Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Rate of Convergence for Consensus with Delays

Pierre-Alexander Bliman
INRIA
Domaine de Voluceau
Rocquencourt BP105
78153 Le Chesnay Sedex, France
Email: pierre-alexandre.bliman@inria.fr

Angelia Nedić
Department of Industrial and
Enterprise Systems Engineering
University of Illinois
Urbana-Champaign, IL 61801
Email: angelia@illinois.edu

Asuman Ozdaglar
Department of Electrical Engineering
and Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02142
Email: asuman@mit.edu

Abstract— **We study the problem of reaching a consensus in the values of a distributed system of agents with time-varying connectivity in the presence of delays. We consider a widely studied consensus algorithm, in which every agent forms a weighted average of its own value with the values received from its neighboring agents. We study an asynchronous operation of this algorithm using delayed agent values. Our focus is on establishing convergence rate results for this algorithm. In particular, for general network topologies, we provide a bound on the time required to reach consensus, which is an explicit function of the system parameters including the delay bound and the bound on agents' intercommunication intervals.**

I. INTRODUCTION

There has been much recent interest in distributed cooperative control problems, in which several autonomous agents try to collectively accomplish a global objective. Most recent literature in this area focuses on the *consensus problem*, where the objective is to develop distributed algorithms for agents to reach an agreement or consensus on a common decision. Consensus problem arises in a number of applications including coordination of UAVs, information processing in wireless sensor networks, and distributed multi-agent optimization.

A widely studied algorithm, proposed and analyzed by Tsitsiklis [19] (see also [21]) involves at each time step every agent computing a weighted average of its own value with values received from some of the other agents. The convergence properties of the consensus algorithm have further been studied under different assumptions on agent connectivity and information exchange in [9] and [5]. The work [19] established upper bounds on the convergence time of this algorithm, which is exponential in the number of agents m. Under some conditions, including doubly stochasticity of the agent weights and no delay, recent work [12] presented tight upper bounds on the convergence time, which is quadratic in m. Despite much work on the consensus algorithm, there has not been a systematic study of the convergence rate of this algorithm in the presence of delays. The presence of delays is a good model for communication networks where the delays are associated with transmission of agent values. Establishing the rate properties of consensus algorithms in such systems is essential in understanding the robustness of the system against dynamic changes.

In this paper, we study convergence and convergence rate properties of the consensus algorithm in the presence of delays. Our analysis is based on reducing the consensus problem with delays to a problem without delays by using state augmentation (i.e., by introducing a new agent in the system for each delay element). The state augmentation allows us to represent the evolution of agent values using linear dynamics. Under the assumption that all delays are bounded, we provide convergence results and convergence rate estimates for the agents' values. Our rate estimates are per iteration and highlight the dependence on the system parameters including the delay bound.

Other than the works cited above, our paper is also related to the literature on the consensus problem and "average" consensus problem (a special case, where the goal is to reach an agreement on the *average of the initial values of the agents*); see [6], [16], [17], [18]. Recently, the implications of noise and quantization effects on the asymptotic behavior of consensus algorithms have been investigated in [10], [8], [12]. Consensus algorithms also play a key role in the development of distributed optimization methods. The convergence properties of such methods have been studied in [20], [11], [3], and more recently in [15], [14].

There has also been some work on the convergence of consensus algorithms in the presence of delays. In particular, [4] studied convergence of (average) consensus under *symmetric* delays for a continuous model of agent updates, i.e., a model that represents the evolution of agent values using partial differential equations (which is in contrast with the slotted update rule studied in this paper). The convergence of consensus algorithms for non-symmetric delays and a discrete time model has been studied in [7]. Other related works are [1], [2], where the rate of convergence of the consensus algorithms is estimated in the presence of delays, assuming special topologies for agent connectivity; namely the existence of repetitive [1] or permanent [2] spanning-trees in the communication graph.[1]

Recent work [13] studied the convergence rate of delayed

This research was partially supported by the National Science Foundation under CAREER grants CMMI 07-42538 and DMI-0545910.

[1]This assumption of communication "from-one-to-all" is in contrast with the setting in the present paper, where communication "from-all-to-all" is assumed.

consensus algorithms for general network topologies using the convergence properties of infinite products of stochastic matrices. In this paper, we provide a novel analysis which allows us to establish improved bounds on the convergence rate estimates for general network topologies.

The paper is organized as follows: In Section II, we formulate the consensus problem and state our assumptions. In Section III, we discuss an equivalent consensus problem without a delay, but with an enlarged number of agents, and we provide our main convergence and rate of convergence results. In Section IV, we provide concluding remarks.

Regarding notation, for a matrix A, we write $[A]_i^j$ to denote the matrix entry in the i-th row and j-th column. We write $[A]_i$ and $[A]^j$ to denote the i-th row and the j-th column of the matrix A, respectively. A vector a is said to be a *stochastic vector* when $a_i \geq 0$ for all i and $\sum_i a_i = 1$. A square matrix A is said to be a *(row) stochastic matrix* when each row of A is a stochastic vector. For a scalar t, we write $\lfloor t \rfloor$ to denote the largest integer less than or equal to t. For a set of vectors $\{z^i, i \in \mathcal{I}\}$, the vector $\max_{i\in\mathcal{I}} z^i$ (or $\min_{i\in\mathcal{I}} z^i$) denotes the componentwise maximum (or minimum) of the vectors $z^i, i \in \mathcal{I}$.

II. Consensus Problem

In this section, we formulate the consensus problem and state our assumptions on agent connectivity and local information exchange.

A. Statement of the Consensus Problem with Delay

We consider a network with m agents. The neighbors of agent i are the agents j communicating with agent i over a directed link (j, i). Each agent updates and sends its information to its neighbors at discrete times $t_0, t_1, \ldots$. We index agents' estimates and other information at time t_k by k. We use $x^i(k) \in \mathbb{R}^n$ to denote agent i estimate at time t_k.

Each agent i updates its estimate $x^i(k)$ by combining it with the available (potentially delayed) estimates $x^j(s)$ of its neighbors j. In particular, each agent i updates its estimate according to the following relation:

$$x^i(k+1) = \sum_{j=1}^{m} a_j^i(k) x^j(k - t_j^i(k)) \quad \text{for } k = 0, 1, 2, \ldots, \quad (1)$$

where the vector $x^i(0) \in \mathbb{R}^n$ is an initial estimate (or state) of agent i. Agent j sends its estimate $x^j(s)$ at time $s = k - t_j^i(k)$, and the estimate reaches agent i at time k. The time $k - t_j^i(k) \geq 0$ and the integer $t_j^i(k)$ are nonnegative for all i, j and k, where the integer $t_j^i(k)$ represents the delay of a message from agent j to agent i. The scalar $a_j^i(k)$ is a nonnegative weight that agent i assigns to a (delayed) estimate $x^j(s)$ arriving from agent j at time k. The vector $a^i(k) = (a_1^i(k), \ldots, a_m^i(k))'$ is the set of nonnegative weights that agent i uses at time k.

The *consensus problem* involves determining conditions on the agents' connectivity and interactions, including conditions on the weights $a^i(k)$, that guarantee the convergence of the estimates $x^i(k)$ to a common vector $\bar{x} \in \mathbb{R}^n$, as $k \to \infty$.

B. Assumptions

We first describe some rules that govern the dynamic evolution of agent estimates, motivated by the model of Tsitsiklis [19] and the "consensus" setting of Blondel *et al.* [5]. We use the following assumption on the weights $a_j^i(k)$.

Assumption 1: (Weights Rule) *We have:*

(a) *There exists a scalar η with $0 < \eta < 1$ such that for all $i \in \{1, \ldots, m\}$,*
   - (i) *$a_i^i(k) \geq \eta$ for all $k \geq 0$.*
   - (ii) *$a_j^i(k) \geq \eta$ for all $k \geq 0$, and all agents j whose (potentially delayed) information $x^j(s)$ reaches agent i in the time interval (t_k, t_{k+1}).*
   - (iii) *$a_j^i(k) = 0$ for all $k \geq 0$ and j otherwise.*

(b) *The vectors $a^i(k)$ are stochastic.*

Assumption 1(a) states that each agent gives significant weights to its own estimate $x^i(k)$ and the estimates received from its neighbors. Under Assumption 1, for the matrix $A(k)$ with columns $a^1(k), \ldots, a^m(k)$, the transpose $A'(k)$ is a stochastic matrix for all $k \geq 0$.

We now discuss the rules for the information exchange among agents. Here, it is convenient to view the agents as a set of nodes $V = \{1, \ldots, m\}$. At each update time t_k, the agents' communications may be represented by a directed graph (V, E_k) with the set E_k of directed edges given by

$$E_k = \{(j, i) \mid a_j^i(k) > 0\}.$$

We impose a connectivity assumption on the agent system, which ensures that the information state of any agent i influences the state of any other agent infinitely often in time. In formulating this, we use the set E_∞ consisting of directed edges (j, i) such that j is a neighbor of i who communicates with i infinitely often in time.

Assumption 2: (Connectivity) *The graph (V, E_∞) is strongly connected, where E_∞ is the set of edges (j, i) representing agent pairs communicating directly infinitely many times, i.e.,*

$$E_\infty = \{(j, i) \mid (j, i) \in E_k \text{ for infinitely many indices } k\}.$$

Assumption 2 is equivalent to having the composite directed graph $(V, \cup_{l \geq k} E_l)$ strongly connected for all k. This all-to-all communication assumption will be central in the sequel.[2] This assumption is restrictive since it requires that the "steady-state" graph is strongly connected, however it is a standard assumption in the consensus literature.

We assume that the intercommunication intervals are bounded for those agents that communicate infinitely often. In particular, we use the following.

Assumption 3: (Bounded Intercommunication Interval) *There exists an integer $B \geq 1$ such that for every $(j, i) \in E_\infty$, agent j sends information to its neighbor i at least once every B consecutive time slots, i.e., at time t_k or at time t_{k+1} or $\ldots$ or (at latest) at time t_{k+B-1} for any $k \geq 0$.*

Finally, we assume that the delays $t_j^i(k)$ in delivering a message from an agent j to any neighboring agent i are

[2] It implies that the initial value of every agent affects the consensus outcome, or all agents are "distinguished" using the terminology of [3, Section 7.3].

uniformly bounded at all times.[3]

Assumption 4: (Bounded Delays) *Let the following hold:*

(a) $t_i^i(k) = 0$ *for all agents* i *and all* $k \geq 0$.

(b) $t_j^i(k) = 0$ *for all agents* j *whose estimates* x^j *are not available to agent* i *at time* t_{k+1}.

(c) *There is an integer* B_1 *such that* $0 \leq t_j^i(k) \leq B_1 - 1$ *for all agents* i, j, *and all* k.

Assumption 4 (a) states that each agent i has its own estimate available without any delay. Part (b) states that the delay is zero for those agents j whose (delayed) estimates $x^j(s)$ are not available to agent i at an update time. Under Assumption 1 (a), Assumption 4 (b) reduces to the following relation: $t_j^i(k) = 0$ when $a_j^i(k) = 0$. Assumption 4 (c) states that the delays are uniformly bounded at all times and for all agents.

III. Convergence Analysis

In this section, we show that the agents updating their information according to Eq. (1) reach a consensus under the assumptions of Section II-B. In particular, we establish the convergence of agent estimates and provide a convergence rate estimate. Our analysis is based on reducing the consensus problem with delays to a problem without delays.

A. Reduction to a Consensus Problem without Delay

Under the Bounded Delays [cf. Assumption 4], we reduce the original agent system with delays to a system without delays. In particular, we define an enlarged agent system that is obtained by adding new (virtual) agents into the original system in order to deal with delays. With each agent i of the original system, we associate a new agent for each of the possible delay values that a message originating from agent i may experience. In view of the Bounded Delays assumption, it suffices to add $(m-1)B_1$ new agents handling the delays.[4]

To differentiate between the original agents in the system and the new agents, we introduce the notions of computing and noncomputing agents. We refer to the original agents as **computing agents** since these agents maintain and update their estimates. We refer to the new agents as **noncomputing agents** since they do not compute or update any information, but only pass the received information to their neighbors.

In the enlarged system, we enumerate the computing agents first and then the noncomputing agents. In particular, the computing agents are indexed by $1, \ldots, m$ and noncomputing agents are indexed by $m+1, \ldots, (B_1-1)m$. Furthermore, the noncomputing agents are indexed so that the first m of them model the delay of 1 for the computing agents, the next m of them model the delay of 2 for the computing agents, and so on. Formally, for a computing agent i, the noncomputing agents $i+m, \ldots, i+(B_1-1)m$ model the delay values $t = 1, \ldots, (B_1-1)m$, respectively.

We now identify the neighbors of each agent in the new (virtual) system. The computing agents are connected and

[3]The intercommunication interval and the delay bounds are used in our analysis. In the implementation of the algorithm, these bounds need not be available to any agent.

[4]This idea has also been used in the distributed computation model of Tsitsiklis [19], and it motivates our development here.

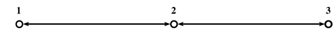

Figure 1(a)

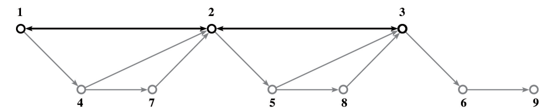

Figure 1(b)

Fig. 1. Figure 1(a) illustrates an agent network with 3 agents and two bidirectional links (1,2) and (2,3). Figure 1(b) illustrates the enlarged network associated with the original network of part (a), when the delay bound is $B_1 = 3$. The noncomputing agents introduced in the system are $4, \ldots, 9$. Agents 4, 5, and 6 model the delay of 1 while agents 7, 8, and 9 model the delay of 2 for the computing nodes 1, 2 and 3, respectively.

communicate in the same way as in the original system. The noncomputing agents corresponding to the delays of different computing agents do not communicate among themselves. Specifically, for t with $1 \leq t < B_1 - 1$, a noncomputing agent $j+tm$ receives the information only from agent $j+(t-1)m$, and sends the same information to either the noncomputing agent $j+(t+1)m$, or to a computing agent i provided that agent j communicates with agent i in the original system. The same rule applies for $t = B_1 - 1$ except that the only possible transmission is to a computing agent. An agent system with 3 agents and a maximum delay of 3, and the corresponding enlarged system are illustrated in Figure 1, together with the enumeration rule for noncomputing agents.

We let $\tilde{x}^i(k)$ denote the estimate of agent i in the enlarged system at time k. Then, the relation in Eq. (1) for the *evolution of estimates of computing agents* is given by:

$$\tilde{x}^i(k+1) = \sum_{h=1}^{mB_1} \tilde{a}_h^i(k)\tilde{x}^h(k) \quad \text{for all } i \text{ and } k, \tag{2}$$

where for all $h \in \{1, \ldots, mB_1\}$ and $k \geq 0$,

$$\tilde{a}_h^i(k) = \begin{cases} a_j^i(k) & \text{if } h = j + tm,\ t = t_j^i(k) \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

and $a_j^i(k)$ are the weights used by the agents in the original system. The *evolution of states for noncomputing agents* is given by: for all $i = m+1, \ldots, mB_1$ and $k \geq 0$,

$$\tilde{x}^i(k+1) = \tilde{x}^{i-m}(k),$$

where the initial values are $\tilde{x}^i(0) = 0$. Therefore, for noncomputing agents i and all $k \geq 0$, we have

$$\tilde{a}_h^i(k) = \begin{cases} 1 & \text{for } h = i - m \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

We can now compactly write the *evolution of estimates* $\tilde{x}^i(k)$ *for all agents* i *in the enlarged system* as follows: for all $i \in \{1, \ldots, mB_1\}$ and $k \geq 0$,

$$\tilde{x}^i(k+1) = \sum_{h=1}^{mB_1} \tilde{a}_h^i(k)\tilde{x}^h(k), \tag{5}$$

where the initial vectors are given by: for $i \in \{1,\ldots,m\}$ and $t \in \{0,\ldots,B_1-1\}$,

$$\tilde{x}^{i+tm}(0) = x^i(0). \tag{6}$$

The weights $\tilde{a}^i_h(k)$ for computing agents $i \in \{1,\ldots,m\}$ and for noncomputing agents $i \in \{m+1,\ldots,mB_1\}$ are given respectively by Eqs. (3) and (4). Thus, for a noncomputing agent i, we always have $\sum_{h=1}^{mB_1} \tilde{a}^i_h(k) = 1$. For a computing agent, we have $\sum_{h=1}^{mB_1} \tilde{a}^i_h(k) = 1$ *if and only if* $\sum_{j=1}^{m} a^i_j(k) = 1$ for the weights $a^i_j(k)$ in the original system.

We next represent the evolution of the estimates $\tilde{x}^i(k)$ of Eq. (5) in a form more suitable for our convergence analysis. Specifically, we introduce matrices $\tilde{A}(s)$ whose i-th column is the vector $\tilde{a}^i(s)$, and we define the *transition matrices* $\tilde{\Phi}(k,s)$ for any s and k with $k \geq s$,

$$\tilde{\Phi}(k,s) = \tilde{A}(s)\tilde{A}(s+1)\cdots\tilde{A}(k-1)\tilde{A}(k). \tag{7}$$

Using these matrices, we relate estimates $\tilde{x}^i(k+1)$ to the estimates $\tilde{x}^j(s)$ for all j and any $s \leq k$. Specifically, it can be seen that for the iterates generated by Eq. (5), we have for any i, and any s and k with $k \geq s$,

$$\tilde{x}^i(k+1) = \sum_{j=1}^{mB_1} [\tilde{\Phi}(k,s)]^i_j\, \tilde{x}^j(s) \tag{8}$$

(for details see [13]). We next establish some properties of the transition matrices $\tilde{\Phi}(k,s)$ that we use later in our convergence analysis of the iterates $\tilde{x}^i(k)$.

B. Properties of the Transition Matrices

Here, we discuss some properties of the matrices $\tilde{\Phi}(k,s)$ under the assumptions imposed on agent interactions in Section II-B. In particular, under the Weights Rule [Assumption 1], from the definition of the weights $\tilde{a}^i_h(k)$ in Eqs. (3) and (4), it follows that each matrix $\tilde{A}(k)'$ is stochastic. Since the product of stochastic matrices is a stochastic matrix, it follows that *the matrices* $\tilde{\Phi}(k,s)'$ *are stochastic*.

In the following lemma, we establish some additional properties of these matrices that will be important in our convergence analysis. Specifically, we show that the entries of the row $[\tilde{\Phi}(s+(m-1)B+mB_1-1,s)]_j$ are uniformly bounded away from zero for all s and for all computing nodes $j \in \{1,\ldots,m\}$.

Lemma 1: Let Weights Rule (a), Connectivity, Bounded Intercommunication Interval, and Bounded Delay assumptions hold for the agents in the original network [cf. Assumptions 1(a), 2, 3, and 4]. Then, the following hold.

(a) *For any computing nodes $i,j \in \{1,\ldots,m\}$, we have for all $s \geq 0$, and all $k \geq s+(m-1)(B+B_1)$,*

$$[\tilde{\Phi}(k,s)]^i_j \geq \eta^{k-s+1}.$$

(b) *For any computing node $j \in \{1,\ldots,m\}$, we have for all nodes i and all $s \geq 0$,*

$$[\tilde{\Phi}(s+(m-1)B+mB_1-1,s)]^i_j \geq \eta^{(m-1)B+mB_1}.$$

Lemma 1(b) states that the information originating from any computing agent contributes to the estimate of any other agent with a "significant" weight provided that a sufficient time has passed by. The proof of Lemma 1 is given in [13].

C. Convergence Result

In this section, we establish the main results of this paper: we prove the convergence of the iterates of Eq. (1) to a consensus and we provide convergence rate estimates. The proof uses the equivalence between the evolution equations (1) for the original system and the evolution equations (5) for the enlarged system.

We first provide a result showing that the differences in the maximum and the minimum values of the agent estimates decrease in time. This result is central in our convergence analysis of the estimates. Lemma 1 plays a key role in establishing this result.

Lemma 2: Let Weights Rule, Connectivity, Bounded Intercommunication Interval, and Bounded Delay assumptions hold [cf. Assumptions 1–4]. Let the sequences $\{\tilde{x}^i(k)\}$, $i = 1,\ldots,mB_1$ be generated by Eq. (5). For all $k \geq 0$, define $M(k), \mu(k) \in \mathbb{R}^n$ as follows:

$$M(k) = \max_{1\leq i\leq mB_1} \tilde{x}^i(k), \quad \mu(k) = \min_{1\leq i\leq mB_1} \tilde{x}^i(k). \tag{9}$$

Then, we have $m\eta^{B_2} \leq 1$ and for all $k \geq 0$,

$$M(k) - \mu(k) \leq \left(1 - m\eta^{B_2}\right)^{\left\lfloor \frac{k}{B_2} \right\rfloor} (M(0) - \mu(0)),$$

where $B_2 = (m-1)B + mB_1$.

Proof: By Eq. (8), we have for all $i \in \{1,\ldots,mB_1\}$,

$$\tilde{x}^i(s+B_2) = \sum_{j=1}^{mB_1} [\tilde{\Phi}(s+B_2,s)]^i_j\, \tilde{x}^j(s) \qquad \text{for all } s \geq 0,$$

where $B_2 = (m-1)B + mB_1$. Distinguishing the terms due to computing agents $j = 1,\ldots,m$ and noncomputing agents $j = m+1,\ldots,mB_1$ in the sum on the right hand-side of the preceding relation, we can write

$$\begin{aligned}\tilde{x}^i(s+B_2) &= \sum_{j=1}^{m} \eta^{B_2}\tilde{x}^j(s) \\ &+ \sum_{j=1}^{mB_1} [\tilde{\Psi}(s+B_2-1,s)]^i_j \tilde{x}^j(s),\end{aligned} \tag{10}$$

where $[\tilde{\Psi}(s+B_2-1,s)]^i_j = [\tilde{\Phi}(s+B_2-1,s)]^i_j - \eta^{B_2}$ for computing agent j, and $[\tilde{\Psi}(s+B_2-1,s)]^i_j = [\tilde{\Phi}(s+B_2-1,s)]^i_j$ for noncomputing agent j. By Lemma 1(b), we have for any agent i and any *computing* agent $j \in \{1,\ldots,m\}$:

$$[\tilde{\Phi}(s+B_2-1,s)]^i_j \geq \eta^{B_2} \qquad \text{for all } k \geq 0.$$

By the definition of the matrix $\tilde{\Psi}(s+B_2-1,s)$, this implies

$$[\tilde{\Psi}(s+B_2-1,s)]^i_j \geq 0 \quad \text{for all } i,j \in \{1,\ldots,mB_1\}.$$

Since the matrix $\tilde{\Phi}'(s+B_2-1,k)$ is stochastic and there are m computing agents, we have the relation

$$\sum_{j=1}^{mB_1} [\tilde{\Psi}(s+B_2-1,k)]^i_j = 1 - m\eta^{B_2}.$$

This relation together with $[\tilde{\Psi}(s+B_2-1,s)]^i_j \geq 0$ for all i,j implies the compatibility relation $1-m\eta^{B_2} \geq 0$ and

$$(1-m\eta^{B_2})\mu(s) \leq \sum_{j=1}^{mB_1}[\tilde{\Psi}(s+B_2-1,s)]^i_j\tilde{x}^j(s) \leq (1-m\eta^{B_2})M(s),$$

where $\mu(s)$ and $M(s)$ are defined in Eq. (9). Combining the preceding relation with Eq. (10), we obtain for all $i \in \{1,\ldots,mB_1\}$ and $s \geq 0$,

$$\left(1-m\eta^{B_2}\right)\mu(s) \leq \tilde{x}^i(s+B_2) - \sum_{j=1}^{m}\eta^{B_2}\tilde{x}^j(s) \leq \left(1-m\eta^{B_2}\right)M(s).$$

Since this relation holds for all i, it follows that

$$\left(1-m\eta^{B_2}\right)\mu(s) \leq \mu(s+B_2) - \sum_{j=1}^{m}\eta^{B_2}\tilde{x}^j(s),$$

$$M(s+B_2) - \sum_{j=1}^{m}\eta^{B_2}\tilde{x}^j(s) \leq \left(1-m\eta^{B_2}\right)M(s).$$

From the preceding two relations we obtain for all $s \geq 0$,

$$M(s+B_2)-\mu(s+B_2) \leq \left(1-m\eta^{B_2}\right)(M(s)-\mu(s)). \quad (11)$$

In view of the stochasticity of the matrices $\tilde{\Phi}(k,s)'$, the sequences $\{M(k)\}$ and $\{\mu(k)\}$ are nonincreasing and nondecreasing, respectively. Therefore, for any $k \geq 0$,

$$M(k)-\mu(k) \leq M(lB_2)-\mu(lB_2) \leq \left(1-m\eta^{B_2}\right)^l(M(0)-\mu(0)),$$

where $l = \lfloor k/B_2 \rfloor$ and the second inequality follows by a recursive application of relation (11). ■

Using the relations between the original system and the enlarged system, the vectors $M(k)$ and $\mu(k)$ defined in Eq. (9) can be also expressed as:

$$M(k) = \max_{\substack{i=1,\ldots,m\\ \tau=0,\ldots,B_1-1}} x^i(k-\tau), \quad \mu(k) = \min_{\substack{i=1,\ldots,m\\ \tau=0,\ldots,B_1-1}} x^i(k-\tau),$$

where $x^i(k-\tau) = x^i(0)$ for all k and τ such that $k-\tau \leq 0$. In view of these relations, Lemma 2 provides a bound on the decrease in the difference between the maximum and the minimum values of the agent current and delayed estimates up to the maximum possible delay (i.e., the delay B_1-1).

We now prove our main results.

Theorem 1: Let Weights Rule, Connectivity, Bounded Intercommunication Interval, and Bounded Delay assumptions hold [cf. Assumptions 1–4]. Then, the following hold.

(a) *The sequences $\{x^i(k)\}, i=1,\ldots,m$ generated by Eq. (1) converge to a consensus, i.e., for an $\bar{x} \in \mathbb{R}^n$,*

$$\lim_{k\to\infty} x^i(k) = \bar{x} \qquad \text{for all } i=1,\ldots,m.$$

(b) *The consensus vector $\bar{x} \in \mathbb{R}^n$ is a nonnegative combination of the agent initial vectors $x^j(0), j=1,\ldots,m$,*

$$\bar{x} = \sum_{j=1}^{m} w_j x^j(0),$$

with scalars $w_j \geq 0$ such that $\sum_{j=1}^m w_j = 1$, and

$$\bar{x} - \min_{1\leq i\leq m} x^i(0) \geq m\eta^{B_2}\left(\frac{1}{m}\sum_{i=1}^m x^i(0) - \min_{1\leq i\leq m} x^i(0)\right),$$

$$\max_{1\leq i\leq m} x^i(0) - \bar{x} \geq m\eta^{B_2}\left(\max_{1\leq i\leq m} x^i(0) - \frac{1}{m}\sum_{i=1}^m x^i(0)\right).$$

(c) *The convergence rate to the consensus is geometric: for all agents $i \in \{1,\ldots,m\}$,*

$$\|x^i(k)-\bar{x}\| \leq 2\left(1-m\eta^{B_2}\right)^{\left\lfloor \frac{k}{B_2}\right\rfloor} \sum_{j=1}^m \|x^j(0)-\bar{x}\|.$$

Proof: For the vectors $M(k)$ and $\mu(k)$ defined in Eq. (9), and for all $i \in \{1,\ldots,mB_1\}$ and all $k \geq 0$, we have

$$\mu(k) \leq \tilde{x}^i(k) \leq M(k). \quad (12)$$

By the stochasticity of the columns $[\tilde{\Phi}(k,s)]^i$, it follows that the sequences $\{M(k)\}$ and $\{\mu(k)\}$ are bounded and monotone; therefore, $\{M(k)\}$ and $\{\mu(k)\}$ are convergent. By Lemma 2, we have $\lim_{k\to\infty}(M(k)-\mu(k)) = 0$, implying that $M(k)$ and $\mu(k)$ converge to the same limit, denoted by $\bar{x} \in \mathbb{R}^n$. Therefore, by Eq. (12), we obtain $\lim_{k\to\infty}\tilde{x}^i(k) = \bar{x}$ for all $i \in \{1,\ldots,mB_1\}$. Since $\tilde{x}^i(k) = x^i(k)$ for all $i=1,\ldots,m$ and all $k \geq 0$, it follows

$$\lim_{k\to\infty} x^i(k) = \bar{x} \qquad \text{for all } i \in \{1,\ldots,m\},$$

establishing the result in part (a).

Letting $s=0$ in Eq. (8), we have for any $i=1,\ldots,mB_1$,

$$\tilde{x}^i(k) = \sum_{j=1}^{mB_1}[\tilde{\Phi}(k-1,0)]^i_j\,\tilde{x}^j(0) \qquad \text{for all } k \geq 0.$$

Using the definition (6) of the initial vectors $\tilde{x}^i(0)$ for the agents in the enlarged system, we have for all $k \geq 0$,

$$\tilde{x}^i(k) = \sum_{j=1}^{m}\left(\sum_{t=0}^{B_1-1}[\tilde{\Phi}(k-1,0)]^i_{j+tm}\right)x^j(0). \quad (13)$$

From part (a), for any initial vectors $x^j(0)$, $j=1,\ldots,m$, the limit $\lim_{k\to\infty}\tilde{x}^i(k)$ exists so that for all $i \in \{1,\ldots,m\}$,

$$\lim_{k\to\infty}\tilde{x}^i(k) = \sum_{j=1}^m \lim_{k\to\infty}\left(\sum_{t=0}^{B_1-1}[\tilde{\Phi}(k-1,0)]^i_{j+tm}\right)x^j(0).$$

For any fixed h, we can take the initial vectors as $x^j(0)=0$ for all $j \neq h$, and $x^h(0) = e$, where $e \in \mathbb{R}^n$ is the vector of all 1's. Then, the preceding relation and part (a) imply that the limit

$$\lim_{k\to\infty}\left(\sum_{t=0}^{B_1-1}[\tilde{\Phi}(k-1,0)]^i_{h+tm}\right)$$

exists for all i,h, and it is independent of i. Denoting this limit by w_h and using Eq. (13), we obtain the first result in part (b), where the properties of the weights w_h follow from the stochasticity of the vectors $[\tilde{\Phi}(k-1,0)]^i$.

From Eq. (10), we have for all $i = 1, \ldots, mB_1$ and $s \geq 0$,

$$\begin{aligned} \tilde{x}^i(s + B_2) &\geq \eta^{B_2} \sum_{j=1}^{m} \tilde{x}^j(s) + (1 - m\eta^{B_2})\mu(s) \\ &= \eta^{B_2} \sum_{j=1}^{m} x^j(s) + (1 - m\eta^{B_2})\mu(s). \end{aligned}$$

This yields the following inequality for all $s \geq 0$,

$$D \begin{pmatrix} \frac{1}{m} \sum_{j=1}^{m} x^j(s) \\ \mu(s) \end{pmatrix} \leq \begin{pmatrix} \frac{1}{m} \sum_{j=1}^{m} x^j(s + B_2) \\ \mu(s + B_2) \end{pmatrix},$$

where D is a matrix given by

$$D = \begin{pmatrix} m\eta^{B_2} & 1 - m\eta^{B_2} \\ m\eta^{B_2} & 1 - m\eta^{B_2} \end{pmatrix}.$$

As the matrix D is idempotent (i.e., equal to its square), using the preceding inequality recursively with $s = 0$, and then by taking the limit, we can see that

$$D \begin{pmatrix} \frac{1}{m} \sum_{j=1}^{m} x^j(0) \\ \mu(0) \end{pmatrix} \leq \begin{pmatrix} \bar{x} \\ \bar{x} \end{pmatrix},$$

which follows in view of $x^j(k) \rightarrow \bar{x}$ for all j. The second inequality yields the lower bound on the difference $\bar{x} - \min_{1 \leq i \leq m} x^i(0)$ in part (b). The lower bound on the difference $\max_{1 \leq i \leq m} x^i(0) - \bar{x}$ in part (b) is proved similarly.

Since $x^i(k) \leq M(k)$ for all $i \in \{1, \ldots, m\}$ and all $k \geq 0$, and $\mu(k) \leq \bar{x}$ for all $k \geq 0$ [which follows because $\{\mu(k)\}$ is nondecreasing], we have

$$\|x^i(k) - \bar{x}\| \leq \|M(k) - \mu(k)\| \qquad \text{for all } k \geq 0.$$

Moreover,

$$\|M(0) - \mu(0)\| \leq \|M(0) - \bar{x}\| + \|\mu(0) - \bar{x}\|. \tag{14}$$

By the definition of $M(0)$ and $\mu(0)$ in Eq. (9), and the definitions of the vectors $\tilde{x}^i(0)$ of Eq. (6), we obtain $M(0) = \max_{1 \leq i \leq m} x^i(0)$ and $\mu(0) = \min_{1 \leq i \leq m} x^i(0)$, implying

$$\|M(0) - \bar{x}\| + \|\mu(0) - \bar{x}\| \leq 2 \sum_{j=1}^{m} \|x^j(0) - \bar{x}\|.$$

The rate result of part (c) follows by combining the preceding relation with Eq. (14) and by using Lemma 2. ■

The result in Theorem 1(b) can be viewed as an error estimate between the final consensus value and the average of the initial agent estimates. In particular, if $m\eta^{B_2} = 1$, then it follows from these relations that $\bar{x} = \frac{1}{m} \sum_{i=1}^{m} x^i(0)$.

The following rate estimate was shown in [13]:

$$\|x^i(k + 1) - \bar{x}\| \leq 2 \frac{1 + \eta^{-B_2}}{1 - \eta^{B_2}} \left(1 - \eta^{B_2}\right)^{\frac{k}{B_2}} \sum_{j=1}^{m} \|x^j(0)\|.$$

The estimate in Theorem 1(c) improves this bound in terms of the much better constant (independent of η) and the decrease factor of $1 - m\eta^{B_2}$ instead of $1 - \eta^{B_2}$.

IV. CONCLUSIONS

We considered an algorithm for the consensus problem in the presence of delays in the multi-agent system. Our analysis relies on reducing the problem to a consensus problem in an enlarged system without delays. We studied properties of the reduced model and used them to establish the convergence and convergence rate properties for the consensus problem with delays. Our convergence rate estimate is geometric and it is explicitly given in terms of the system parameters. Future work includes incorporating the delayed consensus algorithm in the distributed optimization model developed in [15] to account for delays in the agent values.

REFERENCES

[1] D. Angeli and P.-A. Bliman, *Convergence speed of unsteady distributed consensus: decay estimate along the settling spanning-trees*, http://arxiv.org/abs/math.OC/0610854, 2006.

[2] ———, *Tight estimates for convergence of some non-stationary consensus algorithms*, http://arxiv.org/abs/0706.0630, 2007.

[3] D.P. Bertsekas and J.N. Tsitsiklis, *Parallel and distributed computation: Numerical methods*, Athena Scientific, Belmont, MA, 1997.

[4] P.-A. Bliman and G. Ferrari-Trecate, *Average consensus problems in networks of agents with delayed communications*, Automatica (2008).

[5] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis, *Convergence in multiagent coordination, consensus, and flocking*, Proceedings of IEEE CDC, 2005.

[6] M. Cao, A.S. Morse, and B.D.O. Anderson, *Reaching a consensus in a dynamically changing environment: A graphical approach*, SIAM J. on Control and Opt. **47** (2008), no. 2, 575–600.

[7] ———, *Reaching a consensus in a dynamically changing environment: Convergence rates, measurement delays, and asynchronous events*, SIAM J. on Control and Opt. **47** (2008), no. 2, 601–623.

[8] R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri, *Communication constraints in coordinated consensus problems*, Proceedings of IEEE ACC, 2006.

[9] A. Jadbabaie, J. Lin, and S. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Transactions on Automatic Control **48** (2003), no. 6, 988–1001.

[10] A. Kashyap, T. Basar, and R. Srikant, *Quantized consensus*, Proceedings of IEEE CDC, 2006.

[11] S. Li and T. Basar, *Asymptotic agreement and convergence of asynchronous stochastic algorithms*, IEEE Transactions on Automatic Control **32** (1987), no. 7, 612–618.

[12] A. Nedić, A. Olshevsky, A. Ozdaglar, and J.N. Tsitsiklis, *On distributed averaging algorithms and quantization effects*, Tech. Report LIDS 2778, MIT, 2007.

[13] A. Nedić and A. Ozdaglar, *Convergence rate for consensus with delays*, Tech. Report LIDS 2774, MIT, 2007.

[14] ———, *On the rate of convergence of distributed subgradient methods for multi-agent optimization*, Proceedings of IEEE CDC, 2007.

[15] ———, *Distributed subradient methods for multi-agent optimization*, forthcoming in IEEE Transactions on Automatic Control, 2008.

[16] R. Olfati-Saber and R.M. Murray, *Consensus problems in networks of agents with switching topology and time-delays*, IEEE Transactions on Automatic Control **49** (2004), no. 9, 1520–1533.

[17] A. Olshevsky and J.N. Tsitsiklis, *Convergence rates in distributed consensus and averaging*, Proceedings of IEEE CDC, 2006.

[18] ———, *Convergence speed in distributed consensus and averaging*, forthcoming in SIAM Journal on Control and Optimization, 2008.

[19] 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.

[20] J.N. Tsitsiklis and M. Athans, *Convergence and asymptotic agreement in distributed decision problems*, IEEE Transactions on Automatic Control **29** (1984), no. 1, 42–50.

[21] J.N. Tsitsiklis, D.P. Bertsekas, and M. Athans, *Distributed asynchronous deterministic and stochastic gradient optimization algorithms*, IEEE Transactions on Automatic Control **31** (1986), no. 9, 803–812.