Comment on "Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules" *

Dimitri P. Bertsekas
John N. Tsitsiklis
Massachusetts Institute of Technology
e-mail: {dimitrib,jnt}@mit.edu

June 14, 2006

Abstract

We clarify the relation of the model and the convergence results of Jadbabaie et al. [3] to those studied by Bertsekas et al. [6, 5, 1]. We show that the update equations in [3] are a very special case of those in [5]. Furthermore, the main convergence results in [3] are special cases of those in [5], except for a small difference in the connectivity assumptions which, however, does not affect the proof.

*Research supported in part by the NSF under contract ECS-0312921.

1 Introduction

In the 1980s, references [6, 5, 1] studied various models of distributed asynchronous iterations, motivated by the contexts of parallel computation, distributed optimization, and distributed signal processing. An important "subroutine" in this context was the "agreement algorithm" (see Secs. 7.3 and 7.7 of [1]), whereby a set of agents reach consensus on a common value by forming convex combinations of their current values with possibly outdated values possessed by their neighbors. A comprehensive convergence analysis was provided at the time, including examples delineating how convergence to consensus might fail in the absence of certain assumptions on the communication delays, or on the time between consecutive processor communications (see, e.g., Example 1.2, in p. 485 of [1], and Exercise 3.1, in p. 517 of [1]).

Historically, the idea of reaching consensus through repeated averaging was introduced earlier by De Groot [2], for a structured, time-invariant, and synchronous environment. In the absence of communication delays, convergence conditions for the time-varying case were provided by Chatterjee and Seneta [4], although without making the connection with the primitives of the processors' behavior, such as communication frequency, etc.

In recent years, motivated by various contexts such as multiagent coordination and flocking, a large number of papers have been published that propose and analyze consensus algorithms that are closely related to (and oftentimes, special cases of) the agreement algorithm. As the relation with [6, 5, 1] has not been sufficiently elucidated in the recent literature, we think it is important to clarify it.

Since this is only a "note" (rather than a survey paper), we will not attempt to cover the recent literature, and we will limit ourselves to a discussion of the paper [3], which received the 2005 George S. Axelby Outstanding Paper Award from the IEEE Control Systems Society. The information contained in this note should be sufficient for understanding the relation with other recent papers.

2 The Model in [5]

Consider a set $N = \{1, \ldots, n\}$ of agents that try to reach agreement on a common scalar value by exchanging tentative values, and combining them by forming convex combinations. In particular, each agent i starts with a scalar value $x_i(0)$. The vector $x(t) = (x_1(t), \ldots, x_n(t))$, with the values held by the agents at time t, is updated according to the equation $x(t+1) = A(t)x(t)$, where $A(t)$ is a stochastic matrix with entries $a_{ij}(t)$. In more detail, let T_{ij} be the set of times at which agent i receives a message from agent j, containing the value of $x_j(t)$, with the convention that $t \in T_{ii}$ for every i and t. The update equation is of the form

$$x_i(t+1) = \sum_{j=1}^{n} a_{ij}(t)x_j(t), \tag{1}$$

where the coefficients $a_{ij}(t)$ satisfy:

(i) $a_{ij}(t) \geq 0, \; \forall \, i, \; j, \; t;$

(ii) $\sum_{j=1}^{n} a_{ij}(t) = 1, \; \forall \, i, \; t;$

(iii) $a_{ij}(t) = 0, \; \forall \, t \notin T_{ij}, \; i \neq j.$

The update equation (1) corresponds to a simplified version of Eqs. (2.1)-(2.4) in p. 804 of [5]. The model in [5] is actually more general: it allows for each $x_i(0)$ to be a vector, allows messages to incur delays before reaching their destination, and also allows for an additional exogenous input in the right-hand side of (1).

Let D be a subset of the set of agents, interpreted as a set of i for which $x_i(0)$ will have a long-term affect on the limit of every $x_j(t)$. (These are called "computing processors" in [5] and "distinguished processors" in [1].) Finally, let $E(t)$ be the set of ordered pairs (j, i) such that $a_{ij}(t) > 0$, so that an "edge" (j, i) indicates a communication from j to i. Thus, $(N, E(t))$ is a directed graph indicating the influences between agents at time t (the "communication graph"). Let E be the set of (i, j) such that $(i, j) \in E(t)$ for infinitely many t.

Assumption 1:

(a) The set D is nonempty.

(b) There is some $\alpha > 0$ such that if agent i is influenced by agent j at time t [i.e., $(j, i) \in E(t)$], then $a_{ij}(t) \geq \alpha$.

(c) There exists some B such that for every t, we have $E(t+1) \cup \cdots \cup E(t+B) = E$.

(d) The graph (N, E) contains a directed path from every $i \in D$ to every $j \in N$. (In particular if $D = N$, then (N, E) is strongly connected.)

Assumption 1 above is a special case of Assumptions 2.1, 2.2, and 2.4 in [5]. The assumptions in [5] are actually weaker: they allow for communication delays, and also allow $a_{ii}(t)$ to be zero for certain processors. Part (c) of Assumption 1 is a rephrasing of the connectivity assumptions made in [5], which had been stated as follows: (i) for every i, j, the set T_{ij} is empty or infinite; (ii) If $(j, i) \in E$, then the time between consecutive transmissions from j to i is upper bounded by some B.

We now present a convergence result from [5].

Theorem 1: There exist nonnegative coefficients $\phi_1, \ldots, \phi_n$ such that

$$\lim_{t \to \infty} x_i(t) = \sum_{j=1}^{n} \phi_j x_j(0), \qquad \forall \, i,$$

and convergence takes place at the rate of a geometric progression. Furthermore, if $j \in D$, then $\phi_j > 0$.

Theorem 1 shows that all $x_i(t)$ converge to a common limit, resulting in asymptotic consensus. It is a special case of Lemma 2.1 in pp. 805-806 of [5]. (To see this, set all the exogenous driving terms $s_l^j(k)$ in [5] to zero, and identify ϕ_j with $\Phi_l^j(0)$.) The proof of Lemma 2.1 was omitted from [5] as rather straightforward, but can be found in [6]. The proof of a special case of Theorem 1 was included in Sec. 7.3 of [1].

The proof of Theorem 1 (see pp. 248-255 in Appendix A of [6]) proceeds as follows. Consider the matrix $\Phi(k) = \prod_{i=k+1}^{k+nB} A(i)$. An easy induction argument shows that $\Phi(k)$ "is a stochastic matrix with the property that all entries in some column (corresponding to any computing processor) are positive and bounded away from zero by a constant $\overline{\alpha} > 0$ that does not depend on k" (p. 253 of [6]). (The underlying idea is that within nB time steps, a "computing processor" will influence the value held by any other processor.) It then follows, as in [4], that $\prod_{k=1}^{t} \Phi(k)$ converges, as $t \to \infty$, to a matrix with equal rows, and that asymptotic consensus is obtained.

3 Relation with [3]

The model in [3] is the following. At each time t, there is a set $E(t)$ of edges (i, j), and each agent i updates a scalar variable $\theta_i(t)$ according to (cf. Eq. (2) in [3])

$$\theta_i(t+1) = \frac{1}{1+n_i(t)}\Big(\theta_i(t) + \sum_{j\in\mathcal{N}_i(t)} \theta_j(t),\Big) \tag{2}$$

where $\mathcal{N}_i(t)$ is the set $\{j \in N \mid j \neq i,\ (j,i) \in E(t)\}$ of "neighbors" of i, and $n_i(t)$ is the cardinality of $\mathcal{N}_i(t)$. Actually, [3] assumes an undirected graph, which translates to the following symmetry assumption: $(i, j) \in E(t)$ if and only if $(j, i) \in E(t)$.

Clearly, Eq. (2) is a special case of our iteration (1), with the correspondence $D = \{1, \dots, n\}$, $x_i(t) = \theta_i(t)$ and, $a_{ij}(t) = 1/(1 + n_i(t))$ for every $j \neq i$ such that $(j, i) \in E(t)$. Note that our Assumption 1(b) is satisfied with $\alpha = 1/n$.

Finally, [3] makes the following assumption. (The phrasing of this assumption, in Theorem 2 of [3] is somewhat different, but equivalent to what follows.)

Assumption 2. There exists some B such that for every t, the graph with edge set $E(t+1) \cup \dots \cup E(t+B)$ is strongly connected. Furthermore, $(i, j) \in E(t)$ if and only if $(j, i) \in E(t)$.

Reference [3] establishes convergence to consensus under Assumption 2 (Theorem 2 in [3]). Given the discussion above, this convergence result is a special case of the results of [5] (cf. Theorem 1 above), except for the difference between Assumptions 1(c)-(d) and Assumption 2. In particular, reference [5] requires $(V, E(t+1) \cup \dots \cup E(t+B))$ to be the *same* strongly connected graph for all t, whereas [3] allows it to change with t. It turns out that this "same strongly connected graph" restriction is not used anywhere in the proof of Theorem 1 given in [6]. (This should also be clear from the proof outline given above, at the end of Section 2.) The same comments apply to the generalization provided in Section II.A of [3].

Reference [3] then proceeds to consider a variant ("leader following") in which one of the agents (say, agent 1) never updates its own variable, but indirectly influences all of the other agents. This is again a special case of the model of [5], with $D = \{1\}$. Convergence to consensus on the initial value of agent 1 (Theorem 4 in [3]) is again covered by the results of [5] (Theorem 1 above). Once more, the connectivity assumption in [3] is slightly weaker (in [3], $E(t+1) \cup \cdots \cup E(t+B)$ is not required to be the same for all t), but the convergence proof (the proof of Theorem 1) goes through, unaffected by this weaker assumption.

4 Conclusions

In summary, the relation between the results in [6, 5, 1] and the more recent reference [3] are as follows:

(a) The update equations in [3] are a very special case of the agreement algorithm in [5]; the latter allows a much more general form of the coefficients $a_{ij}(t)$, and does not require symmetry in the communication pattern.

(b) The assumption on the intercommunication time intervals are less stringent in [3], but the proof of the result in [5] (as given in [6]) applies without modification.

(c) The results in [5] are more general in the following respect: convergence is established even in the presence of bounded communication delays, for the iteration

$$x_i(t+1) = \sum_{j=1}^{n} a_{ij}(t) x_j\big(\tau_{ij}(t)\big),$$

where $\tau_{ij}(t) \in [t-B, t]$ and where $t - \tau_{ij}(t)$ is interpreted as a communication delay.

References

[1] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Prentice Hall, 1989; republished by Athena Scientific, 1997; available at `http://hdl.handle.net/1721.1/3719`.

[2] M. H. DeGroot, "Reaching a Consensus," *Journal of American Statistical Association*, Vol. 69, No. 345, 1974, pp. 118-121.

[3] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, 2003, pp. 988- 1001.

[4] S. Chatterjee and E. Seneta, "Towards Consensus: Some Convergence Theorems on Repeated Averaging," *J. Applied Probability,* Vol. 14, 1977, pp. 89-97.

[5] J. N. Tsitsiklis, D. P. Bertsekas and M. Athans, “Distributed Asynchronous Deterministic and Stochastic Gradient Optimization Algorithms,” *IEEE Transactions on Automatic Control,* Vol. 31, No. 9, 1986, pp. 803-812.

[6] J. N. Tsitsiklis, *Problems in Decentralized Decision Making and Computation,* Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 1984; available at `http://hdl.handle.net/1721.1/15254`.