Stochastic programming with primal–dual dynamics: a mean-field game approach

Casper T. Röling¹ · Stefanny Ramirez¹ · Dario Bauso²,³ · Hamidou Tembine⁴

Received: 27 March 2021 / Accepted: 7 July 2022 / Published online: 19 August 2022
© The Author(s) 2022

Abstract
This study addresses primal–dual dynamics for a stochastic programming problem for capacity network design. It is proven that consensus can be achieved on the here and now variables which represent the capacity of the network. The main contribution is a heuristic approach which involves the formulation of the problem as a mean-field game. Every agent in the mean-field game has control over its own primal–dual dynamics and seeks consensus with neighboring agents according to a communication topology. We obtain theoretical results concerning the existence of a mean-field equilibrium. Moreover, we prove that the consensus dynamics converge such that the agents agree on the capacity of the network. Lastly, we emphasize the ways in which penalties on control and state influence the dynamics of agents in the mean-field game.

Keywords Mean-field games · Primal–dual dynamics · Stochastic programming · Network optimization

Dario Bauso
d.bauso@rug.nl
Stefanny Ramirez
s.g.ramirez.juarez@rug.nl

¹ Engineering and Technology Institute Groningen (ENTEG), Faculty of Science and Engineering, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
² Jan C. Willems Center for Systems and Control, Engineering and Technology Institute Groningen (ENTEG), Faculty of Science and Engineering, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
³ Dipartimento di Ingegneria, Università di Palermo, Palermo, Italy
⁴ Learning and Game Theory Laboratory, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates
1 Introduction

Motivated by a stochastic program for the network capacity design and flow optimization, we cast for the first time primal–dual dynamics in the framework of mean-field games. We consider a scenario where demand materializes at some geographic spots and goods are transported from warehouses to fulfill such demand. We first model the problem as a two stage stochastic program for supply chain network design in the same spirit as in Ch. 1.5 [28]. Demand materializes at some nodes and goods flow over the edges. The flow in each edge is subject to capacity constraints and the flows in and out of each node must satisfy flow conservation constraints. The resulting optimization model is a constrained stochastic nonlinear (quadratic) two-stage program.

Continuous-time primal–dual gradient dynamics, also known as saddle-point dynamics, were first introduced in [2, 20]. Primal–dual dynamics is used in a number of application domains such as energy resource allocation [17], cyber-physical systems [26], wireless communications networks [11], smart grids [34], just to name a few. Recent works have been dedicated to exponential stability analysis of primal–dual dynamics [12, 24, 27, 30]. Furthermore, analysis on robustness has been performed in [15, 19, 25]. Also, the global/local nature of the convexity-concavity properties of the primal–dual dynamics was examined. Explicit attention has been paid to asymptotic stability properties of the primal–dual dynamics in [13, 14, 16, 29]. Primal–dual dynamics and stochastic gradient descent has been recently studied in [8]. The extension to large scale problems is possible through parallelization and using elastic averaging stochastic gradient descent algorithms [9].

There is less work dedicated to the combination of primal–dual dynamics with mean-field games. However, there have been studies concerning Lagrangian relaxation methods to solve mean-field games, such as in [1], where a planning problem is discussed. Additionally, in [32] and [33] Lagrangian relaxation is used in mean-field game power control. More recently a primal–dual approach has been used to solve constrained mean-field games [10]. Mean-field games in the form of coupled partial differential equations originated in the works of Lasry and Lions [21–23]. Explicit closed-form expressions for mean-field equilibria were first proposed for linear-quadratic mean-field games [3] and have been extended to more general cases in [18]. This paper mainly follows the linear-quadratic mean-field game formulation as in [4–6].

The main contribution is a heuristic approach which involves the reformulation of the original problem into a large number of stochastic primal–dual dynamics which are coupled in the same spirit as in mean-field games. To be more specific, we consider the instance of a large number of agents, each assigned with a primal–dual dynamics subject to a specific realization of the demand. The players of the mean-field game have to obtain consensus over the here and now decision variables: the capacity of each edge of the supply chain network. We also consider wait and see decision variables for flow and the Lagrange multipliers coming from the primal–dual dynamics. Note the analogy with a two-player
zero-sum game whereby the first player, the minimizer, sets the here and now variables required to accommodate for any potential scenario concerning flow and demand realization. Subsequently, the second player, the maximizer, sets the wait and see variables to maximize demand satisfaction given the decision of player 1 and the realized demand. The condition for convergence to a consensus value is that the communication network of agents needs to be connected, i.e., for any pair of nodes there exists a path connecting them. We first transform the primal–dual dynamics into a mean-field game, after which theoretical results are provided concerning a mean-field equilibrium solution in the same spirit as in [4–6]. Lastly, we provide numerical simulations to prove once again that consensus is obtained. Since the proposed methodology is heuristic in nature, the convergence values are in general sub-optimal.

The proposed model and methodology is original as it involves turning the micro-network model into a set of primal–dual dynamics and after that into a mean-field game in which each agent faces a different realization of the uncertain demand. We see this as a value and believe that the link between primal–dual dynamics and mean-field games expands the significance, meaning, and potential of mean-field games approaches beyond the ones already in the literature. Note also that the obtained mean-field game depends strongly on the micro-network optimization parameters. Actually, the incidence matrix, the penalty coefficients and the uncertain demand enter into the mean-field dynamics of the optimization flow and capacity variables as well as of the Lagrange multipliers.

This paper is organized as follows. In Sect. 2, we formulate the stochastic programming for network design. In Sect. 3, we present the primal–dual dynamics. In Sect. 4, we present the corresponding mean-field game. In Sect. 5, we state theoretical results on mean-field equilibrium and convergence. In Sect. 6, we provide a simulation example to corroborate our results. Lastly, in Sect. 7, we provide concluding remarks and discuss future works.

2 Stochastic programming for network design

Let us consider a physical flow network (referred to as micro-network) denoted as \( G(V, E) \). Let the set of nodes be \( V \in \{V_1, V_2, \ldots, V_n\} \), where each node corresponds to an agent in the network. The parameter \( n \) denotes the number of nodes present in the network. Let the set of edges be \( E \in \{E_1, E_2, \ldots, E_m\} \). The parameter \( m \) denotes the number of edges in the network. The demand is described by a vector \( w \in \mathbb{R}^{n \times 1} \) and is an uncertain random parameter. The capacity of the edges are the here and now variables and are denoted by \( c \in \mathbb{R}^{m \times 1} \). The transported goods in each edge are the wait and see variables and are denoted by \( u \in \mathbb{R}^{m \times 1} \). A graphical representation of the micro-network can be found in Fig. 1.

Let \( f_1(c) = \frac{1}{2}c^T \tilde{Q}_1 c + \tilde{f}_1^T c \) be the cost associated to the capacity of the edges in the micro-network, and \( f_2(u) = \frac{1}{2}u^T \tilde{Q}_2 u + \tilde{f}_2^T u \) be the cost charged for using the edges to transport goods such that the demand is satisfied. In the above costs we have a quadratic and a linear term. The matrices \( \tilde{Q}_1, \tilde{Q}_2 \in \mathbb{R}^{m \times m} \) are the cost coefficients for the quadratic term, and \( \tilde{f}_1, \tilde{f}_2 \in \mathbb{R}^m \) are the linear cost coefficients. We
model the supply chain network design problem as a two-stage stochastic program as follows:

\[
\begin{align*}
\min & \quad f_1(c) + E_w \hat{Q}(c, w) \\
\text{s.t.} & \quad u, c > 0.
\end{align*}
\]

(1)

Note that the objective function that we wish to minimize describes the total cost of transporting goods. The optimization variables in this problem are the capacity of the edges \( c \in \mathbb{R}^{m \times 1} \), and the flow in the edges \( u \in \mathbb{R}^{m \times 1} \). The capacities \( c \) are the first-stage decision variables, and they are obtained before the realization of the demand. The optimal capacities are independent of the demand, however they have to accommodate any realization of it. On the other hand, the flows \( u \) are the second-stage decision variables, they depend on the values of the first-stage variables and they are obtained after observing the realized demand. In addition, \( \bar{B} \in \mathbb{R}^{n \times m} \) is the incidence matrix of the micro-network, \( E_w(.) \) denotes expectation with respect to \( w \) and \( \hat{Q}(u, w) \) denotes the sub-optimal value of the second-stage problem, namely

\[
\hat{Q}(c, w) := \min_u \{f_2(u) | \bar{B}u = w, u \leq c \}.
\]

Observe that as \( \hat{Q}(c, w) \) depends on the capacities of the edges \( c \) and on the random variable \( w \), the optimal value of the second-stage problem is also a random variable. Hence, we can obtain its expected value with respect to the demand, namely \( E_w \hat{Q}(c, w) \), which is included in the objective function of the two-stage stochastic program (1).

### 3 Primal–dual dynamics

In this section we introduce the stochastic primal–dual dynamic related to the two-stage stochastic program (1).
Let \( \omega \) be a constant demand vector randomly extracted from the same distribution of \( w \). To obtain the primal–dual dynamics associated with this particular realization of the two-stage program (1) we first derive the Lagrangian function as follows:

\[
\mathcal{L}(u, c, \lambda, \mu) = f_1(c) + \lambda^T (Bu - \omega) + \mu^T (u - c) + f_2(u)
\]

\[
= f_1(c) + [\tilde{B}^T \lambda + \mu]_T u - \mu^T c - \lambda^T \omega + f_2(u)
\]

\[
= f_1(c) + \begin{bmatrix} I & 0 \\ -I & 0 \end{bmatrix} \mu + \begin{bmatrix} \tilde{B}^T \\ 0 \end{bmatrix} \lambda_+^T \begin{bmatrix} u \\ c \end{bmatrix} - \omega^T \lambda + f_2(u),
\]

where \( \mu \in \mathbb{R}^{m \times 1} \) and \( \lambda \in \mathbb{R}^{n \times 1} \) are the Lagrange multipliers.

The Lagrange dual problem can then be formulated as follows:

\[
\inf_{u, c} \left\{ f_1(c) + \begin{bmatrix} I & 0 \end{bmatrix} \mu + \begin{bmatrix} \tilde{B}^T \\ 0 \end{bmatrix} \lambda_+^T \begin{bmatrix} u \\ c \end{bmatrix} - \omega^T \lambda + f_2(u) \right\}.
\]

By imposing that the gradient of the Lagrangian with respect to \( u \) and \( c \) vanishes, we obtain the following Karush–Kuhn–Tucker (KKT) conditions:

\[
\begin{bmatrix} \nabla_u f_2(u) \\ \nabla_c f_1(c) \end{bmatrix} + \begin{bmatrix} I & -I \end{bmatrix} \mu + \begin{bmatrix} \tilde{B}^T \\ 0 \end{bmatrix} \lambda_+ = 0.
\]

(2)

The corresponding primal–dual dynamics are obtained from the gradient descent and gradient ascent dynamics over \( u \) and \( c \) respectively as shown next:

\[
\begin{cases}
\dot{u} = -\nabla_u \mathcal{L}_+ = \begin{bmatrix} -\tilde{Q}_2 u - \tilde{f}_2 - \mu - \tilde{B}^T \lambda \end{bmatrix}_+,
\dot{c} = -\nabla_c \mathcal{L}_+ = \begin{bmatrix} -\tilde{Q}_1 c - \tilde{f}_1 + \mu \end{bmatrix}_+,
\dot{\lambda} = +\nabla_{\lambda} \mathcal{L} = \tilde{B} u - \omega,
\dot{\mu} = +\nabla_{\mu} \mathcal{L} = \begin{bmatrix} u - c \end{bmatrix}_+.
\end{cases}
\]

(3)

In compact form, the above set of differential equation can be written as

\[
\begin{bmatrix}
\dot{u} \\
\dot{c} \\
\dot{\lambda} \\
\dot{\mu}
\end{bmatrix} = \begin{bmatrix}
-\tilde{Q}_2 & 0 & -\tilde{B}^T & -I \\
0 & -\tilde{Q}_1 & 0 & I \\
B & 0 & 0 & 0 \\
I & -I & 0 & 0
\end{bmatrix} \begin{bmatrix}
u \\
c \\
\lambda \\
\mu
\end{bmatrix} + \begin{bmatrix}
-\tilde{f}_2 \\
-\tilde{f}_1 \\
-\omega \\
0
\end{bmatrix}.
\]

(4)

In the two-stage program (1) the capacity vector \( c \) needs to be the same for any realization of the demand \( w \), namely for any value of \( \omega \) in the support of \( w \). This implies that the primal–dual dynamics (4) obtained for different realization \( \omega \) need to reach consensus on \( c \). We show that this is possible by developing an ad-hoc mean-field game model in the following section.
4 Construction of the mean-field game

Let us now introduce a two-layer network as in Fig. 2. Layer 1 involves the communication topology among different populations (henceforth referred to as macro-network) denoted as $\tilde{G}(\tilde{V}, \tilde{E})$, whereas layer 2 comprises the physical flow network (micro-network) as introduced in Sect. 2, denoted as $G(V, E)$. Each node in the macro-network $\tilde{G}(\tilde{V}, \tilde{E})$ represents a different population. All populations in $\tilde{G}$ have the same topology of their micro-network $G$. Hence, they are characterized by the same incidence matrix $\tilde{B}$ and the same cost functions $f_1$ and $f_2$. However, each agent in the population has a different realization of the demand. The aim is to obtain a feasible solution to the two-stage stochastic program (1) by reaching consensus on the capacities of the edges between any agent in the population and across all populations.

We use the macro-network to model a distributed optimization setting. In this section we present a heuristic approach that turns the two-stage stochastic program (1) into a mean-field game, where each player corresponds to an agent in each population, and it is assigned to a primal–dual dynamics according to its demand realization. This new approach represents a distributed optimization problem, which enables the agents to reach consensus on the edges’ capacities, and it provides a feasible solution to problem (1) by taking into account the communication between populations.

Consider $|\tilde{V}| = p$ populations such that each agent in our game belongs to a population $k \in \{1, \ldots, p\}$ and is characterized by a primal–dual dynamics (4) obtained for a particular realization $\omega$ of the demand. The generic agent in population $k$ is characterized by its state $x(t)$ which involves the decision variables (flows and capacities) and the Lagrange multipliers of problem (1), namely $x(t) := [u^T \ c^T \ \lambda^T \ \mu^T]^T \in \mathbb{R}^{3m+n}$. At each time $t$ and for a given time horizon window $[0, T]$, $x(t)$ evolves according to the primal–dual dynamics in (4) under the control variable $v$ as follows:

---

**Fig. 2** Two-layer network: macro-network $\tilde{G}(\tilde{V}, \tilde{E})$ in layer 1 and micro-network $G(V, E)$ in layer 2
Here, $\omega$ is a constant sample randomly extracted from the same distribution of the demand in (1) and $v$ is an additional control input that we use to force consensus on the edge capacities $c$. Note that (5) can be written in compact form as
\[
\dot{x} = Ax + Bv + C = : f(x, v).
\]
Now let us consider a probability density function that describes the density of the agents in a population in state $x$ at time $t$, $m_k(x, t)$, with the property
\[
\int_{\mathbb{R}^3} m_k(x, t) \, dx = 1.
\]
Then the mean states can be computed following
\[
m_k(t) = \int_{\mathbb{R}^3} m_k(x, t) \, dx.
\]
So far, we have considered the populations separately. To describe the interaction between populations we take into account the topology of the macro-network $\hat{G}$. Let us associate population $k$ with agent $k$. Then we can introduce an interaction topology between agents, say $\hat{G} = \{\hat{V}, \hat{E}\}$, and we can define the neighbors of an agent $k$ in $\hat{G}$ as:
\[
N(k) = \{ j \in \hat{V} \mid (k, j) \in \hat{E} \}.
\]
In order to reach consensus, each population is interested in knowing the local average state of its neighbors, which is given by:
\[
\rho_k = \frac{\sum_{j \in N(k)} m_j}{|N(k)|}.
\]
Here, $|N(k)|$ denotes the cardinality of neighbor set $k$. Let us now consider a running cost function $g(x, \rho_k, v)$ and a terminal cost function $\psi(\rho_k, x)$, which are defined as follows:
\[
g(x, \rho_k, v) = \frac{1}{2}(\rho_k - x)^T Q(\rho_k - x) + \frac{1}{2} v^T R v,
\]
\[
\psi(\rho_k, x) = \frac{1}{2} S(\rho_k - x)^2.
\]
The first term in the expression for $g$ assigns a penalty to the state deviating from the mean $\rho_k$. The second term assigns a penalty on control. The matrices $Q$, $S$ and $R$ are diagonal matrices of compatible dimensions.

Every agent in population $k$ wishes to solve the following problem:
\[
\min_{v(t)} \mathbb{E} \int_0^T \left[ g(x, \rho_k, v) \right] dt + \psi(\rho_k(T), x(T))
\]
subject to
The optimal control problem as in (7) has the structure of an optimal tracking problem. The objective function in (7) captures the minimization of the total expected cost. The aim is to reach consensus by minimizing the deviations of the state $x$ from the local average of each population $\rho_k$. This is denoted by the quadratic error term $(\rho_k - x)^T Q(\rho_k - x)$ in the running cost function $g$, and by $S(\rho_k - x)^2$ in the terminal cost function $\psi$. The term $\frac{1}{2} \nu^T R \nu$ in the cost $g$ penalises the energy of the control. The affine dynamics $\dot{x} = Ax + Bv + C$ in the constraints comes from the primal–dual dynamics explained in the previous section.

For every population $k$, denote the value of the optimization problem starting at time $t$ and state $x$ by $\mathbb{E}_k(x, t)$. Let us also denote by $m_k(x, 0) = m_{k_0}(x) \in \mathbb{R}$ the initial density of the agents of population $k$ in state $x$. This results in the following mean-field game in $\mathbb{E}_k(x, t)$ and $m_k(x, t)$:

$$\begin{align*}
\partial_t \mathbb{E}_k(x, t) + \{f(x, v^*)^T \partial_v \mathbb{E}_k(x, t) + g(x, \rho_k, v^*)\} &= 0, \\
\mathbb{E}_k(x, T) &= \mathbb{E}_k(T, x), \\
\partial_t m_k(x, t) + \text{div}(m_k(x, t) f(\cdot)) &= 0.
\end{align*}$$

Any solution of (8) is referred to as the mean-field equilibrium, which provides the sub-optimal values for the wait and see variables. Note that the second and fourth equation from (8) are the boundary conditions, while the third equation is the advection equation.

The optimal time-varying state-feedback control can be computed for every single agent in population $k$ and is given by:

$$v^*(x, t) \in \arg \min_v \{(Ax + Bv + C)^T \partial_x \mathbb{E}_k(x, t) + g(x, \rho_k(t), v)\}.$$  

5 Mean-field equilibrium and convergence

In this section, we obtain an expression for the mean-field control and provide results for the mean-field equilibrium dynamics.

**Lemma 1** The mean-field game takes the form:

$$\begin{align*}
\partial_t \mathbb{E}_k(x, t) - \frac{1}{2} (\partial_x \mathbb{E}_k(x, t))^T \left[ BR^{-1} B^T \right] \partial_x \mathbb{E}_k(x, t) + \frac{1}{2} (\rho_k - x)^T Q(\rho_k - x) &= 0, \\
\mathbb{E}_k(x, T) &= \mathbb{E}_k(T, x), \\
\partial_t m_k(x, t) + \partial_x \left[ m_k(x, t) (Ax - BR^{-1} B^T \partial_x \mathbb{E}_k(x, t) + C) \right] &= 0.
\end{align*}$$

Additionally, the optimal control is:
In this set of equations, the first equation corresponds to the Hamilton–Jacobi–Isaacs equation, and the third equation corresponds to the Fokker-Planck-Kolmogorov equation. The proof of Lemma 1 can be found in Appendix A.

We assume that the time evolution of the common state is known and, subsequently, investigate the solution of the Hamilton–Jacobi equation. Consider the following problem with known $\rho_k$:

$$\min_{v(\cdot)} \mathbb{E} \int_0^T [g(X(t), \rho_k(t), v(t))] dt$$

(12)

where

$$\dot{x} = Ax + Bv + C \in \mathbb{R}_+^m \times \mathbb{R}_+^m \times \mathbb{R}_+^n \times \mathbb{R}_+^m.$$  

The next theoretical result presents the mean-field equilibrium control. In preparation to that, let us consider a probability density function that describes the density of the agents of a population in state $c$ for the capacity of the edges at time $t$, $m_k(c, t)$, with the property $\int_{\mathbb{R}^m} m_k(c, t) dc = 1$. Then the mean states can be computed following $\bar{m}_k(t) = \int_{\mathbb{R}^m} c m_k(c, t) dc$.

We define an agent’s objective in population $k$ based on the aggregate $k$th state as:

$$\rho_k = \frac{\sum_{j \in N(k)} \bar{m}_j}{|N(k)|}.$$  

(13)

In the following, given a generic matrix $A \in \mathbb{R}^{(3m+n) \times (3m+n)}$ we denote by $A_c \in \mathbb{R}^m$ the matrix obtained extracting the rows and columns associated with the only variables $c \in \mathbb{R}^{m \times m}$. In addition, we consider the following value function:

$$\sigma_k(x, t) = \frac{1}{2} x^T \Phi(t) x + H(t)^T x + \chi(t)$$  

(14)

and denote

$$\tilde{Z} := [A^T - 2BR^{-1}B^T]^{-1}.$$  

Theorem 1 A mean-field equilibrium for the dynamics of (10) is obtained from the following set of equations:

$$\left\{ \begin{array}{l} \sigma_k(x, t) = \frac{1}{2} x^T \Phi(t) x + H(t)^T x + \chi(t), \\ \bar{m}_k(t) = [A - BR^{-1}B^T \Phi(t)] \bar{m}_k(t) - BR^{-1}B^T H(t) + C, \end{array} \right.$$  

(15)

where
\[
\begin{align*}
\dot{\Phi}(t) + A^T \Phi(t) + \Phi^T \begin{bmatrix} -BR^{-1}B^T \end{bmatrix} \Phi(t) + Q &= 0 \\
\in [0,T], \quad \Phi(T) &= S, \\
\dot{H}(t) - 2\Phi(t)BR^{-1}B^TH(t) + A^TH(t) + \Phi(t)^T C - Q\rho_k(t) &= 0 \in [0,T], \\
H(T) &= -S\rho_k(T), \\
\dot{x} + H(t)^T \begin{bmatrix} BR^{-1}B^T \end{bmatrix} H(t) + H(t)^T C + \frac{1}{2\rho_k(t)} Q\rho_k(t) &= 0 \in [0,T], \quad \chi(T) = \frac{1}{2\rho_k(T)} S\rho_k(T).
\end{align*}
\] (16)

Additionally, if the inverse matrix $\tilde{Z}$ exists, the mean-field equilibrium control is:
\[
v^* = -R^{-1}B^T(\Phi(t)^T x + H(t)).
\] (17)

If the inverse matrix $\tilde{Z}$ does not exist, let us denote by $\hat{H}(t)$ the approximation of matrix $H(t)$ obtained by applying the least squares method. Then the mean-field equilibrium control is:
\[
v^* = -R^{-1}B^T(\Phi(t)^T x + \hat{H}(t)).
\] (18)

Furthermore, for infinite time horizon $T \to \infty$, and for all aggregate states $\tilde{m} = (\tilde{m}_1, \tilde{m}_2, \ldots, \tilde{m}_p)$, we have the following consensus-type dynamics:
\[
\tilde{m} = -\tilde{Q}_1 \tilde{m} + H(t) + \delta,
\] (19)

where $\tilde{Q}_1 := \text{diag}(\tilde{Q}_i)$ (diagonal matrix with block entry $\tilde{Q}_i$), $\delta = R^{-1}_c \left( \tilde{Z}_c \Phi^T c \mu + (-\tilde{Z}_c Q_c - \Phi^T c \rho_k) + (\mu - \tilde{f}) \right)$ if the inverse matrix $\tilde{Z}$ exists, or $\delta = R^{-1}_c \left( \hat{H} - \Phi^T c \rho_k \right) + (\mu - \tilde{f})$ if the inverse does not exist, and $L$ is defined as the graph-Laplacian matrix where the $kj$th entry is the block matrix:
\[
L_{kj} = \begin{cases} 
R^{-1}_c \Phi^T c & j = k, \\
\frac{R^{-1}_c}{|N(k)|} R^{-1}_c \Phi^T c & j \neq k, j \in N(k), \\
0 & \text{otherwise}.
\end{cases}
\] (20)

This result is relevant since we can solve (15) in closed form and we show that consensus is ultimately achieved among agents. The proof of Theorem 1 can be found in Appendix B.

**Remark 1** Note that based on the way in which we have defined the mean-field game it is not possible to guarantee convergence to the optimal solution. We can only guarantee convergence of the agents to a feasible solution. It is important to highlight that the mean-field game introduced in this manuscript has similarities with the distributed subgradient algorithm used to solve distributed optimization problems [26]. From (19) we can observe that the term $\tilde{Q}_1 \tilde{m} (t)$, which corresponds to the costs of the edges capacities, is related to the gradient dynamics of variable $c$. On the other hand, the term $L\tilde{m} (t)$ enforces consensus, as it aims to minimize the
deviation of $\tilde{m}_k^c$ from $\rho_k^c$. It is worth mentioning that in the distributed subgradient method a step-size $\gamma(t)$ is applied, and convergence to a feasible solution is only guaranteed for a diminishing step size. Nevertheless, based on (19), one can assume that the step-size for the approach introduced in this paper is always $\gamma = 1$. There exist other algorithms that can be applied to solve distributed optimization problems. One of these algorithms is the gradient tracking method [26]. This method tracks the gradient of the cost function and not only the value of the optimization variables. By applying the gradient tracking algorithm it has been proven that it is possible to reach consensus to the optimal value. However, the study of the gradient tracking algorithm is outside of the scope of this research.

6 Simulation example

In this section we present a numerical example of a capacity network design for a supply chain network. We show that the agents in the macro-network reach consensus on the edges’ capacities of their respective micro-networks. We analyse the ways in which penalties on control and state influence the dynamics of the agents in the mean-field game. Furthermore, we run various simulations under different scenarios to investigate the sub-optimality of the algorithm built on the mean-field game.

Consider a scale-free network where $p = 1000$ agents. The agents only consider the local average computed over its neighbors according to (6). Furthermore, we generate stochastic demand according to the normal distribution at every time instance for the nodes of the micro-networks displayed in Fig. 1. This makes the constant term vector $C$ change over time since $\omega$ is now renewed at every time instance. Prior to the simulation, we parametrize the system as follows:

$$\omega = [0 0 23 7 0 0]^T, \quad \tilde{Q}_1 = I^{9 \times 9}, \quad \tilde{Q}_2 = I^{9 \times 9}$$

$$\tilde{f}_1 = [1 1 1 1 1 1 1 1 1]^T, \quad \tilde{f}_2 = [1 1 1 1 1 1 1 2 1]^T$$

$$C = \begin{bmatrix} -\tilde{f}_2 \cr -\tilde{f}_1 \cr -\omega \cr \mathbb{O}^{9 \times 1} \end{bmatrix}, \quad B = \begin{bmatrix} \mathbb{O}^{9 \times 1} \\ \mathbb{I}^{9 \times 1} \\ \mathbb{I}^{15 \times 1} \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} 1 0 -1 -1 0 0 0 0 0 \\ 0 1 0 0 -1 -1 0 0 0 \\ 0 0 0 0 0 1 1 0 0 \\ 0 0 1 0 0 0 0 0 1 \\ 0 0 0 1 1 0 0 -1 0 \\ 0 0 0 0 0 0 -1 1 0 \end{bmatrix}$$

$$A = \begin{bmatrix} -\tilde{Q}_2 & \mathbb{O}^{9 \times 9} & -B_l^T & -I^{9 \times 9} & \mathbb{O}^{9 \times 9} & -B_l & -I^{9 \times 9} \\ \mathbb{O}^{9 \times 9} & -\tilde{Q}_1 & \mathbb{O}^{9 \times 6} & \mathbb{I}^{9 \times 9} & \mathbb{O}^{9 \times 6} & \mathbb{I}^{9 \times 9} & \mathbb{O}^{9 \times 6} \\ I^{9 \times 9} & -I^{9 \times 9} & \mathbb{O}^{9 \times 6} & \mathbb{I}^{9 \times 9} & \mathbb{O}^{9 \times 6} & \mathbb{I}^{9 \times 9} & \mathbb{O}^{9 \times 6} \end{bmatrix}$$

$$R = 1, \quad Q = \begin{bmatrix} \mathbb{O}^{9 \times 9} & \cdots & \cdots & \mathbb{O}^{9 \times 9} \\ \cdots & \mathbb{I}^{9 \times 9} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \mathbb{O}^{9 \times 9} & \cdots & \cdots & \mathbb{O}^{9 \times 9} \end{bmatrix} \in \mathbb{R}^{33 \times 33}.$$

Note that in our example, the penalty of using flow $u_5$ is twice the penalty on the other flows since we wish to keep the flows balancing $w_4$ and $w_3$ as much distinct as possible. In general, any route can be incentivised by adjusting penalties and this emphasizes the versatility of our approach. Furthermore, in Table 1 the following additional simulation parameters are defined: time step $\delta t$, mean and standard
deviation of initial states of the primal–dual dynamics $\bar{\sigma}_0$ and $\bar{\mu}_0$; mean and standard deviation of demand at node 3, denoted $\tilde{\mu}_3$ and $\tilde{\sigma}_3$; and mean and standard deviation of demand at node 4, denoted $\tilde{\mu}_4$ and $\tilde{\sigma}_4$. The multi-population considered allows the simulation algorithm to be initialized with different initial conditions and also allows to accommodate different realizations of the \textit{wait and see} variables.

We use the parameters mentioned above as input for the following simulation algorithm by which we compute the states of each agent over time.

**Algorithm 1 Simulation algorithm**

\textbf{Input:} Simulation parameters as in Table 1, penalties $Q$ and $R$, vectors and matrices $\bar{f}_1, \bar{f}_2, \bar{Q}_1, \bar{Q}_2, \bar{B}$

\textbf{Output:} State $x(t) := [u^T, c^T, \lambda^T, \mu^T]^T \in \mathbb{R}^{3m + n}$

1: Define matrices $A$, and $B$ as in (5)
2: Compute initial state $x_0$ from a random normal distribution
3: Initialize $\rho$ according to (6), and the optimal control $v^*$ according to (17) or (18)
4: Initialize $\omega$ by a random normal distribution, and use this value to initialize $C$ as in (5)
5: Solve the continuous time Riccati equation with $A, B, Q, R$ for $\Phi$ in (16)
6: for Every time instance do
7: \hspace{0.5cm} for Every agent do
8: \hspace{1cm} Compute average state of neighbors $\rho$ according to (6)
9: \hspace{1cm} Compute demand $\omega$ by the random normal distribution
10: \hspace{1cm} Compute vector $C$ with new vector $\omega$ as in (5)
11: \hspace{1cm} Solve for $H$ the system of linear equations in (16)
12: \hspace{1cm} Compute optimal control $v^*$ according to (17) or (18)
13: \hspace{1cm} Compute new state $x(t)$ according to (5)
14: \hspace{0.5cm} end for
15: end for

Figures 3a–c depict the evolution of the edges’ capacities over time for all agents. Figure 3a shows a simulation where the penalty on control equals the penalty on state deviation, i.e., $Q = R = 1$. Figure 3b presents a simulation of the mean-field

![Fig. 3 Simulation of the mean-field game for different values of the penalty on control $R$ and the penalty on state deviations $Q$](image-url)
game where the penalty on state deviation is increased, namely \(Q = 10, R = 1\). On the contrary, the results of a mean-field game simulation where \(Q = 1, R = 10\) can be found in Fig. 3c.

By setting the penalty on state deviation sufficiently high as it was done in the simulation of Fig. 3b, the states converge to closer values compared to the case where \(Q\) and \(R\) are equal. This is in accordance to intuition as in the case of Fig. 3b states deviation from \(\rho_k\) incur in a higher penalty. Conversely, by setting the penalty on control sufficiently high, convergence is obtained quicker in comparison to the simulation where \(Q\) and \(R\) are equal. In this situation, the system is penalized more to make adjustments and this explains the \textit{thinner} lines in the converged region. Another observation is that, in this case, the best strategy to pursue the minimization problem is to choose a different solution compared to the simulations of Fig. 3a, b.

From the examples presented above, let us consider the best strategy for the minimization, which is when the penalty on the states deviations is \(Q = 1\) and the penalty on the control is \(R = 10\). To analyse the sub-optimality of the algorithm for the mean-field game, we compare the flows and capacities obtained from the consensus in the mean-field game with the values obtained by solving the primal–dual dynamics as it is explained in Sect. 3, and with the real optimal values obtained by solving (1). We compare the capacities and flows obtained from the different methods by evaluating the values of the objective function in (1). Furthermore, for the mean-field game we obtain a lower bound of the objective function by relaxing the constraint on consensus on the capacities of the edges.

In Fig. 4a we display the evolution over time of the objective function for the different methods applied to solve the two-stage stochastic program (1). Note that when the agents reach consensus on the capacities, the value of the objective function obtained from the mean-field game becomes an upper bound of the real optimal value. On the other hand, we can observe that the value of the objective function when we relax the constraint on consensus becomes a lower bound of the value of the objective function of the mean-field game when we take into account the consensus constraint.

Fig. 4 Values of the objective function and % error
Let us denote by $F_{MFG}$, $F_{NC}$, $F_{PD}$, $F_{O}$ the value of the objective function obtained from the mean field game, the value of the objective function obtained by relaxing the constraints on consensus, the value of the objective function obtained from the primal–dual dynamics and the value of the objective function obtained from the real optimum, respectively. To analyse the differences between the different approaches, we compute the percentage errors between the values of the objective function obtained from the mean-field game and the values of the objective function using the other methods as follows:

\[
\text{% Error MFG consensus vs Non Consensus} = \frac{F_{MFG} - F_{NC}}{F_{NC}},
\]
\[
\text{% Error MFG consensus vs Primal–dual} = \frac{F_{MFG} - F_{PD}}{F_{PD}},
\]
\[
\text{% Error MFG consensus vs Optimal value} = \frac{F_{MFG} - F_{O}}{F_{O}}.
\]

In Fig. 4b we plot the percentage errors between the values of the objective function obtained from the mean-field game and the values of the objective function using the other methods. In the long term, the percentage error when we compare the results of the mean-field game with the results from the primal–dual dynamics is 8.47%. On the other hand, compared to the optimal value obtained by solving (1), the percentage error in the long term is 1.96%.

We develop the same procedure explained above for fifteen different scenarios. In these scenarios we consider different distributions of the demand, different topology for the micro-network and for the macro-network, different vectors $\tilde{f}_1$ and $\tilde{f}_2$, and different values for the penalty on the states deviations and penalty on the control. We run fifty simulations of each scenario and we obtain the average of the percentage error over all simulations.

In the first three scenarios we change the mean and standard deviation of the demand at nodes 3 and 4, and we keep the rest of the simulation parameters as in the first example. For the first scenario the mean and standard deviation are $\tilde{\mu}_3 = 10$, $\tilde{\mu}_4 = 20$, $\tilde{\sigma}_3 = 3$ and $\tilde{\sigma}_4 = 1$. For the second scenario the parameters are $\tilde{\mu}_3 = 30$, $\tilde{\mu}_4 = 30$, $\tilde{\sigma}_3 = 1$ and $\tilde{\sigma}_4 = 3$. For the third scenario the mean and standard deviation are $\tilde{\mu}_3 = 5$, $\tilde{\mu}_4 = 5$, $\tilde{\sigma}_3 = 3$ and $\tilde{\sigma}_4 = 3$. The percentage error for each scenario is depicted in Fig. 5. Observe that for lower mean and higher variance, the average percentage error increases.

For the next three scenarios we change the topology of the micro-network, and we keep the rest of the simulation parameters as in the first example. In the first scenario the network consists of $n = 6$ nodes and $m = 11$ edges. For the second scenario the micro-network consists of $n = 6$ nodes and $m = 7$ edges. For the third scenario we increase the number of nodes to $n = 10$ and the number of edges to $m = 16$. The percentage error for each scenario is depicted in Fig. 6. Note that when we reduce the connections between the nodes, the average percentage error increases.

Now we consider three different scenarios by changing the topology of the macro-network, and we keep the rest of the simulation parameters as in the first
example. We design the topology of the macro-network by using a power law distribution, in which few nodes have high connectivity and most nodes have very low connectivity. In the first scenario the average degree of the nodes is four. In the second scenario the average degree is two. In the third scenario the average degree is six. The complex networks that we consider for these scenarios are depicted in Fig. 7a–c. As one can observe in Fig. 7d–f there is not a significant difference in the average percentage error between these three scenarios.

The next three scenarios are generated by changing the cost coefficients $\tilde{f}_1$ and $\tilde{f}_2$ and by keeping the rest of the simulation parameters as in the first example. In the first scenario the cost coefficients are $\tilde{f}_1^T = [1 1 1 1 1 1 1 1 1 1]$ and $\tilde{f}_2^T = [1 1 1 1 1 1 1 2 1 2 1]$. In the second scenario we have $\tilde{f}_1^T = [1 1 1 1 1 1 1 1 2 1 2 1]$ and $\tilde{f}_2^T = [1 1 2 1 1 1 1 2 1 2 1]$. In the third scenario these coefficients are $\tilde{f}_1^T = [1 1 1 2 2 1 2 3 1]$ and $\tilde{f}_2^T = [1 1 1 2 2 1 2 3 1]$. We can observe in Fig. 8 that when we increase the cost coefficients for the capacities and flows, the average percentage error with respect to the primal–dual dynamics and with respect to the real optimal value are very similar.

The last three scenarios consist in changing the penalty on state deviation $Q$ and the penalty on control $R$, and keep the rest of the simulation parameters as in the first example. In the first scenario the penalty on state deviation and the penalty on control are the same, namely $R = Q = 1$. In the second scenario the
The penalty on state deviation is $Q = 10$ and the penalty on control is $R = 1$. In the third scenario we take the penalty on state deviation $Q = 1$ and the penalty on control $R = 10$. In accordance with the first examples analysed at the beginning of this section, we observe that when the penalty on control is sufficiently high the convergence is obtained quicker, and the percentage error is smaller than in the other two scenarios (Fig. 9).

The average percentage errors, in the long-term, for the different scenarios when we compare the results of the mean-field game with the results from the primal–dual dynamics and with the optimal value obtained by solving (1) are summarized in Table 2.
In the long term, the total average of the percentage error over all scenarios and simulations when we compare the results of the mean-field game with the results from the primal–dual dynamics is 11.46%. On the other hand, compared to the optimal value obtained by solving (1), the average percentage error in the long term is 7.43%.

### Table 2: Average percentage error over different scenarios

| Scenario                        | % Error primal–dual | % error optimal value |
|---------------------------------|----------------------|-----------------------|
| Changing demand distribution    | 11.51                | 5.40                  |
| Changing micro-network topology | 7.15                 | 6.12                  |
| Changing macro-network topology | 7.89                 | 1.84                  |
| Changing cost coefficients      | 9.73                 | 9.46                  |
| Changing penalty on control and state deviation | 20.99 | 14.33 |

**Fig. 9** Average percentage error of the objective function over fifty simulations for three scenarios by changing the penalty on state deviation and the penalty on control.

**7 Concluding remarks and future work**

In this paper, we have provided a heuristic method to solve a two-stage stochastic program. We have illustrated the method on a supply-chain network design problem. The underlying idea of the method is to replace the original stochastic program with a mean-field game with a large number of agents. Each agent controls a primal–dual dynamics with the double aim of moving along the gradient descent/ascent while at the same time reaching consensus on the *here and now* variables, which are represented by the edge capacities. We show that the resulting mean dynamics mimics elastic averaging stochastic gradient algorithms. The heuristics has been implemented on a numerical example for different values of the parameters such as the penalties on state deviation and control.

Future work involves the generalization of the method to nonlinear convex problems, the analysis of the sub-optimality performance at least for specific scenarios,
the extension to mean-field dynamics with implemented (online) forecasting algorithms based on regression learning models [7, 9].

A Proof of Lemma 1

To obtain the optimal control in (11), from (9) let us write the minimized Hamiltonian as:

$$\mathcal{H}(x, \partial_x \sigma_k, m_k) = \inf_v \left\{ \frac{1}{2} v^T R v + \frac{1}{2} (\rho_k - x)^T Q (\rho_k - x) + (Ax + Bv + C)^T \partial_x \sigma_k \right\}.$$  \hfill (21)

By differentiating with respect to $v$ and equalizing the gradient to zero we obtain:

$$Rv + B^T \partial_x \sigma_k(x, t) = 0.$$  

Solving for $v$ yields the expression for $v^*$ from (11). Let us rewrite (8) in terms of the minimized Hamiltonian as:

$$\begin{cases}
\partial_t \sigma_k(x, t) + H(x, \partial_x \sigma_k(x, t), m_k) = 0 \in \mathbb{R}, \\
\sigma_k(x, T) = \frac{1}{2} (\rho_k - x)^T S (\rho_k - x) \in \mathbb{R}, \\
\partial_t m_k + \partial_x (m_k \partial_x H(x, \partial_x \sigma_k(x, t), m_k)) = 0 \in \mathbb{R}.
\end{cases}$$  \hfill (22)

Now by substituting the optimal value $v^*$ in (11), in the Hamiltonian, we obtain:

$$\begin{aligned}
\mathcal{H}(x, \sigma_k(x, t), m_k) &= \frac{1}{2} \left[ x^T Q x + v^* R v^* \right] + (\partial_x \sigma_k(x, t))^T (Ax + Bv^* + C) \\
&= - \frac{1}{2} (\partial_x \sigma_k(x, t))^T \left[ BR^{-1} B^T \right] \partial_x \sigma_k(x, t) (\partial_x \sigma_k(x, t))^T (Ax + C) + \frac{1}{2} (\rho_k - x)^T Q (\rho_k - x).
\end{aligned}$$  \hfill (23)

Using this expression for the Hamiltonian in (22), we obtain the Hamilton-Jacobi-Bellman equation in (10). The third equation from (10) can be obtained by substituting the optimal control (11) in (22). This concludes the proof. $\square$

B Proof of Theorem 1

We start the proof by isolating the Hamilton–Jacobi–Bellman part with fixed $\rho_k$ from (10) and we obtain:

$$\begin{aligned}
\partial_t \sigma_k(x, t) - \frac{1}{2} (\partial_x \sigma_k(x, t))^T \left[ BR^{-1} B^T \right] \partial_x \sigma_k(x, t) + (\partial_x \sigma_k(x, t))^T (Ax + C) \\
+ \frac{1}{2} (\rho_k - x)^T Q (\rho_k - x) = 0 \in \mathbb{R} \times [0, T], \\
\sigma_k(x, T) = \psi (\rho_k(T), x) \in \mathbb{R}.
\end{aligned}$$  \hfill (24)

Now consider the following value function:
\[ \sigma_k(x, t) = \frac{1}{2} x^T \Phi(t)x + H(t)^T x + \chi(t). \] (25)

Then (24) can be rewritten as:

\[
\frac{1}{2} x^T \Phi(t)x + \dot{H}(t)^T x + \dot{\chi}(t) - \frac{1}{2} \left[ (\Phi(t)^T x + H(t))^T \left[ -BR^{-1}B^T \right] (\Phi(t)^T x + H(t)) + \left[ \Phi(t)^T x + H(t) \right]^T (Ax + C) + \frac{1}{2} (\rho_k - x)^T Q(\rho_k - x) \right] = 0 \in \mathbb{R} \times [0, T].
\] (26)

Since the equation above is an identity in \( x \) it reduces to the following set of three equations:

\[
\begin{cases}
\Phi(t) + A^T \Phi(t) + C^T [-BR^{-1}B^T] \Phi(t) + Q = 0 \in [0, T], \quad \Phi(T) = S, \\
H(t) - 2\Phi(t)BR^{-1}B^T H(t) + A^T H(t) + \Phi(t)^T C - Q(\rho_k) = 0 \in [0, T], \quad H(T) = -S\rho_k(T), \\
\dot{\chi} + H(t)^T [BR^{-1}B^T] H(t) + H(t)^T C + \frac{1}{2} (\rho_k)^T Q(\rho_k) = 0 \in [0, T], \quad \chi(T) = \frac{1}{2} \rho_k^2(T)S.
\end{cases}
\] (27)

Furthermore, the optimal control is given by:

\[ v^* = -R^{-1}B^T (\Phi(t)^T x + H(t)). \] (28)

The existence of a solution for (27) is assured by the standard convexity-concavity assumptions, which justifies the choice for the quadratic value function [31].

Lastly, we average the expression for optimal control and by substitution in \( \frac{d}{dt} \bar{m}_k(t) = A\bar{m}_k + B\bar{v}^* + C \), we obtain the expression

\[ \bar{m}_k(t) = [A - BR^{-1}B^T \Phi(t)] \bar{m}_k(t) - BR^{-1}B^T H(t) + C, \] (29)

which is the second equation from (15). Now considering the stationary case with \( T \rightarrow \infty \), we obtain the following set of equations:

\[
A^T \Phi + C^T [-BR^{-1}B^T] \Phi + Q = 0, \\
-2\Phi BR^{-1}B^T H + A^T H + \Phi^T C - Q(\rho_k) = 0, \\
H^T [BR^{-1}B^T] H + H^T C + \frac{1}{2} (\rho_k)^T Q(\rho_k) = 0.
\]

Solving for \( H \), when the inverse of \( [A^T - 2\Phi BR^{-1}B^T] \) exists, yields the following expression:

\[ H = [A^T - 2\Phi BR^{-1}B^T]^{-1} [Q(\rho_k) - \Phi^T C]. \] (30)

We now substitute (30) in (28) and we obtain:

\[
v^* = -R^{-1}B^T \left( \Phi^T x + [A^T - 2\Phi BR^{-1}B^T]^{-1} [Q(\rho_k) - \Phi^T C] \right)
= -R^{-1}B^T \left( \Phi^T x + \bar{Z} Q(\rho_k) - \bar{Z} \Phi^T C \right),
\] (31)

where we denote \( \bar{Z} := [A^T - 2\Phi BR^{-1}B^T]^{-1} \).
Averaging over population \( k \) we obtain for the mean control the following expression
\[
\bar{v}^* = -R^{-1}B^T \left( \Phi^T \bar{m}_k + \bar{Z}Q\rho_k - \bar{Z}\Phi^T C \right).
\]

This implies that the interaction of neighbors consists of local averaging and local adjustment as in the following mean state dynamics:
\[
\dot{\bar{m}}_k(t) = A\bar{m}_k(t) - BR^{-1}B^T \left( \Phi^T \bar{m}_k - \bar{Z}\Phi^T C + \bar{Z}Q\rho_k \right) + C
\]
\[
= A\bar{m}_k(t) + BR^{-1}B^T \left( \Phi^T (\rho_k - \bar{m}_k) + \bar{Z}\Phi^T C + (-\bar{Z}Q - \Phi^T)\rho_k \right) + C
\]
\[
= A\bar{m}_k(t) + BR^{-1}B^T \Phi^T (\rho_k - \bar{m}_k) + BR^{-1}B^T \left( \bar{Z}\Phi^T C + (-\bar{Z}Q - \Phi^T)\rho_k \right) + C.
\]

Isolating the dynamics for the here and now \( c \) vector components we obtain
\[
\dot{m}_k^c(t) = -\bar{Q}_c \bar{m}_k^c(t) + R_c^{-1} \Phi^c (\rho_k^c - \bar{m}_k^c) + R_c^{-1} \left( \bar{Z}_c^c \Phi^c \mu + (-\bar{Z}_c^c Q_c^c - \Phi^c)\rho_k^c \right) + (\mu - \bar{f}_1).
\]

If the inverse matrix \( \bar{Z} \) does not exist, we can approximate the solution of \( H \) by finding the least squares solution. Let us define by \( Z := -2\Phi BR^{-1}B^T H + A^T H + \Phi^T C - Q\rho_k = (-2\Phi BR^{-1}B^T + A^T)H - (Q\rho_k - \Phi^T C) \in \mathbb{R}^{3m+n} \). Hence, finding the least squares solution consists of minimizing the sum of the squares of the difference \( (-2\Phi BR^{-1}B^T + A^T)H - (Q\rho_k - \Phi^T C) \). This can be formally represented as follows:
\[
\min_H \sum_{i=1}^{3m+n} Z_i^2 = \min_H \sum_{i=1}^{3m+n} \left( [(-2\Phi BR^{-1}B^T + A^T)H - (Q\rho_k - \Phi^T C)]_i \right)^2.
\]

The solution of the least squares method returns an approximation of \( H \), denoted by \( \hat{H} \). The mean-field equilibrium control is then given by:
\[
\bar{v}^* = -R^{-1}B^T (\Phi(t)^T \bar{m}_k + \hat{H}(t)).
\]

Hence, the mean state dynamics is as follows:
\[
\dot{\bar{m}}_k(t) = A\bar{m}_k(t) - BR^{-1}B^T \left( \Phi^T \bar{m}_k + \hat{H} \right) + C
\]
\[
= A\bar{m}_k(t) + BR^{-1}B^T \left( \Phi^T (\rho_k - \bar{m}_k) + \hat{H} - \Phi^T \rho_k \right) + C
\]
\[
= A\bar{m}_k(t) + BR^{-1}B^T \Phi^T (\rho_k - \bar{m}_k) + BR^{-1}B^T \left( \hat{H} - \Phi^T \rho_k \right) + C.
\]

By isolating the dynamics for the here and now \( c \) vector components, we obtain:
\[
\dot{m}_k^c(t) = -\bar{Q}_c \bar{m}_k^c(t) + R_c^{-1} \Phi^c (\rho_k^c - \bar{m}_k^c) + R_c^{-1} \left( \hat{H} - \Phi^c \rho_k^c \right) + (\mu - \bar{f}_1).
\]

For all aggregate states \( \bar{m}^c = (\bar{m}_1^c, \bar{m}_2^c, \ldots, \bar{m}_\rho^c) \), we have the following consensus-type dynamics:
\[ \dot{m}^c = -\tilde{Q} \tilde{m}^c(t) - Lm^c(t) + \delta. \] (38)

Here, \( \tilde{Q}_1 := \text{diag}(\tilde{Q}_1) \) (diagonal matrix with block entry \( \tilde{Q}_1 \)), \( \delta := R_{c}^{-1}(\tilde{Z}_c \Phi^T \mu + (-\tilde{Z}_c Q_c - \Phi^T p_c^k) + (\mu - \tilde{f}_1) \) if the inverse matrix \( \tilde{Z} \) exists, or \( \delta = R_{c}^{-1}(\tilde{H} - \Phi^T p_c^k) \) + \( (\mu - \tilde{f}_1) \) if the inverse does not exist, and \( L \) is defined as the graph-Laplacian matrix where the \( k \)th entry is the block matrix:

\[
L_{kj} = \begin{cases} 
R_{c}^{-1} \Phi^T, & j = k, \\
\frac{1}{|N(k)|} R_{c}^{-1} \Phi^T, & j \neq k, j \in N(k), \\
0, & \text{otherwise}.
\end{cases}
\] (39)

This concludes the proof of Theorem 1. \( \square \)

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. 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)
2. Arrow, K.J., Hurwicz, L., Uzawa, H.: Studies in Linear and Non-linear Programming. Stanford University Press, Redwood City (1958)
3. Bardi, M.: Explicit solutions of some linear-quadratic mean field games. Netw. Heterog. Media 7, 243–261 (2012)
4. Bauso, D.: Consensus via multi-population robust mean-field games. Syst. Control Lett. 107, 76–83 (2017)
5. Bauso, D.: Dynamic demand and mean-field games. IEEE Trans. Autom. Control 62(12), 6310–6323 (2017)
6. Bauso, D., Mylvaganam, T., Astolfi, A.: Crowd-averse robust mean-field games: approximation via state space extension. IEEE Trans. Autom. Control 61(7), 1882–1894 (2016)
7. Bauso, D., Namerikawa, T.: 20. Electric vehicles and Mean-field, from Advanced Data Analytics for Power Systems edited by Ali Tajer; Samir M. Perlaza; H. Vincent Poor. Cambridge University Press (in print)
8. Boffi, N.M., Slotine, J.-J.: A continuous-time analysis of distributed stochastic gradient. Neural Comput. 32, 36–96 (2020)
9. Boyd, S., Parikh, N., Chu, E.: Distributed Optimization and Statistical Learning Via the Alternating Direction Method of Multipliers. Now Publishers Inc, New York (2011)
10. Briceno, A.L., Kalise, D., Kobeissi, Z., Laurière, M., González, A.M., Silva, F.J.: On the implementation of a primal-dual algorithm for second order time-dependent mean field games with local couplings. ESAIM Proc. Surv. 65, 330–348 (2019)
11. Chen, J., Lau, V.K.N.: Convergence analysis of saddle point problems in time varying wireless systems-control theoretical approach. IEEE Trans. Signal Process. 60(1), 443–452 (2011)
12. Chen, X., Li, N.: Exponential stability of primal-dual gradient dynamics with non-strong convexity. In: Proceedings of the 2020 American Control Conference (ACC). arXiv:1905.00298 (2020)
13. Cherukuri, A., Gharesifard, B., Cortes, J.: Saddle-point dynamics: conditions for asymptotic stability of saddle points. SIAM J. Control. Optim. 55(1), 486–511 (2017)
14. Cherukuri, A., Mallada, E., Cortés, J.: Asymptotic convergence of constrained primal-dual dynamics. Syst. Control Lett. 87, 10–15 (2016)
15. Cherukuri, A., Mallada, E., Low, S., Cortés, J.: The role of convexity in saddle-point dynamics: Lyapunov function and robustness. IEEE Trans. Autom. Control 63(8), 2449–2464 (2017)
16. Feijer, D., Paganini, F.: Stability of primal-dual gradient dynamics and applications to network optimization. Automatica 46(12), 1974–1981 (2010)
17. Ferragut, A., Paganini, F.: Network resource allocation for users with multiple connections: fairness and stability. IEEE ACM Trans. Netw. (TON) 22(2), 349–362 (2014)
18. Gomes, D.A., Sáide, J.: Mean field games models—a brief survey. Dyn. Games Appl. 4(2), 110–154 (2014)
19. Jönsson, U.T.: Primal and dual criteria for robust stability and their application to systems interconnected over a bipartite graph. In: Proceedings of the 2010 American Control Conference, pp. 5458–5464. IEEE (2010)
20. Kose, T.: Solutions of saddle value problems by differential equations. Econom. J. Econom. Soc. 24(1), 59–70 (1956)
21. Lasry, J.-M., Lions, P.-L.: Jeux à champ moyen. i le cas stationnaire. Comptes Rendus Mathematique 343(9), 619–625 (2006)
22. Lasry, J.-M., Lions, P.-L.: Jeux à champ moyen. ii horizon fini et controle optimal. Comptes Rendus Mathematique 343(10), 679–684 (2006)
23. Lasry, J.-M., Lions, P.-L.: Mean field games. Jpn. J. Math. 2, 229–260 (2007)
24. Liang, S., Wang, L.Y., Yin, G.: Exponential convergence of distributed primal-dual convex optimization algorithm without strong convexity. Automatica 105, 298–306 (2019)
25. Nguyen, H.D., Vu, T.L., Turitsyn, K., Slotine, J.-J.: Contraction and robustness of continuous time primal-dual dynamics. IEEE Control Syst. Lett. 2(4), 755–760 (2018)
26. Notarstefano, G., Notarnicola, I., Camisa, A.: Distributed optimization for smart cyber-physical networks. Found. Trends Syst. Control 7(3), 253–383 (2019)
27. Qu, G., Li, N.: On the exponential stability of primal-dual gradient dynamics. IEEE Control Syst. Lett. 3, 43–48 (2019)
28. Shapiro, A., Dentcheva, D., Andrzej, R.: Modeling and Theory. MOS-SIAM Series on Optimization, Lectures on Stochastic Programming (2009)
29. Stolyar, A.L.: Greedy primal-dual algorithm for dynamic resource allocation in complex networks. Queueing Syst. 54(3), 203–220 (2006)
30. Tang, Y., Guannan, Q., Li, N.: Semi-global exponential stability of primal-dual gradient dynamics for constrained convex optimization. Syst. Control Lett. 144, 104754 (2020)
31. Tannenbaum, A.: 𝔇∞-optimal control and related minimax design problems (tamer basar and pierre bernhard). SIAM Rev. 35(3), 538–540 (1993)
32. Yang, C., Dai, H., Li, J., Zhang, Y., Han, Z.: Distributed interference-aware power control in ultra-dense small cell networks: a robust mean field game. IEEE Access 6, 12608–12619 (2018)
33. Yang, C., Li, J., Semasinghe, P., Hossain, E., Perlaza, S.M., Han, Z.: Distributed interference and energy-aware power control for ultra-dense d2d networks: a mean field game. IEEE Trans. Wirel. Commun. 16(2), 1205–1217 (2016)
34. Zhao, C., Topcu, U., Low, S.: Swing dynamics as primal-dual algorithm for optimal load control. In: 2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm), pp. 570–575. IEEE (2012)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.