University of Groningen

Recurrent averaging inequalities in multi-agent control and social dynamics modeling

Proskurnikov, Anton; Calafiore, Giuseppe; Cao, Ming

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

Link to publication in University of Groningen/UMCG research database

Recurrent Averaging Inequalities
in Multi-Agent Control and Social Dynamics Modeling

Anton V. Proskurnikov[a,b,1,*], Giuseppe Calafiore[a], Ming Cao[c]

[a]*Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy*
[b]*Institute for Problems of Mechanical Engineering of the Russian Academy of Sciences (IPME RAS), St. Petersburg, Russia*
[c]*Faculty of Science and Engineering, University of Groningen, Groningen, The Netherlands*

Abstract

Many multi-agent control algorithms and dynamic agent-based models arising in natural and social sciences are based on the principle of *iterative averaging.* Each agent is associated to a value of interest, which may stand, e.g., for an opinion of an individual in a social group, velocity vector of a mobile robot in a flock or measurement of a sensor within a sensor network. This value is updated, at each iteration, to a convex combination of itself and the values of the adjacent agents. It is well known that, under natural assumptions on the network's graph connectivity, this local averaging procedure eventually leads to global consensus, or synchronization of the values at all nodes. Applications of iterative averaging include, but are not limited to, algorithms for distributed optimization, for solution of linear and nonlinear equations, for multi-robot coordination and for opinion formation in social groups. Although these algorithms have similar structures, mathematical techniques used for their analysis are manifold, and the conditions for their convergence and stability differ from case to case. In this review paper, we show that the properties of such algorithms can be analyzed in a unified way by using a novel tool based on recurrent averaging inequalities (RAI). We develop a theory of RAI and apply it to the analysis of several important multi-agent algorithms recently proposed in the literature.

Keywords: Distributed algorithm, opinion dynamics, multi-agent control, consensus

1. Introduction

In recent decades, problems of distributed control and coordination in multi-agent systems (MAS) (Mesbahi and Egerstedt, 2010; Shamma, 2007; Ren and Beard, 2008; Ren and Cao, 2011; Lewis et al., 2014) have attracted enormous attention of the research community. The foundational problem of MAS theory is how a complex cooperative goal (coordinated motion such as flocking or swarming, uniform deployment over some area, collective decision making) can be reached by a team of (relatively) simple autonomous agents. Usually, agents have restricted computational capabilities and limited information about the team as a whole; each agent can interact only to a few neighboring agents. The neighborhood relations are conveniently represented by the interaction graph, or "topology" of the multi-agent system. Agent-based models help to study complex phenomena in nature and society (Leonard, 2014; Proskurnikov and Tempo, 2017; Bullo, 2016).

Many multi-agent algorithms considered in the literature are based, explicitly or implicitly, on the idea of *iter-*

ative averaging. Discrete-time dynamics of iterative averaging date back to the seminal model of opinion diffusion in social groups introduced by French Jr. (1956) and thoroughly studied in (Harary, 1959; Harary et al., 1965). In French's model, each agent (social actor) corresponds to a node of a graph and keeps an *opinion* – a real number standing for some quantity of interest or cognitive orientation towards some object (Friedkin, 2015). At each step of the opinion iteration, the actors simultaneously display their opinions to the neighbors in the graph and update their opinions based on the information displayed to them: a new value of an actor's opinion is computed as the mean of its previous value and the opinions displayed by the neighbors (Fig. 1). French's model can be written as

$$x_i(k+1) = \sum_{j=1}^{n} w_{ij} x_j(k), \ i = 1, \ldots, n; \ k = 0, 1, \ldots \quad (1)$$

where $x_i(k)$ denotes the opinion of actor i at time k, and w_{ij} may be interpreted as the relative strength of influence of agent j on agent i (the diagonal entries measure self-influence of the agents). The weights w_{ij} are nonnegative and $\sum_{j=1}^{n} w_{ij} = 1$ for all i. As shown in (Harary, 1959; Harary et al., 1965), typically the iteration of process (1) establishes eventual consensus (unanimity) of opinions.

Collecting the weights in a (row-stochastic) matrix $W = (w_{ij})$, and defining the vector $x(k) \in \mathbb{R}^n$ of opinions at

---

[*]Corresponding author.
Email addresses: `anton.p.1982@ieee.org` (Anton V. Proskurnikov), `giuseppe.calafiore@polito.it` (Giuseppe Calafiore), `m.cao@rug.nl` (Ming Cao)
[1]Some results of the paper were reported (Proskurnikov and Cao, 2017b) on IEEE Conference on Decision and Control CDC 2017, Melbourne, Australia.
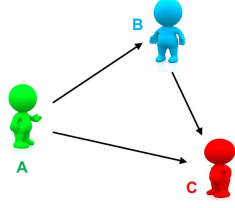
Figure 1: French's model of opinion formation in a group of $n = 3$ actors. A's opinion is displayed to B and C, B's opinion is displayed to C but is hidden from A, C's opinion is displayed to nobody. The corresponding dynamical system obeys the equations $x_A(k+1) = x_A(k)$, $x_B(k+1) = \frac{x_A(k) + x_B(k)}{2}$, $x_C(k+1) = \frac{x_A(k) + x_B(k) + x_C(k)}{3}$.

time k, we rewrite model (1) in vector format as

$$x(k+1) = Wx(k), \quad k = 0, 1, \ldots. \qquad (2)$$

Model (2), with a general stochastic matrix W, was later independently proposed by DeGroot (1974) and Lehrer (1976) as a method for reaching rational agreement in expert communities. Consensus in (2) is equivalent to the "full regularity" (Gantmacher, 2000), or SIA (stochastic, indecomposable, aperiodic) property (Wolfowitz, 1963) of the matrix W or, equivalently, regularity of the Markov chain generated by this matrix (Kemeny and Snell, 1976). In particular, the iterations in (2) always reach consensus if W is aperiodic and irreducible (or primitive) (DeGroot, 1974; Gantmacher, 2000). In graph-theoretic language, the necessary and sufficient criterion of consensus is the existence of an agent that communicates (directly or indirectly) to all other agents, that is, the graph has a directed spanning tree (Ren and Beard, 2008; Bullo, 2016).

The behavior of iterative averaging algorithms such as (2), however, is more complicated in the case of *dynamic* interaction graphs, where the matrix of influence weights $W = W(k)$ varies with time

$$x(k+1) = W(k)x(k), \quad k = 0, 1, \ldots. \qquad (3)$$

Although such models were studied in relation to non-stationary Markov processes (Dobrushin, 1956; Hajnal, 1958; Wolfowitz, 1963; Seneta, 1981), the explosive growth of interest in them has been driven by agent-based models of synchronous collective motion based on the "nearest neighbor" rules (Reynolds, 1987; Vicsek et al., 1995; Olfati-Saber, 2006; Lin et al., 2007). In such models, the interactions between the agents are range-restricted, and the set of an agent's neighbors evolves as the agents move in the space. Consensus in the time-varying algorithm (3) is equivalent to the ergodicity[2] of the backward infinite matrix products $W(k) \ldots W(0)$ Seneta (1981). Despite the interest and efforts devoted to the problem of matrix products ergodicity in the literature on probability theory and matrix analysis (Seneta, 1981; Leizarowitz, 1992), a complete solution to this problem remains elusive, and a gap

between necessary and sufficient ergodicity conditions exists (Touri and Nedic, 2012; Bolouki and Malhame, 2016).

In recent years, the iterative averaging model has attracted enormous attention from the research community as a simple algorithm for multi-agent coordination (Jadbabaie et al., 2003; Blondel et al., 2005; Moreau, 2005; Olfati-Saber et al., 2007; Mesbahi and Egerstedt, 2010; Ren and Beard, 2008; Ren and Cao, 2011; Cao et al., 2008). Besides the important problem of multi-agent consensus, procedures of iterative averaging lie at the heart of many distributed numerical algorithms (D.Bertsekas and Tsitsiklis, 1989) such as, e.g., techniques of distributed estimation and filtering (Borkar and Varaiya, 1982; Garin and Schenato, 2010; Bullo, 2016), deterministic and stochastic optimization (Tsitsiklis et al., 1986; Nedic et al., 2010; Lin and Ren, 2014; Yang et al., 2019), load balancing (Amelina et al., 2015), and solving systems of linear equations (Liu et al., 2014; Mou et al., 2015; You et al., 2016; Wang et al., 2019). Many models of opinion formation (Proskurnikov and Tempo, 2017, 2018) in social groups stem from the classical French-DeGroot model (2).

While the mentioned algorithms and dynamical models have similar structures, their properties visibly differ, as well as mathematical techniques used for their analysis. For instance, under some natural assumptions, the standard condition for consensus in (2) is the existence of a directed spanning tree in the communication graph (Ren and Cao, 2011; Cao et al., 2008), known also as the quasi-strong connectivity. At the same time, similar algorithms for *constrained* consensus (Nedic et al., 2010) and linear equations solving (Mou et al., 2015; You et al., 2016) that are also based on iterative averaging, require strongly connected graphs. Their convergence properties cannot be directly derived from properties of the standard consensus algorithm, and require additional tools. To the best of the authors' knowledge, there is no unified mathematical theory for averaging-based multi-agent algorithms.

1.1. Systems of recurrent averaging inequalities

In this review, we demonstrate that many of the aforementioned averaging-based algorithms can be examined by means of a novel tool that we call a *recurrent averaging inequality* (RAI), which is defined as an inequality relaxation of the (possibly time-varying) iterations in (2), namely,

$$x(k+1) \leq W(k)x(k), \quad k = 0, 1, \ldots, \qquad (4)$$

where matrices $W(k)$ are row-stochastic. At a first glance, the system of inequalities (4) is too "loose" to entail any interesting properties of the sequence $\{x(k)\}$. However, under rather modest conditions of connectivity (for instance, if every matrix $W(k) + \ldots + W(k+T-1)$, where $k \geq 0$ and T is a fixed period, corresponds to a strongly connected graph) the inequality (4) implies asymptotic consensus of the opinions $x_i(k)$ (which, however, can be achieved at $-\infty$). In many situations where consensus

---

[2] The backward products of stochastic matrices $W(k) \ldots W(0)$ are *ergodic* if they converge, as $k \to \infty$, to a stochastic rank-one matrix.

is not established, (4) provides convergence of the sequence $\{x(k)\}$. Similar properties have been recently obtained (Proskurnikov and Cao, 2017a) for inequalities

$$\dot{x}(t) \leq -L(t)x(t), \qquad (5)$$

where $L(t)$ stands for a time-varying Laplacian matrix. The theory developed in Proskurnikov and Cao (2017a) is not directly applicable to discrete-time inequalities (4), although the proofs of the main results are similar in spirit.

1.2. Contribution and paper organization

In this paper we develop a mathematical theory for the RAI (4). In particular, we establish convergence and consensus criteria for any feasible sequence $\{x(k)\}$. Further, we show that a number of multi-agent algorithms and opinion formation models can be examined in a unified way using the RAI theory, which allows, in particular, to examine Hegselmann-Krause model with informed ("truth-seeking") agents Hegselmann and Krause (2006), and to generalize the recent fundamental results Fullmer and Morse (2018) on distributed algorithms that compute a common fixed point for a family of paracontractions (solving thus a special system of nonlinear equations).

The paper is organized as follows: Section 2 introduces the notation and concepts from graph theory. Section 3 recapitulates some classical results concerned with consensus in iterative averaging procedures. Section 4 presents main results of this paper, establishing convergence of solutions of RAI (4). In Section 5, we illustrate them by applications to some models of opinion formation and distributed algorithms. Section 6 contains proofs of the main results. Section 7 concludes the paper.

2. Preliminaries and notation

For positive integers m, n, with $n \geq m$, denote $[m : n] \triangleq \{m, m+1, \ldots, n\}$. The cardinality of a finite set I is denoted by $|I|$. We use $\mathbb{1}_n = (1, 1, \ldots, 1)^\top \in \mathbb{R}^n$ to denote a (column) vector of all ones, and $\mathbf{e}_1 = (1, 0, \ldots, 0)^\top, \ldots, \mathbf{e}_n = (0, 0, \ldots, 1)^\top$ to denote the basis of the unit coordinate vectors. For two vectors $x, y \in \mathbb{R}^n$ we write $x \leq y$ if $x_i \leq y_i \, \forall i$ and $x \leq 0$ if $x_i \leq 0 \, \forall i$; the reverse inequalities $\geq$ are defined analogously. We define the sign of a real number as

$$\text{sgn}(t) = \begin{cases} 1, & t > 0 \\ -1, & t < 0 \\ 0, & t = 0. \end{cases}$$

The spectral radius of a matrix A is denoted by $\rho(A)$.

A graph is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a finite set referred to as the set of *nodes* and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is a set of *arcs*. The arc (i, j) is also denoted by $i \rightarrow j$. We call a graph *undirected* if $\mathcal{E}$ is symmetric in the sense that $(i, j) \in \mathcal{E} \Leftrightarrow (j, i) \in \mathcal{E}$, otherwise the graph is *directed*.

A sequence of arcs $i_0 \rightarrow i_1 \rightarrow i_2 \rightarrow \ldots \rightarrow i_s$ is called a *walk* connecting i_0 to i_s, the number of arcs s is the walk's *length*. A walk that starts and ends at the same node $i_0 = i_s$ is said to be a *cycle*. A graph is said to be *periodic* if an integer $p > 1$ exists that divides the length of any cycle; otherwise the graph is *aperiodic*. A graph is *strongly connected* if every two distinct nodes are connected by a walk. A graph is *quasi-strongly connected*, or has a directed (out-branching) *spanning tree* (Ren and Beard, 2008) if some node ("root") is connected by walks to all other nodes. For undirected graphs, strong and quasi-strong connectivity are equivalent (such a graph is said to be *connected*).

A graph $(\mathcal{V}', \mathcal{E}')$, where $\mathcal{V}' \subseteq \mathcal{V}$ and $\mathcal{E}' \subseteq \mathcal{E}$ is referred to as a *subgraph* of the graph $(\mathcal{V}, \mathcal{E})$. A subgraph is said to be a *strongly connected* (or simply *strong*) *component* if it is strongly connected and maximal in the sense that no other node or arc can be added to it without destroying the subgraph's strong connectivity. Each node of $\mathcal{G}$ belongs to at most one strong component. If $\mathcal{G}$ is strongly connected, then it has only one strong component ($\mathcal{G}$ itself). Otherwise, it contains multiple strong components, among which at least one component is a *source component* (no arcs enter it) and at least one component is a *sink component* (no arcs leave it), see Fig. 2. A graph is quasi-strongly connected if and only if it has a single source component. A strong component can be *isolated*, i.e., to have neither incoming nor outcoming arcs and thus be both a source and a sink. Strong components of undirected graphs are always isolated.
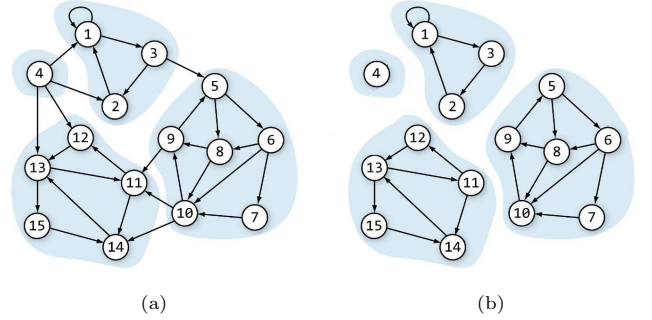
Figure 2: Strong components of a directed graph: (a) non-isolated; (b) isolated. In (a), $\{4\}$ is a single source component, $\{11, \ldots, 15\}$ is a single sink component.

An arbitrary matrix $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ can be associated with a *signed* weighted graph, being the triple $\mathcal{G}[A] = ([1 : n], E[A], A)$, where $[1 : n]$ is the set of nodes, $E[A] \triangleq \{(j, i) : a_{ij} \neq 0\}$ is the set of arcs and a_{ij} stands for the (signed) weight or value of arc[3] (j, i). In this paper, we mainly deal with graphs generated by non-negative

---

[3]In the models considered below, entry a_{ij} usually quantifies the strength of influence agent j exerts on agent i. Following the tradition of multi-agent systems theory (Ren and Beard, 2008; Ren and Cao, 2011) and the original work by French Jr. (1956), such an influence is depicted by an arc $j \rightarrow i$ rather than $i \rightarrow j$.

matrices: in such a situation, the entry of a matrix is considered as a weight (or value) of the corresponding arc. For matrices $B_1, B_2 \in \mathbb{R}^{n \times n}$ being non-negative, one has $\mathcal{E}[B_1 + B_2] = \mathcal{E}[B_1] \cup \mathcal{E}[B_2]$; the resulting graph $\mathcal{G}[B_1 + B_2]$ is called the *union* of the graph $\mathcal{G}[B_1]$ and $\mathcal{G}[B_2]$. A non-negative matrix B is *irreducible* if $\mathcal{G}[B]$ is strongly connected and *aperiodic* if $\mathcal{G}[B]$ is aperiodic.

The subdivision of nodes into two non-empty disjoint sets I and $J = I^c \triangleq [1 : n] \setminus I$ is referred to as a *cut* in $\mathcal{G}[B]$. The graph $\mathcal{G}[B]$, corresponding to a non-negative matrix B, is said to be *weight-balanced* if the total weights of incoming and outcoming arcs are same for all nodes $\sum_j b_{ij} = \sum_j b_{ji}$ $\forall i$. For any cut (I, J) in such a graph, the balance condition holds as follows

$$\sum_{i \in I, j \in J} b_{ij} = \sum_{i \in I, j \in J} b_{ji}.$$

Following (Hendrickx and Tsitsiklis, 2013), we call a graph *cut-balanced* if a constant $C \geq 1$ exists such that

$$C^{-1} \sum_{i \in I, j \in J} b_{ji} \leq \sum_{i \in I, j \in J} b_{ij} \leq C \sum_{i \in I, j \in J} b_{ji}. \quad (6)$$

Lemma 1. *For a nonnegative matrix B, the following statements are equivalent:*

1. *graph $\mathcal{G}[B]$ is cut-balanced;*

2. *flows from I to J and from J to I are either both positive or both zero:*

$$\sum_{i \in I, j \in J} b_{ij} > 0 \Longleftrightarrow \sum_{i \in I, j \in J} b_{ji} > 0; \quad (7)$$

3. *all strongly connected components of $\mathcal{G}[B]$ are isolated;*

4. *a walk from i to j exists if and only if a walk from j to i exists $(i, j \in [1 : n])$.*

PROOF. The implications $1 \Longrightarrow 3$ and $3 \Longleftrightarrow 4$ follow from a more general result (Hendrickx and Tsitsiklis, 2013, Lemma 1). To prove $4 \Longrightarrow 2$, note that the left-hand inequality in (7) holds if and only if an arc (j, i) from $j \in J$ to $i \in I$ exists. The walk from i to j (existing due to 4) contains an arc connecting a node from I to a node from J, therefore, the right-hand side inequality in (7) is also valid. The right-hand side inequality in (7) entails the left-hand side one for the same reason. To prove that $2 \Longrightarrow 1$, it suffices to notice that (6) holds with the constant

$$C \triangleq \max_{(I,J)} \frac{\sum_{i \in I, j \in J} b_{ij}}{\sum_{i \in I, j \in J} b_{ji}},$$

where the maximum is taken over all possible cuts such that the ratio is well-defined.

3. Convergence of classical averaging algorithms

We start by reviewing some basic results on stability of the iterative averaging algorithm (3), where $W(k)_{k \geq 0}$ is a sequence of $n \times n$ row-stochastic matrices, corresponding to the dynamic communication graph $\mathcal{G}(k) = \mathcal{G}[W(k)]$.

3.1. Time invariant dynamics

In the special case of constant $W(k) = W$ for all k, consensus conditions are well-known and dual to the conditions of regularity in a stationary discrete-time Markov chain, discussed e.g. in (Proskurnikov and Tempo, 2017).

Theorem 1. *For $W(k) = W$, $k = 0, 1, \ldots$, where the constant matrix W is row-stochastic, the following conditions are equivalent:*

1. *for any initial condition, the opinions $x_j(k)$ converge to some consensus value:*

$$\forall x(0) \, \exists c = c(x(0)) : \quad x(k) \xrightarrow[k \to \infty]{} c \mathbb{1}_n. \quad (8)$$

2. *the matrix W is SIA[4], that is, there exists a nonnegative vector $\pi \in \mathbb{R}^n$ such that $\lim_{k \to \infty} W^k = \mathbb{1}_n \pi^\top$;*

3. *$\mathcal{G}[W]$ is a quasi-strongly connected graph whose (unique) source component is aperiodic[5].*

If these conditions hold, then π is the unique Perron-Frobenius left eigenvector such that $\pi^\top W = \pi^\top$ and $\sum_i \pi_i = 1$. The consensus opinion of the group is $c = \pi^\top x(0)$.

If the source component of the graph contains a set of nodes I, then the opinions of the agents from I evolve independently of the remaining group:

$$w_{ij} = 0 \quad \forall i \in I, j \notin I \Longrightarrow x_i(k+1) = \sum_{j \in I} w_{ij} x_j(k) \quad \forall i \in I.$$

This explains the impossibility of consensus in presence of two such components. If strong component I being a source is periodic, then the submatrix $(W_{ij})_{i,j \in I}$ has an eigenvalue $\lambda \in \mathbb{C}$ such that $\lambda \neq 1$ but $|\lambda| = 1$, so that for almost all initial conditions the opinions periodically oscillate (Proskurnikov and Tempo, 2017).

---

[4] SIA matrices are also called regular (Seneta, 1981) or fully regular (Gantmacher, 2000), since they correspond to regular (ergodic) Markov chains Seneta (1981) that "forget" their history and converge to a unique stationary distribution. Matrix W is regular if and only if some column of W^k is strictly positive for some k (Seneta, 1981).

[5] We recall that a strong component of a graph is a source component if no arc enters it, a quasi-strongly connected graph has only one such component. A dual formulation in terms of Markov chains is: a chain is regular if and only if has only one essential (recurrent) class, which is aperiodic (Gantmacher, 2000; Seneta, 1981).

3.2. Time-varying averaging: necessary conditions

We now turn the attention to the more sophisticated case of non-stationary averaging procedure (3), with time-varying $W(k)$. Notice first that consensus in the sense of definition (8) can be established by "degenerate" procedures of iterative averaging. For instance, if $W(0) = \mathbb{1}_n \pi^\top$ is a trivial rank-one stochastic matrix, then the opinion iteration terminates in one step: $x(1) = x(2) = \ldots = (\pi^\top x(0))\mathbb{1}_n$. Similarly, if $W(k)W(k-1)\ldots W(0) = \mathbb{1}_n \pi^\top$ is a rank-one matrix, the iterative averaging procedure terminates and establishes consensus in no more than k steps. Notice that the dynamics of matrices $W(s)$, $s \geq k$ play no role: consensus remains invariant even if the agents do not communicate after the first k steps. Although the problem of finite-time consensus is of self-standing interest (Hendrickx et al., 2015), the aforementioned situation is non-generic and it is usually ruled out by a stronger requirement (Moreau, 2005; Shi and Johansson, 2013b) of establishing consensus for each *starting time*:

$$
\begin{aligned}
\forall k_0 \geq 0, \ \forall x(k_0) \ \exists c = c(k_0, x(k_0)): \\
W(k)\ldots W(k_0)x(k_0) \xrightarrow[k\to\infty]{} c\mathbb{1}_n.
\end{aligned}
\tag{9}
$$

As has been already mentioned, consensus in (3) is equivalent to *strong ergodicity* (Seneta, 1981) of the backward matrix products $\mathcal{W}^{k_0,k} \triangleq W(k)W(k-1)\ldots W(k_0)$, that is, the existence of a limit $\mathcal{W}^{k_0,\infty} = \lim_{k\to\infty} \mathcal{W}^{k_0,k}$, which is a one-rank stochastic matrix $\mathcal{W}^{k_0,\infty} = \mathbb{1}_n \pi_{k_0}^\top$. To the best of the authors knowledge, no necessary and sufficient condition for consensus (or for strong ergodicity of backward products) has been obtained in the literature. Many consensus criteria available in the literature are based on various *ergodicity coefficients* (see e.g. (Dobrushin, 1956; Hajnal, 1958; Seneta, 1981; Leizarowitz, 1992; Cao et al., 2008)) and so-called properties of infinite flow (Touri and Nedić, 2011; Touri and Nedic, 2012; Bolouki and Malhame, 2016). In this review, we confine ourselves to a few criteria that seem to be most convenient as they admit simple graph-theoretic interpretations.

An important *necessary* condition for consensus is the quasi-connectivity of the so-called *persistent* graph (Shi and Johansson, 2013b; Xia et al., 2019). In continuous time, this property is called integral or essential connectivity (Martin and Hendrickx, 2016; Matveev et al., 2013).

Definition 1. A persistent graph[6] of the matrix sequence $\{W(k)\}$ is the graph $\mathcal{G}_p = ([1:n], \mathcal{E}_p)$, whose set of arcs is defined as follows

$$
\mathcal{E}_p = \left\{ (j,i): \sum_{k=0}^\infty w_{ij}(k) = \infty \right\}.
\tag{10}
$$

---

[6]In many works on consensus Blondel et al. (2005); Moreau (2005), an additional assumption is stipulated that non-zero influence weights w_{ij} are uniformly positive, see the condition (11). Then $(j,i) \in \mathcal{E}_p$ if and only if the influence of j on i "persists" in the sense that arc (j,i) appears in infinitely many graphs $\mathcal{G}[k]$.

Lemma 2. *(Shi and Johansson, 2013b, Proposition 3.1) If consensus in the sense* (9) *is established, then the persistent graph $\mathcal{G}_p$ is quasi-strongly connected.*

The necessary condition from Lemma 2 implies another necessary condition proposed by Touri and Nedic (2012) under the name of *infinite flow property*: for every set of indices $S \subsetneq [1:n]$, $S \neq \emptyset$, the graph $\mathcal{G}_p$ contains an arc (i,j) between S and S^c, that is, either $i \in S, j \in S^c$ or $i \in S^c$ and $j \in S$. This is obvious, since either S or S^c should contain the spanning tree's root. At the same time, Touri and Nedic (2012) establish a stronger necessary condition for consensus (ergodicity of the matrix products), which is called *absolute flow property* and require that the infinite flow property remains invariant under special "rotational" transformations of the matrices $W(k)$; an extension of this result obtained by (Bolouki and Malhame, 2016) states the necessity of a so-called *infinite jet-flow property*. To verify these properties is, however, a self-standing non-trivial problem, which has been solved only in special cases (Touri and Nedic, 2012; Bolouki and Malhame, 2016).

In the case of static matrix $W(k) \equiv W$, $\forall k$, $\mathcal{G}_p = \mathcal{G}[W]$, the necessary condition from Lemma 2 becomes "almost" sufficient in view of Theorem 1 (modulo the aperiodicity assumption, which holds e.g. if $w_{ii} > 0 \forall i$). For non-stationary matrices, even completeness of the graph $\mathcal{G}_p$ does not always imply consensus if some persistent interactions are much "weaker" than others and the divergence rates of the series in (10) are different (Moreau, 2005).

To guarantee consensus, the quasi-strong connectivity of $\mathcal{G}_p$ has to be supplemented by additional assumptions. Three typical conditions of this type are (*i*) *repeated* (uniform) connectivity, (*ii*) uniformly bounded ratios of influence weights on persistent arcs ("arc-balance"), and (*iii*) a uniform version of the cut-balance condition (6). We focus only on convergence to consensus and do not consider here additional properties of the algorithms such as, e.g., their convergence rates (Blondel et al., 2005; Cao et al., 2008; Olshevsky and Tsitsiklis, 2011; Xia et al., 2019).

3.3. Sufficient conditions: repeated connectivity

Conditions of the first type are known as the *repeated* (periodic, uniform) quasi-strong connectivity. Typically, these conditions are formulated under additional assumption of uniform positivity of non-zero weights $w_{ij}(k)$. The following consensus criterion is known from (Blondel et al., 2005, Theorem 1), (Moreau, 2005, Theorem 2) and (Ren and Beard, 2008, Theorem 2.39).

Lemma 3. *Suppose that all non-zero entries of $W(k)$ are uniformly positive:*

$$
w_{ij}(k) \in \{0\} \cup [\eta, 1] \qquad \forall i,j \in [1:n]\,\forall k \geq 0,
\tag{11}
$$

furthermore, $w_{ii}(k) > 0$ for any i, k. Assume also that a period $T > 0$ exists such that the graphs

$$
\mathcal{G}_{k,T} = \mathcal{G}[W(k) + \ldots + W(k+T-1)]
$$

(arising as unions of T consecutive graphs $\mathcal{G}[W(k)], \ldots, \mathcal{G}[W(k+T-1)]$) are quasi-strongly connected for all $k \geq 0$. Then consensus (9) is established.

The assumptions of Lemma 3 are known as the *repeated quasi-strong connectivity* and, obviously, imply the quasi-strong connectivity of the persistent graph $\mathcal{G}_p$.

3.4. Sufficient conditions: arc-balance and cut-balance

Sufficient conditions of the second kind have been proposed in (Shi and Johansson, 2013b) under name of "arc-balance". This condition requires that all persistent interactions occur simultaneously, moreover, the ratios of corresponding influence rates on *every two persistent arcs* $(j, i), (m, l) \in \mathcal{E}_p$ is uniformly bounded

$$C^{-1} w_{lm}(k) \leq w_{ij}(k) \leq C w_{lm}(k) \; \forall k \geq 0. \tag{12}$$

A "non-instantaneous" relaxation of this condition was introduced in (Xia et al., 2019): for each pair of persistent arcs $(j, i), (m, l) \in \mathcal{E}_p$ and time instant $k_0 \geq 0$ one has

$$C^{-1} \sum_{k=k_0}^{k_0+L} w_{lm}(k) \leq \sum_{k=k_0}^{k_0+L} w_{ij}(k) \leq C \sum_{k=k_0}^{k_0+L} w_{lm}(k). \tag{13}$$

Here $C \geq 1, L \geq 0$ are, respectively, a real and an integer numbers that are independent of $(j, i), (m, l)$ and k_0.

Lemma 4. *(Xia et al., 2019) Let the diagonal entries of $W(k)$ be uniformly positive $w_{ii}(k) \geq \eta > 0 \, \forall i, k$ and (13) hold for some $C \geq 1, L \geq 0$. Then consensus (9) is established if and only if $\mathcal{G}_p$ is quasi-strongly connected.*

An alternative consensus condition, introduced in (Hendrickx and Tsitsiklis, 2013), and, later in (Bolouki and Malhame, 2015; Martin and Hendrickx, 2016) is the *uniform cut-balance*: for any cut (I, J) and any k,

$$\sum_{i \in I, j \in J} w_{ij}(k) \leq C \sum_{i \in I, j \in J} w_{ji}(k), \tag{14}$$

where constant C is independent of k and of the cut (for static graphs, this condition coincides with (7)). A more general condition introduced in (Xia et al., 2019) replaces (14) by its "non-instantaneous" version: for any cut (I, J) and any $k_0 \geq 0$ one has

$$\sum_{k=k_0}^{k_0+L} \sum_{i \in I, j \in J} w_{ij}(k) \leq C \sum_{k=k_0}^{k_0+L} \sum_{i \in I, j \in J} w_{ji}(k), \tag{15}$$

where a real number $C \geq 1$ and an integer $L \geq 0$ are independent of the choice of (I, J) and k_0.

Lemma 5. *(Xia et al., 2019) Let the diagonal entries of $W(k)$ be uniformly positive $w_{ii}(k) \geq \eta > 0 \, \forall i, k$ and (15) hold for some $C \geq 1, L \geq 0$. Then consensus (9) is established if and only if graph $\mathcal{G}_p$ is quasi-strongly connected[7].*

---

[7]Under the assumption (15), the graph $\mathcal{G}_p$ has isolated strong components, so quasi-strong and strong connectivity are equivalent.

4. New results: convergence of RAI

In this section, we present novel results concerned with the behavior of any feasible solution to the RAI (4). The proofs of these results will be given in Section 6. We are primarily interested in the following two properties.

Definition 2. *RAI (4) is* convergent *if any of its feasible solutions converges*

$$\exists \bar{x} = \lim_{k \to \infty} x(k). \tag{16}$$

RAI (4) establishes consensus *if it is convergent and, for any feasible solution, the terminal opinions are coincident $\bar{x}_1 = \ldots = \bar{x}_n$ (equivalently, $\bar{x} = c\mathbb{1}_n$, where c is a scalar).*

Obviously, $x(k)$ is a feasible solution to (4) if and only if

$$\Delta(k) = W(k)x(k) - x(k+1) \geq 0. \tag{17}$$

A feasible solution to the RAI thus can be considered as a solution to a "forced" equation

$$x(k+1) = W(k)x(k) - \Delta(k),$$

where the forcing term $\Delta(k) \geq 0$ may depend on the trajectory, be unknown and unbounded. In spite of the input-to-state stability of consensus algorithms (Shi and Johansson, 2013a), the exact consensus (9) is destroyed by an arbitrarily small bounded disturbance. The capability of RAI (4) to establish consensus can be thus considered as a counterintuitive robustness property of the associated iterative averaging procedure (3) against *unbounded yet sign-preserving* disturbances.

Some components $\bar{x}_i$ of the limit vector in (16) may be equal to $-\infty$, since (4) does not guarantee the existence of a finite lower bound for the solution; however, the solution always has a finite *upper* bound as implied by the following straightforward proposition.

Proposition 1. *If $x(k)$ is a feasible solution for RAI (4), then the sequence $M(k) = \max_i x_i(k)$ is non-increasing $M(k+1) \leq M(k) \leq \ldots \leq M(0)$.*

Unlike the monotone sequence $M(k)$, the behavior of the minimum opinion $m(k) = \min_i x_i(k)$ and the opinions' "diameter" $d(k) = M(k) - m(k)$ can be very different. Even for bounded solutions, the diameter's non-increasing property fails, and the diameter cannot serve as a natural Lyapunov function. This first principal difference between the RAI (4) and the classical iterative averaging procedure (3) makes inapplicable the plethora of methods based, explicitly or implicitly, on the "pseudo-contracting" properties of iterative averaging models (Fang and Antsaklis, 2008), and estimates of the diameter $d(k)$ (Blondel et al., 2005; Moreau, 2005; Shi and Johansson, 2013b; Xia et al., 2019) cannot be used to prove convergence to consensus. Another principal difference between RAI (4) and

the averaging procedure (3) is the absence of duality between consensus and ergodicity of backward matrix products. Iterating (4), one easily derives the one-sided inequality $x(k+1) \leq W(k) \ldots W(0)x(0)$, however, the convergence of the right-hand side says nothing about the behavior of the solution. This is not surprising since, as will be shown, the strong ergodicity of backward products $W(k) \ldots W(0)$ (equivalently, consensus in (3)) is *not* sufficient for the convergence of solutions in (4). For this reason, methods based on matrix analysis and Markov chain theory (Seneta, 1981; Ren and Beard, 2008; Cao et al., 2008; Touri and Nedic, 2012; Bolouki and Malhame, 2016) are also inapplicable to analysis of the RAI.

Notice that all results presented in this section are applicable (with straightforward modifications) to the RAI

$$y(k+1) \geq W(k)y(k)$$

that reduce to (4) by the transform $y(k) \mapsto x(k) = -y(k)$.

4.1. Time-invariant RAI: convergence and consensus

We start with our first result, being a counterpart of Theorem 1.

Theorem 2. *RAI* (4) *with* $W(k) \equiv W$, $\forall k$, *where* W *is a row-stochastic matrix, is convergent if and only if all strong components of* $\mathcal{G}[W]$ *are isolated and aperiodic. If this condition holds, then*

1. $\bar{x}_i = \bar{x}_j$ *whenever* i *and* j *belong to the same strongly connected component of the graph* $\mathcal{G}[W]$ *("partial consensus" is established);*

2. *if the opinion* $x_i(k)$ *convergence to a finite limit, the corresponding residual vanishes*

$$\Delta_i(k) \xrightarrow[k \to \infty]{} 0. \tag{18}$$

The RAI establishes consensus if and only if $\mathcal{G}[W]$ *is a strongly connected aperiodic graph, that is,* W *is a primitive (irreducible aperiodic) matrix.*

Comparing Theorems 1 and 2, one notes that consensus in the RAI requires the *strong* connectivity, whereas inequalities over quasi-strongly connected graphs cannot ensure convergence of each solution. For instance, the RAI

$$x_1(k+1) \leq x_1(k), \quad x_2(k+1) \leq \frac{x_1(k) + x_2(k)}{2},$$

has a non-converging solution $x_1(k) \equiv 1$, $x_2(k) = (-1)^k$.

4.2. Time-varying case: convergence under reciprocal interactions

Convergence of the solutions in the time-varying case will be established under several assumptions. Our first assumption, which is typically adopted in the works on consensus (Blondel et al., 2005; Ren and Beard, 2008; Shi

and Johansson, 2013b; Xia et al., 2019), can be considered as a counterpart of the aperiodicity assumption. Theorem 2 shows that, even for the static graph case, this assumption can be relaxed yet not discarded completely.

Assumption 1. *(Self-influence) The diagonal entries of* $W(k)$ *are uniformly positive: a constant* $\eta > 0$ *exists such that* $w_{ii}(k) \geq \eta > 0 \,\forall i, k$.

The second assumption requires uniform positivity of non-zero weights w_{ij}. This assumptions holds in many interesting applications (see Section 5) and allows to visibly simplify the proofs. At the same time, this assumption is in fact not necessary for convergence and consensus and can be dropped in some situations (see Theorem 5 below).

Assumption 2. *The entries of* $W(k)$ *satisfy the condition* (11), *i.e., all non-zero entries are uniformly positive.*

Remark 1. *Under Assumption 2, persistent arc* $(j, i) \in \mathcal{E}_p$ *stands for a pair of agents that interacts infinitely often, i.e.* $w_{ij}(k_s) > 0$ *for an infinite sequence* $k_s \to \infty$.

We start with a counterpart of Lemma 2, which gives a *necessary* condition for consensus and shows that the results of Lemmas 3 and 4 do not retain their validity for the inequalities. Even for repeated quasi-strong connectivity, some solutions to RAI (4) may fail to converge.

Lemma 6. *Under Assumption 2, RAI* (4) *can be convergent (respectively, establishes consensus) only if all strong components of the graph* $\mathcal{G}_p$ *are isolated (respectively,* $\mathcal{G}_p$ *is a strongly connected graph).*

As has been discussed already, this necessary condition is not sufficient for consensus, even for the classical iterative averaging procedure (3) (see e.g. the counterexample in Moreau (2005)). To ensure consensus, some additional assumption of *reciprocity* of the interactions is needed. We now introduce such an assumption, whose detailed discussion will be given in Section 4.3.

Definition 3. *For two non-empty subsets* $I, J \subseteq [1 : n]$, *let* $a_{I,J}(k_0 : k_1)$ *denote the number of arcs connecting* J *to* I *over the time window* $[k_0 : k_1]$, *that is,*

$$a_{I,J}(k_0 : k_1) \triangleq |\{(i, j) : w_{ij}(k) > 0 \text{ for some } k \in [k_0 : k_1].\}|.$$

Assumption 3. *(Reciprocity) There exist integer numbers* $\mathfrak{M} \geq 1, T \geq 0$ *such that for any cut* (I, J) *the following implication holds:*

$$a_{I,J}(k_0 : k_1) \geq \mathfrak{M} \implies a_{J,I}(k_0 : (k_1 + T)) \geq 1. \tag{19}$$

Thus a sufficiently large cumulative influence of group J *onto group* I *during some time interval* $[k_0 : k_1]$ *triggers the response from* I *to* J *(possibly, retarded by* T *steps).*

We are now in position to formulate our second result.

Theorem 3. *Let Assumptions 1 ,2 and 3 hold. Then*

1. *RAI (4) is convergent;*

2. *opinions of the agents $i, j \in [1 : n]$ from the same strong component of $\mathcal{G}_p$ the terminal opinions coincide $\bar{x}_i = \bar{x}_j$ (partial consensus is established);*

3. *if the opinion $x_i(k)$ convergence to a finite limit, then the corresponding residual vanishes $\Delta_i(k) \xrightarrow[k \to \infty]{} 0$.*

RAI (4) establishes consensus if and only if $\mathcal{G}_p$ is strong.

Theorem 3 can be further extended: both convergence and consensus appear to be robust against bounded *communication delays*. For classical consensus algorithm, such a robustness is known under the assumption of repeated quasi-strong connectivity (Angeli and Bliman, 2006).

Theorem 4. *Consider a sequence of matrices $(d_{ij}(k))_{i,j=1}^n$, whose diagonal entries are zero[8] $d_{ii}(k) = 0$, the other entries being uniformly bounded $0 \leq d_{ij}(k) \leq d_* < \infty$. Then the criteria from Theorem 3 retain their validity for the inequalities*

$$x_i(k+1) \leq \sum_{j=1}^n w_{ij}(k) x_j(k - d_{ij}(k)), \quad i \in [1 : n]. \quad (20)$$

Namely, the delayed RAI (20) is convergent if Assumptions 1,2 and 3 hold. Also, if the component $x_i(k)$ is bounded, then the corresponding "residual" vanishes

$$\Delta_i(k) = \sum_{j=1}^n w_{ij}(k) x_j(k - d_{ij}(k)) - x_i(k+1) \xrightarrow[k \to \infty]{} 0 \; \forall i.$$

If, additionally, the graph $\mathcal{G}_p$ is strongly connected, then the RAI establishes consensus.

If both weights w_{ij} and delays d_{ij} are constant, Assumption 1 can be relaxed: it suffices that each strongly connected component of $\mathcal{G}[W]$ contains a self-arc (standing for the diagonal entry $w_{ii} > 0$).

Notice that the result of Theorem 2, replacing Assumption 1 by *aperiodicity* of each strong component of $\mathcal{G}[W]$, does not retain its validity even if the delays are constant. Such a robustness cannot be proved even for the algorithm (3). A trivial counterexample is

$$x_1(k+1) = \frac{x_2(k-1) + x_3(k)}{2}, \quad (21)$$
$$x_2(k+1) = x_1(k-1), \; x_3(k+1) = x_2(k).$$

The corresponding graph $\mathcal{G}[W]$ is strongly connected and aperiodic since it contains two cycles $1 \to 2 \to 1$ (length 2) and $1 \to 2 \to 3 \to 1$ (length 3). At the same time, the system obviously has infinitely many periodic solutions:

---

[8] In other words, an agent has access to its own *undelayed* state.

choosing $x_1(k)$ to be a sequence of period 4, the vector $(x_1(k), x_1(k-2), x_1(k-3))$ is a solution to (21).

In presence of time-varying delays, the strong connectivity and positivity of a single diagonal element w_{ii} can be insufficient, as shown by the following trivial example. Let $n = 2$ and $w_{11} = 0, w_{12} = 1, w_{21} = w_{22} = 1/2$. The graph $\mathcal{G}[W]$ is then strongly connected and aperiodic since node 2 has a self-loop. Choosing the delays $d_{11} = d_{22} = d_{12} = 0$ and $d_{21}(k) = k \mod 2 \in \{0, 1\}$, the sequence

$$x_1(k) = \begin{cases} 1, & k \text{ odd} \\ 0, & k \text{ even} \end{cases}, \quad x_2(k) \equiv 1$$

obey RAI (20). Indeed, $x_1(k) \leq x_2(k) = w_{11}x_1(k) + w_{12}x_2(k)$. Also, $x_1(k - d_{21}(k)) = 1$ for any k and hence $x_2(k) = w_{21}x_1(k) + w_{22}x_2(k)$. Therefore, the delayed RAI (20) fails to be convergent.

4.3. Reciprocity condition: discussion

In this subsection, we discuss the relation between our results and previously known consensus criteria.

4.3.1. Consensus under repeated strong connectivity

As it has been already discussed, Lemma 3 does not retain its validity for the RAI, even for constant $W(k) \equiv W$. Notice, however, that if all graphs $G[W(k) + \ldots + W(k + T - 1)]$, $k \geq 0$ (equivalently, unions of T consecutive graphs) are *strongly* connected, then the implication (19) holds with an arbitrary $\mathfrak{M} \geq 0$, since the statement on its right-hand side is always true (during T consecutive steps, some agent from I has to communicate to some agent from $\bar{I}$). Theorem 3 implies the following counterpart of Lemma 3, which has been reported in our previous work (Proskurnikov and Cao, 2017b).

Corollary 1. *Assume that Assumptions 1 and 2 hold and a period $T > 0$ exists such that all graphs $G[W(k) + \ldots + W(k+T-1)]$, $k \geq 0$ (equivalently, unions of T consecutive graphs) are strongly connected. Then RAI (4) establishes consensus and (18) holds for any bounded solution.*

4.3.2. Non-instantaneous type-symmetry

Another example where the assumptions of Theorem 3 hold is *type-symmetry* (Lorenz, 2005; Hendrickx and Tsitsiklis, 2013) of the interaction weights

$$w_{ij}(k) \leq C w_{ji}(k) \quad \forall i, j \in [1 : n] \forall k \geq 0. \quad (22)$$

where $C \geq 1$ is a constant. Under Assumption 2, the type-symmetry can be reformulated as a condition of *bidirectional* communication: if j communicates to i at time k (that is, $w_{ij}(k) > 0$, then i communicates to j (i.e. $w_{ji}(k) > 0$). A natural extension (Blondel et al., 2005) of the latter condition allows a *delayed* response:

$$\forall k_0 \geq 0 \quad w_{ij}(k_0) > 0 \implies \sum_{k=k_0}^{k_0 + T} w_{ji}(k) > 0. \quad (23)$$

Condition (22), as well as (23), obviously implies Assumption 3 with $\mathfrak{M} = 1$. Theorem 3 thus extends the result on convergence of type-symmetric consensus algorithms (3) (Lorenz, 2005, Theorem 2), (Blondel et al., 2005, Theorem 5) to the systems of RAI (4).

4.3.3. Uniform cut-balance and arc-balance

Under Assumption 2, the condition of uniform cut-balance (14), obviously, implies the validity of Assumption 3 (with $\mathfrak{M} = 1$, $T = 0$). More general condition (15) also entails the validity of Assumption 3 with $\mathfrak{M} = 1$, $T = L$. Indeed, suppose that $a_{I,J}(k_0 : k_1) \geq 1$ and let $k \in [k_0 : k_1]$ be the first instant when $a_{I,J}(k_1') > 0$. Then (15) implies that $a_{J,I}(k : (k+L)) > 0$, which proves the implication (19) with $T = L$.

In view of Remark 1, under Assumption 2 the condition of "arc-balance" (12) implies that, starting from some instant k_* one has either $\mathcal{E}[W(k)] = \emptyset$ or $\mathcal{E}[W(k)] = \mathcal{E}_p$. This obviously implies $(1,0)$-reciprocity of the sequence $\{W(k)\}_{k \geq k_*}$. Similarly, the non-instantaneous arc-balance (13) implies that for $k \geq k_*$ either $\mathcal{E}[W(k) + \ldots + W(k+L)] = \emptyset$ or $\mathcal{E}[W(k) + \ldots + W(k+L)] = \mathcal{E}_p$. Similar to the uniform cut-balance case, one shows that Assumption 3 holds with $\mathfrak{M} = 1$, $T = L$.

Theorem 3 yields in the following counterpart of Lemmas 4 and 5.

Corollary 2. *Let Assumptions 1, 2 be valid and one of the conditions (13) or (15) hold. Then RAI (4) is convergent and consensus in each strong component of $\mathcal{G}_p$ is established and (18) holds for any bounded solution. The RAI establishes consensus if and only if $\mathcal{G}_p$ is strong.*

4.3.4. Periodic gossiping with intermittent communication

We now propose a simple example where the implication (19) does not reduce to any of the aforementioned type-symmetry or balance conditions. A special class of iterative averaging policies (3) is constituted by so-called *gossiping* algorithms, where at each stage of the iteration at least one pair of agents communicates. In the case of deterministic unidirectional *periodic* gossip (Anderson et al., 2010) the sequence of graphs $\mathcal{G}[W(k)]$ is periodic and each graph $\mathcal{G}[W(k)]$ contains a single arc $\mathcal{E}[W(k)] = \{(j_k, i_k)\}$. Suppose now that these gossiping interactions are separated by arbitrarily long periods of "silence" where the agents do not interact. Formally, suppose that there exists a sequence $k_1 < k_2 < \ldots < k_s < \ldots$ and a *periodic* sequence of arcs (j_s, i_s) such that

$$W(k) = \begin{cases} (1 - \alpha_s)\mathbf{e}_{i_s}\mathbf{e}_{j_s}^{\top} + \alpha_s I_n, & k = k_s \\ I_n, & k \neq k_s \; \forall s. \end{cases} \quad (24)$$

For the iterative averaging algorithm (3), the periods of silence do not change the asymptotic behavior of the solution (except for its convergence rate) since the vector of opinions $x(k)$, obviously, remains unchanged for $k = k_s + 1, \ldots, k_{s+1} - 1$. The RAI allows the opinion

vector to evolve during the silence periods (the the only restriction is the inequality $x(k+1) \leq x(k)$), so the convergence is not straightforward.

All the aforementioned reciprocity conditions (type-symmetry, arc-balance and cut-balance), obviously, fail to hold if $\sup_s(k_{s+1} - k_s) = \infty$. At the same time, the sequence of matrices (24) satisfies the condition (19), where $T = 0$ and $\mathfrak{M}$ is the period of the sequence (j_s, i_s). If $a_{I,J}(t_0 : t_1) \geq \mathfrak{M}$ for some time window $[t_0 : t_1]$, then during this time window each arc of the strongly connected graph $\mathcal{G}_p$ has appeared at least once, and hence $a_{J,I}(t_0 : t_1) > 0$. Theorem 3 now implies the following.

Corollary 3. *Let $W(k)$ have the representation (24), where (j_s, i_s) is a periodic sequence of arcs constituting a strongly connected graph $\mathcal{G}_p$ and constants α_s obey the inequality $\eta \leq \alpha_s \leq 1 - \eta$ for some $\eta > 0$. Then RAI (4) establishes consensus and the residual vanishes (18) for any bounded component.*

4.4. Assumption 2 is not necessary

As we have mentioned, Assumption 2 is, in fact, *not necessary* for convergence and consensus and can be discarded in some situations. In particular, this assumption can be dropped in the case of uniform cut-balance (14) and absence of delays.

Theorem 5. *Assume that the graph $\mathcal{G}[W(k)]$ is uniformly cut-balanced (14) and Assumption 1 holds. Then all statements of Theorem 3 retain their validity, furthermore, if $x_i(k)$ is a bounded sequence, then $\sum_{k=0}^{\infty} \Delta_i(k) < \infty$.*

Notice that Theorem 5 for the case of iterative averaging procedure (3) has been proved (in a more general stochastic formulation) in (Touri and Nedić, 2011; Touri and Langbort, 2014).

5. Applications

In this section, we apply the main results to analysis of several agent-based models of opinion formation and distributed algorithms.

5.1. Stability of some positive delay systems

A non-negative $n \times n$ matrix $A = (a_{ij})$ is *substochastic* if $\sum_{j=1}^{n} a_{ij} \leq 1 \, \forall i$. Unlike a stochastic matrix, always having eigenvalue at 1, a substochastic matrix may be Schur stable $\rho(A) < 1$. Theorem 2 leads to an elegant stability criterion (Frasca et al., 2013; Parsegov et al., 2015).

Lemma 7. *For a substochastic matrix A, consider the set of "deficiency" indices $I_d = \{i : \sum_j a_{ij} < 1\}$. If every node i in the graph $\mathcal{G}[A]$ is reachable from the set I_d by a walk (formally, i is reachable from at least one node $j \in I_d$) then $\rho(A) < 1$. In particular, if A is irreducible ($\mathcal{G}[A]$ is strongly connected) and $I_d \neq \emptyset$, then $\rho(A) < 1$.*

More generally, if a constant matrix $(d_{ij})_{i,j=1}^n$ has zero diagonal entries $d_{ii} = 0$, then the following linear delay system is globally asymptotically stable

$$x_i(k+1) = \sum_{j=1}^n a_{ij}x_j(k-d_{ij}), \quad i \in [1:n]. \quad (25)$$

PROOF. We are going to show that every solution $(x(k))$ of (25) obeys RAI (20) with a special constant row-stochastic matrix W, "dominating" the matrix A in the sense that $w_{ij} \geq a_{ij}$ for all i,j.

Consider first the undelayed case, where $x(k) = A^k x(0)$ for every k. Let $V = \text{diag}(A\mathbb{1}_n)$ be the diagonal matrix, whose entries $v_{ii} = \sum_{l=1}^n a_{il} \leq 1$ stand for the sums of matrix A's rows. Obviously, the matrix

$$W = A + \frac{1}{n}(I - V)\mathbb{1}_n\mathbb{1}_n^\top \quad (26)$$

is stochastic since its entries are non-negative and $W\mathbb{1}_n = A\mathbb{1}_n + (I - V)\mathbb{1}_n = \mathbb{1}_n$. Notice also that $w_{ij} > a_{ij} \geq 0 \forall j$ for every $i \in I_d$. Hence in the graph $\mathcal{G}[W]$ each node j is connected to every node from I_d; in particular, all nodes from I_d have self-loops. By assumption, from I_d every node is reachable by a walk, therefore $\mathcal{G}[W]$ is strongly connected. Choosing an arbitrary non-negative vector $x_0 \geq 0$, the vectors $x(k) = A^k x_0$ are non-negative for any $k \geq 0$ and satisfy the inequality (4) with $W(k) \equiv W$. Thanks to Theorem 2, $x(k) \to c\mathbb{1}$, where $c \geq 0$ and

$$\Delta(k) = Wx(k) - x(k+1) = n^{-1}(I-V)\mathbb{1}_n\mathbb{1}_n^\top x(k) \xrightarrow[k\to\infty]{} 0. \quad (27)$$

The latter condition implies that $c(I - V)\mathbb{1}_n = 0$, which is only possible for $c = 0$ (by assumption, $I_d(A) \neq \emptyset$). Hence $A^k x_0 \xrightarrow[k\to\infty]{} 0$ for any vector $x_0 \geq 0$. Since every vector is a difference of two non-negative vectors, A is Schur stable.

The second statement is proved similarly. Introducing W as in (26) and consider a solution to (25) with non-negative initial condition $x_i(\tau) \geq 0$ for $-d_* \leq \tau \leq 0$. This solution is non-negative and obeys the inequalities

$$0 \leq x_i(k+1) \leq \sum_{j=1}^n w_{ij}x_j(k-d_{ij}) \quad \forall i \in [1:n].$$

Theorem 4 ensures consensus $x(k) \xrightarrow[k\to\infty]{} c\mathbb{1}$ and

$$\Delta_i(k) = \sum_{j=1}^n w_{ij}x_j(k-d_{ij}) - x_i(k+1) = \\ = n^{-1}(1-v_{ii})\sum_{i=1}^n x(k) \xrightarrow[k\to\infty]{} 0, \quad (28)$$

i.e. $c(1-v_{ii}) = 0 \forall i$ and thus $c = 0$. Hence, solutions with non-negative initial conditions vanish as $k \to \infty$. Due to the linearity of (25), the same holds for an arbitrary solution, that is, the system is asymptotically stable. ∎

The Schur stability criterion from Lemma 7 is not only sufficient but also necessary (Parsegov et al., 2017). Lemma 7 implies, in particular, the condition of opinion convergence in the *Friedkin-Johnsen* model of opinion formation (Frasca et al., 2013). To obtain an explicit estimate of the $\rho(A)$ (that is, the solution's convergence rate) in terms of the weighted graph $\mathcal{G}[A]$ is a non-trivial problem; some results are available in Proskurnikov et al. (2017).

5.2. The Hegselmann-Krause model with "truth seekers"

One of the seminal models describing opinion formation in social networks is known as the Hegselmann-Krause *bounded confidence* model. Its simplest version (Krause, 2000; Hegselmann and Krause, 2002) arises as a modification of the DeGroot algorithm (3), where the matrix $W(k)$ co-evolves with the opinion vector and is defined as follows

$$W(k) = \bar{W}(x(k)), \quad \bar{w}_{ij}(x) \triangleq \begin{cases} \frac{1}{|N_i(x)|}, & j \in N_i(x), \\ 0, & j \notin N_i(x), \end{cases} \quad (29)$$

$$N_i(x) \triangleq \{j : |x_j - x_i| < \varepsilon\}.$$

In other words, at each stage of the opinion iteration, an individual replaces his/her opinion by the average of its own ($N_i(x) \ni i$) opinion and the opinions of *like-minded* individuals that belong to the *confidence set*, that is, the ball $B_\varepsilon(x_i) = \{x : |x - x_i| < \varepsilon\}$. The agent ignores dissimilar opinions, lying outside his/her confidence set.

Notice that the matrix $\bar{W}(x)$ is stochastic and satisfies the condition (11) with $\eta = n^{-1}$ since $1 \leq |N_i(x)| \leq n \forall i \in [1:n]$. Also, $\bar{w}_{ii}(x) \geq n^{-1} \forall i$. In particular, the sequence of matrices $W(k) = \bar{W}(x(k))$ satisfies the *type-symmetry* condition (22) with $C = n$. The corresponding graph $\mathcal{G}_p$ is thus undirected and $(j,i) \in \mathcal{E}_p$ if and only if $w_{ij}(k) > 0$ (equivalently, $|x_i(k)-x_j(k)| < \varepsilon$) for infinitely many $k \geq 0$. One arrives at a simple proposition.

Proposition 2. *For any sequence of vectors $x(k)$, the sequence $W(k) = \bar{W}(x(k))$ satisfies Assumptions 1,2 and 3.*

In view of Proposition (2) and Theorem 3 every solution to the Hegselmann-Krause model (3),(29) converges.[9] Since $(j,i) \notin \mathcal{E}_p$ if and only if $|x_i(k) - x_j(k)| \geq \varepsilon$ for large k, the steady opinions $\bar{x}_i = \lim x_i(k)$ and $\bar{x}_j = \lim x_j(k)$ are either coincident or sufficiently distant $|\bar{x}_i - \bar{x}_j| \geq \varepsilon$.

In this subsection, we consider a more general model with *"truth seekers"* (Hegselmann and Krause, 2006)

$$x(k+1) = (I - A)\bar{W}(x(k))x(k) + tA\mathbb{1}_n. \quad (30)$$

Here $A = \text{diag}(a_{11}, \ldots, a_{nn})$ is a diagonal matrix, $0 \leq a_{ii} \leq 1$, $\bar{W}(x)$ is defined in (29) and $t \in \mathbb{R}$ is a constant variable, referred to as *truth* value. The presence of an

---

[9]For a group of n agents, the opinion evolution actually terminates in $O(n^3)$ steps for the case of scalar opinions and $O(n^4)$ in the case of multidimensional opinions, see (Proskurnikov and Tempo, 2018) for a historical survey of the relevant results.

additional term in the right-hand side of (30) is explained by availability of some information to the agents, making their opinions closer to the truth value. The coefficient a_{ii} measure the level of agent i's "awareness of" the truth value (Hegselmann and Krause, 2006). If $a_{ii} = 1$, the agent is able to find the truth at a single step without communicating to the other agents. If $a_{ii} = 0$, agent i updates his/her opinion in accordance with the usual Hegselmann-Krause model, and the dynamics of his/her opinions are not (directly) influenced by the truth. If $0 < a_{ii} < 1$, the opinion of individual i is driven by both the truth and the opinions of the other individuals.

The presence of the static term in the right-hand side of (30) visibly changes the dynamics and makes analysis of the model sophisticated. In the original work (Hegselmann and Krause, 2006) it was shown that if $a_{ii} > 0 \,\forall i$, then all opinions converge to the truth $x(k) \xrightarrow[k\to\infty]{} t\mathbb{1}_n$. Later Kurz and Rambau (2011) showed that the opinions of the *truth-seekers* ($a_{ii} > 0$) always converge to t even if the group has *ignorant* agents (with $a_{ii} = 0$). At the same time, the convergence of the ignorant agents' opinions remained an open problem, which has been solved in Chazelle (2011) by using a method of power series ("s-energy"). The aforementioned results employ different and highly non-trivial mathematical techniques. Theorem 3 allows to examine convergence properties of (30) in a simpler way.

Theorem 6. *If $a_{ii} > 0$ for some i, then the following statements hold:*

1. *the limit $\bar{x}_i = \lim_{k\to\infty} x_i(k)$ exists for any i;*

2. *$\bar{x}_i = t$ if and only if agent i persistently interacts with some truth-seeker: there exists j such that $a_{jj} > 0$ and $(i,j) \in \mathcal{E}_p$ (this holds e.g. if agent i is a truth-seeker him/herself since $(i,i) \in \mathcal{E}(W(k)) \,\forall k$);*

3. *if agent i does not satisfy the condition from statement 2, then $x_i(k) \equiv \bar{x}_i$ for sufficiently large k (the opinion evolution terminates after finite number of iterations);*

4. *for any i,j either $\bar{x}_i = \bar{x}_j$ or $|\bar{x}_i - \bar{x}_j| \geq \varepsilon$ (clustering).*

PROOF. The proof exploits the vector $\xi(k) = (\xi_i(k))_{i=1}^n$ whose elements $\xi_i(k) \stackrel{\Delta}{=} |x_i(k) - t|$ stand for the distances from the agents' opinions to the truth. The sequence $\xi(k)$ appears to be a feasible solution to RAI (4), which enables us to use the result of Theorem 3. Indeed,

$$
\begin{aligned}
|x_i(k+1) - t| &= (1 - a_{ii}) \left| \sum_{j=1}^n w_{ij}(k)(x_j(k) - t) \right| \\
&\leq (1 - a_{ii}) \sum_{j=1}^n w_{ij}(k)|x_j(k) - t| \quad \forall i \,\forall k.
\end{aligned}
\tag{31}
$$

Step 1. We are going to prove the sufficiency part of statement 2. In view of Proposition 2 and Theorem 3,

the sequence $\xi(k)$ converges and remains bounded, since $\xi(k) \geq 0$. Therefore, the residual term (17) vanishes, i.e.,

$$
\Delta_i(k) = \sum_{j=1}^n w_{ij}(k)\xi_j(k) - \xi_i(k+1) \xrightarrow[k\to\infty]{} 0 \quad \forall i.
$$

In view of (31), $\Delta_i(k) \geq a_{ii}\sum_j w_{ij}(k)\xi_j(k) \geq a_{ii}w_{ii}(k)\xi_i(k)$. If agent i is a truth-seeker ($a_{ii} > 0$), then $\xi_i(k) = |x_i(k) - t| \xrightarrow[k\to\infty]{} 0$. If i persistently interacts to a truth-seeker j, then $\xi_j(k)$ and $\xi_i(k)$ reach "consensus"

$$
\lim_{k\to\infty} \xi_i(k) = \lim_{k\to\infty} \xi_j(k) = 0,
$$

i.e. $\lim_{k\to\infty} x_i = t$ for truth-seekers and agents persistently interacting to them.

Step 2. To prove statements 1 and 3, consider the set I of all agents that do not obey the condition of statement 2. For any $i \in I$, one has $a_{ii} = 0$. Also, if $i \in I$ and $a_{jj} > 0$, then $(i,j) \notin \mathcal{E}_p$ and hence $w_{ij}(k) = 0$ for large $k \geq 0$. An integer $k_0 \geq 0$ thus exists such that

$$
x_i(k+1) = \sum_{j \in I} w_{ij}(k)x_j(k) \quad \forall k \geq k_0.
$$

In other words, the subvector $\tilde{x}(k) = (x_i(k))_{i \in I}$ (after k_0 steps) obeys the conventional Hegselmann-Krause model without truth-seekers (29), which terminates in finite time. We have proved statement 3 and, due to statement 1, the existence of limits $\bar{x}_i$ for all i. From the aforementioned properties of the Hegselmann-Krause model (29), we know that $|\bar{x}_i - \bar{x}_j| \in \{0\} \cup [\varepsilon, \infty]$ for any $i,j \in I$.

Step 3. To prove statement 4 and necessity in statement 2, it remains to show that $|\bar{x}_i - t| \geq \varepsilon$ for any $i \in I$. Suppose, on the contrary, that $|\bar{x}_i - t| < \varepsilon$. For every truth-seeker j, we have $\lim_{k\to\infty} |x_i(k) - x_j(k)| < \varepsilon$ and thus $w_{ij}(k) > 0$ for k being sufficiently large, which contradicts to the assumption $(i,j) \notin \mathcal{E}_p$.∎

Notice that the condition from statement 2 is not easy to check, since $\mathcal{E}_p$ depends on a specific trajectory, which, in turn, nonlinearly depends on the initial condition. To disclose an explicit relation between the terminal opinion profile $\bar{x}$ and the initial condition $x(0)$ is a non-trivial open problem that remains beyond the scope of this paper.

5.3. The discrete-time model of bimodal polarization

Altafini's model of opinion formation, originally proposed in Altafini (2012) in its continuous-time form, portrays polarization, or "bipartite consensus", of opinions over *structurally balanced* signed graphs (Shi et al., 2019). In this subsection, we consider a discrete-time modification of the Altafini model, examined in (Meng et al., 2016; Liu et al., 2017). This model is similar to the consensus protocol (3), allowing, however, negative influence weights

$$
x(k+1) = A(k)x(k) \in \mathbb{R}^n. \tag{32}
$$

The matrix $A(k)$ satisfies the following assumption.

Assumption 4. *For any $k \geq 0$, the diagonal entries are non-negative $a_{ii}(k) \geq 0$. The non-negative matrix of absolute values $W(k) = (|a_{ij}(k)|)$ is row-stochastic.*

The non-diagonal entries $a_{ij}(k)$ in (32) may be both positive and negative. Considering the elements $x_i(k)$ as opinions, the positive value $a_{ij}(k) > 0$ can be treated as trust or friendship between agents i and j. In this case, agent i shifts his/her opinion towards the opinion of agent j. Similarly, the negative value $a_{ij}(k) < 0$ stands for distrust or enmity among the agents: the ith agent's opinion is shifted away from the opinion of agent j. This rule is motivated by the theory of structural balance in social networks, originating from (Heider, 1944, 1946).

The central question concerned with the model (32), as well as its continuous-time counterparts (Altafini, 2013; Proskurnikov and Cao, 2014) is reaching modulus consensus or consensus in absolute value:

$$\lim_{k \to \infty} |x_1(k)| = \ldots = \lim_{k \to \infty} |x_n(k)| \quad \forall x(0). \quad (33)$$

Consensus in absolute value implies that the opinions either reach consensus or polarize: some opinions converge to $M \geq 0$, the other reaching $(-M)$, where the value M depends on the initial condition. It is possible that $M = 0$ for all initial conditions, such a situation is considered as degenerate, since the "consensus" opinion is independent of the individual opinions of the agents. The following criterion of consensus in absolute value extends the results established in (Meng et al., 2016; Liu et al., 2017).

Theorem 7. *Let the sequence $\{W(k)\}$ obey Assumptions 1,2 and 3. Then the finite limit $\bar{x}_i = \lim_{k \to \infty} x_i(k) \in \mathbb{R}$ exists for any i. If $\mathcal{G}_p$ is strongly connected, then consensus in absolute value (33) is established and, furthermore,*

- *either the system (32) is globally asymptotically stable $x(k) \xrightarrow[k \to \infty]{} 0 \, \forall x(0)$,*

- *or the sequence of signed graphs $\mathcal{G}[A(k)]$ is jointly structurally balanced (Liu et al., 2017), that is, a diagonal matrix D with entries $d_{ii} \in \{-1, 1\}$ such that $W(k) = DA(k)D$ for k sufficiently large[10].*

PROOF. The columns $\xi(k) = (\xi_i(k))$ of absolute values $\xi_i(k) = |x_i(k)|$ obey RAI (4) with $w_{ij}(k) = |a_{ij}(k)|$, since

$$\xi_i(k+1) \leq \sum_{j=1}^{n} |a_{ij}(k)|\xi_j(k) \quad \forall i. \quad (34)$$

---

[10]In the case of joint structural balance, one has $\operatorname{sgn} a_{ij}(k) = \operatorname{sgn} d_{ii} \operatorname{sgn} d_{jj}$ whenever $a_{ij} \neq 0$. In other words, the relation between agents i and j is positive if d_{ii}, d_{jj} have the same sign and negative if the signs are different. The sets of agents with $d_{ii} = 1$ and $d_{ii} = -1$ thus constitute two opposing factions in the signed graph $\mathcal{G}[A(k)]$ (Altafini, 2013; Proskurnikov et al., 2016), which condition is referred to as *structural balance*. Opinions in the two factions converge to the opposite values $\bar{\xi}$ and $-\bar{\xi}$, where $\bar{\xi}$ depends on the initial condition; such a property is called "bipartite consensus" Altafini (2013) or bimodal polarization.

Thanks to Theorem 3, the finite limits exist $\bar{\xi}_i = \lim_{k \to \infty} \xi(k) \geq 0$ and the residuals vanish

$$\Delta_i(k) = \sum_{j=1}^{n} |a_{ij}(k)|\xi_j(k) - \xi_j(k+1) \xrightarrow[k \to \infty]{} 0. \quad (35)$$

We are going to show that $x_i(k)$ converges. In the case where $\bar{\xi}_i = 0$, this is obvious. Assume that $\bar{\xi}_i \neq 0$ and let $\sigma_i(k) = \operatorname{sgn} x_i(k)$. Notice that for any two numbers $c, d \in \mathbb{R}$ such that $c + d \geq 0$ one has $|c| + |d| - (c+d) \geq \min(-2c, 0)$. Since

$$\xi_i(k+1) = x_i(k+1)\sigma_i(k+1) \leq \underbrace{a_{ii}(k)\xi_i(k)\sigma_i(k)\sigma_i(k+1)}_{=c} +$$
$$+ \underbrace{\sum_{j \neq 0} |a_{ij}(k)|\xi_j(k)}_{=d},$$

we obtain that $\Delta_i(k) \geq |c| + |d| - (c + d) \geq \min(-2a_{ii}(k)\xi_i(k)\sigma_i(k)\sigma_i(k+1), 0)$. According to Assumption 1, for any k such that $\sigma_i(k)\sigma_i(k+1) = -1$ (that is, x_i changes its sign) $\Delta_i(k) \geq 2\eta \xi_i(k)$. By assumption, $\bar{\xi}_i > 0$ and thus for large k one has $\sigma_i(k)\sigma_i(k+1) = 1$ in view of (35), in other words, $\sigma_i(k) \equiv \bar{\sigma}_i$ is constant and therefore $x_i(k) \xrightarrow[k \to \infty]{} \bar{x}_i = \bar{\xi}_i\sigma_i$. If $\mathcal{G}_p$ is strongly connected, then (33) holds due to Theorem 3.

To prove the final statement, we first establish the following relation

$$\max(-a_{ij}(k) \operatorname{sgn} \bar{x}_i \operatorname{sgn} \bar{x}_j, 0) \xrightarrow[k \to \infty]{} 0 \quad \forall i, j. \quad (36)$$

Obviously, one can suppose that $\bar{x}_i, \bar{x}_j \neq 0$. Since

$$0 \leq \xi_i(k+1) = x_i(k+1)\sigma_i(k+1) \leq$$
$$\leq \underbrace{a_{ij}(k)\xi_j(k)\sigma_j(k)\sigma_i(k+1)}_{=c} + \underbrace{\sum_{m \neq j} |a_{im}(k)|\xi_m(k)}_{=d},$$

for any $k \geq 0$ one obtains that $\Delta_i(k) \geq 2\min(a_{ij}(k)\xi_j(k)\sigma_j(k)\sigma_i(k+1), 0)$. Recalling that $\xi_j(k) \xrightarrow[k \to \infty]{} \bar{\xi}_j \neq 0$ and $\sigma_i(k+1) \equiv \bar{\sigma}_i = \operatorname{sgn} \bar{x}_i$, $\sigma_j(k+1) \equiv \bar{\sigma}_j = \operatorname{sgn} \bar{x}_j$ for k being large, one obtains (36). Assume now that $\mathcal{G}_p$ is strongly connected and at least one solution $x(k)$ does not vanish as $k \to \infty$. Due to (33), all components $x_i(k)$ converge to limits $\bar{x}_i \neq 0$. Let $d_{ii} \triangleq \operatorname{sgn} \bar{x}_i$. Due to (36) and Assumption 2, for large k either $a_{ij}(k) = 0$ or $\operatorname{sgn} a_{ij}(k) = d_{ii}d_{jj}$, and thus $(DA(k)D)_{ij} = |a_{ij}(k)| = w_{ij}(k)$ for any i, j.∎

As has been discussed, the reciprocity condition from Assumption 3 holds if the graph is repeatedly strongly connected, which is a standard condition for consensus in absolute value (Liu et al., 2017). In fact, this condition is necessary and sufficient for *exponential* convergence of the opinions (Liu et al., 2017).

5.4. Constrained consensus and common fixed points of paracontractions

In this subsection, we consider another application of the RAI, related to the problem of *constrained* or "optimal" consensus (Nedic et al., 2010; Shi and Johansson, 2012; Lin and Ren, 2014), which includes, as a special case, distributed solving of linear equations (Liu et al., 2014; Mou et al., 2015; You et al., 2016; Wang et al., 2019).

For any closed convex set $\Omega \subset \mathbb{R}^d$ and $x \in \mathbb{R}^d$ the *projection* operator $P_\Omega : \mathbb{R}^d \to \Omega$ maps a point to the closest element of Ω, i.e., $|x - P_\Omega(x)| = \min_{y \in \Omega} |x - y|$. It can be shown that $\angle(y - P_\Omega(x), x - P_\Omega(x)) \geq \pi/2$ (Fig. 3) and

$$\|x - y\|^2 \geq \|x - P_\Omega(x)\|^2 + \|y - P_\Omega(x)\|^2 \quad \forall y \in \Omega. \quad (37)$$

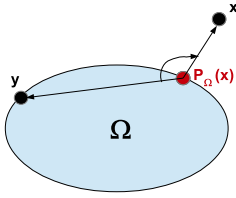The distance $d_\Omega(x) \triangleq \|x - P_\Omega(x)\|$ is a convex function.

Figure 3: The projection onto a closed convex set

Consider a group of n agents. Each agent keeps in its memory some constraints, described by a closed convex set $\Xi_i \subseteq \mathbb{R}^d$ (which can be, e.g., a hyperplane in $\mathbb{R}^d$ or the set of minima of some convex function). The agents aim to find some point $\xi_* \in \Xi_* \triangleq \Xi_1 \cap \ldots \cap \Xi_n$ that satisfies all the constraints, but do not want to communicate the information about sets Ξ_i. Assuming that an agent is able compute the projection $P_i(\xi) = P_{\Xi_i}$ of an arbitrary point $\xi \in \mathbb{R}^d$ onto the set Ξ_i, a point belonging to Ξ_* can be computed by the following modifications of the DeGroot iterative procedure (3)

$$\xi^i(k+1) = P_i \left[\sum_{j=1}^n w_{ij}(k) \xi^j(k) \right], \quad (38)$$

$$\xi^i(k+1) = P_i \left[\sum_{j=1}^n w_{ij}(k) P_j(\xi^j(k)) \right], \quad (39)$$

$$\xi^i(k+1) = w_{ii}(k) P_i(\xi_i(k)) + \sum_{j \neq i} w_{ij}(k) \xi^j(k). \quad (40)$$

The protocol (38) has been proposed in the influential paper Nedic et al. (2010), which deals with distributed optimization problems and then extended (removing some restrictive assumptions) by Lin and Ren (2014). The special cases of protocols (39) and (40) naturally arise in distributed algorithms, solving linear equations[11] (Liu et al., 2014; Mou et al., 2015; You et al., 2016). A randomized version of (40) was examined by Shi and Johansson (2012). In all of the algorithms, $\xi^i(k)$ stands for an approximation to the desired point ξ_*, computed by agent i at step

---

[11]In this case, Ξ_i are linear hyperplanes

k. In algorithms (38),(39) this approximation always satisfies the constraint of agent i ($\xi^i(k) \in \Xi_i$ for $k \geq 1$), whereas (40) provides this constraint only asymptotically.

We say that *constrained consensus* is established if the sequences $\xi^i(k)$ converge and

$$\lim_{k \to \infty} \xi^1(k) = \ldots = \lim_{k \to \infty} \xi^n(k) \in \Xi_*. \quad (41)$$

Recently it has been realized (Fullmer and Morse, 2018) that the problem of constrained consensus can be considered as a special case of a more general problem: to find a common point of a finite family of *paracontractions*, or known as *M*-Fejer mappings (Wang et al., 2019).

Definition 4. *A continuous map $M : \mathbb{R}^d \to \mathbb{R}^d$ with the set of fixed points $\mathcal{F}(M) = \{\zeta : M(\zeta) = \zeta\}$ is a* paracontraction *with respect to some norm $\|\cdot\|$ if*

$$\|M(\xi) - M(\xi_0)\| < \|\xi - \xi_0\| \quad \forall \xi \notin \mathcal{F}(M), \xi_0 \in \mathcal{F}(M). \quad (42)$$

Simple examples of paracontractions are a continuous map without fixed points $\mathcal{F}(M) = \emptyset$ and a contractive mapping ($\|M(\xi) - M(\zeta)\| \leq q\|\xi - \zeta\| \forall \xi, \zeta \in \mathbb{R}^d$ with $q \in (0,1)$). In most interesting situations, however, the fixed point is non-unique. The inequality (37) implies that the orthogonal projection P_Ω onto a closed convex set is a paracontraction in the standard Euclidean norm with $\mathcal{F}(P_\Omega) = P_\Omega$. Other examples include, but are not limited to, proximal mappings and gradient descent mappings corresponding to special convex functions Fullmer and Morse (2018). Obviously, (42) implies that M is non-expansive

$$\|M(\xi) - M(\xi_0)\| \leq \|\xi - \xi_0\| \quad \forall \xi \in \mathbb{R}^d, \xi_0 \in \mathcal{F}(M). \quad (43)$$

Notice that the requirement of constrained consensus (41) can be reformulated as follows: the distributed algorithm converges to a common *fixed point* of the paracontractive projection operators $P_i = P_{\Xi_i}$. A natural question arises whether the algorithms (38)-(40) (under proper assumptions on $W(k)$) are capable of computing a common fixed point of a general family of paracontractions P_i, that is, an element of $\Xi_* = \bigcap \mathcal{F}(P_i)$? For the algorithm (38), the affirmative answer was given in (Fullmer and Morse, 2018), assuming that the graph is repeatedly strongly connected and $\|\cdot\| = \|\cdot\|_p$ with some $1 \leq p \leq \infty$. We extend this result to all algorithms (38)-(40) and an arbitrary norm on $\mathbb{R}^d$.

Analysis of the algorithms (38)-(40) relies on a technical lemma, which holds for repeatedly strongly (and even *quasi-strongly*) connected graphs. This lemma establishes robustness of the iterative averaging procedure (3) against asymptotically vanishing disturbances.

Lemma 8. *(Lin and Ren, 2014)*[12]*. Let the matrices*

---

[12]Formally, Lemma 9 in (Lin and Ren, 2014) adopts some stronger requirements on the matrix $W(k)$ than Lemma 8, however, its proof employs only the exponential convergence of backward matrix products $W(k) \ldots W(0)$, which can be established (Blondel et al., 2005) under the assumptions of Lemma 3. An alternative proof can be given by retracing the arguments from (Shi and Johansson, 2013a), concerned with robustness of continuous-time consensus algorithms.

$W(k)$ satisfy the conditions of Lemma 3 and consider such sequences of vectors $\xi^1(k), \ldots, \xi^n(k) \in \mathbb{R}^d$, $k \geq 0$, that

$$\xi^i(k+1) = \sum_{j=1}^{n} w_{ij}(k)\xi^j(k) + e^i(k) \, \forall i, \; e^i(k) \xrightarrow[k\to\infty]{} 0. \quad (44)$$

Then the sequences $\xi^i(k)$ asymptotically synchronize, that is, $\lim_{k\to\infty} \|\xi^i(k) - \xi^j(k)\| = 0 \, \forall i, j$.

Note that Lemma 8 does not guarantee the convergence of $\xi^i(k)$; the latter property requires stronger assumptions on the "disturbances" $e^i(k)$.

To prove convergence of the algorithms, we also use the following proposition.

Proposition 3. Let M be a paracontraction in some norm $\|\cdot\|$ and ξ^0 be its fixed point. Denote $d(\xi) \triangleq \|\xi - \xi^0\| - \|M\xi - \xi^0\| \geq 0$ and consider a bounded sequence of vectors $\xi(k)$ such that $d(\xi(k)) \xrightarrow[k\to\infty]{} 0$. Then,

$$\|M\xi(k) - \xi(k)\| \xrightarrow[k\to\infty]{} 0.$$

PROOF. Assume, on the contrary, that $\|M\xi(k_r) - \xi(k_r)\| \geq \varepsilon$, $\forall r = 1, 2, \ldots$, for a number $\varepsilon > 0$ and a sequence $k_r \to \infty$. Passing to a subsequence, one can assume without loss of generality that the vectors $\xi(k_r)$ converge to a limit $\xi_* \in \mathbb{R}^d$. Recalling that M is continuous, one has $\|M\xi_* - \xi_*\| \geq \varepsilon$ while $d(\xi_*) = 0$. One arrives at a contradiction with (42), since $\xi_* \notin \mathcal{F}(M)$ whereas $\|M\xi_* - \xi^0\| = \|\xi_* - \xi^0\|$. ∎

We now formulate the main result of this subsection.

Theorem 8. Suppose that maps $P_i : \mathbb{R}^d \to \mathbb{R}^d$ are paracontractions with respect to some common norm $\|\cdot\|$ that have at least one common fixed point $\Xi_* = \bigcap_{i=1}^n \mathcal{F}(P_i) \neq \emptyset$. Let the assumptions of Corollary 1 hold. Then each of the algorithms (38)-(40) finds a common fixed point of $\{P_i\}$, that is, (41) holds.

PROOF. We first introduce some auxiliary notation. Fix an arbitrary point $\xi^0 \in \Xi_*$. For this point, let $\delta_i(\xi) \triangleq \|\xi - \xi^0\| - \|P_i\xi - \xi^0\| \geq 0$. Let $\zeta^i(k) = \sum_{j=1}^n w_{ij}(k)\xi^j(k)$.

The central idea of the proof is to explore the properties of the vectors $x(k) = (x_i(k))_{i=1}^n$ whose components $x_i(k) \triangleq \|\xi^i(k) - \xi^0\|$ stand for the distances of the agents' "opinions" $\xi^i(k)$ to the chosen fixed point. It will be shown (Step 1) that these vectors obey RAI (4), which enable us to use the techniques of Theorem 3.

Step 1. We are going to show first that the sequence $x(k)$ is a feasible solution to RAI (4). In the case of algorithm (38), one has

$$x_i(k+1) = \|P_i(\zeta^i(k)) - \xi^0\| \overset{(43)}{\leq} \|\zeta^i(k) - \xi^0\| =$$
$$= \left\| \sum_j w_{ij}(k)(\xi^j(k) - \xi^0) \right\| \leq \qquad (45)$$
$$\leq \sum_j w_{ij}(k)\|\xi^j(k) - \xi^0\| = \sum_j w_{ij}(k)x_j(k).$$

The case of (39) is considered similarly. Denoting $\bar{\zeta}^i(k) \triangleq \sum_{j=1}^n w_{ij}(k)P_j(\xi^j(k))$,

$$x_i(k+1) = \|P_i(\bar{\zeta}^i(k)) - \xi^0\| \overset{(43)}{\leq} \|\bar{\zeta}^i(k) - \xi^0\| =$$
$$= \left\| \sum_j w_{ij}(k)(P_j(\xi^j(k)) - \xi^0) \right\| \leq$$
$$\leq \sum_j w_{ij}(k)\|P_j(\xi^j(k)) - \xi^0\| \overset{(43)}{\leq} \qquad (46)$$
$$\leq \sum_j w_{ij}(k)\|\xi^j(k) - \xi^0\| = \sum_j w_{ij}(k)x_j(k).$$

In the case of algorithm (40), one has

$$x_i(k+1) = \left\| w_{ii}(k)(P_i(\xi^i(k)) - \xi^0)) + \sum_{j\neq i} w_{ij}(k)(\xi^j(k) - \xi^0) \right\|$$
$$\leq w_{ii}(k)\|P_i(\xi^i(k)) - \xi^0\| + \sum_{j\neq i} w_{ij}(k)\|\xi^j(k) - \xi^0\| \overset{(43)}{\leq}$$
$$\leq w_{ii}(k)\|\xi^i(k) - \xi^0\| + \sum_{j\neq i} w_{ij}(k)x_j(k) = \sum_{j=1}^n w_{ij}(k)x_j(k).$$
$$(47)$$

Using Corollary 1, one shows that the RAI establishes consensus, that is, $x_i(k) \xrightarrow[k\to\infty]{} c \geq 0 \, \forall i$ and $\Delta_i(k) \xrightarrow[k\to\infty]{} 0$.

Step 2. Using (18), we are now going to prove (44), which can be also written as

$$\|\xi^i(k+1) - \zeta^i(k)\| \xrightarrow[k\to\infty]{} 0 \quad \forall i. \qquad (48)$$

Also, it will be shown that

$$\|P_i(\xi^i(k)) - \xi^i(k)\| \xrightarrow[k\to\infty]{} 0 \quad \forall i. \qquad (49)$$

Notice first that the vectors $\xi^i(k)$ and $\zeta^i(k)$ are bounded, since their distances $x_i(k)$ to ξ^0 are bounded. For algorithm (38), the inequality (45) imply that

$$\Delta_i(k) \geq \|\zeta^i(k) - \xi^0\| - \|P_i(\zeta^i(k)) - \xi^0\| = \delta_i(\zeta^i(k)). \quad (50)$$

Applying Proposition 3 to $M = P_i$, (18) leads to $P_i(\zeta^i(k)) - \zeta^i(k) \xrightarrow[k\to\infty]{} 0$, which is equivalent to (48) $(\xi_i(k+1) = P_i(\zeta^i(k)))$. To derive (49), note that

$$\|P_i(\xi^i(k+1)) - \xi^i(k+1)\| \leq \underbrace{\|P_i(\xi^i(k+1)) - P_i(\zeta^i(k))\|}_{\leq \|\xi^i(k+1) - \zeta^i(k)\|} +$$
$$+ \|P_i(\zeta^i(k)) - \zeta^i(k)\| + \|\zeta^i(k) - \xi^i(k+1)\| \xrightarrow[k\to\infty]{} 0.$$

The proof in the case of algorithm (40) is simpler. Note that (47) yields in the following

$$\Delta_i(k) \geq w_{ii}(k)(\|\xi^i(k) - \xi^0\| - \|P_i(\xi^i(k)) - \xi^0\|) \geq$$
$$\geq \eta \delta_i(\xi^i(k)). \qquad (51)$$

Here $\eta > 0$ is the constant from Assumption 1. Applying Proposition 3, one proves (49), entailing also (48) since $\xi^i(k+1) = \zeta^i(k) + w_{ii}(k)(P_i(\xi^i(k)) - \xi^i(k))$.

The proof in the case of algorithm (39) combines the two aforementioned estimates. First, (46) yields in

$$\Delta_i(k) \geq \|\bar\zeta^i(k) - \xi^0\| - \|P_i(\bar\zeta^i(k)) - \xi^0\| = \delta_i(\bar\zeta^i(k)), \ (52)$$

which inequality is similar to (50) and entails that

$$\|\xi_i(k+1) - \bar\zeta_i(k)\| = \|P_i(\bar\zeta^i(k)) - \bar\zeta^i(k)\| \xrightarrow[k\to\infty]{} 0.$$

Second, (46) also implies (51), which in turn implies (49). Thus $\|\bar\zeta^i(k) - \zeta^i(k)\| \xrightarrow[k\to\infty]{} 0$. This proves (48) since $\|\xi_i(k+1) - \bar\zeta^i(k)\| \leq \|\xi_i(k+1) - \bar\zeta^i(k)\| + \|\bar\zeta^i(k) - \zeta^i(k)\|$.

Since (48) is equivalent to (44), Lemma 8 ensures synchronization property $\|\xi^i(k) - \xi^j(k)\| \xrightarrow[k\to\infty]{} 0$ for all i, j.

Step 3. Recalling that the vectors $\xi^i(k)$ are bounded, there exists a sequence $k_r \to \infty$ such that $\xi^1(k_r) \xrightarrow[r\to\infty]{} \xi_* \in \mathbb{R}^d$. In view of (44), we have $\xi^i(k_r) \xrightarrow[r\to\infty]{} \xi_*$ for every i. The property (49) implies that $P_i(\xi_*) = \xi_*$, and thus $\xi_* \in \Xi_*$. It remains to show that $\xi^i(k) \xrightarrow[k\to\infty]{} \xi_*$.

To prove this, notice that at Step 1 we have not specified the choice of ξ^0, which can be an arbitrary point in Ξ_*. We have proved that for *any* such point the distances $x_i(k) = \|\xi^i(k) - \xi^0\|$ converge to some consensus value c, depending on ξ^0 and the initial conditions. In particular, substituting $\xi^0 = \xi_*$, we obtain that the limits exist

$$x_i(k) = \lim_{k\to\infty} \|\xi^i(k) - \xi_*\| = c_* \quad \forall i.$$

Recalling that $x_i(k_r) \xrightarrow[r\to\infty]{} 0$, one has $c_* = 0$, which shows that the limits in (41) exist and are equal to $\xi_* \in \Xi_*$. ∎

Remark 2. *Since Lemma 8 retains its validity (Lin and Ren, 2014) in the case of bounded communication delays, Theorem 4 allows to extend the result of Theorem 8 to networks with delayed communication. The corresponding extension is straightforward and is thus omitted.*

6. Proofs of the technical results

In this section, we prove Theorems 2-5. We start with some auxiliary constructions and technical lemmas.

6.1. Preliminary results

We start with the following simple proposition.

Proposition 4. *For any sequence $a_1, \ldots, a_m \in [0, 1-\eta]$, where $\eta \in (0,1)$, the inequality holds*

$$\pi(a_1, \ldots, a_m) \triangleq \prod_{i=1}^{m} (1 - a_i) \geq exp\left(-\eta \sum_{i=1}^{m} a_i\right). \quad (53)$$

PROOF. Since $\pi \triangleq \pi(a_1, \ldots, a_m) > 0$, one has

$$-\log \pi = \sum_{i=1}^{m} \log\left(\frac{1}{1-a_i}\right) = \sum_{i=1}^{m} \log\left(1 + \frac{a_i}{1-a_i}\right) \overset{(+)}{\leq}$$
$$\leq \sum_{i=1}^{m} \frac{a_i}{1-a_i} \leq \eta \sum_{i=1}^{m} a_i.$$

The inequality denoted $(+)$ uses the well-known fact that $\log(1 + a) \leq a$ for any $a \geq 0$.

We now derive two important estimates for the solutions of RAI (4). For a non-empty set of indices $I \subseteq [1:n]$, let

$$M_I(k) \triangleq \max_{i \in I} x_i(k), \quad M(k) \triangleq \max_i x_i(k) = M_{[1:n]}(k).$$

Along with the number of arcs $a_{I,J}(k_0 : k_1)$ between two non-empty subsets of agents J, I, consider the total flow from J to I, which is defined as follows

$$\mathbf{w}_{I,J}(k_0 : k_1) \triangleq \sum_{k=k_0}^{k_1} \sum_{i \in I, j \in J} w_{ij}(k).$$

Let $\mathbf{w}_{I,J}(k) \triangleq \mathbf{w}_{I,J}(k:k)$, $\mathbf{w}_{i,J}(k_0 : k_1) \triangleq \mathbf{w}_{\{i\},J}(k_0 : k_1)$.

In some sense, the flow $\mathbf{w}_{I,J}(k_0 : k_1)$ measures the influence of agents from $J = I^c$ on agents from I during the time interval $[k_0, k_1]$. In particular, if $\mathbf{w}_{I,J}(k_0 : k_1) = 0$, then the group of agents I evolves independently of J on this interval, so that $M_I(k)$ is non-decreasing (Proposition 1). It can be expected that if $\mathbf{w}_{I,J}(k_0 : k_1)$ is small, the maximal increase $M_I(k_1+1) - M_I(k_0)$ can be positive yet admits a small upper estimate. This fact is established by the following simple lemma.

Lemma 9. *Assume that Assumption 1 holds. For every instants $k_0 \geq 0$, $k_0' \geq k_0$, $k_1 \geq k_0'$ and every cut (I, J) with $I, J \neq \emptyset$ the solution $x(\cdot)$ obeys the inequality*

$$M_I(k_1 + 1) \leq \vartheta M_I(k_0') + (1 - \vartheta)M(k_0),$$
$$\vartheta \triangleq \exp(-\eta \mathbf{w}_{I,J}(k_0' : k_1)), \quad (54)$$

PROOF. We know (Proposition 1) that $x_i(k) \leq M(k) \leq M(k_0) \forall k \geq k_0 \forall i$. Denoting $\rho(k) \triangleq \max_{i \in I} \mathbf{w}_{i,J}(k)$, for each $k \geq k_0$, one obtains the inequality

$$x_i(k+1) \leq \sum_{j \in I} w_{ij}(k) \underbrace{x_j(k)}_{\leq M_I(k)} + \sum_{j \in J} w_{ij}(k) \underbrace{x_j(k)}_{\leq M(k_0)} =$$
$$= (1 - \mathbf{w}_{i,J}(k))M_I(k) + \mathbf{w}_{i,J}(k)M(k_0) = \quad (55)$$
$$= M(k_0) - (1 - \mathbf{w}_{i,J}(k))(M(k_0) - M_I(k)) \leq$$
$$\leq M(k_0) - (1 - \rho(k))(M(k_0) - M_I(k)) \quad \forall i \in I.$$

Therefore, $M(k_0) - M_I(k+1) \geq (1 - \rho(k))(M(k_0) - M_I(k)) \forall k \geq k_0$ and thus

$$M(k_0) - M_I(k_1 + 1) \geq (M(k_0) - M_I(k_0')) \prod_{k=k_0'}^{k_1} (1 - \rho(k)).$$

By noticing that $\mathbf{w}_{i,J}(k) \le 1 - w_{ii}(k) \le 1 - \eta$ and also $\mathbf{w}_{i,J}(k) \le \mathbf{w}_{I,J}(k)$ for each $i \in I$, Proposition 4 yields in

$$M(k_0) - M_I(k_1 + 1) \ge (M(k_0) - M_I(k_0'))e^{-\eta \sum\limits_{k=k_0'}^{k_1} \rho(k)} \ge$$

$$\ge (M(k_0) - M_I(k_0'))e^{-\eta \sum\limits_{k=k_0'}^{k_1} \mathbf{w}_{I,J}(k)} = (M(k_0) - M_I(k_0'))\theta,$$

which inequality is equivalent to (54).∎

Lemma 9 can be extended to delayed RAI (20), although the estimate is more conservative.

Lemma 10. *Let Assumption 1 hold. For a solution of the RAI (20), denote*

$$\bar{M}(k) \triangleq \max\{M(k), M(k-1), \ldots, M(k-d_*)\},$$
$$\bar{M}_I(k) \triangleq \max\{M_I(k), \ldots, M_I(k-d_*)\}.$$

For every instants $k_0 \ge 0$, $k_0' \ge k_0$ and $k_1 \ge k_0$ and every cut (I,J) with $I, J \ne \emptyset$ one has

$$\bar{M}_I(k_1 + 1) \le \bar{\vartheta}\bar{M}_I(k_0') + (1 - \bar{\vartheta})\bar{M}(k_0),$$
$$\bar{\vartheta} \triangleq \eta^{d^*} \exp(-\eta^{d_*+1}\mathbf{w}_{I,J}(k_0' : k_1)). \tag{56}$$

PROOF. Using (20), it is easy to show that the sequence $\bar{M}(k)$ is non-increasing. Denoting $\rho(k) \triangleq \min_{i \in I} \mathbf{w}_{i,J}(k)$, for each $k \ge k_0$ and $i \in I$, the following inequality is derived from (20) similar to (55)

$$x_i(k+1) \le (1 - \mathbf{w}_{i,J}(k))\bar{M}_I(k_0) + \mathbf{w}_{i,J}(k)\bar{M}(k_0)$$
$$\le \bar{M}(k_0) - \mathbf{w}_{i,J}(k)(\bar{M}(k_0) - \bar{M}_I(k)) \le \tag{57}$$
$$\le \bar{M}(k_0) - \rho(k)(\bar{M}(k_0) - \bar{M}_I(k)).$$

Besides this, notice that for any $k' \ge k_0$ and any i

$$x_i(k'+1) \le w_{ii}(k')x_i(k') + (1 - w_{ii}(k'))\bar{M}(k_0) =$$
$$= \bar{M}(k_0) - w_{ii}(k')[\bar{M}(k_0) - x_i(k')] \le$$
$$\le \bar{M}(k_0) - \eta(\bar{M}(k_0) - x_i(k')),$$

entailing the estimate

$$\bar{M}(k_0) - x_i(k' + s) \ge \eta^s(\bar{M}(k_0) - x_i(k')) \, \forall s \ge 1. \tag{58}$$

Combining (58) with (57), one proves that for $i \in I$ and $k \ge k_0$ the inequality holds as follows

$$x_i(k+s) \le \bar{M}(k_0) - \eta^{d_*}\rho(k)(\bar{M}(k_0) - \bar{M}_I(k))$$
$$\forall s = 1, \ldots, d_*, \tag{59}$$

in particular, $\bar{M}(k_0) - \bar{M}_I(k+d_*) \ge (1 - \eta^{d_*}\rho(k))(\bar{M}(k_0) - \bar{M}_I(k))$. Denote now $m \triangleq \lfloor(k_1 - k_0')/d_*\rfloor$, so that $k_0' +$

$md_* \le k_1 < k_0' + (m+1)d_*$. Then,

$$\bar{M}(k_0) - \bar{M}_I(k_0' + md_*) \ge$$
$$\ge (\bar{M}(k_0) - \bar{M}_I(k_0'))\prod_{j=0}^{m-1}(1 - \eta^{d_*}\rho(k_0' + jd_*)) \ge$$
$$\ge e^{-\eta^{(d_*+1)}\sum\limits_{j=0}^{m-1}\rho(k_0'+jd_*)}(\bar{M}(k_0) - \bar{M}_I(k_0')) \ge \tag{60}$$
$$\ge e^{-\eta^{(d_*+1)}\mathbf{w}_{I,J}(k_0':k_1)}(\bar{M}(k_0) - \bar{M}_I(k_0')) =$$
$$= \bar{\vartheta}\eta^{-d_*}(\bar{M}(k_0) - \bar{M}_I(k_0')),$$

or, equivalently, $\bar{M}(k_0) - x_i(k_0' + md_*) \ge \bar{\vartheta}\eta^{-d_*}(\bar{M}(k_0) - \bar{M}_I(k))$ for each $i \in I$. Substituting $k' = k_0' + md_*$ and $s = 1, \ldots, k_1 + 1 - k' \le d_*$ into (59), we have

$$\bar{M}(k_0) - x_i(k) \ge \bar{\vartheta}(\bar{M}(k_0) - \bar{M}_I(k_0')) \, \forall i \in I, \tag{61}$$

whenever $k \in [(k' + 1) : (k_1 + 1)]$. In view of (60), (61) holds also for all $k \in [(k_1 - d_*) : k_1]$, entailing thus (56).

Combining the results of Lemmas 9,10 with Assumptions 2 and 3, we are now able to derive the following corollary, lying in the heart of the proofs of Theorems 3 and 4. Informally, these corollary mean that for a strongly connected persistent graph, the maximum of $\bar{M}(k)$ (or $M(k)$ in the undelayed case) will eventually decrease until the opinions become unanimous. We formulate the corresponding result for RAI (20), which includes the undelayed RAI (20) as a special case $(d_* = 0)$.

Corollary 4. *Suppose that Assumptions 1,2 and 3 hold and $\mathcal{G}_p$ is strongly connected. Assume that for some $i \in [1 : n]$ and some $k_0 \ge 0$, a feasible solution of RAI (20) satisfies the condition $\bar{M}(k_0) - x_i(k_0) \ge \varepsilon > 0$. Then there exist $k_* \ge k_0$ and constant $\varrho \in (0, 1)$, independent of ε, k_0, such that inequality holds as follows*

$$\bar{M}(k_* + 1) \le \bar{M}(k_0) - \varrho\varepsilon. \tag{62}$$

PROOF. **Step 1.** First, we prove the following statement: *Consider a cut (I,J) with $I, J \ne \emptyset$ and two instants $k_0 \ge 0$, $k_0' \ge k_0$ such that $\bar{M}(k_0) - x_i(k_0') \ge \varepsilon > 0$ for any $i \in I$. Then there exist $k_1 \ge k_0'$ and $j_0 \in J$ such that*

$$x_i(k_1 + 1) \le \bar{M}(k_0) - \varrho_0\varepsilon \quad \forall i \in I_1 = I \cup \{j_0\}, \tag{63}$$

where the constant ϱ_0 is determined by η from Assumptions 1, 2 and $\mathfrak{M}, T$ from Assumption 3.

Notice first that, due to (58),

$$\bar{M}(k_0) - x_i(k_0' + s) \ge \eta^s(\bar{M}(k_0) - x_i(k_0')) \ge \eta^{d_*}\varepsilon$$
$$\forall i \in I \quad \forall s = 0, \ldots, d_*. \tag{64}$$

Thus, for $k_0'' \triangleq k_0' + d_*$ one has $\bar{M}_I(k_0'') \le \bar{M}(k_0) - \eta^{d_*}\varepsilon$.

Since $\mathcal{G}_p$ is strongly connected, $a_{J,I}(k_0' : \infty) = \infty$. Let $k_1 \ge k_0''$ be the first instant such that $a_{J,I}(k_0'' : k_1) > 0$. In view of Assumption 3 and (19), either $k_1 - k_0'' < T$ and

thus $a_{I,J}(k_0'' : k_1) \leq |I||J|T \leq (|I| + |J|)^2 T/4 = n^2 T/4$ or $a_{I,J}(k_0'' : (k_1 - T)) < \mathfrak{M}$. In both cases, we obtain that $\mathbf{w}_{I,J}(k_0'' : k_1) \leq a_{I,J}(k_0'' : k_1) < \mathfrak{M}_0 \triangleq \max(\mathfrak{M}, n^2 T/4)$.

In view of the estimates on $\bar{M}_I(k_0'')$ and $\mathbf{w}_{I,J}(k_0'' : k_1)$,

$$\bar{M}_I(k) \leq \bar{M}(k_0) - \bar{\theta}' \eta^{d_*} \varepsilon \ \forall k \in [k_0'' : (k_1 + 1)]$$
$$\bar{\vartheta}' \triangleq \eta^{d^*} \exp(-\eta^{d_*+1} \mathfrak{M}_0). \tag{65}$$

Inequality (65) is obvious if $k = k_0''$ (since $\bar{\theta}' < 1$ and $\bar{M}_I(k_0'') \leq \bar{M}(k_0) - \eta^{d_*} \varepsilon$), otherwise it follows from Lemma 10 applied to the time interval $[k_0'' : (k - 1)]$.

By definition of k_1, there exists an arc (i_0, j_0) in $\mathcal{G}[W(k_1)]$ connecting $i_0 \in I$ to some $j_0 \in J$. Denoting $d_0 \triangleq d_{j_0 i_0}(k_1)$ and $w_0 \triangleq w_{j_0 i_0}(k_1) \geq \eta$, one arrives at

$$x_{j_0}(k_1 + 1) \leq w_0 \underbrace{x_{i_0}(k_1 - d_0)}_{\leq M_I(k_1)} + (1 - w_0(k_1))\bar{M}(k_0) \leq$$
$$\leq \bar{M}(k_0) - \varrho_0 \varepsilon,$$
$$\varrho_0 \triangleq \eta^{d_*+1} \bar{\theta}' = \eta^{2d_*+1} e^{-\eta^{d_*} \mathfrak{M}_0}. \tag{66}$$

Along with (65), the latter inequality implies (63).

Step 2. Corollary 4 is now proved by the following inductive procedure. Applying the statement proved at Step 1 to $k_0' = k_0$ and $I_0 = \{i\}$, one proves the existence of $k_1 \geq k_0$ and the set $I_1 \supsetneq I_0$ of cardinality $|I_1| = 2$ such that (63) holds. Applying the same statement to $I = I_1$, $k_0' = k_1 + 1$ one shows the existence of an instant $k_2 \geq k_1 + 1$ and a set $I_2 \supsetneq I_1$ of cardinality $|I_2| = 3$ such that

$$x_i(k_2 + 1) \leq M(k_0) - \varrho_0^2 \varepsilon \quad \forall i \in I_2, \tag{67}$$

and so on. Iterating this argument $n - 1$ times, one finally arrives at the existence of $k_n \geq k_{n-1} + 1 > k_0$ such that

$$x_i(k_n + 1) \leq M(k_0) - \varrho_0^{n-1} \varepsilon \quad \forall i \in I_n = [1 : n], \tag{68}$$

entailing that $\bar{M}(k_n + d_*) \leq \bar{M}(k_0) - \eta^{d^*} \varrho_0^{n-1} \varepsilon$ due to (58). Hence, (62) holds with $\varrho = \eta^{d^*} \varrho_0^{n-1}$. ∎

6.2. Proof of Lemma 6

To prove the first statement, assume that the RAI (4) is convergent yet $\mathcal{G}_p$ contains a strong component with outcoming arcs. Then a *source* component exists that has an outcoming arc but no incoming ones. [13] Denote the set of its nodes by $I \subsetneq [1 : n]$ and let $J \triangleq I^c$. Since $(j, i) \notin \mathcal{E}_p$ for any $i \in I, j \in J$, one has $\mathbf{w}_{I,J}(0 : \infty) < \infty$. Denote $a_k = \max_{i \in I} \mathbf{w}_{i,J}(k)$. In view of Assumption 1, $\mathbf{w}_{i,J}(k) \leq 1 - w_{ii}(k) \leq 1 - \eta$. Defining $\zeta(0) \triangleq 1$ and

$$\zeta(m) \triangleq \prod_{k=0}^{m-1} (1 - a_k), \quad 1 \leq m \leq \infty.$$

the sequence $\zeta(m)$ is decreasing and converges to a limit $\zeta(\infty)$; Proposition 4 entails that that $\zeta(\infty) > 0$ since $\sum_k a_k < \infty$. Notice also that

$$\zeta(k+1) = (1 - a_k)\zeta(k) \leq \zeta(k) - \sum_{j \in J} w_{ij}(k)\zeta(k) =$$
$$= \sum_{j \in I} w_{ij}(k)\zeta(k) \quad \forall i \in I. \tag{69}$$

By assumption, there exists a persistent arc $(i_0, j_0) \in \mathcal{E}_p$, where $i_0 \in I$ and $j_0 \in J$, therefore, $w_{j_0 i_0}(k) \geq \eta$ for an infinite sequence of instants $k = k_1, k_2, \ldots$. We will now construct a solution to (4) that does not converge. Let

$$x_i(k) = \begin{cases} \zeta(k), & i \in I, \\ 0, & i \in J \setminus \{j_0\}, \\ 0, & (i = j_0) \wedge (k \neq k_s + 1 \forall s), \\ \frac{1 + (-1)^s}{2} \eta \zeta(\infty), & (i = j_0) \wedge (k = k_s + 1). \end{cases}$$

Obviously, $x_{j_0}(k_s + 1)$ does not converge as $s \to \infty$, so $x(k)$ fails to have a limit as $k \to \infty$. We are going to prove that $x(k)$ is a solution to the RAI, that is,

$$x_i(k+1) \leq \sum_{j=1}^n w_{ij}(k)x_j(k). \tag{70}$$

For $i \in I$, the latter inequality follows from (69) since

$$x_i(k+1) \overset{(69)}{\leq} \sum_{j \in I} w_{ij}(k)x_j(k) \leq \sum_{j \in I} w_{ij}(k)x_j(k) +$$
$$+ \sum_{j \in J} w_{ij}(k) \underbrace{x_j(k)}_{\geq 0} = \sum_{j=1}^n w_{ij}(k)x_j(k).$$

Notice also that $0 \leq x_{j_0}(k+1) \leq w_{j_0 i_0}(k)x_{i_0}(k)$ for any k. For $k \neq k_s$, this is obvious since $x_{i_0}(k) \geq 0 = x_{j_0}(k+1)$. For $k = k_s$, one has $w_{j_0 i_0}(k) \geq \eta$ and $x_{i_0}(k) > \zeta(\infty)$. Since all components $x_i(k)$ are non-negative, (70) holds also for $i = j_0$. For $i \in J \setminus \{j_0\}$, (70) is obvious since $x_j(k) \geq 0 \forall j$.

The second statement is immediate from Lemma 2. If consensus is established by the RAI (4), it is automatically established by the iterative averaging procedure (3), and hence $\mathcal{G}_p$ is quasi-strongly connected. Since all strong components of $\mathcal{G}_p$ are isolated, $\mathcal{G}_p$ has to be strong.

Remark 3. *The construction of the oscillatory solution works also for delayed RAI* (20), *adding the initial conditions:* $x_i(-1) = \ldots = x_i(-d_*) = \zeta(0) = 1$ *for* $i \in I$, $x_i(-1) = \ldots = x_i(-d_*) = 0$ *for* $i \notin I$. *In this case,* $x_i(k - d) \geq \xi(k) > \xi(\infty)$ *for any* $k \geq 0$, $0 \leq d \leq d_*$, $i \in I$.

Remark 4. *In the proof of Lemma 6, we in fact used a relaxed form of Assumption 1:* $w_{ii}(k) \geq \eta$ *for each node of the graph, belonging to a* source *component. This will be used in the proof of Theorem 2.*

---

[13] In other words, if the acyclic *condensed graph* Harary et al. (1965) has at least one arc, it should have a source node with at least one arc coming from it.

6.3. Proof of Theorem 3

We first notice that Assumptions 2 and 3 imply that the strong components of $\mathcal{G}_p$ are isolated. Indeed, (19) entails that if $a_{I,J}(0:\infty) = \infty$, then $a_{J,I}(0:\infty) = \infty$ for any cut (I,J). Thus, if a path from J to I exists in $\mathcal{G}_p$ (which automatically contains an arc (j,i) with $j \in J, i \in I$) then a path (and in fact, an arc) from I to J exists as well.

In view of Remark 1, starting from some instant $k \geq k_0$ only persistent arcs exist, and, renumbering the agents, the matrix $W(k)$ becomes block-diagonal $W(k) = \mathrm{diag}(W_{11}(k), \dots, W_{ss}(k))$, where $\{W_{jj}(k)\}_{k \geq k_0}$ are sequences of stochastic matrices corresponding to strongly connected graphs $\mathcal{G}_p^j$. Since the RAI decomposes into several independent inequalities, it suffices to consider the case of a strongly connected persistent graph.

It remains to show that if $\mathcal{G}_p$ is strongly connected, then the RAI establishes consensus and (18) holds whenever the solution is bounded. This statement is immediate from Corollary 4. Indeed, we know that the maximum $M(k)$ is non-increasing and hence converges to a limit $M_* = \lim_{k \to \infty} M(k)$. If $M_* = -\infty$, then all opinions reach consensus at $-\infty$. Otherwise, we have $x_i(k) \leq M(k)$ and thus $\overline{\lim}_{k \to \infty} x_i(k) \leq M_*$. It suffices to show now that $\underline{\lim}_{k \to \infty} x_i(k) \geq M_*$. Assume, on the contrary, that $\underline{\lim}_{k \to \infty} x_i(k) < M_* - \varepsilon$, where $\varepsilon > 0$. Then a sequence $k_m \to \infty$ exists such that $x_i(k_m) < M_* - \varepsilon \leq M(k_m) - \varepsilon$. In view of Corollary 4, there exists a sequence $k'_m > k_m$ such that $M(k'_m) \leq M(k_m) - \varrho\varepsilon$. Passing to the limit as $m \to \infty$, one arrives at $M_* \leq M_* - \varrho\varepsilon$, which contradicts to the assumption $M_* > -\infty$. We have proved that $x_i(k) \xrightarrow[k \to \infty]{M} {}_*$ for every i, that is, $x(k) \xrightarrow[k \to \infty]{} M_* \mathbb{1}_n$. Also, if $M_* > -\infty$, the residual can be represented as

$$0 \leq \Delta(k) = W(k)x(k) - x(k+1) =$$
$$= W(k)[x(k) - M_*\mathbb{1}_n] - (x(k+1) - M_*) \xrightarrow[k \to \infty]{} 0.$$

The "only if" part in the last statement is immediate from Lemma 6. ∎

6.4. Proof of Theorem 2

To prove the "if" part of the first statement, it suffices to consider the case where $\mathcal{G}[W]$ is strongly connected (indeed, if the strong components of $\mathcal{G}[W]$ are isolated, then RAI (4) splits into several independent RAI). If $\mathcal{G}[W]$ is strongly connected and aperiodic, then W is a *primitive* matrix (Gantmacher, 2000), and hence W^s has positive entries for sufficiently large $s \geq 0$, satisfying thus Assumptions 1 and 2. Since $x(k+s) \leq W^s x(k) \quad \forall k \geq 0$, Theorem 3 (applied to $W(k) \equiv W^s$) entails that each subsequence $x(sm + j)$, $m = 0, 1, \dots$ and $j = 0, \dots, s-1$, converges to a consensus vector $x(sm + j) \xrightarrow[m \to \infty]{} c_j \mathbb{1}_n$, $c_j \in [-\infty, \infty)$. By noticing that $x(sm + j + 1) \leq Wx(sm + j)$, one obtains that $c_0 \geq c_1 \geq \dots c_{s-1} \geq c_0$, that is, $c_0 = \dots = c_{s-1} = c$ and the RAI establishes consensus. The proof of (18) in the case where $c > -\infty$ is the same

as in the time-varying case. This finishes the proof of the "if" part in the first and the last statements, as well as the two remaining statements of Theorem 1.

To prove the converse statement, notice first that if the RAI (4) is convergent, the same holds for the RAI

$$z(m+1) \leq W^s z(m), \tag{71}$$

where $s \geq 1$ is a fixed integer number. Indeed, if some solution of the latter RAI fails to have a limit, the same holds for the sequence

$$x(k) = \begin{cases} z(m), & k = ms \text{ for some integer } m \\ Wx(k-1), & s \text{ does not divide } k, \end{cases}$$

which is a feasible solution to the RAI (4). Also, any solution to the DeGroot's system (2) is a feasible solution to the RAI (4). Hence, the RAI can be convergent only if all solutions of (2) converge, which means that each *source* component of $\mathcal{G}[W]$ is an aperiodic graph (Bullo, 2016; Proskurnikov and Tempo, 2017). For a source component of $\mathcal{G}[W]$ with the set of nodes I, $w_{ij} = 0$ for any $i \in I$ and $j \notin I$. Hence, the corresponding submatrix $W_I \triangleq (w_{i,j})_{i,j \in I}$ is row-stochastic, irreducible and aperiodic (primitive). In particular, for s being sufficiently large, the matrix $(W_I)^s$ (being a submatrix of W^s) has strictly positive entries. Obviously, I remains a source component in $\mathcal{G}[W^s]$; also, if there is an arc coming out of I in $\mathcal{G}[W]$, that is, $w_{ji} > 0$ for some $j \notin I$ and $i \in I$, the same arc exists in the graph $\mathcal{G}[W^s]$ for large s.

In view of Remark 4, Lemma 6 is applicable to RAI (71) in spite of potential violation of Assumption 1. Hence, all strong components of $\mathcal{G}[W^s]$ (in particular, all source components) are isolated for every sufficiently large s. Therefore, the *source* components of $\mathcal{G}[W]$ are also isolated. As discussed in the proof of Lemma 6, this implies that in fact all components of $\mathcal{G}[W]$ are isolated (being simultaneously sources and sinks) and have to be aperiodic. If there is more than one strong component in the graph, then DeGroot's model (2) (and also the RAI) cannot provide consensus. This finishes the proof of "only if" parts in the first and the last statements.

6.5. Proof of Theorem 4

The proof of Theorem 4, except for its final statement, retraces the proof of Theorem 3, where $M(k)$ is to be replaced by $\bar{M}(k)$ and the proof of (18) for the case of a persistent graph $\mathcal{G}_p$ and a bounded solution has to be modified in the following way: if $x_i(k) \to M_* > -\infty$, then

$$\Delta_i(k) = \sum_j w_{ij}(k)(x_j(k - d_{ij}(k)) - M_*) -$$
$$- (x_i(k+1) - M_*) \xrightarrow[k \to \infty]{} 0.$$

Notice that Lemma 6 is applicable to the delayed case in view of Remark 3.

To prove the final statement, concerned with the time-invariant situation, it suffices to consider the case where $\mathcal{G}[W]$ is strongly connected (the isolated strong components correspond to independent subsystems of inequalities). One can get rid of the delays by using a trick proposed by Xiao and Wang (2006): consider the vector $y(k) \in \mathbb{R}^{d_* n}$ obtained by stacking the vectors $x(k), x(k-1), \ldots, x(k-d_*)$ one on top of another. Then it can be easily checked (Xiao and Wang, 2006) that

$$y(k+1) \le \Xi y(k), \quad \Xi = \begin{bmatrix} W_0 & W_1 & \ldots & W_{d_*-1} & W_{d_*} \\ I_n & 0 & & & 0 \\ 0 & I_n & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & I_n & 0 \end{bmatrix}$$

Here W_i are non-negative matrices such that $W = W_0 + \ldots + W_{d_*}$; $(W_r)_{ij} = w_{ij}$ if and only if $d_{ij} = r$. In particular, $\operatorname{diag} W_0 = \operatorname{diag} W$ and, if $\operatorname{diag} W \ne 0$, the graph $\mathcal{G}[\Xi]$ has a self-arc and is aperiodic. It remains to notice that each arc (j,i) in the graph $\mathcal{G}[W]$ corresponds to a walk $j \to (j+d_*) \to (j+2d_*) \to \ldots \to (j+d_{ij}d_*) \to i$ in the graph $\mathcal{G}[\Xi]$, therefore, the graph $\mathcal{G}[\Xi]$ is strongly connected. Applying Theorem 2 to the matrix Ξ, one shows that RAI (20) establishes consensus. $\blacksquare$

6.6. Proof of Theorem 5

The proof is based on the idea of ordering that was employed in analysis of continuous-time consensus algorithms over type-symmetric and cut-balanced weighted graphs (Hendrickx and Tsitsiklis, 2013; Martin and Girard, 2013; Matveev et al., 2013) and later extended to the discrete-time case (Bolouki and Malhame, 2015). Note that the proofs from the aforementioned papers essentially use the boundedness of the solutions and their extension to a general feasible solution of RAI (20) is not straightforward. For this reason, we give an independent proof.

Let $\sigma(k) = (\sigma_1(k), \ldots, \sigma_n(k))$ denote the permutation of indices $1, \ldots, n$, sorting vector $x(k)$ in the ascending order

$$y_1(k) \overset{\triangle}{=} x_{\sigma_1}(k) \le y_2(k) \overset{\triangle}{=} x_{\sigma_2}(k) \le \ldots \le y_n(k) \overset{\triangle}{=} x_{\sigma_n}(k).$$

Obviously, the sorted vector $y(k)$ obeys RAI

$$y(k+1) \le V(k)y(k), \quad v_{ij}(k) \overset{\triangle}{=} w_{\sigma_i(k+1)\sigma_j(k)}(k). \quad (72)$$

As shown in (Bolouki and Malhame, 2015; Xia et al., 2019), the stochastic matrix $V(k)$ also satisfies the uniform cut-balanced condition (14) (with a different constant C). Notice that this fact heavily relies on Assumption 1.

Step 1. Using induction on $r = n, \ldots, 1$, we will prove the following statements: *1) the sequence $y_r(k)$ has a limit*

$\bar{y}_r \ge -\infty$ *as $k \to \infty$; 2) if $\bar{y}_r > -\infty$, then*

$$\sum_{k=0}^{\infty} v_{rs}(k)|y_r(k) - y_s(k)| < \infty \quad \forall s \in [1:n], \quad (73)$$

$$r > 1 \Rightarrow \sum_{k=0}^{\infty} v_{sp}(k)(y_r(k) - y_{r-1}(k)) < \infty \, \forall s < r \, \forall p \ge r. \quad (74)$$

To prove the induction base $r = n$, recall that $y_n(k) = M(k)$ is a non-increasing function, which thus has a limit $\bar{y}_n$ as $k \to \infty$. Notice also that

$$y_n(k) - y_n(k+1) \ge \sum_{s=1}^{n-1} v_{ns}(k)(y_n(k) - y_s(k)) \ge 0, \quad (75)$$

If $\bar{y}_n > -\infty$, the latter inequality, obviously, implies (73) with $r = n$. To prove (74), we apply the uniform cut-balance property to $I = [1:(n-1)], J = \{n\}$ and note that $y_n(k) - y_s(k) \ge y_n(k) - y_{n-1}(k) \, \forall s < n$, therefore

$$0 \le (y_n(k) - y_{n-1}(k)) \sum_{s<n} v_{sn}(k) \overset{(14)}{\le}$$
$$\le C(y_n(k) - y_{n-1}(k)) \sum_{s<n} v_{ns}(k).$$

To prove the induction step, assume that our statements have been already proved for $r = m+1, \ldots, n$. Our goal is to prove them for $r = m$. If $\bar{y}_{m+1} = -\infty$, then, obviously, $y_m(k) \xrightarrow[k \to \infty]{} -\infty$. Otherwise, (73),(74) hold for all $r > m$. Using (74) for $s = m$ and any $p > m$, $r \in [(m+1):p]$,

$$v_{mp}(y_p - y_m) = v_{mp}(y_{m+1} - y_m) + v_{mp}(y_{m+2} - y_{m+1}) +$$
$$+ \ldots + v_{mp}(y_p - y_{p-1}) \in \ell_1 \quad \forall p > m$$

(ss usual, we use ℓ_1 to denote the set of all sequences $(a(k))$ such that $\sum_k |a(k)| < \infty$.) Similar to (75), we obtain that

$$y_m(k) - y_m(k+1) \ge \underbrace{\sum_{p>m} v_{mp}(k)(y_m(k) - y_p(k))}_{\in \ell_1} +$$
$$+ \sum_{r \le m} \underbrace{v_{mr}(k)(y_m(k) - y_r(k))}_{\ge 0},$$

therefore, the exists the limit (possibly, infinite)

$$\lim_{K \to \infty} (y_m(0) - y_m(K+1)) = \sum_{k=0}^{\infty} (y_m(k) - y_m(k+1)) \le +\infty,$$

in other words, $y_m(k)$ has a limit as $k \to \infty$. Furthermore, if this limit is finite, then also $v_{mr}(y_m - y_r) \in \ell_1$ for $r < m$, which proves (73) for $r = m$. Applying (73) to all $s < m$ and $r = m, \ldots, n$ and noticing that $|y_r(k) - y_s(k)| \ge y_m(k) - y_{m-1}(k)$, one shows that

$$(y_m - y_{m-1}) \sum_{s<m} \sum_{r \ge m} v_{rs}(y_m - y_{m-1}) \in \ell_1,$$

19

and thus, applying (14) for $I = [1 : (m-1)]$, $J = [m : n]$, we also have $v_{sp}(y_m - y_{m-1}) \in \ell_1$ if $s < m \le p$. This proves (74) for $r = m$ and finishes the induction step.

Step 2. Notice that the convergence of the sorted vectors $y(k)$ in general *does not* imply the convergence of the original vector $x(k)$. A trivial counterexample is the 2-periodic sequence $x(0) = (1,0)^\top, x(1) = (0,1)^\top, x(k) = x(k-2) \forall k \ge 2$. Obviously, $y(k) \equiv x(0)$, whereas $x(k)$ fails to converge. In order to prove convergence of $x(k)$, we have to use the aperiodicity property (Assumption 1).

Suppose first that $\bar{y}_1 > -\infty$, that is, the solution is bounded. Then, (73) holds for every $r \ge 1$, entailing that $\sum_{k=0}^\infty |y_r(k+1) - y_r(k)| < \infty$ and thus

$$\sum_{k=0}^\infty \sum_{i,j=1}^n w_{ij}(k)|x_i(k+1) - x_j(k)| \le$$
$$\le \sum_{k=0}^\infty \sum_{s,r=1}^n v_{rs}(k)|y_r(k+1) - y_s(k)| < \infty$$

(recall that $v_{rs}(k) = w_{ij}(k)$ if and only if $r = \sigma_i(k+1)$ and $s = \sigma_i(k)$, for this reason we have $x_i(k+1) = y_r(k+1)$ and $x_j(k) = y_s(k)$). Since $w_{ii}(k)$ are uniformly positive, we have $\sum_{k=0}^\infty |x_i(k+1) - x_i(k)| < \infty$ (therefore, each $x_i(k)$ converges to a finite limit $\bar{x}_i$) and $w_{ij}(x_i - x_j) \in \ell_1$ for all i, j, which implies that $\bar{x}_i = \bar{x}_j$ if $(j, i) \in \mathcal{E}_p$, that is, consensus in each strongly connected component of $\mathcal{G}_p$ is established. To prove (18), notice that

$$\Delta_i(k) = \sum_j w_{ij}(k)(x_j(k) - x_i(k)) - (x_i(k+1) - x_i(k)),$$
(76)

which sequence belongs to ℓ_1.

The case of unbounded solution is more sophisticated. Assume that $\bar{y}_1 = \ldots = \bar{y}_m = -\infty < \bar{y}_{m+1}$. Notice first that if a sequence $k_m \to \infty$ such that the limit $\xi = \lim_{m\to\infty} x_i(k_m)$ exists, then ξ is one of the values $\bar{y}_1, \ldots, \bar{y}_n$. Indeed, for every m we have $x_i(k_m) = y_{j_m}(k_m)$ with for some j_m. Passing to the subsequence, we may assume, without loss of generality, that $j_m \equiv j$, and thus $\xi = \bar{y}_j$. For the same reason, for every $\bar{y}_j$ there exists at least one sequence $x_i(k_m) \xrightarrow[m\to\infty]{} \bar{y}_j$: indeed, $y_j(k) = x_{\sigma_j(k)}(k)$ for some reason, and we can find a subsequence k_m such that $\sigma_j(k_m) \equiv i$ for all m.

Let J stand for the set of indices i such that

$$\varliminf_{k\to\infty} x_i(k) = -\infty,$$

and $I = J^c$. It can be easily seen that $x_j(k) \to -\infty$ for every $j \in I$ and $\varliminf_{k\to\infty} x_i(k) \ge \bar{y}_{m+1}$ for all $i \in I$. Indeed, choose two constants $b < \bar{y}_{m+1}$ and $a < b$ in such a way that $\eta a + (1-\eta)\bar{y}_n < b$. For large k, we have $y_m(k) < a < b < y_{m+1}(k)$, therefore, no component $x_i(k)$ can belong to $[a, b]$. Hence, if $x_j(k) < a$ for k sufficiently large, RAI (4) and Assumption 1 imply that

$$x_j(k+1) < \eta a + (1-\eta)y_n(k) < b \implies x_j(k+1) < a.$$

that is, the opinion cannot leave the interval $(-\infty, a)$. Therefore, for every $j \in J$ one has $\varlimsup_{k\to\infty} x_j(k) \le a$, which is possible only if $x_j(k) \to -\infty$. For $i \in I$, we have $\varliminf_{k\to\infty} x_i(k) \ge a$, which is possible only if $\varliminf_{k\to\infty} x_i(k) \ge y_{m+1}$. Obviously, $J = \{\sigma_1(k), \ldots, \sigma_m(k)\}$ for k being sufficiently large, whereas $I = \{\sigma_{m+1}(k), \ldots, \sigma_n(k)\}$. Now, we can repeat the argument we used in the bounded solution case with a minor modification. Applying (73) for every $r > m$, one has

$$\sum_{i\in I}\sum_{j=1}^n w_{ij}(k)|x_i(k+1) - x_j(k)| =$$
$$= \sum_{r>m}\sum_{s=1}^n v_{rs}(k)|y_r(k+1) - y_s(k)|,$$

for k being sufficiently large, and therefore the latter sequence belongs to ℓ_1. Since $w_{ii}(k) \ge \eta$, we have $\sum_{k=0}^\infty |x_i(k+1) - x_i(k)| < \infty$ for $i \in I$ (therefore, $x_i(k)$ converges to a finite limit $\bar{x}_i$ for all $i \in I$) and $\sum_{k=0}^\infty w_{ij}(k)|x_j(k) - x_i(k)| < \infty$ for every $i \in I, j \in [1:n]$. Using (76), one proves that $\Delta_i \in \ell_1$ for $i \in I$. Obviously, if $i \in I$ and $j \in J$, then $(j, i) \notin \mathcal{E}_p$, that is, there are no arcs between I and J in $\mathcal{E}_p$ (recall that $\mathcal{G}_p$ automatically has isolated strongly connected components in view of the cut-balance property). We know that $\bar{x}_j = -\infty$ for $j \in J$, if $i, j \in I$ and $(j, i) \in \mathcal{E}_p$, then $\bar{x}_i = \bar{x}_j$, so that consensus in each strongly connected component of $\mathcal{G}_p$ is established.∎

Remark 5. *Using techniques from Xia et al. (2019), the consensus criterion from Theorem 5 can be easily extended to a more general situation of non-instantaneous cut-balance (15). Namely, if (15) holds, Assumption 1 is valid and $\mathcal{G}_p$ is strongly connected, then RAI (4) establishes consensus. Indeed, matrices $B(k) = W(k+L-1)\ldots W(k+1)W(k)$ satisfy the uniform cut-balanced condition (14) and Assumption 1 (Xia et al., 2019). Since*

$$x(j + (k+1)L) \le B(j + kL)x(j + kL),$$

for each $j = 0, \ldots, L-1$, the sequence $x(j + kL)$ converges to consensus vector $c_j \mathbb{1}_n$ as $k \to \infty$. Similar to the proof of Theorem 1, one proves that $c_0 = c_1 = \ldots = c_{L-1}$.

7. Conclusions and Future Works

The recurrent averaging inequalities (RAI) introduced in (4) can be considered as a relaxed form of the conventional dynamics of iterative averaging (or the French-DeGroot model) (2). While in the standard iterative averaging each agent updates its opinion with a weighted average of its own and the others' opinions, in the RAI each agent is allowed to choose *any* opinion which does not exceed that linear combination. Such a constraint seems non-restrictive, however, under some connectivity assumptions, all feasible solutions converge and even reach consensus. Similarly to the classical iterative averaging, consensus is robust to bounded communication delays. Especially interesting are bounded solutions to RAI, for which

the residual vector (the difference between the right-hand and the left-hand sides) vanishes asymptotically.

The systematic study of RAI appears to be relevant, since they naturally arise, implicitly or explicitly, in a number of distributed algorithms and multi-agent models based on iterative averaging. Using the general results on consensus in RAI, we have derived a number of known and new results in the unified way, namely, a graph-theoretic stability criterion for substochastic matrices and its novel extension for time-delay systems, the convergence of opinions in the Hegselmann-Krause model with informed agents ("truth-seekers") and the discrete-time model of opinion polarization over a signed network, the convergence of algorithms for finding a common fixed point for a family of paracontractions. The latter algorithms are very important and can be used, in particular, to solve large-scale systems of equations (Wang et al., 2019), as well as strongly convex optimization problems (Fullmer and Morse, 2018), in a distributed way. Our results can be extended to RAI with random matrices $W(k)$ and nonlinear RAI, obtained from nonlinear procedures of iterative averaging (Moreau, 2005; Angeli and Bliman, 2006; Fang and Antsaklis, 2008; Proskurnikov and Matveev, 2015), however, the corresponding results require additional mathematical tools and are thus beyond the scope of this review.

Although the theory of RAI allows to study many multi-agent algorithms and models, based on the seminal idea of iterative averaging, some consensus-based algorithms are currently beyond its scope, and their analyses still require special theory. One class of the algorithms is constituted by distributed optimization procedures employed, in particular, in machine learning, multi-sensor tracking, control of cyber-physical systems and smart power grids, see e.g. (Nedic et al., 2010; Bajovic et al., 2017; Tatarenko and Touri, 2017; Mokhtari and Ribeiro, 2016; Molzahn et al., 2017; Erofeeva and Granichin, 2018; Notarstefano et al., 2019; Scaman et al., 2019; Yang et al., 2019) and references therein. Another examples is proximal dynamics in multi-agent games (Grammatico, 2018; Cenedese et al., 2018) that are similar in spirit to the algorithms from Section 5.4, however, the corresponding proximal maps do not have a common fixed point. We believe, however, that the extensions of RAI theory towards such algorithms and models are possible; the relevant extensions are subjects of ongoing research.

References

Altafini, C., 2012. Dynamics of opinion forming in structurally balanced social networks. PLoS ONE 7, e38135.

Altafini, C., 2013. Consensus problems on networks with antagonistic interactions. IEEE Trans. Autom. Control 58, 935–946.

Amelina, N., Fradkov, A., Jiang, Y., Vergados, D., 2015. Approximate consensus in stochastic networks with application to load balancing. IEEE Trans. Inform. Theory 61, 1739–1752.

Anderson, B.D.O., Yu, C., Morse, A.S., 2010. Convergence of Periodic Gossiping Algorithms. Springer, Berlin, Heidelberg. pp. 127–138.

Angeli, D., Bliman, P., 2006. Stability of leaderless discrete-time multi-agent systems. Math. Control, Signals, Syst. 18, 293–322.

Bajovic, D., Jakovetic, D., Krejic, N., Jerinkic, N.K., 2017. Newton-like method with diagonal correction for distributed optimization. SIAM Journal on Optimization 27, 1171–1203.

Blondel, V., Hendrickx, J., Olshevsky, A., Tsitsiklis, J., 2005. Convergence in multiagent coordination, consensus, and flocking, in: Proc. IEEE Conf. Decision and Control, pp. 2996 – 3000.

Bolouki, S., Malhame, R.P., 2015. Linear consensus algorithms based on balanced asymmetric chains. IEEE Trans. Autom. Control 60, 2808–2812.

Bolouki, S., Malhame, R.P., 2016. Consensus algorithms and the decomposition-separation theorem. IEEE Transactions on Automatic Control 61, 2357–2369.

Borkar, V., Varaiya, P., 1982. Asymptotic agreement in distributed estimation. IEEE Trans. Autom. Control 27, 650–655.

Bullo, F., 2016. Lectures on Network Systems. published online at http://motion.me.ucsb.edu/book-lns. With contributions by J. Cortes, F. Dorfler, and S. Martinez.

Cao, M., Morse, A., Anderson, B., 2008. Reaching a consensus in a dynamically changing environment. SIAM J. Control Optim. 47, 575–623.

Cenedese, C., Kawano, Y., Grammatico, S., Cao, M., 2018. Towards time-varying proximal dynamics in multi-agent network games, in: IEEE Conference on Decision and Control (CDC), pp. 4378–4383.

Chazelle, B., 2011. The total s-energy of a multiagent system. SIAM J. Control Optim. 49, 1680–1706.

D.Bertsekas, Tsitsiklis, J., 1989. Parallel and Distributed Computation: Numerical Methods. Prentice-Hall, Englewood Cliff, NJ.

DeGroot, M., 1974. Reaching a consensus. Journal of the American Statistical Association 69, 118–121.

Dobrushin, R., 1956. Central limit theorem for non-stationary Markov chains I, II. Theorem Probab. Appl. 1, 63–80,233–246.

Erofeeva, V., Granichin, O., 2018. Distributed state estimation of moving targets using cyclic simultaneous perturbation stochastic approximation. IFAC-PapersOnLine 51, 218 – 223.

Fang, L., Antsaklis, P., 2008. Asynchronous consensus protocols using nonlinear paracontractions theory. IEEE Trans. Automat. Control 53, 2351–2355.

Frasca, P., Ravazzi, C., Tempo, R., Ishii, H., 2013. Gossips and prejudices: Ergodic randomized dynamics in social networks, in: Proc. of IFAC NecSys 2013 Workshop, Koblenz, Germany. pp. 212–219.

French Jr., J., 1956. A formal theory of social power. Physchol. Rev. 63, 181–194.

Friedkin, N., 2015. The problem of social control and coordination of complex systems in sociology: A look at the community cleavage problem. IEEE Control Systems 35, 40–51.

Fullmer, D., Morse, A.S., 2018. A distributed algorithm for computing a common fixed point of a finite family of paracontractions. IEEE Transactions on Automatic Control 63, 2833–2843.

Gantmacher, F., 2000. The Theory of Matrices. volume 2. AMS Chelsea Publishing.

Garin, F., Schenato, L., 2010. A Survey on Distributed Estimation and Control Applications Using Linear Consensus Algorithms. Springer London, London. pp. 75–107.

Grammatico, S., 2018. Proximal dynamics in multiagent network games. IEEE Transactions on Control of Network Systems 5, 1707–1716.

Hajnal, J., 1958. Weak ergodicity in non-homogeneous Markov chains. Proc. Cambridge Philos. Soc. 54, 233–246.

Harary, F., 1959. A criterion for unanimity in French's theory of social power, in: Cartwright, D. (Ed.), Studies in Social Power. Univ. of Michigan Press, Ann Arbor, MI, pp. 168–182.

Harary, F., Norman, R., Cartwright, D., 1965. Structural Models. An Introduction to the Theory of Directed Graphs. Wiley & Sons, New York, London, Sydney.

Hegselmann, R., Krause, U., 2002. Opinion dynamics and bounded confidence models, analysis, and simulation. Journal of Artifical Societies and Social Simulation (JASSS) 5, 2.

Hegselmann, R., Krause, U., 2006. Truth and cognitive division of

labour: First steps towards a computer aided social epistemology. Journal of Artifical Societies and Social Simulation (JASSS) 9, 1.

Heider, F., 1944. Social perception and phenomenal causality. Psychol. Rev. 51, 358–374.

Heider, F., 1946. Attitudes and cognitive organization. J. Psychol. 21, 107–122.

Hendrickx, J., Tsitsiklis, J., 2013. Convergence of type-symmetric and cut-balanced consensus seeking systems. IEEE Trans. Autom. Control 58, 214–218.

Hendrickx, J.M., Shi, G., Johansson, K.H., 2015. Finite-time consensus using stochastic matrices with positive diagonals. IEEE Transactions on Automatic Control 60, 1070–1073.

Jadbabaie, A., Lin, J., Morse, A., 2003. Coordination of groups of mobile autonomous agents using nearest neighbor rules. IEEE Trans. Autom. Control 48, 988–1001.

Kemeny, F., Snell, J., 1976. Finite Markov Chains. Springer Verlag, New York, Berlin, Heidelberg, Tokyo.

Krause, U., 2000. A discrete nonlinear and non-autonomous model of consensus formation. Commun. in Difference Equations , 227–236.

Kurz, S., Rambau, J., 2011. On the Hegselmann–Krause conjecture in opinion dynamics. J. Difference Equations and Applications 17, 859–876.

Lehrer, K., 1976. When rational disagreement is impossible. Noûs 10, 327–332.

Leizarowitz, A., 1992. On infinite products of stochastic matrices. Linear Algebra and its Applications 168, 189 – 219.

Leonard, N.E., 2014. Multi-agent system dynamics: Bifurcation and behavior of animal groups. Annual Reviews in Control 38, 171 – 183.

Lewis, F., Zhang, H., Hengster-Movris, K., Das, A., 2014. Cooperative Control of Multi-Agent Systems. Optimal and Adaptive Design Approaches. Springer-Verlag, London.

Lin, J., Morse, A., Anderson, B., 2007. The multi-agent rendezvous problem (2 parts). SIAM J. Control Optim. 46, 2096–2147.

Lin, P., Ren, W., 2014. Constrained consensus in unbalanced networks with communication delays. IEEE Trans. Autom. Control 59, 775–781.

Liu, J., Chen, X., Basar, T., Belabbas, M., 2017. Exponential convergence of the discrete- and continuous-time Altafini models. IEEE Trans. Autom. Control 62, 6168–6182.

Liu, J., Morse, A., Nedić, A., Başar, T., 2014. Stability of a distributed algorithm for solving linear algebraic equations, in: Proc. of IEEE Conference on Decision and Control (CDC), pp. 3707–3712.

Lorenz, J., 2005. A stabilization theorem for dynamics of continuous opinions. Physica A 353, 217–223.

Martin, S., Girard, A., 2013. Continuous-time consensus under persistent connectivity and slow divergence of reciprocal interaction weights. SIAM J. Control. Optim. 51, 2568–2584.

Martin, S., Hendrickx, J., 2016. Continuous-time consensus under non-instantaneous reciprocity. IEEE Trans. Autom. Control 61, 2484–2495.

Matveev, A., Novinitsyn, I., Proskurnikov, A., 2013. Stability of continuous-time consensus algorithms for switching networks with bidirectional interaction, in: Proc. Europ. Control Conference (ECC), pp. 1872–1877.

Meng, Z., Shi, G., Johansson, K., Cao, M., Hong, Y., 2016. Behaviors of networks with antagonistic interactions and switching topologies. Automatica 73, 110–116.

Mesbahi, M., Egerstedt, M., 2010. Graph Theoretic Methods in Multiagent Networks. Princeton University Press, Princeton and Oxford.

Mokhtari, A., Ribeiro, A., 2016. DSA: Decentralized double stochastic averaging gradient algorithm. Journal of Machine Learning Research 17, 1–35.

Molzahn, D.K., Dörfler, F., Sandberg, H., Low, S.H., Chakrabarti, S., Baldick, R., Lavaei, J., 2017. A survey of distributed optimization and control algorithms for electric power systems. IEEE Transactions on Smart Grid 8, 2941–2962.

Moreau, L., 2005. Stability of multiagent systems with time-dependent communication links. IEEE Trans. Autom. Control

50, 169–182.

Mou, S., Liu, J., Morse, A., 2015. A distributed algorithm for solving a linear algebraic equation. IEEE Trans. Autom. Control 60, 2863–2878.

Nedic, A., Ozdaglar, A., Parrilo, P., 2010. Constrained consensus and optimization in multi-agent networks. IEEE Trans. Autom. Control 55, 922–938.

Notarstefano, G., Notarnicola, I., Camisa, A., 2019. Distributed optimization for smart cyber-physical networks. Foundations and Trends in Systems and Control 7, 253–383. URL: http://dx.doi.org/10.1561/2600000020.

Olfati-Saber, R., 2006. Flocking for multi-agent dynamic systems: algorithms and theory. IEEE Trans. Autom. Control 51, 401–420.

Olfati-Saber, R., Fax, J., Murray, R., 2007. Consensus and cooperation in networked multi-agent systems. Proceedings of the IEEE 95, 215–233.

Olshevsky, A., Tsitsiklis, J., 2011. Convergence speed in distributed consensus and averaging. SIAM Rev. 53, 747–772.

Parsegov, S., Proskurnikov, A., Tempo, R., Friedkin, N., 2015. A new model of opinion dynamics for social actors with multiple interdependent attitudes and prejudices, in: Proc. of IEEE Conference on Decision and Control (CDC), pp. 3475 – 3480.

Parsegov, S., Proskurnikov, A., Tempo, R., Friedkin, N., 2017. Novel multidimensional models of opinion dynamics in social networks. IEEE Trans. Autom. Control 62, 2270–2285.

Proskurnikov, A., Cao, M., 2014. Opinion dynamics using Altafini's model with a time-varying directed graph, in: Proceedings of IEEE ISIC 2014 (Part of IEEE MSC 2014), Antibes. pp. 849–854.

Proskurnikov, A., Cao, M., 2017a. Differential inequalities in multi-agent coordination and opinion dynamics modeling. Automatica 85, 202–210.

Proskurnikov, A., Cao, M., 2017b. Modulus consensus in discrete-time signed networks and properties of special recurrent inequalities, in: IEEE Conf. Decision and Control, pp. 2003–2008.

Proskurnikov, A., Matveev, A., 2015. Popov-type criterion for consensus in nonlinearly coupled networks. IEEE Trans. Cybernetics 45, 1537–1548.

Proskurnikov, A., Matveev, A., Cao, M., 2016. Opinion dynamics in social networks with hostile camps: Consensus vs. polarization. IEEE Trans. Autom. Control 61, 1524–1536.

Proskurnikov, A., Tempo, R., 2017. A tutorial on modeling and analysis of dynamic social networks. Part I. Annual Reviews in Control 43. 65-79.

Proskurnikov, A., Tempo, R., 2018. A tutorial on modeling and analysis of dynamic social networks. Part II. Annual Reviews in Control 45. 166–190.

Proskurnikov, A., Tempo, R., Cao, M., Friedkin, N., 2017. Opinion evolution in time-varying social influence networks with prejudiced agents. IFAC PapersOnline 50, 11896–11901.

Ren, W., Beard, R., 2008. Distributed Consensus in Multi-Vehicle Cooperative Control: Theory and Applications. Springer-Verlag, London.

Ren, W., Cao, Y., 2011. Distributed Coordination of Multi-agent Networks. Springer.

Reynolds, C., 1987. Flocks, herds, and schools: a distributed behavioral model. Computer Graphics 21, 25–34.

Scaman, K., Bach, F., Bubeck, S., Lee, Y.T., Massoulié, L., 2019. Optimal convergence rates for convex distributed optimization in networks. Journal of Machine Learning Research 20, 1–31.

Seneta, E., 1981. Non-Negative Matrices and Markov Chains, 2nd ed. Springer-Verlag, New York.

Shamma, J. (Ed.), 2007. Cooperative Control of Distributed Multi-Agent Systems. Wiley.

Shi, G., Altafini, C., Baras, J., 2019. Dynamics over signed networks. SIAM Review 61, 229–257.

Shi, G., Johansson, K., 2012. Randomized optimal consensus of multi-agent systems. Automatica 48, 3018–3030.

Shi, G., Johansson, K., 2013a. Robust consensus for continuous-time multi-agent dynamics. SIAM J. Control Optim 51, 3673–3691.

Shi, G., Johansson, K., 2013b. The role of persistent graphs in the agreement seeking of social networks. IEEE J. Selected Areas

Commun. 31, 595–606.

Tatarenko, T., Touri, B., 2017. Non-convex distributed optimization. IEEE Transactions on Automatic Control 62, 3744–3757.

Touri, B., Langbort, C., 2014. On endogenous random consensus and averaging dynamics. IEEE Trans. Control of Network Syst. 1, 241–248.

Touri, B., Nedić, A., 2011. On ergodicity, infinite flow, and consensus in random models. IEEE Trans. Autom. Control 56, 1593–1605.

Touri, B., Nedic, A., 2012. On backward product of stochastic matrices. Automatica 48, 1477 – 1488.

Tsitsiklis, J., Bertsekas, D., Athans, M., 1986. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. IEEE Trans. Autom. Control 31, 803–812.

Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, O., Shochet, I., 1995. Novel type of phase transition in a system of self-driven particles. Phys. Rev. Letters 75, 1226–1229.

Wang, P., Mou, S., Lian, J., Ren, W., 2019. Solving a system of linear equations: From centralized to distributed algorithms. Annual Reviews in Control 47, 306 – 322.

Wang, P., Ren, W., Duan, Z., 2019. Distributed algorithm to solve a system of linear equations with unique or multiple solutions from arbitrary initializations. IEEE Transactions on Control of Network Systems 6, 82–93.

Wolfowitz, J., 1963. Products of indecomposable, aperiodic, stochastic matrices. Proceedings of Amer. Math. Soc. 15, 733–737.

Xia, W., Shi, G., Meng, Z., Cao, M., Johansson, K.H., 2019. Persistent flows in deterministic chains. IEEE Trans. Autom. Control 64, 2766–2781.

Xiao, F., Wang, L., 2006. State consensus for multi-agent systems with switching topologies and time-varying delays. International Journal of Control 79, 1277–1284.

Yang, T., Yi, X., Wu, J., Yuan, Y., Wu, D., Meng, Z., Hong, Y., Wang, H., Lin, Z., Johansson, K.H., 2019. A survey of distributed optimization. Annual Reviews in Control 47, 278 – 305.

You, K., Song, S., Tempo, R., 2016. A networked parallel algorithm for solving linear algebraic equations, in: Proc. of IEEE Conf. Decision and Control, pp. 1727–1732.