Abstract. Models of iterative averaging constitute an important and extensively studied class of discrete-time dynamical systems over networks. In such models, each node of a network is associated to a value (or opinion), which is updated, at each iteration, to a convex combination of itself and the opinions of the neighboring nodes. It is well known that, under natural assumptions on the network’s graph connectivity, this local averaging procedure eventually leads to global consensus of opinions at all nodes. The idea of iterative averaging actually lies at the heart of many algorithms for distributed optimization, for solution of linear and nonlinear equations, for multi-agent coordination and for opinion formation in social groups. Although these algorithms have similar structures, the mathematical techniques used for their analysis are manifold, and the conditions for their convergence and stability differ from case to case. We show in this paper that the properties of such algorithms can be analyzed in a unified way by using a novel tool based on recurrent averaging inequalities (RAI). In this paper, we develop a theory for RAI, and apply it to the analysis of several important multi-agent algorithms and models of opinion formation recently proposed in the literature.

Key words. Distributed algorithms, opinion dynamics, multi-agent system, consensus

AMS subject classifications. 39B72, 39A06, 90B10

1. Introduction. In his seminal work on social power theory, French [17] introduced a simple model of opinion diffusion in social groups based on the principle of iterative averaging. Each 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 [18]. 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. Namely, a new value of an actor’s opinion is computed as the mean of its previous value and the opinions displayed by the actor’s neighbors, see Fig. 1. Typically, [23, 24], the iteration of such process establishes eventually consensus (unanimity) of opinions. Consensus is reached if and only if some actor communicates (directly or indirectly) to all other group members, that is, the graph has a directed spanning tree [53]. French’s model can be written as

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

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 $\sum_{j} w_{ij} = 1$ for all $i$. Collecting the weights in a (row-stochastic) matrix $W = (w_{ij})$, and defining the vector $x(k) \in \mathbb{R}^n$ of opinions at time $k$, we rewrite model (1.1) in vector format as

$$x(k + 1) = Wx(k), \quad k = 0, 1, \ldots$$
\((x_A(k+1), x_B(k+1), x_C(k+1)) = \left( x_A(k), \frac{x_A(k)+x_B(k)}{2}, \frac{x_A(k)+x_B(k)+x_C(k)}{3} \right)\)

Fig. 1: French’s model of opinion formation in a small 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.

Model (1.2), with a general stochastic matrix \(W\), was later independently proposed by DeGroot [13] and Lehrer [32] as a method for reaching rational agreement in expert communities. Consensus in (1.2) is equivalent [51] to the “full regularity” [20], or SIA (stochastic, indecomposable, aperiodic) [64] property, of the matrix \(W\) or, equivalently, regularity of the Markov chain generated by this matrix. In particular, the iterations in (1.2) always reach consensus if \(W\) is aperiodic and irreducible (or primitive) [13, 20]. The behavior of iterative averaging algorithms such as (1.2), however, is more complicated in the case of dynamic communication graphs, where the matrix of influence weights \(W = W(k)\) varies with time. Consensus in (1.2), with time varying \(W\), is equivalent to the ergodicity\(^3\) of the backward infinite matrix products \(W(k) \ldots W(0)\) [55]. Despite the interest and efforts devoted to the ergodicity problem in the literature on probability theory and matrix analysis [22, 33, 55, 64], a complete solution to this problem remains elusive, and a substantial gap between necessary and sufficient conditions for ergodicity still exists [7, 61].

In recent years, the iterative averaging model has attracted enormous attention from the research community, as a simple algorithm for multi-agent coordination [5, 10, 29, 41, 42, 45, 53, 54]. Besides the important problem of multi-agent consensus, procedures of iterative averaging lie at the heart of many distributed numerical algorithms [12] such as, e.g., techniques of distributed estimation and filtering [8, 9, 21], deterministic and stochastic optimization [34, 44, 62], load balancing [3], and equation solving [36, 43, 67]. Many models of opinion formation [51, 52] in social groups stem from the classical French-DeGroot model (1.2).

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 (1.2) is the existence of a directed spanning tree in the communication graph (quasi-strong connectivity) [10, 54], whereas similar algorithms for constrained consensus [44] and linear equations solving [43, 67] 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 the study of multi-agent algorithms based on iterative averaging.

In this paper, 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-

\(^3\)The backward products of stochastic matrices \(W(k) \ldots W(0)\) are said to be ergodic if they converge, as \(k \to \infty\), to a stochastic rank-one matrix.
1. Contribution and paper organization. In this paper we develop a mathematical theory for the RAI (1.3). 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 \([26]\), and to generalize the recent fundamental results \([19]\) 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 (1.3). 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 \( n \geq m \), we let \([m : n] \triangleq \{m, m+1, \ldots, n\} \). The cardinality of a finite set \( I \) is denoted by \(|I|\). We use \( \mathbf{1}_n = (1,1,\ldots,1)^\top \in \mathbb{R}^n \) to denote a (column) vector of all ones, and \( \mathbf{e}_i = (0,0,\ldots,1)^\top \) to denote 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 \((i, j) \in \mathcal{E} \iff (j, i) \in \mathcal{E}\), 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

\[
x(k + 1) \leq W(k)x(k), \quad k = 0, 1, \ldots,
\]

where matrices \( W(k) \) are row-stochastic. At a first glance, the system of inequalities (1.3) 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 (1.3) implies asymptotic consensus of the opinions \( x_i(k) \) (which, however, can be achieved at \(-\infty\)). In many situations where consensus is not established, (1.3) provides convergence of the sequence \( \{x(k)\} \). Similar properties have been recently obtained in \([48]\) for differential inequalities

\[
\dot{x}(t) \leq -L(t)x(t),
\]

where \( L(t) \) stands for a time-varying Laplacian matrix; the theory developed in \([48]\) is however inapplicable to discrete-time inequalities (1.3).
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 [53] 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 \((V', E')\), where \(V' \subseteq V\) and \(E' \subseteq E\) is referred to as a *subgraph of the graph* \((V, E)\). 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 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-strong if and only if it has a single source component. A strong component can be *isolated*, i.e., to have neither incoming nor outgoing arcs and thus be both a source and a sink. Strong components of an undirected graph are always isolated.

\[\text{(a)} \quad \text{Fig. 2: Strong components of a directed graph: (a) non-isolated; (b) isolated. In (a), \{4\} is a single source component, \{11, \ldots, 15\} is a single sink component.}\]

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\(^2\) \((j, i)\). In this paper, we mainly deal with graphs generated by non-negative 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 \(E[B_1 + B_2] = E[B_1] \cup E[B_2]\); 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 outgoing arcs are the same for all nodes \(\sum_j b_{ij} = \sum_j b_{ji}\) for all \(i\). It can be easily shown that for

\(^2\)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 [54] and the original work by French [17], such an influence is depicted by an arc \(j \rightarrow i\) rather than \(i \rightarrow j\).
any cut \((I, J)\) in such 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 [27], we call a graph \(\text{cut-balanced}\) if a constant \(C \geq 1\) exists such that

\[
C^{-1} \sum_{i \in I, j \in J} b_{ij} \leq \sum_{i \in I, j \in J} b_{ji} \leq C \sum_{i \in I, j \in J} b_{ji}.
\]

**Lemma 2.1.** 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:

\[
\sum_{i \in I, j \in J} b_{ij} \geq 0 \iff \sum_{i \in I, j \in J} b_{ji} > 0;
\]
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\(\Rightarrow\)3 and 3\(\iff\)4 follow from a more general result [27, Lemma 1]. To prove 4\(\Rightarrow\)2, note that the left-hand inequality in (2.2) 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 (2.2) is also valid. Similarly, the right-hand side inequality in (2.2) entails the left-hand side one. To prove that 2\(\Rightarrow\)1, introduce the constant

\[
C = \max_{(I, J)} \frac{\sum_{i \in I, j \in J} b_{ij}}{\sum_{i \in I, j \in J} b_{ji}},
\]

where the maximum is taken over all possible cuts such that the ratio is well-defined. Obviously, (2.1) holds in this case. \(\square\)

3. **Classical results on iterative averaging algorithms.** We start by reviewing some basic results on stability of the iterative averaging algorithm

\[
x(k + 1) = W(k)x(k) \in \mathbb{R}^n,
\]

where \(W(k)\) for all \(k\), consensus conditions are well-known and dual to the conditions of regularity in a stationary discrete-time Markov chain (see e.g. [51]); the following theorem holds.

**Theorem 3.1.** 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:

\[
\forall x(0) \exists c = c(x(0)) : x(k) \xrightarrow[k \to \infty]{} c \mathbb{1}_n.
\]
2. the matrix \(W\) is SIA\(^3\), that is, \(\lim_{k \to \infty} W^k = \mathbb{1}_n \pi^\top\), \(\pi \in \mathbb{R}^n\);

---

\(^3\)SIA matrices are also called regular [55] or fully regular [20]. They correspond to regular Markov chains [55] that “forget” their history and converge to a unique stationary distribution. An equivalent definition is the existence of \(k\) such that \(W^k\) has a column with strictly positive entries [55].
3. the graph $\mathcal{G}[W]$ is quasi-strongly connected (has a directed spanning tree), and the (unique) source component is aperiodic\(^4\).

If these conditions hold, then $\pi$ 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:

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

This explains the impossibility of consensus in presence of two such components. Also, if a component being a source is periodic, the corresponding submatrix $(W_{ij})_{i,j \in I}$ always has an eigenvalue $\lambda \in \mathbb{C}$ such that $\lambda \neq 1$ but $|\lambda| = 1$ [51], so that for almost all initial conditions the opinions periodically oscillate.

### 3.2. Time-varying averaging: necessary conditions.

We now turn the attention to the more sophisticated case of non-stationary averaging procedure (3.1), with time-varying $W(k)$.

Notice first that consensus in the sense of definition (3.2) 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 [28], the aforementioned situation is non-generic and it is usually ruled out by a stronger requirement [42, 59] of establishing consensus for each starting time:

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

Similar to the static case, in the case of time-varying matrix $W(k)$ consensus in (3.1) is equivalent to strong ergodicity [55] of the backward matrix products $\mathcal{W}_{k_0,k} \triangleq W(k)W(k-1)\ldots W(0)$, 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. In this section we do not survey strong ergodicity criteria based on various ergodicity coefficients (see e.g. [10, 14, 22, 33, 55]) and so-called properties of infinite flow [7, 60, 61], confining ourselves to a few criteria, admitting simple graph-theoretic interpretations.

A widely known necessary condition for consensus is the quasi-connectivity of so-called “persistent” graph [59, 65] (in continuous time, a similar property is known also as integral connectivity [38] or essential connectivity [39]).

**Definition 3.2.** A persistent graph\(^5\) of the matrix sequence $\{W(k)\}$ is the graph

\[^4\]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 [55].

\[^5\]In many works on consensus [5, 42], an additional assumption is stipulated that non-zero influence weights $w_{ij}$ are uniformly positive, see the condition (3.5). Then $(j,i) \in E_P$ if and only if the influence of $j$ on $i$ “persists” in the sense that arc $(j,i)$ appears in infinitely many graphs $\mathcal{G}[k]$. 

$G_p = ([1 : n], E_p)$, whose set of arcs is defined as follows

$$E_p = \left\{ (j, i) : \sum_{k=0}^{\infty} w_{ij}(k) = \infty \right\}. \quad (3.4)$$

**Lemma 3.3.** [59, Proposition 3.1] If consensus in the sense (3.3) is established, then the persistent graph $G_p$ is quasi-strongly connected.

In the case of static matrix $W(k) ≡ W$, $\forall k$, $G_p = G[W]$, the necessary condition from Lemma 3.3 becomes sufficient, under the assumption that $w_{ii} > 0 \forall i$ (or under any other condition that guarantees the aperiodicity of the unique source component of the graph). For non-stationary matrices, even completeness of the graph $G_p$ does not always imply consensus if some persistent interactions are much “weaker” than others (the series in (3.4) diverge with different speeds), see, e.g., [42, Sec.IV-C].

To guarantee consensus, the quasi-strong connectivity of $G_p$ 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 (2.1). We focus only on convergence to consensus and do not consider here additional properties of the algorithms such as, e.g., their convergence rates [5, 10, 46, 65].

**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 well-known consensus criterion follows from, e.g., [5] (Theorem 1), [42] (Example 1 and Theorem 2), and [53] (Theorem 2.39).

**Lemma 3.4.** Suppose that all non-zero entries of $W(k)$ are uniformly positive:

$$w_{ij}(k) \in \{0\} \cup [\eta, 1] \quad \forall i, j \in [1 : n] \forall k \geq 0, \quad (3.5)$$

Furthermore, $w_{ii}(k) > 0$ for any $i, k$. Assume that 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 quasi-strongly connected. Then consensus (3.3) is established.

**3.4. Sufficient conditions: “arc-balance” and uniform cut-balance.** Sufficient conditions of the second kind have been proposed in [59] under name of “arc-balance”. This condition requires that all persistent interactions occur simultaneously, moreover, the ratios of corresponding influence rates is uniformly bounded

$$C^{-1}w_{lm}(k) \leq w_{ij}(k) \leq Cw_{lm}(k) \quad \forall (j, i), (m, l) \in E_p \forall k \geq 0. \quad (3.6)$$

A “non-instantaneous” relaxation of this condition was introduced in [65]

$$C^{-1} \sum_{k=k_0}^{k_0+L} w_{lm}(k) \leq \sum_{k=k_0}^{k_0+L} w_{ij}(k) \leq C \sum_{k=k_0}^{k_0+L} w_{lm}(k) \quad \forall (j, i), (m, l) \in E_p \forall k_0 \geq 0. \quad (3.7)$$

Here $C \geq 1$ is a real constant and $L \geq 0$ is an integer constant.

**Lemma 3.5.** [65, Theorem 1] Let the diagonal entries of $W(k)$ be uniformly positive $w_{ii}(k) \geq \eta > 0$ for any $i, k$ and (3.7) hold for some $C \geq 1, L \geq 0$. Then consensus (3.3) is established if and only if graph $G_p$ is quasi-strongly connected.
An alternative consensus condition, introduced in [27] (see also [6] and [38, Theorem 5]) is the uniform cut-balance: for any cut \((I, J)\) and any \(k\),

\[
\sum_{i \in I, j \in J} w_{ij}(k) \leq C \sum_{i \in I, j \in J} w_{ji}(k),
\]

where constant \(C\) is independent of \(k\) and of the cut (for static graphs, this condition coincides with (2.2)). A more general condition introduced in [65] replaces (3.8) by its “non-instantaneous” version: for some real \(C \geq 1\) and integer \(L \geq 0\),

\[
\sum_{k=k_0}^{k_0+L} \sum_{i \in I, j \in J} w_{ij}(k) \leq C \sum_{k=k_0}^{k_0+L} \sum_{i \in I, j \in J} w_{ji}(k), \quad \forall k_0 \geq 0.
\]

**Lemma 3.6.** [65, Theorem 2] Let the diagonal entries of \(W(k)\) be uniformly positive \(w_{ii}(k) \geq \eta > 0\) for any \(i, k\) and (3.9) hold for some \(C \geq 1, L \geq 0\). Then consensus (3.3) is established if and only if graph \(G_p\) is quasi-strongly connected\(^6\).

4. Main results on RAI. In this section, we present our main results concerned with the behavior of any feasible solution to the RAI (1.3). For ease of reading, the proofs of the results presented in this section are reported in Section 6. We are primarily interested in the following two properties of the RAI.

**Definition 4.1.** RAI (1.3) is convergent if any of its feasible solutions converges

\[
\exists \bar{x} = \lim_{k \to \infty} x(k).
\]

RAI (1.3) establishes consensus if it is convergent and, for any feasible solution, the terminal opinions are coincident \(\bar{x}_1 = \ldots = \bar{x}_n\).

Obviously, \(x(k)\) is a feasible solution to (1.3) if and only if

\[
\Delta(k) = W(k)x(k) - x(k+1) \geq 0.
\]

A 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 \(\Delta(k) \geq 0\) depends on the trajectory, it is unknown and it is generally unbounded. In spite of the input-to-state stability of consensus algorithms [58], the exact consensus (3.3) can be destroyed by an arbitrarily small bounded disturbance. The capability of RAI (1.3) to establish consensus can be thus considered as a counterintuitive robustness property of the associated iterative averaging procedure (3.1) against unbounded yet sign-preserving disturbances.

Some components \(\bar{x}_i\) of the limit vector in (4.1) may be equal to \(-\infty\), since (1.3) 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 straightforward proposition.

**Proposition 4.2.** If \(x(k)\) is a feasible solution for RAI (1.3), then the sequence \(M(k) = \max_i x_i(k)\) is non-increasing \(M(k+1) \leq M(k) \leq \ldots \leq M(0)\).

\(^6\)As can be shown, under the assumption (3.9) the graph \(G_p\) has isolated strong components, so quasi-strong and strong connectivity are equivalent.
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 (1.3) and the classical iterative averaging procedure (3.1) makes inapplicable the plethora of methods based, explicitly or implicitly, on the “pseudo-contracting” properties of iterative averaging models [15], and estimates of the diameter \( d(k) \) [5, 42, 59, 65] cannot be used to prove convergence to consensus. The second principal difference between RAI (1.3) and consensus algorithm (3.1) is the absence of duality between consensus and ergodicity of backward matrix products. Iterating (1.3), 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 (3.1)) is not sufficient for the convergence of solutions in (1.3). For this reason, methods based on matrix analysis [7, 10, 53, 55, 61] are also inapplicable to analysis of the RAI.

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

\[
y(k+1) \geq W(k)y(k)
\]

that reduce to (1.3) by the transformation \( y(x(k)) = -x(k) \).

4.1. Time-invariant RAI: convergence and consensus. We start with our first main result, being a counterpart of Theorem 3.1.

**Theorem 4.3.** RAI (1.3) with \( W(k) \equiv W, \forall k \), where \( W \) is a row-stochastic matrix, is convergent if and only if all strong components of \( G[W] \) 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 \( G[W] \) (“partial consensus” is established);
2. if the solution is bounded, then the residuals (4.2) vanish

\[
\Delta_i(k) \to 0 \quad \forall i.
\]

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

Comparing Theorems 3.1 and 4.3, one notes that consensus in the RAI requires the strong connectivity, whereas inequalities over quasi-strongly connected graphs are too “loose” to guarantee 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 bounded 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 [5, 53, 59, 65], can be considered as a counterpart of the aperiodicity assumption. Theorem 4.3 shows that this assumption can be relaxed yet not discarded completely.

**Assumption 4.4.** (Self-influence) The diagonal entries of \( W(k) \) are uniformly positive \( w_{ii}(k) \geq \eta > 0 \), where \( \eta \) is a constant.
The second assumption requires uniform positivity of non-zero weights \( w_{ij} \). This assumption holds in many interesting applications (see Section 5) and, although it can be relaxed in some situations, we adopt it in order to simplify the proofs.

**Assumption 4.5.** The entries of \( W(k) \) satisfy the condition (3.5).

**Remark 4.6.** Under Assumption 4.5, persistent arc \((j, i) \in E_p\) 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 3.3, showing that the results of Lemmas 3.4 and 3.5 in general, do not retain their validity for the inequalities. Even for repeated quasi-strong connectivity, some solutions to RAI (1.3) may fail to converge.

**Lemma 4.7.** Under Assumption 4.5, the RAI (1.3) can be convergent (respectively, establishes consensus) only if the graph \( G_p \) has isolated strong components (respectively, is strongly connected).

We now introduce the last assumption, which will be discussed in Section 4.3.

**Definition 4.8.** 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,

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

**Assumption 4.9.** (Reciprocity) There exist integer numbers \( \mathfrak{M} \geq 1, T \geq 0 \) such that for any cut \((I, J)\) the following implication holds:

\[
a_{I,J}(k_0 : k_1) \geq \mathfrak{M} \Rightarrow a_{J,I}(k_0 : (k_1 + T)) \geq 1.
\]

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 4.10.** Let Assumptions 4.4, 4.5 and 4.9 hold. Then

1. the RAI (1.3) is convergent;
2. opinions of the agents \( i, j \in [1 : n] \) from the same strong component of \( G_p \) the terminal opinions coincide \( \bar{x}_i = \bar{x}_j \) (partial consensus);
3. if the solution is bounded, then the residual vanishes \( \Delta(k) \to 0 \).

The RAI (1.3) establishes consensus if and only if \( G_p \) is strongly connected.

Theorem 4.10 can be further extended: both convergence and consensus appear to be robust against bounded communication delays.

**Theorem 4.11.** Consider a sequence of matrices \((d_{ij}(k))_{i,j=1}^{p}\), whose diagonal entries are zero\(^7\) \( d_{ii}(k) = 0 \), the other entries being uniformly bounded \( 0 \leq d_{ij}(k) \leq d_* < \infty \). Then the criteria from Theorem 4.10 retain their validity for the inequalities

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

Namely, the delayed RAI (4.5) is convergent if Assumptions 4.4, 4.5 and 4.9 hold. Also, if the solution is bounded, then the “residual” vanishes in the sense that

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

\(^7\)In other words, an agent has access to its own undelayed state.
If, additionally, the graph $G_p$ is strongly connected, then the RAI establishes consensus.

If both weights $w_{ij}$ and delays $d_{ij}$ are constant, Assumption 4.4 can be relaxed: it suffices that each strong component of $G[W]$ contains a self-arc (standing for the diagonal entry $w_{ii} > 0$).

Notice that the result of Theorem 4.3, replacing Assumption 4.4 by aperiodicity of each strong component of $G_1[W]$, does not retain its validity even if the delays are constant. A trivial counterexample is the iterative averaging procedure

$$x_1(k+1) = \frac{x_2(k-1) + x_3(k)}{2}, \quad x_2(k+1) = x_1(k-1), \quad x_3(k+1) = x_2(k). \tag{4.6}$$

The corresponding graph $G[W]$ 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: 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 (4.6).

In presence of time-varying delays, even 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 $G[W]$ 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 RAI (4.5) has a non-converging solution

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

Obviously, $x_1(k) \leq x_2(k) = w_{11}x_1(k) + w_{12}x_2(k)$. Also, $x_1(k - d_{21}(k)) = 1$ for any $k$ and hence $x_2(k) = w_{21}x_1(k) + w_{22}x_2(k)$.

### 4.3. Reciprocity condition: discussion

In this subsection, we discuss the relation between our main results and consensus criteria, summarized in Section 3.

#### 4.3.1. Consensus under repeated strong connectivity

As it has been already discussed, Lemma 3.4 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 (4.4) 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 $I$). Theorem 4.10 implies the following counterpart of Lemma 3.4.

**Corollary 4.12.** Assume that Assumptions 4.4 and 4.5 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 (1.3) establishes consensus and (4.3) holds for any bounded solution.

#### 4.3.2. Non-instantaneous type-symmetry

Another example where the assumptions of Theorem 4.10 hold is type-symmetry [27] of matrices

$$w_{ij}(k) \leq CW_{ji}(k) \quad \forall i, j \in [1 : n] \forall k \geq 0. \tag{4.7}$$

where $C \geq 1$ is a constant. Under Assumption 4.5, 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 [5] of the latter condition allows a delayed response:

$$\forall k_0 \geq 0 \quad w_{ij}(k_0) > 0 \implies \sum_{k=k_0}^{k_0+T} w_{ji}(k) > 0. \tag{4.8}$$
Condition (4.7), as well as (4.8), obviously implies Assumption 4.9 with $\mathfrak{M} = 1$. Theorem 4.10 thus extends the result on convergence of type-symmetric consensus algorithms (3.1) [37, Theorem 2], [5, Theorem 5] to RAI (1.3).

4.3.3. Uniform cut-balance and arc-balance. Under Assumption 4.5, the condition of uniform cut-balance (3.8), obviously, implies the validity of Assumption 4.9 (with $\mathfrak{M} = 1$, $T = 0$). More general condition (3.9) also entails the validity of Assumption 4.9 with $\mathfrak{M} = 1$, $T = L$. Indeed, suppose that $a_{1,j}(k_0 : k_1) \geq 1$ and let $k \in [k_0 : k_1]$ be the first instant when $a_{1,j}(k_1') > 0$. Then (3.9) implies that $a_{1,j}(k : (k + L)) > 0$, which proves the implication (4.4) with $T = L$.

In view of Remark 4.6, under Assumption 4.5 the condition of “arc-balance” (3.6) implies that, starting from some instant $k_s$ 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 $\{W(k)\}_{k \geq k_s}$. Similarly, the non-instantaneous arc-balance (3.7) implies that for $k \geq k_s$ 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 4.9 holds with $\mathfrak{M} = 1$, $T = L$.

Theorem 4.10 yields in the following counterpart of Lemmas 3.6 and 3.5.

**Corollary 4.13.** Let Assumptions 4.4, 4.5 hold and (3.9) or (3.7) be valid. Then RAI (1.3) is convergent and consensus in each strong component of $\mathcal{G}_p$ is established and (4.3) holds for any bounded solution. The RAI establishes consensus if and only if $\mathcal{G}_p$ is strongly connected.

4.3.4. Periodic gossiping with intermittent communication. We now propose a simple example where the implication 4.4 does not reduce to any of the aforementioned type-symmetry or balance conditions. A special class of iterative averaging policies (3.1) 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 [4] 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

$$W(k) = \begin{cases} (1 - \alpha_s) e_s e_s^\top + \alpha_s I_n, & k = k_s \\ I_n, & k \neq k_s \forall s. \end{cases} \tag{4.9}$$

For the iterative averaging algorithm (3.1), 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 only restriction is the inequality $x(k + 1) \leq x(k)$), so the convergence is not obvious.

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 (4.9) satisfies the condition (4.4), where $T = 0$ and $\mathfrak{M}$ is the period of the sequence $(j_s, i_s)$. If $a_{1,j}(t_0 : t_1) \geq \mathfrak{M}$ 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 $a_{1,j}(t_0 : t_1) > 0$. Theorem 4.10 now implies the following.

**Corollary 4.14.** Let $W(k)$ have the representation (4.9), where $(j_s, i_s)$ is a periodic sequence of arcs constituting a strongly connected graph $\mathcal{G}_p$ and constants $\alpha_s$ obey the inequality $\eta \leq \alpha_s \leq 1 - \eta$ for some $\eta > 0$. Then the RAI (1.3) establishes consensus and (4.3) holds for any bounded solution.
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 substochastic matrices and associated 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 4.3 leads to an elegant stability criterion [16, 47].

**Lemma 5.1.** 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 $G[A]$ 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 ($G[A]$ 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

$$x_i(k + 1) = \sum_{j=1}^{n} a_{ij}x_j(k - d_{ij}), \quad i \in [1 : n]. \quad (5.1)$$

**Proof.** We are going to show that the vector of opinions obeys RAI (1.3) (or, respectively, (4.5)) with a special constant row-stochastic matrix $W$, “dominating” the matrix $A$ in the sense that $w_{ij} \geq a_{ij}$ for all $i,j$.

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

$$W = A + \frac{1}{n}(I - V)\mathbb{1}_n\mathbb{1}_n^\top \quad (5.2)$$

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 $G[W]$ 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 $G[W]$ is strongly connected. Choosing an arbitrary non-negative vector $x_0 \geq 0$, the vectors $x(k) = A^kx_0$ are non-negative for any $k \geq 0$ and satisfy the inequality (1.3) with $W(k) \equiv W$. Thanks to Theorem 4.3, $x(k) \rightarrow c\mathbb{1}$, where $c \geq 0$ and

$$\Delta(k) = Wx(k) - x(k + 1) = n^{-1}(I - V)\mathbb{1}_n\mathbb{1}_n^\top x(k) \xrightarrow{k \rightarrow \infty} 0. \quad (5.3)$$

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 $A^kx_0 \xrightarrow{k \rightarrow \infty} 0$ for any vector $x_0 \geq 0$. Since every vector is a difference of two non-negative vectors, this entails Schur stability of $A$.

The second statement is proved similarly. Introducing $W$ as in (5.2) and consider a solution to (5.1) with non-negative initial condition is non-negative ($x_i(\tau) \geq 0$ for $-d_s \leq \tau \leq 0$), the solution remains non-negative and obeys the inequalities

$$0 \leq x_i(k + 1) \leq \sum_{j=1}^{n} w_{ij}x_j(k - d_{ij}) \quad \forall i \in [1 : n]. \quad (5.4)$$

In view of Theorem 4.11, consensus is established $x_i(k) \xrightarrow{k \rightarrow \infty} c \geq 0 \forall i$ and

$$\Delta_i(k) = \sum_{j=1}^{n} w_{ij}x_j(k - d_{ij}) - x_i(k + 1) = n^{-1}(1 - v_{ii})\sum_{i=1}^{n} x(k) \xrightarrow{k \rightarrow \infty} 0,$$
i.e. \( c(1 - v_{ij}) = 0 \forall i \) and thus \( c = 0 \). Hence, solutions with non-negative initial conditions vanish as \( k \to \infty \). Due to the linearity of (5.1), the same holds for an arbitrary solution, that is, the system is asymptotically stable. \( \square \)

The condition from Lemma 5.1 is not only sufficient but also necessary for the Schur stability [47]. Lemma 5.1 implies, in particular, the condition of opinion convergence in the Friedkin-Johnsen model of opinion formation [47, 49, 51].

5.2. Hegselmann-Krause model with informed agents (“truth seekers”). One of the seminal models describing opinion formation in social networks is known as the Hegselmann-Krause model of bounded confidence. Its simplest version [25, 30] is a nonlinear modification of the DeGroot algorithm (3.1), where the matrix \( W(k) \) co-evolves with the opinion vector and is defined as follows

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

\[
N_i(x) \triangleq \{ j : |x_j - x_i| < \varepsilon \}.
\] (5.5)

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, ignoring the “deviant” opinions. The constant \( \varepsilon > 0 \) is known as the confidence bound of the agents. Notice that the matrix \( \bar{W}(x) \) is stochastic and satisfies the condition (3.5) 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 (4.7) with \( C = n \). The corresponding graph \( G_p \) is thus undirected and \((j, i) \in E_p\) 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 5.2.** For any sequence of vectors \( x(k) \), the sequence \( W(k) = \bar{W}(x(k)) \) satisfies Assumptions 4.4,4.5 and 4.9.

In view of Proposition (5.2) and Theorem 4.10 every solution to the Hegselmann-Krause model (3.1),(5.5) converges. In fact, the opinion evolution terminates in \( O(n^3) \) steps (see [52] for a historical survey of the relevant results). Also, since \((j, i) \notin E_p\) if and only if \( |x_i(k) - x_j(k)| \geq \varepsilon \) for \( k \) being large (Remark 4.6), terminal opinions \( \bar{x}_i = \lim x_i(k) \) and \( \bar{x}_j = \lim x_j(k) \) either coincide or are sufficiently distant \( \bar{x}_i - \bar{x}_j \geq \varepsilon \).

In this subsection, we consider a more general model with “truth seekers” [26]

\[
x(k + 1) = (I - A)\bar{W}(x(k))x(k) + tA1_n.
\] (5.6)

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 (5.5) and \( t \in \mathbb{R} \) is a constant variable, referred to as truth value. The presence of an additional term in the right-hand side of (5.6) 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 [26]. 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 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 (5.6) visibly changes the dynamics and makes analysis of the model sophisticated. In the original work [26] it
was shown that if \( a_{ii} > 0 \forall i \), then all opinions converge to the truth \( x(k) \xrightarrow{k \to \infty} t \mathbf{1}_n \). Later it was shown \([31]\) that 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 \([11]\) by using a method of power series (“s-energy”). The aforementioned results employ different and highly non-trivial mathematical techniques. Theorem 4.10 allows to examine convergence properties of (5.6) in a simpler way.

**Theorem 5.3.** 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 \( i \) is either a truth-seeker \( (a_{ii} > 0) \) or some truth-seeker \( j \) communicates to it infinitely many times (formally, there exists \( j \) such that \( a_{jj} > 0 \) and \( w_{ij}(k_r) > 0 \) for an infinite sequence \( k_1 < k_2 < \ldots \));
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 central idea of the proof is to exploit the vector \( \xi(k) = (\xi_i(k))_{i=1}^n \) whose elements \( \xi_i(k) \triangleq |x_i(k) - t| \) stand for the distances from the agents’ opinions to the truth. The sequence \( \xi(k) \) appears to be a feasible solution to RAI (1.3), which enables us to use the result of Theorem 4.10. Indeed,

\[
\xi_i(k + 1) = |x_i(k + 1) - t| = (1 - a_{ii}) \left| \sum_{j=1}^n w_{ij}(k)(x_j(k) - t) \right| \leq
\]

\[
(1 - a_{ii}) \sum_{j=1}^n w_{ij}(k)\xi_j(k) \leq \sum_{j=1}^n w_{ij}(k)\xi_j(k) \forall i.
\]

(5.7)

**Step 1.** We are going to prove the sufficiency part of statement 2. In view of Proposition 5.2 and Theorem 4.10, the sequence \( \xi(k) \) converges and remains bounded, since \( \xi(k) \geq 0 \). Therefore, the residual term (4.2) 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 (5.7), \( \Delta_i(k) \geq a_{ii} \sum_{j=1}^n w_{ij}(k)\xi_j(k) \geq a_{ii}w_{ii}(k)\xi_i(k) \). If agent \( i \) is a truth-seeker \( (a_{ii} > 0) \), then \( \xi_i(k) = |x_i(k) - t| \xrightarrow{k \to \infty} 0 \). If a truth-seeker \( j \) communicates to \( i \) infinitely many times, then \((j, i) \in E_p\) and thus \( \xi_j(k) \) 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 communicating to them infinitely often.

**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 \( w_{ij}(k) = 0 \) 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 $\bar{x}(k) = (x_i(k))_{i \in I}$ (after $k_0$ steps) obeys the conventional Hegselmann-Krause model without truth-seekers, and thus its evolution terminates after finite number of steps \cite{11,52}. We have proved statement 3 and the existence of limits $\bar{x}_i$ for all $i \in [1:n]$ (statement 1). Also, from the aforementioned properties of Hegselmann-Krause model 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 now 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 $(i,j) \notin \mathcal{E}_p$. $\square$

Notice that the condition from statement 2 is not easy to check, since the dynamics (5.6) are highly nonlinear. 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 Altafini model of bipartite consensus.

 Altafini’s model of opinion formation, originally proposed in \cite{1} in its continuous-time form, portrays polarization, or “bipartite consensus”, of opinions over structurally balanced signed graphs \cite{56}. In this subsection, we consider a discrete-time modification of the Altafini model, examined in \cite{35,40}. This model is similar to the consensus protocol (3.1), allowing, however, negative influence weights

$$x(k+1) = A(k)x(k) \in \mathbb{R}^n. \tag{5.8}$$

Here the matrix $A(k) = (a_{ij}(k))$ satisfies the following assumption.

**Assumption 5.4.** For any $k \geq 0$, the diagonal entries are non-negative $a_{ii}(k) \geq 0$. The non-negative matrix $W(k) = ([a_{ij}(k)])$ is row-stochastic.

The non-diagonal entries $a_{ij}(k)$ in (5.8) may be both positive and negative. Considering the elements $\xi_i(k)$ as “opinions” of $n$ agents, the positive value $a_{ij}(k) > 0$ can be treated as trust or attraction among agents $i$ and $j$. In this case, agent $i$ shifts its opinion towards the opinion of agent $j$. Similarly, the negative value $a_{ij}(k) < 0$ stands for distrust or repulsion among the agents: the $i$th agent’s opinion is shifted away from the opinion of agent $j$. The central question concerned with the model (5.8) is reaching consensus in absolute value that is, for any $x(0) \in \mathbb{R}^n$ the limits exist

$$\lim_{k \to \infty} |x_1(k)| = \ldots = \lim_{k \to \infty} |x_n(k)|. \tag{5.9}$$

Consensus in absolute value implies that the opinions either reach consensus or polarize: some opinions converge to $M > 0$, the other reaching $(-M)$, where the value $M$ depends on the initial condition. The following criterion of consensus in absolute value extends the results established in \cite{35,40}.

**Theorem 5.5.** Let the sequence $\{W(k)\}$ obey Assumptions 4.4, 4.5 and 4.9. 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 (5.9) is established and, furthermore,

- either the system (5.8) is globally asymptotically stable $x(k) \underset{k \to \infty}{\longrightarrow} 0 \forall x(0)$,

- or the sequence of signed graphs $\mathcal{G}[A(k)]$ is jointly structurally balanced \cite{35},

that is, a diagonal matrix $D$ with entries $d_{ii} \in \{-1,1\}$ such that $W(k) = DA(k)D$ for $k$ sufficiently large$^8$.

$^8$In the case of joint structural balance, one has $\text{sgn} a_{ij}(k) = \text{sgn} d_{ii} \text{sgn} d_{jj}$ whenever $a_{ij} \neq 0$. 


Proof. The absolute values $\xi_i(k) = |x_i(k)|$ obey the inequalities

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

and hence the vector $\xi(k) \geq 0$ composed of them is a solution to the RAI (1.3). In view of Theorem 4.10, the finite limits exist $\bar{\xi}_i = \lim_{k \to \infty} \xi(k) \geq 0$ and

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

We are going to show that $x_i(k)$ converges. In the case where $\bar{\xi}_i = 0$, this is obvious. Assume that $\bar{\xi}_i \not= 0$ and let $\sigma_i(k) = \text{sgn} x_i(k)$. 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

$$
0 \leq \xi_i(k + 1) = x_i(k + 1)\sigma_i(k + 1) \leq a_{ii}(k)\xi_i(k)\sigma_i(k)\sigma_i(k + 1) + \sum_{j \neq i} a_{ij}(k)\xi_j(k),
$$

we obtain that $\Delta_i(k) \geq \min(-2a_{ii}(k)\xi_i(k)\sigma_i(k)\sigma_i(k + 1), 0)$. According to Assumption 4.4, 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, $\xi_i > 0$ and thus for large $k$ one has $\sigma_i(k)\sigma_i(k + 1) = 1$ in view of (5.11), in other words, $\sigma_i(k) \equiv \bar{\sigma}_i$ is constant and $x_i(k) \xrightarrow[k \to \infty]{} \bar{x}_i = \bar{\xi}_i\sigma_i$. If $G_p$ is strongly connected, then (5.9) holds due to Theorem 4.10.

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

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

Obviously, one can suppose that $\bar{x}_i, \bar{x}_j \not= 0$. Since

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

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 $\xi_j(k) \xrightarrow[k \to \infty]{} \bar{\xi}_j \not= 0$ and $\sigma_i(k + 1) \equiv \bar{\sigma}_i = \text{sgn} \bar{x}_i$, $\sigma_j(k + 1) \equiv \bar{\sigma}_j = \text{sgn} \bar{x}_j$ for $k$ being large, one obtains (5.12). Assume now that $G_p$ is strongly connected and at least one solution $x(k)$ does not vanish as $k \to \infty$. Due to (5.9), all components $x_i(k)$ converge to limits $\bar{x}_i \not= 0$. Let $d_{ii} \equiv \text{sgn} \bar{x}_i$. Then, thanks to (5.12) and Assumption 4.5, for $k$ being large, either $a_{ij}(k) = 0$ or $\text{sgn} a_{ij}(k) = d_{ii}d_{jj}$, and thus $(DAD)_{ij} = |a_{ij}(k)| = w_{ij}(k)$ for any $i, j$, that is, $DA(k)D = W(k)$.

As has been discussed, the reciprocity condition from Assumption 4.9 holds if the graph is repeatedly strongly connected, which is a standard condition for consensus in absolute value [35]. In fact, this condition is necessary and sufficient for exponential convergence of the opinions [35], which cannot be examined by using our techniques.

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 $G[A(k)]$ [2, 50], which condition is referred to as structural balance. Opinions in the two factions converge to the opposite values $\xi$ and $-\xi$, where $\xi$ depends on the initial condition; such a property is called “bipartite consensus” [2] or bimodal polarization.
5.4. Constrained consensus and fixed points of paracontractions. In this subsection, we consider another application of the RAI, related to the problem of constrained or “optimal” consensus [34, 44, 57]. As a special case, the problem of constrained consensus includes distributed solving of linear equations [36, 43, 67].

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 $\Omega$, 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

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

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

![Fig. 3: The projection onto a closed convex set](image)

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 $\xi^* \in \Xi = \Xi_1 \cap \ldots \cap \Xi_n$ that satisfies all the constraints, but do not want to communicate the information about sets $\Xi_i$. 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 $\Xi_i$, such a point can be computed by the following modifications of the DeGroot iterative procedure (3.1)

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

(5.14)  

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

(5.15)  

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

(5.16)

The protocol (5.14) has been proposed in the influential paper [44] dealing with distributed optimization problems and then extended (removing some restrictive assumptions) in [34]. The special cases of protocols (5.15) and (5.16) naturally arise in distributed algorithms, solving linear equations\(^9\), see respectively [36, 43] and [67]. A randomized version of (5.16) was examined in [57]. In all of the algorithms, $\xi^i(k)$ stands for an approximation to the desired point, which agent $i$ computes at step $k$. In algorithms (5.14),(5.15) this approximation always satisfies the constraint of agent $i$ ($\xi^i(k) \in \Xi_i$ for $k \geq 1$), whereas (5.16) provides this constraint only asymptotically. We say that constrained consensus is established if the sequences $\xi^i(k)$ converge and

$$\lim_{k \to \infty} \xi^1(k) = \ldots = \lim_{k \to \infty} \xi^n(k) \in \Xi^* \quad \forall \xi^1(0), \ldots, \xi^n(0) \in \mathbb{R}^d.$$  

\(^9\)In this case, $\Xi_i$ are linear hyperplanes
Recently it has been realized [19] 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 (also known as $M$-Fejer mappings [63]).

**Definition 5.6.** 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

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

(5.18)

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 (5.13) implies that the orthogonal projection $P_\Omega$ 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 [19]. Obviously, (5.18) implies

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

(5.19)

Notice that the requirement of constrained consensus (5.17) 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 (5.14)-(5.16) (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 $\Xi_\ast = \bigcap \mathcal{F}(P_i)$? For the algorithm (5.14), the affirmative answer was given in [19], assuming that $\| \cdot \| = \| \cdot \|_p$ with some $1 \leq p \leq \infty$. We prove this result for all algorithms (5.14)-(5.16) and an arbitrary norm on $\mathbb{R}^d$.

Analysis of the algorithms (5.14)-(5.16) relies on a technical lemma, which holds for repeatedly quasi-strongly connected graphs yet fails to hold, in general, under Assumption 4.9. This lemma establishes robustness of the iterative averaging procedure (3.1) against asymptotically vanishing disturbances.

**Lemma 5.7.** Let the matrices $W(k)$ satisfy the conditions of Lemma 3.4 and consider such sequences of vectors $\xi^i(k), \ldots, \xi^n(k) \in \mathbb{R}^d, k = 0, 1, \ldots$ that

$$\xi^i(k + 1) - \sum_{j=1}^n w_{ij}(k)\xi^j(k) \xrightarrow{k \to \infty} 0.$$

(5.20)

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

The proof of Lemma 5.7 retracts\(^{10}\) the proof of Lemma 9 in [34]. The continuous-time counterpart of Lemma 5.7 has been established in [58]. Note that Lemma 5.7 does not guarantee the convergence of $\xi^i(k)$; the latter property requires stronger assumptions on the “disturbance” $e(k)$.

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

**Proposition 5.8.** Let $M$ be a paracontraction in some norm $\| \cdot \|$ and $\xi^0$ 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,

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

\(^{10}\)Formally, [34] adopts some stronger requirements on the matrix $W(k)$ than Lemma 5.7. The proof of Lemma 9 in [34], however, employs only the exponential convergence of backward matrix products $W(k)W(k-1)\ldots W(k_0)$, which can be established [5] under the assumptions of Lemma 3.4.
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 \( \|M\xi_* - \xi_*\| \geq \varepsilon \) yet \( d(\xi_*) = 0 \). One arrives at a contradiction with (5.18), since \( \xi_* \notin \mathcal{F}(M) \) whereas \( \|M\xi_* - \xi_0\| = \|\xi_* - \xi_0\| \). \( \Box \)

We now formulate the main result of this subsection.

**Theorem 5.9.** 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 \( \mathcal{Z}_* = \bigcap_{i=1}^{n} \mathcal{F}(P_i) \neq \emptyset \). Let the assumptions of Corollary 4.12 hold. Then each of the algorithms (5.14)-(5.16) computes (5.17) a common fixed point of \( \{P_i\} \).

**Proof.** We first introduce some auxiliary notation. Fix an arbitrary point \( \xi^0 \in \mathcal{Z}_* \).

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 (1.3), which enable to use the techniques of Theorem 4.10.

**Step 1.** We are going to show first that the sequence \( x(k) \) is a feasible solution to RAI (1.3). In the case of algorithm (5.14), one has

\[
(5.21) \quad x_i(k + 1) = \|P_i(\zeta^i(k)) - \xi^0\| \overset{(5.19)}{\leq} \|\zeta^i(k) - \xi^0\| = \left\| \sum_{j} w_{ij}(k)(\xi^j(k) - \xi^0) \right\| \\
\leq \sum_{j} w_{ij}(k)\|\xi^j(k) - \xi^0\| = \sum_{j} w_{ij}(k)x_j(k).
\]

The case of (5.15) is considered similarly. Denoting \( \tilde{\zeta}^i(k) \overset{\Delta}{=} \sum_{j=1}^{n} w_{ij}(k)P_j(\xi^j(k)) \),

\[
(5.22) \quad x_i(k + 1) = \|P_i(\tilde{\zeta}^i(k)) - \xi^0\| \overset{(5.19)}{\leq} \|\tilde{\zeta}^i(k) - \xi^0\| = \left\| \sum_{j} w_{ij}(k)(P_j(\xi^j(k)) - \xi^0) \right\| \leq \sum_{j} w_{ij}(k)\|P_j(\xi^j(k)) - \xi^0\| \overset{(5.19)}{\leq} \\
\leq \sum_{j} w_{ij}(k)\|\xi^j(k) - \xi^0\| = \sum_{j} w_{ij}(k)x_j(k).
\]

In the case of algorithm (5.16), one has

\[
(5.23) \quad x_i(k + 1) = \left\| w_{ii}(k)(P_i(\xi^i(k)) - \xi^0)) + \sum_{j \neq i} w_{ij}(k)(\xi^j(k) - \xi^0) \right\| \\
\leq w_{ii}(k)\|P_i(\xi^i(k)) - \xi^0\| + \sum_{j \neq i} w_{ij}(k)\|\xi^j(k) - \xi^0\| \overset{(5.19)}{\leq} \\
\leq w_{ii}(k)\|\xi^i(k) - \xi^0\| + \sum_{j \neq i} w_{ij}(k)x_j(k) = \sum_{j=1}^{n} w_{ij}(k)x_j(k).
\]

Using Corollary 4.12, one shows that the RAI establishes consensus, that is, \( x_i(k) \xrightarrow{k \to \infty} c \geq 0 \forall i \) and the residuals \( \Delta_i(k) \) vanish (4.3).
Step 2. Using (4.3), we are going to prove (5.20) or, equivalently,

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

Also, it will be shown that

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

Notice first that the vectors \(\xi^i(k)\) and \(\zeta^i(k)\) are bounded, since their distances \(x_i(k)\) to \(\xi^0\) are bounded. For algorithm (5.14), the inequality (5.21) imply that

\[ \Delta_i(k) \geq \|\zeta^i(k) - \xi^0\| - \|P_i(\zeta^i(k)) - \xi^0\| = \delta_i(\zeta^i(k)). \]

Applying Proposition 5.8 to \(M = P_i\), (4.3) leads to \(P_i(\zeta^i(k)) - \zeta^i(k) \xrightarrow[k \to \infty]{} 0\), which is equivalent to (5.24). To derive (5.25), note that

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

The proof in the case of algorithm (5.16) is simpler. Note that (5.23) yields in

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

Here \(\eta > 0\) is the constant from Assumption 4.4. Applying Proposition 5.8, one proves that (5.25), entailing (5.24) since \(\xi^i(k + 1) = \zeta^i(k) + w_{ii}(k)(P_i(\xi^i(k)) - \zeta^i(k))\).

The proof in the case of algorithm (5.15) combines the two aforementioned estimates. First, similar to (5.26), one derives from (5.22) that

\[ \Delta_i(k) \geq \|\tilde{\xi}^i(k) - \xi^0\| - \|P_i(\tilde{\xi}^i(k)) - \xi^0\| = \delta_i(\tilde{\xi}^i(k)), \]

entailing that \(\|\xi_i(k + 1) - \tilde{\zeta}_i(k)\| = \|P_i(\tilde{\xi}^i(k)) - \tilde{\zeta}^i(k)\| \xrightarrow[k \to \infty]{} 0\). Second, (5.22) also implies (5.27), which in turn implies (5.25). Thus \(\|\tilde{\xi}^i(k) - \zeta^i(k)\| \xrightarrow[k \to \infty]{} 0\). This proves (5.24) since \(\|\xi_i(k + 1) - \tilde{\zeta}^i(k)\| \leq \|\xi_i(k + 1) - \tilde{\zeta}^i(k)\| + \|\tilde{\zeta}^i(k) - \zeta^i(k)\|\).

Since (5.24) is equivalent to (5.20), Lemma 5.7 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 \(\xi^4(k_r) \xrightarrow[r \to \infty]{} \xi_* \in \mathbb{R}^d\). In view of (5.20), we have \(\xi^i(k_r) \xrightarrow[r \to \infty]{} \xi_*\) for every \(i\). The property (5.25) implies that \(P_i(\xi_*) = \xi_*\), and thus \(\xi_* \in \Xi_*\). We are going 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 \(\xi^0\), which can be an arbitrary point in \(\Xi_*\). We have proved that for any such point the distances \(x_i(k) = \|\xi^i(k) - \xi^0\|\) converge to some consensus value \(c_i\), depending on \(\xi^0\) and the initial conditions. Substituting \(\xi^0 = \xi_*\), we obtain that

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

At the same time, we know that \(x_i(k_r) \xrightarrow[r \to \infty]{} 0\). Therefore, \(c_* = 0\), which shows that the limits in (5.17) exist and are equal to \(\xi_* \in \Xi_*\). \(\square\)
Remark 5.10. The assumption on repeated strong connectivity in Theorem 5.9 can be relaxed to the combination of repeated quasi-strong connectivity (required by Lemma 5.7) and Assumption 4.9. Since Lemma 5.7 retains its validity \([34]\) in the case of bounded communication delays, Theorem 4.11 allows to extend the result of Theorem 5.9 to networks with delayed communication, generalizing thus the result of \([34]\) to paracontractions. The corresponding extensions are straightforward and are omitted for the sake of brevity.

6. Proofs of the main results. In this section, we prove Theorems 4.3, 4.10 and 4.11. We start with some auxiliary constructions and technical lemmas.

6.1. Preliminary results. We start with the following simple proposition.

**Proposition 6.1.** For any sequence \(a_1, \ldots, a_m \in [0, 1 - \eta]\), where \(\eta \in (0, 1)\), the inequality holds

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

**Proof.** Since \(\pi \triangleq \pi(a_1, \ldots, a_m) > 0\), one has

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

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

We now derive two important estimates for the solutions of the RAI \((1.3)\). For a non-empty subset of indices \(I \subseteq \{1 : n\}\), denote

\[
M_I(k) \triangleq \max_{i \in I} x_i(k).
\]

Obviously, \(M_I(k) \leq M(k) \triangleq \max_{i \in [1 : n]} x_i(k)\).

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 \(I\) defined as follows

\[
w_{I,J}(k_0 : k_1) \triangleq \sum_{k=k_0}^{k_1} \sum_{i \in I, j \in J} w_{ij}(k).
\]

For brevity, we denote \(w_{I,J}(k) \triangleq w_{I,J}(k : k)\) and \(w_{i,J}(k_0 : k_1) \triangleq w_{i,J}(k_0 : k_1)\).

**Lemma 6.2.** Assume that Assumption 4.4 holds. For any instants \(k_0 \geq 0, k_1 \geq k_0\) and the sets \(I \subseteq \{1 : n\}, J = I^c\) the solution \(x(\cdot)\) obeys the inequality

\[
M_I(k_1 + 1) \leq \vartheta M_I(k_0) + (1 - \vartheta)M(k_0), \quad \vartheta \triangleq \exp(-\eta w_{I,J}(k_0 : k_1)),
\]

**Proof.** We know (Proposition 4.2) that \(x_i(k) \leq M(k) \leq M(k_0)\) \(\forall k \geq k_0 \forall i \in \{1 : n\}\). Denoting \(\rho(k) \triangleq \max_{i \in I} w_{i,J}(k)\), for each \(k \geq k_0\) and \(i \in I\), one has

\[
x_i(k + 1) \leq \left(\sum_{j \in I} w_{ij}(k) x_j(k) + \sum_{j \in J} w_{ij}(k) x_j(k)\right) = (1 - w_{i,J}(k))M_I(k) + w_{i,J}(k)M(k_0) = M(k_0) - (1 - w_{i,J}(k))(M(k_0) - M_I(k)) \leq M(k_0) - (1 - \rho(k))(M(k_0) - M_I(k)).
\]
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 \( w_{i,J}(k) \leq 1 - w_{i,J}(k) \leq \eta \) and also \( w_{i,J}(k) \leq w_{I,J}(k) \) for each \( i \in I \), Proposition 6.1 implies (6.2), since
\[
M(k_0) - M_I(k_1 + 1) \geq (M(k_0) - M_I(k_0)) \exp \left( -\eta \sum_{k=k_0}^{k_1} \rho(k) \right) \geq (M(k_0) - M_I(k_0)) \exp \left( -\eta \sum_{k=k_0}^{k_1} w_{I,J}(k) \right) = (M(k_0) - M_I(k_0)) \theta.
\]
Lemma 6.2 retains its validity for the delayed RAI (4.5) with a minor modification.

**Lemma 6.3.** Let Assumption 4.4 hold. For a solution of the RAI (4.5), denote \( \bar{M}(k) \overset{\Delta}{=} \max \{ M(k), M(k-1), \ldots, M(k-d_s) \} \), \( \bar{M}_I(k) \overset{\Delta}{=} \max \{ M_I(k), \ldots, M_I(k-d_s) \} \).

For any instants \( k_1 \geq k_0 \geq 0 \) and cut \( (I, J) \) the inequality holds
\[
(6.4) \quad x_i(k+1) \leq \bar{\theta} \bar{M}_I(k_0) + (1 - \bar{\theta}) \bar{M}(k_0) \forall i \in I, \quad \bar{\theta} \equiv \eta^{d_s} \exp(-\eta^{d_s+1} w_{I,J}(k_0 : k_1)).
\]

**Proof.** Similar to Proposition 4.2, it is easy to show that the sequence \( \bar{M}(k) \) is non-increasing. Denoting \( \rho(k) \overset{\Delta}{=} \min_{i \in J} w_{i,J}(k) \), for each \( k \geq k_0 \) and \( i \in I \), the following inequality is derived similar to (6.3)
\[
(6.5) \quad x_i(k+1) \leq (1 - w_{i,J}(k)) M_I(k_0) + w_{i,J}(k) \bar{M}(k_0) \leq \bar{M}(k_0) - w_{i,J}(k) \bar{M}_I(k_0) \leq \bar{M}(k_0) - \rho(k) (\bar{M}(k_0) - \bar{M}_I(k_0)).
\]
Besides this, notice that for any \( k' \geq k_0 \) and any \( i \)
\[
(6.6) \quad x_i(k' + 1) \leq w_{i,J}(k') x_i(k') + (1 - w_{i,J}(k')) \bar{M}(k_0) \leq \bar{M}(k_0) - \eta(\bar{M}(k_0) - x_i(k')), \quad \bar{M}(k_0) - x_i(k' + s) \geq \eta^s (\bar{M}(k_0) - x_i(k')) \forall s \geq 1.
\]
Combining this with (6.5), one proves that for \( i \in I \) and \( k \geq k_0 \),
\[
x_i(k + s) \leq \bar{M}(k_0) - \eta^{d_s} \rho(k) (\bar{M}(k_0) - \bar{M}_I(k)) \forall s = 1, \ldots, d_s,
\]
that is, \( \bar{M}(k_0) - \bar{M}_I(k + d_s) \geq (1 - \eta^{d_s} \rho(k)) (\bar{M}(k_0) - \bar{M}_I(k)) \). Denote now \( m \overset{\Delta}{=} \lfloor (k_1 - k_0)/d_s \rfloor \), so that \( k_0 + m d_s \leq k_1 < k_0 + (m + 1)d_s \). Then,
\[
\bar{M}(k_0) - \bar{M}_I(k_0 + m d_s) \geq \prod_{j=0}^{m-1} (1 - \eta^{d_s} \rho(k_0 + jd_s)) (\bar{M}(k_0) - \bar{M}_I(k_0)) \geq \exp \left[ -\eta^{d_s+1} \sum_{j=0}^{m-1} \rho(k + jd_s) \right] (\bar{M}(k_0) - \bar{M}_I(k_0)) \geq \exp \left( -\eta^{d_s+1} \bar{w}_{I,J}(k_0 : k_1) \right) (\bar{M}(k_0) - \bar{M}_I(k_0)) = \bar{\theta} \eta^{-d_s} (\bar{M}(k_0) - \bar{M}_I(k_0)),
\]
or, equivalently, \( \bar{M}(k_0) - x_i(k_0 + m d_s) \geq \bar{\theta} \eta^{-d_s} (\bar{M}(k_0) - \bar{M}_I(k)) \) for each \( i \in I \). Choosing now \( m \) in such a way that \( m d_s \leq k_1 - k_0 < (m + 1)d_s \) and substituting \( k' = k_0 + m d_s \) and \( s = k_1 + 1 - k' \leq d_s \) into (6.6), one arrives at \( \bar{M}(k_0) - x_i(k_1 + 1) \geq \eta^{k_1 + 1 - k_0 - m d_s} (\bar{M}(k_0) - x_i(k_0 + m d_s)) \geq \bar{\theta} (\bar{M}(k_0) - \bar{M}_I(k)). \]
COROLLARY 6.4. Suppose that Assumptions 4.4,4.5 and 4.9 hold and $G_p$ is strongly connected. Assume that for some cut $(I,J)$ and $k_0 \geq 0$, the solution of the RAI (1.3) satisfies the condition $M(k_0) - x_i(k_0) \geq \varepsilon > 0$ for any $i \in I$. Then there exist $k_s \geq k_0$, $j_0 \in J = I^*$ and constant $\varrho = \varrho(\eta, W, T) \in (0, 1)$ such that the inequality holds

\[ x_i(k_s + 1) \leq M(k_0) - \varrho \varepsilon \quad \forall i \in I \cup \{j_0\}. \]

(6.7)

Analogously, if the solution of the delayed RAI (4.5) satisfies the condition $\bar{M}(k_0) - x_i(k_0) \geq \varepsilon > 0$ for any $i \in I$, then for some constant $\bar{\varrho} \in (0,1)$ and $j_0 \in J$ one has

\[ x_i(k_s + 1) \leq \bar{M}(k_0) - \bar{\varrho} \varepsilon \quad \forall i \in I \cup \{j_0\}. \]

(6.8)

Proof. Consider first a feasible solution to the undelayed RAI (1.3) such that $M_I(k_0) \leq M(k_0) - \varepsilon$. By using Lemma 6.2 and noticing that $\theta \geq \vartheta \triangleq e^{-\eta M_0} \in (0, 1)$, one has $M_I(k) \leq M(k_0) - \vartheta \varepsilon$ for any $k \in [k_0 : k_s]$. Since $G_p$ is strongly connected, $a_{I,J}(k_0 : \infty) = \infty$. Let $k_s \geq k_0$ be the first instant such that $a_{I,J}(k_s : k_s) > 0$. In view of Assumption 4.9 and (4.4), either $k_s - k_0 < T$ or $a_{I,J}(k_s : k_s) < W$. In both situations, $a_{I,J}(k_s : k_s) < W_0 \triangleq W + n^2 T$ and thus $w_{I,J}(k_s : k_s) < W_0$ since $w_{I,J} \leq 1$. By assumption, at time $k_s$ there exists an arc $(i_0, j_0)$ in $G[W(k)]$ connecting $i_0 \in I$ to $j_0 \in J$, that is, $w_{j_0 i_0}(k_s) \geq \eta$. Therefore,

\[ x_{j_0}(k_s + 1) \leq w_{j_0 i_0}(k_s) x_{i_0}(k_s) + (1 - w_{j_0 i_0}(k_s)) M(k_0) = M(k_0) - w_{j_0 i_0}(k_s)(M(k_0) - x_{i_0}(k_s)) = M(k_0) - \varrho \varepsilon. \]

(6.9)

Therefore, (6.7) holds with $\varrho \triangleq \eta e^{-\eta M_0}$.

The proof in the delayed case is similar with several modifications. First, the condition $\bar{M}(k_0) - x_i(k_0) \geq \varepsilon > 0$ for any $i \in I$ does not imply, in general, that $\bar{M}(k_0) - \bar{M}_I(k_0) \geq \varepsilon$. Notice, however, that in view of (6.6),

\[ \bar{M}(k_0) - x_i(k_0 + s) \geq \eta^\ast(\bar{M}(k_0) - x_i(k_0)) \geq \eta^\ast \varepsilon \quad \forall i \in I \forall s = 0, \ldots, d_\ast. \]

(6.10)

Therefore, denoting $k'_0 = k_0 + d_\ast$, one shows that $M_I(k'_0) \leq M(k_0) - \eta^\ast \varepsilon$. Now let $k_s \geq k'_0$ be the first instant such that $a_{I,J}(k_s : k_s) > 0$, as we have seen, in this situation $w_{I,J}(k'_0 : k_s) < W_0 = W + n^2 T$. Substituting $k_0 = k'_0$ and an arbitrary $k_1 \in [k'_0 : k_s]$ into (6.4), recalling that $\bar{M}(k'_0) \leq \bar{M}(k_0)$ and noticing that $\bar{\theta} \geq \bar{\vartheta} \triangleq \eta^\ast \exp(-\eta^\ast W_0)$, one arrives at the inequality

\[ x_i(k_1 + 1) \leq \bar{M}(k_0) - \bar{\varrho} \eta^\ast \varepsilon \quad \forall i \in I \forall k_1 \in [k'_0 : k_s]. \]

(6.11)

By assumption, at time $k_s$ there exists an arc $(i_0, j_0)$ in $G[W(k)]$ connecting $i_0 \in I$ to $j_0 \in J$, that is, $w_{j_0 i_0}(k_s) \geq \eta$. For $d_0 \triangleq d_{j_0 i_0}(k_s)$, a counterpart of (6.9) holds

\[ x_{j_0}(k_s + 1) \leq w_{j_0 i_0}(k_s) x_{i_0}(k_s - d_0) + (1 - w_{j_0 i_0}(k_s)) M(k_0) = \bar{M}(k_0) - w_{j_0 i_0}(k_s)(\bar{M}(k_0) - x_{i_0}(k_s - d_0)). \]

(6.12)

Obviously, $k_s - d_0 \geq k'_0 - d_0 \geq k'_0 - d_\ast$. Two situations are now possible: either $k_s - d_0 \leq k'_0$ or $k_s - d_0 \geq k'_0 + 1$. In the first situation, one has $x_{i_0}(k_s - d_0) \leq M_I(k'_0) \leq M(k_0) - \eta^\ast \varepsilon \leq M(k_0) - \eta^\ast \bar{\vartheta} \varepsilon$. In the second situation, (6.11) entails that $x_{i_0}(k_s - d_0) \leq \bar{M}(k_0) - \eta^\ast \bar{\vartheta} \varepsilon$. The inequality (6.12) entails now (6.8) with $\bar{\varrho} = \eta^\ast \bar{\varrho} = \eta^\ast \bar{\vartheta} \exp(-\eta^\ast W_0)$.
6.2. Proof of Lemma 4.7. To prove the first statement, assume that the RAI (1.3) 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. Denote the set of its nodes by \( I \subseteq \{1 : n\} \) and let \( J \triangleq I^c \). Since \((j, i) \notin \mathcal{E}_p\) for any \( i \in I, j \in J \), one has \( w_{i,j}(0 : \infty) < \infty \). Denote \( a_k = \max_{i \in I} w_{i,j}(k) \). In view of Assumption 4.4, \( w_{i,j}(k) \leq 1 - w_{i,i}(k) \leq 1 - \eta \). 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 6.1 entails that that \( \zeta(\infty) > 0 \) since \( \sum_k a_k < \infty \). Notice also that

\[
(6.13) \quad \zeta(k + 1) = (1 - a_k)\zeta(k) \leq \zeta(k) - \sum_{j \in J} w_{i,j}(k)\zeta(k) = \sum_{j \in J} w_{i,j}(k)\zeta(k) \quad \forall i \in I.
\]

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 are now ready to construct a solution to (1.3) that does not converge. Let

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

Obviously, \( x_i(k_s) \) does not converge as \( s \to \infty \), so \( x(k) \) fails to have a limit. We are going to prove that \( x(k) \) is a solution to the RAI, that is,

\[
(6.14) \quad x_i(k + 1) \leq \sum_{j=1}^{n} w_{ij}(k)x_j(k).
\]

For \( i \in I \), the latter inequality follows from (6.13) since

\[
x_i(k + 1) \leq \sum_{j \in I} w_{ij}(k)x_j(k) \leq \sum_{j \in I} w_{ij}(k)x_j(k) + \sum_{j \in J} w_{ij}(k)x_j(k) \geq \sum_{j=1}^{n} w_{ij}(k)x_j(k).
\]

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, (6.14) holds also for \( i = j_0 \). For \( i \in J \setminus \{j_0\} \), (6.14) is obvious since \( x_j(k) \geq 0 \forall j \).

The second statement is immediate from Lemma 3.3. If consensus is established by the RAI (1.3), it is automatically established by the iterative averaging procedure (3.1), and hence \( \mathcal{G}_p \) is quasi-strongly connected. Since all strong components of \( \mathcal{G}_p \) are isolated, \( \mathcal{G}_p \) has only one such component.

Remark 6.5. Notice that the construction of the oscillatory solution works also for the delayed RAI (4.5), 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 situation, \( x_i(k - d) \geq \xi(k) > \zeta(\infty) \) for any \( k \geq 0, 0 \leq d \leq d_s \) and \( i \in I \).

\[11\]In other words, if the acyclic condensed graph \[24\] has at least one arc, it should have a source node with at least one arc coming from it.
Remark 6.6. In the proof of Lemma 4.7, we in fact used a relaxed form of Assumption 4.4: \( w_i(k) \geq \eta \) for each node of the graph, belonging to a source component. This will be used in the proof of Theorem 4.3.

6.3. Proof of Theorem 4.10. We first notice that Assumptions 4.5 and 4.9 imply that the strong components of \( G_p \) are isolated. Indeed, (4.4) entails that if \( a_{j,i}(0 : \infty) = \infty \), then \( a_{j,i}(0 : \infty) = \infty \) for any cut \((I, J)\). Thus, if a path from \( J \) to \( I \) exists in \( G_p \) (which automatically contains an arc \((j,i)\) with \( j \in J, i \in E \)) then a path (and in fact, an arc) from \( I \) to \( J \) exists as well.

In view of Remark 4.6, starting from some instant \( k \geq k_0 \) only persistent arcs exist, and, renumbering the agents, the matrix \( W(k) \) becomes block-diagonal \( W(k) = \text{diag}(W_{11}(k), \ldots, W_{ss}(k)) \), where \( \{W_{jj}(k)\}_{k \geq k_0} \) are sequences of stochastic matrices corresponding to strongly connected graphs \( G_p^j \). Since the RAI decomposes into several independent inequalities, it suffices to consider the case of a strongly connected persistent graph. We are going to show that if \( G_p \) is strongly connected, then the RAI establishes consensus and (4.3) holds for all bounded solutions.

For each vector \( x(k) \in \mathbb{R}^n \), let \( \sigma_1(k), \ldots, \sigma_n(k) \) be a permutation of the indices sorting its component in the ascending order, that is

\[
y_1(k) \overset{\Delta}{=} x_{\sigma_1(k)}(k) \leq y_2(k) \overset{\Delta}{=} x_{\sigma_2(k)}(k) \leq \ldots \leq y_n(k) \overset{\Delta}{=} x_{\sigma_n(k)}(k) = M(k).
\]

Since \( y_n(k) = M(k) \) is non-increasing (Proposition 4.2), there exists the limit

\[
y_* \overset{\Delta}{=} \lim_{k \to \infty} y_n(k) \geq -\infty
\]

To prove consensus, it suffices to show that

\[
y_j(k) \overset{\Delta}{\longrightarrow} y_* \quad \forall j.
\]

For \( y_* = -\infty \), the statement is obvious since \( y_j(k) \leq y_n(k) \). For \( y_* > -\infty \), we will now prove (6.16) via backward induction on \( j = n, n-1, \ldots, 1 \). The induction base \( j = n \) is true by definition of \( y_* \). Assume that the statement is proved for \( j = r+1, r+2, \ldots, n \), then \( \lim_{k \to \infty} y_r(k) \leq \lim_{k \to \infty} y_{r+1}(k) = y_* \). It remains to show that \( \lim_{k \to \infty} y_r(k) \geq y_* \). Assume, on the contrary, that a sequence \( k_1 < k_2 < \ldots \) and \( \varepsilon > 0 \) exist such that \( y_r(k_j) \leq y_* - \varepsilon \). For a fixed \( s \), let \( I_s = \{\sigma_1(k_s), \ldots, \sigma_r(k_s)\} \) and \( J_s = I_s^c \). By construction, \( M_{I_s}(k_s) = y_r(k_s) \leq M(k_s) - \varepsilon \). Corollary 6.4 (applied to \( k_0 = k_s \) and \( I = I_s \)) ensures the existence of \( k_s^* \geq k_s \) and an index \( j_s \in J_s \) such that

\[
x_i(k_s^* + 1) \leq M(k_s) - \varepsilon \quad \forall i \in I_s^c \overset{\Delta}{=} I_s \cup \{j_s\},
\]

and therefore, since \( |I_s^c| = r + 1 \), we have \( y_{r+1}(k_s^* + 1) \leq M(k_s) - \varepsilon \). When \( s \to \infty \), we have \( M(k_s) \to y_* \) and, due to the induction hypothesis, \( y_{r+1}(k_s^* + 1) \to y_* \), which leads to a contradiction. The induction step is proved.

If the solution is bounded, then \( y_* > -\infty \) and \( x(k) \overset{\Delta}{\longrightarrow} y_* \|_n \). Therefore,

\[
0 \leq \Delta(k) = W(k)x(k) - x(k+1) = W(k)[x(k) - y_* \|_n] - (x(k+1) - y_*) \overset{\Delta}{\longrightarrow} 0
\]

(in the latter equation, we use the fact that \( W(k) \) is stochastic, thus \( W(k) \|_n = \|_n \) and all entries of \( W(k) \) are uniformly bounded \( 0 \leq w_{ij}(k) \leq 1 \)).

The “only if” part in the last statement is immediate from Lemma 4.7.
\textbf{6.4. Proof of Theorem 4.3.} 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 (1.3) splits into several independent RAI). If \( \mathcal{G}[W] \) is strongly connected and aperiodic, then \( W \) is a \textit{primitive} matrix \([20, 51]\), and hence \( W^* \) has positive entries for sufficiently large \( s \geq 0 \), satisfying thus Assumptions 4.4 and 4.5. Since \( x(k + s) \leq W^* x(k) \quad \forall k \geq 0 \), Theorem 4.10 (applied to \( W(k) \equiv W^* \)) entails that each subsequence \( x(sm + j) \), \( m = 0, 1, \ldots \) and \( j = 0, \ldots, s-1 \), converges to a consensus vector \( x(sm + j) \xrightarrow{m \to \infty} c_j \mathbb{1}_n \), \( c_j \in [-\infty, \infty) \). Since
\[
x(sm + j + 1) \leq Wx(sm + j),
\]
one obtains that \( c_0 \geq c_1 \geq \ldots c_{s-1} \geq c_0 \), that is, \( c_0 = \ldots = c_{s-1} = c \) and the RAI establishes consensus. The proof of (4.3) 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 3.1.

To prove the converse statement, notice first that if the RAI (1.3) is convergent, the same holds for the RAI
\[
z(m + 1) \leq W^* z(m),
\](6.18)
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 (1.3). Also, any solution to the DeGroot’s system (1.2) is a feasible solution to the RAI (1.3). Hence, the RAI can be convergent only if all solutions of (1.2) converge, which means that each \textit{source} component of \( \mathcal{G}[W] \) is an aperiodic graph \([9, 52]\). For a source component of \( \mathcal{G}[W] \) with the set of nodes \( I \), \( w_{ij} = 0 \) for any \( i \in I \) and \( j \not\in I \). Hence, the corresponding submatrix \( W_I \triangleq (w_{ij})_{i,j \in I} \) is row-stochastic, irreducible and aperiodic (primitive). In particular, for \( s \) being sufficiently large, the matrix \( (W_I)^s \) (being a submatrix of \( W^* \)) has strictly positive entries. Obviously, \( I \) remains a source component in \( \mathcal{G}[W^*] \); also, if there is an arc coming out of \( I \) in \( \mathcal{G}[W] \), that is, \( w_{ji} > 0 \) for some \( j \not\in I \) and \( i \in I \), the same arc exists in the graph \( \mathcal{G}[W^*] \) for large \( s \).

In view of Remark 6.6, Lemma 4.7 is applicable to RAI (6.18) in spite of potential violation of Assumption 4.4. Hence, all strong components of \( \mathcal{G}[W^*] \) (in particular, all source components) are isolated for every sufficiently large \( s \). Therefore, the \textit{source} components of \( \mathcal{G}[W] \) are also isolated. As discussed in the proof of Lemma 4.7, 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 (1.2) (and also the RAI) cannot provide consensus. This finishes the proof of “only if” parts in the first and the last statements.

\textbf{6.5. Proof of Theorem 4.11.} The proof of Theorem 4.11 for the time-varying case retraces the proof of “if” part in Theorem 4.10 with the following minor change.

The value \( M(k) = y_n(k) \) is no longer monotone, for this reason we define
\[
y_* \triangleq \lim_{k \to \infty} \overline{M}(k),
\]
where \( \bar{M}(k) \) is introduced in Lemma 6.3. To show that \( y_j(k) \xrightarrow[k \to \infty]{} y_* \), we again use induction on \( j = n, \ldots, 1 \). To prove the induction base \( j = n \), notice that \( y_n(k) = \bar{M}(k) \leq \bar{M}(k_0) \), so the induction base is true if \( y_n = -\infty \). Otherwise, one can use the inequalities (6.6) (valid for any \( k_0 \geq 0 \) and \( k' \geq k_0 \)) in order to show that

\[
\bar{M}(k_0 + d_*) \leq \bar{M}(k_0) - y^{d_*}(\bar{M}(k_0) - y_n(k_0))
\]

Passing to the limit as \( k_0 \to \infty \), one shows that \( \bar{M}(k_0) - y_n(k_0) \xrightarrow[k \to \infty]{} 0 \) if \( y_n > -\infty \).

The induction step is proved in the same way as in Theorem 4.10 with the only difference that (6.17) is replaced by

\[
(6.19) \quad x_i(k_0^n + 1) \leq \bar{M}(k_s) - \bar{c} \quad \forall i \in I_s \triangleq I_s \cup \{j_s\},
\]

which holds due to Corollary 6.4. The proof of (4.3) has to be modified in the following way: if a solution is bounded and \( x_i(k) \to c > -\infty \), then

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

In the time-invariant case, it suffices to consider the case where \( G[W] \) 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 in [66]: consider the vector \( y(k) \in \mathbb{R}^{d \times n} \) 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 [66] that

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

Here \( W_i \) are non-negative matrices such that \( W = W_0 + \ldots + W_{d_*} \); \( (W_i)_{ij} = w_{ij} \) if and only if \( d_{ij} = r \). In particular, \( \text{diag } W_0 = \text{diag } W \) and, if \( \text{diag } W \neq 0 \), the graph \( G[\Xi] \) has a self-arc and is aperiodic. It remains to notice that each arc \((j, i)\) in the graph \( G[W] \) corresponds to a walk \( j \to (j + d_*) \to (j + 2d_*) \to \ldots \to (j + d_{ij}d_*) \to i \) in the graph \( G[\Xi] \), therefore, the graph \( G[\Xi] \) is strongly connected. Applying Theorem 4.3 to the matrix \( \Xi \), one shows that the RAI (4.5) establishes consensus.

7. Conclusion. The recurrent averaging inequalities (RAI) introduced in (1.3) can be considered as a relaxed form of the conventional dynamics of iterative averaging (or the French-DeGroot model) (1.2). While in the standard iterative averaging each agent updates its opinion with a weighted average of its own and the neighbors’ opinions, in the RAI each agent is allowed to choose any opinion which does not exceed that linear combination. Such a constraint may seem non-restrictive however, under some connectivity assumptions, all feasible solutions of the RAI converge and even reach consensus. Similarly to the classical iterative averaging, consensus is robust to bounded communication delays. Especially interesting are bounded solutions to the RAI, for which the residual vector (defined as 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 derived a number of known and new results in the unified way, in particular:
1. a known graph-theoretic stability criterion for substochastic matrices and its novel extension for time-delay systems;
2. convergence of opinions in the Hegselmann-Krause model with informed agents (“truth-seekers”);
3. convergence of opinions in the discrete-time Altafini model;
4. a convergence criterion for constrained consensus, linear equation solving and finding a common fixed point for a family of paracontractions, where we showed that all types of algorithms proposed in the literature can be examined in a unified way and, furthermore, some restrictive assumptions imposed in the previous papers can be discarded.

It should be noticed that some of the assumptions adopted throughout the paper, in particular, the uniform positivity of non-zero weights (Assumption 4.5) can be relaxed or completely discarded in special situations, e.g., when the type-symmetry condition holds (4.7). Another possible extension is concerned with nonlinear RAI, obtained from nonlinear averaging consensus algorithms [42]. These extensions are beyond the scope of this paper and are the object of current investigation.

References.

[1] C. Altafini, Dynamics of opinion forming in structurally balanced social networks, PLoS ONE, 7 (2012), p. e38135.
[2] C. Altafini, Consensus problems on networks with antagonistic interactions, IEEE Trans. Autom. Control, 58 (2013), pp. 935–946.
[3] N. Amelina, A. Fradkov, Y. Jiang, and D. Vergados, Approximate consensus in stochastic networks with application to load balancing, IEEE Trans. Inform. Theory, 61 (2015), pp. 1739–1752.
[4] B. D. O. Anderson, C. Yu, and A. S. Morse, Convergence of Periodic Gossiping Algorithms, Springer, Berlin, Heidelberg, 2010, pp. 127–138.
[5] V. Blondel, J. Hendrickx, A. Olshevsky, and J. Tsitsiklis, Convergence in multiagent coordination, consensus, and flocking, in Proc. IEEE Conf. Decision and Control, 2005, pp. 2996 – 3000.
[6] S. Bolouki and R. P. Malhame, Linear consensus algorithms based on balanced asymmetric chains, IEEE Trans. Autom. Control, 60 (2015), pp. 2808–2812.
[7] S. Bolouki and R. P. Malhame, Consensus algorithms and the decomposition-separation theorem, IEEE Transactions on Automatic Control, 61 (2016), pp. 2357–2369.
[8] V. Borkar and P. Varaiya, Asymptotic agreement in distributed estimation, IEEE Trans. Autom. Control, 27 (1982), pp. 650–655.
[9] F. Bullo, Lectures on Network Systems, published online at http://motion.me.ucsb.edu/book-lns, 2016. With contributions by J. Cortes, F. Dorfler, and S. Martinez.
[10] M. Cao, A. Morse, and B. Anderson, Reaching a consensus in a dynamically changing environment, SIAM J. Control Optim., 47 (2008), pp. 575–623.
[11] B. Chazelle, The total s-energy of a multiagent system, SIAM J. Control Optim., 49 (2011), pp. 1680–1706.
[12] D. Bertsekas and J. Tsitsiklis, Parallel and Distributed Computation: Numerical Methods, Prentice-Hall, Englewood Cliff, NJ, 1989.
[13] M. DeGroot, Reaching a consensus, Journal of the American Statistical Association, 69 (1974), pp. 118–121.
[14] R. Dobrushin, Central limit theorem for non-stationary Markov chains I, II,
Theorem Probab. Appl., 1 (1956), pp. 63–80,233–246.

[15] L. Fang and P. Antsaklis, Asynchronous consensus protocols using nonlinear paracontractions theory, IEEE Trans. Automat. Control, 53 (2008), pp. 2351–2355.

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

[17] J. French Jr., A formal theory of social power, Phychol. Rev., 63 (1956), pp. 181–194.

[18] N. Friedkin, The problem of social control and coordination of complex systems in sociology: A look at the community cleavage problem, IEEE Control Syst. Mag., 35 (2015), pp. 40–51.

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

[20] F. Gantmacher, The Theory of Matrices, vol. 2, AMS Chelsea Publishing, 2000.

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

[22] J. Hajnal, Weak ergodicity in non-homogeneous Markov chains, Proc. Cambridge Philos. Soc., 54 (1958), pp. 233–246.

[23] F. Harary, A criterion for unanimity in French’s theory of social power, in Studies in Social Power, D. Cartwright, ed., Univ. of Michigan Press, Ann Arbor, MI, 1959, pp. 168–182.

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

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

[26] R. Hegselmann and U. Krause, Truth and cognitive division of labour: First steps towards a computer aided social epistemology, Journal of Artificial Societies and Social Simulation (JASSS), 9 (2006), p. 1.

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

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

[29] A. Jadlabaie, J. Lin, and A. Morse, Coordination of groups of mobile autonomous agents using nearest neighbor rules, IEEE Trans. Automat. Control, 48 (2003), pp. 988–1001.

[30] U. Krause, A discrete nonlinear and non-autonomous model of consensus formation, Commun. in Difference Equations, (2000), pp. 227–236.

[31] S. Kurz and J. Rambau, On the Hegselmann–Krause conjecture in opinion dynamics, J. Difference Equations and Applications, 17 (2011), pp. 859–876.

[32] K. Lehrer, When rational disagreement is impossible, Noûs, 10 (1976), pp. 327–332.

[33] A. Leizarowitz, On infinite products of stochastic matrices, Linear Algebra
and its Applications, 168 (1992), pp. 189 – 219.

[34] P. Lin and W. Ren, Constrained consensus in unbalanced networks with communication delays, IEEE Trans. Autom. Control, 59 (2014), pp. 775–781.

[35] J. Liu, X. Chen, T. Başar, and M. Belabbas, Exponential convergence of the discrete- and continuous-time Altafini models, IEEE Trans. Autom. Control, 62 (2017), pp. 6168–6182.

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

[37] J. Lorenz, A stabilization theorem for dynamics of continuous opinions, Physica A, 353 (2005), pp. 217–223.

[38] S. Martin and J. Hendrickx, Continuous-time consensus under non-instantaneous reciprocity, IEEE Trans. Autom. Control, 61 (2016), pp. 2484–2495.

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

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

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

[42] L. Moreau, Stability of multiagent systems with time-dependent communication links, IEEE Trans. Autom. Control, 50 (2005), pp. 169–182.

[43] S. Mou, J. Liu, and A. Morse, A distributed algorithm for solving a linear algebraic equation, IEEE Trans. Autom. Control, 60 (2015), pp. 2863–2878.

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

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

[46] A. Olshevsky and J. Tsitsiklis, Convergence speed in distributed consensus and averaging, SIAM Rev., 53 (2011), pp. 747–772.

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

[48] A. Proskurnikov and M. Cao, Differential inequalities in multi-agent coordination and opinion dynamics modeling, Automatica, 85 (2017), pp. 202–210.

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

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

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

[52] A. Proskurnikov and R. Tempo, A tutorial on modeling and analysis of dynamic social networks. Part II, Annual Reviews in Control, 45 (2018), 166–190.
[53] W. Ren and R. Beard, *Distributed Consensus in Multi-Vehicle Cooperative Control: Theory and Applications*, Springer-Verlag, London, 2008.

[54] W. Ren and Y. Cao, *Distributed Coordination of Multi-agent Networks*, Springer, 2011.

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

[56] G. Shi, C. Altafini, and J. Baras, *Dynamics over signed networks*, SIAM Review, 61 (2019), pp. 229–257.

[57] G. Shi and K. Johansson, *Randomized optimal consensus of multi-agent systems*, Automatica, 48 (2012), pp. 3018–3030.

[58] G. Shi and K. Johansson, *Robust consensus for continuous-time multi-agent dynamics*, SIAM J. Control Optim, 51 (2013), pp. 3673–3691.

[59] G. Shi and K. Johansson, *The role of persistent graphs in the agreement seeking of social networks*, IEEE J. Selected Areas Commun., 31 (2013), pp. 595–606.

[60] B. Touri and A. Nedić, *On ergodicity, infinite flow, and consensus in random models*, IEEE Trans. Autom. Control, 56 (2011), pp. 1593–1605.

[61] B. Touri and A. Nedić, *On backward product of stochastic matrices*, Automatica, 48 (2012), pp. 1477 – 1488.

[62] J. Tsitsiklis, D. Bertsekas, and M. Athans, *Distributed asynchronous deterministic and stochastic gradient optimization algorithms*, IEEE Trans. Autom. Control, 31 (1986), pp. 803–812.

[63] P. Wang, W. Ren, and Z. Duan, *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 (2019), pp. 82–93.

[64] J. Wolfowitz, *Products of indecomposable, aperiodic, stochastic matrices*, Proceedings of Amer. Math. Soc., 15 (1963), pp. 733–737.

[65] W. Xia, G. Shi, Z. Meng, M. Cao, and K. H. Johansson, *Persistent flows in deterministic chains (publ. online)*, IEEE Trans. Autom. Control, (2019).

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

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