Game-theoretic learning and allocations in robust dynamic coalitional games

Bauso, Dario; Tembine, Hamidou

Published in:
SIAM Journal of Control and Optimization

GAME-THEORETIC LEARNING AND ALLOCATIONS IN ROBUST DYNAMIC COALITIONAL GAMES*

M. SMYRNAKIS †, D. BAUSO ‡, AND H. TEMBINE §

Abstract. The problem of allocation in coalitional games with noisy observations and dynamic environments is considered. The evolution of the excess is modelled by a stochastic differential inclusion involving both deterministic and stochastic uncertainties. The main contribution is a set of linear matrix inequality conditions which guarantee that the distance of any solution of the stochastic differential inclusions from a predefined target set is second-moment bounded. As a direct consequence of the above result we derive stronger conditions still in the form of linear matrix inequalities to hold in the entire state space, which guarantee second-moment boundedness. Another consequence of the main result are conditions for convergence almost surely to the target set, when the Brownian motion vanishes in proximity of the set. As further result we prove convergence conditions to the target set of any solution to the stochastic differential equation if the stochastic disturbance has bounded support. We illustrate the results on a simulated intelligent mobility scenario involving a transport network.

Key words. Coalitional Games, Transferable Utility (TU), Second-moment boundedness, Intelligent mobility network, Robust control.

AMS subject classifications. 68Q25, 68R10, 68U05

1. Introduction.

The theory of coalitional games with transferable utility studies stable allocations for groups of agents who decide to cooperate (Osborne, 2004; Shapley, 1953; Aumann et al., 1960; Schmeidler, 1969; Aumann, 1961; Luce and Raiffa, 1957; Maschelr et al., 1979). Cooperation materializes in different forms such as sharing facilities, sharing costs, placing joint bids. Coalitional games arise in many areas such as: communication networks (Saad et al., 2009), smart grids (Saad et al., 2012), reconfigurable robotics (Ramaekers et al., 2011), swarm robotics (Cheng et al., 2008), multi-robot task allocation (Bayram et al., 2016).

A research area where coalitional games are an active topic is robust control (Bauso and Timmer, 2012; Wada and Fujisaki , 2017; Fele et al. , 2017). A widely used approach to solve robust control problems, (Bauso , 2017; Garud , 2005; Bauso et al. , 2015), is approachabilitty theorem (Blackwell , 1956). In Lehrer (2003), Blackwell's approachability theorem was used in order to analyse an allocation process based on coalitional games. Another technique which has been used in order to analyse game-theoretic learning algorithms is stochastic approximation. In his seminal pepar Benaim et al. (2005) showed that stochastic approximation methods can be seen as a continuous asymptotic version of approachability theorem. Based on this result in this article stochastic approximation methods are used in order to analyse coallitional games.

The results we provide collocate within the learning, control and optimisation research areas. This research direction finds applications in various problems such as wind energy (Opathella and Venkatesh , 2013; Bayens et al. , 2013), and the inventory control problem Bauso et al. (2008); Bauso et al. (2010).

In accordance with the classification provided in (Saad et al., 2009), this paper

*Submitted to the editors DATE.

Funding:

†Learning and Game Theory Laboratory, New York University Abu Dhabi (m.smyrnakis@nyu.edu).

‡Department of Automatic Control and Systems Engineering (d.bauso@sheffield.ac.uk).

§Learning and Game Theory Laboratory, New York University Abu Dhabi (tembine@nyu.edu).

answers most of the questions arising in canonical coalitional games with transferable utility (TU) within the framework of robust stabilizability. The underlying idea is that the cooperative agents, now viewed as players, form a coalition which includes all players, namely the grand coalition and need to reach agreement on how to redistribute the reward deriving from forming such a grand coalition in a way that makes the grand coalition stable. Stability is generally linked to the possibility of allocating to each sub-coalition a quantity greater than the reward itself that the sub-coalition could guarantee for itself without coalizing with the rest of the players (players outside that sub-coalition). When this occurs, we say that no players or subsets of players gain from quitting the grand coalition. This corresponds to saying that the *excess*, namely the difference between the allocated rewards and the value of the coalition is non-negative. In the broad context of coalitional games, consider the possibility that the reward of a coalition is divided among the players of the coalitions. By value of the coalition we mean the reward produced by that coalition. The procedure to allocate the reward which needs to be agreed by the players, constitutes the so-called allocation rule. Under the assumption that the values of the coalitions are time-varying and uncertain, and the allocation process occurs continuously in time, the resulting game is called *robust coalitional game*. Such a game was first formulated by (Bauso and Timmer, 2009, 2012). The evolution of the excesses is also captured by a fluid flow system of the type discussed in (Bauso et al., 2010).

The contribution of this paper is three-fold. We first formulate the problem of allocation in TU games with noisy observations and dynamic environments. In the considered scenario the evolution of the excess is subjected to both deterministic and stochastic uncertainty. The resulting dynamics can be expressed in the form of a stochastic differential inclusion, involving also a Brownian motion. For this game, as main result we provide conditions which guarantee that the distance of any solution of the stochastic differential inclusion from a predefined target set is second-moment bounded. We show that these conditions can take the form of a linear matrix inequality to be verified in different regions of the state space (Boyd et al., 1994, Chapter 6). As direct consequence of the above result we derive stronger conditions still in the form of linear matrix inequalities to hold in the entire state space, which guarantee second-moment boundedness. Further to the above main result we provide conditions for convergence almost surely to the target set, when the influence of the Brownian motion vanishes with decreasing distance from the set. The resulting dynamics mimics a geometric Brownian motion. As further result we prove convergence conditions to the target set of any solution to the stochastic differential equation if the stochastic disturbance has bounded support.

The rest of the paper is organised as follows. Section 2 introduces preliminaries on coalitional games. Section 4 discusses the model and states the problem. Section 5 links the model to saturated control and population game dynamics. Section 6 includes the main results of the paper. Section 7 specializes the model to an intelligent mobility scenario. Section 8 contains numerical examples. Finally, Section 9 provides conclusions and future works.

2. Preliminaries on TU games. This section overviews coalitional games with transferable utility (TU). Let a set $N = \{1, \dots, n\}$ of players be given and a function $\eta : S \mapsto \mathbb{R}$ defined for each non-empty coalition $S \in \mathcal{S}$, where $\mathcal{S}$ is the set of all possible non-empty coalitions, with cardinality $|\mathcal{S}| = 2^n - 1$. We denote by $< N, \eta >$ the TU game with players set N and characteristic function η, which quantifies the gain of coalition S.

Let us introduce some arbitrary mapping of $\mathcal{S}$ into $M := \{1, \dots, q\}$ where $q = 2^n - 1$, is the number of non-empty 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 labelling of the j_{th} element of $\mathcal{S}$, say S_j, according to some arbitrary but fixed ordering. Let the grand coalition be denoted by N. Furthermore, let η_j be the value of the characteristic function η associated with a non-empty coalition $S_j \in \mathcal{S}$.

Given a TU game, we wish first to investigate if the grand coalition is stable, i.e. if it is possible for the players to get better rewards by choosing a smaller coalition.

A partial answer to the above question lies in the concept of *imputation set*. The imputation set $I(\eta)$ 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 there is no individual which is benefited, increase his reward, by splitting from the grand coalition and playing alone.

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

$$I_\eta = \{\tilde{u} \in \mathbb{R}^n | \overbrace{\sum_{i \in N} \tilde{u}_i = \eta_N}^{\text{Efficiency}}, \quad \underbrace{\tilde{u}_i \geq \eta_{S_i}, \forall n_i \in \mathcal{S}'}_{\text{individual rationality}} \},$$

where $\tilde{u}_i$ is the reward allocated to player i, N here represents the grand coalition where all the players participate, $\mathcal{S}'$ is the set of all coalitions which consist of a single player and η_{S_j} is the gain of coalition S_j.

A stronger solution concept than the imputation set is the *core*. Given any allocation in the core, the players do not benefit from not only quitting the grand coalition and playing alone, but also from creating any sub-coalition. In this sense the core strengthens the conditions valid for the imputation set. Thus the core is still a polyhedral set which is included in the imputation set.

DEFINITION 2.1. *The core of a game $\langle N, u \rangle$ is the set of allocations that satisfy i) efficiency, ii) individual rationality, and iii)* super-additivity*, i.e. stability with respect to sub-coalitions:*

$$C_\eta = \{\tilde{u} \in I(\eta) | \quad \underbrace{\sum_{i \in S_j} \tilde{u}_i \geq \eta_{S_j}, \forall S_j \in \mathcal{S}}_{\textit{stability w.r.t. subcoalitons}} \quad \}.$$

Even though the core is a fundamental concept in coalitional games, it is not necessary that the core will be a non-empty set. Two broad categories of coalitional games with non-empty core are: convex (Shapley, 1971) and balanced games (Bondareva, 1963; Shapley, 1967).

DEFINITION 2.2. *A coalitional game $< N, \eta >$ is convex if the following inequality is satisfied.*

$$\eta_{S_i} + \eta_{S_j} \leq \eta_{S_i \cap S_j} + \eta_{S_i \cup S_j}, \forall S_i, S_j \subset N.$$

DEFINITION 2.3. *A coalitional game $< N, \eta >$ is balanced if for any balanced map α we have:*

$$\sum_{j \in \mathcal{S}} \alpha_{S_j} \eta_{S_j} \leq \eta_N.$$

In order to overcome the problem of an empty core in (Shapley and Shubik, 1966) the notion of ϵ-core was introduced

DEFINITION 2.4. *For a real number ϵ the ϵ-core is defined as:*

$$C_\eta = \{\tilde{u} \in I(\eta) | \sum_{i \in S_j} \tilde{u}_i \geq \eta_{S_j} - \epsilon, \forall S_j \in \mathcal{S}\}.$$

In order to assess stability of the grand coalition, the core, both its value η_N, and the reward allocated to each player is needed. Therefore, there is a need to define an allocation mechanism of the coalition's rewards among the players. One of the most used allocation mechanisms is the Shapley value (Shapley, 1953, 1971). An additional reason for choosing Shapley's value is its connection with feedback control and uncertainty as it was shown in (Bauso and Timmer, 2012)

DEFINITION 2.5. *The Shapley value of player i, given a coalitional game $< N, \eta >$ is defined as:*

$$\phi_i(\eta) = \sum_{S_j \subset N \setminus \{i\}} \frac{|S_j|!(|N| - |S_j| - 1)!}{|N|!} (\eta_{S_j \cup \{i\}} - \eta_{S_j}).$$

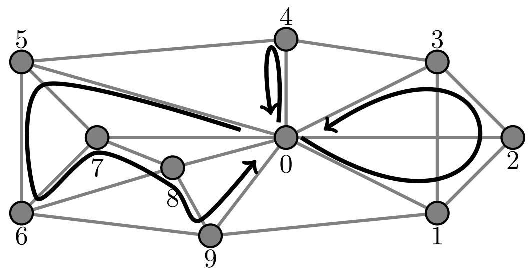
The Shapley value can be interpreted as the expected weighted contribution of player i when it joins the grand coalition in a random order.

3. Motivating example. Various applications of the TU games have been considered in literature. Examples include Market games (Shapley and Shubic, 1969), public good games (Bodwin, 2017), the bankruptcy problem (Aumann and Maschler, 1985) and inventory problems (Chinchuluun et al., 2008). Applications which combine TU games with optimisation and learning include micro-grid problems (Saad et al. , 2013) and coordinated replenishment (Bauso and Timmer, 2009).

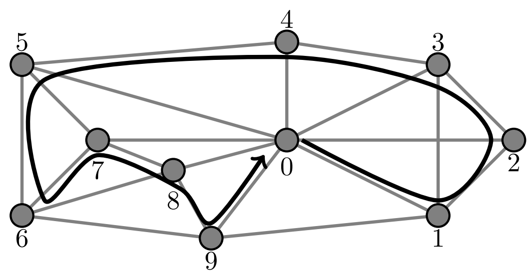
The case study which is considered in this article, the intelligent mobility network application, falls in the category of the inventory problems. Players should decide if it is more beneficial to create a coalition and share the cost of the inventory or it is better to bear the cost alone.

Intelligent mobility deals with the smart transport of items, goods or individuals from source to destination nodes using shared facilities like buses, trams, electric vehicles. Suppose that items are initially stored in the supply centre indexed by 0 and need to be transported to different destination centres generically indexed by i, $i = 1, \ldots, n$. Destination centres are characterized by a time-varying demand which is independent identically distributed across time and centres.

Note here that the capacitate vehicle routing problem is usually solved in two parts. In the first one the assignment problem is solved, i.e. one makes decisions about the sites that should be visited. In the second part the optimal route is found through traveller salesman algorithms for example. In this article we focus on the first part, where the network topology is not playing a significant role. The manager of destination center i bids the quantity to be transported from the supply center and terminating in center i based on his forecast of the future demand. Managers can collaborate and place joint bids with the advantage of compensating potential fluctuation of the their demand. This can be represented using a graph and a cycle, namely, a closed path with source and destination in node zero, see for instance the three transport cycles originating from and terminating in 0 and touching destination centres $\{1, 2, 3\}$, $\{4\}$, and $\{5, \ldots, 9\}$ in the network of Figure 1(a).



(a) Three transport cycles originating from and terminating in 0 and touching destination centers $\{1, 2, 3\}$, $\{4\}$, and $\{5, \ldots, 9\}$.



(b) One transport cycle originating from and terminating in 0 and touching all destination centers.

FIG. 1. *Example of a distribution network*

When all managers act jointly, we say that they form a grand coalition. In such case a single cycle will touch all destination centres as described by the transport cycle originating from and terminating in 0 and touching all destination centres in Figure 1(b).

In stable environments, in cases where the cost function of players is deterministic, and it possible to obtain observations without noise the conventional analysis of TU games can be applied, i.e. results about the existence of the core, or the evaluation of nucleus or Shapley's value.

In particular, consider the scenario where $N = \{1, \ldots, n\}$ be the set of receiving centres. For each coalition $S \in \mathcal{S}$, let D_S be a random variable representing the aggregate demand faced by that coalition. Let us assume that D_S has continuous probability density function $f(D_S)$. In other words, the probability that the aggregate demand is between a and b is

$$\mathbb{P}(a \leq D_S \leq b) = \int_a^b f(D_S)\, \mathrm{d}D_S.$$

The continuous cumulative distribution function (CDF) is $F(b)$, and represents the probability that the aggregate demand is less than or equal to b:

$$F(b) := \mathbb{P}(D_S \leq b) = \int_0^b f(D_S)\, \mathrm{d}D_S.$$

Let Θ be the order quantity, p in $\mathbb{R}_+$ be the sale price, s in $\mathbb{R}_+$ be the penalty price for shortage, when demand exceeds supply, and let h in $\mathbb{R}_+$ be the penalty price for holding, when supply exceeds demand.

Introduce the stock variable $Z_S = \Theta - D_S$. Denote the indicator function by

$$\mathbf{I}_{\mathbb{R}_+}(Z_S) = \begin{cases} 1 & \text{if } Z_S \in \mathbb{R}_+ \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Then, the expected profit for the generic coalition $S \in \mathcal{S}$ under the order quantity Θ is given by

$$\langle \mathcal{P}_S(D_S, \Theta) \rangle = \mathbb{E}\Big[p \min(\Theta, D_S) - c\Theta \quad - [s\mathbf{I}_{\mathbb{R}_+}(Z_S) - h\mathbf{I}_{\mathbb{R}_+}(-Z_S)]|Z_S|\Big]. \tag{2}$$

In the above we express the expected profit as function of the expected shortage and expected holding, which are given by

$$\begin{aligned} \mathbb{E}\Big[\mathbf{I}_{\mathbb{R}_+}(-Z_S)|Z_S|\Big] &= \textstyle\int_\Theta^\infty f(D_S)(D_S - \Theta)\,\mathrm{d}D_S, \\ \mathbb{E}\Big[\mathbf{I}_{\mathbb{R}_+}(Z_S)|Z_S|\Big] &= \textstyle\int_0^\Theta f(D_S)(\Theta - D_S)\,\mathrm{d}D_S. \end{aligned} \tag{3}$$

We can then rewrite the expected profit as

$$\begin{aligned} \langle \mathcal{P}_S(D_S, \Theta) \rangle &= \mathbb{E}[p \min(\Theta, D_S)] - c\Theta \\ -s\mathbb{E}\Big[\mathbf{I}_{\mathbb{R}_+}(-Z_S)|Z_S|\Big] &- h\mathbb{E}\Big[\mathbf{I}_{\mathbb{R}_+}(Z_S)|Z_S|\Big]. \end{aligned} \tag{4}$$

The following relation between the expected shortage $\mathbf{E}_s$ and the expected holding $\mathbf{E}_h$ holds:

$$\begin{aligned} \mathbb{E}\Big[\mathbf{I}_{\mathbb{R}_+}(Z_S)|Z_S|\Big] &= \textstyle\int_0^\Theta f(D_S) Z_S \,\mathrm{d}D_S \\ &= \textstyle\int_0^\infty f(D_S) Z_S \,\mathrm{d}D_S - \int_\Theta^\infty f(D_S) Z_S \,\mathrm{d}D_S \\ &= \Theta - \langle D_s \rangle + \mathbb{E}\Big[\mathbf{I}_{\mathbb{R}_+}(-Z_S)|Z_S|\Big], \end{aligned}$$

where $\langle y_s \rangle$ is the mean demand and is given by $\int_0^\infty f(D_S) D_S \,\mathrm{d}D_S$. The problem faced by the coalition is the one of maximizing the expected profit with respect to the order quantity Θ, which is the decision variable:

$$\begin{aligned} \max_\Theta \Big\{ &\mathbb{E}[p \min(\Theta, D_S)] - c\Theta \\ &-s\mathbb{E}\Big[\mathbf{I}_{\mathbb{R}_+}(-Z_S)|Z_S|\Big] - h\mathbb{E}\Big[\mathbf{I}_{\mathbb{R}_+}(Z_S)|Z_S|\Big]\Big\}. \end{aligned}$$

Assuming concavity of $\langle \mathcal{P}_S(D_S, \Theta) \rangle$ the optimal order quantity Θ^* is obtained by computing the derivative of $\langle \mathcal{P}_S(D_S, \Theta) \rangle$ with respect to Θ and taking it equal to zero. To do this, after rearranging the first term $\mathbb{E} \min(\Theta, D_S)$ in the above equation as below

$$\begin{aligned} \mathbb{E} \min(\Theta, D_S) &= \textstyle\int_0^\Theta D_S f(D_S)\,\mathrm{d}D_S + \int_\Theta^\infty \Theta f(D_S)\,\mathrm{d}D_S \\ &= \langle D_S \rangle - \textstyle\int_\Theta^\infty D_S f(D_S)\,\mathrm{d}D_S + \int_\Theta^\infty \Theta f(D_S)\,\mathrm{d}D_S \end{aligned}$$

we can rewrite the expected profit as

$$\begin{aligned} \langle \mathcal{P}_S(D_S, \Theta) \rangle &= p\langle D_S \rangle - c\Theta \\ &-s\Theta \textstyle\int_0^\Theta f(D_S)\,\mathrm{d}D_S + s\int_0^\Theta D_S f(D_S)\,\mathrm{d}D_S \\ &+(p+h)\Theta \textstyle\int_\Theta^\infty f(D_S)\,\mathrm{d}D_S - (p+h)\int_\Theta^\infty D_S f(D_S)\,\mathrm{d}D_S. \end{aligned}$$

Then for the derivative we have

$$\begin{aligned}
&\tfrac{d}{dC}(\langle \mathcal{P}_S(D_S,\Theta)\rangle)\\
&= -c - s\int_0^{\Theta} f(D_S)\,\mathrm{d}D_S - s\Theta f(\Theta) + s\Theta f(\Theta)\\
&+(p+h)\int_{\Theta}^{\infty} f(D_S)\,\mathrm{d}D_S - (p+h)\Theta f(\Theta) + (p+h)\Theta f(\Theta)\\
&= -c - s\int_0^{\Theta} f(D_S)\,\mathrm{d}D_S + (p+h)\int_{\Theta}^{\infty} f(D_S)\,\mathrm{d}D_S\\
&\quad = -c - sF(\Theta) + (p+h)[1-F(\Theta)],
\end{aligned}$$

where F is the cumulative distribution function (CDF) of y. The optimal order quantity is given by:

$$F(\Theta_S^*) = \frac{p+h-c}{p+h+s}. \tag{5}$$

Let F^{-1} be the inverse function of F then it holds

$$\Theta_S^* = F^{-1}\Big(\frac{p+h-c}{p+h+s}\Big). \tag{6}$$

Then, the optimal expected profit is

$$\begin{aligned}
\langle \mathcal{P}_S(D_S,\Theta_S^*)\rangle &= p\mu - c\Theta_S^* - s\int_0^{\Theta_S^*}(\Theta_S^* - D_S) f(D_S)\,\mathrm{d}D_S\\
&\quad -(p+h)\int_{\Theta_S^*}^{\infty}(D_S - \Theta_S^*) f(D_S)\,\mathrm{d}D_S\\
&= p\mu - c\Theta_S^* - s(\Theta_S^* - \mu + \mathbf{E}_h^*) - (p+h)\mathbf{E}_h^*\\
&= p\mu - cF^{-1}\Big(\tfrac{p+h-c}{p+h+s}\Big) - s\Big(F^{-1}\Big(\tfrac{p+h-c}{p+h+s}\Big)\\
&\quad -\mu + \mathbf{E}_h^*\Big) - (p+h)\mathbf{E}_h^*,
\end{aligned} \tag{7}$$

where we denote by $\mathbf{E}_h^*$ the expected surplus under the optimal order quantity Θ_S^*.

Consider a sequence of sampling intervals indexed by $k = 0, 1, \ldots$. We build on the results for the optimal order quantity (6) and expected profit (7), which we have obtained above. We assume that the demand at interval k has a Normal distribution with mean $D_S(k-1)$ and variance σ^2:

$$D_S(k) - D_S(k-1) \sim \mathcal{N}(0, \sigma^2). \tag{8}$$

We can rewrite the optimal order quantity in terms of the number of standard deviations away from the mean:

$$\Theta_S^* = D_S(k-1) + k^*\sigma,$$

where k has standard Normal distribution. Denote by $\Phi(k)$ the CDF of a standard Normal distribution, from (5) we have

$$\Phi(k^*) = \frac{p+h-c}{p+h+s}.$$

To obtain (6) from (5), we introduced the inverse function F^{-1}. We follow the same procedure here and consider the inverse function Φ^{-1} of Φ. Then, for the optimal k^* it holds

$$k^* = \Phi^{-1}\Big(\frac{p+h-c}{p+h+s}\Big).$$

Denote the expected surplus of k as

$$G(k) = \int_k^\infty (D_S - k) f(D_S)\,\mathrm{d}D_S.$$

Then, from (7) the optimal expected profit is

$$\begin{aligned}\langle \mathcal{P}_S(D_S, \Theta_S^*)\rangle &= p\mu - c(D_S(k-1) + k^*\sigma) \\ &\quad - s[k^*\sigma + \sigma G(k^*)] - (p+h)\sigma G(k^*) \\ &= p\mu - c y_{k-1} \underbrace{-\sigma(c+s)}_{<0} k^* \underbrace{-\sigma(s+p+h)}_{<0} G(k^*).\end{aligned}$$

Note that the expected profit decreases with the standard deviation σ, namely, the volatility of the demand.

Coalition games that are subject to probabilistic demand/ characteristic function, as in the aforementioned example, have been also studied in the context of stochastic cooperative games (Suijs et al., 1997; Toriello and Nelson, 2017). In that context conditions for a stable core were devised. Similarly the news agent problem (Muller et al., 2002; Hartman and Dror, 2005; Slikker et al., 2005) is a coalition problem where probabilistic utilities emerge. The literature concerning this problem also focuses on conditions for non-empty core and fair allocations.

In the current article a different approach is adopted. The control of the stochastic process in order to be bounded around the core is considered, instead of trying to define suitable conditions for the core of the game to be non-empty. As a result a formulation of TU games with dynamically changing characteristic function, which allows its representation as a stochastic process is provided. A saturated controller is used in order for the process to be bounded around the core. The proposed controller resembles the "Best response" decision making process. Hence, stochastic differential inclusions emerge from the control process. Therefore, analysis of a stochastic process which can be occured through the TU game formulation is provided, based on the theory of stochastic differential inclusions Benaim et al. (2005).

Since the cost function is not constant throughout the game any more and in each time step of the decision making process a fluctuated version of the cost function is available because either of changes in the environment or noisy observations. This analysis focuses on the control of the outcome of the stochastic process either to be in the core or bounded in the ϵ-core based on the volatility of the perturbations.

4. Model and problem statement. This section is separated into two parts. The fist contains the description of the dynamic TU model and provides an illustrative example of a 3-player game. The second part contains the representation of the dynamic TU game as a stochastic process and a proposed control strategy which allows an a solution bounded in the e-core of the dynamic TU-game. The distance ϵ from the core depends on the volatility of the stochastic process.

4.1. TU Games with noisy observations. A *dynamic TU game* is described by $< N, \eta(t) >$, where $\eta(t)$ is a time-varying characteristic function representing the values of different coalitions. In real life applications there are many uncontrollable processes which introduce uncertainty either on the rewards of the coalitional games or the observations of the other players' decisions. In the intelligent mobility network problem, of the previous section, managers can have an estimate of the ordering capacities of the other managers. This estimate can be of the form of a probability distribution which changes over time. Therefore, the uncertainty can be modelled as a stochastic process.

It possible to represent a dynamic TU game in Matrix form. In addition, following the dynamic programming paradigm, all the constraints which arise from the definition of the core can be represented as inequalities. In particular, let $B_{\mathcal{H}}$ be a $((q-1)\times n)$-matrix whose rows are the characteristic vectors $y^{S_j} \in \mathbb{R}^n$ of each coalition other than the grand coalition, i.e., $S_j \in \mathcal{S}$, $S_j \neq N$. In other words $B_{\mathcal{H}} = \{(y^{S_j})^T\}_{S_j \in \mathcal{S},\, S_j \neq N}$.

The characteristic vectors are in turn binary vectors representing the participation or not of a player i in the coalition S_j, whereby $y_i^{S_j} = 1$ if $i \in S_j$ and $y_i^{S_j} = 0$ if $i \notin S_j$. For any allocation in the *core* of the game $C(\eta(t))$ we have:

$$\tilde{u}(t) \in C(\eta(t)) \quad \Leftrightarrow \quad B_{\mathcal{H}}\tilde{u}(t) \geq \eta(t), \tag{9}$$

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

Let

$$B = \begin{bmatrix} B_{\mathcal{H}} & -I \\ \mathbf{1}^T & \mathbf{0}^T \end{bmatrix} \in \{-1,0,1\}^{q\times n+(q-1)}. \tag{10}$$

Inequality (9) can be rewritten as an equality by using an augmented allocation vector given by $u := \left[\begin{smallmatrix}\tilde{u}\\ s\end{smallmatrix}\right] \in \mathbb{R}^{n+q-1}$ where s is a vector of $q-1$ non-negative surplus variables. Then, we have

$$Bu(t) = \eta(t). \tag{11}$$

For a 3-player coalitional game equation (11) takes the form

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

Remark Note here that in general TU coalitional games, as well as the formulation which is proposed in this article, suffer from the curse of dimensionality. In particular, the dimensionality of B will exponentially increase with the number of players and possible actions. In that case a distributed solution as the one in (Nedich and Bauso , 2013) can be used in order to cluster the problem to smaller sub-problems which are feasible to be solved.

4.2. TU games as a stochastic process. Let us assume that the perturbations of the characteristic function are bounded in an ellipsoid. Let $w(t)$ denote the perturbed observation of the players at time t, $w_0(t)$ being the time-varying characteristic function and $\tilde{w}(t)$ the perturbation term, such as a bias in the estimator of the characteristic function $w_0(t)$. In the case of an additive perturbation term the drift from $w_0(t)$ can be expressed as $w(t) = [w_0(t) + \tilde{w}(t)]$. The analysis of the dynamic TU games which follows in the rest of this article is based on the assumption that the perturbations are bounded in an ellipsoid, i.e $w(t)$ can be written as:

$$w(t) \in \mathcal{W} = \{w \in \mathbb{R}^q : w^T R w \leq 1\}. \tag{12}$$

The changes in the characteristic function as they are realised by the players can be written then as

$$d\eta(t) = w(t)dt - \Sigma d\mathcal{B}(t), \quad \text{in } \mathbb{R}^q, \tag{13}$$

where $\Sigma d\mathcal{B}(t)$ is a random noise with zero mean and $\Sigma = diag((\Sigma_{ii})_{i=1,\dots,q}) \in \mathbb{R}^{q\times q}$ for given scalars Σ_{ii}, all full column rank, and $\mathcal{B}(t) \in \mathbb{R}^q$ is a q-dimensional Brownian motion, which is independent across its components, independent of the initial state η_0, and independent across time.

Instead of studying the evolution of the characteristic function in order to solve a TU game the surpluses s_j can be studied. Note that the difference between the allocated value and the coalitional S_j, corresponds to surplus variable s_j and is described as,

$$s_j(t) = \sum_{i\in S_j} \tilde{u}_i(t) - \eta_j(t).$$

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 insight is that *if all the surpluses are non-negative, then the total allocation to any coalition exceeds the value of the coalition itself and the allocation vector lies in the core.* Also, note 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.

Let $x(t) \in \mathbb{R}^q$, denote the cumulative excess which is obtained as follows. In essence, every component of vector $Bu(t)$ is the total reward given to the members of a coalition at time t, and the drift from this reward, $w(t)$, is subtracted. Then, a positive $x(t)$ means positive cumulative excess.

Let us denote the controler in linear state feedback form as:

$$u(x) = K(x,t)x, \tag{14}$$

where $K(x,t) \in \mathbf{co}\{K^{(i)}\}_{i\in I}$.

Then the problem of stabilising the core can be cast as a problem of solving the following stochastic differential inclusion:

$$dx(t) \in F(x)dt + \Sigma d\mathcal{B}(t). \tag{15}$$

Also,

$$\begin{aligned} F(x) := \{\xi \in \mathbb{R}^q |\, \xi = (BK(x,t) - I)x - w, \\ K(x,t) \in \mathbf{co}\{K^{(i)}\}_{i\in I},\, w \in \mathcal{W}\}, \end{aligned} \tag{16}$$

for assigned polytopic sets $\mathbf{co}\{K^{(i)}\}_{i\in I}$, and ellipsoidal set $\mathcal{W}$, and where $\mathcal{B}(t)$ is a Brownian motion weighted by a matrix Σ and B defined as in (10).

The stability, well-posedness and existence of solution to (15), when saturated linear controllers are used has been studied in Hu et al. (2006); Cai et al. (2009); Hu et al. (2005); Jokic et al. (2008); Grammatico et al. (2014).

For any symmetric positive definite matrix $P \in \mathbb{R}^{n\times n}$, define the function $V(x) = x^T Px$ and the ellipsoidal target set $\Pi = \{x \in \mathbb{R}^n : V(x) \leq 1\}$. We are interested in studying convergence of the solutions of (15) to the target set.

5. Examples. The stochastic differential inclusion (15) arises in the case of saturated controls, and in the case of two-population games. We discuss next these three examples.

5.1. Example 1: saturated controls. Assume that controls are bounded within polytopes

$$u(t) \in \mathcal{U} = \{u \in \mathbb{R}^{(q-1)+n} : u^- \leq u \leq u^+\}, \tag{17}$$

where u^+, u^- are assigned vectors. Note that we can assume the characteristic function centred at zero as in (12) as we can always center the hypercube of $u(t)$ around any desired value.

In addition, for any matrix $K \in \mathbb{R}^{n+(q-1)\times q}$, define as saturated linear state feedback control any policy

$$u = -sat\{Kx\} = \begin{cases} -Kx & \text{if } Kx \in \mathcal{U} \\ u(x) \in \partial\mathcal{U} & \text{otherwise,} \end{cases} \tag{18}$$

where $\partial\mathcal{U}$ indicates the frontier of set $\mathcal{U}$.

In the above, the $sat\{.\}$ operator has to be interpreted component-wise, namely

$$u_i = sat_{[u_i^-, u_i^+]}\{-K_{i\bullet}x\}, \tag{19}$$

where $K_{i\bullet}$ denotes the ith row of K and where, for any given scalar a and b

$$sat_{[a,b]}\{\zeta\} = \begin{cases} b, & \text{if } \zeta > b, \\ \zeta, & \text{if } a \leq \zeta \leq b, \\ a, & \text{if } \zeta < a. \end{cases}$$

Henceforth we omit the indices of the sat function.

Under the control $u = sat\{-Kx\}$, the closed-loop dynamics mimics the differential inclusion (15) as follows

$$dx \in \{(-x + Bsat\{-Kx\} - w)dt + \Sigma d\mathcal{B}(t), w \in \mathcal{W}\}.$$

5.2. Example 2: distribution network. Consider a distribution network problem where there is a demand for a specific commodity and the reward for supplying it is suitably described by our control law. When the demands are based on a diffusion process, their evolution can be written as:

$$\dot{d} = w(t) - \sum d\mathcal{B}(t). \tag{20}$$

Then (13) can be written with respect to $\dot{d}$ as:

$$d\eta(t) = [w_0(t) + \dot{d}(t) + \Sigma d\mathcal{B}(t)]dt - \Sigma d\mathcal{B}(t).$$

The excess then can be written as

$$dx(t) = (-x(t) + B_{\mathcal{H}}u(t))dt - d\eta(t), \tag{21}$$

where u is the control vector as defined in (18).

$u^{(j)} \backslash w^{(k)}$	$w^{(1)}$	$\cdots$	$w^{(\tilde{q})}$
$u^{(1)}$	$Bu^{(1)} - w^{(1)}$	$\cdots$	$Bu^{(1)} - w^{(\tilde{q})}$
$\vdots$	$\vdots$		$\vdots$
$u^{(\tilde{p})}$	$Bu^{(\tilde{p})} - w^{(1)}$	$\cdots$	$Bu^{(\tilde{p})} - w^{(\tilde{q})}$

TABLE 1
The possible vector payoffs.

5.3. Example 3: approachability. Equation (15) is in the same spirit as in Hart and Mas-Colell's paper (Hart and Mas-Colell, 2003) on continuous-time approachability.

In particular (15), can be obtained when a 2-player repeated game with vector payoffs as displayed in Table 1, is considered. Let $A_1 = \{u^{(1)}, \dots, u^{(\tilde{p})}\}$ and $A_1 = \{w^{(1)}, \dots, w^{(\tilde{q})}\}$ be the actions sets of player 1 and 2. Denote $a_1 = [a_{11}, \dots, a_{1\tilde{p}}]^T$ and $a_2 = [a_{21}, \dots, a_{2\tilde{q}}]^T$ the mixed strategies of player 1 and 2, respectively. Introduce the mixed extension mapping $\Delta(A_1) \times \Delta(A_2) \to \mathcal{U} \times \mathcal{W}$, such that $(a_1, a_2) \mapsto (u, w)$ where

$$u = \sum_{j=1}^{\tilde{p}} a_{1j} u^{(j)}, \quad w = \sum_{k=1}^{\tilde{q}} a_{2j} w^{(k)}.$$

Consider the time-average expected (over opponent's play) payoff defined as

$$\Gamma(s) = \frac{1}{s} \int_0^s (Bu - w)\, \mathrm{d}\tau \in \mathbb{R}^q.$$

If we rescale the time window using $s = e^t$, take $x(t) = \Gamma(e^t)$ and differentiate with respect to t, we obtain the differential equation (15). Note that, after rescaling the time window, we have

$$x(0) = \int_0^1 (Bu - w)\, \mathrm{d}\tau \in \mathbb{R}^q.$$

Adopting a "population-game dynamics" perspective, the state $x(t) \in \mathbb{R}^q$ represents the current average payoff over the population.

6. Main results. In this section it is shown that the second moment of the deviations from the core, $x(t)$, is bounded, when a saturated linear feedback controller is used. This is achieved by the use of polytopic techniques (Mayne , 2003). Polytopic constraints are widely used in order to model problems related to robust control problems when the transition matrix of the process is state-dependent, i.e. $\dot{x} = A(x)x$. In addition, because no further constraints have been imposed on (15), the proposed methodology can be used to control dynamic TU games when (15) describes the dynamics of the game.

Our idea is to rewrite the above dynamics in the following polytopic form

$$dx \in \{(BK(x,t) - I)x(t) - w(t)dt + \Sigma d\mathcal{B}(t),\, w \in \mathcal{W}\}, \tag{22}$$

where the time varying matrices $K(x,t)$ are expressed as convex combinations of $|I|$ matrices $K^{(i)}$, $i \in I$. More precisely the expressions for $K(x,t)$ are

$$K(x,t) = \sum_{i \in I} \tilde{\sigma}_i(x,t) K^{(i)}, \quad \sum_{i \in I} \tilde{\sigma}_i(x,t) = 1. \tag{23}$$

The control policy is then

$$u = Kx = (\sum_{i \in I} \tilde{\sigma}_i(x,t) K^{(i)})x, \quad \sum_{i \in I} \tilde{\sigma}_i(x,t) = 1.$$

In the case of saturated controls the procedure to derive the weights in the above control policy are discussed in (Gomes da Silva, 2001).

THEOREM 6.1. *The distance of any solution of the stochastic differential inclusion (15) from the target set* Π *is second-moment bounded if for all* $x \in X_j$, $j \in I$

$$x^T \Big[Q(\Psi^{(i)})^T + \Psi^{(i)} Q + \alpha Q + \frac{1}{\beta} R^{-1} \Big] x \le 0, \tag{24}$$

where $\Psi^{(i)} = [BK^{(i)} - I]$ *and* X_j *is any subspace where* $K^{(i)}$ *is in the support* $\mathcal{S}_j$ *of* K*, i.e., the control is*

$$u = Kx = (\sum_{i \in \mathcal{S}_j} \tilde{\sigma}_i(x,t) K^{(i)})x, \quad \sum_{i \in \mathcal{S}_j} \tilde{\sigma}_i(x,t) = 1.$$

Proof. The analysis is then performed within the framework of stochastic stability theory (Loparo and Feng, 1996). To this end, consider the infinitesimal generator

$$\mathcal{L}[\cdot] = \lim_{dt \to 0} \frac{\frac{1}{2}\mathbb{E} \sum_{i \in I} dx^T \nabla^2_{xx}[\cdot] dx + \mathbb{E} dx^T \nabla_x [\cdot]}{dt}, \tag{25}$$

and the Lyapunov function $V(x) = x^T P x$. The stochastic derivative of $V(x)$ is obtained by applying (25) to $V(x)$, which yields

$$\begin{aligned} \mathcal{L}V(x(t)) &= \lim_{dt \to 0} \frac{\mathbb{E}V(x(t+dt)) - V(x(t))}{dt} \\ &= \lim_{dt \to 0} \frac{\frac{1}{2}\mathbb{E}\sum_{i \in I} dx^T \nabla^2_{xx}[V(x)] dx + \mathbb{E} dx^T \nabla_x [V(x)]}{dt} \\ &= \frac{1}{2} \sum_{i \in I} \Sigma^2_{ii}(x) (\nabla^2_{xx}[V(x)])_{ii} + [BK(\cdot)x - x - w]^T \cdot \\ &\quad \cdot \nabla_x [V(x)] + \nabla_x [V(x)]^T [BK(\cdot)x - x - w]. \end{aligned}$$

Using $\nabla^2_{xx}[V(x)] = P$ and $\nabla_x[V(x)] = Px$ the above can be rewritten as follows, for all $x \notin \Pi$, and $w \in \mathcal{W}$

$$\begin{array}{l} \mathcal{L}V(x) = [-x + BK(x,t)x - w]^T P x \\ +x^T P[-x + BK(x,t)x - w] + \sum_{i=1}^q \Sigma^2_{ii}(x) P_{ii} \\ = x^T [BK(x,t) - I]^T P x + x^T P [BK(x,t) - I] x \\ -w^T P x - x^T P w + \sum_{i=1}^q \Sigma^2_{ii} P_{ii} < 0. \end{array} \tag{26}$$

Let $\overline{\Pi} = \mathbb{R}^q \setminus \Pi$. From the $\mathcal{S}$-procedure, we know that for all $x \in \overline{\Pi}$, and $w \in \mathcal{W}$ condition (26) holds if there exist $\alpha, \beta \ge 0$, such that for all $(x,w) \in \overline{\Pi} \times \mathcal{W}$

$$\begin{array}{rl} \mathcal{L}V(x) & = x^T [BK(x,t) - I]^T P x \\ & +x^T P [BK(x,t) - I] x \\ & -w^T P x - x^T P w + \sum_{i=1}^q \Sigma^2_{ii} P_{ii} \\ & \le \alpha(1 - V(x)) + \beta(\|w\|^2_R - 1) \le 0. \end{array} \tag{27}$$

The last inequality is obtained from observing that

$$\overline{\Pi} \times \mathcal{W} := \{(\xi, \omega) : 1 - V(\xi) \leq 0, \; \|\omega\|_R^2 - 1 \leq 0\}.$$

Let $\Psi(x,t) = [BK(x,t) - I]$, inequality (27) can be rewritten as

$$\begin{bmatrix} x \\ w \end{bmatrix}^T \begin{bmatrix} \Psi(x,t)^T P + P\Psi(x,t) + \alpha P & -P \\ -P & -\beta R \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} \\ -\alpha + \beta + \sum_{i=1}^q \Sigma_{ii}^2 P_{ii} \leq 0.$$

Trivially it must hold $\beta \leq \alpha$. Assume without loss of generality that $\beta = \alpha - \sum_{i=1}^q \Sigma_{ii}^2 P_{ii}$.[1] Recall that α and β can be chosen arbitrarily. After pre and post-multiplying by $Q = P^{-1}$, the above condition becomes

$$\begin{bmatrix} x \\ w \end{bmatrix}^T \begin{bmatrix} Q\Psi(x,t)^T + \Psi(x,t)Q + \alpha Q & -I \\ -I & -\beta R \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} \leq 0.$$

Now, as the state never leaves the region $S(\psi^\theta)$, i.e., $x(t) \in S(\psi^\theta)$, we can always express $A(x(t))$ as a convex combination of the A_js as in (23).

By convexity, the above condition is true if it holds, for all $j = 1, \dots, 2^n$,

$$\begin{bmatrix} x \\ w \end{bmatrix}^T \begin{bmatrix} Q(\Psi^{(i)})^T + \Psi^{(i)}Q + \alpha Q & -I \\ -I & -\beta R \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} \leq 0, \tag{28}$$

where $\Psi^{(i)} = [BK^{(i)} - I]$. Using the Shur complement condition (28) is implied by (24).

Based on the above stated theorem we can infer that the solution of a dynamic TU game when (15) is used will lie in the ϵ-core. This is because even if the disturbance in 13 is a q-dimensional unbounded Brownian motion, the dynamics of the process are bounded in the second moment.

Stronger conditions are established in the following corollary.

COROLLARY 6.2. *The distance of any solution of the stochastic differential inclusion (15) from the target set* Π *is second-moment bounded, if there exists a scalar* $\alpha \geq 0$ *such that, for all* $K^{(i)}, i \in I$

$$Q[BK^{(i)} - I]^T + [BK^{(i)} - I]Q + \alpha Q + \frac{1}{\beta} R^{-1} < 0. \tag{29}$$

Proof. Straightforward from observing that (29) implies (24).

Note that conditions (24) simply impose that each one of the conditions (29) (for fixed j) holds only in a specific region of the state space and not over the entire $\mathbb{R}^n$. In this sense, condition (24) is weaker than (29).

Let $d(x, \Pi)$ be the distance of any given $x \in \mathbb{R}^q$ from the target set Π. Consider a modified stochastic differential inclusion

$$dx(t) \in F(x)dt + \Sigma(x)d\mathcal{B}(t), \tag{30}$$

where $\Sigma(x)$ is the weight of the random noise which is now upper bounded by the distance of x from the target set, i.e., $\Sigma(x) \leq d(x, \Pi)$. We are in a position to establish the next result relating to the case where the variance of the stochastic process vanishes the closer the trajectory is to the target set.

[1] P_{ii} is not known a priori so we need to implement a guess method

COROLLARY 6.3. *Let $\Sigma(x) \leq d(x,\Pi)$ and let $\Psi^{(i)} = [BK^{(i)} - I]$. Any solution of the stochastic differential inclusion (30) converges to the target set Π almost surely if for all $x \in X_i$, $i \in I$*

$$x^T\Big[Q(\Psi^{(i)})^T + \Psi^{(i)}Q + \alpha Q + \frac{1}{\beta}R^{-1}\Big]x \leq 0. \tag{31}$$

Proof. The underlying idea is that for all $x \notin \Pi$, and $w \in \mathcal{W}$

$$\begin{array}{l} \lim_{x\to\Pi} \mathcal{L}(V(x)) \\ = \lim_{x\to\Pi} \Big\{[-x + BK(x,t)x - w]^T Px \\ +x^T P[-x + BK(x,t)x - w] + \sum_{i=1}^q \Sigma_{ii}^2(x)P_{ii}\Big\} \\ = x^T[BK(x,t) - I]^T Px + x^T P[BK(x,t) - I]x \\ -w^T Px - x^T Pw < 0. \end{array} \tag{32}$$

We then look for $\alpha, \beta \geq 0$, such that for all $(x,w) \in \overline{\Pi} \times \mathcal{W}$

$$\begin{array}{rl} \mathcal{L}V(x) & = x^T[BK(x,t) - I]^T Px \\ & +x^T P[BK(x,t) - I]x \\ & -w^T Px - x^T Pw \\ & \leq \alpha(1 - V(x)) + \beta(\|w\|_R^2 - 1) \leq 0, \end{array} \tag{33}$$

which is equivalent to setting $\beta \leq \alpha$ and solving

$$\begin{array}{l} \left[\begin{array}{c} x \\ w \end{array}\right]^T \left[\begin{array}{cc} \Psi(x,t)^T P + P\Psi(x,t) + \alpha P & -P \\ -P & -\beta R \end{array}\right] \left[\begin{array}{c} x \\ w \end{array}\right] \\ \quad -\alpha + \beta \leq 0. \end{array}$$

After pre and post-multiplying by $Q = P^{-1}$, and using convexity, the above condition leads to (28), and this concludes the proof.

Let $\underline{\mathcal{B}(t)}$ be a zero-mean random noise such that $\int d\underline{\mathcal{B}(t)}$ has bounded support. For instance, think of $\int d\underline{\mathcal{B}(t)}$ as a truncated Gaussian noise with bounded support in the interval $[-\bar{\kappa}\sigma, \bar{\kappa}\sigma]$ for a positive scalar $\bar{\kappa}$. The counterpart of (15) is then

$$dx(t) \in F(x)dt + \Sigma d\underline{\mathcal{B}(t)}. \tag{34}$$

Assume $\underline{\mathcal{B}(t)} \in [-\Sigma, \Sigma]$ and let $\tilde{W} := \{\omega : \omega = w + \tilde{\sigma}, w \in \mathcal{W}, \tilde{\sigma} \in [-\Sigma, \Sigma]\}$. Also, let $\tilde{R}$ be such that

$$\tilde{W} \subseteq \bar{W} := \{\omega : \|\omega\|_{\tilde{R}}^2 - 1 \leq 0\}.$$

We are in a position to state the following main result.

THEOREM 6.4. *Any solution of the stochastic differential inclusion (15) converges to the target set Π if for all for all $K^{(i)}, i \in I$*

$$\Big[Q(\Psi^{(i)})^T + \Psi^{(i)}Q + \alpha Q + \frac{1}{\beta}\tilde{R}^{-1}\Big] \leq 0. \tag{35}$$

Proof. For all $x \notin \Pi$,

$$\begin{array}{l} \dot{V}(x) \in \Big\{[-x + BK(x,t)x - w \pm \Sigma]^T Px \\ +x^T P[-x + BK(x,t)x - w \pm \Sigma],\, w \in \mathcal{W}\Big\} \\ = \Big\{x^T[BK(x,t) - I]^T Px + x^T P[BK(x,t) - I]x \\ -(w \pm \Sigma)^T Px - x^T P(w \pm \Sigma),\, w \in \mathcal{W}\Big\} < 0. \end{array} \tag{36}$$

Recall that $\tilde{W} := \{\omega : \omega = w + \tilde{\sigma}, w \in \mathcal{W}, \tilde{\sigma} \in [-\Sigma, \Sigma]\}$. From the above we have that for all $x \notin \Pi$ it must hold

$$\begin{aligned}&\dot{V}(x) \leq \max_{\omega \in \tilde{W}} \\ &\Big\{x^T[BK(x,t) - I]^T Px + x^T P[BK(x,t) - I]x \\ &-\omega^T Px - x^T P\omega\Big\} < 0.\end{aligned} \tag{37}$$

For all $x \in \overline{\Pi}$, and $\omega \in \tilde{W}$ the above condition holds if there exist $\alpha, \beta \geq 0$, such that for all $(x, w) \in \overline{\Pi} \times \mathcal{W}$

$$\begin{aligned}\dot{V}(x) \quad &= x^T[BK(x,t) - I]^T Px \\ &+ x^T P[BK(x,t) - I]x \\ &-\omega^T Px - x^T P\omega \\ &\leq \alpha(1 - V(x)) + \beta(\|w\|_R^2 - 1) \leq 0.\end{aligned} \tag{38}$$

From the definition of $\tilde{R}$ it holds

$$\tilde{W} \subseteq \bar{W} := \{\omega : \|\omega\|_{\tilde{R}}^2 - 1 \leq 0\}.$$

For all (x, w) in

$$\overline{\Pi} \times \bar{W} := \{(\xi, \omega) : 1 - V(\xi) \leq 0, \|\omega\|_{\tilde{R}}^2 - 1 \leq 0\},$$

condition (38) can be rewritten as

$$\begin{bmatrix} x \\ \omega \end{bmatrix}^T \begin{bmatrix} Q(\Psi^{(i)})^T + \Psi^{(i)}Q + \alpha Q & -I \\ -I & -\beta\tilde{R} \end{bmatrix} \begin{bmatrix} x \\ \omega \end{bmatrix} \leq 0. \tag{39}$$

and this concludes our proof.

7. Intelligent Mobility Network.

In this section the stability analysis of the case study of the intelligent mobility network of Section 3 is presented.

Initially the deterministic version of dynamics (15) is decomposed as

$$\begin{aligned}dx(t) \in \{&(-x(t) + Bu(t) - \tilde{w}(t))dt \\ &+\Sigma d\mathcal{B}(t),\ \tilde{w}(t) \in \tilde{W}\},\end{aligned} \tag{40}$$

where $\tilde{w}(t)$ is an uncertain but bounded deviation from the expected profit, given by

$$\begin{aligned}\tilde{w}(t) \quad &= [\mathcal{P}_S(y, \Theta_S^*) - \mathbb{E}\mathcal{P}_S(y, \Theta_S^*)]_{S \in \mathcal{S}} \\ &\in W^{(2)} := \{w \in \mathbb{R}^m |\, \underline{\delta} \leq w \leq \overline{\delta}\}.\end{aligned} \tag{41}$$

In the above expression $\overline{\delta}$ and $\underline{\delta}$ are upper and lower bounds respectively, and are obtained as

$$\overline{\delta} := \mathcal{P}_S(\overline{D}_S, \Theta_S^*) - \mathbb{E}\mathcal{P}_S(y, \Theta_S^*), \tag{42}$$

$$\underline{\delta} := \mathcal{P}_S(\underline{D}_S, \Theta_S^*) - \mathbb{E}\mathcal{P}_S(y, \Theta_S^*). \tag{43}$$

Before we calculate $\overline{\delta}^j$ and $\underline{\delta}^j$, note that to derive (40), we simply write the real profit as combination of expected profit $w_0(t)$ and deviation from the expected profit $\tilde{w}(t)$, namely $w(t) = w_0(t) + \tilde{w}(t)$. The expected profit is a priori known and given by $w_0(t) = [\langle \mathcal{P}_S(D_S, \Theta_S^*) \rangle]_{S \in \mathcal{S}}$.We can then design a first control input $u_0(t)$ based on the Shapley allocation to compensate the optimal expected profit. To do this, let $u_0(t)$ be obtained from the following equation:

$$Bu_0(t) = w_0(t) = [\mathbb{E}_S J(y, \Theta_S^*)]_{S\in\mathcal{S}}. \tag{44}$$

To obtain an expression for $\overline{\delta}^j$ let us maximize the profit of the corresponding coalition S with respect to y, namely

$$\begin{array}{l}\overline{D}_S := \arg\max_{D_S} \mathcal{P}_S(D_S, \Theta_S^*) \\ = \arg\max_{D_S}\{p\mu - c\Theta_S^* - s\max(0, \Theta_S^* - D_S) \\ \quad -(p+h)\max(0, D_S - \Theta_S^*)\} = \Theta_S^*.\end{array}$$

Then, the maximal profit for coalition S is

$$\max_y \mathcal{P}_S(y, \Theta_S^*) = \mathcal{P}_S(\overline{D}_S, \Theta_S^*) = \mathcal{P}_S(\Theta_S^*, \Theta_S^*) = p\mu - c\Theta_S^*.$$

Substituting the above in (42), we have

$$\overline{\delta}^j := p\mu - c\Theta_S^* - \langle \mathcal{P}_S(D_S, \Theta_S^*)\rangle.$$

Similarly, to obtain $\underline{\delta}^j$ used in (43), let us minimize the profit of the corresponding coalition S with respect to y, namely

$$\begin{array}{l}\underline{D}_S := \arg\min_{D_S} \mathcal{P}_S(D_S, \Theta_S^*) \\ = \arg\min_{D_S}\{p\mu - c\Theta_S^* - s\max(0, \Theta_S^* - D_S) \\ \quad -(p+h)\max(0, D_S - \Theta_S^*)\} = 0.\end{array}$$

The above means that the minimal profit is obtained when the power output is zero, which leads to

$$\min_y \mathcal{P}_S(y, \Theta_S^*) = \mathcal{P}_S(\underline{D}_S, \Theta_S^*) = \mathcal{P}_S(0, \Theta_S^*) = p\mu - (s+c)\Theta_S^*.$$

Substituting the above in (43), we have

$$\underline{\delta}^j := p\mu - (s+c)\Theta_S^* - \langle \mathcal{P}_S(D_S, \Theta_S^*)\rangle.$$

We can conclude that

$$\begin{array}{l}\tilde{w}(t) \in \tilde{W} := \{w \in \mathbb{R}^m | \\ \quad [p\mu - (s+c)\Theta_S^* - \langle \mathcal{P}_S(D_S, \Theta_S^*)\rangle]_{S\in\mathcal{S}}\} \leq w \\ \quad \leq [p\mu - c\Theta_S^* - \langle \mathcal{P}_S(D_S, \Theta_S^*)\rangle]_{S\in\mathcal{S}}\}.\end{array}$$

As last step we define the parametrized ellipsoid

$$\Pi_k = \{\omega \in \mathbb{R}^m : \, k^2\omega^T\Phi\omega \leq 1\},$$

where Φ is a matrix in $\mathbb{R}^{m\times m}$ and consider the problem of finding the smallest ellipsoid Π_k which contains $\mathcal{W}^{(2)}$:

$$k^* = \max_k\{k | \, \Pi_k \supset \mathcal{W}^{(2)}\}.$$

The dynamic model we obtain is then

$$dx(t) \in \{(-x(t) + Bu(t) - \omega)dt + \Sigma d\mathcal{B}(t), \, \omega \in \Pi_{k^*}\},$$

which is of the same form as in (15).

8. Simulations. An application of the multi-inventory coalitional model, which was described in the previous section, can be found in the electricity trade market. Consider the case of n electricity producers which should meet the electricity demands of a central distributor. The expected profit of a generic coalition is described by (2) under the following two assumptions (Baeyens et al., 2013):

- The structure of the network does not affect the prices and the demand of electricity.
- The electricity market system comprises of a single ex-ante forward penalty and a single ex-post imbalance penalty for variations from the contracted values.

The dynamic demand of such system can be defined as the diffusion process of (20) and the excess is defined as in (21). In the simulations of this section a saturated controller of the form of (18) is used here $K = kB^{-1}$ and $k = \frac{2}{3}$. In our simulations we consider the case of four players/energy producers that should decide if they will be part of a coalition and share the costs and profits from energy production. The initial demand was set to $[0.1693 \quad 0.2019 \quad 0.1304 \quad 0.0562]^T$. The drift parameter w was bounded in $w^T Rw \leq 1$ and R was set to be the identity matrix. Figures 2-4 depict the evolution of the excess, the variance of the excess and the Shapley value respectively.

As it is evident from Figure 2 the excess is always non-negative for all the coalitions which is an indication of a non-empty core. In addition the excess is grouped according to the number of the coalition's members. In particular, the excess for the coalitions with one member have greater excess than the coalitions with two members and the coalitions with two members have greater excess than the coalitions with three members. The grand coalition has excess near to zero.

Figure 3 depicts the variance of the excess of all possible coalitions. As it can be seen from Figure 3 the variances of all coalitions converge to a constant value smaller than one.

Figure 4 depicts the Shapley's value for all players over time. Since the excess value is always positive we can conclude that the core is non-empty.

9. Conclusion. The problem of controlling the allocations in dynamic TU games is considered. Stochastic differential inclusions are used to model the uncertainty of dynamic TU games, which can be occurred either as a result of a dynamic environment or noisy observations. A model is proposed, which extends the results of Bauso et al. (2010) that allows allocation to be controlled by taking into account the deterministic and stochastic uncertainty which exists in the evolution of the excess of a coalition. In particular based on linear matrix inequality conditions it is shown that the stochastic differential inclusion solutions are second-moment bounded. An intelligent mobility scenario is used to show the applicability of the proposed methodology. Additionally simulations in a distribution network are employed which support the theoretical results, by showing stability of the core and bounded variance of the coalitions' excesses.

Future work could include a distributed version of the proposed model. This will increase the efficiency of the proposed methodology's applicability in scenarios which include thousand of players. In addition the performance of the proposed methodology and limitation which may arise from the usage of real distribution network's data in the simulations will be considered.

References.

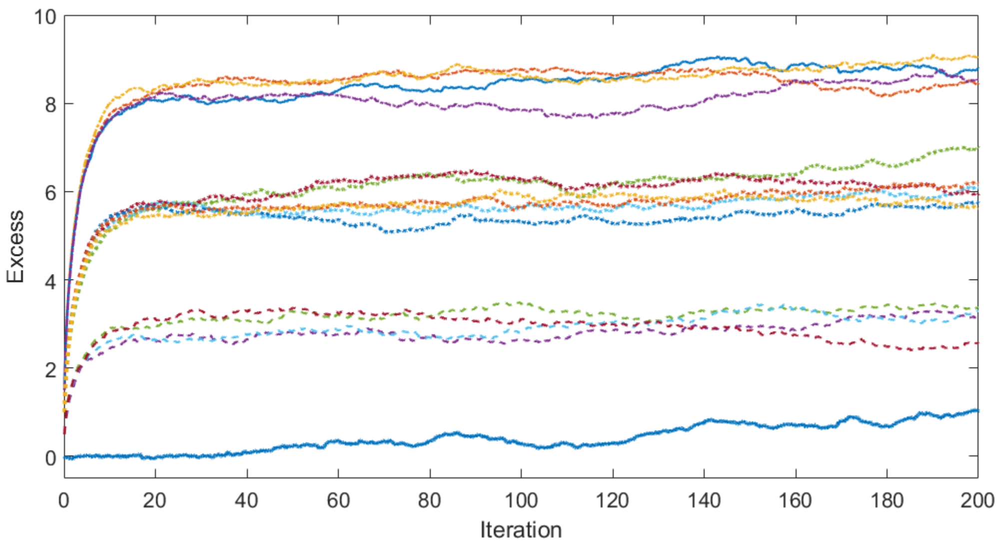


FIG. 2. *Evolution of excess. The combined dotted and dashed lines depict the coalitions with a single member, the dotted lines depict the coalitions with two members, the dashed lines depict the coalitions with three members and the solid line depicts the grand coalition.*

MJ. Osborne. *An introduction to game theory*. New York: Oxford University Press, 2004.

W. Saad, Z. Han, M. Debbah, A. Hjørungnes and T. Başar. Coalitional game theory for communication networks. *Signal Processing Magazine*, IEEE, 26(5): 77–97, 2009.

W. Saad, Z. Han, H. V. Poor and T. Başar. Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications. *Signal Processing Magazine*, IEEE 29(5): 86-105, 2012.

Z. Ramaekers, R. Dasgupta, V. Ufimtsev, S. G. M. Hossain and Carl A. Nelson. Self-Reconfiguration in Modular Robots Using Coalition Games with Uncertainty. *In Automated Action Planning for Autonomous Mobile Robots*, 1462–1468. 2011.

K. Cheng and P. Dasgupta. Coalition game-based distributed coverage of unknown environments by robot swarms. *In Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems* 3: 1191–1194, 2008.

H. Bayram and H. I. Bozma. Coalition formation games for dynamic multi-robot tasks. *The International Journal of Robotics Research*, 35(5): 514–527, 2016.

D. Bauso, L. Giarré and R. Pesenti. Robust control of uncertain multi-inventory systems via Linear Matrix Inequality. *International Journal of Control*, 83(8): 1727–1740, 2010.

S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

J. M. Gomes da Silva, Jr. and S. Tarbouriech. Local Stabilization of Discrete-Time Linear Systems with Saturating Controls: An LMI-based Approach. *IEEE Transactions on Automatic Control*, 46(1): 119–124, 2001.

S. Hart and A. Mas-Colell. Regret-based continuous-time dynamics. *Games and*

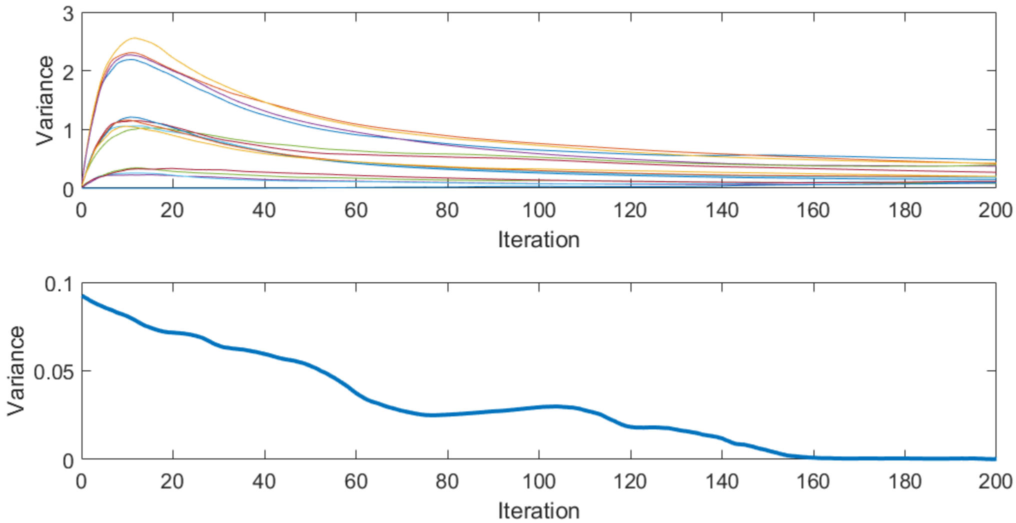


Fig. 3. *Variance of the excess for each coalition. The top plot depicts the variance of all coalitions. The bottom panel depicts the variance of the grand coalition.*

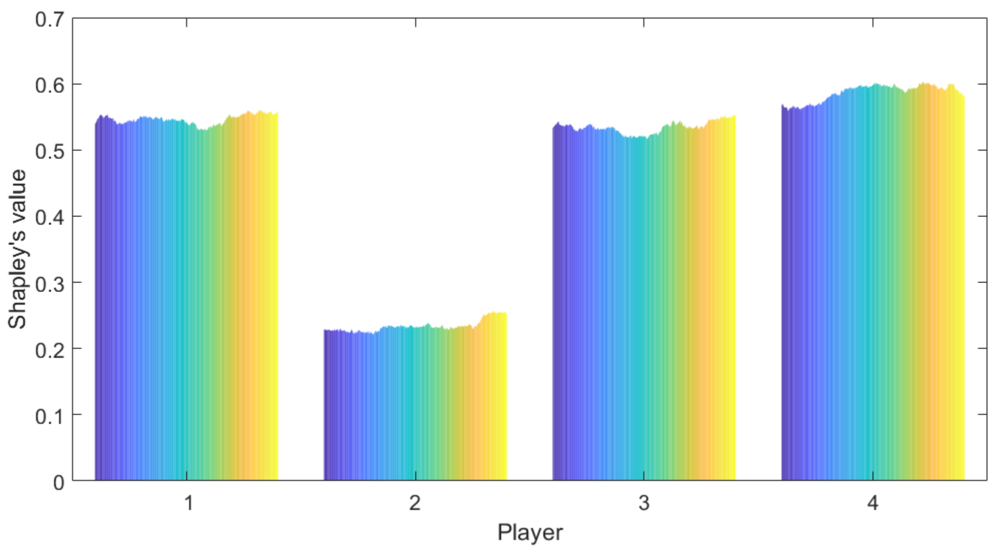


Fig. 4. *Evolution of Shapley's value for the four players.*

Economic Behavior, 45:375–394, 2003.
L. S. Shapley. Cores of convex games. *International Journal of Game Theory*, 1:11–26, 1971.
O.N. Bondareva. Some applications of linear programming methods to the theory of cooperative games. *Problemy Kybernetiki*, 10:119–139, 1963.

L. S. Shapley. On balance sets and cores. *Naval Research Logistics Quarterly*, 14:453–460, 1967.

K. A. Loparo and X. Feng. Stability of stochastic systems. *The Control Handbook*, CRC Press, pp. 1105-1126, 1996.

D. Bauso and J. Timmer. On robustness and dynamics in (un)balanced coalitional games. *Automatica*, 48(10): 2592-2596, 2012.

D. Bauso and J. Timmer. Robust Dynamic Cooperative Games. *International Journal of Game Theory*, 38(1): 23-36, 2009.

E. Baeyens, E.Y. Bitar, P. P. Khargonekar and K.Poolla. Coalitional aggregation of wind power, *IEEE Transactions on Power Systems*, 28(4): 3774-3784, 2013.

L. S. Shapley A value for n-person games, *in Kuhn, H.; Tucker, A.W., Contributions to the Theory of Games II*, Princeton, New Jersey: Princeton University Press, 307-317, 1953.

R.J. Aumann and B. Peleg. Von Neumann - Morgenstern solutions to cooperative games without side payments. *Bul of the Amer Math Society*, 66, 173- 9, 1960.

D. Schmeidler The nucleolus of a characteristic function game, *SIAM Journal of Applied Mathematics*, 17 (6): 1163-1170, 1969.

R.J. Aumann. The core of a cooperative game without side payments. *Transactions of the American Mathematical Society*, 98(3): 539-552, 1961.

R.D. Luce and H. Raiffa. *Games and Decisions: An Introduction and Critical Survey.* Wiley & Sons, 1957.

M. Maschler, B. Peleg and L.S. Shapley, Geometric properties of the kernel, nucleolus, and related solution concepts, *Mathematics of Operations Research*, 4(4): 303-338, 1979.

T. Hu, A.R.Teel and L. Zaccarian. Stability and performance for saturated systems via quadratic and nonquadratic Lyapunov functions. *IEEE Transactions on Automatic Control*, 51(11): 1770-1786, 2006.

X. Cai, L. Liu and W. Zhang. Saturated control design for linear differential inclusions subject to disturbance. *Nonlinear Dynamics*, 58(3): 487-496, 2009.

T. Hu, A.R. Teel and L. Zaccarian. Performance analysis of saturated systems via two forms of differential inclusions. *In 44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05*, 8100-8105, 2005.

A. Jokic, M. Lazar, and P.P.J Van den Bosch. Complementarity systems in constrained steady-state optimal control. *International Workshop on Hybrid Systems: Computation and Control*. Springer, Berlin, Heidelberg, 2008.

S. Grammatico, F. Blanchini and A. Caiti. Control-sharing and merging control Lyapunov functions. *IEEE Transactions on Automatic Control*. 59(1): 107-119, 2014.

L. Shapley and M. Shubik. Quasi-cores in a monetary economy with nonconvex preferences. *Econometrica: Journal of the Econometric Society*. 805–827, 1966.

J. Suijs, P. Borm, A. De Waegenaere and S. Tijs. Cooperative games with stochastic payoffs. *European Journal of Operational Research*. 113(1), 193–205, 1997.

Dynamic linear programming games with risk-averse players. *Mathematical Programming*. 163(1), 25–56, 2017.

M. Benaïm, J. Hofbauer and S. Sorin. Stochastic approximations and differential inclusions. *SIAM Journal on Control and Optimization*. 44(1),328-48, 2005.

L.S. Shapley and M. Shubik On market games. *Journal of Economic Theory*. 1(1), 9-25, 1969.

G. Bodwin Testing Core Membership in Public Goods Economies. *arXiv preprint arXiv:1705.01570*. 2017.

R.J. Aumann and M. Maschler Game theoretic analysis of a bankruptcy problem from the Talmud. *Journal of Economic Theory.* 36(2), 195-213, 1985.

A. Müller, M. Scarsini and M. Shaked. The newsvendor game has a nonempty core. *Games and Economic Behavior.* 38(1), 118-26, 2002.

B.C. Hartman and M. Dror. Allocation of gains from inventory centralization in newsvendor environments. *IIE Transactions.* 37(2),93-107, 2005.

M. Slikker, J. Fransoo and M. Wouters. Cooperation between multiple news-vendors with transshipments. *European Journal of Operational Research.* 167(2), 370-80,2005.

A. Nedich and D. Bauso. Dynamic Coalitional TU Games: Distributed Bargaining among Players' Neighbors. *IEEE Trans on Automatic Control.* 58(6), 1362–1376, 2013.

A. Chinchuluun, A. Karakitsiou and A. Mavrommati. Game theory models and their applications in inventory management and supply chain. *Pareto Optimality, Game Theory And Equilibria.* 833-865, 2008.

T. Wada and Y. Fujisaki. A stochastic approximation for finding an element of the core of uncertain cooperative games. *11th Asian Control Conference (ASCC).* 2071-2076, 2017.

D. Blackwell. Pacific Journal of Mathematics, A Non-profit Corporation. *Pacific J. Math.* 6(1) 1–8,1956.

F. Fele and J. M. Maestre and E. F. Camacho. Coalitional Control: Cooperative Game Theory and Control. *IEEE Control Systems Magazine.* 37(1), 53-69, 2017

E. Lehrer. Allocation processes in cooperative games. *International Journal of Games Theory.* 31,341-351, 2003.

D. Bauso Adaptation, coordination, and local interactions via distributed approachability *Automatica.* 84, 48-55, 2017.

I. Garud. Robust Dynamic Programming. *Mathematics of Operations Research.* 30(2), 257-280, 2005.

D. Bauso and H. Tembine and T. Basar. Robust Mean Field Games *Dynamic Games and Applications.* 6(06), 2015.

C. Opathella and B. Venkatesh. Managing Uncertainty of Wind Energy With Wind Generators Cooperative. *IEEE Transactions on Power Systems.* 28(08), 2918-2928, 2013.

E. Baeyens and Y. E. Bitar and P. Khargonekar, and K. Poolla. Coalitional Aggregation of Wind Power. *IEEE Transactions on Power Systems.* 28, 3774-3784, 2013.

W. Saad and H. Zhu Han and H. V. Poor. Coalitional game theory for cooperative micro-grid distribution networks. *IEEE International Conference on Communications*, 2013.

D. Bauso and L. GiarrÉ and R. Pesenti. Consensus in Noncooperative Dynamic Games: A Multiretailer Inventory Application. *IEEE Transactions on Automatic Control.* 53(4), 998-1003, 2008.

D.Q. Mayne. Constrained Control: Polytopic Techniques. In: Gong W., Shi L. (eds) Modeling, Control and Optimization of Complex Systems. *The International Series on Discrete Event Dynamic Systems.* 14,2003.