2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

A Fixed-Neighbor, Distributed Algorithm for Solving a Linear Algebraic Equation

S. Mou A. S. Morse
Yale University
shaoshuai.mou@yale.edu, as.morse@yale.edu

***Abstract*— This paper presents a distributed algorithm for solving a linear algebraic equation of the form $Ax = b$ where A is an $n \times n$ nonsingular matrix and b is an $n-$vector. The equation is solved by a network of n agents assuming that each agent knows exactly one distinct row of the partitioned matrix $[\,A \quad b\,]$, the current estimates of the equation's solution generated by its neighbors, and nothing more. Each agent recursively updates its estimate of $A^{-1}b$ by utilizing the current estimates generated by each of its neighbors. Neighbor relations are characterized by a simple, undirected graph $\mathbb{G}$ whose vertices correspond to agents and whose edges depict neighbor relations. It is shown that for any nonsingular matrix A and any connected graph $\mathbb{G}$, the proposed algorithm causes all agents' estimates to converge exponentially fast to the desired solution $A^{-1}b$.**

I. INTRODUCTION

Linear algebraic equations arise quite naturally when modeling many phenomena, such as forecasting, estimation, or approximating non-linear equations [1]. Efforts to develop distributed algorithms to solve such equations have been under way for a long time especially in the parallel processing community where the main objective is to achieve efficiency by somehow decomposing a large system of linear equations into smaller ones which can be solved on parallel processors more accurately or faster than direct solutions of the original equations would allow [2]–[4]. In this paper, we are interested in solving linear equations of the form $Ax = b$ using a network of n autonomous agents; here $A \in \mathbb{R}^{n\times n}$ is non-singular and $b \in \mathbb{R}^n$.

Each agent in the network is able to exchange information with certain other agents called its "neighbors". We write $\mathcal{N}_i$ for the labels of agent i's neighbors, and we always take agent i to be a neighbor of itself. Neighbor relations can be conveniently characterized by an undirected graph $\mathbb{G}$ with n vertices and a set of undirected edges defined so that there is an edge (i, j) in $\mathbb{G}$ just in case that agent i and j are neighbors. Each agent i updates a state vector $x_i(t)$ taking values in $\mathbb{R}^n$, and we assume that the information agent i receives from neighbor j at time t is $x_j(t)$. We also assume that each agent i knows $A_i \in \mathbb{R}^{1\times n}$ and $b_i \in \mathbb{R}$, where $[\,A_i \quad b_i\,]$ is a distinct row of the partitioned matrix $[\,A \quad b\,]$. Note that $\mathbb{G}$ limits the information flow across the network, which consequently precludes centralized processing. With this in mind we are led to the problem of devising a local algorithm for each agent which causes its state to converge to $A^{-1}b$.

There are a number of classical parallel algorithms for solving linear equations. Among these are Jacobi iterations, so called "successive over relaxations" method [5], and Kaczmart's method [6]. Although these are parallel algorithms, they either require A to be diagonally dominant or positive definite, or rely on " relaxation factors", which are not easily determined in a distributed setting. Reference [7] and [8] develop several algorithms for a special version of the problem and give sufficient conditions for them to work correctly.

It is easy to see that problem under consideration here can be recast as a distributed optimization problem to which the methods in [9] can be applied. Although we do not do this, we do exploit an especially useful observation from [9] which makes quite clever use of the idea of consensus. Here's the *key idea*: Observe that if $x_1, x_2, \ldots, x_n$ are vectors satisfying $A_i x_i = b_i$ for $i \in \{1, 2, \ldots, n\}$, and in addition, if a consensus is reached in that all x_i are equal, then automatically all x_i satisfy $Ax_i = b$. So one can address the overall problem of interest by having each agent i solve its own equation and at the same time making sure that a consensus is reached. The algorithm proposed in [9] almost accomplishes this. However it is based on gradient descent and thus has the disadvantages normally associated with that methodology: e.g. slow convergence as the optimal solution is approached.

The approach taken in this paper is as follows: suppose time is discrete in that t takes values in $\{1, 2, \ldots\}$. Suppose each agent i initializes its state x_i at clock time $t = 1$ by picking $x_i(1)$ to be any solution to the equation $A_i x_i = b_i$. Suppose that K_i is a matrix whose column span is the kernel of A_i. If we restrict the updating of $x_i(t)$ to iterations of the form $x_i(t + 1) = x_i(t) + K_i u_i(t), \quad t \geq 1$, then no matter what $u_i(t)$ is, each $x_i(t)$ will obviously satisfy $A_i x_i(t) = b_i,\ t \geq 1$. Thus all we need to do to solve the problem is to come up with a good way to choose the u_i so that a consensus is ultimately reached. Capitalizing on what's known about consensus algorithms based on averaging [10]–[13], we choose each $u_i(t)$ to minimize the difference $x_i(t) + K_i u_i(t) - \frac{1}{d_i}\left(\sum_{j\in\mathcal{N}_i} x_j(t)\right)$ in the least square sense, where d_i denote the number of neighbors of agent i. Doing this

The research of S. Mou and A. S. Morse was supported by the US Air Force Office of Scientific Research and the by National Science Foundation.

leads at once to an iteration for agent i of the form

$$x_i(t+1) = x_i(t) - \frac{1}{d_i} P_i \left(d_i x_i(t) - \sum_{j \in \mathcal{N}_i} x_j(t) \right) \quad (1)$$

where P_i is the readily computable orthogonal projection on the kernel of A_i. Note right away that the algorithm does not involve a step-size or a relaxation factor and is totally distributed. The main result of this paper is as follows.

Theorem 1: Suppose each agent i updates its state $x_i(t)$ according to rule (1). If A is non-singular and $\mathbb{G}$ is connected, then there exists an non-negative constant $\rho < 1$ for which all $x_i(t)$ converges to $A^{-1}b$ as t goes to infinity as fast as ρ^t converges to 0.

II. Analysis

In this section we explain why Theorem 1 is true. Toward to this end, we orient $\mathbb{G}$ by associating with each edge (i,j) a " head" i and a "tail" j thereby converting (i,j) into a directed arc from i to j. Let m denote the number of edges in $\mathbb{G}$. Define the *incidence matrix* of $\mathbb{G}$ as an $n \times m$ matrix $H = [h_{ik}]$ with entries given by

$$h_{ik} = \begin{cases} -1 & \text{if} \quad i \text{ is the tail of the } k\text{th edge} \\ 1 & \text{if} \quad i \text{ is the head of the } k\text{th edge} \\ 0 & \quad \text{otherwise} \end{cases}$$

Since $\mathbb{G}$ is connected, one has $\ker H' = \text{span}\,\{\mathbf{1}\}$, where $\mathbf{1}$ is the $n \times 1$ vector with all elements equal to 1.

Suppose the kth edge in the oriented $\mathbb{G}$ is with i as the head and j as the tail and define $e_k = x_i - x_j$, $k = 1, 2, ..., m$. Let $e(t) = [\, e_1(t)' \quad e_2(t)' \quad \cdots \quad e_m(t)' \,]'$ and $x(t) = [\, x_1(t)' \quad x_2(t)' \quad \cdots \quad x_n(t)' \,]'$. Then

$$e(t) = \bar{H}' x(t), \quad (2)$$

where $\bar{H} = H \otimes I_n$ and $\otimes$ denotes the Kronecker product. Let $\bar{D} = D \otimes I_n$, where $D = \text{diag}\,\{d_1, d_2, ..., d_n\}$. Let $P = \text{diag}\,\{P_1, P_2, ..., P_n\}$. Since each P_i is an orthogonal projection matrix, so is P. By (1), one has

$$x(t+1) = x(t) - P\bar{D}^{-1}\bar{H}e(t). \quad (3)$$

Equation (3) bridges the connection between $e(t) \to 0$ and $x_i(t) \to A^{-1}b$, $i = 1, 2, ..., n$, as shown in the following lemma.

Lemma 1: There exists an non-negative constant $\rho < 1$ such that $e(t) \to 0$ as fast as $\rho^t \to 0$. Then each $x_i(t)$ tends to $A^{-1}b$ as fast as $\rho^t \to 0$.

Proof of Lemma 1 will be given in the Appendix. Theorem 1 is a direct subsequence of the following proposition.

Proposition 1: Suppose A is non-singular and $\mathbb{G}$ is connected. Then there exists an non-negative constant $\rho < 1$ such that $e(t) \to 0$ as fast as $\rho^t \to 0$

In order to prove Proposition 1, let I denote the $mn \times mn$ identity matrix. From (2) and (3),

$$e(t+1) = (I - Q)e(t), \quad (4)$$

where

$$Q = \bar{H}' P \bar{D}^{-1} \bar{H}. \quad (5)$$

Note that

$$P_i(\frac{1}{d_i} \otimes I_n) = (\frac{1}{\sqrt{d_i}} \otimes I_n) P_i (\frac{1}{\sqrt{d_i}} \otimes I_n),$$

thus

$$P\bar{D}^{-1} = \bar{D}^{-\frac{1}{2}} P \bar{D}^{-\frac{1}{2}}. \quad (6)$$

Since $P^2 = P$,

$$P\bar{D}^{-1} = \bar{D}^{-\frac{1}{2}} P P \bar{D}^{-\frac{1}{2}}.$$

It follows that

$$Q = (P\bar{D}^{-\frac{1}{2}}\bar{H})'(P\bar{D}^{-\frac{1}{2}}\bar{H}). \quad (7)$$

Thus Q is symmetric and [1]

$$Q \geq 0.$$

More is true:

Lemma 2:

$$Q < 2I. \quad (8)$$

and

Lemma 3:

$$\ker Q = \ker \bar{H}. \quad (9)$$

Proofs of Lemma 2 and Lemma 3 will be given in the appendix. In the following, we'll explain why Proposition 1 is true by considering the case $m = n - 1$ and the case $m \geq n$ separately.

A. $m = n - 1$

In the case that $m = n - 1$, H is an $n \times (n-1)$ matrix. Since $\mathbb{G}$ is connected, $\text{rank}\, H = n - 1$. Then $\ker H = 0$. It follows that $\ker \bar{H} = 0$. By Lemma 3, one has $\ker Q = 0$, which together with $Q \geq 0$ implies

$$Q > 0 \quad (10)$$

By (8) and (10), one has

$$0 < Q < 2I$$

then

$$-I < I - Q < I \quad (11)$$

Since Q is symmetric, then

$$||I - Q|| < 1$$

By (4) one has $e(t) \to 0$ as fast as $||I - Q||^t \to 0$. Thus Proposition 1 is true in the case that $m = n - 1$.

[1]By $B \geq A$ we mean $B - A$ is positive semi-definite; similarly $B > A$ means $B - A$ is positive definite.

B. $m \geq n$

In the case that $m \geq n$, one has $\ker H \neq 0$ and then $\ker Q \neq 0$ by Lemma 3. Then (10) can not be obtained as in the case $m = n-1$. In order to prove Proposition 1 is still true, we introduce a matrix J whose columns form a basis for $\ker Q$. Then $QJ = 0$. Since $\ker Q = \ker \bar{H}$,

$$J'\bar{H}' = 0 \tag{12}$$

Let

$$S = J(J'J)^{-1}J' \tag{13}$$

From $e(t) = \bar{H}'x(t)$, (12) and (13), one has

$$Se(t) = 0 \tag{14}$$

which together with (4) implies

$$e(t+1) = (I - Q - S)e(t) \tag{15}$$

Before proceeding on, we need the following lemmas, whose proofs will be given in the Appendix.

Lemma 4: Let $\lambda_{\max}(\cdot)$ denote the largest eigenvalue of a matrix whose eigenvalues are all real. Then

$$\lambda_{\max}(Q+S) \leq \max\{\lambda_{\max}(Q), \quad \lambda_{\max}(S)\}. \tag{16}$$

Lemma 5:

$$\ker Q \cap \ker S = 0. \tag{17}$$

Now we are ready to explain why Proposition 1 is true in the case that $m \geq n$. Since $Q \geq 0$ and $S \geq 0$,

$$Q + S \geq 0.$$

Suppose 0 is an eigenvalue of $Q+S$. If q is a corresponding eigenvector. Then $q \neq 0$ and

$$q'Qq + q'Sq = 0$$

Note that $Q \geq 0$ and $S \geq 0$, one has

$$Qq = 0, \quad Sq = 0$$

which contradict to Lemma 5. Therefore 0 is not an eigenvalue of $Q+S$. It follows that

$$Q + S > 0 \tag{18}$$

By (13) one has S is an orthogonal projection matrix. Then

$$\lambda_{\max}(S) \leq 1$$

which together with $Q < 2I$ by Lemma 2 and (16) implies

$$\lambda_{\max}(Q+S) < 2$$

Thus

$$Q + S < 2I \tag{19}$$

From (18) and (19) one has

$$-I < I - Q - S < I$$

Note that $I - Q - S$ is symmetric, then

$$||I - Q - S|| < 1$$

Then $e(t) \to 0$ as fast as $||I-Q-S||^t \to 0$. Thus Proposition 1 is true in the case that $m \geq n$.

It follows from Proposition 1 and Lemma 1 that Theorem 1 is true.

III. Concluding Remarks

In this paper we have described a simple distributed algorithm for solving a linear algebraic equation over a fixed network. Recently we have been able to show that the same algorithm also works even if $\mathbb{G}$ changes with time and state updating occurs asynchronously [14], [15].

The central idea used in this paper has been to restrict each agent's state updating to points in its own solution space and, in addition, to use distributed averaging to ultimately reach a consensus. This underlying idea has been used with success in a slightly different way to address a family of distributed optimization problems [9]. It is natural wonder how much further the idea can be pushed, and also if there is another standard distributed algorithm such as sorting or ordering which might be used in place of averaging to address still other distributed problems.

IV. Appendix

Proof of Lemma 1: We first prove

$$x_i(t) \to A^{-1}b, \quad i = 1, 2, ..., n \tag{20}$$

From $e(t) \to 0$ and (3), there exists a constant vector $q \in \mathbb{R}^{n^2}$ such that

$$x(t) \to q \tag{21}$$

From $e(t) = \bar{H}'x(t)$ and $e(t) \to 0$, one has

$$\bar{H}'q = 0 \tag{22}$$

From $\bar{H}' = H' \otimes I_n$, $\ker H' = \text{span}\ \{\mathbf{1}\}$ and (22), there must exist a constant vector $\bar{x} \in \mathbb{R}^n$ such that

$$q = \text{column}\ \{\bar{x}, \bar{x}, ..., \bar{x}\}$$

Then

$$x_i(t) \to \bar{x}, \quad i = 1, 2, ..., n$$

From (1), $A_i x_i(1) = b_i$ and $A_i P_i = 0$, one has $A_i x_i(t) = b_i$. Then

$$A_i \bar{x} = b_i, \quad i = 1, 2, ..., n$$

Then $\bar{x}$ is such that $A\bar{x} = b$. Since A is non-singular, one has $\bar{x} = A^{-1}b$. Then (20) is true.

Second, we show that the convergence of (20) is as fast as $\rho^t \to 0$. By (3),

$$\lim_{\tau \to \infty} x(\tau) - x(t) = -P\bar{D}^{-1}\bar{H}\sum_{\tau=t}^{\infty} e(\tau)$$

It follows that

$$||x(t) - q|| = ||P\bar{D}^{-1}\bar{H}\sum_{\tau=t}^{\infty} e(\tau)||$$

Since $e(t) \to 0$ as fast as $\rho^t \to 0$, there exists a non-negative constant c such that $||e(t)|| \leq c\rho^t$. Then

$$||x(t) - q|| \leq c||P\bar{D}^{-1}\bar{H}||\sum_{\tau=t}^{\infty} \rho^{\tau}$$

Note that $0 \leq \rho < 1$, then

$$||x(t) - q|| \leq c||P\bar{D}^{-1}\bar{H}||(1-\rho)^{-1}\rho^t.$$

Then $x(t) \to q$ as fast as $\rho^t \to 0$. Therefore each $x_i(t) \to A^{-1}b$ as fast as $\rho^t \to 0$. We complete the proof . ■

To Prove Lemma 2, one needs the following two lemmas

Lemma 6: Let $L = D^{-\frac{1}{2}}H'HD^{-\frac{1}{2}}$. Then

$$||L|| < 2.$$

Proof of Lemma 6: Note that L is positive semi-definite. Then all eigenvalues of L are non-negative. Moreover, by Rayleigh-Ritz Theorem one has

$$\lambda_{\max}(L) = \max_{q \neq 0, q \in \mathbb{R}^n} \frac{q'Lq}{q'q}.$$

Let $p = D^{-\frac{1}{2}}q$, where the ith element of p is denoted by p_i. Let $\mathcal{E}$ denote the edge set of $\mathbb{G}$. Then

$$\frac{q'Lq}{q'q} = \frac{\sum_{(i,j)\in\mathcal{E}}(p_i - p_j)^2}{\sum_{i=1}^n d_i p_i^2}$$

which reaches its largest value when $p_i = -p_j$ for each $(i,j) \in \mathcal{E}$ as shown in [16]. Since $\mathbb{G}$ is connected, one has all the $p_i^2, i = 1, 2, ..., n$ are equal and non-zero. Then

$$\frac{\sum_{(i,j)\in\mathcal{E}}(p_i - p_j)^2}{\sum_{i=1}^n d_i p_i^2} \leq \frac{4m}{n+2m} < 2$$

Then

$$\lambda_{\max}(L) < 2$$

Note that L is positive-semidefinite, then $||L|| < 2$. ■

Lemma 7: Let A be any $m \times n$ matrix and B be any $n \times m$ matrix. Then the non-zero eigenvalues of AB are the same as the non-zero eigenvalues of BA.

Proof of Lemma 7: Let λ be an non-zero eigenvalue of AB. Then $\det(I_m - \frac{1}{\lambda}AB) = 0$. By Sylvester's determinant theorem, one has $\det(I_n - \frac{1}{\lambda}BA) = 0$. Then λ is also an non-zero eigenvalue of BA. Similarly, one could show that an non-zero eigenvalue of BA is also an eigenvalue of AB. We complete the proof. ■

Proof of Lemma 2: From (5) and (6),

$$Q = \bar{H}'\bar{D}^{-\frac{1}{2}}P\bar{D}^{-\frac{1}{2}}\bar{H}$$

Let $\bar{L} = L \otimes I_n$, where

$$L = D^{-\frac{1}{2}}HH'D^{-\frac{1}{2}}$$

By Lemma 7, the non-zero eigenvalues of Q are the same as the non-zero eigenvalues of $P\bar{L}$. Since Q is positive semi-definite, all the eigenvalues of $P\bar{L}$ are real and non-negative. Then

$$\lambda_{\max}(P\bar{L}) \leq ||P\bar{L}|| \leq ||P||||\bar{L}||$$

Note that $||\bar{L}|| = ||L||$ and $\lambda_{\max}(Q) = \lambda_{\max}(P\bar{L})$. Then

$$\lambda_{\max}(Q) \leq ||P||||L|| \tag{23}$$

which together with $||L|| < 2$ by Lemma 6 and $||P|| \leq 1$ implies

$$\lambda_{\max}(Q) < 2.$$

Since Q is positive semi-definite, one has $Q < 2I$. We complete the proof. ■

Proof of Lemma 3: Since $Q = \bar{H}'P\bar{D}^{-1}\bar{H}$, then to prove $\ker Q = \ker \bar{H}$, it's sufficient to prove the following two equations:

$$\ker \bar{H}' \cap \text{Im } (P\bar{D}^{-1}) = 0 \tag{24}$$

$$\text{Im } \bar{H} \cap \ker(P\bar{D}^{-1}) = 0 \tag{25}$$

We first prove (24). Let u be any vector such that $u \in \ker \bar{H}' \cap \text{Im } (P\bar{D}^{-1})$. From $\ker H' = \text{span } \{\mathbf{1}\}$, $\bar{H} = H \otimes I_n$ and $u \in \ker \bar{H}'$, there exists a vector $q \in \mathbb{R}^n$ such that

$$u = \text{column } \{q, q, ..., q\}$$

In the other hand, since $u \in \text{Im } (P\bar{D}^{-1})$, there must exist $v_i \in \mathbb{R}^n, i = 1, 2, ..., n$ such that

$$u = P\bar{D}^{-1}\text{column } \{v_1, v_2, ..., v_n\}$$

Then

$$\frac{1}{d_i}P_i v_i = q, \; i = 1, 2, ..., n \tag{26}$$

Note that A is non-singular, that is, $A_1, A_2, \cdots, A_n$ are linearly independent, one has $\cap_{i=1}^n \ker A_i = 0$. It follows that

$$\bigcap_{i=1}^n \text{Im } P_i = 0 \tag{27}$$

From (26) and (27), one has $q = 0$ and then $u = 0$. Thus (24) is true.

By (24), one has

$$\text{Im } \bar{H} \cup \ker(P\bar{D}^{-1}) = \mathbb{R}^{n^2} \tag{28}$$

From rank $\bar{H} = n(n-1)$, rank $(P\bar{D}^{-1}) = n(n-1)$ and $P\bar{D}^{-1} \in \mathbb{R}^{n^2 \times n^2}$, one has

$$\dim(\text{Im } \bar{H}) + \dim(\ker(P\bar{D}^{-1})) = n^2 \tag{29}$$

It follows from (28) and (29) that (25) is true. We complete the proof. ■

Proof of Lemma 4: Let λ denote the largest eigenvalue of $Q+S$ with the corresponding eigenvector q. Then $q \neq 0$ and

$$(Q+S)q = \lambda q \tag{30}$$

If $Qq = 0$, one has

$$Sq = \lambda q$$

which implies that λ is an eigenvalue of S; Otherwise $Qq \neq 0$, one multiplies Q to both sides of (30) and gets

$$Q(Q+S)q = \lambda Qq$$

Note that $QJ = 0$, then $QS = 0$. It follows that

$$Q(Qq) = \lambda(Qq).$$

Note that $Qq \neq 0$. Then λ is an eigenvalue of Q. To sum up, the largest eigenvalue of $Q+S$ is either an eigenvalue of S or an eigenvalue of Q. Then Lemma 4 holds. ■

Proof of Lemma 5: Note that $S = J(J'J)^{-1}J'$, one has

$$\ker J' = \text{Im }(I-S), \quad \ker S = \text{Im }(I-S)$$

Then

$$\ker S = \ker J' \tag{31}$$

From $\ker Q = \text{Im } J$ and $\ker J' \cap \text{Im } J = 0$, one has

$$\ker Q \cap \ker J' = 0 \tag{32}$$

By (31) and (32), one has $\ker Q \cap \ker S = 0$. We complete the proof. ■

REFERENCES

[1] R. Mehmood and J. Crowcroft. Parallel iterative solution method of large sparse linear equation systems. *Techinical Report, University of Cambridge*, 2005.

[2] C. K. Koc, A. Guvenc, and B. Bakkagoglu. Exact solution of linear eqauations on distributed memory processors. *In Proc. 14th IMACS World Congress on Computational and Applied Mathematics*, pages 1339–1341.

[3] C. Anderson. Solving linear eqauations on parallel distributed memory architectures by extrapolation. *Technical Report, Royal Institute of Technology*, 1997.

[4] D. M. Young. Iterative methods for solving partial differentce equations of elliptical type. *Ph D Thesis, Harvard University*, 1950.

[5] O. Axelsson. Iterative solution methods. *Cambridge University Press*, 1996.

[6] S. Kaczmarz. Approximate solution of systems of linear equations. *International Journal of Control*, 57:1269–1271, 1993.

[7] J. Lu and Y. Tang. Distributed asynchronous algorithms for solving positive definite linear equations over networks-part i: Agent networks. *Proc. First IFAC Workshop on Estimation and Control of Network Systems*, pages 22–26, 2009.

[8] J. Lu and Y. Tang. Distributed asynchronous algorithms for solving positive definite linear equations over networks-part ii: Wireless networks. *Proc. First IFAC Workshop on Estimation and Control of Network Systems*, pages 258–264, 2009.

[9] A. Nedic and A. Ozdaglar. Distributed sub-gradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, pages 48–61, 2009.

[10] 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 Optimization*, pages 575–600, February 2008.

[11] J. N. Tsitsiklis. *Problems in Decentralized Decision Making and Computation*. PhD thesis, MIT, November 1984.

[12] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 46(6):988–1001, june 2003. also in Proc. 2002 IEEE CDC, pages 2953 - 2958.

[13] V. D. Blondel, J. M. Hendrichkx, A. Olshevsky, and J. N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Proc. of the 44th IEEE Conference on Decision and control*, pages 2996–3000, 2005.

[14] S. Mou, J. Liu, and A. S. Morse. A distributed algorithm for solving a linear algebraic equation. *In Proc. of 52nd IEEE Conference on Decision and Control, Submitted*, 2013.

[15] S. Mou, J. Liu, and A. S. Morse. A distributed algorithm for solving a linear algebraic equation. *IEEE Trans. Automatic Control. Submitted*, 2013.

[16] F. R. K. Chung. Spectral graph theory. *American Mathematical Society*, 1997.