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

Anton V. Proskurnikov\textsuperscript{a,b,1,2,*,} Giuseppe C. Calafiore\textsuperscript{a}, Ming Cao\textsuperscript{c}

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

\section*{A R T I C L E   I N F O}

Article history:
Received 19 January 2020
Accepted 29 April 2020
Available online 16 May 2020

Keywords:
Distributed algorithm
Opinion dynamics
Multi-agent control
Consensus

\section*{A B S T R A C T}

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

\section*{1. Introduction}

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

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

\begin{equation}
    x_i(k+1) = \sum_{j=1}^{n} W_{ij} x_j(k), \quad i = 1, \ldots, n; \quad k = 0, 1, \ldots
\end{equation}
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 $w_{ii}$ are nonnegative and $\sum_{j=1}^{n} w_{ij} = 1$ for all $i$. As shown in Harary (1959) and Harary et al. (1965), the iteration of process (1) typically establishes eventual consensus (unanimity) of opinions.

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

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

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

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

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

While such models were initially studied in relation to non-stationary Markov processes (Dobrushin, 1956; Hajnal, 1958; Seneta, 1981; Wolowitz, 1963), they more recently experienced a resurgence of interest driven by agent-based models of synchronous collective motion based on the "nearest neighbor" rules (Lin, Morse, & Anderson, 2007; Olfati-Saber, 2006; Reynolds, 1987; Vicsek, Czirók, Ben-Jacob, Cohen, & Shochet, 1995). In such models, the interactions between the agents are range-restricted, and the set of an agent’s neighbors evolves as the agents move in the space. Consensus in the time-varying algorithm (3) is equivalent to the ergodicity3 of the backward infinite matrix products $W(k) \cdots W(0)$, see Seneta (1981). Despite the interest and efforts devoted to the problem of matrix products ergodicity in the literature on probability theory and matrix analysis (Leizarowitz, 1992; Seneta, 1981), a complete solution to this problem remains elusive, and a gap between necessary and sufficient ergodicity conditions still exists (Bolouki & Malhame, 2016; Touri & Langbort, 2014; Touri & Nedic, 2012). In recent years, the iterative averaging model has attracted significant attention from the research community as a simple algorithm for multi-agent coordination (Blondel, Hendrickx, Olshevsky, & Tsitsiklis, 2005; Cao, Morse, & Anderson, 2008; Jadbabaie, Lin, & Morse, 2003; Mesbah & Egerstedt, 2010; Moreau, 2005; Olfati-Saber, Fax, & Murray, 2007; Ren & Beard, 2008; Ren & Cao, 2011). Besides the important problem of multi-agent consensus, procedures of iterative averaging lie at the core of many distributed numerical algorithms (Bertsekas & Tsitsiklis, 1989) such as, e.g., techniques of distributed estimation and filtering (Borkar & Varaiya, 1982; Bullo, 2016; Calafiore & Abrate, 2009; Garin & Schenato, 2010), deterministic and stochastic optimization (Lin & Ren, 2014; Nedic, Ozdaglar, & Parrilo, 2010; Tsitsiklis, Bertsekas, & Athans, 1986; Yang et al., 2019), load balancing (Amelina, Fradkov, Jiang, & Vergados, 2015), and solving systems of linear equations (Liu, Morse, Nedić, & Başar, 2014; Mou, Liu, & Morse, 2015; Wang, Mou, Lian, & Ren, 2019; You, Song, & Tempo, 2016). Many models of opinion formation (Proskurnikov & Tempo, 2017; 2018) in social groups stem from the classical model (2).

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

1.1. Systems of recurrent averaging inequalities

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

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

where matrices $W(k)$ are row-stochastic. At a first glance, the system of inequalities (4) is too “loose” to entail any interesting properties of the sequence $\{x(k)\}$. However, under rather modest conditions of connectivity (for instance, if every matrix $W(k) + \cdots + W(k + T - 1)$, where $k \geq 0$ and $T$ is a fixed period, corresponds to a strongly connected graph) the inequality (4) implies asymptotic consensus of the opinions $x_i(k)$ (which, however, can be achieved at $\infty$). In many situations where consensus is not established, (4) provides convergence of the sequence $\{x(k)\}$. Similar properties have been recently obtained (Proskurnikov & Cao, 2017a) for inequalities

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

where $L(t)$ stands for a time-varying Laplacian matrix. The theory developed in Proskurnikov and Cao (2017a) is not directly applicable to discrete-time inequalities (4).

1.2. Contribution and paper organization

In this paper we develop a mathematical theory for the RAI (4). In particular, we establish convergence and consensus criteria for any feasible sequence $x(k)$. Further, we review a number of multi-agent algorithms and opinion formation models, and

---

3 The backward products of stochastic matrices $W(k) \cdots W(0)$ are ergodic if they converge, as $k \to \infty$, to a stochastic rank-one matrix.
show that they can be examined in a unified way using the RAI theory, which allows, in particular, to examine Hegselmann-Krause model with informed (“truth-seeking”) agents (Hegselmann & Krause, 2006), and to generalize the recent fundamental results (Fullmer & Morse, 2018) on distributed algorithms that compute a common fixed point for a family of paractions, thus solving a special system of nonlinear equations.

The paper is organized as follows: Section 2 introduces the notation and some concepts from graph theory. Section 3 recapitulates some classical results concerned with consensus in iterative averaging procedures. Section 4 presents new results, establishing convergence of the solutions of the RAI (4). In Section 5, we illustrate the results showing applications to models of opinion formation and distributed algorithms. Section 6 contains the 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 $e_n = (1, 1, \ldots, 1)^T \in \mathbb{R}^n$ to denote a (column) vector of all ones, and $e_i = (0, 0, \ldots, 1)^T \in \mathbb{R}^n$ to denote the basis of the unit coordinate vectors. For two vectors $x, y \in \mathbb{R}^n$ we write $x \leq y$ if $x_i \leq y_i \forall i$ and $x \leq 0$ if $x_i \leq 0 \forall i$; the reverse inequalities $\geq$ are defined analogously. We define the sign of a real number as follows

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

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

A graph is a pair $\mathcal{G} = (V, E)$, where $V$ is a finite set referred to as the set of nodes and $E \subseteq V \times V$ is a set of arcs. The arc $(i, j)$ is also denoted by $i \rightarrow j$. We call a graph undirected if $E$ is symmetric in the sense that $(i, j) \in E$ implies $(j, i) \in E$, otherwise the graph is directed. A sequence of arcs $i_0 \rightarrow i_1 \rightarrow i_2 \rightarrow \ldots \rightarrow i_n$ is called a walk connecting $i_0$ to $i_n$, the number of arcs $n$ is the walk’s length. A walk that starts and ends at the same node $i_0 = i_n$ is said to be a cycle. A graph is said to be periodic if an integer $p > 1$ exists that divides the length of any cycle; otherwise the graph is aperiodic. A graph is strongly connected if every two distinct nodes are connected by a walk. A graph is quasi-strongly connected, or has a directed (out-branching) spanning tree (Ren & Beard, 2008) if some node (“root”) is connected by walks to all other nodes. For undirected graphs, strong and quasi-strong connectivity are equivalent (such a graph is said to be connected).

A graph $(V', E')$, 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 strong components, among which at least one component is a source component (no arcs enter it) and at least one component is a sink component (no arcs leave it), see Fig. 2. A graph is quasi-strongly connected if and only if it has a single source component. A strong component can be isolated, when it has neither incoming nor outgoing arcs, and thus it is both a source and a sink. Strong components of undirected graphs are always isolated.

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 $^4(j, i)$.

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 nonnegative matrices $B_1, B_2 \in \mathbb{R}^{n \times n}$, one has $\mathcal{G}[B_1 + B_2] = \mathcal{G}[B_1] \cup \mathcal{G}[B_2]$; the resulting graph $\mathcal{G}[B_1 + B_2]$ is called the union of the graphs $\mathcal{G}[B_1]$ 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 $J = F \triangleq [1 : n] \setminus I$ 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 same for all nodes $\sum_{i \in I} b_{ij} = \sum_{j \in J} b_{ij}$, $\forall i$. For any cut $(I, J)$ in such a graph, the balance condition holds as follows

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

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

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

The following result characterizes a cut-balanced graph.

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

1. graph $\mathcal{G}[B]$ is cut-balanced;
2. flows from $I$ to $J$ and from $J$ to $I$ are either both positive or both zero:

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

3. all strongly connected components of $\mathcal{G}[B]$ are isolated;
4. for any $i, j \in [1 : n]$, a walk from $i$ to $j$ exists if and only if a walk from $j$ to $i$ exists.

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

$^4$In the models considered below, entry $a_{ij}$ usually quantifies the strength of influence agent $j$ exerts on agent $i$. Following the tradition of multi-agent systems theory (Ren & Beard, 2008; Ren & Cao, 2011) and the original work by French Jr. (1956), such an influence is depicted by an arc $i \rightarrow j$ rather than $i \rightarrow j$. 

Fig. 2. Strong components of a directed graph: (a) non-isolated; (b) isolated. In (a), $\mathcal{G}$ is a single source component, $\{11, \ldots, 15\}$ is a single sink component.
(6) holds with the constant
\[ C \triangleq \max_{(j)} \frac{\sum_{i \neq j} b_{ij}}{\sum_{i \neq j} b_{ij}} \]
where the maximum is taken over all possible cuts such that the ratio is well-defined. □

3. Convergence of classical averaging algorithms

We start by reviewing some basic results on stability of the iterative averaging algorithm (3), where \( \mathbf{W}(k) = \mathbf{W} \) for all \( k \), consensus conditions are well-known and dual to the conditions of regularity in a stationary discrete-time Markov chain, discussed, e.g., in Proskurnikov and Tempo (2017).

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

1. for any initial condition \( x(0) \), the opinions \( x(k) \) converge to some consensus value, that is
   \[ \forall x(0), \exists c = c(x(0)) : x(k) \xrightarrow[k \to \infty]{} c \mathbf{1}_n. \]
2. the matrix \( \mathbf{W} \) is \( \text{SIA}^5 \), that is, there exists a nonnegative vector \( \pi \in \mathbb{R}^n \) such that \( \mathbf{W}^k \pi = \pi \mathbf{1} \); \( \mathbf{1} \) is the vector whose entries are all equal to 1.
3. \( g(\mathbf{W}) \) is a quasi-strongly connected graph whose (unique) source component is aperiodic\(^6\).

If these conditions hold, then \( \pi \) is the Perron-Frobenius left eigenvector of \( \mathbf{W} \), such that \( \pi^T \mathbf{W} = 1 \mathbf{1} \mathbf{^T} \) and \( \sum \pi = 1 \). The consensus opinion of the group is \( c = \pi^T x(0) \).

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:
\[ w_{ij} = 0 \quad \forall i \in I, j \not\in I \Rightarrow x_i(k+1) = \sum_{j \in I} w_{ij} x_j(k) \quad \forall i \in I. \]
This explains the impossibility of consensus in presence of two such components. If a strong component \( I \) is a source and periodic, then the submatrix \( (W_{ij})_{i,j \in I} \) has an eigenvalue \( \lambda \in \mathbb{C} \) such that \( \lambda \neq 1 \) but \( |\lambda| = 1 \), so that for almost all initial conditions the opinions periodically oscillate (Proskurnikov & Tempo, 2017).

3.2. Time-varying averaging: necessary conditions

We now turn the attention to the more sophisticated case of non-stationary averaging procedure (3), with time-varying \( \mathbf{W}(k) \). Notice first that consensus in the sense of definition (8) can be established by “degenerate” procedures of iterative averaging. For instance, if \( \mathbf{W}(0) = \mathbf{1}_n \mathbf{\pi} \mathbf{^T} \) is a trivial rank-one stochastic matrix, then the opinion iteration terminates in one step: \( x(1) = x(2) = \ldots = (\pi^T x(0)) \mathbf{1}_n \).

\(^5\) SIA matrices are also called regular (Seneta, 1981) or fully regular (Gantmacher, 2000), since they correspond to regular Markov chains (Seneta, 1981) that “forget” their history and converge to a unique stationary distribution. Matrix \( \mathbf{W} \) is regular if and only if some column of \( \mathbf{W}^n \) is strictly positive for some \( k \) (Seneta, 1981).

\(^6\) 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 it has only one essential (recurrent) class, which is aperiodic (Gantmacher, 2000; Seneta, 1981).

\[ \mathbf{W}(k) = \begin{bmatrix} \mathbf{W}(k-1) & \mathbf{W}(k-2) & \cdots & \mathbf{W}(0) \end{bmatrix} \mathbf{\pi} \mathbf{^T} \]

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 \( \mathbf{W}(k) \), \( s < k \) play no role: consensus remains invariant even if the agents do not communicate after the first \( k \) steps. Although the problem of finite-time consensus is of self-standing interest (Hendrickx, Shi, & Johansson, 2015), the aforementioned situation is non-generic and it is usually ruled out by a stronger requirement (Moreau, 2005; Shi & Johansson, 2013b) of establishing consensus for each starting time:
\[ \forall x_0 \geq 0, \forall x(k_0), \exists c = c(x_0, x(k_0)) : \mathbf{W}(k) \cdots \mathbf{W}(k_0) x(k_0) \xrightarrow[k \to \infty]{} c \mathbf{1}_n. \]
e.g., if $w_{ij} > 0 \forall i$. For non-stationary matrices, even completeness of the graph $g_\pi$ does not always imply consensus if some persistent interactions are much “weaker” than others and the divergence rates of the series in (10) are different (Moreau, 2005).

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

3.3. Sufficient conditions: repeated connectivity

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

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

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

(11)

furthermore, $w_{ij}(k) > 0$ for any $i, k$. Assume also that a period $T > 0$ exists such that the graphs $g_k = \{W(k) + \ldots + W(k + T - 1)\}$ (arising as unions of $T$ consecutive graphs $\{W(k), \ldots, W(k + T - 1)\}$) are quasi-strongly connected for all $k \geq 0$. Then consensus (7) is established.

The assumptions of Lemma 3 are known as the repeated quasi-strong connectivity and, obviously, imply the quasi-strong connectivity of the persistent graph $g_\pi$.

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

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

$$C^{-1}w_{im}(k) \leq w_{ij}(k) \leq Cw_{im}(k) \forall k \geq 0.$$  

(12)

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

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

(13)

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

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

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

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

(14)

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

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

(15)

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

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

4. New results: convergence of RAl

In this section, we present novel results concerned with the behavior of any feasible solution to the RAI (4). The proofs of these results will be given in Section 6. The RAI constitutes a relaxation of the averaging Eq. (3), and clearly any trajectory of (3) is also a feasible trajectory for (4). Hence, any result that holds for all feasible trajectories of a RAI also holds for the trajectories of (3).

We next develop a theory for the RAI model, and then discuss in Section 5 several relevant applications. We are primarily interested in the properties of convergence and consensus, as formalized in the following

**Definition 2.**

1. The RAI (4) is convergent if all of its feasible solutions converge, that is, for any feasible solution $\{x(k)\}$ there exists a limit

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

(16)

2. The RAI (4) establishes consensus if it is convergent and the terminal opinions are coincident $\bar{x}_1 = \ldots = \bar{x}_n$ (equivalently, $\bar{x} = c\mathbf{1}_n$, where $c$ is a scalar).

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

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

(17)

A feasible solution to the RAI thus can be considered as a solution for the “forced” recursion

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

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

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

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

\footnote{Under the assumption (15), the graph $g_\pi$ 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 (4) and the classical iterative averaging procedure (3) makes inapplicable the plethora of methods based, explicitly or implicitly, on the “pseudo-contracting” properties of iterative averaging models (Fang & Antsaklis, 2008), and estimates of the diameter $d(k)$ (Blondel et al., 2005; Moreau, 2005; Shi & Johansson, 2013b; Xia et al., 2019) cannot be used to prove convergence to consensus. Another principal difference between RAI (4) and the averaging procedure (3) is the absence of duality between consensus and ergodicity of backward matrix products. Iterating (4), one easily derives the one-sided inequality $x(k+1) \leq W(k) \cdots W(0)x(0)$, however, the convergence of the right-hand side says nothing about the behavior of the solution. This is not surprising since, as it will be shown, the strong ergodicity of backward products $W(k) \cdots W(0)$ (equivalently, consensus in (3)) is not sufficient for the convergence of solutions in (4). For this reason, methods based on matrix analysis and Markov chain theory (Bolouki & Malhame, 2016; Cao et al., 2008; Ren & Beard, 2008; Seneta, 1981; Touri & Nedic, 2012) are also inapplicable to the analysis of the RAI.

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

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

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

4.1. Time-invariant RAI: convergence and consensus

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

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

1. $x_i = \bar{x}_j$ whenever $i$ and $j$ belong to the same strongly connected component of the graph $G[W]$ (“partial consensus” is established);
2. if the opinion $x_i(k)$ converges to a finite limit, the corresponding residual vanishes, i.e.,

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

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

$$x_i(k+1) \leq x_i(k), \quad x_j(k+1) \leq \frac{x_i(k) + x_j(k)}{2},$$

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

4.2. Time-varying case: convergence under reciprocal interactions

Convergence of the solutions in the time-varying case will be established under several assumptions. Our first assumption, which is typically adopted in the works on consensus (Blondel et al., 2005; Ren & Beard, 2008; Shi & Johansson, 2013b; Xia et al., 2019), can be considered as a counterpart of the aperiodicity assumption.

**Theorem 2** shows that, even for the static graph case, this assumption can be relaxed yet not discarded completely.

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

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

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

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

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

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

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

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

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

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

$$a_{ij}(k_0: k_1) \geq M \implies a_{ji}(k_0: (k_1 + T)) \geq 1. \quad (19)$$

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

We are now in position to formulate our second result.

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

1. the RAI (4) is convergent;
2. the terminal opinions of the agents $i, j \in [1:n]$ from the same strong component of $G_p$ coincide $\bar{x}_i = \bar{x}_j$ (i.e., partial consensus is established);
3. if the opinion $x_i(k)$ converges to a finite limit, then the corresponding residual vanishes $\Delta_i(k) \xrightarrow[k \to \infty]{} 0$.

The RAI (4) establishes consensus if and only if $G_p$ is strong.

**Theorem 3** can be further extended: both convergence and consensus appear to be robust against bounded communication delays. For a classical consensus algorithm, such a robustness is known under the assumption of repeated quasi-strong connectivity (Angeli & Bliman, 2006).
Theorem 4. Consider a sequence of matrices \((d_{ij}(k))_{ij=1}^n\), whose diagonal entries are zero\(^9\), \(d_{ii}(k) = 0\), the other entries being uniformly bounded \(0 \leq d_{ij}(k) \leq \delta < \infty\). Then the criteria from Theorem 3 retain their validity for the inequalities

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

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

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

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

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

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

\[
x_1(k + 1) = \frac{x_2(k - 1) + x_2(k)}{2},
\]

\[
x_2(k + 1) = x_1(k - 1), \quad x_2(k + 1) = x_2(k).
\]

The corresponding graph \(\mathcal{G}[W]\) is strongly connected and aperiodic since it contains two cycles \(1 \to 2 \to 1\) (length 2) and \(1 \to 2 \to 3 \to 1\) (length 3). At the same time, the system obviously has infinitely many periodic solutions: choosing \(x_1(k)\) to be a sequence of period 4, the vector \((x_1(k), x_1(k - 2), x_2(k), x_2(k - 2))\) is a solution to (21).

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

\[
x_1(k) = \begin{cases} 1, & k \text{ odd}, \\ 0, & k \text{ even}, \end{cases}
\]

\[
x_2(k) \equiv 1
\]

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

4.3. Reciprocity condition: discussion

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

4.3.1. Consensus under repeated strong connectivity

As it has been already discussed, Lemma 3 does not retain its validity for the RAI, even for constant \(\mathcal{W}(k) \equiv \mathcal{W}\). Notice, however, that if all graphs \(\mathcal{G}[W(k) + \ldots + W(k + T - 1)]\), \(k \geq 0\) (equivalently, unions of \(T\) consecutive graphs) are strongly connected, then the implication (19) holds with an arbitrary \(\mathcal{M} \geq 0\), 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 3 implies the following counterpart of Lemma 3, which has been reported in our previous work (Proskurnikov & Cao, 2017b).

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

4.3.2. Non-instantaneous type-symmetry

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

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

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

\[
\forall k_0 \geq 0 \quad w_{ij}(k_0) > 0 \Rightarrow \sum_{k=k_0}^{k_0+T} w_{ji}(k) > 0.
\]

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

4.3.3. Uniform cut-balance and arc-balance

Under Assumption 2, the condition of uniform cut-balance (14) implies the validity of Assumption 3 (with \(\mathcal{M} = 1\), \(T = 0\). The more general condition (15) also implies the validity of Assumption 3 with \(\mathcal{M} = 1\), \(T = L\). Indeed, suppose that \(a_{ij}(k_i)\), \(k_i > 1\) and let \(K \in [k_0, k_1]\) be the first instant when \(a_{ij}(k_i) > 0\). Then (15) implies that \(a_{ij}(k : (k + L)) > 0\), which proves (19) with \(T = L\).

In view of Remark 1, under Assumption 2 the condition of “arc-balance” (12) implies that, starting from some instant \(k\), one has either \(\mathcal{E}[W(k)] = \emptyset\) or \(\mathcal{E}[W(k)] = \mathcal{E}_p\). This obviously implies (10)-reciprocity of the sequence \(\{W(k)\}_{k=0}^{\infty}\). Similarly, the non-instantaneous arc-balance (13) implies that for \(k \geq k_0\) either \(\mathcal{E}[W(k) + \ldots + W(k + L)] = \emptyset\) or \(\mathcal{E}[W(k) + \ldots + W(k + L)] = \mathcal{E}_p\). Similar to the uniform cut-balance case, one shows that Assumption 3 holds with \(\mathcal{M} = 1\), \(T = L\).

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

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

4.3.4. Periodic gossiping with intermittent communication

We now propose a simple example where the implication (19) does not reduce to any of the aforementioned type-symmetry or balance conditions. A special class of iterative averaging polices (3) is constituted by so-called gossiping algorithms, where at each stage of the iteration at least one pair of agents communicates. In the case of deterministic unidirectional periodic gossip (Anderson, Yu, & Morse, 2010) the sequence of graphs \(\mathcal{G}[W(k)]\) is periodic and each graph \(\mathcal{G}[W(k)]\) contains a single arc \(\mathcal{E}[W(k)] = \{u_k, v_k\}\). 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 \(k_1 < k_2 < \ldots < k_0 < \ldots\) and a periodic sequence of arcs \((j_k, i_k)\) such

\(^9\) In other words, an agent has access to its own undelayed state.
that
\[ W(k) = \begin{cases} 
(1 - \alpha_k) e_i^j e_j^k + \alpha_k I_n, & k = k_i \\
I_n, & k \neq k_i \forall s.
\end{cases} \]  

(24)

For the iterative averaging algorithm (3), the periods of silence do not change the asymptotic behavior of the solution (except for its convergence rate) since the vector of opinions \( x(k) \) obviously, remains unchanged for \( k = k_i + 1, \ldots, k_i + 1 - 1 \). 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 straightforward.

All the aforementioned reciprocity conditions (type-symmetry, arc-balance and cut-balance) fail to hold if \( \sup_i (k_{s+1} - k_i) = \infty \). At the same time, the sequence of matrices (24) satisfies the condition (19), where \( T = 0 \) and \( \Omega \) is the period of the sequence \( (k_i, j) \). If \( a_{ij}(t_0 : t_1) \geq \Omega \) for some time window \( [t_0 : t_1] \), then during this time window each arc of the strongly connected graph \( G_\Omega \) has appeared at least once, and hence \( a_{ij}(t_0 : t_1) > 0 \). Theorem 3 now implies the following.

**Corollary 3.** Let \( W(k) \) have the representation (24), where \( (j, k) \) is a periodic sequence of arcs constituting a strongly connected graph \( G_\Omega \) and constants \( \alpha_k \) obey the inequality \( \eta \leq \alpha_k \leq 1 - \eta \) for some \( \eta > 0 \). Then the RAI (4) establishes consensus and the residual vanishes (18) for any bounded component.

4.4. Releases Assumption 2

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

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

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

5. Applications

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

5.1. Stability of some positive delay systems

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

**Lemma 7.** For a substochastic matrix \( A \), consider the set of “deficiency” indices \( \ell_k = \{1 : \sum a_{ij} < 1\} \). If every node \( i \) in the graph \( G(A) \) is reachable from the set \( \ell_k \) by a walk (formally, \( i \) is reachable from at least one node \( j \in \ell_k \) then \( \rho(A) < 1 \). In particular, if \( A \) is irreducible \( (G(A) \) is strongly connected) and \( \ell_k \neq \emptyset \), 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
\[ x_i(k + 1) = \sum_{j=1}^{n} a_{ij} x_j(k - d_{ij}), \quad i \in [1 : n]. \]  

(25)

**Proof.** We are going to show that every solution \( x(k), k = 0, 1, \ldots, \) of (25) obeys the RAI (20), with a special constant row-stochastic matrix \( W \) “dominating” the matrix \( A \) in the sense that \( w_{ij} \geq a_{ij} \) for all \( i, j \).

Consider first the undelayed case, where \( x(k) = A^k x(0) \) for every \( k \). Let \( V = \text{diag}(A_{1n}) \) be the diagonal matrix, whose entries \( v_{ii} = \sum_{j=1}^{n} a_{ij} \leq 1 \) stand for the sums of matrix \( A \)'s rows. Obviously, the matrix
\[ W = A + \frac{1}{n} (I - V) 1_n 1_n^T \]  

(26)

is stochastic since its entries are non-negative and \( W_{1n} = A_{1n} + (I - V) 1_n = 1_n \). Notice also that \( w_{ij} > a_{ij} \geq 0 \forall j \). Hence in the graph \( G(W) \) each node \( j \) is connected to every node from \( I_{j} \); in particular, all nodes from \( I_{j} \) have self-loops. By assumption, from \( I_{j} \) 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 \( x(k) = A^k x_0 \) are non-negative for any \( k \geq 0 \) and satisfy the inequality (4) with \( W(k) \equiv W \). Thanks to Theorem 2, \( x(k) \to c \), where \( c \geq 0 \) and
\[ \Delta(k) = W(x(k)) - x(k + 1) = n^{-1} (I - V) 1_n \sum_{k=0}^{\infty} x(k) \to 0. \]  

(27)

The latter condition implies that \( c(I - V) 1_n = 0, \) which is only possible for \( c = 0 \) (by assumption, \( I_{j}(A) \neq 0 \)). Hence \( A_{1n} = 0 \) for any vector \( x_0 \geq 0 \). Since every vector is a difference of two non-negative vectors, \( A \) is Schur stable.

The second statement is proved similarly. Let \( W \) be as in (26), and consider a solution to (25) with non-negative initial condition \( x_i(\tau) \geq 0 \) for \( -d_i \leq \tau \leq 0 \). This solution is non-negative and obeys the inequalities
\[ 0 \leq x_i(k + 1) \leq \sum_{j=1}^{n} w_{ij} x_j(k - d_{ij}) \quad \forall i \in [1 : n]. \]

**Theorem 4** ensures consensus \( x(k) \to c \) and
\[ \Delta_i(k) = \sum_{j=1}^{n} w_{ij} x_j(k - d_{ij}) - x_i(k + 1) \]
\[ = n^{-1} (1 - v_{ii}) \sum_{i=1}^{n} x(k) \to 0, \]  

(28)

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

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

5.2. The Hegselmann-Krause model with “truth seekers”

One of the seminal models describing opinion formation in social networks is known as the Hegselmann-Krause bounded confidence model. Its simplest version (Hegselmann & Krause, 2002; Krause, 2000) arises as a modification of the DeGroot algorithm (3), where the matrix \( W(k) \) co-evolves with the opinion vector and is defined as follows
\[ W(k) = \tilde{W}(x(k)), \quad \tilde{w}_{ij}(x) \triangleq \begin{cases} 
\frac{1}{|N_i(x)|}, & j \in N_i(x), \ j \neq N_j(x), \\
0, & \text{otherwise}
\end{cases} \]  

(29)

\[ N_i(x) \triangleq \{ j : |x_j - x_i| < \varepsilon \}. \]
In other words, at each stage of the opinion iteration, an individual replaces his/her opinion by the average of its own \(N(x)\) opinion and the opinions of like-minded individuals that belong to the confidence set, that is, the ball \(B_n(x_j) = \{x : |x - x_j| < \epsilon\}\). Agents ignore dissimilar opinions, i.e., those lying outside their confidence sets.

Notice that the matrix \(W(x)\) is stochastic and satisfies the condition (11) with \(\eta = n^{-1}\) since \(1 < r(x) \leq n\) \(\forall i \in \{1, n\}\). Also, \(W(x)\) is \(n^{-1}\). In particular, the sequence of matrices \(W(x) = W(x(k))\) satisfies the type-symmetry condition (22) with \(C = n\). The corresponding graph \(G_p\) is thus undirected and \((j, i) \in \mathcal{E}_p\) if and only if \(w_{ij}(k) > 0\) (equivalently, \(|\xi_i(k) - \xi_j(k)| < \epsilon\)) for infinitely many \(k > 0\). One thus arrives at a simple proposition.

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

In view of Proposition 2 and Theorem 3 every solution to the Hegselmann-Krause model (31),(29) converges. \(^{10}\) Since \((j, i) \notin \mathcal{E}_p\) and only if \(|\xi_i(k) - \xi_j(k)| \geq \epsilon\) for large \(k\), the steady opinions \(\bar{x}_i = \lim x_i(k)\) and \(\bar{x}_j = \lim x_j(k)\) are either coincident or sufficiently distant \(|\bar{x}_i - \bar{x}_j| \geq \epsilon\).

In this subsection, we consider a more general model with “truth seekers” (Hegselmann & Krause, 2006)

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

Here \(A = \text{diag}(a_1, \ldots, a_m)\) is a diagonal matrix, \(0 \leq a_i \leq 1\), \(\bar{W}(x)\) is defined in (29) and \(t \in \mathbb{R}\) is a constant variable, referred to as truth value. The presence of an additional term in the right-hand side of (30) is explained by availability of some information to the agents, making their opinions closer to the truth value. The coefficient \(a_i\) measures the level of agent \(i\)’s “awareness” of the truth value (Hegselmann & Krause, 2006). If \(a_i = 1\), the agent is able to find the truth at a single step without communicating to the other agents. If \(a_i = 0\), agent \(i\) updates his/her opinion in accordance with the usual Hegselmann-Krause model, and the dynamics of his/her opinions are not (directly) influenced by the truth. If \(0 < a_i < 1\), the opinion of individual \(i\) is driven by both the truth and the opinions of the other individuals.

The presence of the static term in the right-hand side of (30) visibly changes the dynamics and makes analysis of the model sophisticated. In the original work (Hegselmann & Krause, 2006) it was shown that if \(a_i > 0\) \(\forall i\), then all opinions converge to the truth \(x(k) \longrightarrow t \epsilon_n\). Later Kurz and Rambau (2011) showed that the opinions of the truth-seekers \((a_i > 0)\) always converge to \(t\) even if the group has ignorant agents \((a_i = 0)\). At the same time, the convergence of the ignorant agents’ opinions remained an open problem, which has been solved in Chazelle (2011) by using a method of power series (“s-energy”). The aforementioned results employ diverse and highly non-trivial mathematical techniques. Theorem 3 allows to examine convergence properties of (30) in a simpler way, as formalized in the following theorem.

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

1. the limit \(\bar{x}_i = \lim_{k \to \infty} x_i(k)\) exists for any \(i\);
2. \(\bar{x}_i = t\) if and only if agent \(i\) persistently interacts with some truth-seeker: there exists \(j\) such that \(a_j > 0\) and \((i, j) \in \mathcal{E}_p\) (this holds e.g. if agent \(i\) is a truth-seeker him/herself since \((i, i) \in \mathcal{E}(\bar{W}(x)) \forall k\));
3. if agent \(i\) does not satisfy the condition from statement 2, then \(x_i(k) = \bar{x}_i\) 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 \(|\bar{x}_i - \bar{x}_j| \geq \epsilon\) (clustering).

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

\[ |x_i(k + 1) - t| = |1 - a_i|^n \sum_{j=1}^n w_{ij}(k) (x_i(k) - t) | \leq (1 - a_i) | \sum_{j=1}^n w_{ij}(k) | x_i(k) - t | \forall i \in \mathcal{V}. \]

(31)

**Step 1.** We are going to prove the sufficiency part of statement 2. In view of Proposition 2 and Theorem 3, the sequence \(\xi(k)\) converges and remains bounded, since \(\xi(k) \geq 0\). Therefore, the residual term (17) vanishes, i.e.,

\[ \Delta_i(k) = \sum_{j=1}^n w_{ij}(k) \xi_i(k) - \xi_i(k + 1) \longrightarrow 0 \quad \text{as} \quad k \to \infty. \]

In view of (31), \(\Delta_i(k) \geq a_i \sum_{j=1}^n w_{ij}(k) \xi_j(k) \geq a_i w_{ii}(k) \xi_i(k)\). If agent \(i\) is a truth-seeker \((a_i > 0)\), then \(\xi_i(k) = |x_i(k) - t| \longrightarrow 0 \quad \text{as} \quad k \to \infty\). If \(i\) persists interactively to a truth-seeker \(j\), then \(\xi_j(k)\) and \(\xi_i(k)\) reach “consensus”

\[ \lim_{k \to \infty} \xi_i(k) = \lim_{k \to \infty} \xi_j(k) = 0. \]

Therefore, \(\lim x_i(t) \to t\) for truth-seekers and agents persistently interacting to them.

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

\[ x_i(k + 1) = \sum_{j=1}^n w_{ij}(k) x_j(k) \quad \forall k \geq k_0. \]

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

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

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

5.3. The discrete-time model of bimodal polarization

Altafini’s model of opinion formation, originally proposed in Altafini (2012) in its continuous-time form, portrays polarization, or “bipartite consensus”, of opinions over structurally balanced signed graphs (Shi, Altafini, & Baras, 2019). In this subsection, we consider a discrete-time modification of the Altafini model, examined in Meng, Shi, Johansson, Cao, and Hong (2016) and Liu, Chen, Basar, and Belabbas (2017). This model is similar to the consensus protocol (3), allowing, however, negative influence weights

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

(32)

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

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

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

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

(33)

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

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

- either the system (32) is globally asymptotically stable, so that $x(k) \longrightarrow \Omega \forall x(0)$, or
- the sequence of signed graphs $g[A(k)]$ is jointly structurally balanced (Liu et al., 2017), that is, a diagonal matrix $D$ with entries $d_{ii} \in [-1,1]$ such that $W(k) = DA(k)D$ for $k$ sufficiently large.\footnote{In the case of joint structural balance, one has $sgn a_{ii}(k) = sgn d_{ii}$, $sgn d_{ij}$. If $a_{ij} \neq 0$. In other words, the relation between agents $i$ and $j$ is positive if $d_{ii}$ have the same sign and negative if the signs are different. The sets of agents with $d_{ii} = 1$ and $d_{ij} = -1$ thus constitute two opposing factions in the signed graph $g[A(k)]$ (Altafini, 2013; Proskurnikov, Matveev, & Cao, 2016), 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” Altafini (2013) or bimodal polarization.}

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

$$\bar{x}_i(k + 1) = \frac{1}{n} \sum_{j=1}^{n} a_{ij}(k)|\xi_j(k)| \quad \forall i.$$ 

(34)

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

$$\Delta_i(k) = \sum_{j=1}^{n} a_{ij}(k)|\xi_j(k) - \xi_i(k + 1)| \longrightarrow 0.$$ 

(35)

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

$$\xi_i(k + 1) = x_i(k + 1) + \frac{1}{n} \sum_{j=1}^{n} a_{ij}(k)|\xi_j(k)| \geq \min_{j=1}^{n} (\xi_j(k)),$$

we obtain that $\Delta_i(k) \geq |c + d| - (c + d) \geq \min_{j=1}^{n} (\xi_j(k))$. Moreover, $\bar{x}_i = 0$ and for large $k$ one has $\bar{x}_i = 0$. In view of (35), in other words, $x_i(k) \longrightarrow x_i$ is constant and therefore $x_i(k) \longrightarrow x_i = \bar{x}_i$. If $G_P$ is strongly connected, then (33) holds due to Theorem 3.

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

$$\text{max} (a_{ij}(k) sgn x_i(k), 0) \longrightarrow 0 \quad \forall i, j,$$

(36)

which can be assumed that $x_i, x_j \neq 0$. Since $0 \leq \xi_i(k + 1) = x_i(k + 1) + \frac{1}{n} \sum_{j=1}^{n} a_{ij}(k)|\xi_j(k)| \text{sgn} x_i(k + 1) \geq \frac{1}{n} \sum_{j=1}^{n} a_{ij}(k)|\xi_j(k)| \text{sgn} x_i(k + 1) + \sum_{j=1}^{n} a_{ij}(k)|\xi_j(k)| 

we obtain that $\Delta_i(k) \geq 2\xi_i(k)$. By assumption, $\xi_i > 0$ and thus for large $k$ one has $\bar{x}_i = 0$. In view of (35), in other words, $x_i(k) \longrightarrow x_i$ is constant and therefore $x_i(k) \longrightarrow x_i = \bar{x}_i$. If $G_P$ is strongly connected, then (33) holds due to Theorem 3.

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

5.4. Constrained consensus and common fixed points of paracointractions

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

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.,

$$||x - P_\Omega(x)|| = \min_{x \in \Omega} ||x - y||.$$ 

It can be shown that $\langle x - P_\Omega(x), x - P_\Omega(x) \rangle \geq \pi/2$ (Fig. 3) and

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

(37)

The distance $d_\Omega(x) = \frac{1}{2} ||x - P_\Omega(x)||$ is a convex function.
Consider a group of \( n \) agents. Each agent keeps in its memory some constraints, described by a closed convex set \( \mathbb{S}_i \subseteq \mathbb{R}^d \) (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 \( \xi_i \in \mathbb{S}_i \) that satisfies all the constraints, but do not want to communicate the information about sets \( \mathbb{S}_i \). Assuming that an agent is able to compute the projection \( P_i(\xi) = P_{\mathbb{S}_i} \) of an arbitrary point \( \xi \in \mathbb{R}^d \) onto the set \( \mathbb{S}_i \), a point belonging to \( \mathbb{S}_i \) can be computed by one of the following modifications of the DeGroot iterative procedure (3):

\[
\xi_i(k + 1) = P_i \left[ \sum_{j=1}^{n} w_{ij}(k) \xi_j(k) \right].
\] (38)

\[
\xi_i(k + 1) = P_i \left[ \sum_{j=1}^{n} w_{ij}(k) P_j(\xi_j(k)) \right].
\] (39)

\[
\xi_i(k + 1) = w_i(k) P_i(\xi_i(k)) + \sum_{j \neq i} w_{ij}(k) \xi_j(k).
\] (40)

The protocol (38) has been proposed in the influential paper (Nedic et al., 2010), which deals with distributed optimization problems and then extended (removing some restrictive assumptions) by Lin and Ren (2014). The special cases of protocols (39) and (40) naturally arise in distributed algorithms, solving linear equations (Liu et al., 2014; Mou et al., 2015; You et al., 2016). A randomized version of (40) was examined by Shi and Johansson (2012). In all of the algorithms, \( \xi_i(k) \) stands for an approximation to the desired point \( \xi_i \), computed by agent \( i \) at step \( k \). In algorithms (38),(39) this approximation always satisfies the constraint of agent \( i \) (\( \xi_i(k) \in \mathbb{S}_i \) for \( k \geq 1 \)), whereas (40) provides this constraint only asymptotically.

We say that constrained consensus is established if the sequences \( \xi_i(k) \) converge and

\[
\lim_{k \to \infty} \xi_i(k) = \ldots = \lim_{k \to \infty} \xi_i(k) \in \mathbb{S}_i.
\] (41)

It has been recently realized in Fuller and Morse (2018) that the problem of constrained consensus can be considered as a special case of the more general problem of finding a common point for a finite family of paracontractions, also known as M-Fejer mappings (Wang, Ren, & Duan, 2019).

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

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

Simple examples of paracontractions are a continuous map without fixed points \( \mathcal{F}(M) = \emptyset \) and a contractive mapping \( \| M(\xi) - M(\xi_0) \| \leq q \| \xi - \xi_0 \| \quad \forall \xi, \xi_0 \in \mathbb{R}^d \) with \( q \in (0, 1) \). In the most interesting situations, however, the fixed point is non-unique. The inequality (37) implies that the orthogonal projection \( P_{\mathbb{S}_i} \) onto a closed convex set is a paracontraction in the standard Euclidean norm with \( \mathcal{F}(P_{\mathbb{S}_i}) = P_{\mathbb{S}_i} \). Other examples include, but are not limited to, proximal mappings and gradient descent mappings corresponding to special convex functions, see Fuller and Morse (2018). Obviously, (42) implies that \( M \) is non-expansive

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

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

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

**Lemma 8.** Lin and Ren (2014)12 Let the matrices \( W(k) \) satisfy the conditions of Lemma 3 and consider such sequences of vectors \( \xi^i(1), \ldots, \xi^n(1) \in \mathbb{R}^d, k \geq 0 \), that

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

Then the sequences \( \xi^i(k) \) asymptotically synchronize, that is,

\[
\lim_{k \to \infty} \| \xi^i(k) - \xi^j(k) \| = 0 \quad \forall i, j.
\]

Note that Lemma 8 does not guarantee the convergence of \( \xi^i(k) \); the latter property requires stronger assumptions on the “disturbances” \( e^i(k) \). To prove convergence of the algorithms, we also use the following proposition.

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

**Proof.** Assume, on the contrary, that \( \| M \xi(k) - \xi(k) \| \geq \epsilon, \forall \epsilon = 1, 2, \ldots, \) for a sequence \( \epsilon_0 \to 0 \). Passing to a subsequence, one can assume without loss of generality that the vectors \( \xi(k_\epsilon) \) converge to a limit \( \xi_\epsilon \in \mathbb{R}^d \). Recalling that \( M \) is continuous, one has \( \| M \xi_\epsilon - \xi_\epsilon \| \geq \epsilon \) while \( d(\xi_\epsilon) = 0 \). One arrives at a contradiction with (42), since \( \xi_\epsilon \notin \mathcal{F}(M) \) whereas \( \| M \xi_\epsilon - \xi^0 \| = \| \xi_\epsilon - \xi^0 \| \). □

We now formulate the main result of this subsection.

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

**Proof.** We first introduce some auxiliary notation. Fix an arbitrary point \( \xi^0 \in \mathbb{S}_* \). For this point, let \( \delta(\xi) = \| \xi - \xi^0 \| - \| P_i \xi - \xi^0 \| \geq 0 \). Let \( \xi^i(k) = \sum_{j=1}^{n} w_{ij}(k) \xi^j(k) \) be the central idea of the proof is to explore the properties of the vectors \( x(k) = (x_i(k))_{i=1}^{n} \) whose components \( x_i(k) = \| \xi^i(k) - \xi^0 \| \) 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 satisfy the RAI (4), which enables us to use the techniques of Theorem 3.

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

---

12 In this case, \( \mathbb{S}_i \) are linear hyperplanes.
has
\[ x_i(k+1) = \| P_i(\xi_i(k)) - \xi^0 \| \leq \| \xi_i(k) - \xi^0 \| \]
\[ = \sum_j w_{ij}(k) \| P_j(\xi_j(k)) - \xi^0 \| \]
\[ \leq \sum_j w_{ij}(k) \| \xi_j(k) - \xi^0 \| = \sum_j w_{ij}(k) x_j(k). \tag{45} \]

The case of (39) is considered similarly. Denoting \( \tilde{\xi}^i(k) \triangleq \sum_{j=1}^n w_{ij}(k) P_j(\xi_j(k)) \),
\[ x_i(k+1) = \| P_i(\tilde{\xi}^i(k)) - \xi^0 \| \leq \| \tilde{\xi}^i(k) - \xi^0 \| \]
\[ = \sum_j w_{ij}(k) \| P_j(\xi_j(k)) - \xi^0 \| \]
\[ \leq \sum_j w_{ij}(k) \| \xi_j(k) - \xi^0 \| \leq \sum_j w_{ij}(k) x_j(k). \tag{46} \]

In the case of algorithm (40), one has
\[ x_i(k+1) = \| w_i(k) (P_i(\xi_i(k)) - \xi^0) + \sum_{j \neq i} w_{ij}(k) (\xi_j(k) - \xi^0) \|
\[ \leq w_i(k) \| P_i(\xi_i(k)) - \xi^0 \| + \sum_{j \neq i} w_{ij}(k) \| \xi_j(k) - \xi^0 \| \leq \sum_{j \neq i} w_{ij}(k) x_j(k) = \sum_{j=1}^n w_{ij}(k) x_j(k). \tag{47} \]

Using Corollary 1, one shows that the RAI establishes consensus, that is, \( x_i(k) \rightarrow c \geq 0 \forall i \) and \( \Delta_i(k) \rightarrow 0 \).

Step 2. Using (18), we are now going to prove (44), which can be also written as
\[ \| \Delta_i(k+1) - \xi^0 \| \rightarrow 0 \quad \forall i. \tag{48} \]

Also, it will be shown that
\[ \| P_i(\xi_i(k)) - \xi^0 \| \rightarrow 0 \quad \forall i. \tag{49} \]

Notice first that the vectors \( \xi^i(k) \) and \( \tilde{\xi}^i(k) \) are bounded, since their distances \( x_i(k) \) to \( \xi^0 \) are bounded. For algorithm (38), the inequality (45) imply that
\[ \Delta_i(k) \geq \| \xi_i(k) - \xi^0 \| - \| P_i(\xi_i(k)) - \xi^0 \| = \delta_i(\xi^i(k)). \tag{50} \]

Applying Proposition 3 to (18), leads to \( \xi_i(k) \rightarrow 0 \) as \( k \rightarrow \infty \), which is equivalent to (48) \( \xi_i(k+1) = P_i(\xi_i(k)) \). To derive (49), note that
\[ \| P_i(\xi_i(k+1)) - \xi^i(k+1) \| \leq \| P_i(\xi_i(k+1)) - P_i(\xi^i(k)) \| \]
\[ + \| P_i(\xi^i(k)) - \xi^i(k) \| + \| \xi^i(k+1) - \xi^i(k) \| \rightarrow 0. \]

In the case of algorithm (40), we notice that (47) yields
\[ \Delta_i(k) \geq \| w_i(k) (\tilde{\xi}^i(k) - \xi^0) - \| P_i(\tilde{\xi}^i(k)) - \xi^0 \| \geq \eta \delta_i(\tilde{\xi}^i(k)). \tag{51} \]

Here \( \eta > 0 \) is the constant from Assumption 1. Applying Proposition 3, one proves (49), entailing also (48) since \( \xi_i(k+1) = \xi^i(k) \), which proves (48) since \( \tilde{\xi}^i(k) = \xi^i(k) \), which proves (48) since \( \tilde{\xi}^i(k) = \xi^i(k) \), which proves (48) since \( \tilde{\xi}^i(k) = \xi^i(k) \), which proves (48).

In the proof of the case (39) combines the aforementioned estimates. First, (46) yields in
\[ \Delta_i(k) \geq \| \tilde{\xi}^i(k) - \xi^0 \| - \| P_i(\tilde{\xi}^i(k)) - \xi^0 \| = \delta_i(\tilde{\xi}^i(k)). \]

which is similar to (50) and entails that
\[ \| \Delta_i(k+1) - \xi^0 \| \rightarrow 0 \quad \forall i. \]

Second, (46) also implies (51), which in turn implies (49). Thus \( \| \tilde{\xi}^i(k) - \xi^0 \| \rightarrow 0 \). This proves (48) since \( \| \xi_i(k+1) - \xi^0 \| \rightarrow 0 \). Since (48) is equivalent to (44), Lemma 8 ensures synchronization property \( \| \xi_i(k) - \xi^0 \| \rightarrow 0 \) for all \( i \).

Step 3. Recalling that the vectors \( \xi^i(k) \) are bounded, there exists a sequence \( k_i \rightarrow \infty \) such that \( \xi^i(k_i) \rightarrow \xi_i \in \mathbb{R}^d \). In view of (44), we have \( \xi^i(k_i) \rightarrow \xi_i \) for every \( i \). The property (49) implies that \( P_i(\xi_i) = \xi_i \), and thus \( \xi_i \in \Xi \). It remains to show that \( \xi^i(k) \rightarrow \xi_i \).

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 \( c \), depending on \( \xi^0 \) and the initial conditions. In particular, substituting \( \xi^0 = \xi_i \), we obtain that the limits exist
\[ x_i(k) = \lim_{k \rightarrow \infty} \| \xi^i(k) - \xi^0 \| = c_i \quad \forall i. \]

Recalling that \( x_i(k) \rightarrow 0 \), one has \( c_i = 0 \), which shows that the limits in (41) exist and are equal to \( \xi_i \in \Xi \).

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

6. Proofs of the technical results

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

6.1. Preliminary results

We start with the following simple proposition.

Proposition 4. For any sequence \( a_1, \ldots, a_m \in [0, 1 - \eta] \), where \( \eta \in (0, 1) \), the following inequality holds
\[ \pi(a_1, \ldots, a_m) \triangleq \prod_{i=1}^m (1 - a_i) \geq \exp \left( -\eta \sum_{i=1}^m a_i \right). \tag{53} \]

Proof. Since \( \pi \geq \pi(a_1, \ldots, a_m) > 0 \), one has
\[ -\log \pi = \sum_{i=1}^m \log \frac{1}{1 - a_i} = \sum_{i=1}^m \log \left( 1 + \frac{a_i}{1 - a_i} \right) \leq \sum_{i=1}^m \frac{a_i}{1 - a_i} \leq \eta \sum_{i=1}^m a_i. \]

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

We now derive two important estimates for the solutions of the RAI (4). For a non-empty set \( \mathcal{I} \subset \{1 \ldots n\} \), denote
\[ M_t(k) \triangleq \max_{i \in \mathcal{I}} x_i(k), \quad M(k) \triangleq \max_{i \in \mathcal{I}} x_i(k) = M_{1:|\mathcal{I}|}(k). \]

Along with the number of arcs \( a_{ij}(k_0, k_1) \) between two non-empty subsets of agents \( I, J \), consider the total flow from \( J \) to \( I \), which is defined as follows
\[ w_{ij}(k_0 : k_1) \triangleq \sum_{k=k_0}^{k_1} w_{ij}(k). \]
Let $w_{ij}(k) \overset{\Delta}{=} w_{ij}(k : l)$, $w_{ij}(k : 0) \overset{\Delta}{=} w_{ij}(k : 0)$. In some sense, the flow $w_{ij}(k : l)$ measures the influence of agents from $l = I$ on agents from $k$ during the time interval $[k_0, k_1]$. In particular, if $w_{ij}(k : 0) = 0$, then the group of agents $i$ evolves independently of $j$ on this interval, so that $M(k)$ is non-decreasing (Proposition 1). It can be expected that if $w_{ij}(k : 0) = 1$, the maximal increase $M(k_0 + 1) - M(k_0)$ can be positive yet admits a small upper estimate. This fact is established by the following simple lemma.

**Lemma 9.** Assume that Assumption 1 holds. For every instant $k_0 \geq 0$, $k'_0 \geq k_0$, $k_1 \geq k_0$, and every cut $(I, J)$ with $I, J \neq 0$ the solution $x(k)$ obeys the inequality

$$M(k_1 + 1) \leq \theta M(k_0) + (1 - \theta) M(k_0).$$

$$\theta \overset{\Delta}{=} \exp(-\eta w_{ij}(k : k_1)).$$

(54)

**Proof.** We know (Proposition 1) that $x_i(k) \leq M(k_0) \leq M(k_0) \forall k \geq k_0$. Denoting $\rho(k) \overset{\Delta}{=} \max_i w_{ij}(k)$, for each $k \geq k_0$, one obtains the inequality

$$\rho(k) \overset{\Delta}{=} \max_i w_{ij}(k).$$

By noticing that $w_{ij}(k) \leq 1 - w_{ij}(k) \leq 1 - \eta$ and also $w_{ij}(k) \leq w_{ij}(k)$ for each $i \in I$, Proposition 4 yields in

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

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

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

(55)

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

**Corollary 4.** Suppose that Assumptions 1, 2 and 3 hold and $G$ is strongly connected. Assume that for some $i \in \{1, \ldots, n\}$ and some $k_0 \geq 0$, a feasible solution of RAI (20) satisfies the condition $M(k_0) - x_i(k_0) \geq \epsilon > 0$. Then there exist $k > k_0$ and constant $C \in (0, 1)$, independent of $x, k_0$, that inequality holds as follows

$$\tilde{M}(k_0 + 1) \leq \tilde{M}(k_0) - \rho \epsilon.$$  

(62)

**Proof.** Step 1. First, we prove the following statement: Consider a cut $(I, J)$ with $I, J \neq 0$ and two instances $k_0 \geq k_0$ such that $M(k_0) - x_i(k_0) \geq \epsilon > 0$ for any $i \in I$. Then there exist $k_i \geq k_0$ and $j_0 \in J$ such that

$$x_i(k_i + 1) \leq \tilde{M}(k_0) - \rho \epsilon \quad \forall i \in I \cup \{j_0\}.$$  

(63)

where the constant $C$ is determined by $\eta$ from Assumptions 1, 2 and $\tilde{M}, \tilde{T}$ from Assumption 3.

Notice first that, due to (58),

$$\tilde{M}(k_0) - x_i(k'_0 + s) \geq \eta^i (\tilde{M}(k_0) - x_i(k'_0)) \geq \eta^i \epsilon \quad \forall i \in I \cup \{j_0\}.$$  

(64)

Thus, for $k'_0 \overset{\Delta}{=} k'_0 + d$, one has $\tilde{M}(k'_0) \geq \tilde{M}(k_0) - \rho \epsilon d$. Since $\tilde{M}$ is strongly connected, $\tilde{M}(k'_0) \overset{\Delta}{=} \infty$. Let $k_1 \geq k'_0$ be the first instant such that

$$a_{ij}(k'_1 : k_1) \leq 0.$$  

Then $k_1 \geq k'_0$ be the first instant such that

$$a_{ij}(k'_1 : k_1) \leq 0.$$  

(59)

(60)

Thus, for $k'_0 \overset{\Delta}{=} k'_0 + d$, one has $\tilde{M}(k'_0) \geq \tilde{M}(k_0) - \rho \epsilon d$. Since $\tilde{M}$ is strongly connected, $\tilde{M}(k'_0) \overset{\Delta}{=} \infty$. Let $k_1 \geq k'_0$ be the first instant such that

$$a_{ij}(k'_1 : k_1) \leq 0.$$  

Then $k_1 \geq k'_0$ be the first instant such that

$$a_{ij}(k'_1 : k_1) \leq 0.$$  

(59)

(60)
In view of the estimates on $\bar{M}_1(k'_0)$ and $w_{ij}(k'_0; k_1)$,
\[
\bar{M}_1(k) \leq \bar{M}(k_0) - \theta' \eta^{d-1} \varepsilon \quad \forall k \in [k'_0; (k_1 + 1)]
\]  
\[
\bar{x}_i(k+1) \leq \bar{x}_i(k) + (1 - \bar{w}_0(k_1)) \bar{M}(k_0) \leq \frac{\bar{M}(k_0) - \theta' \eta^d}{\varepsilon} \quad \forall i \in I_2.
\]  
Equation (65) is obvious if $k = k'_0$ (since $\theta' < 1$ and $\bar{M}(k'_0) \leq \bar{M}(k_0) - \eta^d \varepsilon$), otherwise it follows from Lemma 10 applied to the time interval $[k'_0; (k_1 - 1)]$.

By definition of $k_1$, there exists an arc $(i_0, j_0)$ in $G(W_{[k_1]})$ connecting $i_0 \in l$ to some $j_0 \in J$. Denoting $d_0 = d_{i_0j_0}(k_1)$ and $w_0 = w_{i_0j_0}(k_1) \geq \eta$, one arrives at
\[
x_{i_0}(k_1 + 1) \leq \bar{x}_i(k_0) + (1 - \bar{w}_0(k_1) + \bar{w}(k_0) - \bar{w}_0(k_1)) \bar{M}(k_0) \leq \frac{\bar{M}(k_0) - \theta' \eta^d}{\varepsilon} \tag{66}
\]
and so on. Iterating this argument $n - 1$ times, one finally arrives at the existence of $k_n \geq k_{n-1} + 1 > k_0$ such that
\[
x_{i_0}(k_n + 1) \leq \bar{M}(k_0) - \theta' \eta^d \varepsilon \quad \forall i \in I_n.
\]  
(67)

entailing that $\bar{M}(k_n + 1) \leq \bar{M}(k_0) - \theta' \eta^d \varepsilon$ to due to (58). Hence, (62) holds with $\bar{M} = \frac{\bar{M}(k_0) - \theta' \eta^d}{\varepsilon}$. □

6.2. Proof of Lemma 6

To prove the first statement, assume that the RA (4) is convergent yet $\varphi_p$ contains a strong component with outgoing arcs. Then a source component exists that has an outgoing arc but no incoming ones. 14

Denote the set of its nodes by $I_2 \subseteq [1: n]$ and let $j \notin F$. Since $(j, i) \notin \varphi_p$ for any $i \in l, j \in J$, one has $w_{ij}(0; \infty) < \infty$. Denote $a_k = \max_{i \in I_2} w_{ij}(k)$. In view of Assumption 1, $w_{ij}(k) \leq \bar{w}_{ii}(k) - a_k$. Defining $\zeta(0) = \frac{1}{1}$ and $\zeta(m) = \prod_{k=0}^{m-1} (1 - a_k) \quad 1 \leq m \leq \infty$

the sequence $\zeta(m)$ is decreasing and converges to a limit $\zeta(\infty)$; Proposition 4 entails that $\zeta(\infty) > 0$. Hence, $\zeta(k_0) < \infty$. Notice then that
\[
\zeta(k+1) = (1 - a_k) \zeta(k) - \sum_{j \in l} w_{ij}(k) \zeta(k) = \sum_{j \in l} w_{ij}(k) \zeta(k) \quad \forall i \in I.
\]  
(69)

By assumption, there exists a persistent arc $(i_0, j_0) \in \varphi_p$, where $i_0 \in l$ and $j_0 \notin J$ therefore, $w_{i_0j_0}(k) \geq \eta$ for an infinite sequence of instants $k = k_1, k_2, \ldots$. We will now construct a solution to (4) that does not converge. Let
\[
\chi(k) = \begin{cases} 
\zeta(k), & i \in I, \\
0, & i \in J \setminus \{j_0\}, \\
\frac{1}{k_0 + 1}^{(k - 1)} \zeta(\infty), & (i = j_0) \wedge (k = k_0 + 1) \\
\end{cases}
\]

Obviously, $x_{i_0}(k_1 + 1)$ does not converge as $s \rightarrow \infty$, so $x(k)$ fails to have a limit as $k \rightarrow \infty$. We are going to prove that $x(k)$ is a solution to the RA, that is,
\[
x_{i_0}(k + 1) \leq \sum_{j \in l} w_{ij}(k) x_j(k).
\]  
(70)

For $i \in l$, the latter inequality follows from (69) since
\[
x_{i_0}(k + 1) \leq \sum_{j \in l} w_{ij}(k) x_j(k) = \sum_{j \in l} w_{ij}(k) x_j(k) + \sum_{j \in l} w_{ij}(k) x_j(k) \geq_{\geq 0} 0.
\]  
(69)

Notice also that $0 \leq x_{j_0}(k + 1) \leq w_{i_0j_0}(k) x_{j_0}(k)$ for any $k$. For $k \neq k_0$, this is obvious since $x_{j_0}(k) \geq 0 = x_{j_0}(k + 1)$. For $k = k_0, \ldots$, one has $w_{i_0j_0}(k) \geq \eta$ and $x_{j_0}(k) \geq \zeta(\infty)$. Since all components $x_{j_0}(k)$ are non-negative, (70) holds also for $i = j_0$. For $i \in J \setminus \{j_0\}$, (70) is obvious since $x_{i_0}(k) \geq 0 \forall j$.

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

Remark 3. The construction of the oscillatory solution works also for delayed RA (20), adding the initial conditions: $x_i(-1) = \ldots = x_{i_0}(-d) = \zeta(0) = 1$ for $i \in l, x_{i_0}(-1) = \ldots = x_{i_0}(-d) = 0$ for $i \notin l$. In this case, $x_{i_0}(k - d) \geq \zeta(\infty) \geq \zeta(\infty)$ for any $k > 0, 0 \leq d \leq d_0, i \in l$.

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

6.3. Proof of Theorem 3

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

In view of Remark 1, starting from some instant $k > k_0$ only persistent arcs exist, and, renumbering the agents, the matrix $W(k)$ becomes block-diagonal $W(k) = \text{diag}(W_1(k), \ldots, W_n(k))$, where $W_{ij}(k) = k_{ij}(k) \leq M(k)$ are sequences of stochastic matrices corresponding to strongly connected graphs $\varphi_p$. Since the RA decomposes into several independent inequalities, it suffices to consider the case of a strongly connected persistent graph.

It remains to show that if $\varphi_p$ is strongly connected, then the RA establishes consensus and (18) holds whenever the solution is bounded. This statement is immediate from Corollary 4. Indeed, we know that the maximum $M(k)$ is non-increasing and hence converges to a limit $M_0 = \lim_{k \rightarrow \infty} M(k)$. If $M_0 = \infty$, then all opinions reach consensus at $\infty$. Otherwise, we have $x_{i_0}(k) \leq M(k)$ and thus $\lim_{k \rightarrow \infty} x_{i_0}(k) \leq M_0$. It suffices to show now that $\lim_{k \rightarrow \infty} x_{i_0}(k) \leq M_0$. Assume, on the contrary, that $\lim_{k \rightarrow \infty} x_{i_0}(k) > M_0$. In view of Corollary 4, there exists a sequence $k_m \rightarrow \infty$ such that $M(k_m) \leq M(k_0) - \alpha \varepsilon$. Passing to the limit as $m \rightarrow \infty$, one arrives at $M_0 \leq M(k_0) - \alpha \varepsilon$, which contradicts to the assumption $M_0 > \infty$. We have proved that $x_{i_0}(k) \rightarrow \infty$ for every $i$, that is, $x(k) \rightarrow M_0 \infty$. Also, if $M_0 > \infty$, the residual can be represented as
\[
0 \leq \Delta(k) = (W(k) - x(k) + 1) \rightarrow 0.
\]  
(70)

The “only if” part in the last statement is immediate from Lemma 6. □
6.4. Proof of Theorem 2

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

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

$$z(m + 1) \leq W^s z(m),$$

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

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

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

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

6.5. Proof of Theorem 4

The proof of Theorem 4, except for its final statement, retraces the proof of Theorem 3, where $M(k)$ is to be replaced by $M(k)$ and the proof of (18) for the case of a persistent graph $G^p$ and a bounded solution has to be modified in the following way: if $x_i(k) \to M_i$ as $k \to \infty$, then

$$\Delta_i(k) = \sum_{j} w_{ij}(k)(x_j(k) - d_j(k)) - M_i - (x_i(k+1) - M_i) \to 0.$$ 

Lemma 6 is applicable to the delayed case (Remark 3).

To prove the final statement, concerned with the time-invariant situation, 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 by Xiao and Wang (2006): consider the vector $y(k) \in \mathbb{R}^{n \times m}$ obtained by stacking the vectors $x(k), x(k-1), \ldots, x(k-d_n)$ one on top of another. Then it can be easily checked (Xiao & Wang, 2006) that

$$y(k+1) \leq \mathbb{E}[y(k)].$$

Here $W_i$ are non-negative matrices such that $W = W_0 + \cdots + W_d$; $(W_i)_{ij} = w_{ij}$ if and only if $d_{ij} = r$. In particular, $diag W_i = diag W$ and, if $diag W \neq 0$, the graph $g[\mathbb{E}]$ 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 $j \to (j+d_i) \to (j+2d_i) \to \cdots \to (j+id_i) \to i$ in the graph $g[\mathbb{E}]$. Therefore, the graph $g[\mathbb{E}]$ is strongly connected. Applying Theorem 2 to the matrix $\mathbb{E}$, one shows that RAI (20) establishes consensus.

6.6. Proof of Theorem 5

The proof is based on the idea of ordering that was employed in analysis of continuous-time consensus algorithms over symmetric and cut-balanced weighted graphs (Hendrickx & Tsitsiklis, 2013; Martin & Girard, 2013; Matveev et al., 2013) and later extended to the discrete-time case (Bolouki & Malhame, 2015). Note that the proofs from the aforementioned papers essentially use the boundedness of the solutions and their extension to a general feasible solution of RAI (20) is straightforward. For this reason, we give an independent proof.

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

$$y_1(k) \leq y_2(k) \leq \cdots \leq y_n(k) \leq y_s(k) \leq \cdots \leq y_{n-s+1}(k).$$

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

$$y(k+1) \leq V(k)y(k), \quad v_i(k) \leq w_{\sigma_i(k-1)\sigma_i(k)}(k).$$

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

**Step 1.** Using induction on $r = n, \ldots, 1$, we will prove the following statements: 1) the sequence $y_r(k)$ has a limit $\bar{y}_r(\infty) \leq -\infty$ as $k \to \infty$; 2) if $\bar{y}_r(\infty) > -\infty$, then

$$\sum_{k=0}^{\infty} v_s(k)(y_r(k) - y_{r+1}(k)) < \infty \quad \forall s \in [1 : n].$$

Then

$$\sum_{k=0}^{\infty} v_p(k)(y_r(k) - y_{r-1}(k)) < \infty \quad \forall s < r \forall p \geq r.$$ (74)

The condition (74) holds if $r > 1$.

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

$$y_n(k) - y_n(k+1) \geq \sum_{k=1}^{n-1} v_m(k)(y_n(k) - y_{n-1}(k)) \geq 0.$$ (75)

If $\bar{y}_n > -\infty$, the latter inequality, obviously, implies (73) with $r = n$. To prove (74), we apply the uniform cut-balance property.
to \( l = [1: (n - 1)], J = [n] \) and note that \( y_n(k) = y_n(k) - y_n-1(k) \), \( k < n \), therefore
\[
0 \leq (y_n(k) - y_n-1(k)) \sum_{s=m}^{t} v_{ms}(k) \leq C(y_n(k) - y_n-1(k)) \sum_{s=m}^{t} v_{ms}(k).
\]

To prove the induction step, assume that our statements have been already proved for \( r = m + 1, \ldots, n \). Our goal is to prove them for \( r = m \). If \( y_{m+1} = -\infty \), then, obviously, \( y_{m}(k) \rightarrow -\infty \). Otherwise, \((73),(74)\) hold for all \( r > m \). Using \((74)\) for \( s = m \) and any \( p > m \), \( r \in \{m + 1 : p\} \),
\[
v_{mp}(y_p - y_m) = v_{mp}(y_{m+1} - y_m) + v_{mp}(y_{m+2} - y_{m+1}) + \ldots + v_{mp}(y_p - y_{p-1}) \in \epsilon_1
\]
and thus, the existence of the limit (possibly, infinite)
\[
\lim_{k \rightarrow \infty} (y_m(0) - y_m(K + 1)) = \sum_{k=0}^{\infty} \frac{(y_m(k) - y_m(k + 1))}{1 - \epsilon_1} \leq +\infty.
\]
Furthermore, if this limit is finite, then also \( v_{mp}(y_m - y_J) \in \epsilon_1 \) for \( r < m \), which proves \((73)\) for \( r = m \). Applying \((73)\) to all \( s < m \) and \( r = m, \ldots, n \) and noticing that \( y_r(k) - y_r(k) = y_r(k) \) for \( r = m, \ldots, n \), one shows that
\[
(y_m(k) - y_m(k - 1)) \sum_{s=m}^{t} v_{ms}(y_m(k) - y_m(k - 1)) \in \epsilon_1
\]
and thus, applying \((14)\) for \( l = [1 : (n - 1)], J = [n] \), we also have \( v_{mp}(y_m - y_m) \in \epsilon_1 \) if \( s < m \). This proves \((74)\) for \( r = m \) and finishes the induction step.

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

**Suppose first** that \( y_1 > -\infty \), that is, the solution is bounded. Then, \((73)\) holds for every \( r \geq 1 \), entailing that \( \sum_{k=0}^{\infty} (y_r(k) - y_r(k)) \) and
\[
\sum_{k=0}^{\infty} \sum_{s=m}^{t} v_{ms}(k) |y_r(k+1) - y_r(k)| \leq \sum_{k=0}^{\infty} \sum_{s=m}^{t} v_{ms}(k) |y_r(k+1) - y_r(k)| < \infty
\]
(recall that \( v_{rs}(k) = w_{rs}(k) \) if and only if \( r = \sigma_{j}(k+1) \) and \( s = \sigma_{j}(k+1) \)).

\( x(j+1) \rightarrow -\infty \), Notice that first if a sequence \( k_m \rightarrow \infty \) such that the limit \( \xi = \lim_{k \rightarrow \infty} x(k) \) exists, then \( \xi \) is one of the values \( y_1, \ldots, y_n \). Indeed, for every \( m \) we have \( x(k_m) = y_{k_m}(k_m) \) with some \( j_m \). Passing to the subsequence, we may assume, without loss of generality, that \( J_m \equiv j_m \), and thus \( \xi = y_j \). For the same reason, for every \( y_j \), there exists at least one sequence \( x(k_m) \rightarrow y_j \); indeed, \( y_j(k) = x_{j,k}(k) \) for some reason, and we can find a subsequence \( k_m \) such that \( \sigma_{j}(k_m) = i \) for all \( m \).

Let \( J \) stand for the set of indices \( i \) such that
\[
\lim_{k \rightarrow \infty} x_j(k) = -\infty,
\]
and \( J = J^f \). It can be easily seen that \( x_j(k) \rightarrow -\infty \) for every \( j \in J \) and \( \lim_{k \rightarrow \infty} x_j(k) \geq y_{m+1} \) for all \( i \in J \). Indeed, choose two constants \( b < y_{m+1} \) and \( a < b \) in such a way that \( \eta a + (1 - \eta) \) \( y_n < b \). For large \( k \), we have \( y_n(k) < a < b < y_{m+1}(k) \), therefore, no component \( x(k) \) can belong to \( [a, b] \). Hence, if \( x(k) < a \) for \( k \) sufficiently large, \( RAI \) and Assumption 1 imply that
\[
x_j(k+1) < \eta a + (1 - \eta) y_n(k) < b \Rightarrow x_j(k+1) < a.
\]
that is, the opinion cannot leave the interval \((a, \infty)\). Therefore, for every \( j \in J \) one has \( \lim_{k \rightarrow \infty} x_j(k) < a \), which is possible only if \( x_j(k) \rightarrow -\infty \). For \( i \in J \), we have \( \lim_{k \rightarrow \infty} x_j(k) \geq J_m \), which is possible only if \( \lim_{k \rightarrow \infty} x_j(k) \geq y_{m+1} \). Obviously, \( J = [\sigma_{j}(k), \ldots, \sigma_{j}(k)] \) for \( k \) being sufficiently large, whereas \( I = [\sigma_{j}(k+1), \ldots, \sigma_{j}(k)] \). Now, we can repeat the argument we used in the bounded solution case with a minor modification. Applying \((73)\) for every \( r > m \), one has
\[
\sum_{i \in J} \sum_{j=1}^{n} w_{ij}(k) |x_{j}(k+1) - x_{j}(k)|
\]
for \( k \) being sufficiently large, and therefore the latter sequence belongs to \( \epsilon_1 \). Since \( w_{ij}(k) \) \( \geq \), we have \( \sum_{i \in J} \frac{x_{j}(k+1) - x_{j}(k)}{\epsilon_2} < \infty \) for every \( i \in J \) (therefore, \( x(k) \) converges to a finite limit \( \bar{x} \) for all \( i \in J \) and \( \sum_{i \in J} \frac{x_{j}(k+1) - x_{j}(k)}{\epsilon_2} < \infty \) for every \( i \in J \), \( i \in J \), \( i \in J \), \( i \in J \)). Using that \( \epsilon_2 \) is a constant, \( \epsilon_2 \) is a constant, \( \epsilon_2 \) is a constant, \( \epsilon_2 \) is a constant, one proves that \( x(k+1) = \bar{x} \). Consequently, \( x(k+1) = \bar{x} \). Obviously, that is, there are no arcs between \( I \) and \( J \) (recall that \( \bar{x} \) has already disconnected strongly connected components in view of the cut-balanced property). We know that \( \bar{x} = \bar{x} \), \( \bar{x} = \bar{x} \), \( \bar{x} = \bar{x} \), \( \bar{x} = \bar{x} \), \( \bar{x} = \bar{x} \), so that consensus in each strongly connected component of \( \bar{x} \) is established.

**Remark 5.** Using techniques from Xia et al. (2019), the consensus criterion from Theorem 5 can be easily extended to a more general situation of non-instantaneous cut-balanced (15). Namely, if \( (15) \) holds, Assumption 1 is valid and \( \bar{x} \) is strongly connected, then RAI (4) establishes consensus. Indeed, matrices \( B(k) = W(k) + L^{-1} \ldots W(k+1)W(k) \) satisfy the uniform cut-balanced condition \((14)\) and Assumption 1 (Xia et al., 2019). Since \( x(k+1) \leq B(k+1) \), \( x(k+1) \leq B(k+1) \), \( x(k+1) \leq B(k+1) \), \( x(k+1) \leq B(k+1) \), \( x(k+1) \leq B(k+1) \), the sequence \( x(k+1) \) converges to consensus vector \( \bar{x} \) as \( k \rightarrow \infty \). Similarly to the proof of Theorem 1, one proves that \( c_0 = c_1 = \ldots = c_{L-1} \).

**7. Conclusions and perspectives**

The recurrent averaging inequalities (RAIs) introduced in (4) can be considered as a relaxed form of the conventional dynamics of iterative averaging (2). While in the standard iterative averaging each agent updates its opinion with a weighted average of its own and the others’ opinions, in the RAI each agent is allowed to
choose any opinion which does not exceed that linear combination. Such a constraint seems only loosely restrictive, however, under some connectivity assumptions, all feasible solutions converge and even reach consensus. Similarly to the classical iterative averaging, the consensus is robust to bounded communication delays. Of particular interest is the bounded solutions to the RAI for which the residual vector (the difference between the right-hand and the left-hand sides) vanishes asymptotically.

The systematic study of RAI appears to be relevant since they naturally arise, implicitly or explicitly, in a number of distributed algorithms and multi-agent models based on iterative averaging. Using the general results on consensus in RAI, we derive a number of known and new results in a unified way, namely, a stability criterion for sub stochastic matrices and its novel extension to time-delay systems, the convergence of opinions in the Hegselmann-Krause model with “truth-seekers” and the discrete-time model of opinion polarization over a signed network, and the convergence of algorithms for finding a common fixed point for a family of paracontractions. The latter algorithms are quite relevant and can be used, in particular, for solving in a distributed way large-scale systems of equations and strongly convex optimization problems (Fuller & Morse, 2018; Wang et al., 2019).

The results in this paper can potentially be extended to RAI with random matrices W(k), and to RAI obtained from nonlinear procedures of iterative averaging (Angeli & Bliman, 2006; Fang & Antsaklis, 2008; Moreau, 2005; Proskurnikov, 2014; Proskurnikov & Matveev, 2015). However, these results require additional mathematical tools and are thus beyond the scope of the present work.

Although the theory of RAI allows a unified analysis of many multi-agent algorithms and models that are based on iterative averaging, some consensus-based algorithms still require an ad hoc theory. One class of such algorithms is constituted by distributed optimization procedures employed, in particular, in machine learning, multi-sensor tracking, control of cyber-physical systems and smart power grids, see, e.g., Nedic et al. (2010), Bajovic, Jakovetic, Krejc, and Jerinck (2017), Tataranko and Touri (2017), Mokhtari and Ribeiro (2016), Molzahn et al. (2017), Erofeeva and Granichin (2018), Notarstefano, Notarnicola, and Camisa (2019), Scaman, Bach, Bubeck, Lee, and Massoulié (2019) and Yang et al. (2019) and the references therein. Another example is given by proximal dynamics in multi-agent games (Cenedese, Kawano, Grammatico, & Cao, 2018; Grammatico, 2018), which are similar in spirit to the algorithms in Section 5.4, but the corresponding proximal maps do not have a common fixed point. We believe, however, that extensions of RAI theory towards such algorithms and models are possible and should be the subject of future research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References


