Distributed n-player approachability and consensus in coalitional games

Dario Bauso and Giuseppe Notarstefano

***Abstract*—We study a distributed allocation process where, at each time, every player i) proposes a new bid based on the average utilities produced up to that time, ii) adjusts such allocations based on the inputs received from its neighbors, and iii) generates and allocates new utilities. 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 makes the grand coalition stable.**

I. Introduction

We consider a two-step distributed allocation process where, at every time, the players first adjust their average allocation vectors based on the inputs received from their *neighbor players* and second generate a new utility 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 target set even in the presence of adversarial disturbances.

Motivations. The problem arises in the context of dynamic coalitional games with Transferable Utilities (TU games) [**?**]. 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 players' payoffs. By payoff we mean the share allocated to the player. Therefore, no coalition has incentives to leave the grand coalition and receive a larger payoff.

Highlights of contributions. We analyze conditions under which the average allocations: (i) *approach* a given target set (Theorem 1), (ii) reach *consensus*, in which case we also compute the consensus value (Theorem 2), and (iii) are robust against disturbances (Theorem 3). Validation of such conditions implies the solution of a quadratic program (projection on a convex set) and a feasibility linear program (find at least a feasible solution of a set of inequalities).

Related literature. TU games were first introduced by von Neumann and Morgenstern [**?**]. Here, a main issue is to study whether the core is an "approachable" set, and which allocation processes can lead to stable payoff allocations in that set. Approachability theory was developed by Blackwell in the early '56 in [**?**], and is captured in the well-known Blackwell's Theorem. The geometric (approachability) principle upon which the Blackwell's Theorem is built, is used in several application domains such as allocation processes ([**?**], Section 3.2), regret minimization ([**?**], Equation C.1), and machine learning (see [**?**] Section 2).

The discrete-time dynamics considered in this paper involves a distributed averaging process (see, e.g., [**?**] and references therein) and this is an element in common with distributed multi-agent optimization [**?**], [**?**], [**?**], [**?**], [**?**], [**?**].

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

Notation. For a vector x, we use x_j or $[x]_j$ to denote its jth 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^i_j and $\sum_{j=1}^n a^i_j = 1$ for all $i = 1, \ldots, n$. For a matrix A, we use a^i_j to denote its ijth entry. A matrix A is doubly stochastic if both A and its transpose A' are row-stochastic. 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 $\mathrm{dist}(x, X)$ for the distance from x to X, i.e., $P_X[x] = \arg\min_{y \in X} \|x - y\|$ and $\mathrm{dist}(x, X) = \|x - P_X[x]\|$, 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 utility allocation algorithm

Every player in a set $N = \{1, \ldots, n\}$ proposes a distribution of the utilities which is given by the average allocation vector $\hat{x}_i(t+1) \in \mathbb{R}^n$. The jth component of $\hat{x}_i(t+1)$ defines the share that player i would allocate to player j on average up to time $t+1$. At every time, each player first adjusts its average allocation vector based on the inputs received from its *neighbor players* and then generates a new allocation vector $x_i(t+1)$.

Let a *communication graph* $\mathcal{G}(t) = (N, \mathcal{E}(t))$ be given. A link $(j, i) \in \mathcal{E}(t)$ exists if player j is a neighbor of player i at time t. Each player adjusts its average allocation vector

A preliminary conference version of this paper has appeared as [**?**]. The current paper includes, in addition: i) more detailed and revised proofs of the main results, ii) analysis of adversarial disturbances; iii) analysis of the connections with approachability theory in its strategic version for two-player repeated games, and iii) numerical studies. The work of D. Bauso was supported by PRIN 20103S5RN3 "Robust decision making in markets and organizations, 2013-2016". The second author wants to thank J. Hendrickx for the helpful discussion on the proof of Theorem 2.

D. Bauso is with DICGIM, Università di Palermo, Italy, email: *dario.bauso@unipa.it*.

G. Notarstefano is with Department of Engineering, Università del Salento (University of Lecce), Via per Monteroni, 73100, Lecce, Italy *giuseppe.notarstefano@unisalento.it*

so that the new current average allocation, we call it *space average* and denote it by $w_i(t)$, is in the convex hull of its neighbors' average allocations, i.e.,

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

where $a^i = (a_1^i, \ldots, a_n^i)'$ is a vector of nonnegative weights consistent with the sparsity of $\mathcal{G}$, i.e., $a_j^i(t) \neq 0$ if and only if $(j,i) \in \mathcal{E}(t)$. Thus, the time-averaged allocation $\hat{x}_i(t)$ evolves according to

$$\begin{aligned}\hat{x}_i(t+1) &= \frac{t}{t+1}\left[\sum_{j=1}^{n} a_j^i(t)\hat{x}_j(t)\right] + \frac{1}{t+1}x_i(t+1) \\ &= \frac{t}{t+1}w_i(t) + \frac{1}{t+1}x_i(t+1).\end{aligned} \quad (2)$$

Problem. Our goal is to study under what conditions the average allocation vectors converge to a unique value lying in a predefined target set X: for all $i, j \in N$,

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

A. Main assumptions

Following [?] (see also [?]) we can make the following assumptions on the information structure. 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 $a_j^i(t) > 0$.*

In addition to this, 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 joint strong connectivity is the weakest possible assumption to guarantee persistent circulation of the information through the graph.

The above assumptions imply that the weights of the communication graph are determined exogenously but accordingly to some predefined rules.

The following assumption characterizes the target set X.

Assumption 3. *The target set X is nonempty, convex and compact.*

Finally, the next assumption indicates how the new utility vector has to be generated in order to obtain approachability, i.e., convergence of the average allocations to the target set.

Assumption 4. *For each $i \in N$ the new utility vector 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,*

$$\left(w_i(t) - P_X[w_i(t)]\right)'\left(x_i(t+1) - P_X[w_i(t)]\right) \leq 0.$$

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 utility vector $x_i(t+1)$ lies in the half-space not containing $w_i(t)$.

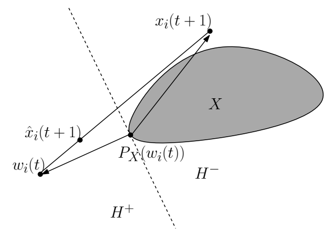

Fig. 1. Approachability principle.

B. Motivating example: coalitional game

The set X introduced above can be thought of as the core of 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.

Consistently with Assumptions 3, a common assumption in TU games is that the core set is nonempty, convex and bounded, and the utilities generated at each time are bounded, [?], [?]. The core set is usually assumed to be known by the players. Whenever the core set is empty and not known by all the players, the grand coalition is not stable and the players tend to form sub-coalitions according to more or less complex bargaining processes. Coalition formation [?] is a separate and independent body of literature, which is far from the scope of this paper. Further, in approachability theory, given a two-player repeated game with vector payoffs, a condition similar to the ones stated in Assumptions 4 and 5, is used to prove that a given set is approachable by player 1, independently of the strategy used by player 2. Such condition have also been shown to be sufficient condition for the existence of attractors in nonlinear analysis and in the theory on differential inclusion (see also [?], Corollary 5.1). We borrow such a condition for design purposes. In other words we wish to design our distributed algorithm in order to let the players reach consensus. This is in accordance with the idea of *mechanism design*, where a central planner designs incentives for the player to cooperate.

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). To improve the readability of the paper all the proofs are collected in Appendix section at the end of the paper.

A. Approachability and consensus

Before stating the first theorem, we need to introduce two lemmas. The first lemma establishes that the space averaging step in (2) reduces the total distance (i.e. the sum of distances) of the average allocations 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.,*

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

Proof. See the Appendix for a proof. □

Observing that the distance of a point from a convex set is equal to the distance from its projection (which is by definition smaller than the distance from any other point in the set) and using (2) and (1), it holds

$$\begin{aligned}
\operatorname{dist}(\hat{x}_i(t+1), X)^2 &= \|\hat{x}_i(t+1) - P_X[\hat{x}_i(t+1)]\|^2 \\
&\leq \|\hat{x}_i(t+1) - P_X[w_i(t)]\|^2 \\
&= \Big\| \frac{t}{t+1}\left(w_i(t) - P_X[w_i(t)]\right) \\
&\quad + \frac{1}{t+1}\left(x_i(t+1) - P_X[w_i(t)]\right) \Big\|^2 \\
&= \left(\frac{t}{t+1}\right)^2 \|w_i(t) - P_X[w_i(t)]\|^2 \\
&\quad + \left(\frac{1}{t+1}\right)^2 \|x_i(t+1) - P_X[w_i(t)]\|^2 \\
&\quad + \frac{2t}{(t+1)^2}(w_i(t) - P_X[w_i(t)])'(x_i(t+1) - P_X[w_i(t)]).
\end{aligned} \tag{3}$$

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.,*

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

Proof. See the Appendix for a proof. □

The proof of Theorem 1 is constructive in the sense that it provides also a guideline on how to select the new iterate so that the conditions in Assumption 4 are satisfied. In particular, the new iterate is obtained by projecting the current point on the approachable set, by identyifing the supporting hyperplane and by selecting a point on the opposite half-space than the one containing the current point. All steps involve solving convex programs or linear inequalities. s

Next, let us introduce the barycenter of the average allocations and the utility vectors respectively

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

Consistently, let us denote by $\bar{x}_b(t)$ 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 average allocations evolves as the time average $\bar{x}_b(t)$ of the barycenter of the utility vectors generated by the players.

Lemma 2. *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 utility vectors $\bar{x}_b(t)$.*

Proof. See the Appendix for a proof. □

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 utility vectors generated by each player, $\bar{x}_b(t)$, i.e.,*

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

Proof. See the Appendix for a proof. □

Summarizing the two main results, we have proven that the players' allocations converge asymptotically to the time-average of the barycenter of the generated utility 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 by

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

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 $\{x \in \mathbb{R}^n \colon \lambda' x \geq 0\}$ (respectively half-space $\{x \in \mathbb{R}^m \colon \lambda' x > 0\}$).

Let us adapt Assumption 4 to the worst-case setting introduced in this section.

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 utility vector is bounded, i.e. there exists $L > 0$ s.t. $\forall t \geq 0$ $\|x_i(t+1)\| \leq L$, and satisfies*

$$(w_i(t) - P_X[w_i(t)])' \left(\mathbb{E}x_i(t+1) - P_X[w_i(t)]\right) \leq 0,$$

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., [?] and also [?], chapter 7). Here we adapt the above condition to the multi-agent and distributed scenario under study.

Corollary III.1 (see [?], 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.,*

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

We conclude this section by observing that Theorem 2 still holds and therefore all players' average allocations reach consensus on the time-average of the barycenter of the utility 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).[1]

[1]The values that are not specified are assumed to be irrelevant for the core definition. For example $\eta_{\{2,3\}} \leq 4$.

That is, each player expects to receive at least a utility of 2 which is its own value. 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

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

We initialize the assignments assuming that each player allocates to itself the entire utility. 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 utility vector $x_i(t+1)$ according to the approachability principle. In particular, we set $x_i(t+1) = P_X[w_i(t)] + \alpha\,(P_X[w_i(t)] - w_i(t)) + v^\top$, where α is a random number uniformly distributed in $[0,1]$ and $v^\perp$ a random vector belonging to the hyperplane orthogonal to the vector $w_i(t) - P_X[w_i(t)]$ with coordinates (with respect to the basis vectors) uniformly chosen in $[0,1]$. The temporal evolution of the local average allocation vectors is depicted in Figure 2. As expected the local average allocations converge to the same average assignment which is the point of the core $[2\ 3\ 2.5\ 2.5]'$.

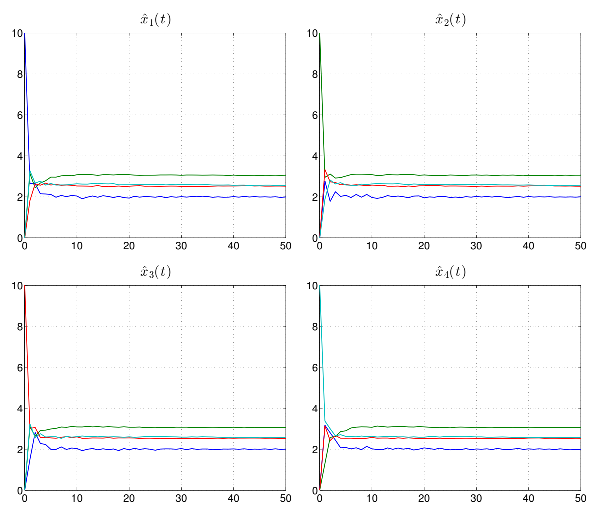

Fig. 2. Local average allocation vectors (first 50 time-instants)

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.

Appendix

Proof of Lemma 1

By convexity of the distance function $\mathrm{dist}(\cdot, X)$ and from (1) 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 $i = 1, \dots, n$ both sides of the above inequality we obtain

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

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

Proof of Theorem 1

Recall from (3) that

$$\|\hat{x}_i(t+1) - P_X[\hat{x}_i(t+1)]\|^2 \le \left(\frac{t}{t+1}\right)^2 \|w_i(t) - P_X[w_i(t)]\|^2 + \left(\frac{1}{t+1}\right)^2 \|x_i(t+1) - P_X[w_i(t)]\|^2 + 2\frac{t}{(t+1)^2}(w_i(t) - P_X[w_i(t)])'(x_i(t+1) - P_X[w_i(t)]).$$

From Lemma 1 and rearranging the above inequality, we have

$$\sum_{i=1}^{n}\left[(t+1)^2\|\hat{x}_i(t+1) - P_X[\hat{x}_i(t+1)]\|^2 - t^2\|\hat{x}_i(t) - P_X[\hat{x}_i(t)]\|^2\right]$$
$$\le \sum_{i=1}^{n}\Big[\|x_i(t+1) - P_X[w_i(t)]\|^2 + 2t(w_i(t) - P_X[w_i(t)])'(x_i(t+1) - P_X[w_i(t)])\Big]$$
$$\le \sum_{i=1}^{n}\|x_i(t+1) - P_X[w_i(t)]\|^2,$$

where the last inequality is due to Assumption 4. The right hand side $\sum_{i=1}^{n}\|x_i(t+1) - P_X[w_i(t)]\|^2$ is bounded (from Assumption 3 and the boundedness of $x_i(t+1)$) by some $M > 0$. Summing over $t = 0, \dots, \tau - 1$, and noting that $\sum_{t=0}^{\tau-1} g(t+1) - g(t) = g(\tau)$ with $g(t) = \sum_{i=1}^{n} t^2\|\hat{x}_i(t) - P_X[\hat{x}_i(t)]\|^2$, we obtain

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

from which $\|\hat{x}_i(\tau) - P_X[\hat{x}_i(\tau)]\|^2 \le \frac{M}{\tau}$, 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 2

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

$$\bar{x}_b(t+1) = \frac{t}{t+1}\bar{x}_b(t) + \frac{1}{t+1}x_b(t+1). \tag{4}$$

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 (4), 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 average allocation $\hat{x}_i(t)$ 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

$$\hat{z}(t+1) = \big(A(t)\otimes I_n\big)\hat{z}(t) + e(t+1), \tag{5}$$

with $\hat{z}(t) = [z_1(t)\ \dots\ z_n(t)]'$, $e(t) = [e_1(t)\ \dots\ e_n(t)]'$, 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

$$[\hat{z}]_\ell(t+1) = A(t)[\hat{z}]_\ell(t) + [e]_\ell(t+1). \tag{6}$$

Thus, we can simply work on each component separately. With a slight abuse of 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 (6), and so (5), 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 (5) are bounded.

First, let us observe that, multiplying both sides of (5) by the vector $\mathbf{1}' = [1 \ \dots \ 1]$, we get

$$\begin{aligned} \mathbf{1}'\hat{z}(t+1) &= \mathbf{1}'A(t)\hat{z}(t) + \mathbf{1}'e(t+1) \\ &= \mathbf{1}'\hat{z}(t). \end{aligned} \tag{7}$$

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 [**?**], let $P \in \mathbb{R}^{(n-1)\times n}$ be a matrix defining an orthogonal projection on 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 (7) 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 $PA(t) = \bar{A}(t)P$. Multiplying both sides of equation (5) by P, we get

$$\begin{aligned} P\hat{z}(t+1) &= PA(t)\hat{z}(t) + Pe(t+1) \\ &= \bar{A}(t)P\hat{z}(t) + Pe(t+1). \end{aligned} \tag{8}$$

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 [**?**]). Thus, the state trajectories of (8) 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 from noting that $\|P\hat{z}(t)\|_2 = \|\hat{z}(t)\|_2$ and that $\hat{z}(t) = t\hat{e}(t)$.

Proof of Theorem 3

From (3), invoking Lemma 1 and using Assumption 5 we have

$$\begin{aligned} &\sum_{i=1}^{n} \left[(t+1)^2\|\hat{x}_i(t+1) - P_X[\hat{x}_i(t+1)]\|^2 \right. \\ &\qquad \left. -t^2\|\hat{x}_i(t) - P_X[\hat{x}_i(t)]\|^2\right] \\ &\leq \sum_{i=1}^{n} \Big[\|x_i(t+1) - P_X[w_i(t)]\|^2 \\ &+2t(w_i(t) - P_X[w_i(t)])'(x_i(t+1) - \mathbb{E}x_i(t+1))\Big]. \end{aligned}$$

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

$$\begin{aligned} &\sum_{i=1}^{n} \|\hat{x}_i(\tau) - P_X[\hat{x}_i(\tau)]\|^2 \\ &\leq \frac{M}{\tau} + \frac{1}{\tau}\sum_{t=0}^{\tau-1}\sum_{i=1}^{n} K_t^i\|x_i(t+1) - \mathbb{E}x_i(t+1)\|, \end{aligned}$$

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 (2) and (1) 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.