Chapter 3
Averaging in Time-Varying Networks

Abstract This chapter studies averaging dynamics in which the update matrix and, possibly, the underlying graph may be different at each time step. This extension is particularly important for the applications. Indeed, in realistic models of sensor and robotic networks, units, and links may be occasionally off due to environmental reasons or for energy saving purposes. Similarly, social dynamics may involve complex patterns of interactions that change over time. We are going to show that time-dependent consensus algorithms converge under relatively mild assumptions involving suitable notions of connectivity. Actually, the underlying graph needs not to be connected at any time, but the sequence of graphs must be "sufficiently connected" over time. More specifically, in Sects. 3.1 and 3.2 we provide two families of results, corresponding to two sufficient connectivity assumptions. Our presentation also includes, in Sect. 3.3, cases when the matrix evolves randomly in time. These randomized dynamics encompass the so-called *gossip* algorithms, which have attracted much attention in the last decade.

3.1 Time-Varying Updates: Uniform Connectivity

Given a set of nodes V of cardinality N, we consider a distributed state $x(t) \in \mathbb{R}^V$ evolving according to a system of the form

$$x(t+1) = P(t)x(t) \qquad t \in \mathbb{Z}_{\geq 0}, \qquad (3.1)$$

where $P(t)$ is a stochastic matrix for each $t \geq 0$. We will use the following notation

$$P(s,s) = I, \quad P(t,s) = P(t-1)\dots P(s), \quad 0 \leq s < t$$

so that $x(t) = P(t,s)x(s)$ for every $s \leq t$.

We start with a preliminary result which is a simple consequence of the contraction principle Lemma 2.1 already used in the time-invariant context.

Lemma 3.1 *Consider system* (3.1). *Assume that*

(i) *there exists $\alpha \in (0, 1]$ such that, for every $t \geq 0$ and $u, v \in V$, $P_{uv}(t) > 0$ implies $P_{uv}(t) \geq \alpha$;*
(ii) *there exists a sequence of times $\{t_k \in \mathbb{Z}_{\geq 0} : k \in \mathbb{Z}_{\geq 0}\}$ such that*

    (a) *there exists $B \in \mathbb{N}$ such that $t_{k+1} - t_k \leq B$ for all k and*
    (b) *for every k, there exists $v^* \in V$ such that, $P(t_{k+1}, t_k)_{uv^*} > 0$ for every $u \in V$.*

Then, $x(t)$ converges to a point in span$\{\mathbf{1}\}$ *from every initial condition in $\mathbb{R}^V$.*

Proof Notice first of all that thanks to assumptions (i) and (ii), we have that

$$P(t_{k+1}, t_k)_{uv^*} > \alpha^{t_{k+1} - t_k} \geq \alpha^B, \quad \forall u \in V. \tag{3.2}$$

Define now, for every $t \geq 0$,

$$x_{\min}(t) = \min_u \{x_u(t)\} \qquad x_{\max}(t) = \max_u \{x_u(t)\}$$

and notice that Lemma 2.1 together with (3.2) implies that, for every $k \geq 0$,

$$x_{\max}(t_{k+1}) - x_{\min}(t_{k+1}) \leq (1 - \alpha^B)\big(x_{\max}(t_k) - x_{\min}(t_k)\big) \tag{3.3}$$

Considering that $x_{\max}(t)$ and $x_{\min}(t)$ are two monotonic sequences, thus admitting limit, it follows that $x_{\max}(t) - x_{\min}(t) \to 0$ as $t \to +\infty$. This yields the thesis. $\quad\square$

    The above result is not very appealing for application as condition (ii). needs in principle to consider large products of the matrices $P(t)$ determining the dynamics. As in the time-invariant case, we would like to have results whose assumptions are at the level of the associated graphs $\mathscr{G}_{P(t)}$. The reader may recall that, on static networks, consensus was proved (see Theorem 2.2) under two conditions: a connectivity condition and an aperiodicity condition. On time-varying networks, we are going to make suitable assumptions on the connectivity over time, while the aperiodicity condition is replaced by the following assumption.

Definition 3.1 (*Nondegeneracy*) A set $\mathscr{P}$ of stochastic matrices over V is *nondegenerate* if

(i) for every $P \in \mathscr{P}$ and for every $u \in V$, $P_{uu}(t) > 0$;
(ii) there exists $\alpha \in (0, 1]$ such that, for every $P \in \mathscr{P}$ and $u, v \in V$, $P_{uv}(t) > 0$ implies $P_{uv}(t) \geq \alpha$.

It is clear that the assumption of nondegeneracy relates to aperiodicity: Indeed, if $\mathscr{P}$ is nondegenerate, then each $P \in \mathscr{P}$ is aperiodic. Notice that the converse is not true, because of the strong positivity condition expressed in the definition. Moreover, the mere aperiodicity of each matrix $P(t)$ is not sufficient for consensus; see Exercise 3.1.

    The following result shows some fundamental consequences of nondegeneracy. If P is a stochastic matrix over V, below we will use the notation $\mathscr{G}_P = (V, E_P)$ for the graph associated with P.

Proposition 3.1 *Suppose that $P(t)$, for $t \in \mathbb{Z}_{\geq 0}$ is a non degenerate sequence of stochastic matrices. Fix $t_1 \leq t_2 \leq t_3 \leq t_4$. Then,*

(i) $E_{P(t_3,t_2)} \subseteq E_{P(t_4,t_1)}$;
(ii) If $(u, v) \in E_{P(t_4,t_3)}$ and $(v, w) \in E_{P(t_2,t_1)}$, then $(u, w) \in E_{P(t_4,t_1)}$.

Proof Both claims follow immediately by combining the two inequalities

$$P(t_4, t_1)_{uw} \geq P(t_4, t_3)_{uu} P(t_3, t_2)_{uw} P(t_2, t_1)_{ww}$$
$$P(t_4, t_1)_{uw} \geq P(t_4, t_3)_{uv} P(t_3, t_2)_{vv} P(t_2, t_1)_{vw}$$

and property (i) of nondegeneracy. $\square$

We are now ready to state the main convergence result of this section which is a generalization of Theorem 2.2.

Theorem 3.1 (Time-dependent consensus I) *Consider system* (3.1). *Assume that*

(i) the set of matrices $\{P(t)\}$ is nondegenerate;
(ii) there exists a duration $T \in \mathbb{N}$ such that, for all $t_0 \in \mathbb{Z}_{\geq 0}$, the graph

$$\bigcup_{s=0}^{T-1} \mathcal{G}_{P(t_0+s)}$$

contains a globally reachable node.

Then, $x(t)$ converges to a point in span$\{\mathbf{1}\}$ *from every initial condition in $\mathbb{R}^V$.*

The reader should note that the connectivity condition does not imply anything on each single graph $\mathcal{G}_{P(t)}$, which may well be not connected at any time t. The following example illustrates the application of the theorem.

Example 3.1 (*Sequences of graphs*) Consider the following sequences composed of the graphs represented in Fig. 3.1.

(i) $\mathcal{S}_1(t) = \begin{cases} G_a & \text{if } t \text{ is a square number} \\ G_b & \text{otherwise} \end{cases}$

(ii) $\mathcal{S}_2(t) = \begin{cases} G_b & \text{if } t \text{ is a square number} \\ G_a & \text{otherwise} \end{cases}$

(iii) $\mathcal{S}_3(t) = \begin{cases} G_c & \text{if } t \text{ is a square number} \\ G_d & \text{otherwise} \end{cases}$

(iv) $\mathcal{S}_4(t) = \begin{cases} G_d & \text{if } t \text{ is a square number} \\ G_a & \text{otherwise} \end{cases}$
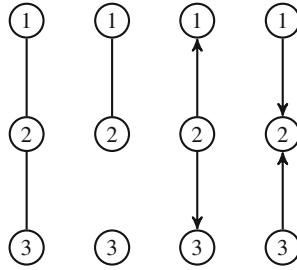
Fig. 3.1 The graphs G_a, G_b, G_c, and G_d used in Example 3.1

Let now $P_i(t)$ denote the sequence of SRW matrices constructed on the sequence of graphs $\mathscr{S}_i(t)$. By Theorem 3.1, we conclude that $P_2(t)$ and $P_3(t)$ lead to consensus. Instead, nothing can be concluded regarding $P_1(t)$ or $P_4(t)$ because assumption (ii) in Theorem 3.1 is not satisfied. Both cases actually lead to a consensus. For $P_4(t)$, consensus is trivial by observing that the SRW related to G_d leads to consensus in one step, while for $P_1(t)$ consensus will follow by Corollary 3.2 later on.

The proof of Theorem 3.1 relies on the results proven so far as well on a classical combinatorial argument reported below for the convenience of the reader.

Lemma 3.2 (Pigeonhole principle) *If n discrete objects are to be allocated to m containers, then at least one container must hold no fewer than $\lceil \frac{n}{m} \rceil$ objects.*

Proof (of Theorem 3.1) Notice first of all that since

$$\bigcup_{s=0}^{T-1} \mathscr{G}_{P(t_0+s)} \subseteq \mathscr{G}_{P(t_0+T,t_0)}$$

by virtue of Proposition 3.1, it follows that the graph $\mathscr{G}_{P(t_0+T,t_0)}$ contains a globally reachable node for every t_0. Let $N = |V|$ and let M be the number of distinct graphs over V possessing a globally reachable node. Consider the time interval $[0, NMT[$ split into subintervals $[0, T[$, $[T, 2T[$ and so on. By the pigeonhole principle, there must exist a graph G over V possessing a globally reachable node, which is repeated at least N times among the sequence of graphs $\mathscr{G}_{P(jT,(j-1)T)}$ for $j = 0, \ldots, NM$. Denote by v^* the globally reachable node inside G. Since every node $u \in V$ is connected to v^* in G with a path of length $l \leq N$, it follows by a repeated application of Proposition 3.1 that $\mathscr{G}_{P(NMT,0)}$ contains every edge of type (u, v^*) for $u \in V$. This implies that $P(NMT, 0)_{uv^*} > 0$ for every $u \in V$. Arguing similarly on every matrix $P(kNMT, (k-1)NMT)$, we can see that the assumptions of Lemma 3.1 are satisfied if we take $t_k = kNMT$.   □

Remark 3.1 (*Convergence time*) Note that the construction in the proof of Theorem 3.1 is "worst-case" in nature and gives little clue about the actual convergence time for the algorithm. This issue is investigated in an exemplary case in Exercise 3.2.

Remark 3.2 (*Consensus point*) In general, the final value upon which all the states x_u agree in the limit is unknown. This final value depends on the initial condition and the specific sequence of matrices defining the time-dependent linear algorithm. Only in few cases, one can compute the final value. One instance are time-independent consensus algorithms, which have been considered in the previous chapter. Another instance are time-dependent algorithms (3.1) involving doubly stochastic matrices. Indeed, it is clear that whenever $P(t)$ is doubly stochastic for every $t \geq 0$, then $x_{\text{ave}}(t) = x_{\text{ave}}(0)$. Then, provided $x(t)$ converges, it converges to $x_{\text{ave}}(0)\mathbf{1}$. This simple remark can be immediately extended to any sequence of matrices which share their dominant left eigenvector.

Theorem 3.1 requires a *uniform* connectivity assumption: The union of graphs, over time, must be connected within a fixed window. Later on, we will present results where the connectivity assumption is weaker. The following result shows that some connectivity condition for consensus will, however, be necessary.

Proposition 3.2 (Connectivity is necessary) *Consider system* (3.1). *If, for every initial condition $x(0)$, the state $x(t)$ converges to a point in* $\text{span}\{\mathbf{1}\}$, *then there exists a node which is globally reachable in the graph*

$$G = \bigcup_{s \geq 0} \mathscr{G}_{P(s)}.$$

Proof By contradiction, assume that G does not possess a globally reachable node. This implies that the correspondent condensation graph has two leaves that correspond to two strongly connected subgraphs of G, denoted by $G_i = (V_i, E_i)$ for $i = 1, 2$ and such that $V_1 \cap V_2 = \emptyset$ and there is no path from V_1 to V_2 or from V_2 to V_1. Consider now an initial condition $x(0)$ such that $x(0)_v = 0$ for all $v \in V_1$ and $x(0)_v = 1$ for all $v \notin V_1$. Since there is no edge outgoing G_1 and G_2, clearly, if Q is a stochastic matrix adapted to G, it follows that $(Qx(0))_v = 0$ for all $v \in V_1$ and $Qx(0)_v = 1$ for all $v \in V_2$. This implies that $x(t)_v = 0$ for all $v \in V_1$ and $x(t)_v = 1$ for all $v \in V_2$, and thus, $x(t)$ cannot converge to a consensus. $\square$

As the necessary condition in Proposition 3.2 is weaker than the sufficient condition in Theorem 3.1, it is natural to ask whether the former is sufficient as well. The answer is negative, as shown by the example proposed in Exercise 3.3.

3.2 Time-Varying Updates: Cut-Balanced Interactions

Theorem 3.1 requires a *uniform* connectivity assumption: The union of graphs, over time, must be connected within a fixed window. In this section, we seek conditions under which this uniform connectivity requirement can be dropped, while maintaining convergence.

To this goal, we need to introduce two new concepts, *cut-balanced* graph and *limit graph*. A graph $G = (V, E)$ is said to be *cut-balanced* when for any nonempty proper subset $S \subset V$, there exist $v \in S$ and $w \notin S$ with $(v, w) \in E$ if and only if there exist $v' \notin S$ and $w' \in S$ with $(v', w') \in E$. Clearly, if a graph is symmetric or strongly connected, then it is also cut-balanced. More precisely, cut-balanced graphs allow for the following characterization.

Lemma 3.3 *The graph G is cut-balanced if and only if every weakly connected component of G is strongly connected.*

Proof Assume that the graph is cut-balanced and let $W \subseteq V$ be a weakly connected component of G. If W is not strongly connected, then there exist a node $u \in W$ from which not all nodes in W can be reached. Let S be the subset of nodes which are reachable from u. Clearly, S is a proper subset of W and, necessarily, there is no edge from S to $W \setminus S$, while there must be edges from $W \setminus S$ to S; otherwise, W would not be a weakly connected component. This clearly contradicts the fact that G was cut-balanced. The proof of the reverse implication is left to the reader. $\square$

The second key ingredient is the definition of *limit graph*. Given a sequence of graphs $(G_t)_{t \in \mathbb{N}}$ such that $G_t = (V, E_t)$ for all $t \in \mathbb{N}$, we say that the limit graph of this sequence is the graph $G_\infty = (V, E_\infty)$ where the edge set E_∞ equals to the set-theoretic limit superior of the sequence $(E_t)_t$, that is,

$$E_\infty = \limsup_{t \in \mathbb{N}} E_t := \bigcap_{t \geq 0} \bigcup_{s \geq 0} E_{t+s}.$$

Equivalently, an edge (u, v) is in E_∞ when $(u, v) \in E_n$ for infinitely many n. This limit graph "forgets" transient interactions and focuses on those interactions that occur infinitely often and thus affect the convergence behavior.

We are now ready to state and prove the main result of this section.

Theorem 3.2 (Convergence) *Consider system* (3.1). *Assume that*

 (i) *the set of matrices $\{P(t)\}$ is nondegenerate;*
 (ii) *the associated graph $\mathscr{G}_{P(t)}$ is cut-balanced for every $t \geq 0$.*

Then, $x(t)$ converges to a limit point $\tilde{x} \in \mathbb{R}^V$ such that $\tilde{x}_u \in [x_{\min}(0), x_{\max}(0)]$ for all $u \in V$. Furthermore, let G_∞ be the limit graph of the sequence $(\mathscr{G}_{P(t)})_t$. If two nodes v and w belong to the same connected component of G_∞, then $\tilde{x}_v = \tilde{x}_w$.

Proof Since $\mathscr{G}_{P(t)}$ is cut-balanced for every $t \geq 0$, then also G_∞ is cut-balanced: This implies, by Lemma 3.3, that all weakly connected components of G_∞ are strongly connected. Let $C \subset V$ denote the node set of one such connected component and observe that there exists a time $t_0 \geq 0$ such that $P_{vw}(t) = P_{wv}(t) = 0$ for all $v \in C$, $w \notin C$ and $t \geq t_0$. Then, without loss of generality we disregard the dynamics before t_0 and we study the dynamics for $t \geq t_0$ over the component C.

Let $m \in C$ be a node such that $x_m(t_0) = \max\{x_u(t_0) : u \in C\}$. Define $S_{t_0} = \{m\}$ and, iteratively for $t \geq t_0$, a sequence of subsets $S_t \subseteq V$, by

$$S_{t+1} = \{u \in C : \exists w \in S_t \text{ such that } P_{vw}(t) > 0\}.$$

The sequence of sets S_t collects those nodes whose states at time t are influenced by the state of node m at time t_0. Note that because of the nondegeneracy assumption, the inclusion $S_{t+1} \supseteq S_t$ holds for every $t \geq t_0$. Let $t^\star$ be the time at which S_t is maximal and assume by contradiction that $S_{t^\star} \neq C$. Then, there is no vertex outside $S_{t^\star}$ that is connected to any vertex in $S_{t^\star}$ for any time $t \geq t^\star$: So C is not strongly connected in the graph $\cup_{s \geq t^\star} \mathscr{G}_{P(s)}$, contradicting the assumptions. Hence, $S_{t^\star} = C$.

We claim that for every $t_0 \leq t \leq t^\star$ and every $v \in S_t$, it holds

$$x_v(t) \geq \min_{u \in C} x_u(t_0) + \alpha^{|S_t|-1} \left(\max_{u \in C} x_u(t_0) - \min_{u \in C} x_u(t_0) \right), \tag{3.4}$$

where $\alpha > 0$ is the nondegeneracy constant. This fact can be shown by induction on t. For $t = t_0$, we have $S_{t_0} = \{m\}$ and so (3.4) trivially holds. For the induction step, we need to consider two cases. If $S_{t+1} = S_t$, then at time t every $w \in S_t$ only influences nodes in S_t. By the cut-balance assumption, every $v \in S_t$ is then only influenced by nodes in S_t. Hence, for every $v \in S_{t+1}$

$$
\begin{aligned}
x_v(t+1) &= \sum_{w \in S_t} P_{vw}(t) x_w(t) \\
&\geq \sum_{w \in S_t} P_{vw}(t) \left(\min_{u \in C} x_u(t_0) + \alpha^{|S_t|-1} \left(\max_{u \in C} x_u(t_0) - \min_{u \in C} x_u(t_0) \right) \right) \\
&= \min_{u \in C} x_u(t_0) + \alpha^{|S_{t+1}|-1} \left(\max_{u \in C} x_u(t_0) - \min_{u \in C} x_u(t_0) \right).
\end{aligned}
$$

If instead $S_{t+1} \neq S_t$, we note that for every $v \in S_{t+1}$, there is at least one $w \in S_t$ and such that $P_{vw}(t) > 0$. Indeed, if $v \notin S_t$, then v is, by construction, connected to at least one node $w \in S_t$, whereas every $v \in S_t$ is by the hypothesis always connected to itself. Since all (positive) entries $P_{vw}(t)$ are by hypothesis lower-bounded by α, this together with the induction hypothesis implies that

$$
\begin{aligned}
x_v(t+1) &= \sum_{w \in V} P_{vw}(t) x_w(t) \\
&\geq \sum_{w \in S_t} P_{vw}(t) \left(\min_{u \in C} x_u(t_0) + \alpha^{|S_t|-1} \left(\max_{u \in C} x_u(t_0) - \min_{u \in C} x_u(t_0) \right) \right) + \sum_{w \notin S_t} P_{vw}(t) \min_{u \in C} x_u(t_0) \\
&\geq \min_{u \in C} x_u(t_0) + \alpha \, \alpha^{|S_t|-1} \left(\max_{u \in C} x_u(t_0) - \min_{u \in C} x_u(t_0) \right) \\
&\geq \min_{u \in C} x_u(t_0) + \alpha^{|S_{t+1}|-1} \left(\max_{u \in C} x_u(t_0) - \min_{u \in C} x_u(t_0) \right),
\end{aligned}
$$

thus proving (3.4). As $\max_{u \in C} x_u(t)$ is not increasing, inequality (3.4) implies that

$$\max_{u \in C} x_u(t^\star) - \min_{u \in C} x_u(t^\star) \le (1 - \alpha^{|C|-1})(\max_{u \in C} x_u(t_0) - \min_{u \in C} x_u(t_0)).$$

By repeating all the above reasoning starting from $t_1 = t^\star$ and so on, we can construct a sequence of times $\{t_k \ : \ k \ge 0\}$ such that for every k it holds that

$$\max_{u \in C} x_u(t_{k+1}) - \min_{u \in C} x_u(t_{k+1}) \le (1 - \alpha^{|C|-1})(\max_{u \in C} x_u(t_k) - \min_{u \in C} x_u(t_k)).$$

This fact implies that all nodes $u \in C$ converge to consensus and thus proves the result. □

Remarkably, Theorem 3.2 does not contain any connectivity assumption other than cut-balance. As a consequence, it does not guarantee consensus among all states, but only convergence and "local" consensus inside each connected component of the limit graph. Global consensus can instead be obtained by restoring an assumption of global connectivity, as in the following two results which immediately follow from Theorem 3.2.

Corollary 3.1 (Time-dependent consensus II) *Consider system* (3.1). *Assume that*

 (i) *the set of matrices* $\{P(t)\}$ *is nondegenerate;*
 (ii) *for every* $t \ge 0$, *the graph* $\mathcal{G}_{P(t)}$ *is cut-balanced; and*
 (iii) *for every* $t \ge 0$, *the graph* $\bigcup_{s \ge 0} \mathcal{G}_{P(t+s)}$ *is weakly connected.*

Then, $x(t)$ *converges to a point in* span$\{\mathbf{1}\}$ *from every initial condition in* $\mathbb{R}^V$.

Note that condition (iii) is equivalent to G_∞ being strongly connected (as the reader may verify). As a special case, we recover the following result on convergence for symmetric graphs. We note that this theorem does not require the matrix $P(t)$ to be symmetric but just its induced graph to be symmetric, i.e., to encode *reciprocal* communications.

Corollary 3.2 (Time-dependent consensus III) *Consider system* (3.1). *Assume that*

 (i) *the set of matrices* $\{P(t)\}$ *is nondegenerate;*
 (ii) *for every* $t \ge 0$, *the graph* $\mathcal{G}_{P(t)}$ *is symmetric; and*
 (iii) *for every* $t \ge 0$, *the graph* $\bigcup_{s \ge 0} \mathcal{G}_{P(t+s)}$ *is connected.*

Then, $x(t)$ *converges to a point in* span$\{\mathbf{1}\}$ *from every initial condition in* $\mathbb{R}^V$.

We stress that these two results do not require any uniform connectivity assumption, and indeed, their proofs do not rely on Lemma 2.1, as opposed to Theorem 3.1. As an example, Corollary 3.2 implies that the sequence $P_1(t)$ in Example 3.1 leads to a consensus.

The interest in a convergence result such as Theorem 3.2, which avoids connectivity assumptions, becomes more apparent in those contexts where checking

connectivity is difficult, for instance because the evolution of $P(t)$ depends on the current state $x(t)$. We are going to illustrate this difficulty with a very popular example, known as *Krause's model*.

Example 3.2 (*Krause's model*) In this dynamics, agent v trusts, i.e., takes into account for its update, only those agents w whose current state $x_w(t)$ is close enough to $x_v(t)$. More precisely, we fix a threshold $\varepsilon > 0$ and, for all $t \in \mathbb{Z}_{\geq 0}$ and all $v \in V$, we let $N_v(t) = \{u \in V : |x_v(t) - x_u(t)| \leq \varepsilon\}$. Given $\rho \in (0, 1]$, we then define the dynamics

$$x_v(t+1) = x_v(t) + \frac{\rho}{|N_v(t)|} \sum_{w \in N_v(t)} (x_w(t) - x_v(t)) \qquad v \in V. \qquad (3.5)$$

Convergence of dynamics (3.5) can be deduced from Theorem 3.2. Notice indeed that if we define

$$P_{vw}(t) = \begin{cases} 1 - \rho \frac{|N_v(t)| - 1}{|N_v(t)|} & \text{if } w = v \\ \frac{\rho}{|N_v(t)|} & \text{if } w \in N_v(t), w \neq v \\ 0 & \text{if } w \notin N_v(t) \end{cases}$$

it is immediate to check that $x(t+1) = P(t)x(t)$. Notice that $\mathscr{G}_{P(t)}$ is symmetric so that assumption (ii) is verified. Nondegeneracy follows easily from the definition as $P_{vv}(t) \geq 1 - \rho \frac{N-1}{N}$ for every $v \in V$ while, if $v \neq w$ and $P_{vw}(t) > 0$, it follows that $P_{vw}(t) = \rho/|N_v(t)| \geq \rho/N$.

Notice that, in this model, the matrix $P(t)$ actually depends on the state of agents at time t and, as a consequence, the model is *nonlinear*. There is no way to guarantee a priori uniform connectivity conditions on the sequence of corresponding graphs, so
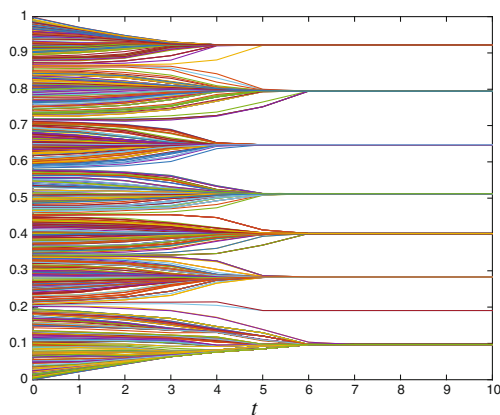
Fig. 3.2 A typical evolution under the time-varying dynamics (3.5) for a large number of agents: In this case, $N = 1000$ and $\varepsilon = 0.05$. Observe that clusters are approximately 2ε apart

that Theorem 3.1 cannot be applied. Simulations (see Fig. 3.2) demonstrate that the limit state is not a consensus point, but instead a collection of disconnected "clusters," composed of agents which share the same limit opinion. Moreover, if $\rho = 1$, then convergence is attained in finite time (see Exercise 3.4).

Krause's dynamics was originally proposed to model opinion dynamics with bounded confidence, but can also represent a simple model of one-dimensional vehicle rendezvous with limited visibility. This model has received much attention in the last years and many generalizations have been proposed: In Exercise 3.5, we study one of these.

3.3 Randomized Updates

This section presents time-varying consensus algorithms, where the update matrix is selected at each time step by a random process. Given a set of nodes V of finite cardinality N, we consider for every time $t \in \mathbb{Z}_{\geq 0}$ a random vector $x(t) \in \mathbb{R}^V$ evolving according to a random discrete-time system of the form

$$x(t + 1) = P(t)x(t) \qquad t \in \mathbb{Z}_{\geq 0}, \tag{3.6}$$

where $P(t)$ is a stochastic matrix for each $t \geq 0$ and $\left(P(t)\right)_{t \geq 0}$ is a sequence of independent and identically distributed random variables. Note that the initial condition is unknown but fixed (not random) and that all the randomness originates from generating the sequence of $P(t)$s. Consequently, in what follows the phrase "almost surely" means "with probability 1" with respect to the matrix selection process.

We begin our discussion from the following example of randomized dynamics. Let a symmetric graph $G = (V, E)$ be given, and for each time step $t \geq 0$, let an edge (v, w) be chosen in E, according to a uniform distribution over E. Define

$$x_v(t + 1) = \frac{1}{2} x_v(t) + \frac{1}{2} x_w(t),$$
$$x_w(t + 1) = \frac{1}{2} x_w(t) + \frac{1}{2} x_v(t),$$
$$x_u(t + 1) = x_u(t), \quad \text{for } u \neq v, w.$$

This dynamics can be written in the form (3.6) by defining

$$P^{(v,w)} = I - \frac{1}{2}(e_v e_v^* - e_v e_w^* - e_w e_v^* + e_w e_w^*),$$

where e_u is the uth vector of the canonical basis of $\mathbb{R}^V$, and $\mathbb{P}[P(t) = P^{(v,w)}] = \frac{1}{|E|}$.

We shall refer to this dynamics as the uniform symmetric gossip (USG).

Proposition 3.3 (USG convergence) *A uniform symmetric gossip dynamics converges almost surely to the average of the initial conditions, provided the underlying graph is connected.*

Proof For any $t_0 \geq 0$ and any edge $(v, w) \in E$, where E is the edge set of the underlying graph, we evaluate the probability of the event "the edge (v, w) is not selected for update at any time larger than t_0." Since the probability that (v, w) is not selected at any time t is $1 - \frac{1}{|E|}$, the probability that (v, w) is not selected for all times s such that $t_0 \leq s < t$ is $\left(1 - \frac{1}{|E|}\right)^{t-t_0}$. Since $\lim_{t \to +\infty} \left(1 - \frac{1}{|E|}\right)^{t-t_0} = 0$, we argue that (v, w) is selected infinitely often after t_0 with probability 1. This fact implies that $G_\infty = G$ almost surely. Since G is connected, convergence can be deduced from Corollary 3.2. □

In the following, we are going to present a more general convergence result that subsumes Proposition 3.3. We will rely on the results of Sect. 3.1. To begin with, let us go back to system (3.6) and let us study the *expected* dynamics, that is, the dynamics of $\mathbb{E}[x(t)]$. Equation (3.6) and the independence among $P(t)$s imply $\mathbb{E}[x(t+1)|x(t)] = \mathbb{E}[P(t)]x(t)$ for all t. Then, denoting $\bar{P} := \mathbb{E}[P(t)]$, we have

$$\mathbb{E}[x(t+1)] = \bar{P}\,\mathbb{E}[x(t)].$$

Note that $\bar{P}$ is a stochastic matrix. If the graph associated with $\bar{P}$ has a globally reachable aperiodic node, by Theorem 2.2 we have that $\mathbb{E}[x(t)]$ converges to a consensus point $c\mathbf{1}$. Moreover, the convergence rate is given by $\rho_2(\bar{P})$ and $c = v^*x(0)$, where v is the normalized dominant left eigenvector of $\bar{P}$. In principle, convergence of the expected dynamics does not, by itself, guarantee convergence of the random dynamics. However, the next result provides general and intuitive conditions for the convergence of (3.6), which are indeed based on the convergence properties of the expected dynamics.

Theorem 3.3 (Almost sure convergence to consensus) *Consider the dynamical system (3.6) and assume that the matrices $P(t)$ are independently and randomly sampled from an ensemble $\mathscr{P}$ of nondegenerate stochastic matrices equipped with a fixed probability distribution. Then, the following three facts are equivalent:*

 (i) *for every initial condition, there exists a scalar random variable x_∞ such that $x(t)$ converges almost surely to $x_\infty \mathbf{1}$;*
 (ii) *$\rho_2(\bar{P}) < 1$;*
 (iii) *the "expected graph" $\mathscr{G}_{\bar{P}}$ has a globally reachable node.*

Proof $(i) \Rightarrow (ii)$: Being $x(t)$ a bounded sequence, if $x(t)$ converges almost surely to $x_\infty \mathbf{1}$, then also $\mathbb{E}[x(t)]$ converges to $\mathbb{E}[x_\infty]\mathbf{1}$ by the dominated convergence theorem. As $\mathbb{E}[x(t+1)] = \bar{P}\,\mathbb{E}[x(t)]$, then necessarily $\rho_2(\bar{P}) < 1$.

$(ii) \Rightarrow (iii)$: Under the assumption on the diagonal of $P(t)$, the graph $\mathscr{G}_{\bar{P}}$ has a globally reachable node if and only if $\rho_2(\bar{P}) < 1$.

$(iii) \Rightarrow (i)$: If $\mathscr{G}_{\bar{P}}$ has a globally reachable node, say $k \in V$, then there exists $m \in \mathbb{N}$ such that the kth column of $\bar{P}^m$ is positive. This implies that, for every t_0, the entry $(P(t_0 + m) \dots P(t_0 + 1)P(t_0))_{vk}$ has a positive probability of being positive: Denote such probability by p_v. Now, recall that $P_{kk}(t) > 0$ almost surely: This implies that the kth column of matrix $P(t_0 + Nm) \dots P(t_0 + 1)P(t_0)$ is positive with probability $\prod_{u \in V} p_u$. Consequently, there exists $\alpha > 0$ such that, with positive probability, each element of this column is larger than α. Now, let us define the sequence of times $t_h = mNh$ for $h \in \mathbb{Z}_{\geq 0}$: Reasoning as in the proof of Lemma 3.1, we can apply Lemma 2.1 to argue that

$$\max_{v \in V} x_v(t_{k+1}) - \min_{v \in V} x_v(t_{k+1}) \leq (1 - \alpha)\left(\max_{v \in V} x_v(t_k) - \min_{v \in V} x_v(t_k)\right)$$

with a positive probability which does not depend on t_k. Hence, this inequality almost surely holds for infinitely many k and then almost surely $x(t)$ converges to consensus. $\square$

It is remarkable that Theorem 3.3 translates on the expected graph of the network the same condition for consensus that holds for time-invariant networks. Following this analogy, one would expect that the essential spectral radius of $\mathbb{E}[P(t)]$ determines the speed of convergence of the algorithm. A result in this direction can be found by a suitable mean-square analysis. Let us denote the current empirical variance as

$$x_{\text{var}}(t) := \frac{1}{N}||x(t) - x_{\text{ave}}(t)\mathbf{1}||^2 = \frac{1}{N}||\Omega x(t)||^2,$$

where $\Omega = I - \frac{1}{N}\mathbf{1}\mathbf{1}^*$, and define the mean-square rate of convergence as

$$R := \sup_{x(0)} \limsup_{t \to +\infty} \mathbb{E}[x_{\text{var}}(t)]^{1/t}. \tag{3.7}$$

Notice that

$$\mathbb{E}[x_{\text{var}}(t)] = \frac{1}{N}\mathbb{E}[x(t)^*\Omega x(t)] = \frac{1}{N}x(0)^*\Delta(t)x(0),$$

where

$$\Delta(t) := \mathbb{E}[P(0)^* P(1)^* \dots P(t-1)\Omega P(t-1) \dots P(1)P(0)]$$

if $t \geq 1$ and $\Delta(0) := \Omega$. Clearly,

$$\Delta(t+1) = \mathbb{E}[P(0)^*\Delta(t)P(0)].$$

This recursion shows that $\Delta(t)$ is the solution of a linear dynamical system, which can be written in the form

$$\Delta(t+1) = \mathscr{L}(\Delta(t))$$

where $\mathscr{L} : \mathbb{R}^{V \times V} \to \mathbb{R}^{V \times V}$ is given by $\mathscr{L}(M) = \mathbb{E}[P(0)^* M P(0)]$. The knowledge of the operator $\mathscr{L}$, in principle, provides all information about the mean-square analysis. For instance, we have that R is the spectral radius of the operator $\mathscr{L}$, restricted to the smallest $\mathscr{L}$-invariant subspace of $\mathbb{R}^{V \times V}$ containing Ω. This characterization, however, is not very useful, because the operator $\mathscr{L}$ is difficult to compute in the applications. The next result provides rate estimates that are easier to compute.

Proposition 3.4 (Mean-square convergence rate) *Consider* (3.6) *and the convergence rate R as in* (3.7). *Then,*

$$\rho_2(\bar{P})^2 \le R \le \text{sr}\left(\mathbb{E}[P(t)^* \Omega P(t)]\right), \tag{3.8}$$

where we recall that $\Omega = I - N^{-1}\mathbf{1}\mathbf{1}^*$ and $\text{sr}(\cdot)$ *denotes the spectral radius of a matrix.*

Proof We start from the first inequality. We define $Q(t) = P(t-1)\dots P(0)$ and notice that

$$\mathbb{E}[x^*(t)\Omega x(t)] = \mathbb{E}[||\Omega x(t)||^2] = \mathbb{E}[||\Omega Q(t)x(0)||^2].$$

Now using Jensen's inequality, we have that

$$\mathbb{E}[||\Omega Q(t)x(0)||^2] \ge ||\mathbb{E}[\Omega Q(t)x(0)]||^2 = ||\Omega \bar{P}^t x(0)||^2,$$

which proves the inequality.

In order to prove the second inequality, let $y \in \mathbb{R}^V$ and note

$$\begin{aligned}
y^*\mathbb{E}[P(0)^* \Omega P(0)]y &= \mathbb{E}[y^* \Omega P(0)^* \Omega P(0) \Omega y] \\
&= y^* \Omega \mathbb{E}[P(0)^* \Omega P(0)] \Omega y \\
&\le ||\mathbb{E}[P(0)^* \Omega P(0)]|| y^* \Omega y,
\end{aligned}$$

by the symmetry of the matrix. We deduce that $\mathscr{L}(\Omega) \le ||\mathscr{L}(\Omega)||\Omega$. This fact, together with the remark that if $M_1 \le M_2$, then $\mathscr{L}(M_1) \le \mathscr{L}(M_2)$, implies that

$$\mathscr{L}^t(\Omega) = \mathscr{L}^{t-1}(\mathscr{L}(\Omega)) \le \mathscr{L}^{t-1}(||\mathscr{L}(\Omega)||\Omega) = ||\mathscr{L}(\Omega)||\mathscr{L}^{t-1}(\Omega).$$

By iterating this reasoning, we get $\mathscr{L}^t(\Omega) \le ||\mathscr{L}(\Omega)||^t \Omega$, which gives the thesis. $\qquad\square$

Note that if all matrices $P(t)$ are symmetric, then $\mathbb{E}[P(t)^* \Omega P(t)] = \mathbb{E}[P^2(t)] - \frac{1}{N}\mathbf{1}\mathbf{1}^*$. In the special case of the USG algorithm, we further have $\mathbb{E}[P^2(t)] - \frac{1}{N}\mathbf{1}\mathbf{1}^* = \mathbb{E}[P(t)] - \frac{1}{N}\mathbf{1}\mathbf{1}^*$ and we can thus argue that $\rho_2(\mathbb{E}[P(t)])^2 \le R \le \rho_2(\mathbb{E}[P(t)])$.

We now introduce two examples of randomized averaging algorithms, which can be studied by the above results. The first example generalizes the USG and features the

activation of *one pair* of connected nodes per time step: The two nodes communicate with each other and both update their states.

Example 3.3 (symmetric gossip algorithm (SG)) Let a weighted graph $G=(I, E, W)$ and $q \in (0, 1)$ be given, such that W is *symmetric* and $\mathbf{1}^*W\mathbf{1} = 1$. For every $t \geq 0$, one edge $(v, w) \in E$ is sampled from a distribution such that the probability of selecting (v, w) is W_{vw}. Then,

$$x_v(t + 1) = (1 - q)\, x_v(t) + q\, x_w(t)$$
$$x_w(t + 1) = (1 - q)\, x_w(t) + q\, x_v(t)$$
$$x_u(t + 1) = x_u(t) \quad \text{for } u \neq v, w.$$

Both W and q can be considered in principle as design parameters, with respect to which one can optimize the performance.

In order to analyze the SG algorithm, for every (v, w) we let

$$P^{(v,w)} := I - q(e_v - e_w)(e_v - e_w)^* = I - q(e_v e_v^* - e_w e_v^* - e_v e_w^* + e_w e_w^*),$$

where e_u is the uth vector of the canonical basis of $\mathbb{R}^V$. Note that trivially $W = \sum_{(v,w)} W_{vw} e_v e_w^*$. Then, the distribution of $P(t)$ is concentrated on these matrices and $\mathbb{P}[P(t) = P^{(v,w)}] = W_{vw}$. We have that (using the notation for the Laplacian of a matrix)

$$\mathbb{E}[P(t)] = \sum_{(v,w)} W_{vw} P^{(v,w)}$$
$$= I - q \sum_{(v,w)} W_{vw}(e_v - e_w)(e_v - e_w)^*$$
$$= I - 2q\, L(W).$$

Note that if the graph associated with W is strongly connected, then the average graph is automatically strongly connected. Since in the SG all the diagonal elements of $P(t)$ are nonzero with probability 1 and all the $P(t)$s are symmetric, we can apply Theorem 3.3 and conclude that this algorithm yields average consensus almost surely. Moreover, noting that $\|e_v - e_w\|_2^2 = 2$ and

$$(P^{(v,w)})^2 = I - 2q(e_v - e_w)(e_v - e_w)^* + q^2(e_v - e_w)(e_v - e_w)^*(e_v - e_w)(e_v - e_w)^*$$
$$= I - 2q(1 - q)(e_v - e_w)(e_v - e_w)^*,$$

we argue that

$$\mathbb{E}[P(t)^* \Omega P(t)] = \mathbb{E}[P(t)^2] - \frac{1}{N}\mathbf{11}^* = \Omega - 4q(1 - q)L(W).$$

Then, by applying Proposition 3.4, we can estimate the convergence rate as

$$\rho_2(I - 2qL(W))^2 \le R \le \mathrm{sr}(\Omega - 4q(1 - q)L(W))$$

and, provided we denote by λ the smallest nonzero eigenvalue of $L(W)$, as

$$1 - 4q\lambda \le R \le 1 - 4q(1 - q)\lambda.$$

The next example features the activation of *one node* per time step. The activated node communicates its current state to all its neighbors, which in turn update their states. We note that the algorithm is inherently asymmetric: As a consequence, the average of the initial states is not preserved.

Example 3.4 (*Broadcast gossip algorithm (BG)*) Let there be $q \in (0, 1)$ and a directed graph $G = (V, E)$ whose adjacency matrix is denoted by $A \in \{0, 1\}^{V \times V}$. For every $t \ge 0$, one node w is sampled from a uniform distribution over V. Then,

$$
\begin{aligned}
x_v(t + 1) &= (1 - q)\,x_v(t) + q\,x_w(t) && \text{if } A_{vw} > 0 \\
x_v(t + 1) &= x_v(t) && \text{otherwise.}
\end{aligned}
$$

In other words, one randomly selected node broadcasts its value to all its neighbors, which update their values accordingly.

For the analysis of this algorithm, we define

$$P^{(w)} = I - q \sum_{v:A_{vw}>0} (e_v e_v^* - e_v e_w^*)$$

and note that $\mathbb{P}[P(t) = P^{(w)}] = \frac{1}{N}$. Then,

$$\mathbb{E}[P(t)] = I - \frac{q}{N}L.$$

If the graph G is strongly connected, then the algorithm converges to consensus almost surely. Before we further investigate the properties of the BG algorithm, we assume that the graph is topologically balanced, i.e., $A\mathbf{1} = A^*\mathbf{1}$. This property in particular implies that $\mathbf{1}^*L = \mathbf{1}^*$ and then $\mathbb{E}[x_\infty] = x_{\mathrm{ave}}(0)$. Moreover, the reader may compute that

$$\mathbb{E}[P(t)^* P(t)] = I - \frac{q(1 - q)}{N}(L + L^*) \tag{3.9a}$$

$$\mathbb{E}[P(t)^*\mathbf{1}\mathbf{1}^* P(t)] = \mathbf{1}\mathbf{1}^* + \frac{q^2}{N}LL^*. \tag{3.9b}$$

As a consequence, the convergence rate can be estimated using Proposition 3.4 as

$$\rho_2\left(I - \frac{q(1 - q)}{N}(L + L^*)\right)^2 \le R \le \mathrm{sr}\left(\Omega - \frac{q(1 - q)}{N}(L + L^*) - \frac{q^2}{N^2}LL^*\right)$$

If we denote by 2λ the smallest nonzero eigenvalue of $L + L^*$, and we remark that LL^* is positive semidefinite, the above bounds can be simplified to

$$1 - 4\frac{q(1-q)}{N}\lambda \le R \le 1 - 2\frac{q(1-q)}{N}\lambda.$$

We have seen that, provided the graph is balanced, the BG algorithm yields $\mathbb{E}[x_\infty] = x_{\text{ave}}(0)$. Considering that x_∞ is a random variable, its spreading around the mean value needs to be evaluated. To this aim, we introduce the mean-square error $\mathbb{E}\left[(x_\infty - x_{\text{ave}}(0))^2\right]$ and below we provide a technical tool to estimate it. In order to state the result, it is again convenient to denote the empirical variance as

$$x_{\text{var}}(t) := \frac{1}{N}\sum_{i=1}^{N}\left(x_i(t) - x_{\text{ave}}(t)\right) = \frac{1}{N}x^*\Omega x(t).$$

Theorem 3.4 (Accuracy condition) *Consider dynamics* (3.6) *and assume that* $1^*\bar{P} = 1^*$ *and that there exists* $\gamma > 0$ *such that*[1]

$$\mathbb{E}[P^*11^*P] - 11^* \le \gamma\left(I - \mathbb{E}[P^*P]\right). \tag{3.10}$$

Then,

$$\mathbb{E}\left[(x_{\text{ave}}(t) - x_{\text{ave}}(0))^2\right] \le \frac{\gamma}{N + \gamma}\mathbb{E}\left[x_{\text{var}}(0) - x_{\text{var}}(t)\right]. \tag{3.11}$$

If additionally $\mathscr{G}_{\bar{P}}$ *has a globally reachable node, then*

$$\mathbb{E}\left[(x_\infty - x_{\text{ave}}(0))^2\right] \le \frac{\gamma}{N + \gamma}x_{\text{var}}(0). \tag{3.12}$$

Proof In the proof, we shall use the notation $x_{\text{ave}} = N^{-1}1^*x$ and $x_{\text{var}} = N^{-1}x^*\Omega x$ to denote the empirical average and variance of a generic vector $x \in \mathbb{R}^V$. We let

$$C(x) := N(\gamma + N)x_{\text{ave}}^2 + N\gamma\, x_{\text{var}}$$
$$= \frac{N(\gamma + N)}{N^2}x^*11^*x + \frac{N\gamma}{N}x^*\left(I - \frac{1}{N}11^*\right)x$$
$$= x^*\left(11^* + \gamma I\right)x.$$

Then, for a generic stochastic matrix P, we have that

$$C(Px) - C(x) = x^*\left(P^*11^*P + \gamma P^*P - 11^* - \gamma I\right)x.$$

---

[1]In this result, inequalities between matrices like $A \le B$ have to be intended as $A - B$ being negative semidefinite.

Consequently, condition (3.10) implies that for dynamics (3.6)

$$\mathbb{E}[C(x(t+1)) - C(x(t))|x(t)] = x(t)^* \Big(\mathbb{E}[P^* \mathbf{11}^* P + \gamma \, P^* P] - \mathbf{11}^* - \gamma \, I \Big) x(t) \leq 0$$

and $\mathbb{E}[C(x(t))] \leq C(x(0))$ for all $t \in \mathbb{N}$. This inequality can be rewritten as

$$\mathbb{E}\left[x_{\text{ave}}(t)^2 - x_{\text{ave}}(0)^2\right] \leq \frac{\gamma}{N + \gamma} \mathbb{E}\left[x_{\text{var}}(0) - x_{\text{var}}(t)\right].$$

This inequality implies (3.11) if $x_{\text{ave}}(0) = 0$: The general case follows by applying this special case to the translated dynamics $x - x_{\text{ave}}(0)\mathbf{1}$. Finally, inequality (3.12) is an immediate corollary of convergence. $\qquad\square$

In the case of the broadcast gossip algorithm, (3.9) implies that (3.10) reads

$$\frac{q^2}{N} LL^* \leq \gamma \, \frac{q(1-q)}{N} (L + L^*),$$

which holds true for $\gamma = d_{\text{max}} \frac{q}{1-q}$ because for balanced graphs $LL^* \leq d_{\text{max}}(L+L^*)$. Consequently,

$$\mathbb{E}\left[(x_\infty - x_{\text{ave}}(0))^2\right] \leq \frac{q}{1-q} \frac{d_{\text{max}}}{N} x_{\text{var}}(0).$$

Remarkably, as long as $d_{\text{max}} = o(N)$, this upper bound goes to zero as N goes to infinity, that is, *the error committed by the algorithm in approximating the average becomes negligible on large networks.*

Exercises

Exercise 3.1 (*Strong positivity* [35]) Consider the sequence of matrices

$$P(t) = \begin{pmatrix} 1 - \alpha_t & \alpha_t \\ \alpha_t & 1 - \alpha_t \end{pmatrix}$$

where $\alpha_t \in [0, 1]$ is a given sequence. Observe that all $P(t)$ are aperiodic and irreducible.

(i) Prove that, if $\alpha_t \geq \alpha > 0$ for all t, then the sequence $P(t)$ leads to a consensus.
(ii) For sequences $\alpha_t \to 0$ when $t \to +\infty$, find sufficient conditions on the speed of convergence which guarantee that $P(t)$ leads to a consensus.
(iii) Find an explicit example of a sequence $\alpha_t \to 0$ for $t \to +\infty$, for which $P(t)$ does not lead to a consensus.

Exercise 3.2 (*Time-varying consensus on the line graph*) Let $V = \{0, \ldots, N-1\}$ and consider the directed line graph $\mathbf{L} = (V, E)$. Let the vector $e_u \in \mathbb{R}^V$ be such that

the vth component of e_u is equal to 1 if $v = u$ and to 0 otherwise, and define the matrix $P^{(u,v)} = I - \frac{1}{2}(e_u e_u^* - e_u e_v^*)$. Consider a time-dependent consensus algorithm (3.1) with

$$P(t) = P^{(k,k+1)} \quad \text{where } k = t \pmod{N}.$$

(i) Verify that the dynamics satisfies the assumptions of Theorem 3.1, and find the minimal T for the connectivity assumption.
(ii) Verify that the dynamics satisfies the assumptions of Lemma 3.1, finding the suitable value of B. Compare B with the value of T found in (i).

Exercise 3.3 (*Uniform connectivity* [35]) Consider (3.1) with $x(0) = (0, 1, 1)^*$ and the sequence $\{P(t)\}_t$ defined as follows. Let

$$P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad P_2 = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$P_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \end{bmatrix} \qquad P_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

and $Q_s = \underbrace{P_1, \dots, P_1}_{2s}, P_2, \underbrace{P_3, \dots, P_3}_{2s+1}, P_4$. Assume the sequence $P(t)$ is the concatenation of $Q_0, Q_1, Q_2, \dots$. Then, show that $x(t)$ does not converge to a consensus.

Exercise 3.4 (*Krause's convergence time*) Consider Krause's dynamics (3.5).

(i) Show that the order between opinions is preserved, i.e., for all $t \geq 0$, if $x_v(t) \leq x_w(t)$, then $x_v(t + 1) \leq x_w(t + 1)$. This implies that we can assume (without loss of generality) that the agents are sorted, i.e., if $v < w$, then $x_v < x_w$.
(ii) Show that if at some time t the distance between two consecutive agent opinions $x_v(t)$ and $x_{v+1}(t)$ is larger than ε, then it remains so for all time $s > t$.
(iii) Assume from now on that $\rho = 1$. Show that there exist $T \in \mathbb{N}$ and $\tilde{x} \in \mathbb{R}^V$ such that $x(t) = \tilde{x}$ for all $t \geq T$.
(iv) Show that for all $v, w \in V$, either $|\tilde{x}_v - \tilde{x}_w| > \varepsilon$ or $\tilde{x}_v = \tilde{x}_w$.
(v) Estimate the worst-case converge time $\bar{T} = \sup_{x(0)} \inf\{t : x(t) = \tilde{x}\}$ ([5, Sect. 4.6.1]).

Exercise 3.5 (*Unbounded confidence*) Consider the following generalized Krause's model. Fix a continuous function $\xi : [0, +\infty) \to [0, +\infty)$ such that $\xi(x) > 0$ for all $x \geq 0$, and define

$$x_v(t + 1) = x_v(t) + \frac{\rho}{\sum\limits_{w \in V} \xi(|x_w(t) - x_v(t)|)} \sum\limits_{w \in V} \xi(|x_w(t) - x_v(t)|)(x_w(t) - x_v(t)),$$

(3.13)

where $\rho \in (0, 1)$. Fix some initial condition $x(0)$ and let $P(t)$ be the sequence of matrix such that $x(t + 1) = P(t)x(t)$.

(i) Prove that the sequence $P(t)$ is nondegenerate and conclude, using Corollary 3.2, that it leads to a consensus.

(ii) Assume that $\xi(x) = e^{-x^2}$ and let $d(t) = \max\{x_u(t)\} - \min\{x_u(t)\}$. Using Lemma 2.1, prove that

$$d(t + 1) \leq \left(1 - e^{-d(t)^2}\right) d(t)$$

(iii) Assuming that $x(0)_v \in [-1, 1]$ for all $v \in V$, find, for fixed $\varepsilon > 0$, an estimate of the convergence time

$$t_\varepsilon := \inf\{t \mid d(t) \leq \varepsilon\}$$

Compare this estimate with explicit simulations of (3.13) for $N = 10$.

Exercise 3.6 (*Broadcast on a star*) Let $S_N = (V, E)$ be a star with N leaves. Consider the following randomized consensus algorithm. For all positive integers t, sample one node w from a uniform distribution over V, and update the states as follows:

$$x_v(t + 1) = (1 - q)x_v(t) + q x_w(t) \quad \text{if } (v, w) \in E$$
$$x_v(t + 1) = x_v(t) \quad \text{if } (v, w) \notin E.$$

The update parameter satisfies $q \in (0, 1)$.

(i) Write down the update rule for the proposed algorithm in matrix form.

(ii) Compute the expected update matrix $\bar{P} = \mathbb{E}[P(t)]$.

(iii) Let $x_{\text{ave}}(t) = \frac{1}{N}\sum_{v \in V} x_v(t)$. Verify that, although in general $x_{\text{ave}}(t + 1) \neq x_{\text{ave}}(t)$, nevertheless $\mathbb{E}[x_{\text{ave}}(t + 1)] = \mathbb{E}[x_{\text{ave}}(t)]$.

(iv) Show that the algorithm ensures almost sure convergence of the states.

(v) Compute the second largest eigenvalue of $\bar{P}$. To this goal, you may use Exercise 2.18.

(vi) Estimate the convergence rate R of the algorithm as a function of q and N, and conclude that $\lim_{N \to \infty} R = 1$ irrespective of q.

Exercise 3.7 (*Triplet-gossip*) Consider the following random dynamics on a complete graph $G = (V, E)$. At every discrete-time step t, three agents u, v, w are uniformly and independently sampled from V, and they update their internal state as follows:

$$x_u(t + 1) = x_v(t + 1) = x_w(t + 1) = \frac{x_u(t) + x_v(t) + x_w(t)}{3}$$

Let $P(t)$ be the corresponding matrix acting on the full vector $x(t)$.

(i) Compute $\bar{P} = \mathbb{E}[P(t)]$, its eigenvalues, and its spectral gap.

(ii) Give an estimation of $\mathbb{E}\|\Omega x(t)\|^2$ analogous to what is done for the pairwise gossip algorithm in Example 3.3.

Exercise 3.8 (*Asynchronous asymmetric gossip algorithm (AAGA)*) This exercise gives an example of randomized algorithm, in which one *directed* edge is activated at each time step, resulting in an asymmetric update rule.

Let a weighted graph $G = (V, E, W)$ and $q \in (0, 1)$ be given, such that $W\mathbf{1} = W^*\mathbf{1}$ and $\mathbf{1}^*W\mathbf{1} = 1$. For every $t \geq 0$, one edge (v, w) is sampled from a distribution such that the probability of selecting (v, w) is W_{vw}. Then, we define

$$x_v(t + 1) = (1 - q)\, x_v(t) + q\, x_w(t) \tag{3.14}$$

and $x_u(t + 1) = x_u(t)$ for $u \neq v$.

(i) Verify that (as proved in [18, Sect. 4])

$$\mathbb{E}[P(t)] = I - qL(W)$$
$$\mathbb{E}[P(t)^*P(t)] = I - q(1 - q)L(W + W^*)$$
$$\mathbb{E}[P(t)^*\mathbf{1}\mathbf{1}^*P(t)] = \mathbf{1}\mathbf{1} + q^2L(W + W^*).$$

(ii) Assume from now on that the graph G is strongly connected. Prove that system (3.14) almost surely converges to a limit value x_∞ such that $\mathbb{E}[x_\infty] = x_{\mathrm{ave}}(0)$ and

$$\mathbb{E}\left[(x_\infty - x_{\mathrm{ave}}(0))^2\right] \leq \frac{q}{1 - q}\frac{1}{N}x_{\mathrm{var}}(0).$$

(iii) Assume moreover that W is symmetric and denote by λ the spectral gap of $L(W)$. Show that the convergence rate is bounded by

$$1 - 2q\lambda \leq R \leq 1 - 2q\left((1 - q) + \frac{q}{N}\right)\lambda,$$

provided N is large enough.

Bibliographical Notes

Deterministic networks. Our choice of results on deterministic time-dependent consensus mostly consist of necessary and sufficient conditions for convergence. While the convergence analysis of time-invariant averaging dates back at least to De Groot [13] in 1974, sufficient conditions for the time-varying case were given by [45] in 1984 and later by [7, 8, 28, 31]. The results on convergence to consensus that we present in Sect. 3.1 appeared, in quite a different formulation, in [35]: The counterexample in Exercise 3.3 is in the original paper. Our version of the results is based on the analysis in [24, 27].

In consensus-seeking systems, results which do make a "global" assumption of connectivity can ensure convergence but possibly not consensus. Such results are

motivated by the difficulty to satisfy, in some applications, connectivity conditions over time. In this spirit, we have presented Theorem 3.2, which is more general than early results as those in [31]. The result can be found in [25, Theorem 2]: We present it here with a proof which is adapted to the rest of our arguments. The works [26, 27] developed the notion of *cut-balance*, which has been used and extended in several subsequent works, including [11, 33, 42, 44]. However, it is clear that in certain applications, e.g., vehicle rendezvous, mere convergence is not satisfactory. For this reason, there has been much work devoted to variations of the consensus algorithm, which inherently guarantee connectivity. A discussion about this *connectivity maintenance* issue may be found, for instance, in [5, Chap. 4]. In the opposite direction, Krause's model is a simple but very interesting example of a consensus-seeking dynamics without a global connectivity assumption. The dynamics was originally proposed in [23, 29], as a model for opinion dynamics with bounded confidence [32]. Krause's dynamics have been the topic of several works [2], which have also considered variations of the dynamics that feature continuous-time evolution [3, 9], multi-dimensional opinions [16, 38], heterogenous thresholds [34], and continua of agents [3, 6].

Instead, we did not investigate much two important issues that have been discussed in detail in Chap. 2 for time-invariant networks: speed of convergence and limit state. Studying the speed of convergence of time-dependent consensus algorithms is indeed quite delicate. First, it is essential to assume connectedness on bounded interval; otherwise, the algorithm can be arbitrarily slowed down by introducing arbitrary sequences of disconnected graphs. Moreover, even if connectivity on bounded intervals is assumed, the convergence time can be large (as in Exercise 3.2), even exponentially large in the interval size and in N: We refer to [37] for a detailed discussion. Results have been proved for specific dynamics, such as Krause's, see Exercise 3.4 and [15]. Stronger results can be found by assuming the matrices $P(t)$ to be nondegenerate and their associated graphs connected at each time step: Recent results in this framework can be found in [36]. Also the issue of determining the consensus (or convergence) point has no simple answer in the literature. In principle, such an analysis reduces to studying the *absolute probability vectors* as defined in [39]. Explicit results, however, are only available in special cases: For instance, see [2] for a partial description of the limit states of Krause's dynamics and [46] for some recent developments on this matter.

Randomized networks. Theorem 3.3 was originally proven in [19, 40]. The proof presented here is new, although inspired by [40], and has been written to seamlessly take advantage of our treatment of the deterministic case. We acknowledge that Theorem 3.3 is not the most general convergence result for randomized consensus dynamics, because it requires positivity of the diagonal and statistical independence of the update matrices. This independence assumption can be significantly relaxed, at the price of using more subtle probabilistic tools. In [41], the condition $\rho_2(\mathbb{E}[P(t)]) < 1$ is proven to be necessary and sufficient for consensus, under the more general assumption that the sequence $P(t)$ is generated by an ergodic stationary process (and has positive diagonals). In [30], it is proved that in fact any adapted stochastic process is suitable, provided certain assumptions of uniform con-

nectivity hold. Also the assumption of positivity of the diagonal can be relaxed. A necessary and sufficient condition for convergence, which does not require to assume that the diagonal is positive, can be found in [19, Theorem 3.1] and is due to Cogburn [10]. An intermediate useful condition is *strong aperiodicity* as defined in [42]. Very general conditions for adapted processes of "balanced" stochastic matrices are provided in [43].

For randomized dynamics, we have presented some estimates on the speed of convergence, by using the mean-square analysis developed in [19]. Another important topic is studying the random variable x_∞ and its distance from $x_{\text{ave}}(0)$. Obtaining a complete characterization of the distribution of x_∞ seems to remain an open problem, but a few practical results are available to estimate its variance. In principle, the variance of the consensus value can be exactly computed by the formula in [41, Eq. (7)], which involves the dominant eigenvectors of the first two moments of the update matrix: However, this characterization can be inconvenient in the applications and does not provide clear insights on the scaling for large networks. More recently, an effective estimate has been derived in [22], providing conditions for the variance to go to zero as N goes to infinity. This result, which we presented in Theorem 3.4, covers a wide class of randomized algorithms that involve asymmetric communication or packet losses [20, 21].

There is a large variety of examples in randomized consensus dynamics, and our selection has focused on two models which have possibly been the most popular in recent literature [14]. The symmetric gossip algorithm in Example 3.3 was popularized in the systems and control community by the influential work [4]. The broadcast gossip algorithm in Example 3.4 has attracted a significant attention, because it involves broadcast communication and thus applies very naturally to wireless networks [1, 17, 19]: Formulas (3.9) can be found in [1, Lemma 4]. As we have shown, the consensus value of BG does not coincide with the average of the initial conditions, but this bias becomes negligible for large networks [22]. Another intuitive algorithm with the same property is the asymmetric gossip algorithm [18], which we introduce in Exercise 3.8. Finally, we would like to mention that there exist dynamics that combine gossiping and bounded confidence, proposed as opinion dynamics models [12, 47].

References

1. Aysal, T.C., Yildiz, M.E., Sarwate, A.D., Scaglione, A.: Broadcast gossip algorithms for consensus. IEEE Trans. Signal Process. **57**(7), 2748–2761 (2009)
2. Blondel, V.D., Hendrickx, J.M., Tsitsiklis, J.N.: On Krause's multi-agent consensus model with state-dependent connectivity. IEEE Trans. Autom. Control **54**(11), 2586–2597 (2009)
3. Blondel, V.D., Hendrickx, J.M., Tsitsiklis, J.N.: Continuous-time average-preserving opinion dynamics with opinion-dependent communications. SIAM J. Control Optim. **48**(8), 5214–5240 (2010)
4. Boyd, S., Ghosh, A., Prabhakar, B., Shah, D.: Randomized gossip algorithms. IEEE Trans. Inf. Theory **52**(6), 2508–2530 (2006)

5. Bullo, F., Cortés, J., Martínez, S.: Distributed Control of Robotic Networks. Applied Mathematics Series. Princeton University Press (2009)
6. Canuto, C., Fagnani, F., Tilli, P.: An Eulerian approach to the analysis of Krause's consensus models. SIAM J. Control Optim. **50**(1), 243–265 (2012)
7. Cao, M., Morse, A.S., Anderson, B.D.O.: Agreeing asynchronously. IEEE Trans. Autom. Control **53**(8), 1826–1838 (2008)
8. Cao, M., Morse, A.S., Anderson, B.D.O.: Reaching a consensus in a dynamically changing environment - convergence rates, measurement delays and asynchronous events. SIAM J. Control Optim. **47**(2), 601–623 (2008)
9. Ceragioli, F., Frasca, P.: Continuous and discontinuous opinion dynamics with bounded confidence. Nonlinear Anal. Real World Appl. **13**(3), 1239–1251 (2012)
10. Cogburn, R.: On products of random stochastic matrices. In: Random Matrices and their Applications, vol. 50. Contemporary Mathematics, pp. 199–213. American Mathematical Society (1986)
11. De Persis, C., Frasca, P., Hendrickx, J.M.: Self-triggered rendezvous of gossiping second-order agents. In: IEEE Conference on Decision and Control, Florence, Italy, pp. 7403–7408 (2013)
12. Deffuant, G., Neau, D., Amblard, F., Weisbuch, G.: Mixing beliefs among interacting agents. Adv. Complex Syst. **3**(1–4), 87–98 (2000)
13. DeGroot, M.H.: Reaching a consensus. J. Am. Stat. Assoc. **69**(345), 118–121 (1974)
14. Dimakis, A.G., Kar, S., Moura, J.M.F., Rabbat, M.G., Scaglione, A.: Gossip algorithms for distributed signal processing. Proc. IEEE **98**(11), 1847–1864 (2010)
15. Etesami, S.R.: Potential-based analysis of social, communication, and distributed networks. Ph.D. thesis, University of Illinois at Urbana-Champaign (2015)
16. Etesami, S.R., Başar, T.: Game-theoretic analysis of the Hegselmann-Krause model for opinion dynamics in finite dimensions. IEEE Trans. Autom. Control **60**(7), 1886–1897 (2015)
17. Fagnani, F., Frasca, P.: Broadcast gossip averaging: interference and unbiasedness in large Abelian Cayley networks. IEEE J. Sel. Top. Signal Process. **5**(4), 866–875 (2011)
18. Fagnani, F., Zampieri, S.: Asymmetric randomized gossip algorithms for consensus. In: IFAC World Congress, Seoul, Korea, pp. 9052–9056, July 2008
19. Fagnani, F., Zampieri, S.: Randomized consensus algorithms over large scale networks. IEEE J. Sel. Areas Commun. **26**(4), 634–649 (2008)
20. Fagnani, F., Zampieri, S.: Average consensus with packet drop communication. SIAM J. Control Optim. **48**(1), 102–133 (2009)
21. Frasca, P., Hendrickx, J.M.: Large network consensus is robust to packet losses and interferences. In: European Control Conference, Zürich, Switzerland, pp. 1782–1787, July 2013
22. Frasca, P., Hendrickx, J.M.: On the mean square error of randomized averaging algorithms. Automatica **49**(8), 2496–2501 (2013)
23. Hegselmann, R., Krause, U.: Opinion dynamics and bounded confidence: models, analysis and simulation. J. Artif. Soc. Soc. Simul. **5**(3), 1–33 (2002)
24. Hendrickx, J.M.: Graphs and Networks for the Analysis of Autonomous Agent Systems. Ph.D. thesis, Université catholique de Louvain (2008)
25. Hendrickx, J.M., Tsitsiklis, J.N.: Convergence of type-symmetric and cut-balanced consensus seeking systems (extended version). http://arxiv.org/abs/1102.2361 (2011)
26. Hendrickx, J.M., Tsitsiklis, J.N.: A new condition for convergence in continuous-time consensus seeking systems. In: IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, USA, pp. 5070–5075 (2011)
27. Hendrickx, J.M., Tsitsiklis, J.N.: Convergence of type-symmetric and cut-balanced consensus seeking systems. IEEE Trans. Autom. Control **58**(1), 214–218 (2013)
28. Jadbabaie, A., Lin, J., Morse, A.S.: Coordination of groups of mobile autonomous agents using nearest neighbor rules. IEEE Trans. Autom. Control **48**(6), 988–1001 (2003)
29. Krause, U.: A discrete nonlinear and non-autonomous model of consensus formation. In: Elaydi, S., Ladas, G., Popenda, J., Rakowski, J. (eds.) Communications in Difference Equations. Gordon & Breach, pp. 227–236 (2000)

30. Liu, B., Lu, W., Chen, T.: Consensus in networks of multiagents with switching topologies modeled as adapted stochastic processes. SIAM J. Control Optim. **49**(1), 227–253 (2011)
31. Lorenz, J.: A stabilization theorem for continuous opinion dynamics. Phys. A **355**(1), 217–223 (2005)
32. Lorenz, J.: Continuous opinion dynamics under bounded confidence: A survey. Int. J. Mod. Phys. C **18**(12), 1819–1838 (2007)
33. Martin, S., Hendrickx, J.M.: Continuous-time consensus under non-instantaneous reciprocity. IEEE Trans. Autom. Control **61**(9), 2484–2495 (2016)
34. Mirtabatabaei, A., Bullo, F.: Opinion dynamics in heterogeneous networks: convergence conjectures and theorems. SIAM J Control Optim. **50**(5), 2763–2785 (2012)
35. Moreau, L.: Stability of multiagent systems with time-dependent communication links. IEEE Trans. Autom. Control **50**(2), 169–182 (2005)
36. Nedich, A., Liu, J.: Lyapunov approach to consensus problems (2014)
37. Olshevsky, A., Tsitsiklis, J.N.: Convergence speed in distributed consensus and averaging. SIAM J. Control Optim. **48**(1), 33–55 (2009)
38. Roozbehani, M., Megretski, A., Frazzoli, E.: Lyapunov analysis of quadratically symmetric neighborhood consensus algorithms. In: IEEE Conference on Decision and Control, Cancún, México, pp. 2252–2257 (2008)
39. Seneta, E.: Non-negative Matrices and Markov Chains. Springer Series in Statistics. Springer (2006)
40. Tahbaz-Salehi, A., Jadbabaie, A.: A necessary and sufficient condition for consensus over random networks. IEEE Trans. Autom. Control **53**(3), 791–795 (2008)
41. Tahbaz-Salehi, A., Jadbabaie, A.: Consensus over ergodic stationary graph processes. IEEE Trans. Autom. Control **55**(1), 225–230 (2010)
42. Touri, B.: Product of Random Stochastic Matrices and Distributed Averaging. Springer Theses. Springer, March 2012
43. Touri, B., Langbort, C.: On endogenous random consensus and averaging dynamics. IEEE Trans. Autom. Control **1**(3), 241–248 (2014)
44. Touri, B., Nedic, A.: Product of random stochastic matrices. IEEE Trans. Autom. Control **59**(2), 437–448 (2014)
45. Tsitsiklis, J.N.: Problems in Decentralized Decision Making and Computation. Ph.D. thesis, Massachusetts Institute of Technology, November 1984
46. Wang, C., Li, Q., Weinan, E., Chazelle, B.: Noisy Hegselmann-Krause systems: phase transition and the 2r-conjecture. J. Stat. Phys. **166**(5), 1209–1225 (2017)
47. Zhang, J., Hong, Y.: Opinion evolution analysis for short-range and long-range Deffuant-Weisbuch models. Phys. A **392**(21), 5289–5297 (2013)