Article:
Bauso, D. orcid.org/0000-0001-9713-677X and Başar, T. (2016) Strategic thinking under social influence: Scalability, stability and robustness of allocations. European Journal of Control, 32. pp. 1-15. ISSN 0947-3580

https://doi.org/10.1016/j.ejcon.2016.04.006

Strategic thinking under social influence: scalability, stability and robustness of allocations

Dario Bauso[a], Tamer Başar[b]

[a] *Department of Automatic Control and Systems Engineering, The University of Sheffield, Mappin Street Sheffield, S1 3JD, United Kingdom, and Dipartimento di Ingegneria Chimica, Gestionale, Informatica, Meccanica, Università di Palermo, V.le delle Scienze, 90128 Palermo, Italy.* d.bauso@sheffield.ac.uk

[b] *CSL, University of Illinois at Urbana-Champaign, US, IL 61801, USA* basar1@illinois.edu

Abstract

This paper studies the strategic behavior of a large number of game designers and studies the scalability, stability and robustness of their allocations in a large number of homogeneous coalitional games with transferable utilities (TU). For each TU game, the characteristic function is a continuous-time stochastic process. In each game, a game designer allocates revenues based on the extra reward that a coalition has received up to the current time and the extra reward that the same coalition has received in the other games. The approach is based on the theory of mean-field games with heterogeneous groups in a multi-population regime.

Keywords: Mean-field games, Coalitional Game Theory, Differential games, Optimal Control

1. Introduction

"System of systems" or "network of networks" were keywords in the FP7 research agenda of the European Community and are still dominating the research agenda of Horizon 2020 (see e.g. http://ec.europa.eu/digital-agenda/en/system-systems). The terms "System of systems" or "network of networks" have often appeared in a variety of scientific papers in social sciences, engineering, and economics [16]. Despite the diversity in their domains of application, these systems have a common characteristic: the elementary units of the system or network are themselves systems or networks (see the expository article [28]). This is illustrated in abstract terms in Fig. 1, where we have multiple groups of people interacting (left). At the higher level, the network topology describes the connections between the groups. At the lower level, a second network topology describes the interactions among people within the same group. Thus, we identify two dimensions for the problem at hand. One is the size of the higher-level network, i.e., the number of groups. The other one is the size of each group, measured by the number of people in that group. As a consequence, the complexity of the system may increase because of both the large number of groups (middle), and the increasing number of people in each group (right). In the latter case, we end up with a finite network of populations, each population involving an infinite number of individuals in the limit. On the other hand, when the number of groups grows, the resulting system can be viewed as a population composed of an infinite number of groups, each of finite size. This paper deals with the second scenario, i.e., a single population made up of an infinite number of homogeneous groups, with each group comprised of a finite number of individuals. Individuals within each group are heterogeneous as they can form coalitions characterized by different values. The number of players within each group is the same.

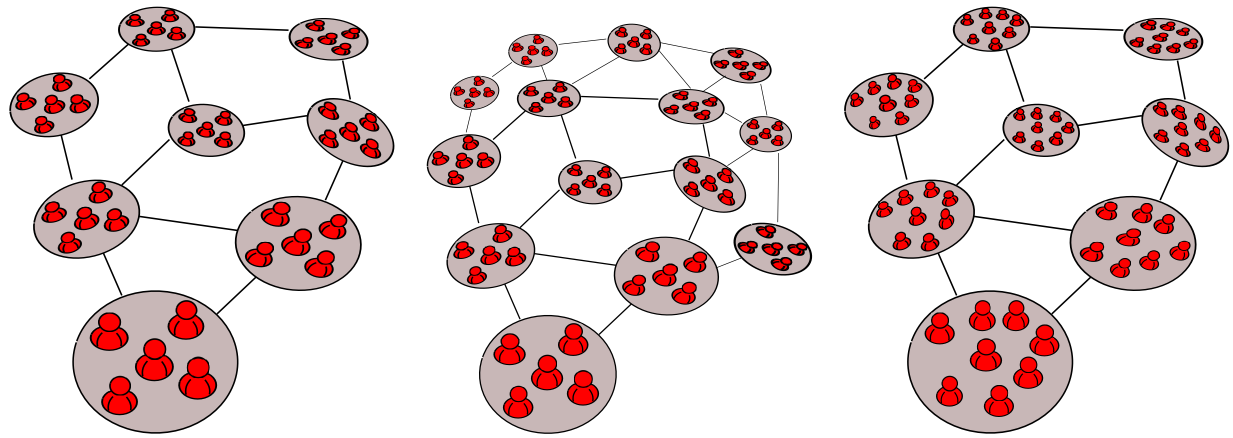

Figure 1: Social networks with multiple groups: 8 groups, each one consisting of 5 players (left); increasing number of groups, i.e., 14 groups, each one consisting of 5 players (middle); increasing number of players in each group, i.e, 8 groups, each one consisting of 8 players (right).

Placing the contribution of this paper in proper context, we consider a large number of the same copy of a coalitional game with transferable utilities (TU game). In other words, each group forms a coalitional TU game and the individuals of each group are the players of such a game. For each TU game, we also have a designer who distributes revenues to the players according to certain criteria introduced in the paper. In other words, the designer is the one who distributes the revenues among the players within each group to make the group (grand coalition) stable. Each group (TU game) has its own designer. In the second part of the paper we also consider the case where we have heterogeneous populations of TU games. By this we mean multiple populations of TU games and a structured environment, where interactions between populations occur locally.

A model with an infinite number of copies of the same TU game is interesting and relevant primarily for three reasons. First multiple copies describe scenarios where the stability of a coalition depends on the revenue expectations and perceptions that the members of a coalition develop. Such expectations and perceptions depend on the revenue allocation observed in the other games, and we refer to this as social pressure factors. Second, as a consequence of the previous point, multiple copies are useful to capture expectations and perceptions in the form of how much the same coalition receives in the other games on the average. Hence, we say that a coalition in one game can be stable if it receives at least as much as the average allocation computed over the other games. This point is also related to inequity aversion as explained in detail in Subsection 1.1.1. Third, multiple copies may be used to model different realizations of a stochastic process.

This paper extends the results presented earlier in the conference versions [6, 7]. One of the main new contributions is the analysis of the multi-population case. The paper also provides more extensive background on the theory of coalitional games with transferable utility and a detailed treatment of the mean-field response. Some of the simulations in Section 8 are also new.

1.1. Highlights of the main points

The novelty of this paper is that the problem of finding stable allocations in a TU game is now approached in the context of an environment that is dynamic and uncertain, with a large number of indistinguishable designers sharing aggregate opinions on the coalitions' average excesses and readjusting their allocations based on such information. We redefine *stability* of allocations in terms of *fairness* as follows: allocations are stable if the accumulated excesses in each single game correspond to the average excesses computed over the population. In other words, by fairness we mean that there is no *inequity* in the allocations in any of the TU games.

The proposed game-theoretic approach provides fundamental insights into the collective behavior deriving from individual rationality on the part of the designers. More specifically, this study focuses on the following three aspects: i) *strategic behavior* of the game designers, ii) *mean-field approach: consistency, scalability, and stability* of the allocation policies, and iii) *robustness* of the allocation policies in the face of inaccurate forecast of excesses or model misspecifications. In the following we further elaborate on these aspects, separately.

1.1.1. Strategic behavior and game theory

Strategic behavior means that each designer is rational and reasons strategically. Strategic behavior is described by the following game-theoretic model.

Each single game has a game designer who allocates rewards or revenues based on the excesses of the coalitions. In the context of repeated interactions, the excess of a coalition is the cumulative deviation of the total amount given to the coalition from the value of the coalition up to the current time. In classical TU games, the coalitions' values are constant and known, and the ultimate goal of the game designer is to stabilize the grand coalition. This occurs when the total revenue assigned to all members of any sub-coalition is greater than the value of the sub-coalition itself (see the notion of "core" in [31]). Differently, in this work the coalitions' values are *time-varying* and thus the excesses evolve according to controlled *uncertain* stochastic differential equations. The objective of the game designer is to align the excesses with the average value computed over the other copies of the same game. Such a phenomenon is known as crowd-seeking behavior in mean-field games and mirrors a typical attitude in macroeconomics known as *inequity aversion*. The latter is a research area in behavioral economics studying people's resistance to inequitable outcomes [15].

In each game, the stochastic differential equation describing the time evolution of its excesses is referred to as *microscopic* dynamics to distinguish it from the dynamics of the aggregate excesses of the whole population, the latter called *macroscopic* dynamics.

In addition to the state dynamics, each TU game is programmed with a given finite-horizon cost functional that accounts for i) deviation of the excesses from the average one, which we call *inequity*, ii) energy expenditure due to the revenue allocations, and iii) energy expenditure due to the allocations' miss-specifications. More formally, the deviation in i) is measured by an error vector used in a cross-coupling term (we will see later that this is a mean-field term) that incentivizes the designer to allocate more revenue to those coalitions whose excesses are below the average value and to decrease the allocations of those coalitions whose excesses are above the average value. In other words, the cross-coupling term models *fairness (the social cost)* in that it forces the designers to shift allocations from high to low-revenued ones.

Given a cost functional of the type discussed above, the designers adjust their allocation policies in order to optimize it. The strategic behavior enters through the contribution of the TU game to the social cost. Actually, strategic thinking involves also the capability of the designer to predict the macroscopic effect on the average excesses produced by all designers acting rationally and selecting a proper best-response allocation policy. The average excesses evolve then based on the current and predicted best-response allocation policies.

1.1.2. Mean-field approach: consistency, scalability, and stability

We consider numerous and indistinguishable copies of the TU game. Indistinguishability means that two designers under identical conditions will react in the same way. The idea is then to build a simulator model that returns an excesses landscape, i.e., the mean-field model, based on past data. Each designer uses such a landscape of excesses to derive an aggregate information on the state of the world, captured by an *average excess vector*. The mean-field game involves a macroscopic description based on the classical forward Kolmogorov partial differential equation which generates the distribution of the excesses over the populations and over the horizon.

In this context, the results obtained in this paper shed light on the existence of *mean-field equilibrium* solutions. By this we mean allocation policies based on the current and forecasted average excesses, which are proven to guarantee fairness in the long run. Allocation strategies are designed as closed-loop feedback strategies on the current excesses. Such strategies are computed over a finite horizon and are therefore based on forecasted average excesses. From another angle, we may say that mean-field equilibrium strategies represent the asymptotic limits of Nash equilibrium strategies, and as such they are the best-response strategies of each single player, given fixed behaviors of the other players. Given such a system, we study the mean-field equilibrium for the underlying deterministic mean-field game. The relevance of such a mean-field equilibrium is that i) it guarantees consistency of the model, ii) equilibrium strategies are scalable, and iii) the resulting dynamics are stable. More specifically,

- (consistency) From the definition of mean-field equilibrium itself, the resulting microscopic and macroscopic models are *consistent*. The mean-field equilibrium is indeed a fixed point of the set of partial differential equations characterizing the best-responses in each single game and the macroscopic evolution of the system as a whole if all the players act rationally. Best-response allocation policies are obtained via dynamic programming while the macroscopic dynamics provide the time history of the distribution of excesses over the population. Best-response allocation policies, which provide consistency, are also called *mean-field equilibrium policies*.

- (scalability) The model developed is *scalable* in the sense that the allocation policies depend on the excesses as well as an aggregate description of the world in terms of past, current, and forecasted average excesses. In other words, the allocation strategies are built on aggregate information and therefore their structures do not depend on the number of players. Such policies are obtained by carrying out a robust control design based on augmentation and regularization of the state space [13].

- (stability) We perform a stability analysis on the microscopic dynamics of the excesses as well as the average excesses. According to this analysis, the stochastic processes describing the excesses are mean square bounded. Asymptotic stability means that both excesses of each TU game converge in a stochastic sense to the reference values. The proven stability of the microscopic dynamics confirms the asymptotic convergence of the TU game excesses to an equilibrium point, this being expressed in terms of average excesses.

We should note that indistinguishability is not a limitation as far as the general approach of this paper goes. In fact, Section 7 of the paper deals with a case of heterogeneity of the TU games, and provides a more complex multi-population model.

1.1.3. Model misspecification and robustness

As noted in the previous subsection, mean-field equilibrium policies are built on local information, as well as some global information. The local information concerns the *local state* of the TU game (the excesses) and its predicted dynamics. The global information is about the average excesses, which can be thought of as a *common state*, which depends on the present and forecasted population behavior. One question of interest is what happens if the estimation of the current local and common states or their forecasted dynamics are misspecified? In other words, whether the above mean-field equilibrium policies determining the allocation policies are still scalable, consistent and stable even in the case where the microscopic and macroscopic dynamics are misspecified and/or are uncertain. In the case of imperfect modeling, model misspecification is considered in both stochastic and deterministic worst-case scenarios. Assuming imperfect models with stochastic or worst-case deterministic disturbances acting on the state dynamics, the paper provides conditions for convergence of the microscopic dynamics. Several cases studied in the paper show that the best-response allocation policies perform well in the cases of both perfect and imperfect modeling.

More specifically, stochastic disturbance has the form of a Brownian motion in the microscopic dynamics. After establishing a mean-field equilibrium, we provide some results on stochastic stability. In the scenario considered, the stochastic disturbance is independent of the state, and the Brownian motion coefficients are constants. This leads to dynamics which resemble what would be generated by the Langevin equation. Following well-known results on the Langevin equation, the dynamic is proven to be stochastically stable in the second-moment. An exposition of stochastic analysis and stability can be found in [22]. A further result deals with robustness for the microscopic dynamics. The dynamics are now influenced by an additional adversarial disturbance, with bounded resource or energy. For this case also, we study the mean-field equilibrium and investigate the conditions that guarantee worst-case stability.

1.2. Related Literature

We discuss here two streams of literature relevant to the topic of this paper, one pertaining to coalitional TU games and the other one to the theory of differential games with a large number of indistinguishable players.

1.2.1. Coalitional TU games

Coalitional games with transferable utilities (TU), introduced first by Von Neuman and Morgenstern [31], have recently sparked much interest in the control and communication engineering communities [27]. In essence, coalitional TU games are comprised of a set of players who can form coalitions, and a characteristic function associating a real number with every coalition. This real number represents the value of the coalition and can be thought of as a monetary value that can be distributed among the members of the coalition according to some appropriate fair allocation rule. The value of a coalition also reflects the monetary benefit demanded by that coalition to be part of the grand coalition. In the context of coalitional TU games, *robustness* and *dynamics* naturally arise in all the situations where the coalition values are uncertain and time-varying; see e.g., [10, 11, 25].

There is also connection with the set invariance theory [3] and stochastic stability theory [1, 17, 22], which provides some useful tools for stability analysis.

1.2.2. Mean-field games

The theory of mean-field games originated in the work of M.Y. Huang, P. E. Caines and R. Malhamé [18, 19] and independently in that of J. M. Lasry and P.L. Lions [21], where the now standard terminology of Mean Field Game (MFG) was introduced (see also [30]). The problem we study in this paper follows in spirit the study on robust dynamical TU coalitional games in [12] with additional mean-field interactions between infinite copies of the same game, which was not present in [12]. Explicit solutions in terms of mean-field equilibria are not common unless the problem has a linear-quadratic structure, see [2]. This is the motivation behind the method presented in this paper, which approximates the original problem by an augmented linear quadratic one. Heterogeneity is studied in mean-field games with major and minor players [20, 26]. Directions for further developments are i) the formulation, analysis and design of Stackelberg mean-field games where we have leaders and followers [24], ii) the analysis of mean-field games over structured environments by using networks [14], and iii) the applications to demand-side management and intelligent mobility [23].

The rest of the paper is organized as follows. In Section 2, we introduce the problem and the model. In Section 3, we present the mean-field game. In Section 5, we describe the solution approach, and show stability in Section 6, with extension to the heterogeneous case in Section 7. In Section 8, we illustrate the main results through simulations. Finally, in Section 9, we draw some conclusions and discuss possible future directions.

Notation. $\mathbb{R}_+$ denotes the set of nonnegative real numbers. Given a random vector ξ, $\mathbb{E}[\xi]$ denotes its expected value. Given a Brownian motion (with drift) $\mathcal{B}(t)$, we denote by $d\mathcal{B}(t)$ its infinitesimal increment, i.e., $\mathcal{B}(t) = \int_0^t d\mathcal{B}(\tau)$, the latter being the Itô integral. We use $\bar{\mathcal{B}}(t) = \frac{\mathcal{B}(t)}{t}$ to indicate the average infinitesimal up to time t. If $a(t)$ is the derivative of an almost everywhere differentiable function, the symbol $\tilde{a}(t)$ denotes the function itself, i.e., $\tilde{a}(t) = \int_0^t a(\tau)d\tau$. We also use $\bar{a}(t) = \frac{\tilde{a}(t)}{t}$ to indicate the average up to time t.

1.3. Preliminary observations

This subsection provides an overview of coalitional TU games and highlights connections with network flow control. In particular, for the first part, we introduce coalitional TU games and the well-known solution concepts of imputation set and core. The material of the first part is based on [5], Chapter 5. The second part recalls recent results on how to turn the allocation problem into a network problem with controlled and uncontrolled flows. This second part builds on recent results obtained in [12].

1.3.1. Coalitional TU game

Given a set $N = \{1, \dots, n\}$ of players and a function $\eta : S \mapsto \mathbb{R}$ defined for each nonempty coalition $S \subseteq N$, we write $< N, \eta >$ to denote the transferable utility (TU) game with players set N and characteristic function η.

Let $\mathcal{S} = 2^N \setminus \emptyset$ be the set of all possible coalitions of n players except for the empty set. Let us introduce some arbitrary mapping of $\mathcal{S}$ into $M := \{1, \dots, q\}$ where $q = 2^n - 1$ is the number of nonempty coalitions, namely, the cardinality of $\mathcal{S}$. Denote a generic element of M by j. In other words, we can see j standing for the labeling of the jth element of $\mathcal{S}$, say S_j, according to some arbitrary but fixed ordering. Let q be the index of the grand coalition N.

We let η_j be the value of the characteristic function η associated with a nonempty coalition $S_j \subseteq N$. Given a TU game, the main question is how to distribute the costs or rewards among the participants of the coalition.

1.3.2. Imputation set

A partial answer to the above question lies in the concept of the *imputation set*. Let v be the characteristic function determining the coalitions' values, the imputation set $I(v)$ is the set of allocations that are

- *efficient*, that is, the sum of the components of the allocation vector is equal to the value of the grand coalition, and
- *individually rational*, namely no individual benefits arise from leaving the grand coalition and playing alone.

More formally, the imputation set is a convex polyhedron defined as

$$I(v) = \{x \in \mathbb{R}^n | \overbrace{\sum_{i \in N} x_i = v(N),}^{\text{Efficiency}} \ \underbrace{x_i \geq v(\{i\}), \forall i \in N}_{\text{individual rationality}} \}.$$

If the imputation set is empty, then, given any efficient allocation, there is always at least one player who is better off by quitting the grand coalition. The imputation set is nonempty if and only if the sum of the values of the single players is not greater than the value of the grand coalition:

$$I(v) \neq \emptyset \quad \text{iff } v(N) \geq \sum_{i \in N} v(\{i\}).$$

It turns out that the imputation set $I(v)$ is the convex hull of the points $f^1, f^2, \ldots, f^n$ where

$$f_k^i = \begin{cases} v(N) - \sum_{k \in N \setminus \{i\}} v(\{k\}) & k = i \\ v(\{k\}) & k \neq i. \end{cases}$$

Note that the generic vector f^i can be interpreted as a tentative allocation recommended by player i and obtained as follows. Given that f_k^i is the revenue that player i is willing to allocate to player k, then player i allocates to any other player the amount of his exact value, that is $f_k^i = v(\{k\})$, for $k \neq i$, and takes all the rest for himself, i.e., $f_i^i = v(N) - \sum_{k \in N \setminus \{i\}} v(\{k\})$.

1.3.3. Core

The *core* is a stronger solution concept than the imputation set and it is common in economic applications. The core strengthens the conditions valid for the imputation set in that the players do not benefit from not only quitting the grand coalition and playing alone, but also creating any sub-coalition. Thus the core is still a polyhedral set which is included in the imputation set.

Definition 1. *The core of a game $\langle N, v \rangle$ is the set of allocations that satisfy i) efficiency, ii) individual rationality, and iii) stability with respect to subcoalitions:*

$$C(v) = \{x \in I(v) | \underbrace{\sum_{i \in S} x_i \geq v(S), \forall S \in 2^N \setminus \emptyset}_{\textit{stability w.r.t. subcoalitons}} \}.$$

Given a TU game $< N, \eta >$, we use $C(\eta)$ to denote the core of the game:

$$C(\eta) = \left\{ x \in \mathbb{R}^n \;\middle|\; \sum_{i \in N} x_i = \eta_q, \ \sum_{i \in S_j} x_i \geq \eta_j \text{ for all nonempty } S \subset N \right\}.$$

We illustrate next the above concepts in a simple example.

Example 1. (Connecting multiple communities to a power source) [29] *Three communities need to be connected to a power source. Connections between the communities and between a community and the power source entail certain costs. This application can be modeled as a minimum spanning tree game. The problem can be described by a graph as in Fig. 2, where the nodes correspond to the communities and the power source, the links represent the connections, and the coefficients on each link is the corresponding cost. Direct connections to the source cost* 100, 90, *and* 80, *respectively. Communities* 1 *and* 2 *can both be connected to the power source at a cost of* $90+40=130$. *Such a cost corresponds to the tree* $\{(source, 2), (2, 1)\}$. *Likewise, the cheapest connection for communities* 1 *and* 3 *has a cost of* $80+30=110$, *which corresponds to the tree* $\{(source, 3), (3, 1)\}$. *Analogously, communities* 2 *and* 3 *can be connected to the power source at a minimum cost of* $80+30=110$, *in which case the solution is represented by the tree* $\{(source, 3), (3, 2)\}$. *If all communities collaborate, they can all be connected to the power source at a minimum cost of* $80+30+30=140$, *which corresponds to the spanning tree* $\{(source, 3), (3, 1), (3, 2)\}$. *The resulting TU game* $\langle N, c\rangle$ *is then given by:* $N=\{1,2,3\}$,

$$\begin{gathered}
c(\{1\})=100, \qquad c(\{2\})=90, \qquad c(\{3\})=80, \\
c(\{1,2\})=130, \quad c(\{1,3\})=110, \quad c(\{2,3\})=110, \\
c(\{1,2,3\})=140.
\end{gathered}$$

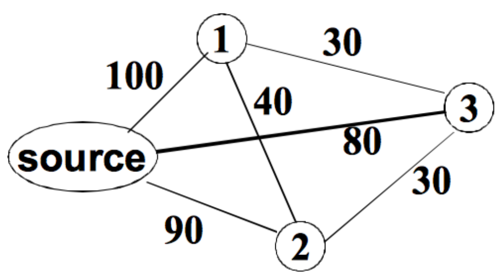

Figure 2: Minimum spanning tree problem as TU game.

We can compute the amounts saved by the three communities when they collude. Thus the value of the coalitions is given by

$$v(S)=\sum_{i\in S} c(\{i\})-c(S).$$

The corresponding cost saving game $\langle N, v\rangle$: $N=\{1,2,3\}$, *is then given by*

$$\begin{gathered}
v(\{1\})=0, \qquad v(\{2\})=0, \qquad v(\{3\})=0, \\
v(\{1,2\})=60, \quad v(\{1,3\})=70, \quad v(\{2,3\})=60, \\
v(\{1,2,3\})=130.
\end{gathered}$$

As for the physical interpretation, the above means that if the three communities sign a contract in which they commit to create a power network with minimum cost $\{(source, 3), (3, 1), (3, 2)\}$, *the three communities will save* $130\,\$$. *Any redistribution of such money that gives players 1 and 2 less than* $60\,\$$ *cannot be accepted by the two players, who will withdraw from the negotiation. We note that in the case where all the players agree to play in the grand coalition, the game takes on the form of a classical minimum spanning tree problem.*

1.3.4. Relation with network flow control

Consider an n-player robust dynamical TU game $< N, \eta(t) >$, where $\eta(t)$ is the characteristic function representing the values of different coalitions. Let $q = 2^n - 1$ be the number of coalitions. The problem of finding stable allocations admits a network representation [12]. For each game, let $\mathcal{H}$ be a corresponding hypergraph involving the vertex set V and edge set E, namely:

$$\mathcal{H} := \{V, E\}, \quad V = \{\mathbf{v}_1, \ldots, \mathbf{v}_q\}, \quad E := \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}.$$

In other words, the vertex set V has one vertex per coalition whereas the edge set E has one edge per player. The incidence relations establish that a generic edge i is incident to a vertex

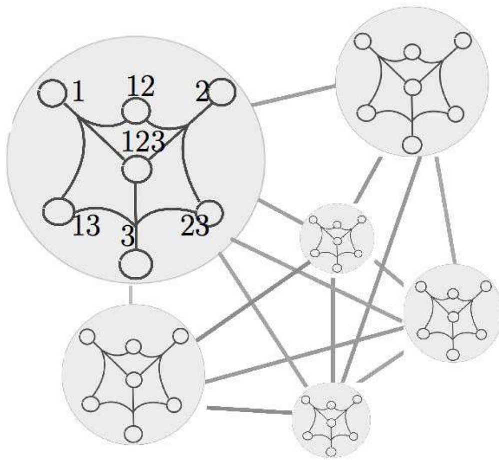

Figure 3: Infinite copies of hypergraph $\mathcal{H} := \{V, E\}$ for a 3-player coalitional game.

$\mathbf{v}_j$ if the player i is a member of the the coalition linked to $\mathbf{v}_j$. The hypergraph can then be described by an incidence matrix $B_{\mathcal{H}}$ whose rows are the characteristic vectors $c^S \in \mathbb{R}^n$. The characteristic vectors are in turn binary vectors where $c_i^S = 1$ if $i \in S$ and $c_i^S = 0$ if $i \notin S$. Figure 3 depicts an example of a hypergraph for a 3-player coalitional game on every single grey node. Following the same approach as in [12], the allocation $\tilde{u}_i(t)$ is represented by the flow on edge $\mathbf{e}_i$ and the coalition value $\eta_j(t)$ of a generic coalition $S_j \in \mathcal{S}$ is the demand in the corresponding vertex $\mathbf{v}_j$. It is apparent then that any allocation in the *core* of the game $C(\eta(t))$ translates into over-satisfying the demand at the vertices. In particular, we have

$$\tilde{u}(t) \in C(\eta(t)) \quad \Leftrightarrow \quad B_{\mathcal{H}}\tilde{u}(t) \geq \eta(t), \tag{1}$$

where the inequality is to be interpreted componentwise, and for the grand coalition is satisfied with equality due to the efficiency condition of the core, i.e, $\sum_{i=1}^n \tilde{u}_i(t) = \eta_q(t)$, where $\eta_q(t)$ denotes the qth component of $\eta(t)$ and is equal to the grand coalition value.

Let

$$B = \left[B_{\mathcal{H}} \left| \begin{array}{c} -I \\ ---- \\ 0 \ldots 0 \end{array} \right. \right] \in \{-1, 0, 1\}^{q \times n + (q-1)}. \tag{2}$$

Inequality (1) can be rewritten as an equality by using an augmented allocation vector given by $u := \begin{bmatrix} \tilde{u} \\ s \end{bmatrix} \in \mathbb{R}^{n+q}$ where s is a vector of q nonnegative surplus variables. Then, we have

$$\left[B_{\mathcal{H}} \left| \begin{array}{c} -I \\ ---- \\ 0 \ldots 0 \end{array} \right. \right] \begin{bmatrix} \tilde{u} \\ s \end{bmatrix} = \begin{bmatrix} \eta_1(t) \\ \vdots \\ \eta_q(t) \end{bmatrix}.$$

	$\mathbf{v_1}$	$\mathbf{v_2}$	$\mathbf{v_3}$	$\mathbf{v_4}$	$\mathbf{v_5}$	$\mathbf{v_6}$	$\mathbf{v_7}$
S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$

$$\underbrace{\left[\begin{array}{ccc|cccccc} 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]}_{B} \underbrace{\begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \\ \tilde{u}_3 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \end{bmatrix}}_{u} = \underbrace{\begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \end{bmatrix}}_{\eta}$$

Table 1: Mapping coalitions into vertices (top) and corresponding dynamical system of a 3-player coalitional game.

Note that each surplus variable s_j corresponds to a coalition S_j of players and describes the difference between the allocated value and the coalitional value, $s_j(t) = \sum_{i \in S_j} \tilde{u}_i(t) - \eta_j(t)$. A positive value for $s_j(t)$ can be interpreted as a debit for the coalition, whereas a negative value can be interpreted as a credit. The main insights we borrow from [12] is that *if all the surpluses are nonnegative, then the total allocation to any coalition exceeds the value of the coalition itself and the allocation vector lies in the core.* Also, notice that there are only $q-1$ surplus variables because coalition N has no surplus ($\sum_{i \in N} \tilde{u}_i - \eta_q = 0$) due to the efficiency condition of the core.

Table 1 depicts the mapping from coalitions into vertices (top) and the corresponding dynamical system of a 3-player coalitional game.

2. Model and problem statement

We consider a large number of replicas of a single coalitional game. The game is an n-player robust dynamical TU game $< N, \eta(t) >$, where $\eta(t)$ is the characteristic function representing the values of different coalitions. The characteristic function is modeled as a diffusion process with drift, and its evolution is described by the stochastic differential equation:

$$\begin{cases} d\eta(t) = w(t)dt - \sigma d\mathcal{B}(t), & \text{in } \mathbb{R}^q, \\ \eta(0) = \eta_0, \end{cases} \tag{3}$$

where $q = 2^n - 1$ is the number of coalitions, $w(t)$ is the vector of variations of coalitions' values, σ is the diffusion parameter, $\mathcal{B}(t)$ is a vector of independent Brownian motions, and η_0 is the vector of initial coalitions' values.

Denote by $x(t) \in \mathbb{R}^q$ the coalition excess; then, we can describe the time evolution of $x(t)$ through the following stochastic differential equation:

$$\begin{cases} dx(t) = (Bu(t) - w(t))dt + \sigma d\mathcal{B}(t), \\ x(0) = x_0, \end{cases} \tag{4}$$

where matrix B is as in (2) and $w(t)$ plays the role of a disturbance. In essence, every component of vector $Bu(t)$ is the total amount given to the members of a coalition at time t, and from this amount the value of the coalition itself, $w(t)$, is subtracted. Then, a positive $x(t)$ means positive cumulative excess.

We assume that controls and disturbances do not have any hard bounds, namely, controls and disturbances are in the sets $U := \mathbb{R}^n$ and $W := \mathbb{R}^q$, respectively.

With the above preamble in mind, and given the infinite copies of the same game, we can derive a probability density function $m : \mathbb{R}^q \times [0, +\infty[\to [0, +\infty[$, $(x, t) \mapsto m(x, t)$, for which $\int_{\mathbb{R}^q} m(x, t)dx = 1$ for every time t. We also denote the mean distribution at time t by $\bar{m}(t) := \int_{\mathbb{R}^q} xm(x, t)dx$.

In the spirit of *inequity aversion*, the designer of each game follows a so-called crowd-seeking law in that it readjusts the allocations by targeting the average distribution of the other games.

This is captured by considering a running cost $g : \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^n \times \mathbb{R}^q \to [0, +\infty[$, $(x, \bar{m}, u, w) \mapsto g(x, \bar{m}, u, w)$ of the quadratic form:

$$g(x, \bar{m}, u, w) = \tfrac{1}{2}\Big[(\bar{m} - x)^T Q (\bar{m} - x) + u^T(t)Ru(t) - w^T(t)\Gamma w(t)\Big], \tag{5}$$

where $Q, R, \Gamma > 0$, that is positive definite.

We also take as terminal cost the function $\Psi : \mathbb{R}^q \times \mathbb{R}^q \to [0, +\infty[$, $(x, \bar{m}) \mapsto \Psi(x, \bar{m})$ of the form

$$\Psi(x, \bar{m}) = \frac{1}{2}(\bar{m} - x)^T S(\bar{m} - x), \tag{6}$$

where $S > 0$. We are then ready to formalize the problem at hand as follows.

Problem 1. *Find the closed-loop optimal control and worst-case disturbance for the problem:*

$$\left\{ \begin{array}{l} \inf_{u(\cdot)\in\mathcal{U}} \sup_{w(\cdot)\in\mathcal{W}} \Big\{ J(x_0, u(\cdot), w(\cdot), m(\cdot)) \\ = \mathbb{E}\Big[\int_0^T g(x, \bar{m}, u, w)dt + \Psi(x(T), \bar{m}(T))\Big]\Big\}, \\ dx(t) = (Bu(t) - w(t))dt - \sigma d\mathcal{B}(t), \end{array} \right. \tag{7}$$

where $\mathcal{U}$ and $\mathcal{W}$ are the sets of all measurable functions $u(\cdot)$ and $w(\cdot)$ from $[0, +\infty[$ to U and W, respectively, and $m(\cdot)$ as a time-dependent function is the evolution of the distribution under the optimal control and the worst-case disturbance.

Note that the problem formulation includes the cases of both constrained and unconstrained controls and disturbances.

Problem 1 models the strategic behavior of each single designer. Given a forecast on the population behavior, which enters the problem through the distribution $m(\cdot)$, each designer minimizes the inequity, represented by the cost functional. In doing this, the designer assumes that an adversarial disturbance provides opposition to this. Evidently, problem (1) describes the system from a microscopic standpoint. The main question is to what extent one can reconstruct the same population behavior that is given as input to the cost functional in (7), assuming that all designers act rationally. This is what we call the macroscopic model. In other words, for model (7) to be consistent, we need to develop a more sophisticated model that integrates the microscopic and macroscopic descriptions in a consistent way. We develop such a model in the next section, and refer to it as the mean-field game representation of the model at hand.

3. The mean-field game

Consistency means that the microscopic and macroscopic models are tied together and are compatible. For the microscopic part, we know from optimal control theory, that the solution to (7) in terms of $u(\cdot)$ and $w(\cdot)$ can be expressed in terms of the *value function*, i.e., the minimum value of the cost functional at each time t and state x. In our case, the value function also depends on the distribution, which makes the problem different from a classical optimal control problem. A classical approach based on dynamic programming, derives the value function as solution of the well-known *Hamilton-Jacobi-Isaacs (HJI)* partial differential equation (PDE), which has to be solved for every state and time. For the macroscopic part, statistical mechanics tells us how the distribution of particles evolves, once they are immersed in a vector field. Viewing each designer as a particle, and the excesses of each TU game as its state, the optimal $u(\cdot)$ and $w(\cdot)$ represent the vector field, and thus we can use the *Fokker-Planck-Kolmogorov (FPK)* equation for a simulated forecast of the evolution distribution. We next study how to arrive at a consistent mean-field game for the problem at hand.

Let us denote by $v(x,t)$ the (upper) value of the robust optimization problem under a worst-case disturbance starting from time t at state x (which in this case also turns out to be the lower value, and hence the *value*, since Isaacs condition [3] holds–see below). Problem 1 results in the following mean-field game system for the unknown functions $v(x,t)$, and $m(x,t)$:

$$\left\{\begin{array}{l} \partial_t v(x,t) + \inf_{u\in U}\sup_{w\in W}\Big\{(Bu-w)^T\partial_x v(x,t) \\ +g(x,\bar{m},u,w)\Big\} + \frac{\sigma^2}{2}Tr\Big(\partial^2_{xx}v(x,t)\Big) = 0 \text{ in } \mathbb{R}^q\times[0,T[, \\ v(x,T)=\Psi(x,\bar{m}(T)) \text{ in } \mathbb{R}^q, \\ \\ \partial_t m(x,t) + div(m(x,t)\cdot(Bu-w)) \\ \qquad -\frac{\sigma^2}{2}Tr(\partial^2_{xx}m(x,t)) = 0, \text{ in } \mathbb{R}^q\times[0,T[, \\ m(0)=m_0, \\ \frac{d}{dt}\bar{m}_t = B\bar{u}^*_t - \bar{w}^*_t, \text{ in } [0,T[, \end{array}\right. \tag{8}$$

where $u^*(t,x)$ and $w^*(t,x)$ are the optimal time-varying state-feedback controls and disturbances, respectively, obtained as

$$\left\{\begin{array}{l} u^*(t,x) \in \arg\min_{u\in U}\{(Bu-w^*)\partial_x v(x,t) + g(x,\bar{m},u,w^*)\}, \\ \\ w^*(t,x) \in \arg\max_{w\in W}\{(Bu-w)\partial_x v(x,t) + g(x,\bar{m},u^*,w)\}. \end{array}\right. \tag{9}$$

Note that the minimization and maximization problems above are completely decoupled, and hence in (8) the inf sup is the same as sup inf (that is, Isaacs condition holds [3]). Further, we have replaced inf and sup in (9) with min and max, respectively, since g is quadratic in u and w.

The first equation in (8) is the *Hamilton-Jacobi-Isaacs (HJI)* equation with variable $v(x,t)$. Given the boundary condition on final state (second equation in (8)), and assuming a given population behavior captured by $m(\cdot)$, the HJI equation is solved backwards and returns the value function and best-response behavior of the individuals (first equation in (9)) as well as the worst adversarial response (second equation in (9)). The HJI equation is coupled with a second PDE, known as the *Fokker-Planck-Kolmogorov (FPK)* equation (third equation in (8)), defined on the variable $m(\cdot)$. Given the boundary condition on initial distribution $m(0)=m_0$

(fourth equation in (8)), and assuming a given individual behavior described by u^*, the FPK equation is solved forward and returns the population behavior time evolution $m(t)$. The last equation in (8) is obtained by averaging the left and right hand sides of the dynamics (4). Any solution of the above system of equations along with (9) is referred to as *worst-disturbance feedback mean-field equilibrium.*

Remark 1. *(On the existence of solutions) Analyzing the existence of solutions for the mean-field system (8) is a challenging task. However, under some restrictive sufficient conditions the existence of classical solutions can be established using a fixed-point theorem argument as in [21]. Indeed, let us assume that the initial measure m_0 is absolutely continuous with a continuous density function with finite second moment. In this case, we note that the running cost is convex in u. With these conditions, the existence of solution is established in Theorem 2.6 in [21]. Existence and uniqueness of a fixed-point solution are also discussed in [19, Theorem 4.3]*

Remark 2. *(On connections with the case with finite players) As the cost is Lipschitz continuous on m given that the control is bounded, the solution to the asymptotic case with infinite number of players relates to the case with a finite number of players as established in [19, 21]. In particular the classical bound of $\frac{1}{\sqrt{N}}$ holds true where N is the number of games. We refer the reader to the ϵ-Nash Equilibrium Theorem [19, Theorem 5.6] for more details.*

The mean-field game (8) can be reformulated as follows. Let us express the optimal $u(\cdot)$ and $w(\cdot)$ explicitly as functions of the value function and its derivatives. To do this, we have to solve (9) and substitute the expressions for $u(\cdot)$ and $w(\cdot)$ in (8). This procedure is the core of this section, and culminates in the mean-field game captured in Theorem 1.

Let the Hamiltonian (without disturbance w) be given by

$$H(x,p,\bar{m}) = \inf_u \left\{ \tilde{g}(x,\bar{m},u) + p^T B u \right\},$$

where p is the co-state and

$$\tilde{g}(x,\bar{m},u) = \frac{1}{2}\Big[(\bar{m}-x)^T Q\, (\bar{m}-x) + u^T R u \Big].$$

The robust Hamiltonian is then

$$\tilde{H}(x,p,\bar{m}) = H(x,p,\bar{m}) + \sup_w \left\{ -p^T w - \frac{1}{2} w^T \Gamma w \right\}.$$

The unique maximizing w is

$$w^* = -\Gamma^{-1} p.$$

Using the Hamiltonian and the expression for w^* in the mean-field system (8), and noting that $p = \partial_x v(x,t)$, we have

$$\left\{ \begin{array}{l} \partial_t v(x,t) + H(x,p,\bar{m}) + \dfrac{1}{2}(\partial_x v(x,t)^T \Gamma^{-1} \partial_x v(x,t)) \\ +\frac{\sigma^2}{2} Tr(\partial^2_{xx} v(x,t)) = 0, \text{ in } \mathbb{R}^q \times [0,T[, \\ \\ v(x,T) = \Psi(x,\bar{m}(T)) \text{ in } \mathbb{R}^q, \\ \\ \partial_t m(x,t) + div\Big(m(x,t)\partial_p H(x,p,\bar{m})\Big) + div\Big(m(x,t)\Gamma^{-1}\partial_x v(x,t)\Big) \\ -\frac{\sigma^2}{2} Tr(\partial^2_{xx} m(x,t)) = 0, \text{ in } \mathbb{R}^q \times [0,T[, \\ \\ m(x,0) = m_0(x) \text{ in } \mathbb{R}^n. \end{array} \right. \tag{10}$$

We are now in a position to specialize the results obtained above to the case under study.

Theorem 1. *The mean-field game is described by*

$$
\begin{cases}
\partial_t v(x,t) + \frac{1}{2}\partial_x v(x,t)^T\Big(- BR^{-1}B^T + \Gamma^{-1}\Big) \cdot \partial_x v(x,t) + \frac{1}{2}(\bar{m}(t) - x)^T Q(\bar{m}(t) - x) \\
+\frac{1}{2}\sigma^2 Tr(\partial^2_{xx} v(x,t)) = 0, \ \ in\ \mathbb{R}^q \times [0,T[, \\
\\
v(x,T) = \Psi(x, \bar{m}(T)) \ \ in\ \mathbb{R}^q, \\
\\
\partial_t m(x,t) + div\Big(m(x,t)(-BR^{-1}B^T + \Gamma^{-1}) \cdot \partial_x v(x,t)\Big) - \frac{1}{2}\sigma^2 Tr(\partial^2_{xx} m(x,t)) = 0, \\
\ in\ \mathbb{R}^n \times [0,T[, \\
\\
m(x,0) = m_0(x) \ \ in\ \mathbb{R}^n.
\end{cases}
\tag{11}
$$

Furthermore, the optimal control and worst-case disturbance are

$$
\begin{cases}
u^*(x,t) = -R^{-1}B^T \partial_x v(x,t), \\
\\
w^*(x,t) = -\Gamma^{-1}\partial_x v(x,t).
\end{cases}
\tag{12}
$$

Proof. w^* in (12) was already obtained earlier. u^* follows from a straightforward minimization of the Hamiltonian. We next prove (11). First notice that the second and last relations are the boundary conditions, and do not require any further justification.

To prove the first equation, which is a PDE corresponding to the HJI, let us replace u^* appearing in the Hamiltonian by its expression (12):

$$
\begin{aligned}
H(x, \partial_x v(x,t), \bar{m}) &= \frac{1}{2}\Big[(\bar{m} - x)^T Q(\bar{m} - x) + u^{*T} R u^*\Big] + \partial_x v(x,t) B u^* \\
= \frac{1}{2}(\bar{m} - x)^T Q(\bar{m} - x) &+ \frac{1}{2}\partial_x v(x,t)^T BR^{-1}B^T \partial_x v(x,t) - \partial_x v(x,t)^T BR^{-1}B^T \partial_x v(x,t) \\
&= \frac{1}{2}(\bar{m} - x)^T Q(\bar{m} - x) - \frac{1}{2}\partial_x v(x,t)^T BR^{-1}B^T \partial_x v(x,t).
\end{aligned}
$$

Using the above expression of the Hamiltonian in the HJI equation in (10), we obtain the HJI equation in (11).

To prove the third equation, which is a PDE representing the FPK equation, we simply substitute (12) in the FPK in (10), and this concludes the proof. □

Equation (12) returns the optimal $u(\cdot)$ and $w(\cdot)$ as explicit functions of the gradient of the value function. Independently of how complicated the computation of such a gradient may be, the expressions prove to be scalable as no direct dependence on the size of the population appears. Scalability is one of the advantages of the provided model as it disconnects the structure of the policies from the dimension of the problem.

4. Mean-field response

By mean-field response we mean the solution to the HJI in (11) for a fixed distribution. By "fixed" we mean that the distribution is known but not constant. In other words, we solve the first PDE in (11), assuming that we are given a solution for the second PDE (the FPK equation). Computing the mean-field response is relevant in that it produces insights on the structure of the mean-field equilibrium allocation policies. The main goal of this section is to show that such mean-field response allocation policies are linear and time-varying functions of the local state. Furthermore, solving the HJI equation turns into solving a set of ordinary differential equations in the parameters of the value function. This simplifies dramatically the tractability of the mean-field equilibrium policies.

To see this, let us isolate the HJI part of (11) for fixed m, which leads to

$$\begin{cases} \partial_t v(x,t) + \frac{1}{2}\partial_x v(x,t)^T\Big(-BR^{-1}B^T + \Gamma^{-1}\Big)\cdot\partial_x v(x,t) + \frac{1}{2}(\bar{m}(t)-x)^T Q(\bar{m}(t)-x) \\ +\frac{1}{2}\sigma^2 Tr(\partial^2_{xx} v(x,t)) = 0, \text{ in } \mathbb{R}^q \times [0,T[, \\ \\ v(x,T) = \Psi(x,\bar{m}(T)) \text{ in } \mathbb{R}^q. \end{cases} \tag{13}$$

Let us consider the following structure for the value function

$$v(x,t) = \frac{1}{2}x^T\phi(t)x + h(t)^T x + \chi(t),$$

where $\phi(t)$ is a square matrix, $h(t)$, $\chi(t)$ are vectors, and all depend on time and have compatible dimensions. Then (13) can be rewritten as

$$\begin{cases} \frac{1}{2}x^T\dot{\phi}(t)x + \dot{h}(t)^T x + \dot{\chi}(t) + \frac{1}{2}[\phi(t)x + h(t)]^T\Big(-BR^{-1}B^T + \Gamma^{-1}\Big) \\ \cdot[\phi(t)x + h(t)] + \frac{1}{2}(\bar{m}(t)-x)^T Q(\bar{m}(t)-x) \\ +\frac{1}{2}\sigma^2\phi(t) = 0 \text{ in } \mathbb{R}^q \times [0,T[, \\ \\ \phi(T) = S,\ h(T) = -S\bar{m}(T),\ \chi(T) = \frac{1}{2}S\bar{m}(T). \end{cases} \tag{14}$$

The advantage of doing this is that the above system represents an identity in x, and therefore it reduces to the following ordinary differential equations assuming that they admit unique continuously differentiable solutions.

$$\begin{cases} \dot{\phi}(t) + \phi(t)^T\Big(-BR^{-1}B^T + \Gamma^{-1}\Big)\phi(t) + Q = 0 \text{ in } [0,T[,\ \phi(T) = S, \\ \\ \dot{h}(t) + h(t)^T\Big(-BR^{-1}B^T + \Gamma^{-1}\Big)\phi(t) - \bar{m}(t)^T Q = 0 \text{ in } [0,T[,\ h(T) = -S\bar{m}(T), \\ \\ \dot{\chi}(t) + \frac{1}{2}h(t)^T\Big(-BR^{-1}B^T + \Gamma^{-1}\Big)h(t) + \frac{1}{2}\bar{m}(t)^T Q\bar{m}(t) + \frac{1}{2}\sigma^2\phi(t) = 0 \\ \text{ in } [0,T[,\ \chi(T) = \frac{1}{2}S\bar{m}(T). \end{cases} \tag{15}$$

For the optimal control and the worst-case disturbance we then have

$$\begin{cases} u^*(x,t) = -R^{-1}B^T(\phi(t)x + h(t)) \\ w^*(x,t) = -\Gamma^{-1}(\phi(t)x + h(t)). \end{cases} \tag{16}$$

The main insight we get from (15)-(16) is that the expressions for the mean-field equilibrium policies may not be very different from the one in (16), obtained for the mean-field responses. In the next section we show that the mean-field equilibrium policies for $u(\cdot)$ and $w(\cdot)$ are still linear in the extended state, which includes both the local and the common states.

It is well known that for (15) to be well posed, it must hold that $-BR^{-1}B^T + \Gamma^{-1} < 0$ (see e.g. Appendix of [4] on the theory of conjugate points). Taking $R = I$ and $\Gamma = \gamma^2 I$ for a given parameter γ, the above condition becomes

$$-BB^T + \frac{1}{\gamma^2} I < 0, \tag{17}$$

which establishes a relation between the *smallest eigenvalue* (this is related to the *connectivity* of the network) of the matrix BB^T and the maximal eigenvalue of the matrix $\frac{1}{\gamma^2}I$, namely $\frac{1}{\gamma^2}$. Condition (17) is particularly interesting as it provides a lower bound on γ^2, namely, a lower bound on the coefficient in the cost functional that weights the disturbance input representing the coalitions' values. A larger value for γ^2 leads to smaller coalitions' values $w(t)$.

5. Mean-field equilibrium strategies

This section describes a simple heuristic approach toward solving the set of equations (8), based on state space augmentation [13]. The augmented state space includes the mean distribution, and thus the augmented state variables evolve according to the equations

$$\begin{bmatrix} dx(t) \\ d\bar{m}(t) \end{bmatrix} = \Big(B \begin{bmatrix} u^*(x,t) \\ \bar{u}^*(t) \end{bmatrix} - \begin{bmatrix} w^*(x,t) \\ \bar{w}^*(t) \end{bmatrix} \Big) dt + \begin{bmatrix} \sigma d\mathcal{B}_t \\ 0 \end{bmatrix}. \tag{18}$$

For this system, we introduce an assumption on the rate of convergence of the state $\bar{m}(t)$.

Assumption 1. *There exists a scalar $\theta > 0$ such that*

$$\frac{d}{dt}\bar{m}(t) = B\bar{u}^*(t) - \bar{w}^*(t) \geq -\theta \bar{m}_t, \text{ for all } t \in [0,T],$$

where the inequality is to be interpreted component-wise.

The above assumption implies that there exists a variable $\tilde{m}(t)$ which approximates the average mean value from below and evolves according to

$$\begin{cases} \frac{d}{dt}\tilde{m}(t) = -\theta \tilde{m}(t), & \text{for all } t \in [0,T], \\ \tilde{m}_0 = \bar{m}_0. \end{cases} \tag{19}$$

As a result, each single game is described by the dynamical closed-loop system depicted in Fig. 4. The block at the top represents the state dynamics (4), it receives as input the control $u(t)$, and the disturbances $w(t)$ and $\mathcal{B}(t)$. The block on the right includes the internal model for $\tilde{m}$ as in (19) from which we obtain the error $e(t) := \tilde{m}(t) - x(t)$. The block at the bottom describes the closed-loop state feedback control obtained solving the linear quadratic tracking problem. The resulting linear quadratic problem is detailed next.
By substituting the current mean value $\bar{m}(t)$ by its estimate $\tilde{m}(t)$, the augmented problem becomes

$$\inf_{u(\cdot)\in\mathcal{U}} \sup_{w(\cdot)\in\mathcal{W}} \int_0^T \frac{1}{2}\Big[(\tilde{m}(t) - x(t))^T Q (\tilde{m}(t) - x(t)) + u^T(t) R u(t) - w^T(t)\Gamma w(t) \Big] dt$$

$$\begin{bmatrix} dx(t) \\ d\tilde{m}(t) \end{bmatrix} = \Big(\begin{bmatrix} 0 & 0 \\ 0 & -\theta I \end{bmatrix} \begin{bmatrix} x(t) \\ \tilde{m}(t) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(t) - \begin{bmatrix} I \\ 0 \end{bmatrix} w(t) \Big) dt + \begin{bmatrix} \sigma d\mathcal{B}(t) \\ 0 \end{bmatrix}.$$

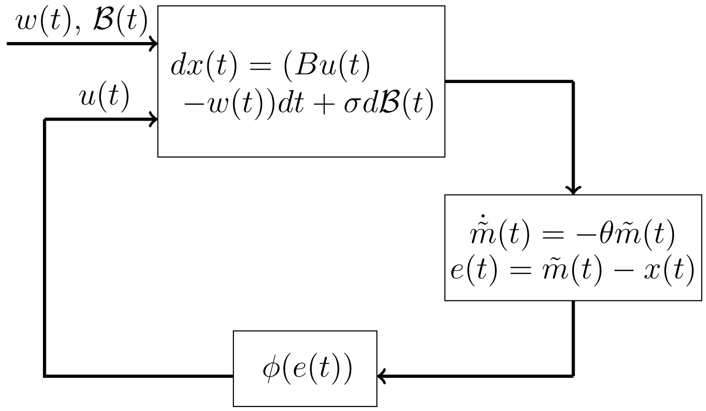

Figure 4: The dynamical closed-loop system of each single game

Reformulating the problem in terms of the augmented state

$$X(t) = \left[\begin{array}{c} x(t) \\ \tilde{m}(t) \end{array} \right],$$

we have the linear quadratic problem:

$$\begin{array}{l} \inf_{u(\cdot)\in\mathcal{U}} \sup_{w(\cdot)\in\mathcal{W}} \displaystyle\int_0^T \Big[\frac{1}{2}(X(t)^T\tilde{Q}X(t) + u^T(t)Ru(t) - w^T(t)\Gamma w(t))\Big] dt + \tilde{\Psi}(X(T)) \\ dX(t) = \Big(FX(t) + Gu(t) + Hw(t)\Big)dt + Ld\mathcal{B}_t, \end{array}$$

where

$$\tilde{Q} = \left[\begin{array}{cc} Q & -Q \\ -Q & Q \end{array} \right], \quad L = \left[\begin{array}{c} \sigma I \\ 0 \end{array} \right],$$

$$F = \left[\begin{array}{cc} 0 & 0 \\ 0 & -\theta I \end{array} \right], \quad G = \left[\begin{array}{c} B \\ 0 \end{array} \right], \quad H = \left[\begin{array}{c} -I \\ 0 \end{array} \right],$$

$R > 0$, $\Gamma > 0$, and $\tilde{\Psi}(X) := \Psi(x, \tilde{m})$.
The idea is therefore to consider a new value function $\mathcal{V}_t(x, \tilde{m})$ (in compact form $\mathcal{V}_t(X)$) in the augmented state space, which satisfies

$$\left\{ \begin{array}{l} \partial_t \mathcal{V}_t(X) + H(X, \partial_X \mathcal{V}_t(X)) + \frac{1}{2}\sigma^2 \mathrm{Tr}\partial^2_{xx}\mathcal{V}_t(X) = 0, \text{ in } \mathbb{R}^{2q} \times [0, T[, \\ \mathcal{V}_T(X) = \tilde{\Psi}(X) \quad \text{in } \mathbb{R}^{2q}, \end{array} \right.$$

where $H(X, \partial_X \mathcal{V}_t(X))$ is the robust Hamiltonian [9]:

$$\begin{array}{l} H(X, \partial_X \mathcal{V}_t(X)) = \frac{1}{2}\, X^T \tilde{Q} X + \partial_X \mathcal{V}_t(X) F X \\ \qquad -\frac{1}{2}\, \partial_X \mathcal{V}_t(X)[GR^{-1}G^T - H\Gamma^{-1}H^T](\partial_X \mathcal{V}_t(X))^T . \end{array}$$

This PDE admits the unique solution given by

$$\mathcal{V}_t(X) = \frac{1}{2}\, X(t)^T \underbrace{\left[\begin{array}{cc} P_{11}(t) & P_{12}(t) \\ P_{12}(t)^T & P_{22}(t) \end{array} \right]}_{P(t)} X(t) + \frac{1}{2}\, p(t),$$

where the symmetric matrix $P(t)$ satisfies (is the unique nonnegative-definite solution of) the generalized (game) Riccati differential equation

$$\begin{aligned} &\dot{P}(t) + P(t)F + F^T P(t) - P(t)(GR^{-1}G^T - H\Gamma^{-1}H^T)P(t) + \tilde{Q} = 0\,, \\ &P(T) = \begin{bmatrix} S & -S \\ -S & S \end{bmatrix}, \end{aligned} \tag{20}$$

and $p(\cdot)$ is solved from

$$\dot{p}(t) + \sigma^2\, \mathrm{Tr} P(t) = 0\,, \quad p(T) = 0\,.$$

Then, the corresponding optimal control is given by

$$\tilde{u}(t) = -R^{-1}G^T P(t)X(t) = -R^{-1}B^T(P_{11}(t)x(t) + P_{12}(t)\bar{m}(t)), \tag{21}$$

and the worst-case disturbance is given by

$$\tilde{w}(t) = \Gamma^{-1}H^T P X(t) = -\Gamma^{-1}(P_{11}(t)x(t) + P_{12}(t)\bar{m}(t)). \tag{22}$$

The Riccati equation (20) is relevant as it returns the parameters identifying the structure of the value function, and as such it replaces the HJI equation. Such parameters are then used in the mean-field equilibrium policies in (21)-(22). From (21)-(22), it is evident that the mean-field equilibrium allocation policies have a linear structure in the extended state.

6. Stability

This section analyzes stability of the microscopic dynamics when the designer implements the mean-field equilibrium allocation policy derived in the previous section. The main question concerns the possibility that the excesses diverge with time from the average value, which would ruin the inequity aversion requirement. We show that this is not the case under the provided allocation policies. More formally, we prove that the excesses are mean square bounded. This means that the second moment of the excesses, viewed as a stochastic process, is bounded. To put it differently, the expected value of the deviation of the excesses in each game from the average excesses is bounded. Successively, we also investigate how the provided allocation policies affect the macroscopic dynamics. In particular we prove that the average excesses converge exponentially to zero in the absence of stochastic disturbances.

6.1. Asymptotic stability and mean-field equilibrium

Using the optimal control and worst-case disturbance (21)-(22) in the SDE (4) we obtain

$$\begin{aligned} dx(t) \;\; = \;\; & (-BR^{-1}B^T + \Gamma^{-1})P_{11}(t)x(t)dt + (-BR^{-1}B^T + \Gamma^{-1})P_{12}(t)\bar{m}(t)dt \\ & +\sigma d\mathcal{B}(t),\; t \in (0,T],\; x_0 \in \mathbb{R}^q. \end{aligned} \tag{23}$$

Remark 3. *If $P_{11}(t) \approx -P_{12}(t)$, then (23) becomes*

$$dx(t) = (BR^{-1}B^T - \Gamma^{-1})P_{11}(t)(\bar{m}(t) - x(t))dt + \sigma d\mathcal{B}(t).$$

This means that the state evolves as a function of the error $(\bar{m}(t) - x(t))$. We will use this fact in the simulation studies.

Assumption 2. *There exists a matrix K, which is Hurwitz, such that the following inequality holds componentwise*

$$Kx(t) \geq (-BR^{-1}B^T + \Gamma^{-1})P_{11}(t)x(t) + (-BR^{-1}B^T + \Gamma^{-1})P_{12}(t)\bar{m}(t) \tag{24}$$

Under the above assumption, the SDE is linear and time-varying, and the corresponding stochastic process can be analyzed in the context of stochastic stability theory [22].

Definition 2 (cf. Definition (11.3.1) in [1]). *(stability in pth moment) The equilibrium solution of a stochastic process $\xi(t)$ is said to be* stable in the pth moment*, $p > 0$, if given $\varepsilon > 0$, there exists a $\delta(\varepsilon, t_0) > 0$ so that $\|x(0)\| \leq \delta$ guarantees that*

$$\mathbb{E}\{\sup_{t\geq t_0} \|x(t)\|^p < \varepsilon\}.$$

When $p = 1$ or 2 we speak of *stability in the mean* or *mean-square*, respectively.

Theorem 2. *The stochastic process (23) describing the time evolution of the excesses is mean-square stable.*

Proof. Let us consider as Lyapunov function the positive-definite quadratic function $V(x) = x^T Sx$, where S is the unique positive-definite solution of the Lyapunov equation: $K^T S + SK = -I$, which exists since K is Hurwitz.

Let the infinitesimal generator be

$$\mathcal{L} = \frac{1}{2}\sigma^2 \sum_{i=1}^{q} \frac{\partial^2}{\partial x_i \partial x_i} + (Kx(t))^T \frac{d}{dx}. \tag{25}$$

We recall that for a Brownian motion we have $\mathbb{E}d\mathcal{B}_t = 0$ and $\mathbb{E}d\mathcal{B}_t^T d\mathcal{B}_t = q\, dt$ and dropping the second-order terms (in dt^2) one obtains (25).

Then the stochastic derivative of $V(x)$ can be obtained by applying the infinitesimal generator to $V(x)$, which yields

$$\begin{aligned}\mathcal{L}V(x(t)) &= \lim_{dt\to 0} \frac{\mathbb{E}V(x(t+dt)) - V(x(t))}{dt} \\ &= q\sigma^2 + x(t)^T K^T Sx(t) + x(t)^T SKx(t) = q\sigma^2 - x(t)^T x(t).\end{aligned}$$

Then we have that $\mathcal{L}V(x) \leq 0$ on $Q_\epsilon := \{x : V(x) \geq \epsilon\}$ for some $\epsilon > 0$. Hence the 2nd moment is bounded and the process is mean square stable. □

The interpretation of the above result is that the variance of the excesses in each game is bounded.

6.2. Mean-field equilibrium

We can approximate the mean-field equilibrium, which is captured by the evolution of $\bar{m}_t$ over the horizon $(0, T]$, as

$$\begin{aligned}\tfrac{d}{dt}\bar{m}_t &= ((-BR^{-1}B^T + \Gamma^{-1})P_{11}(t) \textstyle\int x(t)dm + (-BR^{-1}B^T + \Gamma^{-1})P_{12}(t)\bar{m}(t) \\ &= (-BR^{-1}B^T + \Gamma^{-1})(P_{11}(t) + P_{12}(t))\bar{m}_t \quad t \in (0, T],\ \bar{m}_0 \in \mathbb{R}^q.\end{aligned}$$

Actually, we can derive an expression based on the matrix exponential $e^{\rho t}$ describing the evolution of the mean distribution which represents a bound, namely

$$\begin{cases} \bar{m}_t = \bar{m}_0 e^{Zt} \\ Z = (-BR^{-1}B^T + \Gamma^{-1})(P_{11}(t) + P_{12}(t)). \end{cases}$$

The equation above corresponds to saying that the mean distribution converges exponentially to zero in the absence of the stochastic disturbances (the Brownian motion).

7. Multi-population and local interactions

The results obtained in the previous sections can be extended to the case where we have heterogeneous populations of TU games. By this we mean multiple populations of TU games. Here, we have a structured environment, namely, interactions between populations occur locally, and we use a graph topology to describe them. We recall that each designer is associated with a corresponding TU game and each population is characterized by a different initial distribution of the excesses.

Consider p heterogeneous populations with each population comprised of infinite homogeneous designers; each designer is identified by an index $k \in \{1, \dots, p\}$ which identifies the population he belongs to. Each designer is also characterized by a state $x(t) \in \mathbb{R}^q$ whose dynamics follow (4). As indicated earlier in the paper, the state $x(t)$ represents the excesses of the corresponding TU game at time $t \in [0, T]$.

For every population $k \in \{1, \dots, p\}$, the distribution of the excesses is given by a probability density function $m_k : \mathbb{R} \times [0, +\infty[\to \mathbb{R}$, $(x, t) \mapsto m_k(x, t)$. In other words, such a function represents the density of players of that population in state x at time t. By conservation of mass, $\int_{\mathbb{R}} m_k(x, t) dx = 1$ for every t. We call the *mean state* of population k at time t the value $\bar{m}_k(t) := \int_{\mathbb{R}} x m_k(x, t) dx$.

To model the interactions between populations, we let a graph $\mathcal{H}' = (V', E')$ be given where $V' = \{1, \dots, p\}$ is the set of vertices, one per each population, and $E' = V' \times V'$ is the set of edges. For the sake of simplicity, we henceforth assume that $\mathcal{H}' = (V', E')$ is a balanced graph (or undirected graph) although most results hold even for more general graphs, possibly time-varying. Denote the set of neighbors of k by $N(k) = \{j \in V' | \, (k, j) \in E'\}$.

The signal tracked by each designer in population k is given by the following local average

$$\rho_k = \frac{\sum_{j \in N(k)} \bar{m}_j(t)}{|N(k)|}, \tag{26}$$

where $|N(k)|$ denotes the cardinality of the set $N(k)$, namely the number of neighbors of k.

After replacing $\bar{m}$ with ρ_k in the running cost (5), we have

$$g(x, \rho_k, u, w) = \tfrac{1}{2}\Big[(\rho_k - x)^T Q \, (\rho_k - x) + u^T(t) R u(t) - w^T(t) \Gamma w(t) \Big], \tag{27}$$

where Q, R, Γ are positive definite.

Likewise, for the terminal cost we have

$$\Psi(x, \rho_k) = \frac{1}{2}(\rho_k - x)^T S (\rho_k - x), \tag{28}$$

where $S > 0$. The problem for the multi-population case is then formulated as follows.

Problem 2. *Find the closed-loop optimal control and worst-case disturbance for the problem:*

$$\left\{\begin{array}{l} \inf_{u(\cdot)\in\mathcal{U}} \sup_{w(\cdot)\in\mathcal{W}} \Big\{ J(x_0, u(\cdot), w(\cdot), \rho_k(\cdot)) \\ = \mathbb{E}\Big[\int_0^T g(x, \rho_k, u, w)dt + \Psi(x(T), \rho_k(T))\Big]\Big\}, \\ dx(t) = (Bu(t) - w(t))dt - \sigma d\mathcal{B}(t), \end{array}\right. \tag{29}$$

where $\mathcal{U}$ *and* $\mathcal{W}$ *are the sets of all measurable functions* $u(\cdot)$ *and* $w(\cdot)$ *from* $[0, +\infty[$ *to* U *and* W*, respectively, and* $m(\cdot)$ *as a time-dependent function is the evolution of the distribution under the optimal control and the worst-case disturbance.*

In the following we derive the counterpart of Theorem 1 for the multi-population case. For every population $k \in \{1, 2, \ldots, p\}$, let $v_k(x, t)$ be the (upper) value of the robust optimization problem under worst-case disturbance starting at time t and at state x.

Then we have the following result.

Theorem 3. *The mean-field game is described by*

$$\left\{\begin{array}{l} \partial_t v_k(x,t) + \frac{1}{2}\partial_x v_k(x,t)^T\Big(- BR^{-1}B^T + \Gamma^{-1}\Big) \\ \cdot\partial_x v_k(x,t) + \frac{1}{2}(\rho_k(t) - x)^T Q(\rho_k(t) - x) \\ +\frac{1}{2}\sigma^2 Tr(\partial^2_{xx} v_k(x,t)) = 0, \ \textit{in } \mathbb{R}^q \times [0, T[, \\ \\ v_k(x, T) = \Psi(x, \rho_k(T)) \ \textit{in } \mathbb{R}^q, \\ \\ \partial_t m_k(x,t) + div\Big(m_k(x,t)(-BR^{-1}B^T + \Gamma^{-1}) \\ \cdot\partial_x v_k(x,t)\Big) - \frac{1}{2}\sigma^2 Tr(\partial^2_{xx} m_k(x,t)) = 0, \ \textit{in } \mathbb{R}^q \times [0, T[, \\ \\ m_k(x, 0) = m_{k,0}(x) \ \textit{in } \mathbb{R}^q, \\ \bar{m}_k(t) := \int_{\mathbb{R}^q} x m_k(x,t)dx, \\ \rho_k = \frac{\sum_{j\in N(k)} \bar{m}_j(t)}{|N(k)|}. \end{array}\right. \tag{30}$$

Furthermore, the optimal control and the worst-case disturbance are

$$\left\{\begin{array}{l} u_k^*(x,t) = -R^{-1}B^T\partial_x v_k(x,t), \\ w_k^*(x,t) = -\Gamma^{-1}\partial_x v_k(x,t). \end{array}\right. \tag{31}$$

Proof. Condition (31) is derived as (12) in Theorem 1. We now prove (30). The set of equations (8) adapted to the multi-population case takes the form

$$\left\{\begin{array}{l} \partial_t v_k(x,t) + \inf_{u\in U}\sup_{w\in W}\Big\{(Bu - w)^T\partial_x v_k(x,t) \\ +g(x, \rho_k, u, w)\Big\} + \frac{\sigma^2}{2} Tr\Big(\partial^2_{xx} v_k(x,t)\Big) = 0 \text{ in } \mathbb{R}^q \times [0, T[, \\ v_k(x,T) = \Psi(x, \rho_k(T)) \text{ in } \mathbb{R}^q, \\ \\ \partial_t m_k(x,t) + div(m_k(x,t)\cdot(Bu - w)) \\ \qquad -\frac{\sigma^2}{2} Tr(\partial^2_{xx} m_k(x,t)) = 0, \text{ in } \mathbb{R}^q \times [0, T[, \\ m_k(0) = m_{k,0}. \end{array}\right. \tag{32}$$

Introducing the Hamiltonian and the expression for w^* in the mean-field system (32), we obtain

$$\begin{cases} \partial_t v_k(x,t) + H(x,\tilde{p},\rho_k) + \dfrac{1}{2}(\partial_x v_k(x,t)^T \Gamma^{-1} \partial_x v_k(x,t)) \\ +\frac{\sigma^2}{2} Tr(\partial^2_{xx} v_k(x,t)) = 0, \text{ in } \mathbb{R}^q \times [0,T[, \\ \\ v_k(x,T) = \Psi(x,\rho_k(T)) \text{ in } \mathbb{R}^q, \\ \\ \partial_t m_k(x,t) + div\Big(m_k(x,t)\partial_p H(x,\tilde{p},\rho_k)\Big) + div\Big(m_k(x,t)\Gamma^{-1}\partial_x v_k(x,t)\Big) \\ -\frac{\sigma^2}{2} Tr(\partial^2_{xx} m_k(x,t)) = 0, \text{ in } \mathbb{R}^q \times [0,T[, \\ \\ m_k(x,0) = m_{k,0}(x) \text{ in } \mathbb{R}^q \\ \bar{m}_k(t) := \int_{\mathbb{R}^q} x m_k(x,t) dx, \\ \rho_k = \frac{\sum_{j\in N(k)} \bar{m}_j(t)}{|N(k)|}. \end{cases} \tag{33}$$

To prove the first equation, which is a PDE corresponding to the HJI, let us replace u^* appearing in the Hamiltonian by its expression (31):

$$\begin{aligned} H(x,\partial_x v_k(x,t),\rho_k) &= \frac{1}{2}\Big[(\rho_k - x)^T Q(\rho_k - x) + u^{*T} R u^*\Big] + \partial_x v_k(x,t) B u^* \\ &= \frac{1}{2}(\rho_k - x)^T Q(\rho_k - x) - \frac{1}{2}\partial_x v_k(x,t)^T \cdot BR^{-1}B^T \partial_x v_k(x,t). \end{aligned}$$

Using the above expression of the Hamiltonian in the HJI equation in (33), we obtain the HJI equation in (30).

To prove the third equation, which is a PDE representing the FPK equation, we simply substitute (31) in the FPK in (33), and this concludes the proof. □

The next result provides mean-field equilibrium control strategies and worst-case disturbances for the multi-population case.

Theorem 4. *A mean-field equilibrium for (30) is given by the following set of equation. For all $k \in \{1,2,\dots,p\}$*

$$\begin{cases} v_k(x,t) = \frac{1}{2}x^T\phi(t)x + h(t)^T x + \chi(t), \\ \dot{\bar{m}}_k(t) = (-BR^{-1}B^T + \Gamma^{-1})(\phi(t)\bar{m}_k(t) + h(t)), \end{cases} \tag{34}$$

where

$$\begin{cases} \dot{\phi}(t) + \left(-\frac{1}{c_1} + \frac{1}{\gamma^2}\right)\phi(t)^2 + a = 0 \text{ in } [0,T[, \ \phi(T) = S, \\ \dot{h}(t) + \left(-\frac{1}{2c_1} + \frac{1}{2\gamma^2}\right) 2\phi(t)h(t) - a\rho_k(t) = 0 \text{ in } [0,T[, \\ h(T) = -S\rho_k(T), \\ \dot{\chi}(t) + \left(-\frac{1}{2c_1} + \frac{1}{2\gamma^2}\right) h(t)^2 + \frac{1}{2}a\rho_k(t)^2 + \frac{1}{2}\sigma^2\phi(t) = 0 \\ \text{in } [0,T[, \ \chi(T) = \frac{1}{2}S\rho_k^2(T). \end{cases} \tag{35}$$

The corresponding mean-field equilibrium control and disturbance are

$$\begin{cases} u^*(x,t) = -R^{-1}B^T(\phi(t)x + h(t)) \\ w^*(x,t) = -\Gamma^{-1}(\phi(t)x + h(t)). \end{cases} \tag{36}$$

Furthermore,

$$\tfrac{d}{dt}\bar{m}_k(t) = \Big(- BR^{-1}B^T + \Gamma^{-1}\Big)\phi \tfrac{1}{|N(k)|}(\textstyle\sum_{j\in N(k)}(\bar{m}_j(t) - \bar{m}_k(t))).$$

Proof. Isolating the HJI part of (30) for fixed ρ_k, we have

$$\left\{ \begin{array}{l} \partial_t v_k(x,t) + \frac{1}{2}\partial_x v_k(x,t)^T\Big(- BR^{-1}B^T + \Gamma^{-1}\Big) \\ \cdot\partial_x v_k(x,t) + \frac{1}{2}(\rho_k(t) - x)^T Q(\rho_k(t) - x) \\ +\frac{1}{2}\sigma^2 Tr(\partial^2_{xx} v_k(x,t)) = 0, \text{ in } \mathbb{R}^q \times [0,T[, \\ \\ v_k(x,T) = \Psi(x,\rho_k(T)) \text{ in } \mathbb{R}^q. \end{array} \right. \tag{37}$$

Consider the following structure for the value function:

$$v_k(x,t) = \frac{1}{2}x^T\phi(t)x + h(t)^T x + \chi(t),$$

so that (37) can be rewritten as

$$\left\{ \begin{array}{l} \frac{1}{2}x^T\dot{\phi}(t)x + \dot{h}(t)^T x + \dot{\chi}(t) \\ +\frac{1}{2}[\phi(t)x + h(t)]^T\Big(- BR^{-1}B^T + \Gamma^{-1}\Big)\cdot[\phi(t)x + h(t)] + \frac{1}{2}(\rho_k(t) - x)^T Q(\rho_k(t) - x) \\ +\frac{1}{2}\sigma^2\phi(t) = 0 \text{ in } \mathbb{R}^q\times[0,T[, \\ \\ \phi(T) = S,\ h(T) = -S\rho_k(T),\ \chi(T) = \frac{1}{2}S\rho_k(T). \end{array} \right. \tag{38}$$

Since this is an identity in x, it reduces to three equations:

$$\left\{ \begin{array}{l} \dot{\phi}(t) + \phi(t)^T\Big(- BR^{-1}B^T + \Gamma^{-1}\Big)\phi(t) + Q = 0 \text{ in } [0,T[,\ \phi(T) = S, \\ \dot{h}(t) + h(t)^T\Big(- BR^{-1}B^T + \Gamma^{-1}\Big)\phi(t) - \rho_k(t)^T Q = 0 \text{ in } [0,T[,\ h(T) = -S\rho_k(T), \\ \dot{\chi}(t) + \frac{1}{2}h(t)^T\Big(- BR^{-1}B^T + \Gamma^{-1}\Big)h(t) + \frac{1}{2}\rho_k(t)^T Q\rho_k(t) + \frac{1}{2}\sigma^2\phi(t) = 0 \\ \ \text{in } [0,T[,\ \chi(T) = \frac{1}{2}S\rho_k(T). \end{array} \right. \tag{39}$$

For the mean-field equilibrium control and worst-case disturbance, we then have

$$\left\{ \begin{array}{l} u^*(x,t) = -R^{-1}B^T(\phi(t)x + h(t)) \\ w^*(x,t) = -\Gamma^{-1}(\phi(t)x + h(t)). \end{array} \right. \tag{40}$$

By averaging the above expressions and substituting in $\frac{d}{dt}\bar{m}_k(t) = B\bar{u}_k(t) + \bar{w}_k(t)$, we obtain

$$\dot{\bar{m}}_k(t) = (-BR^{-1}B^T + \Gamma^{-1})(\phi(t)\bar{m}_k(t) + h(t))$$

as in (34). Take $h = -\phi\rho_k$. Substituting in (40), the mean-field equilibrium control and worst-case disturbance take the form

$$\left\{ \begin{array}{l} u^*(x,t) = -R^{-1}B^T\phi(\rho_k - x) \\ w^*(x,t) = -\Gamma^{-1}\phi(\rho_k - x). \end{array} \right.$$

Then, mean states of *neighbor* populations are related by the *local interaction* rule:

$$\frac{d}{dt}\bar{m}_k(t) = B\bar{u}_k(t) + \bar{w}_k(t) = \Big(-BR^{-1}B^T + \Gamma^{-1}\Big)\phi\big(\frac{\sum_{j\in N(k)}\bar{m}_j(t)}{|N(k)|} - \bar{m}_k(t)\big)$$
$$= \Big(-BR^{-1}B^T + \Gamma^{-1}\Big)\phi\frac{1}{|N(k)|}\big(\sum_{j\in N(k)}\cdot(\bar{m}_j(t) - \bar{m}_k(t))\big).$$

and this concludes the proof. □

Remark 4. *Dynamics (34) is a consensus dynamics and as such it guarantees synchronization as required by the design specifics.* □

8. Simulations

In this section, results of Theorems 1 and 3 are illustrated. The simulations show that the populations reach consensus on the excesses under different circumstances, i.e., with or without stochastic disturbances, with local or global interactions, and with first or second-order interaction dynamics in the case of local interactions.

8.1. Single-population example

We provide here simulations of a game with each group consisting of three players as illustrated in Fig. 3 and Table 1. The simulations are obtained using the algorithm displayed in Table 3. The time evolution of the approximate average state is used to simulate the evolution of one of the groups. Matrix $B_{\mathcal{H}} \in \{0,1\}^{7\times 3}$ takes the form

$$B_{\mathcal{H}}^T = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

step size	horizon length	σ	θ
0.1	500	0.01	0.01

Table 2: Single-population: simulations data

Given that (4) is an overdetermined system, where B is a tall matrix, we take the error as $e(t) = \bar{m}(t) - x(t)$ and calculate the least square approximation as $e_{ls}(t) = (B_{\mathcal{H}}^T B_{\mathcal{H}})^{-1} B_{\mathcal{H}}^T e(t)$.

We simulate (4) using the discrete-time expression

$$x(t+dt) = x(t) + (B_{\mathcal{H}}u)dt + \sigma(\mathcal{B}(t+dt) - \mathcal{B}(t)), \tag{41}$$

where the control $u(t) = e_{ls}(t)$ (see Remark 3), the step size $dt = 0.1$ and $\sigma = 0.01$. The initial state is randomly selected, and in this specific example takes the value

$$x(0) = [1\ 3\ 2\ 3\ 2\ 5\ 4],$$

while for the initial average distribution we take

$$\bar{m}(0) = [10\ 20\ 50\ 30\ 20\ 50\ 40].$$

Algorithm
Input: Set of parameters as in Table 2. **Output:** State trajectory $x(t)$ 1 : **Initialize.** Generate x_0 and $\bar{m}_0$ 2 : **for** time $t = 0, 1, \dots, T-1$ **do** 3 : **if** $t > 0$, **then** compute $\bar{m}_t$ 4 : **end if** 5 : compute least-square error $e_{ls}(t)$, 6 : compute new state $x(t+1)$ by executing (41) 7 : **end for** 12 : **STOP**

Table 3: Single-population: simulations algorithm.

We also approximate the time evolution of the average (19) by using the discrete-time expression

$$\begin{cases} \tilde{m}(t+dt) = \tilde{m}(t) - \theta\tilde{m}(t)dt, & \text{for all } t \in [0, T], \\ \tilde{m}_0 = \bar{m}_0. \end{cases}$$

where $\theta = 0.01$.

The temporal evolution of the state is depicted in Figure 5. As to be expected, the state converges to a neighborhood of the origin.

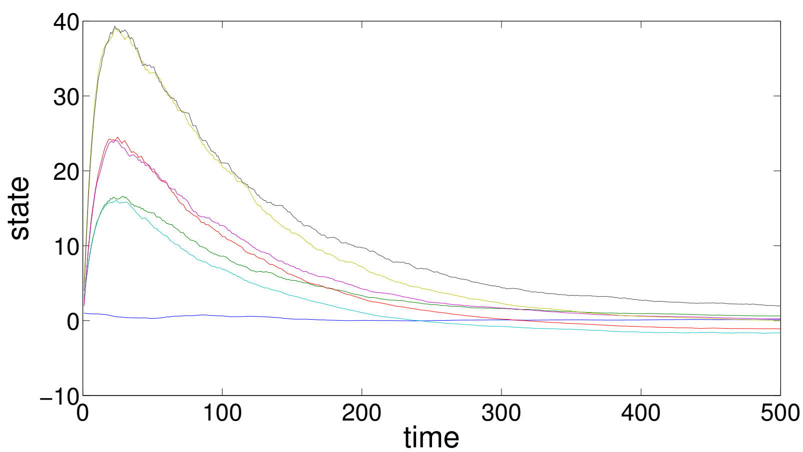

Figure 5: First simulation scenario for the single-population case: time plot of state $x(t)$ [7].

For a second scenario we take $\theta = 0$, which implies that the average $\bar{m}$ is constant and simulate the state evolution in absence of disturbance ($\sigma = 0$). The resulting time-plot is depicted in Fig. 6. We observe that the state converges to the least squares approximation of $\bar{m}(0)$.

8.2. Multi-population example

For the multi-population scenario, the numerical studies have been conducted for 100 copies of the same TU game (with three players in each group) and 5 different populations, i.e.,

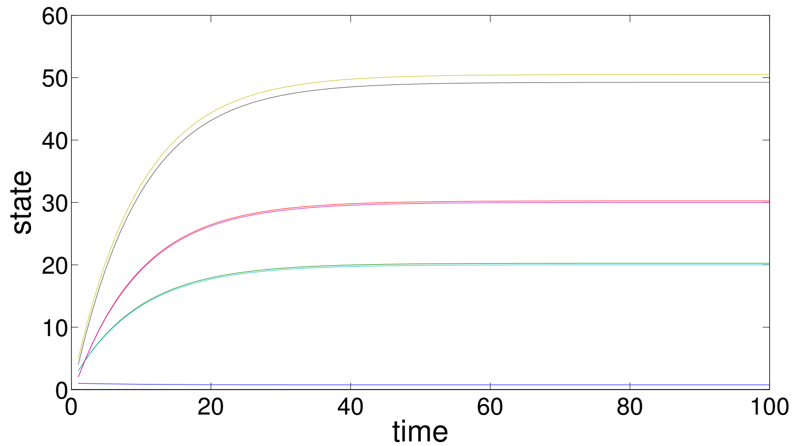

Figure 6: Second simulation scenario for the single-population case: time plot of state $x(t)$ [7].

$n = 10^2$, and $p = 5$. Interactions between populations are modeled by a graph topology $\mathcal{H}' = (V', E')$, which is a chain. The algorithm used for the simulations is as in Table 4, while the simulation parameters are listed in Table 2. We take the step size for the simulation to be $dt = 0.1$ as in the previous example. A horizon length of $T = 60$ is proven to be large enough to highlight the convergence properties of the trajectories of the excesses as a result of each player using a mean-field game strategy.

For each single player, the microscopic dynamics is captured by the following discrete-time equation:

$$x(t + dt) = x(t) + (Bu(t) - w(t))dt + \sigma\, rand[-1, 1], \tag{42}$$

where the initial state $x(0) = x$. It is worth noting that the equation stated above corresponds to (4) after time discretization. Later we explain how we randomly extract the initial state x.

For the macroscopic dynamics let us take

$$\bar{m} = (\bar{m}_1, \ldots, \bar{m}_5, \dot{\bar{m}}_1, \ldots, \dot{\bar{m}}_5)^T.$$

More formally, for each component $j = 1, \ldots q$, let us denote

$$\mu^j_{\bullet 1}(t) = \left(\ \bar{m}_{1j}(t), \ldots, \bar{m}_{5j}(t) \ \right)^T, \quad \mu^j_{\bullet 2}(t) = \left(\ \dot{\bar{m}}_{1j}(t), \ldots, \dot{\bar{m}}_{5j}(t) \ \right)^T,$$

Assuming that the mean-field equilibrium control and disturbance act on the second derivative of the local state, we have the following second-order consensus dynamics

$$\begin{bmatrix} \mu^j_{\bullet 1}(t) \\ \mu^j_{\bullet 2}(t) \end{bmatrix} = \begin{bmatrix} I & I \\ -\theta L & -\tilde{\theta}(L + \hat{\theta} I) + I \end{bmatrix} \begin{bmatrix} \mu^j_{\bullet 1}(t-1) \\ \mu^j_{\bullet 2}(t-1) \end{bmatrix} \ t = 1, 2, \ldots, T; \tag{43}$$

where the initial condition is

$$\mu^j_{\bullet 1}(0) = (\bar{m}_{1j}(0), \ldots, \bar{m}_{5j}(0))^T, \quad \mu^j_{\bullet 2}(t) = (\dot{\bar{m}}_{1j}(t), \ldots, \dot{\bar{m}}_{5j}(t))^T = (0, \ldots, 0)^T,$$

and where the matrix L has one for the entries on the main diagonal, and the reciprocal of the degree of node i for each adjacent node of i in the ith row. L is called the normalized Laplacian matrix of the communication graph $\mathcal{H}' = (V', E')$. The parameters θ, $\hat{\theta}$ and $\tilde{\theta}$ are the elastic and damping coefficients and are selected as illustrated in Table 5.

Input: Set of parameters as in Table 5.
Output: Excesses $x(t)$ for each designer
1 : **Initialize.** Generate $x(0)$ given $\bar{m}_{0k}$ and std(m_{0k}),
   for any $k = 1, \dots, p$
2 : **for** time $iter = 0, 1, \dots, T-1$ **do**
3 :   **if** $iter > 0$, **then** compute m_k, $\bar{m}_k$, and std(m_k)
4 :   **end if**
5 :   **for** player $i = 1, \dots, n$ **do**
6 :   Set $t = iter \cdot dt$ and compute control
7 :   compute $\bar{m}(t) = (\mu_{\bullet 1}(t)\mu_{\bullet 2}(t))^T$ from (43)
8 :   compute new state $x(t + dt)$ by running (42)
8 :   **end for**
9 : **end for**
10 : **STOP**

Table 4: Multi-population: simulations algorithm

set/parameters	n	dt	$std(m_0)$	T	σ	θ	$\tilde{\theta}$	$\hat{\theta}$
1st set	10^2	0.1	180	60	5.5	0.5	0.25	2.6
2nd set	10^2	0.1	180	60	5.5	0.5	0.25	1.6
3rd set	10^2	0.1	180	60	15.5	0.5	0.25	2.6

Table 5: Multi-population: simulations data.

As regards the initial distribution, we assume m_0 to be Gaussian with mean $\bar{m}_{0k} = 100 \cdot (k-1) + [1\ 2\ 3\ 4\ 5\ 6\ 7]^T$. The standard deviation $std(m_{0k})$ is set to 5.5 for all $k = 1, \dots, 5$. Then, the initial state x in (42) is obtained from a random realization with law m_{0k} for all $k = 1, \dots, 5$.

Figure 7 shows the time history of the microscopic evolution of the excesses $x_1(t), \dots, x_6(t)$ for each TU game. The time plot emphasizes two phenomena characterized by two different time scales. First, on a fast time scale, the excesses in each single population $k \in \{1, \dots, 5\}$ reach consensus to the local aggregate state ρ_k. Second, on a slower time scale, the local aggregate states reach consensus via second-order consensus dynamics.

In a second set of simulations we investigate the role of the damping coefficient $\hat{\theta}$. To do this, we analyze a scenario corresponding to a decreasing damping coefficient $\hat{\theta} = 1.6$ rather than $\hat{\theta} = 2.6$. The resulting time plot is displayed in Fig. 8. Oscillations arise, and these are visually clear from the figure.

A third set of simulations are intended to reveal the influence of the stochastic disturbance. This is done by adding a Brownian motion to the microscopic dynamics. We do this by setting the parameter $\sigma = 15.5$. The resulting time plot is displayed in Fig. 9, where one can observe small fluctuations during the transient and at steady-state. In this case, the convergence properties illustrated in this section hold as we have a fully connected graph topology. A different graph topology, perhaps characterized by more than one connected component, would lead to the formation of clusters, but this is left for future analysis.

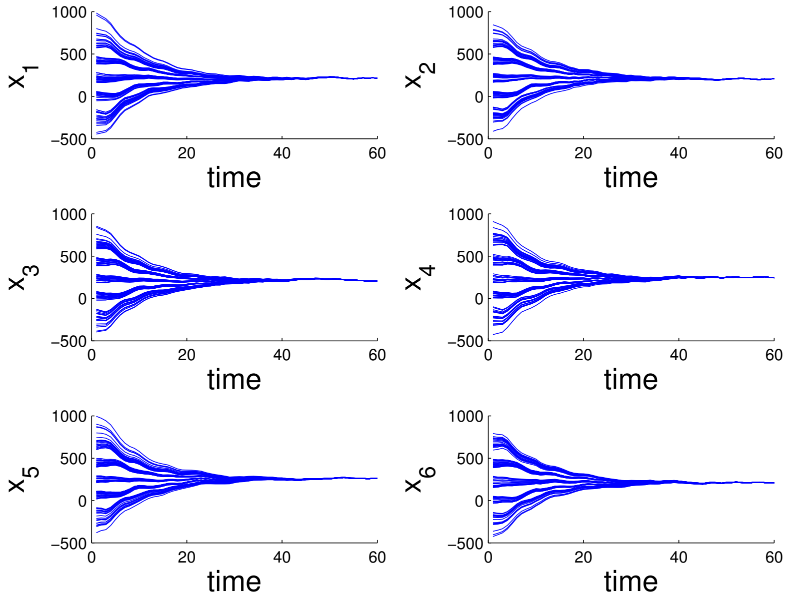

Figure 7: First simulation scenario for the multi-population case: time plot of state $x(t)$.

9. Conclusions and future directions

We have provided a mean-field game formulation of infinite copies of "small worlds", each one described as a TU coalitional game. The problem has connections to recent research on robust dynamic coalitional TU games [12] and robust mean-field games [9, 13]. A quantitative analysis of the approximation error of the solution presented is left as future work.

References

[1] L. Arnold, *Stochastic Differential Equations: Theory and Applications*, A Wiley Interscience publication, John Wiley & sons (1974).

[2] M. Bardi, "Explicit solutions of some linear-quadratic mean field games," *Network and Heterogeneous Media*, 7, pp. 243–261 (2012).

[3] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, SIAM Series in Classics in Applied Mathematics, Philadelphia (1999).

[4] T. Başar and P. Bernhard, *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach* (Birkhäuser, Boston, MA) (1995).

[5] D. Bauso, Game Theory: Models, Numerical Methods, and Applications, A tutorial in *Foundations and Trends in Systems and Control*, vol. 1, no. 4, pp. 379–523 (2014).

[6] D. Bauso and T. Başar, "Mean-field interactions among robust dynamic coalitional games with transferable utilities", *6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, Athens, 2014, pp. 487–490.

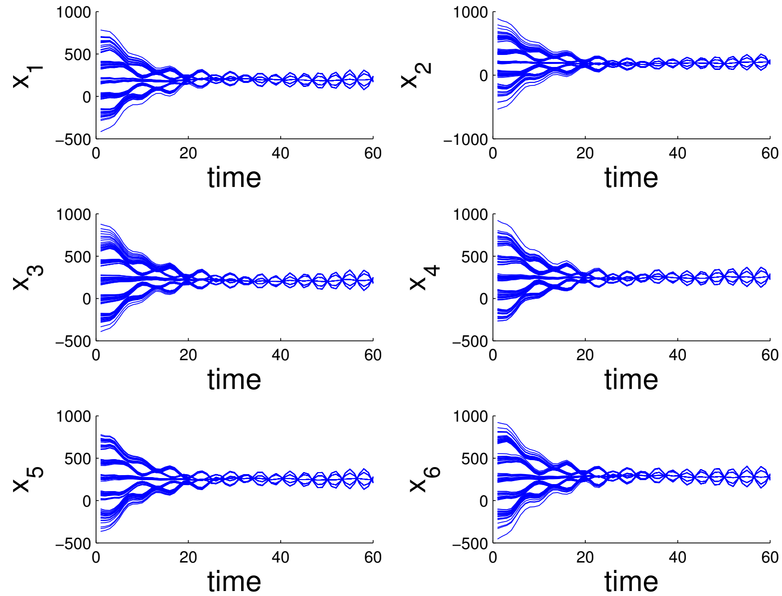

Figure 8: Second simulation scenario for the multi-population case: time plot of state $x(t)$.

[7] D. Bauso and T. Başar, "Opinion dynamics in coalitional games with transferable utilities", *IEEE 53rd Annual Conference on Decision and Control*, Los Angeles, 2014, pp. 2094–2099.

[8] D. Bauso, F. Blanchini, and R. Pesenti, "Robust control strategies for multi inventory systems with average flow constraints," *Automatica*, vol. 42, pp. 1255–1266 (2006).

[9] D. Bauso, H. Tembine, and T. Başar, "Robust mean field games with application to production of an exhaustible resource," *Proceedings of 7th IFAC Symposium on Robust Control Design*, Aalborg, Denmark (2012).

[10] D. Bauso, and J. Timmer, "Robust dynamic cooperative games" *International Journal of Game Theory*, vol. 38, no. 1, pp. 23–36 (2009).

[11] D. Bauso and J. Timmer, "On robustness and dynamics in (un)balanced coalitional games", *Automatica*, vol. 48, no. 10, pp. 2592–2596 (2012).

[12] D. Bauso, P. Viswanadha Reddy and T. Başar, "Lyapunov stochastic stability and control of robust dynamic coalitional games with transferable utilities," preprint (2012).

[13] D. Bauso and T. Mylvaganam and A. Astolfi, "Crowd-averse robust mean-field games: approximation via state space extension", IEEE Transactions on Automatic Control, in print.

[14] D. Bauso and X. Zhang and A. Papachristodoulou, "Density flow over networks: A mean-field game theoretic approach", *Proc. IEEE Conf. on Decision and Control*, Los Angeles, USA, pp. 3469–3474 (2014).

[15] E. Fehr and K.M. Schmidt, "A theory of fairness, competition, and cooperation". The Quarterly Journal of Economics, vol. 114, no. 3, pp. 817–868. doi:10.1162/003355399556151.

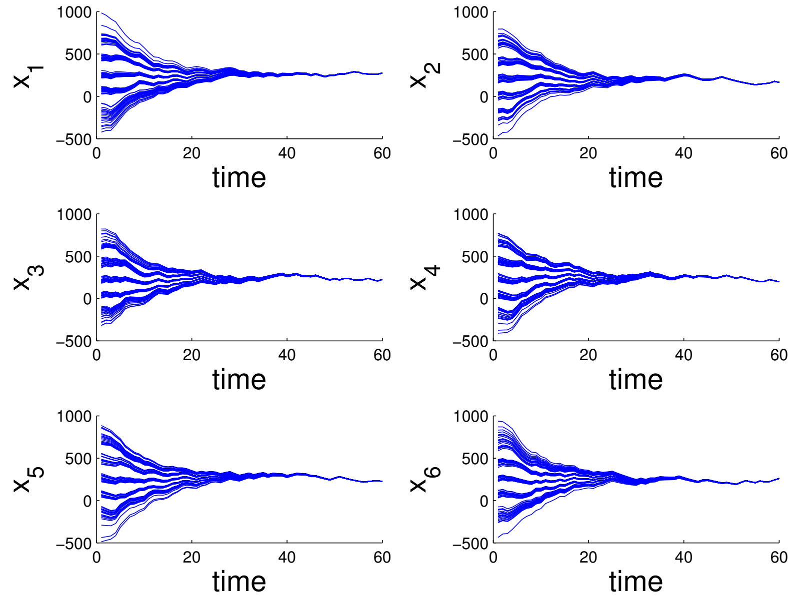

Figure 9: Third simulation scenario for the multi-population case: time plot of state $x(t)$.

[16] J. Gao, D. Li, and S. Havlin, From a single network to a network of networks, *National Science Review* (2014), available at http://nsr.oxfordjournals.org/content/early/2014/07/16/nsr.nwu020.abstract

[17] T. C. Gard, *Introduction to Stochastic Differential Equations*, vol. 114 of Monographs and textbooks in pure and applied mathematics. Marcel Dekker (1988).

[18] M.Y. Huang, P.E. Caines, and R.P. Malhamé, "Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions," *Proc. IEEE Conf. on Decision and Control*, HI, USA, pp. 98–103 (2003).

[19] M.Y. Huang, P.E. Caines, and R.P. Malhamé, "Large population cost-coupled LQG problems with non-uniform agents: individual-mass behaviour and decentralized ϵ-Nash equilibria," *IEEE Trans. on Automatic Control*, vol. 52. no. 9, pp. 1560–1571 (2007).

[20] M.Y. Huang, "Large-Population LQG Games Involving a Major Player: The Nash Certainty Equivalence Principle," *SIAM J. Control Optim.*, vol. 48. no. 5, pp. 3318–3353 (2010).

[21] J.-M. Lasry and P.-L. Lions, "Mean field games," *Jpn. J. Math.*, vol. 2, pp. 229–260 (2007).

[22] K. A. Loparo and X. Feng, Stability of stochastic systems. The Control Handbook, CRC Press, pp. 1105-1126 (1996).

[23] S. Maharjan, Q. Zhu, Y. Zhang, S. Gjessing, and T. Başar, "Demand response management in the smart grid in a large population regime" *IEEE Transactions on Smart Grid*, 7(1), pp.189–199, January 2016.

[24] J. Moon and T. Başar, "Linear-quadratic stochastic differential Stackelberg games with a high population of followers", *Proc. IEEE Conf. on Decision and Control*, Osaka, Japan, pp. 2270–2275 (2015).

[25] A. Nedić and D. Bauso, Dynamic coalitional TU games: Distributed bargaining among players' neighbors. *IEEE Trans. Autom. Control*, 58(6), pp. 1363–1376 (2013).

[26] M. Nourian and P. E. Caines, ϵ-Nash Mean Field Games Theory for Nonlinear Stochastic Dynamical Systems with Major and Minor Agents. *SIAM Journal on Control and Optimization*, 50(5), pp. 2907–2937 (2013).

[27] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Başar, "Coalitional game theory for communication networks: A tutorial", *IEEE Signal Processing Magazine, Special Issue on Game Theory*, vol. 26, no. 5, pp. 77–97 (2009).

[28] T. Samad and T. Parisini, "Systems of systems", http://ieeecss.org/sites/ieeecss.org/files/documents/IoCT-Part3-04SystemsOfSystems.pdf

[29] S. Tijs. *Introduction to Game Theory*, Hindustan Book Agency, 2003.

[30] G. Y. Weintraub, C. Benkard, and B. Van Roy, Oblivious Equilibrium: A Mean Field Approximation for Large-Scale Dynamic Games, Advances in Neural Information Processing Systems, MIT Press (2005).

[31] J. von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton Univ. Press. (1944).