(c) Show that the sequence $\{x(t)\}$ generated by the asynchronous iteration

$$x_i(t+1) = (1-\gamma)x_i\big(\tau_i^i(t)\big) + \gamma h_i\big(x^i(t)\big), \qquad t \in T^i,$$

is guaranteed to be bounded, but does not necessarily converge to a fixed point of h. *Hints:* For part (a), only a minor modification of the proof of Prop. 2.3 is needed. For part (b), use the function $h(x) = -x$ as an example. For part (c), use the example $h(x) = Ax$, where

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

and apply Prop. 3.1 of Section 6.3.

2.2. Let A be a nonnegative and irreducible matrix of dimensions $n \times n$, with the property $\sum_{j=1}^{n} a_{ij} \leq 1$. Suppose that all of the diagonal entries of A are positive. Show that under Assumption 1.1, the asynchronous iteration $x := Ax$ converges to a vector x^* satisfying $Ax^* = x^*$. *Hint:* Let δ be the smallest diagonal entry of A. The iteration $x := Ax$ can be written as [cf. Eq. (2.6)]

$$x := x - (1-\delta)\frac{I-A}{1-\delta}x.$$

Show that the matrix $(I-A)/(1-\delta)$ satisfies Assumption 2.3.

7.3 ALGORITHMS FOR AGREEMENT AND FOR MARKOV CHAIN PROBLEMS

We now consider a set of processors that try to reach agreement on a common scalar value by exchanging tentative values and combining them by forming convex combinations. Although this algorithm is of limited use on its own, it has interesting applications in certain contexts, such as the computation of invariant distributions of Markov chains (Subsection 7.3.2) and an extended model of asynchronous computation (Section 7.7).

7.3.1 The Agreement Algorithm

Consider a set $N = \{1, \ldots, n\}$ of processors, and suppose that the ith processor has a scalar $x_i(0)$ stored in its memory. We would like these processors to exchange messages and eventually agree on an intermediate value y, that is, the agreed upon value should satisfy

$$\min_{i \in N} x_i(0) \leq y \leq \max_{i \in N} x_i(0).$$

A trivial solution to this problem is to have a particular processor (say processor 1) communicate (directly or indirectly) its own value to all other processors and then all processors can agree on the value $x_1(0)$. We shall impose an additional requirement that excludes such a solution. In particular, we postulate the existence of a set $D \subset N$ of *distinguished* processors and we are interested in guaranteeing that the values initially possessed by distinguished processors influence the agreed upon value y. For example, if $k \in D$, $x_k(0) > 0$, and $x_i(0) = 0$ for all $i \neq k$, then we would like y to be influenced by $x_k(0)$ and be positive. Such a requirement can be met if the processors simply cooperate to compute the average of their initial values. This, however, may require a certain amount of coordination between the processors. We shall instead postulate an asynchronous iterative process whereby each processor receives tentative values from other processors and combines them with its own value by forming a convex combination. We let $x_i(t)$ be the value in the memory of the ith processor at time t, and we consider the asynchronous execution of the iteration

$$x_i := \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, \ldots, n,$$

where the coefficients a_{ij} are nonnegative scalars such that

$$\sum_{j=1}^{n} a_{ij} = 1, \qquad \forall i. \tag{3.1}$$

A precise description of the algorithm, referred to as the *agreement algorithm*, is

$$x_i(t+1) = x_i(t), \qquad \text{if } t \notin T^i, \tag{3.2}$$

$$x_i(t+1) = \sum_{j=1}^{n} a_{ij} x_j\big(\tau_j^i(t)\big), \qquad \text{if } t \in T^i. \tag{3.3}$$

Here T^i and $\tau_j^i(t)$ are as in Section 7.1 and will be assumed to satisfy the partial asynchronism Assumption 1.1. In the present context, it would be natural to assume that $0 \leq \tau_j^i(t)$. It is convenient, however, to consider a more general case, allowing $\tau_j^i(t)$ to be negative (as long as Assumption 1.1 is satisfied), and allowing $x_i(t)$, for $t < 0$, to be different from $x_i(0)$.

Let A be the matrix whose ijth entry is equal to a_{ij}. We are then dealing with the special case of the model of Section 7.1, where the mapping f is of the form

$$f(x) = Ax.$$

Notice that any vector $x \in \Re^n$ whose components are all equal is a fixed point of f, because of the condition $\sum_{j=1}^{n} a_{ij} = 1$ for all i. In the sequel, we derive conditions under which the sequence $\{x(t)\}$ of the vectors generated by the partially asynchronous agreement algorithm of Eqs. (3.2) and (3.3) converges to such a fixed point. A result of this type is readily obtained if the matrix A is irreducible, a relaxation parameter $\gamma \in (0,1)$ is employed, and the iteration $x := Ax$ is modified to $x := (1-\gamma)x + \gamma Ax = x - \gamma(I - A)x$. This is because the matrix $I - A$ satisfies Assumption 2.3 of Section

7.2 and Prop. 2.4 in that section applies. Even if no relaxation parameter is used, Prop. 2.4 can be again invoked as long as all of the diagonal entries of A are positive (see Exercise 2.2 in the preceding section). The result derived in this section is more general in that it allows most of the diagonal entries of A to be zero. Let us also mention that the agreement problem can be related to a network flow problem with quadratic costs (see Exercise 5.4 in Section 5.5.) We now consider some examples to motivate our assumptions.

Example 3.1.

Suppose that

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Here the set of fixed points of f is $X^* = \{(x_1, x_2) \mid x_1 = x_2\}$, but the iteration $x := f(x)$ is not guaranteed to converge to X^*, even if it is executed synchronously. To see this, notice that if the synchronous execution is initialized with $x(0) = (1, 0)$, then $x(t)$ alternates between $(0, 1)$ and $(1, 0)$ (Fig. 7.3.1). The possibility of such nonconvergent oscillations will be eliminated by assuming that some diagonal entry of the matrix A is nonzero. Such an entry has the effect of a relaxation parameter and serves as a damping factor.

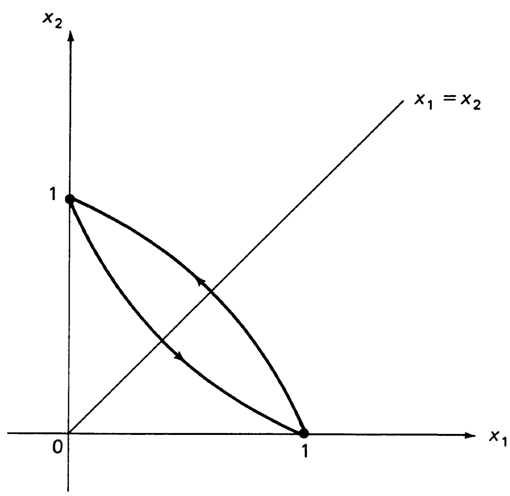


Figure 7.3.1 Nonconvergence in Example 3.1. Here when $x(t)$ is updated synchronously, it alternates between $(1, 0)$ and $(0, 1)$. The oscillation can be eliminated by introducing a relaxation parameter $\gamma \in (0, 1)$ thereby replacing the iteration $x_1(t+1) = x_2(t)$ and $x_2(t+1) = x_1(t)$ with $x_1(t+1) = (1-\gamma)x_1(t) + \gamma x_2(t)$ and $x_2(t+1) = (1-\gamma)x_2(t) + \gamma x_1(t)$. This amounts to introducing positive diagonal elements in the iteration matrix A.

To justify the partial asynchronism assumption, consider the matrix

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

The agreement algorithm for this choice of A is the same as the iteration considered in Example 1.2 of Section 7.1, where it was established that failure to converge is possible if part (b) of the partial asynchronism Assumption 1.1 is violated. There also exist examples that demonstrate that parts (a) and (c) of Assumption 1.1 are also necessary for convergence (Exercise 3.1).

We define a directed graph $G = (N, A)$, where $N = \{1, \ldots, n\}$, and $A = \{(i,j) \mid i \neq j \text{ and } a_{ji} \neq 0\}$. Notice that $(i,j) \in A$ if and only if the value possessed by processor i directly influences the value of processor j. Convergence will be proved under the following assumption on the matrix A.

Assumption 3.1. There exists a nonempty set $D \subset N$ of "distinguished" processors such that:

(a) For every $i \in D$, we have $a_{ii} > 0$.

(b) For every $i \in D$ and every $j \in N$, there exists a positive path from i to j in the previously defined graph G.

This assumption is quite natural. Since we wish the initial values of any distinguished processor to affect the value that is eventually agreed upon, we have imposed the condition that every distinguished processor can indirectly affect the value of every other processor [part (b)]. Part (a) of the assumption ensures that a distinguished processor does not forget its initial value when it executes its first iteration; it also serves to eliminate nonconvergent oscillations (cf. Example 3.1).

Proposition 3.1. Consider the agreement algorithm of Eqs. (3.2) and (3.3) and let Assumptions 1.1 (partial asynchronism) and 3.1 hold. Let $\alpha > 0$ be the smallest of the nonzero entries of A. Then there exist constants $\eta > 0$, $C > 0$, $\rho \in (0,1)$, depending only on the number n of processors, on α, and on the asynchronism measure B of Assumption 1.1, such that for any initial values $x_i(t)$, $t \leq 0$, and for any scenario allowed by Assumption 1.1, the following are true:

(a) The sequence $\{x_i(t)\}$ converges and its limit is the same for each processor i.

(b) There holds

$$\max_i x_i(t) - \min_i x_i(t) \leq C\rho^t \Big(\max_i \max_{-B+1 \leq \tau \leq 0} x_i(\tau) - \min_i \min_{-B+1 \leq \tau \leq 0} x_i(\tau) \Big).$$

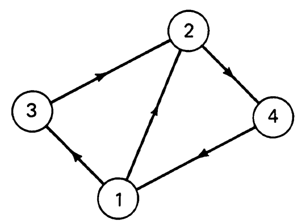
(c) If $0 \leq x_i(\tau)$ for every i and every $\tau \leq 0$, and if $k \in D$, then $y \geq \eta x_k(0)$, where y is the common limit whose existence is asserted in part (a).

Proof. We define

$$M(t) = \max_i \max_{t-B+1 \leq \tau \leq t} x_i(\tau), \tag{3.4}$$

$$m(t) = \min_i \min_{t-B+1 \leq \tau \leq t} x_i(\tau). \tag{3.5}$$

Notice that because of Assumption 1.1, we have $m(t) \leq x_j\left(\tau_j^i(t)\right) \leq M(t)$ for every i, j, $t \in T^i$, a fact that will be often used. The proof consists of showing that the difference $M(t) - m(t)$ is reduced to zero in the course of the algorithm (see Fig. 7.3.2 for an illustration of the main idea of the proof).



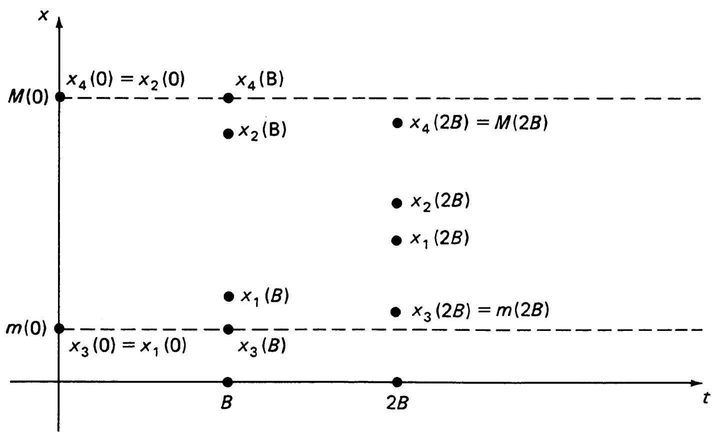


Figure 7.3.2 (a) A possible graph G associated to a matrix A in the agreement algorithm. (b) Illustration of the convergence of the agreement algorithm for the graph of part (a). For simplicity, we assume that $a_{11} > 0$, $a_{44} > 0$, and that information is never outdated. Let the initial conditions be as indicated and let I be the interval $[m(0), M(0)]$. Since $a_{11} > 0$ and $a_{44} > 0$, we have $x_1(t) < M(0)$ and $x_4(t) > m(0)$ for all times t. Within the first B time units, processor 2 performs an iteration and x_2 is pulled by x_1 to a value smaller than $M(0)$. Similarly, x_1 is pulled by x_4 to a value larger than $m(0)$. After an additional B time units, the variables x_3 and x_4 are pulled by x_1 and x_2, respectively, into the interior of I. At that point, all the components of $x(2B)$ lie in the interior of I and the maximum disagreement $M(2B) - m(2B)$ is smaller than the initial maximum disagreement $M(0) - m(0)$.

Lemma 3.1.

(a) For every $t \geq 0$, we have $m(t+1) \geq m(t)$ and $M(t+1) \leq M(t)$.

(b) For every t and $t' \geq t - B + 1$, we have $m(t) \leq x_i(t') \leq M(t)$.

Proof. Fix some i and t. If $t \notin T^i$, then $x_i(t+1) = x_i(t) \geq m(t)$. If $t \in T^i$, then

$$x_i(t+1) = \sum_{j=1}^{n} a_{ij} x_j\big(\tau_j^i(t)\big) \geq \sum_{j=1}^{n} a_{ij} m(t) = m(t).$$

Thus, in either case, $x_i(t+1) \geq m(t)$ for all i, and, using the definition of $m(t+1)$, it is seen that $m(t+1) \geq m(t)$. The proof of the inequality $M(t+1) \leq M(t)$ is similar. Finally, for $t' \geq t - B + 1$, we use the definition of $m(t' + B - 1)$ and part (a) of the lemma to obtain $x_i(t') \geq m(t' + B - 1) \geq m(t)$. The inequality $x_i(t') \leq M(t)$ is proved similarly. **Q.E.D.**

For the remainder of the proof of the proposition, we fix some $k \in D$. We let $D_0 = \{k\}$ and we let D_ℓ be the set of all $i \in N$ such that ℓ is the minimum number of arcs in a positive path, in the graph G, from node k to node i. By Assumption 3.1(b), there exists a path from k to every other processor. It follows that every $i \in N$ belongs to one of the sets $D_0, D_1, \ldots, D_{n-1}$. Furthermore, for every $i \in D_\ell$, there exists some $j \in D_{\ell-1}$, such that $(j, i) \in A$. Let $L \leq n-1$ be such that $D_0 \cup \cdots \cup D_L = N$.

Lemma 3.2. For every $\ell \in \{0, 1, \ldots, L\}$, there exists some $\eta_\ell > 0$ (depending only on n, α, B) such that for every positive integer s, for every $t \in [s + 2\ell B + 1, s + 2LB + B]$, and for every $i \in D_\ell$, we have

$$x_i(t) \geq m(s) + \eta_\ell\big(x_k(s) - m(s)\big) \tag{3.6}$$

and

$$x_i(t) \leq M(s) - \eta_\ell\big(M(s) - x_k(s)\big). \tag{3.7}$$

Proof. Throughout the proof of this lemma, $k \in D$ is fixed. Without loss of generality, we only consider the case where $s = 0$. Suppose that $t \in T^k$. Then

$$x_k(t+1) - m(0) = \sum_{j=1}^{n} a_{kj}\Big(x_j\big(\tau_j^k(t)\big) - m(0)\Big) \geq a_{kk}\big(x_k(t) - m(0)\big) \geq \alpha\big(x_k(t) - m(0)\big),$$

where we made use of the property $\tau_k^k(t) = t$ [cf. Assumption 1.1(c)]. If $t \notin T^k$, then $x_k(t+1) - m(0) = x_k(t) - m(0) \geq \alpha\big(x_k(t) - m(0)\big)$. It follows that for $t \in [0, 2LB + B]$, we have

$$x_k(t) - m(0) \geq \alpha^t\big(x_k(0) - m(0)\big) \geq \eta_0\big(x_k(0) - m(0)\big),$$

where $\eta_0 = \alpha^{2LB+B}$. This proves inequality (3.6) for $i = k$. Since $D_0 = \{k\}$, inequality (3.6) has been proved for all $i \in D_0$.

We now proceed by induction on ℓ. Suppose that inequality (3.6) is true for some $\ell < L$. Let i be an element of $D_{\ell+1}$. We shall prove inequality (3.6) for i.

Let $j \in D_\ell$ be such that $(j,i) \in A$. Suppose now that t belongs to T^i and satisfies $(2\ell+1)B \le t \le 2LB+B$. We then have $2\ell B + 1 \le \tau_j^i(t) \le t \le 2LB+B$ and, by the induction hypothesis,

$$x_j\big(\tau_j^i(t)\big) - m(0) \ge \eta_\ell\big(x_k(0) - m(0)\big).$$

Consequently,

$$\begin{aligned} x_i(t+1) - m(0) &= \sum_{q=1}^{n} a_{iq}\Big(x_q\big(\tau_q^i(t)\big) - m(0)\Big) \ge a_{ij}\Big(x_j\big(\tau_j^i(t)\big) - m(0)\Big) \\ &\ge \alpha\eta_\ell\big(x_k(0) - m(0)\big) = \eta_{\ell+1}\big(x_k(0) - m(0)\big), \end{aligned} \tag{3.8}$$

where $\eta_{\ell+1} = \alpha\eta_\ell$. Let t_i be an element of T^i such that $(2\ell+1)B \le t_i \le 2(\ell+1)B$. Such a t_i exists because of Assumption 1.1(a). Inequality (3.8) has been proved for $t = t_i$, as well for any subsequent $t \in T^i$ such that $t \le 2LB + B$. Furthermore, since $x_i(t) = x_i(t+1)$, if $t \notin T^i$, we conclude that inequality (3.8) holds for all t such that $t_i \le t \le (2L+1)B$. Since $t_i \le 2(\ell+1)B$, we conclude that (3.8) holds for every t such that $2(\ell+1)B \le t \le 2LB + B$. This establishes inequality (3.6) for $i \in D_{\ell+1}$ and for $s = 0$. This completes the induction and the proof of (3.6) for the case $s = 0$. The proof for the case of a general s is identical. Finally, inequality (3.7) is proved by a symmetrical argument. **Q.E.D.**

Proof of Proposition 3.1. (cont.) We now complete the proof of the proposition. We have $m(t) \le M(t) \le M(0)$. Furthermore, the sequence $\{m(t)\}$ is nondecreasing. Since it is bounded above, it converges to a limit denoted by $\bar{m}$. Let $\bar{M}$ be the limit of $M(t)$, which exists by a similar argument. Let $\eta = \min\{\eta_0, \ldots, \eta_L\}$. Using Lemma 3.2 we obtain, for every $t \ge 0$,

$$\begin{aligned} m(t+2LB+B) &= \min_\ell \min_{i \in D_\ell} \min_{t+2LB+1 \le \tau \le t+2LB+B} x_i(\tau) \ge m(t) + \min_\ell \eta_\ell\big(x_k(t) - m(t)\big) \\ &= m(t) + \eta\big(x_k(t) - m(t)\big). \end{aligned}$$

Similarly,

$$M(t+2LB+B) \le M(t) + \eta\big(x_k(t) - M(t)\big).$$

Subtracting these two inequalities, we obtain

$$M(t+2LB+B) - m(t+2LB+B) \le (1-\eta)\big(M(t) - m(t)\big). \tag{3.9}$$

Thus, $M(t) - m(t)$ decreases at the rate of a geometric progression and $\bar{M} = \bar{m}$. Let y be the common value of $\bar{M}$ and $\bar{m}$. Since $m(t) \le x_i(t) \le M(t)$ for every i and

t, it follows that the sequence $\{x_i(t)\}$ also converges to y at the rate of a geometric progression. This establishes parts (a) and (b) of the proposition.

We now prove part (c). We have $m(0) \geq 0$, and using Lemma 3.2, we obtain $y \geq m(2LB + B) \geq \eta x_k(0)$. **Q.E.D.**

It should be emphasized that the value y on which agreement is reached usually depends on the particular scenario.

7.3.2 An Asynchronous Algorithm for the Invariant Distribution of a Markov Chain

Notice that the iteration matrix A in the agreement algorithm was a stochastic matrix. This suggests some similarities between the agreement algorithm and the iterative algorithms of Section 2.8 for computing the invariant distribution of a Markov chain. In this subsection, we let P be an irreducible stochastic matrix and we establish partially asynchronous convergence of the iteration $\pi := \pi P$ by suitably exploiting the convergence result for the agreement algorithm. It should be recalled that the totally asynchronous version of this iteration converges if one of the components of π is not iterated (Subsection 6.3.1). If all of the components of π are iterated, then totally asynchronous convergence is not guaranteed (this can be seen from either Example 3.1 of this section or Example 1.2 of Section 7.1) and, therefore, the partial asynchronism assumption is essential.

Let P be an irreducible and aperiodic stochastic matrix of dimensions $n \times n$ and let p_{ij} denote its ijth entry. Let π^* be the row vector of invariant probabilities of the corresponding Markov chain. Proposition 8.3 of Section 2.8 states that each component π_i^* of π^* is positive and $\lim_{t\to\infty} \pi(0)P^t = \pi^*$ for any row vector $\pi(0)$ whose entries add to 1. This leads to the iterative algorithm $\pi := \pi P$ of Section 2.8 and the corresponding synchronous parallel implementation. We now consider its asynchronous version.

We employ again the model of Section 7.1, except that the vector being iterated is denoted by π and is a row vector. The iteration function f is defined by $f(\pi) = \pi P$. The iteration is described by the equations

$$\pi_i(t+1) = \pi_i(t), \qquad t \notin T^i, \tag{3.10}$$

$$\pi_i(t+1) = \sum_{j=1}^{n} \pi_j\big(\tau_j^i(t)\big)p_{ji}, \qquad t \in T^i. \tag{3.11}$$

Proposition 3.2. Suppose that the matrix P is stochastic, irreducible, and that there exists some i^* such that $p_{i^*i^*} > 0$. Furthermore, suppose that the iteration (3.10)–(3.11) is initialized with positive values [$\pi_i(\tau) > 0$ for $\tau \leq 0$]. Then for every scenario allowed by Assumption 1.1, there exists a positive number c such that $\lim_{t\to\infty} \pi(t) = c\pi^*$. Also, convergence takes place at the rate of a geometric progression.

Proof. We prove this result by showing that it is a special case of the convergence result for the agreement algorithm (Prop. 3.1). We introduce a new set of variables $x_i(t)$ defined by

$$x_i(t) = \frac{\pi_i(t)}{\pi_i^*}.$$

These new variables are well defined because $\pi_i^* > 0$ for all i as a consequence of irreducibility. In terms of the new variables, Eqs. (3.10) and (3.11) become

$$x_i(t+1) = x_i(t), \qquad t \notin T^i, \tag{3.12}$$

$$x_i(t+1) = \sum_{j=1}^{n} \frac{p_{ji}\pi_j^*}{\pi_i^*} x_j\big(\tau_j^i(t)\big), \quad t \in T^i. \tag{3.13}$$

By letting

$$a_{ij} = \frac{p_{ji}\pi_j^*}{\pi_i^*}, \tag{3.14}$$

Eq. (3.13) becomes

$$x_i(t+1) = \sum_{j=1}^{n} a_{ij} x_j\big(\tau_j^i(t)\big), \qquad t \in T^i,$$

which is identical to Eq. (3.3) in the agreement algorithm. Furthermore, notice that $a_{ij} \geq 0$ and that

$$\sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} \frac{p_{ji}\pi_j^*}{\pi_i^*} = \frac{1}{\pi_i^*}\sum_{j=1}^{n} p_{ji}\pi_j^* = \frac{\pi_i^*}{\pi_i^*} = 1,$$

where we have used the property $\pi^* = \pi^* P$. Thus, Eq. (3.1) holds as well.

We now verify that the remaining assumptions in Prop. 3.1 are satisfied. Let i^* be such that $p_{i^*i^*} > 0$, and let $D = \{i^*\}$. Then $a_{i^*i^*} > 0$ and Assumption 3.1(a) holds. Also, since P is irreducible, the coefficients a_{ij} satisfy Assumption 3.1(b).

Therefore, Prop. 3.1 applies and shows that there exists a constant c such that

$$\lim_{t\to\infty} x_i(t) = c, \qquad \forall i.$$

Equivalently,

$$\lim_{t\to\infty} \pi_i(t) = c\pi_i^*, \qquad \forall i.$$

Geometric convergence follows from part (b) of Prop. 3.1.

Let $m(0) = \min_i \min_{-B+1 \leq \tau \leq 0} x_i(\tau)$. Since the algorithm is initialized with positive values, $m(0)$ is positive. As shown in the convergence proof for the agreement algorithm, the limit c of $x_i(t)$ is no smaller than $m(0)$. Thus, c is positive. **Q.E.D.**

As in the agreement algorithm, the constant c whose existence is asserted by Prop. 3.2 depends on the particular scenario. This does not cause any difficulties because π^* can be recovered from the limiting value of $\pi(t)$ by normalizing it so that its entries add to 1.

If P is stochastic and irreducible, but all of its diagonal entries are zero, the iteration of Eqs. (3.10) and (3.11) does not converge, in general (see Example 3.1). However, we can let

$$Q = \gamma P + (1-\gamma)I,$$

and apply the algorithm with P replaced by Q. Here γ is a constant belonging to (0, 1), and I is the identity matrix. Then Q satisfies the assumptions of Prop. 3.2, and we obtain convergence to a multiple of the invariant distribution vector of Q. This is all we need because it is easily seen that P and Q have the same invariant distribution.

As a final extension, suppose that P is nonnegative, irreducible, with $\rho(P) = 1$, but not necessarily stochastic. Then the Perron–Frobenius theorem (Prop. 6.6 in Section 2.6), applied to the transpose of P, guarantees the existence of a positive row vector π^* such that $\pi^* = \pi^* P$, and the proof of Prop. 3.2 remains valid without any modifications whatsoever.

EXERCISES

3.1. **(a)** Suppose that Assumption 1.1(a) is replaced by the requirement that T^i is infinite, for all i. Show that Props. 3.1 and 3.2 are no longer true.

(b) Suppose that Assumption 1.1(c) is replaced by the requirement $t - B + 1 \leq \tau_i^i(t) \leq t$ for all i and $t \in T^i$. Show that Props. 3.1 and 3.2 are no longer true.

Hints: For part (a), let

$$A = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

and show that if each processor in turn executes a large number of iterations, the algorithm behaves similarly with the iteration $x_1 := x_3$, $x_3 := x_2$, and $x_2 := x_1$ executed in Gauss–Seidel fashion. For part (b), let

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Arrange the initial conditions and a scenario so that for every t, we have $x(t) = x(t+2)$, but $x(t+1) \neq x(t)$.

3.2. We consider a variant of the agreement algorithm whereby processors receive messages from other processors, and upon reception, these messages are immediately taken into account by forming convex combinations. Let $G = (N, A)$ be a directed graph, with $N = \{1, \ldots, n\}$ and with $(j, i) \in A$ if and only if processor j communicates to processor i. For any $(j, i) \in A$, let T_j^i be the set of times that processor i receives a message $x_j\big(\tau_j^i(t)\big)$ from processor j. We assume that for any fixed i, the sets T_j^i, $j \neq i$, are disjoint. Let the algorithm be described by the equations

$$x_i(t+1) = a_{ij}x_j\big(\tau_j^i(t)\big) + (1 - a_{ij})x_i(t), \qquad t \in T_j^i,$$

and $x_i(t+1) = x_i(t)$ if t does not belong to any T_j^i. Assume that $0 < a_{ij} < 1$ for every i and j such that $(j, i) \in A$.

(a) Reformulate appropriately Assumptions 1.1 and 3.1, redefine the constant α of Prop. 3.1, and then show that the conclusions of Prop. 3.1 hold for the present algorithm as well.

(b) What happens in this algorithm if one of the processors breaks down and stops transmitting any messages?

(c) Answer the question of part (b) for the original algorithm of Eqs. (3.2) and (3.3).

3.3. For each i we let $\{\epsilon_i(t)\}$ be a sequence of real numbers that converges geometrically to zero. We consider a perturbed version of the agreement algorithm, whereby Eq. (3.3) is replaced by

$$x_i(t+1) = \sum_{j=1}^{n} a_{ij}x_j\big(\tau_j^i(t)\big) + \epsilon_i(t), \qquad \text{if } t \in T^i.$$

Let Assumptions 1.1 and 3.1 hold, and show that for every scenario and each i, the sequence $\{x_i(t)\}$ converges geometrically to a limit independent of i. *Hint:* Fix some positive integer s. For any scenario, define $v(t)$ by letting $v(t) = x(t)$ if $t \leq s$,

$$v_i(t+1) = \sum_{j=1}^{n} a_{ij}v_j\big(\tau_j^i(t)\big), \qquad \text{if } t \geq s,\ t \in T^i,$$

and $v_i(t+1) = v(t)$ if $t \geq s$ and $t \notin T^i$. Let $q(t) = \min_i \min_{t-B+1 \leq \tau \leq t} v_i(\tau)$ and $Q(t) = \max_i \max_{t-B+1 \leq \tau \leq t} v_i(\tau)$. From Prop. 3.1 we have $Q(s+2LB+B) - q(s+2LB+B) \leq \eta\big(Q(s) - q(s)\big)$. Furthermore, $x_i(s+2LB+B) - v_i(s+2LB+B)$ can be bounded by a constant which tends to zero geometrically as s goes to infinity. Combine these two observations to show that $M(s+2B+B) - m(s+2B+B) \leq (1-\eta)\big(M(s) - m(s)\big) + \delta(s)$, where $\{\delta(s)\}$ is a sequence that converges to zero geometrically.

3.4. Let all of the assumptions in Prop. 3.2 hold except for the irreducibility of P. Assume instead that the Markov chain corresponding to P has a single ergodic class and that the nonzero diagonal entry $p_{i^*i^*}$ corresponds to a recurrent state i^*. Show that the sequence $\{\pi(t)\}$ generated by the partially asynchronous iteration (3.10)–(3.11), initialized with positive values, converges geometrically to a positive multiple of the vector of invariant probabilities of the Markov chain. *Hint:* We are dealing with the asynchronous iterations $\pi^{(1)} := \pi^{(1)}P_{11}$ and $\pi^{(2)} := \pi^{(1)}P_{12} + \pi^{(2)}P_{22}$, where $\pi^{(1)}$ and $\pi^{(2)}$ are appropriate subvectors of π, and P_{11}, P_{12} and P_{22} are appropriate submatrices of P. Show that $\rho(P_{11}) < 1$ and that $\pi^{(1)}(t)$ converges to zero geometrically. For the second iteration, use a suitable change of variables and the result of Exercise 3.3.