Distributed n-player approachability and

consensus

in coalitional games

Dario Bauso and Giuseppe Notarstefano

Abstract

We study a distributed allocation process where, repeatedly in time, every player renegotiates past allocations with neighbors and allocates new revenues. The average allocations evolve according to a doubly (over time and space) averaging algorithm. We study conditions under which the average allocations reach consensus to any point within a predefined target set even in the presence of adversarial disturbances. Motivations arise in the context of coalitional games with transferable utilities (TU) where the target set is any set of allocations that make the grand coalitions stable.

I. INTRODUCTION

We consider a two-step distributed allocation process where at every time players first renegotiate their past allocations and second generate a new revenue and allocate it. The time-averaged allocations evolve according to a *doubly (over time and space) averaging*

dynamics. The goal is to let all allocations reach consensus to any value in a predefined set even in the presence of an adversarial disturbance.

Motivations. The problem arises in the context of dynamic coalitional games with Transferable Utilities (TU games) [8]. A coalitional TU game consists in a set of players, who can form coalitions, and a characteristic function that provides a value for each coalition. The predefined set introduced above can be thought of as (but it is not limited to) the core of the game. This is the set of imputations under which no coalition has a value greater than the sum of its members' payoffs. Therefore, no coalition has incentive to leave the grand coalition and receive a larger payoff.

Highlights of contributions. We analyze conditions under which the average allocations: (i) *approach* the set X (Theorem 1), (ii) reach *consensus*, in which case we also compute the consensus value (Theorem 2), and (iii) are robust against disturbances (Theorem 3).

Related literature. Coalitional games with transferable utilities (TU) were first introduced by von Neumann and Morgenstern [14]. Here, a main issue is to study whether the core is an "approachable" set, and which allocation processes can drive the "complaint vector" to that set. Approachability theory was developed by Blackwell in the early '56, [2], and is captured in the well known Blackwell's Theorem. The geometric (approachability) principle that lies behind the Blackwell's Theorem is among the fundamentals in allocation processes in coalitional games [7]. The discrete-time dynamics analyzed in the paper follows the rules of a typical consensus dynamics (see, e.g., [11] and references therein). among multiple agents, where an underlying communication graph for the agents and balancing weights have been used with some variations to reach an agreement on common decision variable in [10], [9], [11], [13], [12], [4] for distributed multi-agent optimization.

The paper is organized as follows. In Section, II, we formulate the problem and discuss motivations and main assumptions. In Section III, we illustrate the main results. In Section IV we provide numerical illustrations. Finally, in Section V, we provide concluding remarks and future directions.

Notation. We view vectors as columns. For a vector x, we use $[x]_j$ to denote its jth coordinate component. We let x' denote the transpose of a vector x, and $\|x\|$ denote its Euclidean norm. An $n \times n$ matrix A is row-stochastic if the matrix has nonnegative entries a_j^i and $\sum_{j=1}^n a_j^i = 1$ for all $i = 1, \ldots, n$. For a matrix A, we use a_j^i or $[A]_{ij}$ to denote its ijth entry. A matrix A

is doubly stochastic if both A and its transpose A' are row-stochastic. Given two sets U and S, we write $U \subset S$ to denote that U is a proper subset of S. We use $|S|$ for the cardinality of a given finite set S. We write $P_X[x]$ to denote the projection of a vector x on a set X, and we write $\text{dist}(x, X)$ for the distance from x to X, i.e., $P_X[x] = \arg \min_{y \in X} \|x - y\|$ and $\text{dist}(x, X) = \|x - P_X[x]\|$, respectively. Given a function of time $x(\cdot) : \mathbb{N} \to \mathbb{R}$, we denote by $\bar{x}(t)$ its average up to time t, i.e., $\bar{x}(t) := \frac{1}{t} \sum_{\tau=1}^{t} x(\tau)$.

II. DISTRIBUTED REWARD ALLOCATION ALGORITHM

Every player in a set $N = \{1, \ldots, n\}$ is characterized by an average allocation vector $\hat{x}_i(t+1) \in \mathbb{R}^n$. At every time he renegotiates with *neighbors* all past allocations and generates a new allocation vector $x_i(t + 1)$. The time-averaged allocation $\hat{x}_i(t)$ evolves as follows:

$$\hat{x}_i(t + 1) = \frac{t}{t+1} \left[\sum_{j=1}^{n} a_j^i(t) \hat{x}_j(t) \right] + \frac{1}{t+1} x_i(t+1), \tag{1}$$

where $a^i = (a_1^i, \ldots, a_n^i)'$ is a vector of nonnegative weights consistent with the sparsity of the *communication graph* $\mathcal{G}(t) = (N, \mathcal{E}(t))$. A link $(j, i) \in \mathcal{E}(t)$ exists if player j is a neighbor of player i at time t, i.e. if player i renegotiates allocations with player j at time t.

Problem. Our goal is to study under what conditions all allocation vectors converge to a unique value and this value belongs to a predefined set X: for all $i, j \in V$,

$$\hat{x}_i(t) = \hat{x}_j(t) \in X, \quad \text{for } t \to \infty. \tag{2}$$

In the sequel, we rewrite equation (1) in the compact form:

$$\hat{x}_i(t + 1) = \frac{t}{t+1} w_i(t) + \frac{1}{t+1} x_i(t+1), \tag{3}$$

where $w_i(t)$ is the *space average* defined as

$$w_i(t) = \left[\sum_{j=1}^{n} a_j^i(t) \hat{x}_j(t) \right]. \tag{4}$$

A. Motivations

The set X introduced above can be thought of as the core of a coalitional game with Transferable Utilities (TU game).

A coalitional TU game is defined by a pair $< N, \eta >$, where $N = \{1, \ldots, n\}$ is a set of players and $\eta : 2^N \to \mathbb{R}$ a function defined for each coalition $S \subseteq N$ $(S \in 2^N)$. The function

η determines the value $\eta(S)$ assigned to each coalition $S \subset N$, with $\eta(\emptyset) = 0$. We let η_S be the value $\eta(S)$ of the characteristic function η associated with a nonempty coalition $S \subseteq N$. Given a TU game $< N, \eta >$, let $C(\eta)$ be the core of the game,

$$C(\eta) = \left\{ x \in \mathbb{R}^n \, \Big| \, \sum_{j \in N}[x]_j = \eta_N, \right.$$

$$\left. \sum_{j \in S}[x]_j \geq \eta_S \text{ for all nonempty } S \subset N \right\}.$$

Essentially, the core of the game is the set of all allocations that make the grand coalition stable with respect to all subcoalitions. Condition $\sum_{j \in N}[x]_j = \eta_N$ is also called efficiency condition. Condition $\sum_{j \in S}[x]_j \geq \eta_S$ for all nonempty $S \subset N$ is referred to as "stability with respect to subcoalitions", since it guarantees that the total amount given to the members of a coalition exceeds the value of the coalition itself.

B. Main assumptions

Following [11] (see also [8]) we can make the following assumptions on the information structure. We let $A(t)$ be the weight matrix with entries $a_j^i(t)$.

Assumption 1: Each matrix $A(t)$ is doubly stochastic with positive diagonal. Furthermore, there exists a scalar $\alpha > 0$ such that $a_j^i(t) \geq \alpha$ whenever $a_j^i(t) > 0$.

At any time, the instantaneous graph $\mathcal{G}(t)$ need not be connected. However, for the proper behavior of the process, the union of the graphs $\mathcal{G}(t)$ over a period of time is assumed to be connected.

Assumption 2: There exists an integer $Q \geq 1$ such that the graph $\left(N, \bigcup_{\tau=tQ}^{(t+1)Q-1} \mathcal{E}(\tau) \right)$ is strongly connected for every $t \geq 0$.

It is worth noting that the above assumptions are fairly standard in the distributed computation literature. In particular, the joint strong connectivity is the weakest possible assumption to guarantee persistent circulation of the information through the graph. The double stochasticity of the matrix $A(t)$ is a common assumption to guarantee average consensus.

Let $X \subset \mathbb{R}^n$ be the core set of the game. A common assumption in approachability theory is that both the core set is convex and bounded, and the payoff (or loss) vectors generated at each time are bounded. Thus, following [2], [5], we borrow and adapt such an assumption to our framework.
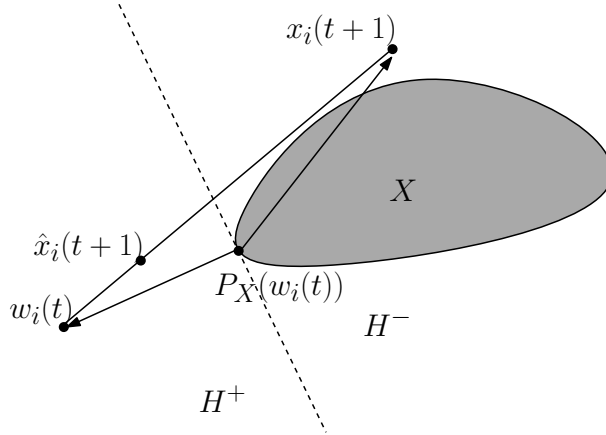
Fig. 1. Approachability principle.

Assumption 3: The core set X is nonempty.

Notice that a nonempty core is a convex and compact set.

The next assumption indicates how the new reward vector has to be generated in order to obtain approachability.

Assumption 4: For each $i \in N$ the new reward vector $x_i(\cdot)$ is bounded, i.e., there exists $L > 0$ s.t. $\forall t \geq 0 \ \|x_i(t+1)\| \leq L$, and satisfies the following inequality, for a scalar negative number, $\phi < 0$,

$$(w_i(t) - P_X(w_i(t)))' (x_i(t+1) - P_X(w_i(t)) \leq \phi < 0.$$

From a geometric standpoint, Assumption 4 requires that, given the two half-spaces identified by the supporting hyperplane of X through $P_X(w_i(t))$, the new reward vector $x_i(t+1)$ lies in the half-space not containing $w_i(t)$.

III. MAIN RESULTS

Next, we provide the main results of the paper. Namely, we prove that the average allocations: (i) *approach* the set X (Theorem 1), (ii) reach *consensus* (Theorem 2), and (iii) are robust against disturbances (Theorem 3).

A. Approachability and consensus

Before stating the first theorem, we need to introduce two lemmas. The next lemma establishes that the space averaging step in (1) reduces the total distance (i.e. the sum of distances) of the estimates from the set X.

Lemma 1: Let Assumption 1 hold. Then the total distance from X decreases when replacing the allocations $\hat{x}_i(t)$ by their space averages $w_i(t)$, i.e.,

$$\sum_{i=1}^{n} \text{dist}(w_i(t), X) \leq \sum_{i=1}^{n} \text{dist}(\hat{x}_i(t), X).$$

As a preliminary step to the next result, observe that, from the definition of $\text{dist}(\cdot, X)$ and from (1) and (4), it holds

$$
\begin{aligned}
\text{dist}(\hat{x}_i(t+1), X)^2 &= \|\hat{x}_i(t+1) - P_X[\hat{x}_i(t+1)]\|^2 \\
&\leq \|\hat{x}_i(t+1) - P_X[w_i(t)]\|^2 \\
&= \left\| \frac{t}{t+1} w_i(t) + \frac{1}{t+1} x_i(t+1) - P_X[w_i(t)] \right\|^2 \\
&= \left\| \frac{t}{t+1} (w_i(t) - P_X[w_i(t)]) \right. \\
&\qquad \left. + \frac{1}{t+1} (x_i(t+1) - P_X[w_i(t)]) \right\|^2 \\
&= \left(\frac{t}{t+1} \right)^2 \|w_i(t) - P_X[w_i(t)]\|^2 \\
&\qquad + \left(\frac{1}{t+1} \right)^2 \|x_i(t+1) - P_X[w_i(t)]\|^2 \\
&\qquad + \frac{2t}{(t+1)^2} (w_i(t) - P_X[w_i(t)])'(x_i(t+1) - P_X[w_i(t)]).
\end{aligned}
\tag{5}
$$

The following lemma states that, under the approachability assumption, the distance of each single estimate from X decreases with respect to the one of the spatial average when applying the time averaging step.

Lemma 2: Let Assumptions 3-4 hold. Then, there exists a positive integer scalar, $\tilde{t} > 0$, such that for all $t \geq \tilde{t} > 0$ the distance of each single $\hat{x}_i(t+1)$ decreases in comparison with the distance of $w_i(t)$, i.e.,

$$\text{dist}(\hat{x}_i(t+1), X) < \text{dist}(w_i(t), X), \quad \forall i = 1, \ldots, n$$

We are now ready to state the first main result.

Theorem 1: Let Assumptions 1-4 hold. Then all average allocations approach set X, i.e.,

$$\lim_{t \to \infty} \sum_{i=1}^{n} \text{dist}(\hat{x}_i(t), X) = 0.$$

Next, let us introduce the barycenter of respectively the estimates and the reward vectors

$$\hat{x}_b(t) := \frac{1}{n} \sum_{i=1}^{n} \hat{x}_i(t) \qquad \text{and} \qquad x_b(t) := \frac{1}{n} \sum_{i=1}^{n} x_i(t).$$

Consistently, let us denote as $\bar{x}_b(t)$ the time average of the barycenter, i.e.

$$\bar{x}_b(t) = \frac{1}{t+1} \sum_{\tau=0}^{t} x_b(\tau).$$

The following lemma establishes that the barycenter of the estimates evolves as the time average $\bar{x}_b(t)$ of the barycenter of the reward vectors generated by the players.

Lemma 3: The barycenter of the local allocations $\hat{x}_b(t)$ coincides at each time t with the time-average of the barycenter of the generated reward vectors $\bar{x}_b(t)$.

The following theorem establishes that all allocations converge to $\bar{x}_b(t)$, which in the limit must belong to X according to Theorem 1.

Theorem 2: (Consensus to the barycenter time-average) Let Assumptions 1-4 hold. Then, all players reach consensus on the time-average of the barycenter of the reward vectors generated by each player, $\bar{x}_b(t)$, i.e.,

$$\lim_{t \to \infty} \|\hat{x}_i(t) - \bar{x}_b(t)\| = 0 \quad \forall i = 1, \dots, n.$$

Summarizing the two main results, we have proven that asymptotically all the players' allocations converge to the time-average of the barycenter of the generated reward vectors and that this vector lies in the core of the game.

B. Adversarial disturbance

Here we analyze the case where, for each player $i \in N$, the input $x_i(\cdot)$ is the payoff of a repeated two-player game between player i (Player i_1) and an (external) adversary (Player i_2). With some slight abuse of notation we denote S_1 and S_2 the finite set of actions of players i_1 and i_2 respectively.

The instantaneous payoff $x_i(t)$ at time t is given by a function $\phi_i : S_1 \times S_2 \to \mathbb{R}^n$ as follows:

$$x_i(t) = \phi(j(t), k(t)),$$

where $j(t) \in S_1$ and $k(t) \in S_2$. We extend x_i to the set of mixed actions pairs, $\Delta(S_1) \times \Delta(S_2)$, in a bilinear fashion. In particular, for every pair of mixed strategies $(p(t), q(t)) \in \Delta(S_1) \times$

$\Delta(S_2)$ for player i_1 and i_2 at time t, the expected payoff is

$$\mathbb{E}x_i(t) = \sum_{j \in S_1} \sum_{k \in S_2} p_j(t)q_k(t)\phi(j,k).$$

For simplicity the one-shot vector-payoff game (S_1, S_2, x_i) is denoted by G_i.

Let $\lambda \in \mathbb{R}^n$. Denote by $\langle \lambda, G_i \rangle$ the zero-sum one-shot game whose set of players and their action sets are as in the game G_i, and the payoff that player 2 pays to player 1 is $\lambda'\phi(j,k)$ for every $(j,k) \in S_1 \times S_2$.

The resulting zero-sum game is described by the matrix

$$\Phi_\lambda = [\lambda'\phi(j,k)]_{j \in S_1, k \in S_2}.$$

As a zero-sum one-shot game, the game $\langle \lambda, G_i \rangle$ has a value, denoted

$$v_\lambda := \min_{p \in \Delta S_1} \max_{q \in \Delta S_2} p'\Phi_\lambda q = \max_{q \in \Delta S_2} \min_{p \in \Delta S_1} p'\Phi_\lambda q.$$

For every mixed action $p \in \Delta(S_1)$ denote $D_1(p)$ the set of all payoffs that might be realized when player i_1 plays the mixed action p:

$$D_1(p) = \{x_i(p,q) \colon q \in \Delta(S_2)\}.$$

If $v_\lambda \geq 0$ (resp. $v_\lambda > 0$), then there is a mixed action $p \in \Delta(S_1)$ such that $D_1(p)$ is a subset of the closed half space $\{x \in \mathbb{R}^n \colon \lambda'x \geq 0\}$ (resp. half space $\{x \in \mathbb{R}^m \colon \lambda'x > 0\}$).

Let us introduce next the counterpart of Assumption 4 in this new worst-case setting.

Assumption 5: For any $w_i(t) \in \mathbb{R}^n$, there exists a mixed strategy $p(t+1) \in \Delta(S_1)$ for Player i_1 such that, for all mixed strategy $q(t+1) \in \Delta(S_2)$ of Player i_2, the new reward vector $x_i(\cdot)$ is bounded, i.e. there exists $L > 0$ s.t. $\forall t \geq 0 \; \|x_i(t+1)\| \leq L$, and satisfies

$$(w_i(t) - P_X(w_i(t)))' (\mathbb{E}x_i(t+1) - P_X(w_i(t))) \leq \phi < 0,$$

where $\mathbb{E}x_i(t+1) = \sum_{j \in S_1} \sum_{k \in S_2} p_j(t+1)q_k(t+1)\phi(j,k)$.

The above condition is among the foundations of approachability theory as it guarantees that the average payoff $\frac{1}{T}\sum_{t=0}^{T-1} x_i(t)$ converges almost surely to X (see, e.g., [2] and also [5], chapter 7). Here we adapt the above condition to the multi-agent and distributed scenario under study.

Corollary 3.1 (see [2], Corollary 2): Any convex set $X \subset \mathbb{R}^n$ is approachable if and only if $v_\lambda < 0$ for any $\lambda \in \mathbb{R}^n$.

Next we show that if the approachability condition expressed above holds true, then $dist(\hat{x}_i, X)$ tends to zero for any X. We write $w.p.1$ to mean "with probability 1".

Theorem 3: Let Assumptions 1-3 and 5 hold. Then all average allocations approach set X, i.e.,

$$\lim_{t \to \infty} \sum_{i=1}^{n} \text{dist}(\hat{x}_i(t), X) = 0, \quad w.p.1.$$

We conclude this section by observing that Theorem 2 still holds and therefore all players' estimates reach consensus on the time-average of the barycenter of the reward vectors generated by each player.

IV. SIMULATIONS

We illustrate the results in a game with four players, $N = \{1, \ldots, 4\}$, communicating according to a fixed undirected cycle graph. That is, $\mathcal{G}(t) = (N, \mathcal{E})$ where $\mathcal{E} = \{(i, j) \mid j = i + 1, i \in \{1, \ldots n - 1\}$ or $(i, j) = (n, 1)\}$.

We set $\eta_{\{1\}} = \ldots = \eta_{\{4\}} = 2$, $\eta_{\{1,2\}} = 5$, $\eta_{\{3,4\}} = 5$, $\eta_{\{1,2,3\}} = 7$ and $\eta_N = 10$ (η_S is the value of coalition S). That is, each player expects to receive at least a reward of 2 which is its value as a singleton coalition. But, for example, players 1 and 2 expect to be more valuable if they form a coalition as well as 3 and 4. Consistently, the core of the game is the polyhedral set given by

$$C(\eta) = \left\{ x \in \mathbb{R}^4 \,\middle|\, x_1 + x_2 + x_3 + x_4 = 10, \right.$$
$$x_1 + x_2 + x_3 \geq 7, \; x_1 + x_2 \geq 5,$$
$$\left. x_3 + x_4 \geq 5, x_1 \geq 2, \ldots, x_4 \geq 2 \right\}.$$

We initialize the assignments assuming each player assign itself the entire reward. That is, denoting $b_i \in \mathbb{R}^n$ the i-th canonical vector (so that, e.g., $b_1 = [1 \; 0 \; \ldots \; 0]'$), we set $\hat{x}_i(0) = 10 \, b_i$ for all $i \in \{1, \ldots, n\}$. At every iteration $t \in \mathbb{N}$, each player chooses the new reward vector $x_i(t+1)$ according to the approachability principle. In particular, we set $x_i(t+1) = P_X[w_i(t)] + \alpha \left(P_X[w_i(t)] - w_i(t)\right) + v^\top$, where α is a random number uniformly distributed in $[0, 1]$ and $v^\top$ a random vector belonging to the hyperplane tangent to the core at $P_X[w_i(t)]$ with coordinates uniformly chosen in $[0, 1]$. The temporal evolution of the local estimates of the average reward vector is depicted in Figure 2. As expected the local estimates converge to the same average assignment which is the point of the core $[3.8 \; 3 \; 2.2 \; 1]'$.
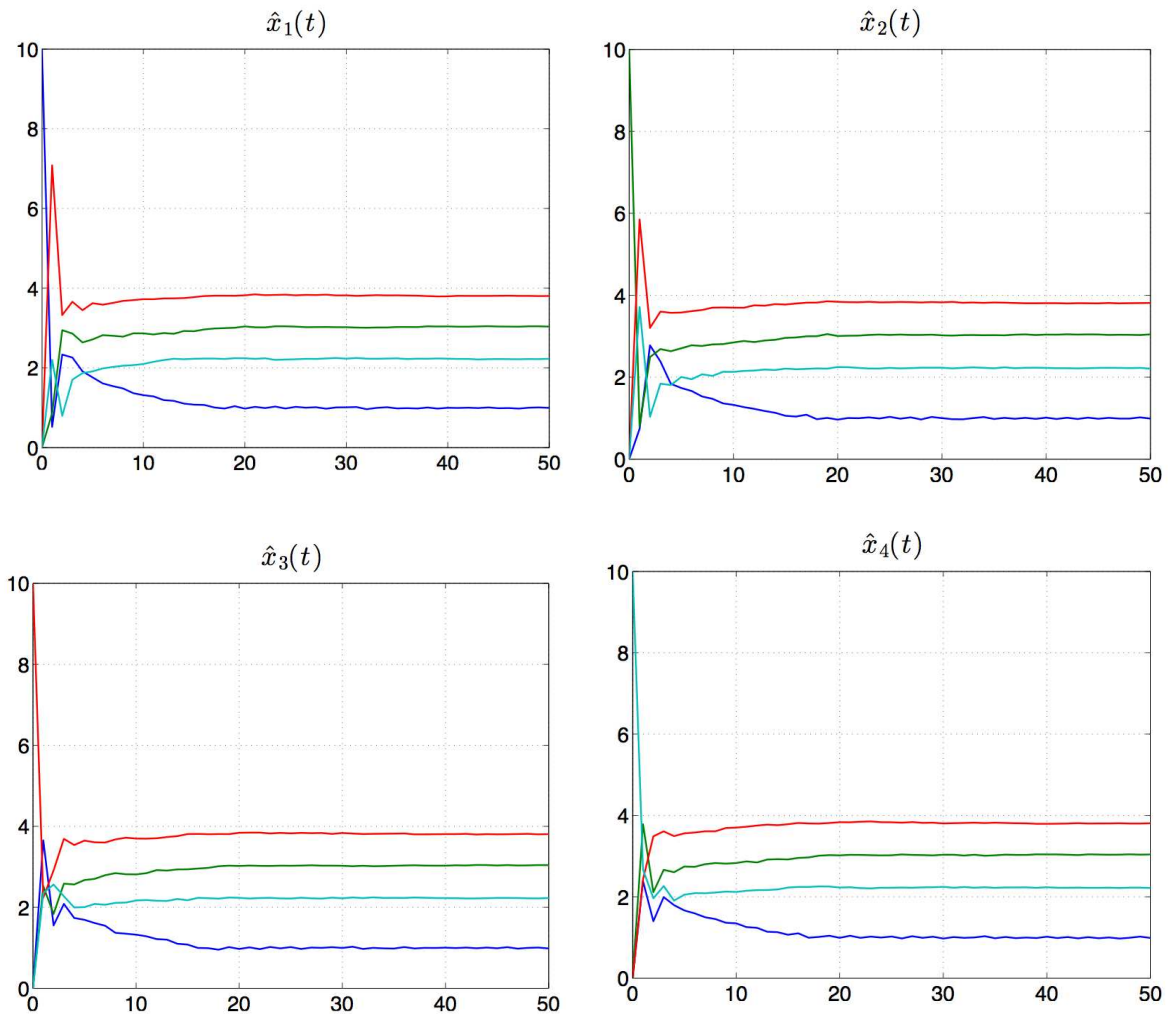
Fig. 2. Local average reward vectors

V. Conclusions

We have analyzed convergence conditions of a distributed allocation process arising in the context of TU games. Future directions include the extension of our results to population games with mean-field interactions, and averaging algorithms driven by Brownian motions.

References

[1] D. Bauso and G. Notarstefano. Distributed n-player approachability via time and space average consensus. In *3rd IFAC Workshop on Distributed Estimation and Control in Networked Systems*, pages 198–203, Santa Barbara, CA, USA, Sept 2012.

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

[3] Vincent D Blondel, Julien M Hendrickx, Alex Olshevsky, and John N Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *44th IEEE Conf. on Decision and Control, 2005*, pages 2996–3000, 2005.

[4] M. Bürger, G. Notarstefano, F. Allgöwer, and F. Bullo. A distributed simplex algorithm for degenerate linear programs and multi-agent assignments. *Automatica*, 48(9):2298–2304, Sept 2012.

[5] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

[6] Julien M Hendrickx. *Graphs and networks for the analysis of autonomous agent systems*. PhD thesis, Ecole Polytechnique, 2008.

[7] E. Lehrer. Allocation processes in cooperative games. *International J. of Game Theory*, 31:341–351, 2002.

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

[9] A. Nedić, A. Olshevsky, A. Ozdaglar, and J.N. Tsitsiklis. On distributed averaging algorithms and quantization effects. *IEEE Trans. Autom. Control*, 54(11):2506–2517, 2009.

[10] A. Nedić and A. Ozdaglar. Distributed subgradient methods for multi-agent optimization. *IEEE Trans. Autom. Control*, 54(1):48–61, 2009.

[11] A. Nedić, A. Ozdaglar, and P.A. Parrilo. Constrained consensus and optimization in multi-agent networks. *IEEE Trans. Autom. Control*, 55(4):922–938, 2010.

[12] G. Notarstefano and F. Bullo. Distributed abstract optimization via constraints consensus: Theory and applications. *IEEE Trans. on Automatic Control*, 56(10):2247–2261, October 2011.

[13] S. Sundhar Ram, A. Nedić, and V.V. Veeravalli. Incremental stochastic subgradient algorithms for convex optimization. *SIAM Journal on Optimization*, 20(2):691–717, 2009.

[14] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton Univ. Press, 1944.

APPENDIX

Proof of Lemma 1

By convexity of the distance function $\mathrm{dist}(\cdot, X)$ and from (4) we have

$$\mathrm{dist}(w_i(t), X) \leq \sum_{j=1}^{n} a_j^i(t)\mathrm{dist}(\hat{x}_j(t), X).$$

Summing over $i = 1, \ldots, n$ both sides of the above inequality we obtain

$$\sum_{i=1}^{n} \mathrm{dist}(w_i(t), X) \leq \sum_{i=1}^{n}\sum_{j=1}^{n} a_j^i(t)\mathrm{dist}(\hat{x}_j(t), X)$$

$$= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_j^i(t)\right)\mathrm{dist}(\hat{x}_j(t), X) = \sum_{j=1}^{n}\mathrm{dist}(\hat{x}_j(t), X),$$

where the last equality follows from the stochasticity of $A(t)$ in Assumption 1. This concludes the proof.

Proof of Lemma 2

Rearranging equation (5) we obtain

$$\|\hat{x}_i(t+1) - P_X[\hat{x}_i(t+1)]\|^2$$

$$-\frac{t^2}{(t+1)^2}\|w_i(t) - P_X[w_i(t)]\|^2 \leq$$

$$\frac{1}{(t+1)^2}\|x_i(t+1) - P_X[w_i(t)]\|^2$$

$$+\frac{2t}{(t+1)^2}(w_i(t) - P_X[w_i(t)])'(x_i(t+1) - P_X[w_i(t)]).$$

(6)

Note that the left hand side in (6) approximates $\text{dist}(\hat{x}_i(t+1), X)^2 - \text{dist}(w_i(t), X)^2$ for increasing t and also that for all t the left hand side upper bounds such a difference, i.e.,

$$\text{dist}(\hat{x}_i(t+1), X)^2 - \text{dist}(w_i(t), X)^2$$

$$\leq \text{dist}(\hat{x}_i(t+1), X)^2 - \frac{t^2}{(t+1)^2}\text{dist}(w_i(t), X)^2 \quad \forall t.$$

It remains to note that there exists a great enough scalar integer $\tilde{t}$ such that the left hand side in (6) is negative for all $t \geq \tilde{t}$. From the boundedness of set X and of vectors $x_i(t)$, there exists $M > 0$ such that $\|x_i(t+1) - P_X[w_i(t)]\|^2 < M$. Thus, we have

$$\text{dist}(\hat{x}_i(t+1), X)^2 - \text{dist}(w_i(t), X)^2$$

$$\leq \text{dist}(\hat{x}_i(t+1), X)^2 - \frac{t^2}{(t+1)^2}\text{dist}(w_i(t), X)^2$$

$$\leq \frac{1}{(t+1)^2}\big(\|x_i(t+1) - P_X[w_i(t)]\|^2$$

$$+2t(w_i(t) - P_X[w_i(t)])'(x_i(t+1) - P_X[w_i(t)])\big)$$

$$\leq \frac{1}{(t+1)^2}(M + 2t\phi) < 0$$

(7)

Taking $\tilde{t} > -M/2\phi > 0$ concludes the proof.

Proof of Theorem 1

Recall from (5) that

$$\|\hat{x}_i(t+1) - P_X[\hat{x}_i(t+1)]\|^2 \leq$$

$$\left(\frac{t}{t+1}\right)^2 \|w_i(t) - P_X[w_i(t)]\|^2$$

$$+ \left(\frac{1}{t+1}\right)^2 \|x_i(t+1) - P_X[w_i(t)]\|^2$$

$$+2\frac{t}{(t+1)^2}(w_i(t) - P_X[w_i(t)])'(x_i(t+1) - P_X[w_i(t)]).$$

From Lemma 1 and rearranging the above inequality, we have

$$\sum_{i=1}^{n} \left[(t+1)^2 \|\hat{x}_i(t+1) - P_X[\hat{x}_i(t+1)]\|^2 \right.$$

$$\left. -t^2 \|\hat{x}_i(t) - P_X[\hat{x}_i(t)]\|^2\right]$$

$$\leq \sum_{i=1}^{n} \left[\|x_i(t+1) - P_X[w_i(t)]\|^2 \right.$$

$$\left. +2t(w_i(t) - P_X[w_i(t)])'(x_i(t+1) - P_X[w_i(t)])\right]$$

$$\leq \sum_{i=1}^{n} \left[\|x_i(t+1) - P_X[w_i(t)]\|^2 \right.,$$

where the last inequality is due to Assumption 4. Summing over $t = 0, \ldots, \tau - 1$, and noting that $\|x_i(t+1) - P_X[w_i(t)]\|$ is bounded (from Assumption 3), so that the right hand side is upper bounded by some $M > 0$, we obtain

$$\sum_{i=1}^{n} \tau^2 \|\hat{x}_i(\tau) - P_X[\hat{x}_i(\tau)]\|^2 \leq M\tau$$

from which $\|\hat{x}_i(\tau) - P_X[\hat{x}_i(\tau)]\|^2 \leq \frac{M}{\tau}$, and therefore $\lim_{\tau \to \infty} \|\hat{x}_i(\tau) - P_X[\hat{x}_i(\tau)]\|^2 = 0$, which concludes the proof.

Proof of Lemma 3

To prove the statement observe that $\bar{x}_b(0) = \hat{x}_b(0) = x_b(0)$. Thus, we prove that $\bar{x}_b(t)$ and $\hat{x}_b(t)$ satisfy the same dynamics. By definition of time-average, $\bar{x}_b(t)$ satisfies the dynamics

$$\bar{x}_b(t+1) = \frac{t}{t+1}\bar{x}_b(t) + \frac{1}{t+1}x_b(t+1). \tag{8}$$

The dynamics of $\hat{x}_b(t)$ is

$$\frac{1}{n}\sum_{i=1}^{n}\hat{x}_i(t+1) = \frac{1}{n}\Big[\frac{t}{t+1}\sum_{i=1}^{n}\sum_{j=1}^{n}a_j^i(t)\hat{x}_j(t)$$

$$+\frac{1}{t+1}\sum_{i=1}^{n}x_i(t+1)\Big].$$

Exchanging the sum signs

$$\hat{x}_b(t+1) = \frac{1}{n}\frac{t}{t+1}\sum_{j=1}^{n}\sum_{i=1}^{n}a_j^i(t)\hat{x}_j(t) + \frac{1}{t+1}x_b(t+1),$$

and, by Assumption 1 ($A(t)$ is doubly stochastic),

$$\hat{x}_b(t+1) = \frac{1}{n}\frac{t}{t+1}\sum_{j=1}^{n}\hat{x}_j(t) + \frac{1}{t+1}x_b(t+1)$$

$$= \frac{t}{t+1}\hat{x}_b(t) + \frac{1}{t+1}x_b(t+1),$$

which is the same dynamics as (8), thus concluding the proof.

Proof of Theorem 2

Using the previous lemma we can show that $\hat{x}_i(t)$ converges to $\hat{x}_b(t)$. Let us introduce the error of the estimate $\hat{x}_i(t)$ from the barycenter, i.e. $\hat{e}_i(t) = \hat{x}_i(t) - \hat{x}_b(t)$. The error dynamics is given by

$$\hat{e}_i(t+1) = \frac{t}{t+1}\left[\sum_{j=1}^{n}a_j^i(t)\hat{e}_j(t) + \sum_{j=1}^{n}a_j^i\hat{x}_b(t)\right]$$

$$+\frac{1}{t+1}e_i(t+1) + \frac{1}{t+1}x_b(t+1)$$

$$-\frac{t}{t+1}\hat{x}_b(t) - \frac{1}{t+1}x_b(t+1),$$

where $e_i(t) = x_i(t) - x_b(t)$. Thus

$$\hat{e}_i(t+1) = \frac{t}{t+1}\Big(\sum_{j=1}^{n}a_j^i(t)\hat{e}_j(t)\Big) + \frac{1}{t+1}e_i(t+1).$$

Multiplying both sides by $(t+1)$ and taking t inside the sum,

$$(t+1)\hat{e}_i(t+1) = \sum_{j=1}^{n}a_j^i(t)t\hat{e}_j(t) + e_i(t+1).$$

Defining $\hat{z}_i(t) = t\,\hat{e}_i(t)$, we have

$$\hat{z}_i(t+1) = \sum_{j=1}^{n} a_j^i(t)\hat{z}_j(t) + e_i(t+1).$$

In vector form the above equation turns to be

$$\hat{z}(t+1) = \big(A(t) \otimes I_n\big)\hat{z}(t) + e(t+1), \tag{9}$$

with $\hat{z}(t) = [z_1(t) \ \dots \ z_n(t)]'$, $\hat{e}(t) = [e_1(t) \ \dots \ e_n(t)]'$, I_n the identity matrix of dimension n and $\otimes$ the Kronecker product. Notice that denoting $[\hat{z}]_\ell = \big[[\hat{z}_1]_\ell \ \dots \ [\hat{z}_n]_\ell\big]$ and $[e]_\ell = \big[[e_1]_\ell \ \dots \ [e_n]_\ell\big]$, $\ell \in \{1, \dots, n\}$, the dynamics of each $[\hat{z}]_\ell$ is given by

$$[\hat{z}]_\ell(t+1) = A(t)[\hat{z}]_\ell(t) + [e]_\ell(t+1). \tag{10}$$

Thus, we can simply work on each component separately. Slightly abusing notation we neglect the subscript of $[\hat{z}]_\ell$ and $[e]_\ell$, and write $\hat{z}(t)$ and $e(t)$.

It is worth noting that the driven system (10), and so (9), is *not* bounded-input-bounded-state stable (when a general input signal is allowed). That is, for general initial condition and input signal the state trajectory may diverge. We show that for the special initial condition ($\hat{z}(t) = 0$ by construction) and class of input signals ($\mathbf{1}'e(t+1) = 0$ by definition) under consideration, the state trajectories of (9) are bounded.

First, let us observe that, multiplying both sides of (9) by the vector $\mathbf{1}' = [1 \ \dots \ 1]$, we get

$$\begin{aligned} \mathbf{1}'\hat{z}(t+1) &= \mathbf{1}'A(t)\hat{z}(t) + \mathbf{1}'e(t+1) \\ &= \mathbf{1}'\hat{z}(t). \end{aligned} \tag{11}$$

Since $\hat{z}(0) = 0$ by construction, it holds $\mathbf{1}'\hat{z}(t) = 0$ for all $t \in \mathbb{N}$. That is, $\hat{z}(t)$ is orthogonal to the vector $\mathbf{1}$ for all t.

Next, we show that the trajectory $\hat{z}(\cdot)$ is bounded. Following [3], let $P \in \mathbb{R}^{(n-1)\times n}$ be a matrix defining an orthogonal projection onto the space orthogonal to span$\{\mathbf{1}\}$. It holds that $P\mathbf{1} = 0$ and $\|Px\|_2 = \|x\|_2$ if $x'\mathbf{1} = 0$. Thus, from equation (11) we have that $\|P\hat{z}(t)\|_2 = \|\hat{z}(t)\|_2$ for all t. Therefore, proving boundedness of $\hat{z}(\cdot)$ is equivalent to showing that $P\hat{z}(\cdot)$ is bounded. For a given P, associated to any $A(t)$ satisfying Assumption 1, there exists $\bar{A}(t)$ satisfying $PA(t) = \bar{A}(t)P$. The spectrum of $\bar{A}(t)$ is the spectrum of $A(t)$ after removing the eigenvalue 1. Multiplying both sides of equation (9) by P, we get

$$\begin{aligned} P\hat{z}(t+1) &= PA(t)\hat{z}(t) + Pe(t+1) \\ &= \bar{A}(t)P\hat{z}(t) + Pe(t+1). \end{aligned} \tag{12}$$

Under Assumptions 1 and 2, the undriven dynamics $y(t + 1) = \bar{A}(t)y(t)$ is uniformly exponentially stable, i.e., $\|y(t)\| < C\rho^t\|y(0)\|$ with C and $\rho < 1$ independent of $y(0)$ and depending only on n, Q and α (see Theorem 9.2 and Corollary 9.1 in [6]). Thus, the state trajectories of (12) are bounded for any bounded signal $Pe(t + 1)$ with $\mathbf{1}'e(t) = 0$. Since $\mathbf{1}'e(t) = 0$ for all t, we have $\|Pe(t)\|_2 = \|e(t)\|_2$ for all t, which is bounded. The proof follows by recalling that $\|P\hat{z}(t)\|_2 = \|\hat{z}(t)\|_2$ and that $\hat{z}(t) = t\hat{e}(t)$.

Proof of Theorem 3

From (5), invoking Lemma 1 and using Assumption 5 we have

$$\sum_{i=1}^{n} \big[(t+1)^2\|\hat{x}_i(t+1) - P_X[\hat{x}_i(t+1)]\|^2$$

$$-t^2\|\hat{x}_i(t) - P_X[\hat{x}_i(t)]\|^2\big]$$

$$\leq \sum_{i=1}^{n} \Big[\|x_i(t+1) - P_X[w_i(t)]\|^2$$

$$+2t(w_i(t) - P_X[w_i(t)])'(x_i(t+1) - \mathbb{E}x_i(t+1)) \Big],$$

Summing over $t = 0, \ldots, \tau - 1$, and noting that $\|x_i(t + 1) - P_X[w_i(t)]\|$ is upper bounded (from Assumption 3) by some $M > 0$, we obtain

$$\sum_{i=1}^{n} \|\hat{x}_i(\tau) - P_X[\hat{x}_i(\tau)]\|^2$$

$$\leq \frac{M}{\tau} + \frac{1}{\tau}\sum_{t=0}^{\tau-1}\sum_{i=1}^{n} K_t^i\|x_i(t+1) - \mathbb{E}x_i(t+1)\|$$

where $K_t^i = \frac{1}{\tau}2t\|w_i(t) - P_X[w_i(t)]\|$. Now, using $\|x_i(t + 1)\| \leq L \ \forall t \geq 0$ from Assumption 5 and from (3) and (4) we have that $w_i(t)$ is bounded which in turn implies that $\|w_i(t) - P_X[w_i(t)]\|$ is bounded. Then, the second term in the right-hand side is an average of bounded zero-mean martingale differences, and therefore the Hoeffding-Azuma inequality (together with the Borel-Cantelli lemma) immediately implies that

$$\lim_{\tau\to\infty} \sum_{i=0}^{n} \|\hat{x}_i(\tau) - P_X[\hat{x}_i(\tau)]\|^2 = 0$$

which concludes the proof.