^{\text{OM}}(t, x) f_2^{\text{OM}}(t, x, \lambda),$$

where for the conditional probability density  $\int h_2^{\text{OM}}(t, x, \lambda) d\lambda = 1$  holds. Upon multiplication of (10b) by  $f_2^{\text{OM}}$ , integration with respect to  $\lambda$  yields:

$$\begin{aligned} 0 = & \nabla_u J \left( u, m_{f_1^{\text{OM}} f_2^{\text{OM}}}(t) \right) \int f_2^{\text{OM}} d\lambda + \alpha u \int f_2^{\text{OM}} d\lambda \\ & + \int D_u G(x, \hat{x}, u)^\top \hat{f}_1^{\text{OM}} \hat{f}_2^{\text{OM}} d\hat{x} d\hat{\lambda} \int \lambda f_2^{\text{OM}} d\lambda, \end{aligned} \quad (12)$$

Hence, if:

$$f_1^{\text{OM}}(t, x) = f^{\text{MO}}(t, x), \quad (13a)$$

$$\int \lambda f_2^{\text{OM}} d\lambda = -\nabla_x \Lambda(t, x) \quad (13b)$$

then Equation (12) coincides with (11c).

Using the assumptions in (13), we obtain that after integration of the Equation (10a) with respect to  $\lambda$ :

$$0 = \partial_t \int f_1^{\text{OM}} f_2^{\text{OM}} d\lambda + \text{div}_x \int f_1^{\text{OM}} f_2^{\text{OM}} G(x, \hat{x}, u) \hat{f}_1^{\text{OM}} \hat{f}_2^{\text{OM}} d\hat{x} d\hat{\lambda} d\lambda,$$

which is equivalent to (11a).

We continue by multiplication of (10a) by  $\lambda$  and insertion of (13b), then integration by parts yields:

$$\begin{aligned} 0 = & -\partial_t (f_1^{\text{OM}} \nabla_x \Lambda) - \text{div}_x \left( f_1^{\text{OM}} \nabla_x \Lambda \int G(x, \hat{x}, u)^\top \hat{f}_1^{\text{OM}} d\hat{x} \right) \\ & + f_1^{\text{OM}} \iint \left[ D_1 G(x, \hat{x}, u)^\top \lambda + D_2 G(\hat{x}, x, \hat{u})^\top \hat{\lambda} \right] \hat{f}_1^{\text{OM}} \hat{f}_2^{\text{OM}} f_2^{\text{OM}} d\hat{x} d\hat{\lambda} d\lambda \\ & + f_1^{\text{OM}} D_x \tilde{m}(x)^\top \nabla_m J(u, \tilde{m}(x)). \end{aligned}$$

This equation is equivalent to (11b) when computing the gradient with respect to  $\nabla_x$ :

$$\begin{aligned} 0 = & -f_1^{\text{OM}} \left\{ \partial_t \nabla_x \Lambda + \nabla_x^2 \Lambda \int G(x, \hat{x}, u) \hat{f}_1^{\text{OM}} d\hat{x} + \int D_1 G(x, \hat{x}, u)^\top \nabla_x \Lambda \hat{f}_1^{\text{OM}} d\hat{x} \right. \\ & \left. + \int D_2 G(\hat{x}, x, \hat{u})^\top \nabla_{\hat{x}} \hat{\Lambda} \hat{f}_1^{\text{OM}} d\hat{x} - D_x \tilde{m}(x)^\top \nabla_m J(u, \tilde{m}(x)) \right\}. \end{aligned}$$

This completes the proof.

### 3 Stackelberg Game

We introduce a linear-quadratic Stackelberg game, consisting of the leader and  $N$  followers. The interchangeability of mean field limit and optimization is discussed using the results from Section 2.

The Stackelberg game to be discussed reads as follows:

$$\begin{aligned}
& \min_v \int_0^T \left[ J^L(v, m(\vec{\xi})) + \frac{\beta}{2} v^2 \right] dt \\
& \text{s.t. } \min_{\vec{w}, \vec{\xi}} \frac{1}{N} \sum_{i=1}^N \int_0^T \left[ J^F(\tilde{m}(\xi_i); v) + \frac{\gamma}{2} w_i^2 \right] dt \\
& \text{s.t. } \dot{\xi}_i = \frac{1}{N} \sum_{j=1}^N P(\xi_i, \xi_j) (\xi_j - \xi_i) + w_i, \quad \xi_i(0) = \xi_{i,0}, \quad i = 1, \dots, N,
\end{aligned} \tag{14}$$

where the leader minimizes the function  $J^L : \mathbb{R}^{n_L} \times \mathbb{R}^{n_F} \rightarrow \mathbb{R}$  regularized by a quadratic term of its control  $v = v(t) \in \mathbb{R}^{n_L}$  with  $\beta > 0$ . The structure of the followers' control problem is similar: Every follower  $i \in \{1, \dots, N\}$  aims to minimize  $J^F : \mathbb{R}^{n_F} \times \mathbb{R}^{n_L} \rightarrow \mathbb{R}$  regularized by its quadratic control  $w_i = w_i(t) \in \mathbb{R}^{n_F}$  with regularization parameter  $\gamma > 0$ . The value  $m = m(\vec{\xi})$  is considered to be a vector of moments of the states  $m : \mathbb{R}^{N n_L} \rightarrow \mathbb{R}^{n_L}$ .

The structure of the followers' problem is a potential game [26], i.e. here the followers' states and controls are not coupled in their objective functions. Therefore they can be summed up which yields cooperation between the followers.

The followers' problem is governed by an ordinary differential equation of every follower's state  $\xi_i = \xi_i(t) \in \mathbb{R}^{n_F}$  which couples to the other followers via the interaction kernel  $P : \mathbb{R}^{n_F} \times \mathbb{R}^{n_F} \rightarrow \mathbb{R}$ . The concatenation of the followers' states and controls is denoted by  $\vec{\xi} \in \mathbb{R}^{N n_F}$  and  $\vec{w} \in \mathbb{R}^{N n_F}$ .

Unlike the optimal control problems in Section 2, Stackelberg games have multiple levels of optimization. Therefore the optimality conditions have to be derived in a systematic order. This results in three different possibilities illustrated in Figure 1.

**Theorem 3.1.** *Consider the Stackelberg game in (14) of a single leader and  $N$  followers.*

*Denote by  $g^{\text{OOM}} : [0, T] \times \mathbb{R}^{n_F} \times \mathbb{R}^{n_F} \times \mathbb{R}^{2n_F} \rightarrow \mathbb{R}$  the probability density of the mean-field limit of the followers after optimization of both, the leader and the followers.*

*Furthermore, let  $g^{\text{OMO}} : [0, T] \times \mathbb{R}^{n_F} \times \mathbb{R}^{n_F} \rightarrow \mathbb{R}$  denote the probability density of the mean field limit after optimization of the leader and  $g^{\text{MOO}} : [0, T] \times \mathbb{R}^{n_F} \rightarrow \mathbb{R}$  the probability density of the followers prior to their optimization.*

*The function  $\Theta : [0, T] \times \mathbb{R}^{n_F} \times \mathbb{R}^{n_F} \rightarrow \mathbb{R}$  is the multiplier to  $g^{\text{OMO}}$  and the function  $\Psi : [0, T] \times \mathbb{R}^{n_F} \rightarrow \mathbb{R}$  is the multiplier to  $g^{\text{MOO}}$ .*

*If the condition:*

$$\int \psi^2 g_2^{\text{OMO}}(t, \xi, \psi) d\psi = \left( \int \psi g_2^{\text{OMO}}(t, \xi, \psi) d\psi \right)^2, \tag{15}$$

*holds for  $t \geq 0$  and all  $\xi \in \mathbb{R}^{n_F}$ , then the three optimality systems are equivalent in the mean-field limit and we have following relations:*

$$g^{\text{OOM}}(t, \xi, \psi, \theta) = g^{\text{OMO}}(t, \xi, \psi) g_2^{\text{OOM}}(t, \xi, \psi, \theta), \tag{16a}$$

$$\nabla_{\xi, \psi} \Theta(t, \xi, \psi) = \begin{bmatrix} \nabla_{\xi} \Theta(t, \xi, \psi) \\ \nabla_{\psi} \Theta(t, \xi, \psi) \end{bmatrix} = - \int \theta g_2^{\text{OOM}}(t, \xi, \psi, \theta) d\theta, \tag{16b}$$

for all  $t \geq 0$  and  $\xi$  in the support of  $g^{\text{OMO}}$  and:

$$g^{\text{OMO}}(t, \xi, \psi) = g^{\text{MOO}}(t, \xi) g_2^{\text{OMO}}(t, \xi, \psi), \quad (16c)$$

$$\nabla_{\xi} \Psi(t, \xi) = - \int \psi g_2^{\text{OMO}}(t, \xi, \psi) d\psi, \quad (16d)$$

for all  $t \geq 0$  and all  $\xi$  in the support of  $g^{\text{MOO}}$ .

Before proving this statement in Section 3.1, we provide additional interpretation.

The variance of a random variable  $X$  with the realization  $x$  and the probability density  $\rho$  is defined as  $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ , where  $\mathbb{E}[X]$  denotes the expectation of  $X$ , which is defined as  $\mathbb{E}[X] = \int x \rho(x) dx$ .

With this, condition (15) is equivalent to requiring that the variance of a random variable  $Y$  with the probability density  $y \mapsto g_2^{\text{OMO}}(t, \xi, y)$  is zero for all  $(t, \xi)$ , i.e. with the realization  $y$  of  $Y$  we have:

$$\text{Var}(Y) = \int y^2 g_2^{\text{OMO}}(t, \xi, y) dy - \left( \int y g_2^{\text{OMO}}(t, \xi, y) dy \right)^2 = 0.$$

That is e.g. the case if the probability density  $g_2^{\text{OMO}}(t, \xi, y)$  coincides with the empirical measure in  $y$  concentrated on  $\mathbb{E}[Y]$ , i.e.  $g_2^{\text{OMO}}(t, \xi, y) = \delta(y - \mathbb{E}[Y])$ .

In the proof for Theorem 3.1, it is shown that the optimal follower control is  $w(t, \xi) = \frac{1}{\gamma} \nabla_{\xi} \Psi$  which is in fact then:

$$w(t, \xi) = -\frac{1}{\gamma} \mathbb{E}[Y].$$

### 3.1 Proof of Theorem 3.1

The proof consists of six steps, in which we follow the graph of Figure 2. For notation convenience, arguments of functions are omitted.

*Step 1.* We begin the analysis on the follower level of the game in Equation (14). Under regularity assumptions, the followers' optimization problem allows using Pontryagin's maximum principle. With this, the optimal response of the followers to the leader can be characterized by a coupled system of ordinary differential equations (ODE). It consists of the state dynamic  $\xi_i$  and the dynamic of the dual  $\psi_i$  for every follower  $i = 1, \dots, N$ . The optimal controls  $w_i$  are available explicitly and substituted in the ODE system. Replacing the optimal control problem of the follower by the ODE system yields the leader's optimal control problem:

$$\begin{aligned} \min_v \int_0^T \left[ J^L(v, m(\vec{\xi})) + \frac{\beta}{2} v^2 \right] dt \\ \text{s.t. } \dot{\zeta}_i = \begin{bmatrix} \dot{\xi}_i \\ \dot{\psi}_i \end{bmatrix} = \frac{1}{N} \sum_{j=1}^N G(\zeta_i, \zeta_j, v), \quad i = 1, \dots, N \\ \xi_i(0) = \xi_{i,0}, \quad \psi_i(T) = 0, \quad i = 1, \dots, N, \end{aligned} \quad (17)$$

where the dynamic is composed as  $G(\zeta_i, \zeta_j, v) = \begin{bmatrix} G_1(\zeta_i, \zeta_j, v) \\ G_2(\zeta_i, \zeta_j, v) \end{bmatrix}$  for:

$$\begin{aligned} G_1(\zeta_i, \zeta_j, v) &= P(\xi_i, \xi_j)(\xi_j - \xi_i) - \frac{1}{\gamma} \psi_i, \\ G_2(\zeta_i, \zeta_j, v) &= -D_{\xi} \tilde{m}(\xi_i)^{\top} \nabla_m J^F(\tilde{m}(\xi_i); v) - D_{\xi_i} [P(\xi_i, \xi_j)(\xi_j - \xi_i)]^{\top} \psi_i \\ &\quad - D_{\xi_j} [P(\xi_j, \xi_i)(\xi_i - \xi_j)]^{\top} \psi_j. \end{aligned}$$

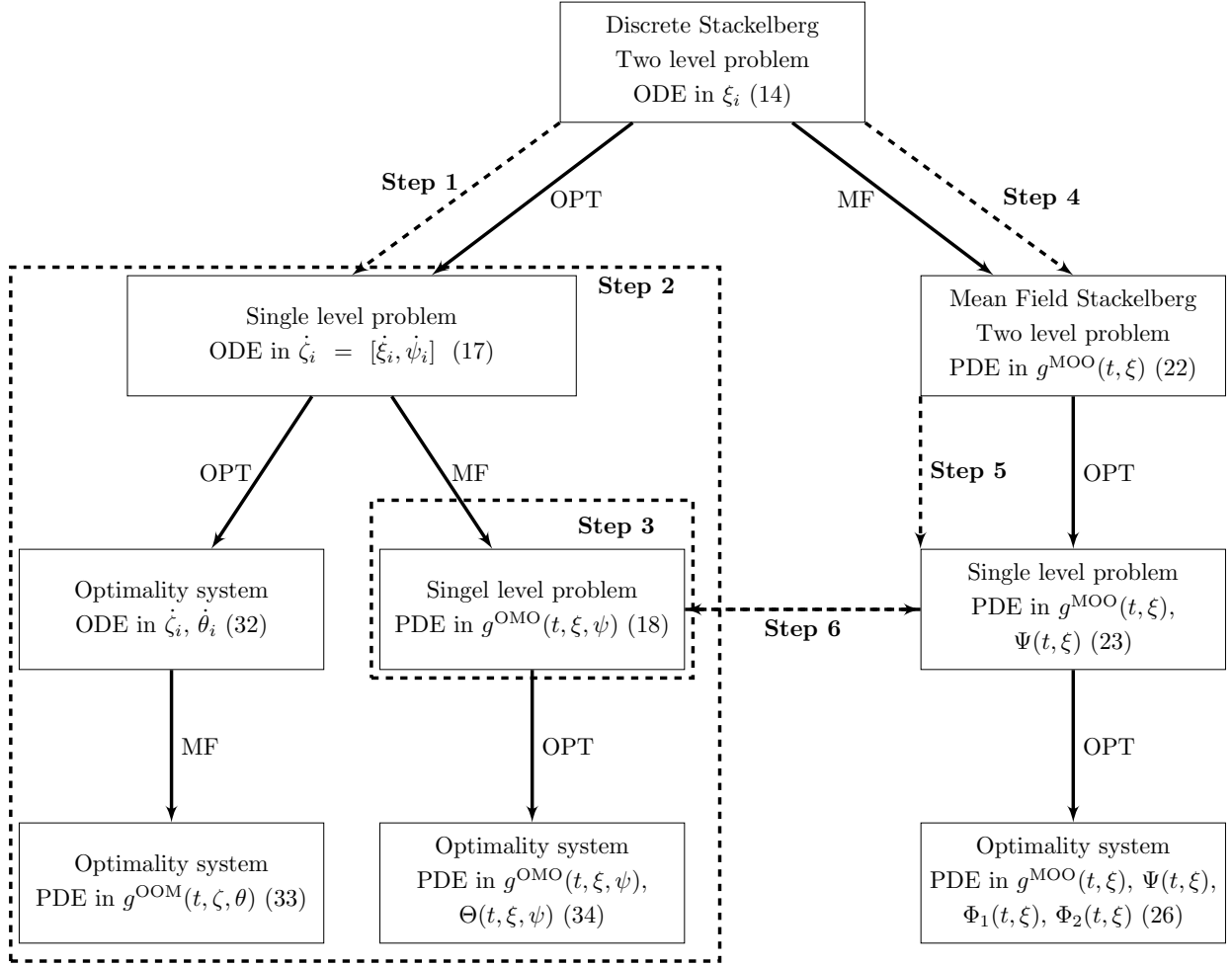


Figure 2: This graph illustrates the proof of Theorem 3.1. The solid lines indicate the order of optimization (OPT) and mean field limit (MF). The dashed lines refer to the steps in the proof.

*Step 2.* Starting from the optimization problem in (17), we now show the interchangeability of optimization and mean field limit for this problem and give the necessary conditions.

We apply Lemma 2.1 to Problem (17) and its formal mean-field limit for the probability density  $g^{\text{OMO}} = g^{\text{OMO}}(t, \zeta)$ :

$$\begin{aligned} \min_v \int_0^T \left[ J^L(v, m_{g^{\text{OMO}}}(t)) + \frac{\beta}{2} v^2 \right] dt \\ \text{s.t. } 0 = \partial_t g^{\text{OMO}} + \text{div}_\zeta \left( g^{\text{OMO}} \int G(\zeta, \hat{\zeta}, v) \hat{g}^{\text{OMO}} d\hat{\zeta} \right) \\ g^{\text{OMO}}(0, \zeta) = g_0^{\text{OMO}}(\zeta) \\ g^{\text{OMO}}(T, \zeta) = g_T^{\text{OMO}}(\zeta). \end{aligned} \quad (18)$$

Hence, provided that we have on the support of  $g^{\text{OMO}}$ :

$$\nabla_\zeta \Theta(t, \zeta) = - \int \theta g_2^{\text{OOM}}(t, \zeta, \theta) d\theta,$$

for the multiplier  $\Theta$  to  $g^{\text{OOM}}$ , we obtain that the optimality systems to (17) and (18) coincides in the mean-field limit. The probability density  $g^{\text{OOM}}$  corresponding to the formal first order optimality system of problem (17) is given in terms of  $g^{\text{OMO}}$  and  $g_2^{\text{OOM}}$  according to:

$$g^{\text{OOM}}(t, \zeta, \theta) = g^{\text{OMO}}(t, \zeta) g_2^{\text{OOM}}(t, \zeta, \theta).$$

*Step 3.* Now, the formal mean-field optimal control problem (18) is reformulated. We denote by  $g^{\text{MOO}}$  the probability density fulfilling the state equation in (22).

We decompose  $g^{\text{OMO}}$  as follows:

$$g^{\text{OMO}}(t, \xi, \psi) = g_1^{\text{OMO}}(t, \xi) g_2^{\text{OMO}}(t, \xi, \psi),$$

where  $\int g_2^{\text{OMO}}(t, \xi, \psi) d\psi = 1$ . Furthermore, we denote the expected value by:

$$\mathfrak{P}(t, \xi) := \int \psi g_2^{\text{OMO}}(t, \xi, \psi) d\psi.$$

a) Inserting this ansatz in the objective in (18) yields:

$$\int_0^T \left[ J^L(v, m_{g_1^{\text{OMO}} g_2^{\text{OMO}}}(t)) + \frac{\beta}{2} v^2 \right] dt = \int_0^T \left[ J^L(v, m_{g_1^{\text{OMO}}}(t)) + \frac{\beta}{2} v^2 \right] dt. \quad (19)$$

b) For the dynamic we have:

$$\begin{aligned} 0 = \partial_t g_1^{\text{OMO}} + \int \text{div}_\xi \left( g_1^{\text{OMO}} g_2^{\text{OMO}} \int G_1(\zeta, \hat{\zeta}, v) \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} d\hat{\zeta} \right) d\psi \\ + \int \text{div}_\psi \left( g_1^{\text{OMO}} g_2^{\text{OMO}} \int G_2(\zeta, \hat{\zeta}, v) \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} d\hat{\zeta} \right) d\psi, \end{aligned}$$

and using the definition of  $G_1$  we have:

$$\begin{aligned} 0 = \partial_t g_1^{\text{OMO}} + \text{div}_\xi \left( g_1^{\text{OMO}} \left[ \int P(\xi, \hat{\xi})(\hat{\xi} - \xi) \hat{g}_1^{\text{OMO}} d\hat{\xi} - \frac{1}{\gamma} \int \psi g_2^{\text{OMO}} d\psi \right] \right), \\ = \partial_t g_1^{\text{OMO}} + \text{div}_\xi \left( g_1^{\text{OMO}} \int g_2^{\text{OMO}} \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) - \frac{1}{\gamma} \psi \right] \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} d\hat{\psi} d\hat{\xi} d\psi \right). \end{aligned}$$

We simplify integrals and have the expression by the definition of  $\mathfrak{P}$ :

$$0 = \partial_t g_1^{\text{OMO}} + \text{div}_\xi \left( g_1^{\text{OMO}} \left[ \int P(\xi, \hat{\xi})(\hat{\xi} - \xi) \hat{g}_1^{\text{OMO}} d\hat{\xi} - \frac{1}{\gamma} \mathfrak{P}(t, \xi) \right] \right). \quad (20)$$

c) The formal equation for  $\mathfrak{P}$  is obtained upon integration of (18):

$$\begin{aligned} 0 &= \partial_t (g_1^{\text{OMO}} \mathfrak{P}) \\ &+ \int \psi \text{div}_\xi \left( g_1^{\text{OMO}} g_2^{\text{OMO}} \int \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) - \frac{1}{\gamma} \psi \right] \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} d\hat{\xi} d\hat{\psi} \right) d\psi \\ &+ \int \psi \text{div}_\psi \left( g_1^{\text{OMO}} g_2^{\text{OMO}} \int G_2(\zeta, \hat{\zeta}, v) \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} d\hat{\xi} d\hat{\psi} \right) d\psi. \end{aligned}$$

Integration by parts yields:

$$\begin{aligned} 0 &= \partial_t (g_1^{\text{OMO}} \mathfrak{P}) + \text{div}_\xi \left( g_1^{\text{OMO}} \left[ \int P(\xi, \hat{\xi})(\hat{\xi} - \xi) \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} g_2^{\text{OMO}} d\hat{\xi} d\hat{\psi} d\psi \right. \right. \\ &\quad \left. \left. - \frac{1}{\gamma} \int \psi^2 \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} g_2^{\text{OMO}} d\hat{\xi} d\hat{\psi} d\psi \right] \right) \\ &\quad - g_1^{\text{OMO}} \int g_2^{\text{OMO}} G_2(\zeta, \hat{\zeta}, v) \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} d\hat{\zeta} d\psi, \\ &= \partial_t (g_1^{\text{OMO}} \mathfrak{P}) + \text{div}_\xi \left( g_1^{\text{OMO}} \left[ \mathfrak{P} \int P(\xi, \hat{\xi})(\hat{\xi} - \xi)^\top \hat{g}_1^{\text{OMO}} d\hat{\xi} - \frac{1}{\gamma} \int \psi^2 g_2^{\text{OMO}} d\psi \right] \right) \\ &\quad - g_1^{\text{OMO}} \int g_2^{\text{OMO}} G_2(\zeta, \hat{\zeta}, v) \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} d\hat{\zeta} d\psi. \end{aligned}$$

Using the assumption in (15), i.e.:

$$\int \psi^2 g_2^{\text{OMO}} d\psi = \left( \int \psi g_2^{\text{OMO}} d\psi \right)^2,$$

we obtain the following equality:

$$\begin{aligned} 0 &= \partial_t (g_1^{\text{OMO}} \mathfrak{P}) + \text{div}_\xi \left( g_1^{\text{OMO}} \mathfrak{P} \left[ \int P(\xi, \hat{\xi})(\hat{\xi} - \xi)^\top \hat{g}_1^{\text{OMO}} d\hat{\xi} - \frac{1}{\gamma} \mathfrak{P} \right] \right) \\ &\quad - g_1^{\text{OMO}} \int g_2^{\text{OMO}} \int G_2(\zeta, \hat{\zeta}, v) \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} d\hat{\zeta} d\psi, \\ &= \mathfrak{P} \left[ \partial_t g_1^{\text{OMO}} + \text{div}_\xi \left( g_1^{\text{OMO}} \left[ \int P(\xi, \hat{\xi})(\hat{\xi} - \xi) \hat{g}_1^{\text{OMO}} d\hat{\xi} - \frac{1}{\gamma} \mathfrak{P} \right] \right) \right] \\ &\quad + g_1^{\text{OMO}} \left[ \partial_t \mathfrak{P} + \nabla_\xi \mathfrak{P} \cdot \left( \int P(\xi, \hat{\xi})(\hat{\xi} - \xi) \hat{g}_1^{\text{OMO}} d\hat{\xi} - \frac{1}{\gamma} \mathfrak{P} \right) \right. \\ &\quad \left. - \int g_2^{\text{OMO}} G_2(\zeta, \hat{\zeta}, v) \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} d\hat{\zeta} d\psi \right]. \end{aligned}$$

Due to (20), we get on the support of  $g_1^{\text{OMO}}$  using the definition of  $G_2$ :

$$\begin{aligned} 0 &= \partial_t \mathfrak{P} + \nabla_\xi \mathfrak{P} \left( \int P(\xi, \hat{\xi})(\hat{\xi} - \xi)^\top \hat{g}_1^{\text{OMO}} d\hat{\xi} - \frac{1}{\gamma} \mathfrak{P} \right) \\ &\quad - \int g_2^{\text{OMO}} \left[ -D_\xi \tilde{m}(\xi)^\top \nabla_m J^F(\tilde{m}(\xi); v) - D_\xi \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) \right]^\top \psi \right. \\ &\quad \left. - D_\xi \left[ P(\hat{\xi}, \xi)(\xi - \hat{\xi}) \right]^\top \hat{\psi} \right] \hat{g}_1^{\text{OMO}} \hat{g}_2^{\text{OMO}} d\hat{\zeta} d\psi. \end{aligned}$$

Using that  $\int \hat{g}_1^{\text{OMO}} d\hat{\xi} = \int \hat{g}_2^{\text{OMO}} d\hat{\psi} = 1$  yields:

$$\begin{aligned} 0 = & \partial_t \mathfrak{P} + \nabla_\xi \mathfrak{P} \left( \int P(\xi, \hat{\xi})(\hat{\xi} - \xi)^\top \hat{g}_1^{\text{OMO}} d\hat{\xi} - \frac{1}{\gamma} \mathfrak{P} \right) + D_\xi \tilde{m}(\xi)^\top \nabla_m J^F(\tilde{m}(\xi); v) \\ & + \int D_\xi \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) \right]^\top \mathfrak{P} \hat{g}_1^{\text{OMO}} d\hat{\xi} + \int D_\xi \left[ P(\hat{\xi}, \xi)(\xi - \hat{\xi}) \right]^\top \hat{\mathfrak{P}} \hat{g}_1^{\text{OMO}} d\hat{\xi}. \end{aligned} \quad (21)$$

We leave this and come back to Equation (21) in Step 6 where we connect it with the optimality system of (23).

*Step 4.* Due to Lemma 2.2, the corresponding mean-field formulation of the optimal control problem of the followers in (14) is obtained and the moment  $m$  of the followers in the leader's objective functional is rewritten by the density  $g^{\text{MOO}}$ . This yields the mean-field Stackelberg game:

$$\begin{aligned} \min_v & \int_0^T \left[ J^L(v, m_{g^{\text{MOO}}}(t)) + \frac{\beta}{2} v^2 \right] dt \\ \text{s.t. } & \min_{w, g^{\text{MOO}}} \int_0^T \int \left[ J^F(\tilde{m}(\xi); v) + \frac{\gamma}{2} w^2 \right] g^{\text{MOO}} d\xi dt \\ & \text{s.t. } 0 = \partial_t g^{\text{MOO}} + \text{div}_\xi \left( g^{\text{MOO}} \int \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) + w \right] \hat{g}^{\text{MOO}} d\hat{\xi} \right) \\ & g^{\text{MOO}}(0, \xi) = g_0^{\text{MOO}}(\xi). \end{aligned} \quad (22)$$

We can identify

$$\begin{aligned} g^{\text{OMO}}(t, \xi, \psi) &= g^{\text{MOO}}(t, \xi) g_2^{\text{OMO}}(t, \xi, \psi), \\ - \int \psi g_2^{\text{OMO}}(t, \xi, \psi) d\psi &= \nabla_\xi \Psi(t, \xi), \end{aligned}$$

and we obtain equivalence of (14) and (22) according to Lemma 2.2 in the mean-field limit.

*Step 5.* The first-order optimality conditions for the follower problem in (22) are given by (23) which yields the mean-field leader's problem as:

$$\begin{aligned} \min_v & \int_0^T \left[ J^L(v, m_{g^{\text{MOO}}}(t)) + \frac{\beta}{2} v^2 \right] dt \\ \text{s.t. } & 0 = \partial_t g^{\text{MOO}} + \text{div}_\xi \left( g^{\text{MOO}} \int \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) + \frac{1}{\gamma} \nabla_\xi \Psi \right] \hat{g}^{\text{MOO}} d\hat{\xi} \right) \\ & 0 = \partial_t \Psi(t, \xi) + \int \left\{ \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) + \frac{1}{\gamma} \nabla_\xi \Psi \right]^\top \nabla_\xi \Psi \right. \\ & \quad \left. + \left[ P(\hat{\xi}, \xi)(\xi - \hat{\xi}) + \frac{1}{\gamma} \nabla_\xi \hat{\Psi} \right]^\top \nabla_\xi \hat{\Psi} \right\} \hat{g}^{\text{MOO}} d\hat{\xi} \\ & \quad - J^F(\tilde{m}(\xi), v) - \frac{1}{2\gamma} (\nabla_\xi \Psi)^2 \\ & g^{\text{MOO}}(0, \xi) = g_0^{\text{MOO}}(\xi), \quad \Psi(T, \xi) = 0, \end{aligned} \quad (23)$$

where the follower control is given by:

$$w(t, \xi) = \frac{1}{\gamma} \nabla_\xi \Psi(t, \xi).$$

*Step 6.* Finally, we show the equivalence of (18) and (23).

a) For  $g_1^{\text{OMO}}(t, \xi) = g^{\text{MOO}}(t, \xi)$ , we have that the objective functionals in (19) and in (23) coincide.

b) The first constraint in (23) and (20) are equivalent provided that:

$$\mathfrak{P}(t, \xi) = -\nabla_\xi \Psi(t, \xi). \quad (24)$$

c) Due to the constraint in (23), the gradient  $\nabla_\xi \Psi$  formally fulfills:

$$\begin{aligned} 0 &= \partial_t \nabla_\xi \Psi(t, \xi) + \nabla_\xi \int \left\{ \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) + \frac{1}{\gamma} \nabla_\xi \Psi \right]^\top \nabla_\xi \Psi \right. \\ &\quad \left. + \left[ P(\hat{\xi}, \xi)(\xi - \hat{\xi}) + \frac{1}{\gamma} \nabla_\xi \hat{\Psi} \right]^\top \nabla_\xi \hat{\Psi} \right\} \hat{g}^{\text{MOO}} d\hat{\xi} \\ &\quad - D_\xi \tilde{m}(\xi)^\top J^F(\tilde{m}(\xi), v) - \frac{1}{\gamma} \nabla_\xi^2 \Psi \nabla_\xi \Psi, \\ &= \partial_t \nabla_\xi \Psi(t, \xi) + \int \left\{ D_\xi \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) \right]^\top \nabla_\xi \Psi + P(\xi, \hat{\xi})(\hat{\xi} - \xi) \nabla_\xi^2 \Psi + \frac{2}{\gamma} \nabla_\xi^2 \Psi \nabla_\xi \Psi \right. \\ &\quad \left. + D_\xi \left[ P(\hat{\xi}, \xi)(\xi - \hat{\xi}) \right]^\top \nabla_\xi \hat{\Psi} \right\} \hat{g}^{\text{MOO}} d\hat{\xi} - D_\xi \tilde{m}(\xi)^\top J^F(\tilde{m}(\xi), v) - \frac{1}{\gamma} \nabla_\xi^2 \Psi \nabla_\xi \Psi, \end{aligned}$$

and therefore:

$$\begin{aligned} 0 &= \partial_t \nabla_\xi \Psi(t, \xi) \\ &\quad + \int \left\{ D_\xi \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) \right]^\top \nabla_\xi \Psi + P(\xi, \hat{\xi})(\hat{\xi} - \xi) \nabla_\xi^2 \Psi \right. \\ &\quad \left. + D_\xi \left[ P(\hat{\xi}, \xi)(\xi - \hat{\xi}) \right]^\top \nabla_\xi \hat{\Psi} \right\} \hat{g}^{\text{MOO}} d\hat{\xi} \\ &\quad - D_\xi \tilde{m}(\xi)^\top J^F(\tilde{m}(\xi), v) + \frac{1}{\gamma} \nabla_\xi^2 \Psi \nabla_\xi \Psi. \end{aligned} \quad (25)$$

Equations (21) and (25) coincide provided that the consistency condition in (24) holds true.

Hence, the optimality conditions coincide under the condition of the Theorem. This finishes the proof.

## 4 Solution Method

We propose an iterative scheme, that alternates between updating the leader control  $v$  and solving a set of PDEs.

In the previous section we have proven that under the given conditions the order of the optimization steps and the mean-field limit can be exchanged. However, the resulting optimization systems differ considerably.

In Option 1, the OOM approach, the optimality system consists of a PDE and an algebraic condition, the unknown is the probability density  $g^{\text{OOM}}(t, \zeta, \theta)$  which is a  $4n_F + 1$ -dimensional function. The optimality system for Option 2 (OMO), consists of two coupled PDEs in the probability density  $g^{\text{OMO}}(t, \zeta)$  and the costate  $\Theta(t, \zeta)$ . Both depend on  $2n_F + 1$  variables.

We therefore develop an algorithm for Option 3, the MOO approach, which consists of four coupled PDEs. The unknowns are  $g^{\text{MOO}}(t, \xi)$ ,  $\Psi(t, \xi)$ ,  $\Phi_1(t, \xi)$ , and  $\Phi_2(t, \xi)$  and they depend on only  $n_F + 1$  variables. The



optimality system derived by the MOO approach reads:

$$0 = \partial_t g^{\text{MOO}} + \text{div}_\xi \left( g^{\text{MOO}} \int \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) + \frac{1}{\gamma} \nabla_\xi \Psi \right] \hat{g}^{\text{MOO}} d\hat{\xi} \right), \quad (26a)$$

$$0 = \partial_t \Psi + \int \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi)^\top \nabla_\xi \Psi + P(\hat{\xi}, \xi)(\xi - \hat{\xi})^\top \nabla_\xi \hat{\Psi} \right] \hat{g}^{\text{MOO}} d\hat{\xi} \\ + \frac{1}{2\gamma} (\nabla_\xi \Psi)^2 - J^F(\tilde{m}(\xi), v), \quad (26b)$$

$$0 = \partial_t \Phi_1 + \frac{1}{\gamma} \nabla_\xi \Phi_1^\top \nabla_\xi \Psi \\ + \int \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi)^\top \nabla_\xi \Phi_1 + P(\hat{\xi}, \xi)(\xi - \hat{\xi})^\top \nabla_\xi \hat{\Phi}_1 \right] \hat{g}^{\text{MOO}} d\hat{\xi} \\ + \int \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi)^\top \nabla_\xi \Psi + P(\hat{\xi}, \xi)(\xi - \hat{\xi})^\top \nabla_\xi \hat{\Psi} \right] \hat{\Phi}_2 d\hat{\xi} \\ - \nabla_m J^L(v, m_{g^{\text{MOO}}}(t)) \tilde{m}(\xi), \quad (26c)$$

$$0 = \partial_t \Phi_2 + \text{div}_\xi \left( \int P(\xi, \hat{\xi})(\hat{\xi} - \xi) \hat{g}^{\text{MOO}} d\hat{\xi} \Phi_2 \right) \\ + \text{div}_\xi \left( \int P(\xi, \hat{\xi})(\hat{\xi} - \xi) \hat{\Phi}_2 d\hat{\xi} g^{\text{MOO}} \right) \\ + \frac{1}{\gamma} \text{div}_\xi (\nabla_\xi \Psi \Phi_2) - \frac{1}{\gamma} \text{div}_\xi (\nabla_\xi \Phi_1 g^{\text{MOO}}), \quad (26d)$$

$$0 = \nabla_v J^L(v, m_{g^{\text{MOO}}}(t)) + \beta v - \int \nabla_v J^F(\tilde{m}(\xi), v) \Phi_2 d\xi, \quad (26e)$$

$$g^{\text{MOO}}(0, \xi) = g_0^{\text{MOO}}(\xi), \quad \Psi(T, \xi) = 0, \quad \Phi_1(T, \xi) = 0, \quad \Phi_2(0, \xi) = 0. \quad (26f)$$

Note, Equations (26a) and (26d) are forward in time, while (26b) and (26c) are backwards in time.

We solve the equation (26b) for  $\nabla_\xi \Psi$  instead of  $\Psi$ , likewise we solve (26c) for  $\nabla_\xi \Phi_1$  instead of  $\Phi_1$ . Then all PDEs are nonlinear conservative transport equations. These PDEs are integro-differential equations and have stiff source terms as  $\gamma \rightarrow 0$ .

Observe, that the equation for the gradient  $\nabla_\xi \Psi$  is independent of the other unknowns in case  $P(\xi, \hat{\xi}) = 0$ . Then we have:

$$0 = \partial_t \nabla_\xi \Psi(t, \xi) + \nabla_\xi \left( \frac{1}{2\gamma} (\nabla_\xi \Psi(t, \xi))^2 - J^F(\tilde{m}(\xi), v(t)) \right), \quad (27a)$$

$$0 = \nabla_\xi \Psi(T, \xi). \quad (27b)$$

This equation also determines the follower's control which we specify in the lemma below.

**Lemma 4.1.** *Consider the Stackelberg game given by (14). Then the follower control is:*

$$w(t, \xi) = \frac{1}{\gamma} \nabla_\xi \Psi(t, \xi).$$

*If the interaction kernel is  $P(\xi, \hat{\xi}) = 0$ , then the evolution of the optimal follower's control  $w$  is uniquely determined by Equation (27) for any given leader control  $v$ .*

*Proof.* Lagrange multiplier theorem is applied to lower level problem of (22), see Step 5 of the proof of Theorem 3.1.  $\square$

For  $P \equiv 0$ , we propose the following sequential solution to (26): The solution to Equation (27) allows to

solve the conservation law for  $g^{\text{MOO}}$ :

$$0 = \partial_t g^{\text{MOO}}(t, \xi) + \text{div}_\xi \left( \frac{1}{\gamma} \nabla_\xi \Psi(t, \xi) g^{\text{MOO}}(t, \xi) \right), \quad (28a)$$

$$g^{\text{MOO}}(0, \xi) = g_0^{\text{MOO}}(\xi). \quad (28b)$$

Also, we solve a balance equation for  $\nabla_\xi \Phi_1$ :

$$0 = \partial_t \nabla_\xi \Phi_1(t, \xi) + \nabla_\xi \left( \frac{1}{\gamma} \nabla_\xi \Phi_1(t, \xi)^\top \nabla_\xi \Psi(t, \xi) \right) \quad (29a)$$

$$- \nabla_m J^L(v(t), m_{g^{\text{MOO}}}(t)) \nabla_\xi \tilde{m}(\xi),$$

$$0 = \nabla_\xi \Phi_1(T, 0). \quad (29b)$$

Finally, we solve for  $\Phi_2$  by (26d):

$$0 = \partial_t \Phi_2(t, \xi) + \frac{1}{\gamma} \text{div}_\xi (\nabla_\xi \Psi(t, \xi) \Phi_2(t, \xi)) - \frac{1}{\gamma} \text{div}_\xi (\nabla_\xi \Phi_1(t, \xi) g^{\text{MOO}}(t, \xi)), \quad (30a)$$

$$0 = \Phi_2(0, \xi). \quad (30b)$$

This procedure sequentially solves the optimality system for any fixed leader control  $v$ .

Then, Equation (26e) implicitly describes the optimal leader control depending on the objective function of the followers and the leader as well as the dual variable  $\Phi_2$ . As last step we update of the leader control:

$$v^{k+1}(t) = v^k(t) + \sigma_k d^k(t)$$

where  $d(t) \in \mathbb{R}^{n_L}$  for  $t \in [0, T]$  is computed by:

$$d(t) = - \left( \nabla_v J^L(v(t), m_{g^{\text{MOO}}}(t)) + \beta v(t) - \int \nabla_v J^F(\tilde{m}(\xi), v(t)) \Phi_2(t, \xi) d\xi \right).$$

We combine this with a backtracking line search for the step size  $\sigma_k > 0$  based on the Armijo condition for the leader objective, c.f. [7, Cor.2]. The complete procedure is summarized in Algorithm 1.

---

**Algorithm 1** Continuous Optimization Method

---

1: **Initialize** Choose initial guess for the leader control  $v^0(t) \in \mathbb{R}^{n_L}$  for  $t \in [0, T]$ .

2: **for**  $k = 0, 1, \dots$  **do**

3:   Solve backward Equation (27) with  $v^k$  to get  $\nabla_\xi \Psi^k$ .

4:   Solve forward Equation (28) to get  $g_k^{\text{MOO}}$ .

5:   Solve backward Equation (29) to get  $\nabla_\xi \Phi_1^k$ .

6:   Solve forward Equation (30) to get  $\Phi_2^k$ .

7:   Compute a decent direction

$$d^k(t) = - \left( \nabla_v J^L(v^k(t), m_{g_k^{\text{MOO}}}(t)) + \beta v^k(t) - \int \nabla_v J^F(\tilde{m}(\xi), v^k(t)) \Phi_2^k(t, \xi) d\xi \right).$$

8:   Choose a step size  $\sigma_k > 0$  according to Armijo rules.

9:   Update control  $v^{k+1}(t) = v^k(t) + \sigma_k d^k(t)$ .

10: **if** The termination condition is satisfied, **then** STOP.

11:   **end if**

12: **end for**

---

For all PDEs we use periodic boundary conditions. The state space is discretized by equidistant cells, we use  $n_\xi$  to denote the number of cells in a state space dimension of  $\xi$ . We use Lax-Friedrich's method to describe the numerical flux across the cell interfaces [24]. The integral in  $J^L$  is discretized by a first order quadrature rule.

The stopping criteria of the algorithm is that the change in the leader control is below a relative tolerance.

## 5 Numerical Results

### 5.1 Setup

In this section we discuss the results of the numerical experiments that were implemented in Matlab. We start with the description of the used parameters: The dimension of the state and control space of the followers is  $n_F = 1$  and we assume  $\xi \in [0, 2]$ . Also the leader control is one-dimensional and we consider the time  $T = 1$ .

The leader objective combines tracking of the leader control with moment of the followers as:

$$J^L(v, m_{g^{\text{MOO}}}(t)) = \frac{1}{2}(v_d + m_{g^{\text{MOO}}}(t) - v)^2,$$

with the desired control:

$$v_d(t) = \sin(2\pi t).$$

The followers objective couple the states of the followers and the leader control as:

$$J^F(\tilde{m}(\xi), v) = -\frac{1}{2}(\tilde{m}(\xi) - v)^2. \quad (31)$$

Both objective functions are regularized by a quadratic term, the regularization parameters are  $\beta, \gamma$  in the experiments. We consider the expectation  $\tilde{m}(\xi) = \xi$ . As initial condition for the probability density  $g^{\text{MOO}}$  we chose the uniform distribution:

$$g_0^{\text{MOO}}(\xi) = \chi_{[0.5, 1.5]}(\xi).$$

The state space is discretized by  $n_\xi$  equidistant cells. The time steps of the PDE solving are chosen to satisfy the Courant-Friedrichs-Lewy condition with the constant  $\text{CFL} = 0.95$ . For the evaluation of the objective functions and the termination conditions the time is discretized by  $n_t$  equidistant points in time. Whenever necessary linear reconstruction is performed by interpolation. Further, we choose a maximum iteration number for Algorithm 1.  $\text{maxIter} = 100$ . The tolerance is related to the space discretization as  $\text{tol} = \frac{1}{100n_\xi}$ . As initial guess:

$$v^0(t) = t,$$

is chosen.

### 5.2 Experiments

In Figure 3 on the left, the behavior of the solutions is illustrated for different discretizations of the leader control  $v$ . We observe monotonically decrease of the  $L^2$  error as the control grid increases. On the right, the implementation of Algorithm 1 is compared to `fminuncon`, a commercial trust-region solver in Matlab where we provide gradient information based on same PDE solvers. The implementation outperforms `fminuncon` in terms of computational time for all mesh sizes and the computational time increases linearly with mesh refinement in Algorithm 1.

All experiments in the plots are performed for the space discretization of  $n_\xi = 500$ .

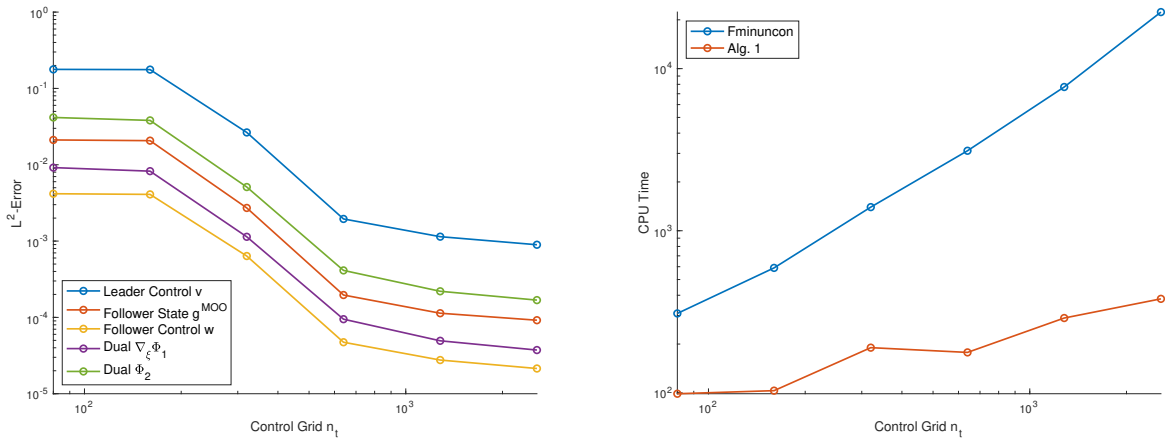


Figure 3: Performance. On the left, the  $L^2$ -error of the leader control and the PDE solutions is plotted for refined mesh  $n_t$  of the leader control. The solutions are compared to the solution of the finest mesh. On the right, CPU times of Matlab fminuncon trust-region with user specified gradient and an implementation of Algorithm 1 are compared.

In Figure 4, the leader control is illustrated. The control action is larger for smaller regularization parameters due to decreasing influence of the objective. For any regularization, the sinodal shape of the optimal leader control  $v$  is clearly recognizable. This corresponds to the expected control  $v_d$ .

In Figure 5 on the left, the control of the followers is shown as surface plot. The final time condition  $\nabla_{\xi}\Psi(T, \xi) = 0$  is visible in the control plot. On the right the resulting evolution of the followers' state is plotted. Note, the choice of the followers' objective in (31), give them the initiative to move away from the leader control. Due to this, we observe that the followers tend to concentrate at  $\xi = 2$  and  $\xi = 0$  as time evolves.

In Figure 6, the solutions of the adjoint equations, that are related to the optimization of the leader, are plotted. The strong relation to the followers' initial condition is visible on the left. On the right, we observe a similar shape as the evolution of the followers' state. Those states are due to the followers' state influence in the source terms of the adjoint equations.

## 6 Outlook

In this article a Stackelberg game with possibly infinitely many of followers is discussed. The formal mean-field limit is derived for the Stackelberg game and consistency conditions are established. Then a numerical method for the full optimality system and numerical results are presented.

Future steps include the extension of the Stackelberg of multiple leaders in order to enable accurate analysis of energy markets and tolling in vehicular traffic models. Also comparisons to particle simulations of the model are subject to further research. From the mathematical perspective, it is interesting, if Theorem 3.1 can be shown also rigorously.

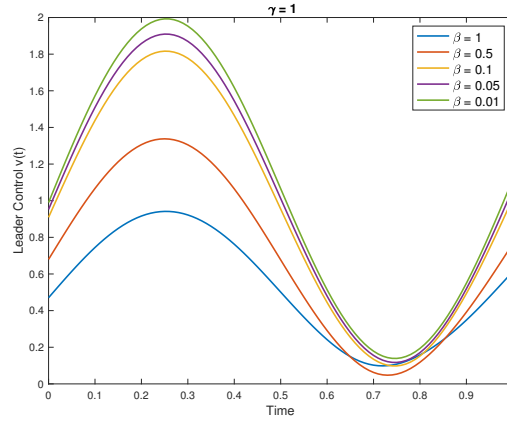


Figure 4: The optimal leader control is shown depending on the choice of the regularization parameter.

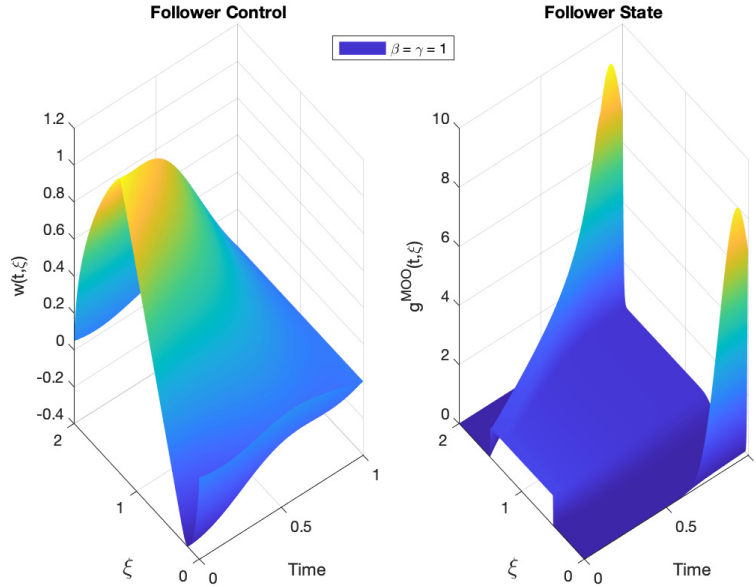


Figure 5: Time evolution of the followers' control and state. On the left, the solution of Equation (27), which is  $w(t, \xi) = \frac{1}{\gamma} \nabla_{\xi} \Psi$ , is shown. On the right, it is the solution of (28).

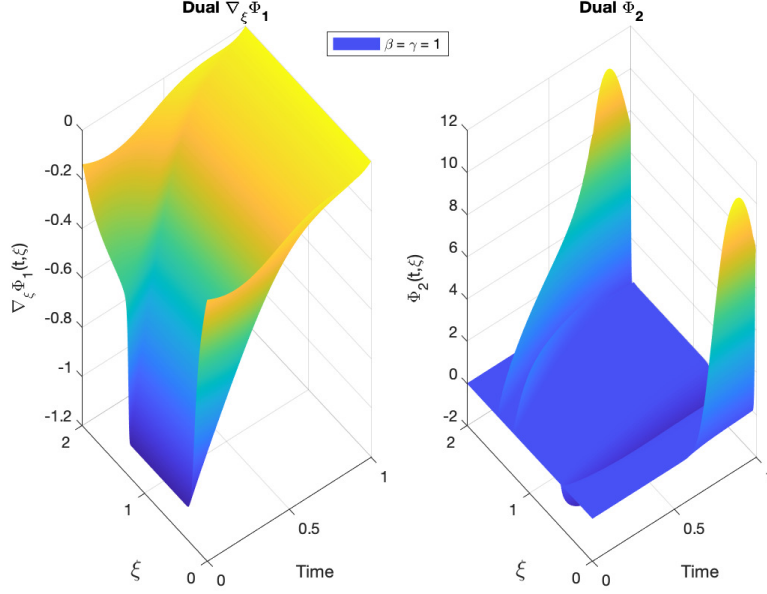


Figure 6: Time evolution of the dual states, on the left is the solution of Equation (29), on the right of (30).

## A Optimality Systems

If the OOM approach is followed, we get as optimality system of (17) for  $i = 1, \dots, N$ :

$$\begin{aligned}
\dot{\zeta}_i &= \frac{1}{N} \sum_{j=1}^N G(\zeta_i, \zeta_j, v) \\
\dot{\theta}_i &= -\frac{1}{N} \sum_{j=1}^N [\mathbf{D}_1 G(\zeta_i, \zeta_j, v) \theta_i + \mathbf{D}_2 G(\zeta_j, \zeta_i, v) \theta_j] + \begin{bmatrix} \mathbf{D}_{\xi_i} \tilde{m}(\xi_i) \\ 0 \end{bmatrix}^\top \nabla_m J^L(v, m(\vec{\xi})) \\
0 &= \nabla_v J^L(v, m(\vec{\xi})) + \beta v - \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \mathbf{D}_v G(\zeta_i, \zeta_j, v) \theta_i
\end{aligned} \tag{32}$$

Computing the mean field behavior yields the following equations with  $g^{\text{OOM}} = g^{\text{OOM}}(t, \zeta, \theta)$ :

$$\begin{aligned}
0 &= \partial_t g^{\text{OOM}} + \text{div}_{\zeta} \left( g^{\text{OOM}} \int G(\zeta, \hat{\zeta}, v) \hat{g}^{\text{OOM}} d\hat{\zeta} d\hat{\theta} \right) \\
&\quad - \text{div}_{\theta} \left( g^{\text{OOM}} \int [\mathbf{D}_1 G(\zeta, \hat{\zeta}, v) \theta + \mathbf{D}_2 G(\hat{\zeta}, \zeta, v) \hat{\theta}] \hat{g}^{\text{OOM}} d\hat{\zeta} d\hat{\theta} \right) \\
&\quad + \begin{bmatrix} \mathbf{D}_{\xi} \tilde{m}(\xi)^\top \nabla_m J^L(v, m_{g^{\text{OOM}}}(t)) \\ 0 \end{bmatrix}^\top \nabla_{\theta} g^{\text{OOM}} \\
0 &= \nabla_v J^L(v, m(\vec{\xi})) + \beta v - \int \mathbf{D}_v G(\zeta, \hat{\zeta}, v) \theta g^{\text{OOM}} \hat{g}^{\text{OOM}} d\hat{\zeta} d\hat{\theta} d\zeta d\theta
\end{aligned} \tag{33}$$

Where we specify

$$\begin{aligned}
D_1 G(\zeta, \hat{\zeta}, v) &= \begin{bmatrix} D_\xi \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) \right] & -\frac{1}{\gamma} & & \\ & & \ddots & \\ & D_\xi G_2(\zeta, \hat{\zeta}, v) & & -\frac{1}{\gamma} \\ & & & -D_\xi \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) \right] \end{bmatrix} \\
D_\xi G_2(\zeta, \hat{\zeta}, v) &= -D_\xi^2 \tilde{m}(\xi)^\top \nabla_m J^F(\tilde{m}(\xi), v) \\
&\quad - (D_\xi \tilde{m}(\xi))^2 \nabla_m^2 J^F(\tilde{m}(\xi), v) - D_\xi^2 \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) \right] \psi \\
D_2 G(\zeta, \hat{\zeta}, v) &= \begin{bmatrix} & & & 0 \\ D_\xi \left[ P(\xi, \hat{\xi})(\hat{\xi} - \xi) \right] & & & \ddots \\ & & & 0 \\ -D_\xi^2 \left[ P(\hat{\xi}, \xi)(\xi - \hat{\xi}) \right] \hat{\psi} & & -D_\xi \left[ P(\hat{\xi}, \xi)(\xi - \hat{\xi}) \right] & \end{bmatrix} \\
D_v G(\zeta, \hat{\zeta}, v) &= \begin{bmatrix} & & & 0 \\ & & \ddots & \\ & & & 0 \\ -D_\xi \tilde{m}(\xi)^\top D_v \left[ \nabla_m J^F(\tilde{m}(\xi), v) \right] & & & \end{bmatrix}
\end{aligned}$$

If in contrast, the OMO approach is chosen, the optimality system to (18) reads with  $g^{\text{OMO}} = g^{\text{OMO}}(t, \zeta)$  and  $\Theta = \Theta(t, \zeta)$ :

$$\begin{aligned}
0 &= \partial_t g^{\text{OMO}} + \text{div}_\zeta \left( g^{\text{OMO}} \int G(\zeta, \hat{\zeta}, v) \hat{g}^{\text{OMO}} d\hat{\zeta} \right) \\
0 &= \partial_t \Theta + \int \left[ G(\zeta, \hat{\zeta}, v)^\top \nabla_\zeta \Theta + G(\hat{\zeta}, \zeta, v)^\top \nabla_{\hat{\zeta}} \Theta \right] \hat{g}^{\text{OMO}} d\hat{\zeta} \\
&\quad - \nabla_m J^L(v, m_{g^{\text{OMO}}}(t))^\top \tilde{m}(\xi) - \frac{\beta}{2} v^2 \\
v(t) &= \frac{1}{\beta} \left[ -\nabla_v J^L(v, m_{g^{\text{OMO}}}(t)) + \int D_v G(\zeta, \hat{\zeta}, v) \nabla_\zeta \Theta \hat{g}^{\text{OMO}} g^{\text{OMO}} d\hat{\zeta} d\zeta \right]
\end{aligned} \tag{34}$$

## Acknowledgments

This work was supported by the DFG under Grant STE2063/2-1 and HE5386/19-1.

## References

- [1] Giacomo Albi, Mattia Bongini, Emiliano Cristiani, and Dante Kalise. Invisible control of self-organizing agents leaving unknown environments. *SIAM Journal on Applied Mathematics*, 76(4):1683–1710, jan 2016.
- [2] Giacomo Albi, Massimo Fornasier, and Dante Kalise. A boltzmann approach to mean-field sparse feedback control. *IFAC-PapersOnLine*, 50(1):2898–2903, jul 2017.
- [3] Giacomo Albi and Lorenzo Pareschi. Modeling of self-organized systems interacting with a few individuals: From microscopic to macroscopic dynamics. *Applied Mathematics Letters*, 26(4):397–401, apr 2013.

- [4] Giacomo Albi, Lorenzo Pareschi, and Mattia Zanella. Boltzmann-type control of opinion consensus through leaders. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2028):20140138, nov 2014.
- [5] Giacomo Albi, Lorenzo Pareschi, and Mattia Zanella. Uncertainty quantification in control problems for flocking models. *Mathematical Problems in Engineering*, 2015:1–14, 2015.
- [6] Elisabetta Allevi, Didier Aussel, and Rossana Riccardi. On an equilibrium problem with complementarity constraints formulation of pay-as-clear electricity market with demand elasticity. *Journal of Global Optimization*, 70(2):329–346, dec 2017.
- [7] Larry Armijo. Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific J. Math.*, 16(1):1–3, 1966.
- [8] Didier Aussel, Pascale Bendotti, and Miroslav Pištěk. Nash equilibrium in a pay-as-bid electricity market: Part 1 – existence and characterization. *Optimization*, 66(6):1013–1025, oct 2016.
- [9] Didier Aussel, Pascale Bendotti, and Miroslav Pištěk. Nash equilibrium in a pay-as-bid electricity market part 2 - best response of a producer. *Optimization*, 66(6):1027–1053, sep 2016.
- [10] Didier Aussel, Michal Červinka, and Matthieu Marechal. Deregulated electricity markets with thermal losses and production bounds: models and optimality conditions. *RAIRO - Operations Research*, 50(1):19–38, aug 2015.
- [11] Didier Aussel and Anton Svensson. Some remarks about existence of equilibria, and the validity of the epcc reformulation for multi-leader-follower games. *Journal of nonlinear and convex analysis*, 19(7):1141–1162, 2018.
- [12] Nicola Bellomo, Pierre Degond, and Eitan Tadmor, editors. *Active Particles, Volume 1*. Springer International Publishing, 2017.
- [13] Alain Bensoussan, Jens Frehse, and Sheung Chi Phillip Yam. On the interpretation of the master equation. *Stochastic Processes and their Applications*, 127(7):2093–2137, jul 2017.
- [14] Pierre Cardaliaguet. The convergence problem in mean field games with local coupling. *Applied Mathematics & Optimization*, 76(1):177–215, jun 2017.
- [15] Emiliano Cristiani, Benedetto Piccoli, and Andrea Tosin. *Multiscale Modeling of Pedestrian Dynamics*. Springer International Publishing, 2014.
- [16] Tobias Harks, Marc Schröder, and Dries Vermeulen. Toll caps in privatized road networks. *European Journal of Operational Research*, 276(3):947–956, aug 2019.
- [17] Rainer Hegselmann and Ulrich Krause. Opinion dynamics and bounded confidence: models, analysis and simulation. *J. Artif. Soc. Soc. Simul.*, 5, 2002.
- [18] René Henrion, Jiří Outrata, and Thomas Surowiec. Analysis of m-stationary points to an EPEC modeling oligopolistic competition in an electricity spot market. *ESAIM: Control, Optimisation and Calculus of Variations*, 18(2):295–317, jan 2012.
- [19] Michael Herty, Gabriella Puppo, and Giuseppe Visconti. From kinetic to macroscopic models and back, 2020.



- [20] Michael Herty and Christian Ringhofer. Consistent mean field optimality conditions for interacting agent systems. *Communications in Mathematical Sciences*, 17(4):1095–1108, 2019.
- [21] Xinmin Hu and Daniel Ralph. Using EPECs to model bilevel games in restructured electricity markets with locational prices. *Operations Research*, 55(5):809–827, oct 2007.
- [22] Andrew Koh and Simon Shepherd. Tolling, collusion and equilibrium problems with equilibrium constraints. *Trasporti Europei*, (n. 44):3–22, 2010.
- [23] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, mar 2007.
- [24] Randall J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhuser Basel, 1992.
- [25] Yan Ma and Minyi Huang. Linear quadratic mean field games with a major player: The multi-scale approach. *Automatica*, 113:108774, mar 2020.
- [26] Dov Monderer and Lloyd S. Shapley. Potential games. *Games and Economic Behavior*, 14(1):124–143, 1996.
- [27] Jun Moon and Tamer Başar. Linear quadratic mean field stackelberg differential games. *Automatica*, 97:200–213, nov 2018.
- [28] Sebastien Motsch and Eitan Tadmor. Heterophilious dynamics enhances consensus. *SIAM Review*, 56(4):577–621, jan 2014.
- [29] Giovanni Naldi, Lorenzo Pareschi, and Giuseppe Toscani, editors. *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*. Birkhäuser Boston, 2010.
- [30] John Nash. Non-cooperative games. *Annals of Mathematics. Second Series*, 54:286–295, 1951.
- [31] Daniel Nowak, Tobias Mahn, Hussein Al-Shatri, Alexandra Schwartz, and Anja Klein. A generalized nash game for mobile edge computation offloading. In *2018 6th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud)*. IEEE, mar 2018.
- [32] Lorenzo Pareschi. *Interacting multiagent systems : kinetic equations and Monte Carlo methods*. Oxford University Press, Oxford, 2014.
- [33] John von Neumann and Oskar Morgenstern. *Theory of games and economic behavior*. Princeton University Press, Princeton, NJ, anniversary edition, 2007. With an introduction by Harold W. Kuhn and an afterword by Ariel Rubinstein.
- [34] Heinrich von Stackelberg. *Market Structure and Equilibrium*. Springer Berlin Heidelberg, 2011.
- [35] Jessie Hui Wang, Dah Ming Chiu, and John C.S. Lui. A game-theoretic analysis of the implications of overlay network traffic on ISP peering. *Computer Networks*, 52(15):2961–2974, oct 2008.