This is a repository copy of *Adaptation, coordination, and local interactions via distributed approachability*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/117654/

Version: Accepted Version

Adaptation, coordination, and local interactions via distributed approachability

Dario Bauso [a,b]

[a] *Department of Automatic Control and Systems Engineering, University of Sheffield, Mappin Street Sheffield, S1 3JD, UK*

[b] *Dipartimento di Ingegneria Chimica, Gestionale, Informatica, Meccanica, Università di Palermo, 90128 Palermo, Italy*

Abstract

This paper investigates the relation between cooperation, competition, and local interactions in large distributed multi-agent systems. The main contribution is the game-theoretic problem formulation and solution approach based on the new framework of distributed approachability, and the study of the convergence properties of the resulting game model. Approachability theory is the theory of two-player repeated games with vector payoffs, and distributed approachability is here presented for the first time as an extension to the case where we have a team of agents cooperating against a team of adversaries under local information and interaction structure. The game model turns into a nonlinear differential inclusion, which after a proper design of the control and disturbance policies, presents a consensus term and an exogenous adversarial input. Local interactions enter in the model through a graph topology and the corresponding graph-Laplacian matrix. Given the above model, we turn the original questions on cooperation, competition, and local interactions, into convergence properties of the differential inclusion. In particular, we prove convergence and exponential convergence conditions around zero under general Markovian strategies. We illustrate our results in the case of decentralized organizations with multiple decision-makers.

Key words: repeated games; approachability; differential games; robust control; network flow.

1 Introduction

Cooperation, competition, and local interactions are three main co-existing elements in large distributed multi-agent systems with humans in the loop, see Fig. 1. The *state* of a decision-maker is captured by a time-varying abstract entity, which contains aggregate information on his past decisions and those of a subset of other decision-makers around him, as well as his cumulative or average payoff.

In abstract terms, *cooperation* refers to the capability of the decision-makers to make decisions to coordinate their states. The decision-makers try to reach consensus by exhibiting reciprocal attraction forces which may lead them to converge to a consensus equilibrium, see [15] and references therein.

By *competition* we refer to the capabilities of the decision-makers to let the collective state, a vector which involves the states of all the decision-makers, converge to a preassigned set or equilibrium point despite the

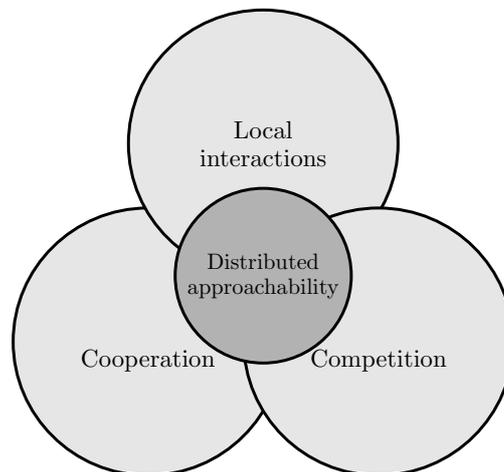*Email address:* `d.bauso@sheffield.ac.uk` (Dario Bauso).

Fig. 1. Three dimensions of distributed decision making reframed within distributed approachability.

presence of disturbances. A natural way to deal with such a scenario is via *approachability* theory, whose traditional formulation involves only two players, the decision-maker (player 1 or row player) and the adver-

sarial disturbance (player 2 or column player) [8]. The two players play repeatedly over time in a continuous- or discrete-time setting, and the outcome of the game at any time is a vector payoff. Both players try to influence the evolution of the average payoff. Existing results show that the approachability problem can be turned into a differential game in which the average payoff appears as the (collective) state of the game [6,7,16]. In particular player 1 plays to make the average payoff converge to a preassigned set, while player 2 tries to contrast him. Equivalence of Blackwell Approachability and No-Regret Learning is studied in [1]. A dynamic programming approach to calculate approachable sets is presented in [14]. Approachability in Stackelberg Stochastic Games is investigated in [13]. Convergence of the cumulative payoff rather than the average implies some variations of the conditions which are formalized in the context of *attainability* in two-player reapeated games with vector payoffs, see e.g. [4] and [3, Ch.11]. The distributed approachability problem that we formulate here assumes that player 1 is indeed made by a team of agents whose cooperation results in attraction forces against a team of adversarial disturbances, referred to as player 2, which exhibit external forces.

Local Coordination captures the idea that the decision-makers have i) local information, namely they know only some state components, and ii) local influence, namely their decisions influence only some state components. To model local coordination we refer to the concept of distributed Markovian or state-feedback control policies. Back to the approachability interpretation, the state of the decision-maker is the subset of payoff components he can monitor and control. As it will be clear later on, the term *distributed approachability* is here used to address such a concept. This term has already appeared in [2] in the context of coalitional games.

Contribution. As main contribution this paper builds a mathematical model involving each of the above dimensions: cooperation, competition, and local interactions. To capture competition the model takes the form of a distributed approachability problem, thus departing in an original way from the traditional two-player approachability formulation. A further contribution is in that the model links in an original way to a stylized model in the literature of decentralized organizations thus contributing to the cross-fertilization of engineering and social science.

Building on existing results [6,7,16], which show that an approachability problem can be turned into a differential game, the game is ultimately transformed into a nonlinear differential inclusion describing the continuous-time evolution of the cumulative or average payoff. Here, the distributed control involves the mixed actions of all the decision-makers (player 1) and the distributed disturbance is the mixed actions of all adversaries (player 2).

The decision-makers coordinate to drive the vector payoff to a preassigned set against the action of the adversaries. Nonlinearity is due to bounds on controls and disturbances. Given such a system, we look at equilibrium points, which represent conditions under which the attraction forces counterbalance the external ones.

We show that cooperation results in a consensus term in the differential inclusion which describes the attraction forces. Under such forces the states of the decision-makers tend to get closer one to each other.

Competition takes the form of an exogenous signal. In other words, the adversary tries to attract the local states by exhibiting some centrifugal force.

Local interaction enters in the model through a graph topology. We study the influence of such topology both on the stationary solution and on the transient dynamics. The graph topology appears in the consensus term, through the graph-Laplacian matrix.

Given the above model, we can turn the original questions on cooperation, competition, and local interactions, into convergence properties of the differential inclusion. In particular, we prove convergence and exponential convergence conditions around zero under general Markovian strategies using approachability theorem by Blackwell. We observe that when we use distributed Markovian strategies, we obtain a robust consensus dynamics and for such a dynamics we study the corresponding convergence properties.

The main assumption is in the form of set inclusion, and represents properties of the action sets of the game. This assumption is borrowed from the literature on robust control of network systems [9,10].

To place the contribution of this paper in proper context, we illustrate our results in the case of decentralized organizations with multiple decision-makers that must perform n specialized tasks [11]. The decision-makers, each one associated to a single task, choose the levels of adaptation and coordination. A higher level of adaptation implies that the workers show higher flexibility to adapt their tasks. A higher level of coordination entails an increase in the communication between workers. The performance of the organization depends on: i) how well each task is adapted to specific market conditions, operational conditions, and consumers' needs and ii) how well all tasks are coordinated with each other.

This paper is organized as follows. In Section 2 we introduce approachability and distributed approachability. In Section 3 we turn the game into a dynamical system. In Section 4 we provide the main results on convergence and exponential convergence. In Section 5 we discuss the results in the context of decentralized organizations. In Section 6 we provide conclusions.

2 Distributed approachability

In this section, we first introduce the traditional approachability setting involving two players and a continuous-time repeated game with vector payoffs. Then, we formulate the problem at hand in the form of a distributed approachability problem with a team of decision-makers playing against a team of adversaries.

2.1 Approachability

The traditional approachability setting involves a two-player repeated game with vector payoffs, which we refer to as Γ. The set of players is $N = \{1, 2\}$, and the finite set of actions of each player i is A_i. The instantaneous payoff is given by a biaffine function $g : A_1 \times A_2 \to \mathbb{R}^m$, where m is a natural number.

We extend g to the set of mixed actions pairs, $\Delta(A_1) \times \Delta(A_2)$, in a bilinear fashion. The one-shot vector-payoff game $(\Delta(A_1), \Delta(A_2), u)$ has compact convex action sets and is denoted by G.

The game Γ is played in continuous-time over the time interval $[0, \infty)$. We assume that the players use non-anticipative behavior strategies, according to the definition provided below.

Denote by $\mathcal{C}_i$ the set of all *actions* of player i, that is, the set of all measurable functions from the time space, $[0, \infty)$, to player i's mixed actions. That is,

$$\mathcal{C}_i := \{a_i : [0, \infty) \to \Delta(A_i), \ a_i \text{ is measurable}\}.$$

Definition 2.1 *A function $\sigma_i = \sigma_i[\cdot] : \mathcal{C}_{-i} \to \mathcal{C}_i$ is a* non-anticipative behavior strategy *for player i, if*

$$a_{-i}(s) = a'_{-i}(s) \quad \forall s \in [0, t]$$
$$\implies \sigma_i[a_{-i}](s) = \sigma_i[a'_{-i}](s) \quad \forall s \in [0, t].$$

Every pair of strategies $\sigma = (\sigma_1, \sigma_2)$ uniquely determines a play path $(a[\sigma](t))_{t \in \mathbb{R}_+}$. The payoff (vector) up to time t associated with the pair of strategies σ is given by

$$x[\sigma](t) = \int_0^t g\left(a[\sigma](s)\right) \mathrm{d}s \in \mathbb{R}^m. \tag{1}$$

The integral in (1) is the *cumulative payoff* up to time t.

We also define the *average payoff* up to time t as

$$\bar{x}[\sigma](t) = \frac{1}{t} \int_0^t g\left(a[\sigma](s)\right) \mathrm{d}s \in \mathbb{R}^m. \tag{2}$$

Let us depart from the traditional setting by introducing the distributed element in our problem. To do this, let the set of actions be given by

$$A_1 = \{a_1^{(1)}, \dots, a_1^{(v)}\}, \quad A_2 = \{a_2^{(1)}, \dots, a_2^{(r)}\},$$

where $a_1^{(i)} \in \mathbb{R}^p$ for all $i = 1, \dots, v$ are the vertices of a hyperbox denoted by $\mathcal{U}$ in $\mathbb{R}^p$. Likewise, $a_2^{(j)} \in \mathbb{R}^q$ for all $j = 1, \dots, r$ are the the vertices of a hyperbox in $\mathbb{R}^q$. Thus, $\Delta(A_1) \subset \mathbb{R}^p$ and $\Delta(A_2) \subset \mathbb{R}^q$. The two-player

$a_1^{(i)}/a_2^{(j)}$	$a_2^{(1)}$	...	$a_2^{(r)}$
$a_1^{(1)}$	$Ba_1^{(1)} - Da_2^{(1)}$	...	$Ba_1^{(1)} - Da_2^{(r)}$
$\vdots$	$\vdots$		$\vdots$
$a_1^{(v)}$			

Table 1
Two-player game with vector payoffs: $\mathcal{A} = [\mathcal{A}_{ij}]$.

game is characterized by the following payoff matrix, for all $i = 1, \dots v$, $j = 1, \dots, r$

$$\mathcal{A} = [\mathcal{A}_{ij}], \quad \mathcal{A}_{ij} = Ba_1^{(i)} - Da_2^{(j)},$$

where $B \in \mathbb{R}^{m \times p}$ and $D \in \mathbb{R}^{m \times q}$ are given matrices.

Table 1 displays the two-player game and the matrix $\mathcal{A}$ with multi-dimensional entries $\mathcal{A}_{ij}$ in $\mathbb{R}^m$. In a centralized setup, at any time t, players 1 and 2 pick vertices of the hyperboxes $\mathcal{U} \subset \mathbb{R}^p$ and $\mathcal{W} \subset \mathbb{R}^q$. In the distributed setup we consider here, the action of player 1 is the path u (see it as a path or as a vector) resulting from different agents selecting simultaneously orthogonal segments $u_1, \dots, u_p$ in the hyperbox $\mathcal{U}$.
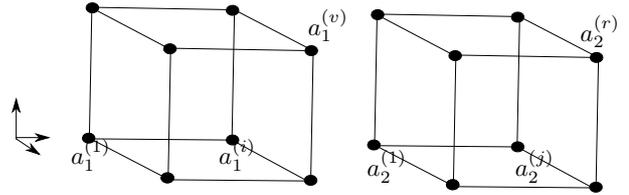

Fig. 2. Sets of actions $A_1 = \{a_1^{(1)}, \dots, a_1^{(v)}\}$ and $A_2 = \{a_2^{(1)}, \dots, a_2^{(r)}\}$. In a centralized setup, at any time t, players 1 and 2 pick vertices of hyperboxes $\mathcal{U} \subset \mathbb{R}^p$ and $\mathcal{W} \subset \mathbb{R}^q$, respectively.

Denote $a_1 = [a_{11}, \dots, a_{1v}]^T$ and $a_2 = [a_{21}, \dots, a_{2r}]^T$. Introduce the mapping $\Delta(A_1) \times \Delta(A_2) \to \mathcal{U} \times \mathcal{W}$, such that $(a_1, a_2) \mapsto (u, w)$ where

$$u = \sum_{i=1}^{v} a_{1i} a_1^{(i)} = [u_1, \dots, u_p]^T,$$
$$w = \sum_{j=1}^{r} a_{2j} a_2^{(j)} = [w_1, \dots, w_q]^T.$$

The instantaneous payoff at time s is given by

$$
\begin{aligned}
g\left(a[\sigma](s)\right) &= \sum_{i=1}^{v} \sum_{j=1}^{r} a_{1i} a_{2j} (B a_1^{(i)} - D a_2^{(j)}) \\
&= \sum_{i=1}^{v} a_{1i} (B a_1^{(i)}) - \sum_{j=1}^{r} a_{2j} (D a_2^{(j)}) \\
&= B \left(\sum_{i=1}^{v} a_{1i} a_1^{(i)} \right) - D \left(\sum_{j=1}^{r} a_{2j} a_2^{(j)} \right) \\
&= Bu - Dw \doteq \hat{g}(u, w).
\end{aligned} \tag{3}
$$

Assume that player 1 involves p distinct agents each one controlling one component of u. In other words, agent i controls u_i, which in turn has effect only on $\hat{g}(\cdot)_j$ and $\hat{g}(\cdot)_k$, these being the jth and kth component of the vector-valued function $\hat{g}(\cdot)$. In addition, agent i knows only $x_j(t)$ and $x_k(t)$ for any pair $j, k = 1, \ldots, m$ at time t. Therefore we set $u_i = f(x_j, x_k)$, where $f(\cdot)$ is a generic function which needs to be designed.

Let a graph $G = (V, E)$ be given, where V is the set of vertices and E is the set of edges. The interaction between the control u_i and the states $x_j(t)$ and $x_k(t)$ is illustrated in Fig. 3. Matrix B is the incidence matrix of
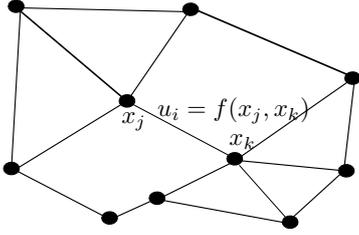

Fig. 3. Graph $G = (V, E)$ illustrating the distributed nature of the problem. Component u_i is function of only $x_j(t)$ and $x_k(t)$, and influences only $\hat{g}(\cdot)_j$ and $\hat{g}(\cdot)_k$.

the above graph.

We can rewrite the cumulative payoff as

$$
\begin{aligned}
x[\sigma](t) &= \int_0^t g\left(a[\sigma](s)\right) \mathrm{d}s \\
&= \int_0^t \hat{g}\left(u(s), w(s)\right) \mathrm{d}s \\
&= \int_0^t (Bu(s) - Dw(s)) \mathrm{d}s \in \mathbb{R}^m.
\end{aligned} \tag{4}
$$

Likewise, the average payoff up to time t is

$$
\begin{aligned}
\bar{x}[\sigma](t) &= \frac{1}{t} \int_0^t g\left(a[\sigma](s)\right) \mathrm{d}s \\
&= \frac{1}{t} \int_0^t \hat{g}\left(u(s), w(s)\right) \mathrm{d}s \\
&= \frac{1}{t} \int_0^t (Bu(s) - Dw(s)) \mathrm{d}s \in \mathbb{R}^m.
\end{aligned} \tag{5}
$$

Both the cumulative or average payoff represent the *collective state* of our system.

3   Uncertain dynamical system

In this section, we build on existing results to turn the repeated game into an uncertain dynamical system or differential game if we review the control as one player and the disturbance as the opponent. Let us consider the following state-feedback control $u(t) = \phi(z(t))$. Let us rescale the time window using $t = e^\tau$ and take $z(\tau) = \bar{x}(e^\tau)$ and differentiate the above expression of $z(\tau)$ with respect to τ. Then, for fixed strategy $u(\tau) = \phi(z(\tau))$, the dynamics is a differential inclusion of type:

$$
\begin{aligned}
\dot{z}(\tau) \in F(z) &:= \{\xi \in \mathbb{R}^m | \\
&\quad \xi = \hat{g}(u(z), w) - z, \forall w \in \mathcal{W}\}.
\end{aligned} \tag{6}
$$

Note that after rescaling the time window, we have

$$
z(0) = \bar{x}(1) = \int_0^1 \hat{g}(u(z(s), s), w(s)) \mathrm{d}s \in \mathbb{R}^m.
$$

Given a compact set $\Lambda \in \mathbb{R}^m$ and $z \in \mathbb{R}^m$ we let $\Pi_\Lambda(z) = \{y \in \Lambda | \ dist^2(z, \Lambda) = \|z - y\|^2 = \langle z - y, z - y \rangle\}$.

Theorem 3.1 (Approachability) *Let $\Lambda \in \mathbb{R}^m$ be a compact set, $r > 0$ and $Z = \{z \in \mathbb{R}^m : dist(z, \Lambda) < r\}$. If for all $z \in Z \setminus \Lambda$ there exists $y \in \Pi_\Lambda(z)$ such that*

$$
\langle z + v - y, z - y \rangle \leq 0, \quad \forall v \in F(z), \tag{7}
$$

then the set Λ is approachable.

Proof. Let $z(t)$, $t \in [0, T]$ be solution of (6) and let $\delta(t) = \|z(t) - y\|^2$. Let $f(\hat{g}(u, w), z) = -z + Bu - Dw$. We have

$$
\begin{aligned}
\dot{\delta}(\tau) &= 2\langle f(p(u, w), x), z(\tau) - y \rangle \\
&= 2\langle \hat{g}(u, w) - z, z(\tau) - y(\tau) \rangle \\
&= 2\langle \hat{g}(u, w) - z(\tau) + z(\tau) - y(\tau), z(\tau) - y(\tau) \rangle \\
&\quad - 2\langle z(\tau) - y(\tau), z(\tau) - y(\tau) \rangle.
\end{aligned} \tag{8}
$$

From (7) we have that

$$
\langle \hat{g}(u, w) - z(\tau) + z(\tau) - y(\tau), z(\tau) - y(\tau) \rangle < 0, \forall t \in (0, T],
$$

which implies

$$
\dot{\delta}(t) \leq -2\langle z(\tau) - y(\tau), z(\tau) - y(\tau) \rangle = -2\delta(t).
$$

By integration of the above inequality, one obtains that

$$
\|z(\tau) - y(\tau)\| \leq \|z(0) - z(\tau)\| e^{-\tau},
$$

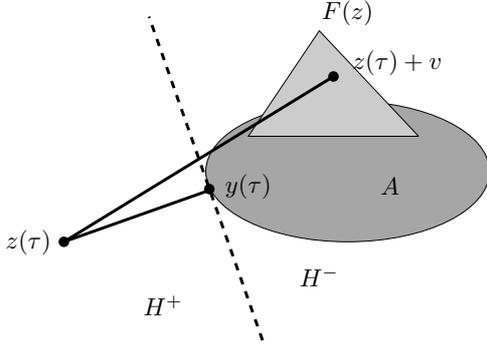and therefore Λ is approachable with exponential rate. ∎

Fig. 4. Geometric illustration of the *Blackwell's Approachability Principle*.

A graphical illustration of condition (7) is as in Fig. 4. There, we have set A, which is the set that the controller wishes to approach. Let $y(t)$ be the projection of $x(t)$ onto set A. The *supporting hyperplane* (dashed line) to A at point $y(\tau)$ is the set of points satisfying

$$H = \{\zeta \in \mathbb{R}^m | \langle \zeta - y(\tau), z(\tau) - y(\tau) \rangle = 0\}.$$

Given the supporting hyperplane H, let us denote by H^+ and H^- the positive and negative half-spaces. That is to say, that for H^+ and H^- it holds

$$H^+ = \{\zeta \in \mathbb{R}^m | \langle \zeta - y(\tau), z(\tau) - y(\tau) \rangle \geq 0\},$$
$$H^- = \{\zeta \in \mathbb{R}^m | \langle \zeta - y(\tau), z(\tau) - y(\tau) \rangle \leq 0\}.$$

Condition (7) essentially states that point $z(\tau) + v$ must lie in the opposite halfspace than the one containing $z(\tau)$, for any $v \in F(z)$.

4 Distributed convergence

This section contains the main results of this paper. We first study approachability of sets of equilibrium points in the case where both u and w are state-feedback strategies. Then we investigate approachability of regions around zero under the worst-case realization of w.

4.1 Approachability of equilibrium sets

We wish to study approachability of equilibrium sets under the assumption that both u and w are obtained from state-feedback strategies. In other words, approachable sets are of the form $\Lambda := \{z | \dot{z} = 0\}$ for given state-feedback control $u(t) = \phi(z(t))$ and disturbance $w(t) = \hat{\phi}(z(t))$.

Remark 4.1 *All the results in this section hold in the special case where $w(t) = \hat{\phi}(z(t)) = \omega \in \mathcal{W}$ where ω is any constant vector.*

For each equilibrium point z in Λ we have that the set $F(z)$ coincides with the zero point,

$$
\begin{aligned}
\dot{z} \in F(z) &= \hat{g}(u, w) - z = \hat{g}(\phi(z), \hat{\phi}(z)) - z \\
&= B\phi(z) - D\hat{\phi}(z) - z = 0.
\end{aligned}
\tag{9}
$$

We can view the term $B\phi(z)$ as the internal (attraction) force, and $D\hat{\phi}(z)$ as the external force. At the equilibrium z is balancing both the internal and the external force. This value for z represents a compromise between internal coordination and adaptation to external conditions.

In the following, we consider equilibrium sets that are convex, namely, the set

$$\Lambda := \{z | B\phi(z) - D\hat{\phi}(z) - z = 0\}$$

is such that given ζ_1 and ζ_2 in Λ, we have that $\bar{\zeta} = \theta\zeta_1 + (1 - \theta)\zeta_2$ for any $0 < \theta < 1$ is also in Λ. This corresponds to saying that

$$
\left.
\begin{aligned}
B\phi(\zeta_1) - D\hat{\phi}(\zeta_1) - \zeta_1 = 0 \\
B\phi(\zeta_2) - D\hat{\phi}(\zeta_2) - \zeta_2 = 0
\end{aligned}
\right\} \\
\Rightarrow \quad B\phi(\bar{\zeta}) - D\hat{\phi}(\bar{\zeta}) - \bar{\zeta} = 0.
\tag{10}
$$

The following theorem restates the approachability conditions in the case of state-feedback strategies.

Theorem 4.1 (Approach. feedback strategies)
Let $\Lambda \in \mathbb{R}^m$ be a compact equilibrium set, i.e., $\Lambda := \{z | \dot{z} = 0\}$ for given state-feedback control $u(t) = \phi(z(t))$ and disturbance $w(t) = \hat{\phi}(z(t))$. Let $r > 0$ and $Z = \{z \in \mathbb{R}^m : dist(z, \Lambda) < r\}$. Assume that for all $z \in Z \setminus \Lambda$ there exists $y \in \Pi_\Lambda(z)$ such that

$$\langle v, z - y \rangle \leq 0, \quad \forall v \in F(z). \tag{11}$$

Then the set Λ is approachable.

Proof. Let $z(t)$, $t \in [0, T]$ be solution of (6). Also, let $\delta(t) = \|z(t) - y\|^2$. Let $f(\hat{g}(u, w), z) = -z + Bu - Dw$. We have

$$
\begin{aligned}
\dot{\delta}(\tau) &= 2\langle f(p(u, w), x), z(\tau) - y \rangle \\
&= 2\langle \hat{g}(u, w) - z, z(\tau) - y(\tau) \rangle \leq 0.
\end{aligned}
\tag{12}
$$

From (11) we have that

$$\langle \hat{g}(u, w) - z(\tau), z(\tau) - y(\tau) \rangle < 0, \forall\, t \in (0, T],$$

which implies

$$\dot{\delta}(t) \leq 0.$$

5

We know that the set of points z for which $\dot{\delta}(t) = 0$ is the set of equilibrium points in Λ. Actually, any point z in Λ is such that

$$\dot{z} = \hat{g}(u, w) - z = \hat{g}(\phi(z), \hat{\phi}(z)) - z \tag{13}$$
$$= B\phi(z) - D\hat{\phi}(z) - z = 0.$$

Then, from LaSalle's Invariance Principle we know that any trajectory originating in Z converges to the largest invariant set in Λ which is Λ itself. ■

Condition (11) is more general than (7) but does not guarantee exponential convergence as for (7). This is established in the next theorem.

Theorem 4.2 *Condition (7) implies (11) but not vice versa.*

Proof. ($\Rightarrow$) Let us first prove that (7) implies (11). Assume that it holds

$$\langle v + z - y, z - y \rangle \leq 0, \quad \forall v \in F(z). \tag{14}$$

Then we have

$$\langle v, z - y \rangle + \langle z - y, z - y \rangle \leq 0, \quad \forall v \in F(z), \tag{15}$$

which in turn implies

$$\langle v, z - y \rangle \leq -\langle z - y, z - y \rangle \leq 0, \quad \forall v \in F(z).$$

($\Leftarrow$ not true) To show that (11) does not imply (7), consider a big enough scalar $\kappa > 0$. Assume that

$$-\kappa \leq \langle v, z - y \rangle \leq 0, \quad \forall v \in F(z).$$

Then, for any z such that $\langle z - y, z - y \rangle > \kappa$ we have

$$\langle v + z - y, z - y \rangle = \langle v, z - y \rangle + \langle z - y, z - y \rangle \tag{16}$$
$$\geq \langle v, z - y \rangle + \kappa \geq 0, \quad \forall v \in F(z).$$

This concludes the proof. ■

In the following, we show that Theorem 4.5 is useful to study conditions for distributed approachability. A state-feedback control strategy for which the set of equilibrium points Λ is compact, is the linear saturated control [5]:

$$u(\tau) = sat \left[-\frac{B^T}{\gamma} z(\tau) \right], \quad \gamma > 0, \tag{17}$$

where the saturation function $sat[.] : \mathbb{R}^p \to \mathbb{R}^p$ is defined componentwise as follows

$$u_i = sat[\xi_i] \doteq \begin{cases} u_i^- & if \ \xi_i < u_i^-, \\ \xi_i & if \ u_i^- \leq \xi_i \leq u_i^+, \\ u_i^+ & if \ \xi_i > u_i^+. \end{cases}$$

In addition, consider the set of equilibrium points

$$\Lambda_\gamma \doteq \left\{ z \mid Bsat \left[-\frac{B^T}{\gamma} z \right] = D\hat{\phi}(z) + z \right\}.$$

The following assumption establishes a set inclusion condition involving the bounding sets of state, control, and disturbance. Such assumption is relevant to approachability of equilibrium points as established next.

Assumption 1 *Matrix $B \in \mathbb{R}^{m \times p}$ is full row rank and set $\mathcal{W}$ is in the interior of $B\mathcal{U}$, that is,*

$$D\mathcal{W} + \Lambda_\gamma \subset int\{B\mathcal{U}\}. \tag{18}$$

Theorem 4.3 *Under Assumption 1, the control (17), with arbitrary $\gamma > 0$ is such that $z(\tau) \to \Lambda_\gamma$.*

Proof. Denote by $y = \Pi_{\Lambda_\gamma}(x)$, where $\Pi_{\Lambda_\gamma}(x)$ is the projection of x onto set Λ_γ. Let us denote $\xi = -B^T z/\gamma$ and $\bar{\xi} = -B^T y/\gamma$. Condition (11) becomes

$$\langle v, z - y \rangle = \left\langle Bsat \left[-\frac{B^T}{\gamma} z \right] - Dw - z, z - y \right\rangle$$
$$= \gamma \cdot \left\langle \frac{B}{\gamma} \left(sat \left[-\frac{B^T}{\gamma} z \right] - sat \left[-\frac{B^T}{\gamma} y \right] \right), z - y \right\rangle$$
$$= \left\langle sat \left[-\frac{B^T}{\gamma} z \right] - sat \left[-\frac{B^T}{\gamma} y \right], \gamma \frac{B^T}{\gamma} (z - y) \right\rangle \tag{19}$$
$$= -\gamma \langle sat[\xi] - sat[\bar{\xi}], \xi - \bar{\xi} \rangle$$
$$= -\gamma \sum_{i=1}^{m} (\xi_i - \bar{\xi}_i) \left(sat[\xi_i] - sat[\bar{\xi}_i] \right) \leq 0.$$

The last inequality derives from $\bar{\xi}_i$ being in the interior of interval $[u_i^-, u_i^+]$, which in turn derives from Assumption 1. ■

Under control (17), dynamics (6) becomes

$$\dot{z}(\tau) \in F(z) := \{\xi \in \mathbb{R}^m \mid$$
$$\xi = Bsat \left[-\frac{B^T}{\gamma} z \right] - D\hat{\phi}(z) - z \}. \tag{20}$$

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

$$\dot{z}(\tau) \in F(z) := \{\xi \in \mathbb{R}^m \mid$$
$$\xi = L(z)z - D\hat{\phi}(z) - z \}, \tag{21}$$

where the time varying matrices $L(z(t))$ are expressed as convex combinations of 2^p matrices L_j, $j = 1, \ldots, 2^p$. More precisely the expressions for $L(z(t))$ are

$$L(z(t)) = \sum_{j=1}^{2^p} \sigma_j(z(t)) L_j, \quad \sum_{j=1}^{2^p} \sigma_j(t) = 1. \quad (22)$$

The procedure to compute matrices L_j's is borrowed from [12] and recalled below. For the control, let us rewrite

$$u_i = \theta_i(z)(-K_{i\bullet}z),$$

where $\theta_i(z)$ are the "degree of saturation" given by

$$\theta_i(z) = \begin{cases} \frac{u_i^-}{-K_{i\bullet}z} & \text{if } -K_{i\bullet}z < u_i^-, \\ 1 & \text{if } u_i^- \leq -K_{i\bullet}z \leq u_i^+, \\ \frac{u_i^+}{-K_{i\bullet}z} & \text{if } -K_{i\bullet}z > u^+. \end{cases} \quad (23)$$

Let $\underline{\theta} = [\underline{\theta}_1, \ldots, \underline{\theta}_p]$ be a vector whose components $\underline{\theta}_i$ are such that $0 \leq \underline{\theta}_i \leq 1$ and represent lower bounds of $\theta_i(z(t))$, for $t \geq 0$. Also let $\psi^\theta = [\psi_1^\theta, \ldots, \psi_p^\theta]$ with $\psi_i^\theta = \frac{u_i^+}{\underline{\theta}_i}$ and the associated portion of the state space (recall the assumption $u_i^+ = -u_i^-$)

$$S(\psi^\theta) = \{z \in \mathbb{R}^m : -\psi^\theta \leq -Kz \leq \psi^\theta\}.$$

Consider now the 2^p vectors $\gamma_j \in \{1, \underline{\theta}_1\} \times \ldots \times \{1, \underline{\theta}_p\}$, with $j = 1, \ldots, 2^p$. In other words, γ_j is a p component vector with ith component γ_{ji} taking value 1 or $\underline{\theta}_i$. Then, each matrix A_j can be expressed as

$$L_j = -Bdiag(\gamma_j)K = -Bdiag(\gamma_j)\frac{B^T}{\gamma}.$$

Roughly speaking each vector γ_j stores the minimum and or maximum degree of saturation of all controls.

Now partition $S(\psi^\theta)$ in subsets X such that for each of them we can define the subset $J_X \subseteq \{1, \ldots, 2^p\}$ of indices j such that, for all $z \in X$, $L(z)$ can be expressed as a convex combination of L_j's with $j \in J_X$. This completes the procedure.

4.2 Attainability

In this section we consider two extensions of the above results. First we focus on attainability rather than approachability, and then we generalize the structure of the state-feedback function by considering the following function, for all arc $(j, k) \in E$:

$$\begin{aligned} u_i = &\min(\alpha_{jk}[z_j - z_k]_+, u_i^+) \\ &+ \max(\alpha_{jk}[z_j - z_k]_-, u_i^-), \end{aligned} \quad (24)$$

where i is the index of the arc $(j, k) \in E$ according to some ordered indexing in E, α_{jk} are nonnegative weights for all arcs $(j, k) \in E$, and $[z_j - z_k]_+$ denotes the positive part of $z_j - z_k$. Note that when the $\alpha_{jk} = \frac{1}{\gamma}$ then we have the saturated function below. Let us rewrite (24) in compact form as

$$\begin{aligned} u = \phi_\alpha(z) := &\min(\mathcal{A}[\Delta z]_+, u^+) \\ &+ \max(\mathcal{A}[\Delta z]_-, u^-), \end{aligned} \quad (25)$$

where $\mathcal{A}$ is a $p \times p$ diagonal matrix with entries α_{jk} for all $(j, k) \in E$ in the main diagonal, Δz in $\mathbb{R}^p$ is the vector of state difference at the two extreme nodes of the each arc, and all operators need to be interpreted component-wise.

In the case of attainability the set of equilibrium points is given by

$$\Lambda_\alpha \doteq \left\{z \mid B\phi_\alpha(z) = D\hat{\phi}(z)\right\}.$$

The following assumption is in the form of set inclusion and turns to be necessary and sufficient to attainability as established in the next theorem.

Assumption 2 *Matrix $B \in \mathbb{R}^{m \times p}$ is full row rank and set $\mathcal{W}$ is in the interior of $B\mathcal{U}$, that is,*

$$D\mathcal{W} \subset int\{B\mathcal{U}\}. \quad (26)$$

Theorem 4.4 *Under Assumption 2, the control (25), under an optimal $\alpha > 0$ is such that $z(\tau) \to \Lambda_\alpha$.*

Proof. We need to prove that there exists an optimal α such that $\langle v, z - y \rangle < 0$. This is true if we rewrite

$$\begin{aligned} &\inf_\alpha \langle v, z - y \rangle \\ &= \inf_\alpha \left\langle B\phi_\alpha(z) - D\hat{\phi}(z), z - y \right\rangle \\ &\leq \inf_\alpha \sup_w \left\langle B\phi_\alpha(z) - Dw, z - y \right\rangle \leq 0, \end{aligned} \quad (27)$$

where the last inequality derives from Assumption 2. ∎

4.3 Approachability of the origin

In this section, we study approachability of the origin under the worst-case realization of the disturbance. To this end, consider a generic hyperbox set $C = \{\zeta \in \mathbb{R}^m \mid z_i^- \leq \zeta_i \leq z_i^+\}$, where z_i^- and z_i^+ are negative and positive scalars respectively.

Equation (20) can be rewritten as

$$\begin{aligned} \dot{z}(\tau) \in F(z) := \{\xi \in \mathbb{R}^m \mid \\ \xi = Bsat\left[-\frac{B^T}{\gamma}z\right] - Dw - z, \forall w \in \mathcal{W}\}. \end{aligned} \quad (28)$$

As the boundary of the set ∂C is nonsmooth, we can approximate the set by introducing the following gauge function.

For any positive integer p let the function $\sigma_p : \mathbb{R} \to \mathbb{R}_+$ be defined as

$$\sigma_p(\zeta) = \begin{cases} \zeta^p & \text{if } \zeta \leq 0, \\ 0 & \text{if } \zeta > 0, \end{cases}$$

and consider a gauge function $\Psi_p : \mathbb{R}^m \mapsto \mathbb{R}^+$ defined as:

$$\Psi_p(z) = \sqrt[p]{\sum_{i=1}^n \sigma_p \left(\frac{z_i}{z_i^+} \right) + \sigma_p \left(\frac{z_i}{z_i^-} \right)}.$$

Note that the unit ball

$$B_{\Psi_p}(\vec{0}, 1) := \{ \xi \in \mathbb{R}^m | \Psi_p(\xi) \leq 1 \}$$

is included in C, i.e., $B_{\Psi_p}(\vec{0}, 1) \subseteq C$, and is such that the boundary $\partial B_{\Psi_p}(\vec{0}, 1)$ is smooth (differentiable). Furthermore $\partial B_{\Psi_p}(\vec{0}, 1)$ tends asymptotically to ∂C for increasing p, and as such $B_{\Psi_p}(\vec{0}, 1)$ represents a good approximation of C. We show next that set $B_{\Psi_p}(\vec{0}, 1)$ is approachable and discuss the approachability strategy.

It turns out that, a possible control strategy is one that pushes the state along the anti-gradient direction of the above function. More formally, if we denote by

$$\Gamma_i(z_i) := \frac{1}{z_i^+} \sigma_{p-1} \left(\frac{z_i}{z_i^+} \right) + \frac{1}{z_i^-} \sigma_{p-1} \left(\frac{z_i}{z_i^-} \right),$$

the gradient for $z \neq 0$ can be expressed as

$$\nabla \Psi_p(z) = \Psi_p(z)^{1-p} [\Gamma_1(z) \quad \Gamma_2(z) \ldots \Gamma_n(z)]. \quad (29)$$

Let the following set of equilibrium points be given:

$$\hat{\Lambda}_\alpha := \left\{ z | \ B\phi_\alpha(z) = D\hat{\phi}(z) \right\}.$$

Consider the following assumption, which is a slight variation of Assumption 1.

Assumption 3 *Matrix $B \in \mathbb{R}^{m \times p}$ is full row rank and set $\mathcal{W}$ is in the interior of $B\mathcal{U}$, that is,*

$$D\mathcal{W} + \hat{\Lambda}_\alpha \subset int\{B\mathcal{U}\}. \quad (30)$$

Theorem 4.5 (Approach. with feedback strategies)
Let Assumption 3 hold. Let a generic hyperbox set $C = \{ \zeta \in \mathbb{R}^m | z_i^- \leq \zeta_i \leq z_i^+ \}$ where z_i^- and z_i^+ are negative and positive scalars. Let $r > 0$ and $Z = \{ z \in \mathbb{R}^m : dist(z, C) < r \}$. Then the set C is approachable.

Proof. Let $z(t)$, $t \in [0, T]$ be solution of (28). The underlying idea of this proof is to show that for all $z \in Z \backslash C$ there exists $\Gamma(z)$ such that the

$$\langle v, \Gamma(z) \rangle < 0, \quad \forall v \in F(z). \quad (31)$$

Now, we have that the derivative of Ψ is given by

$$\min_\alpha \max_{w \in \mathcal{W}} \dot{\Psi}(t) = \min_\alpha \max_{w \in \mathcal{W}} \langle \nabla \Psi_p(z), \dot{z} \rangle$$
$$= \min_\alpha \max_{w \in \mathcal{W}} \Psi_p(z)^{1-p} \langle \Gamma(z), \dot{z}(t) \rangle$$
$$= \min_\alpha \max_{w \in \mathcal{W}} \Psi_p(z(t))^{1-p}$$
$$\cdot \langle \Gamma(z(t)), B\phi_\alpha(z) - Dw - z \rangle < 0.$$

From the above we have that (31) holds true. Now note that the condition $\Psi_p(z(t)) < 1$ implies $z \in C$ and as (32) implies that $\Psi_p(z(t)) \to 0$ then $z(t)$ ultimately reaches C as well. ∎

5 Adaptation and coordination

Centralized organizations entail expensive communication in that one single decision-maker has to process big data sets and coordinate multiple actions. One way to overcome this issue is through decentralization and task specialization. This consists in partitioning the project into tasks and assign them to multiple agents [11]. Decentralization in turn requires *adaptation* and *coordination*. By adaption we mean the capability to adapt to

- market conditions: the actual demand may be higher or lower than forecasted;
- operational conditions: employees may be not available, or unexpected delays may occur;
- consumers' needs: changing characteristics or needs require the products to be continuously redesigned.

In such a scenario each agent must continuously adapt its task to new instances and coordinate the changes with the other agents.

As an example, imagine a large software to be developed by a team of engineers. The first step is to decompose the project in multiple tasks and to assign each task to a different engineer. Think of the software as a proprietary operating system having a task focusing on the process manager, another task relating to the network access and so forth. While each task has to be designed based on the specific needs of the client, all tasks require to be assembled in coherent whole.
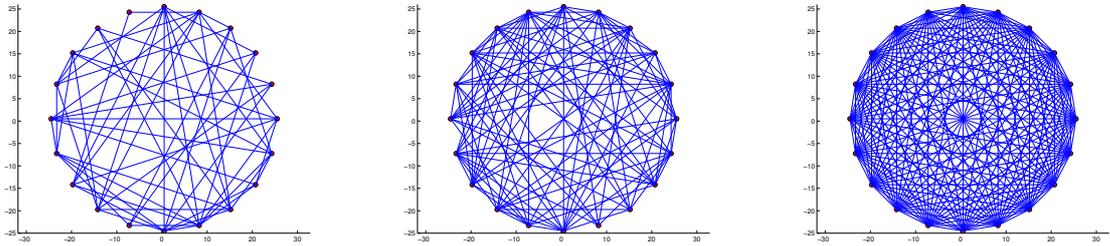
(a) Case I: $G = (V, E)$ with probability of formation of links $h = 0.3$.

(b) Case II: $G = (V, E)$ with probability of formation of links $h = 0.6$.

(c) Case III: $G = (V, E)$ with probability of formation of links $h = 0.99$.

Fig. 5. Topologies

5.1 Numerical example

Consider a decentralized organization, in which a project is decomposed in n tasks, and each task is assigned to an agent. At each time, agents access local information and adapt their tasks consequently. Local information is modeled as an exogenous input w. Coordination is possible via pairwise adjustments u and is visualized by a network described by an incidence matrix B. Nodes are agents, and links are communication channels. Ideally the value of task j, which is indicated by z_j, should be as close as possible to that of task i, denoted by z_i.

The project consists of $n = 20$ tasks. Iterations are $T = 100$. Consider a discrete time version of (20) given by

$$z(t + dt) = z(t) + (-cLz(t) + bw(t))dt, \qquad (32)$$

where c is the coordination weight, and b is the adaptation weight. The parameters are as follows. The step size $dt = 0.01$, the initial state value $z(0)$ is generated as a single uniformly distributed random number in the interval $(0, 1)$ by using the in-built MATLAB command `rand`. The adaptation weight $b = 1.5$, 45.5, 15.5, while the coordination weight $c = 1, 1, 0.5$ for the three simulation sets. For each simulation set, we consider three cases, in which the communication graph is built by fixing a probability of formation of links denoted by $h = 0.3$, 0.6, 0.99, respectively. The exogenous input w is an n-dimensional vector with components uniformly distributed in the set $\{-1, 0, 1\}$. Figures 5a-5c display the graphs in the three cases considered for each simulation set. The three cases differ for the probability of formation of links which is $h = 0.3, 0.6, 0.99$.

The first set of simulations highlights the dominant role of coordination at the expense of adaptation due to an increase in the number of links, which is around 6, 12, and 19 in the three cases. Figure 6 shows the time plot of the task values $z(t)$ in the three cases. The coordination level increases from top to bottom. Thus, investing in a better quality of internal communication benefits the overall coordination capability of the organization.
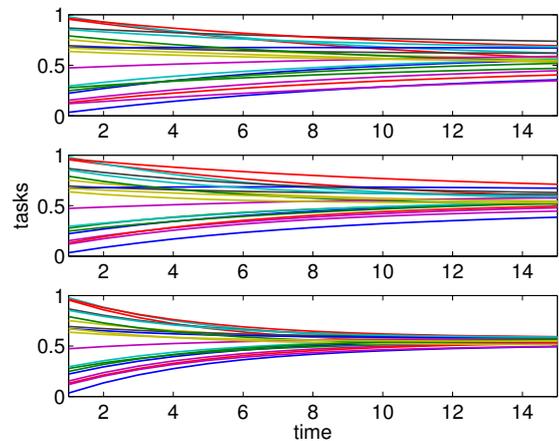

Fig. 6. First set of simulations: time plot of the task values $z(t)$ in the three cases.

The second set of simulations points out how the role of communication may be secondary when the cost of misadaption dominates the cost of miscoordination. Actually, we now set $b = 45.5$ and $c = 1$ in (32), which means that the cost of misadaption is much higher than the cost of miscoordination. If this is the case, investing on internal communication does not benefit much the organization. Figure 7 shows the time plot of the task values $z(t)$ in the three cases. The agents follow the exogenous signal which leads to the formation of three clusters. The approachable set is larger than in previous simulations.

Another scenario where investing on internal communication is not relevant is when the exogenous signal has small volatility. In this case, the agents stick to a priori coordination without compromising the overall coordination of the organization. This is captured in the third set of simulations. We now set w almost constant and equal to 0.5. Even if the cost of adaptation is higher than the cost of coordination, which is obtained by setting $b = 15.5$ and $c = 0.5$, the level of coordination is almost the same for the three graphs. Figure 8 shows the time
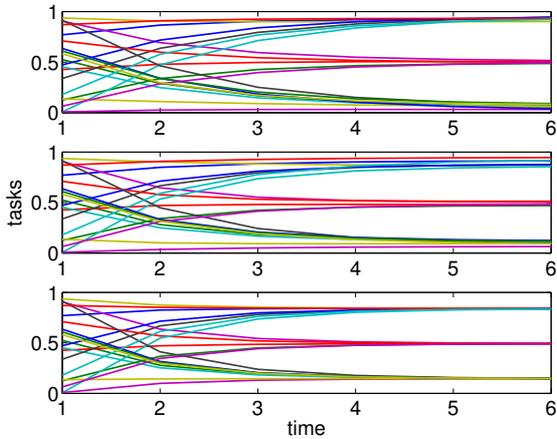
9

Fig. 7. Second set of simulations: time plot of the task values $z(t)$ in the three cases.

plot of the task values $z(t)$ in the three cases. Though the agents follow the exogenous signal, this leads to the formation of one single cluster around 0.5. The set of perfect coordination characterized by $z = \mathbf{1}\mu$ where μ is a scalar is approachable in a distributed way.
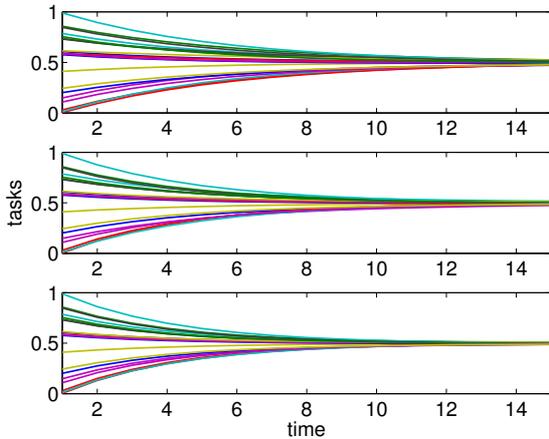

Fig. 8. Third set of simulations: time plot of the task values $z(t)$ in the three cases.

6 Conclusions

This paper has introduced distributed approachability to accommodate cooperation, competition, and local interaction in multi-agent systems. The advantage of such a novel framework is that we can turn the original questions on cooperation, competition and local interactions, into convergence properties of a differential inclusion describing the evolution of the collective state. In particular, we have provided convergence conditions under gen-

eral Markovian strategies. We have specialized our results to the case of decentralized organizations.

References

[1] J. Abernethy, P. L. Bartlett, and E. Hazan. Blackwell approachability and no-regret learning are equivalent. In *Proceedings of 24th Annual Conference on Learning Theory*, 19, 27–46, 2011.

[2] D. Bauso, G. Notarstefano. Distributed n-player approachability and consensus in coalitional games. *IEEE Trans. on Automatic Control*, 60(11): 3107–3112, 2015.

[3] D. Bauso, *Game Theory with Engineering Applications*, SIAM's Advances in Design and Control series, Philadelphia, PA, USA, 2016.

[4] D. Bauso, E. Solan, E. Lehrer, X. Venel. Attainability in Repeated Games with Vector Payoffs. *Math. of Op. Res.*, 40(3):739–755, 2015.

[5] D. Bauso, F. Blanchini, L. Giarrè, R. Pesenti. The linear saturated decentralized strategy for constrained flow control is asymp. optimal. *Automatica*, 49(7): 2206–2212, 2013.

[6] M. Benaïm, J. Hofbauer, and S. Sorin. Stochastic Approximations and Differential Inclusions. *SIAM J. on Control and Optimization*, 44(1): 328–348, 2005.

[7] M. Benaïm, J. Hofbauer, and S. Sorin. Stochastic Approximations and Differential Inclusions, Part II: Applications. *Math. of Op. Res.*, 31(4): 673–695, 2006.

[8] D. Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific J. Math.*, 6(1): 1–8, 1956.

[9] F. Blanchini, S. Miani, W. Ukovich. Control of production-distribution systems with unknown inputs and system failures. *IEEE Trans. on Automatic Control*, 45(6): 1072–1081, 2000.

[10] F. Blanchini, F. Rinaldi, W. Ukovich. Least inventory control of multi-storage systems with non-stochastic unknown input. *IEEE Trans. on Robotics and Automation*, 13: 633–645, 1997.

[11] W. Dessein, T. Santos. Adaptive Organizations. *Journal of Political Economy*, 114: 956– 995, 2006.

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

[13] D. Kalathil, V. S. Borkar, R. Jain. Approachability in Stackelberg Stochastic Games with Vector Costs. *Dynamic Games and Applications* 2016, 1–21.

[14] V. S. Kamble. Games with Vector Payoff: A Dynamic Programming Approach. PhD Dissertation, University of California, Berkeley, 2015.

[15] R. Olfati-Saber, J. A. Fax, R. M. Murray. Consensus and Cooperation in Networked Multi-Agent Systems. *Proceedings of the IEEE*, 95(1): 215–233, 2007.

[16] A. S. Soulaimani, M. Quincampoix, S. Sorin. Approchability theory, discriminating domain and differential games. *SIAM J. of Control and Optimization*, 48(4): 2461–2479, 2009.