CONTINUOUS-TIME DISTRIBUTED ESTIMATION WITH ASYMMETRIC MIXING

Vítor H. Nascimento† *and* *Ali H. Sayed*‡

†Dept. of Electronic Systems Eng.
Univ. of São Paulo, Brazil
vitor@lps.usp.br

‡Dept. of Electrical Engineering
Univ. of California, Los Angeles
sayed@ee.ucla.edu

ABSTRACT

Discrete-time mobile adaptive networks have been successfully used to model self-organization in biological networks. We recently introduced a continuous-time adaptive diffusion strategy with the goal of better modeling physical phenomena governed by continuous-time dynamics. In the present paper we extend our previous work, proposing a new continuous-time diffusion estimation strategy that allows asymmetric mixing matrices. We prove that the new algorithm is stable and has better convergence properties than stand-alone learning for the case of doubly-stochastic mixing matrices.

1. INTRODUCTION

Distributed processing over networks enables the solution of control, estimation and inference tasks in a decentralized manner by relying on in-network localized processing (see, e.g., [1–6]. In [3,5], diffusion adaptation strategies were proposed and studied for the solution of distributed optimization problems; these strategies permit real-time adaptation and learning over static or mobile networks. For example, in [7,8], discrete-time adaptive diffusion methods were used to model the behavior of complex patterns that arise in biological networks (such as fish schooling and bird formations).

We recently introduced a continuous-time adaptive diffusion strategy [9], with the goal of better modeling physical phenomena governed by continuous-time dynamics. Continuous-time diffusion strategies help provide more accurate models for complex systems with large variations in their time constants, and for networks in which the exchange of information between nodes may happen at any instant.

There has been extensive work in the literature on continuous-time consensus strategies for distributed processing; a useful survey article on such techniques is [10]. However, instead of using the error between a desired input and the current local estimate, as in (4a) and (12) below, the consensus methods described in [10] have either no input, or a biasing input that does not depend on the current weight estimates. The adaptive strategy proposed in this paper has similarities to the distributed Kalman filter proposed in [11]. However, there are three main differences: first, the strategy proposed further ahead is adaptive in nature in that the regression data are generally stochastic while the filter of [11] is attached to a deterministic state-space model and requires the solution of a Riccati differential equation. Second, the stability proof in [11] depends on properties of the Kalman Riccati equation, which do not apply in our case. Third, the filter proposed in [11] requires equal weights for all nodes in a neighborhood, while our approach allows for different and asymmetric weights.

We proved stability of our continuous-time diffusion strategy in [9], under the condition that the mixing matrix (the matrix formed by the weights describing the information sharing between nodes) is symmetric and positive-definite. This imposes an important restriction, since in many natural networks the mixing matrix is asymmetric (for example, in the bird flight model of [7], each bird may only use information from the bird in front of it). In this paper, we show that the positive-definiteness condition imposed in our previous paper is indeed necessary for the stability of our original strategy. We then show how to modify the original strategy to allow asymmetric and not necessarily positive-definite mixing matrices. We also establish the stability of the new method.

2. CONTINUOUS-TIME DIFFUSION LMS

In diffusion learning, a network of N nodes senses the environment and collaboratively estimates a vector of unknown parameters. We assume that each node has access to a scalar measurement $d_k(t)$ and a regressor vector $\boldsymbol{u}_k(t) \in \mathbb{R}^M$, related through the linear model

$$d_k(t) = \boldsymbol{u}_k^T(t)\boldsymbol{w}_\mathrm{o} + v_k(t), \tag{1}$$

where $\boldsymbol{w}_\mathrm{o}$ is a parameter vector and $v_k(t)$ is noise.

The continuous-time diffusion LMS proposed in [9] allows the network to collaboratively estimate $\boldsymbol{w}_\mathrm{o}$ as follows. Each node k computes two estimates for $\boldsymbol{w}_\mathrm{o}$, a local estimate $\boldsymbol{w}_k(t)$ and a local mixture $\boldsymbol{\psi}_k(t)$. The latter is the average of the local estimates of the neighborhood $\mathcal{N}_k$ of node k, i.e.,

$$\boldsymbol{\psi}_k(t) = \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \boldsymbol{w}_\ell(t), \tag{2}$$

where $\{a_{\ell k}\}$, $1 \le \ell, k \le N$ is a set of non-negative weights such that $a_{\ell k} = 0$ if $\ell \notin \mathcal{N}_k$, and

$$\sum_{\ell=1}^{N} a_{\ell k} = 1. \tag{3}$$

The work of V. H. Nascimento was partly supported by CNPq and FAPESP grants 303921/2010-2 and 2011/06994-2. The work of A. H. Sayed was supported in part by NSF grants CCF-1011918 and CCF-0942936.

978-1-4673-0183-1/12/$31.00 ©2012 IEEE

That is, expression (2) represents a convex combination of the estimates $\boldsymbol{w}_\ell(t)$ from the neighbors of node k. This requirement guarantees that the estimates remain unbiased [9]. Figure 1 shows an example of a 6-node network. Although not shown in the figure, connections a_{kk} from a node to itself are also allowed.

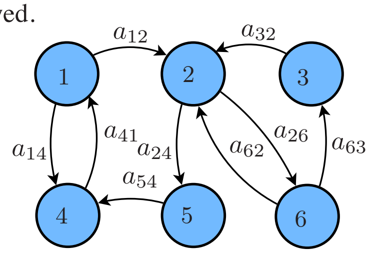

Fig. 1. 6-node network. Self-connections are not drawn.

The local estimates are obtained from a differential equation linking the input data, the local estimate $\boldsymbol{w}_k(t)$, and the local mixture $\boldsymbol{\psi}_k(t)$, as follows [9]:

$$e_k(t) = d_k(t) - \boldsymbol{u}_k^T(t)\boldsymbol{\psi}_k(t), \tag{4a}$$

$$\dot{\boldsymbol{w}}_k(t) = -\gamma_0\left(\boldsymbol{w}_k(t) - \boldsymbol{\psi}_k(t)\right) + \gamma_k e_k(t)\boldsymbol{u}_k(t), \tag{4b}$$

where γ_0 and γ_k are positive constants.

The stability of this system of connected differential equations is more easily studied if we consider a global model. Define the global weight vector:

$$\boldsymbol{w}(t) = \operatorname{col}\{\boldsymbol{w}_1(t), \boldsymbol{w}_2(t), \dots, \boldsymbol{w}_N(t)\},$$

the mixing matrix $\boldsymbol{A} = \begin{bmatrix} a_{ij} \end{bmatrix}$, $i, j = 1 \dots N$, the input vector $\boldsymbol{d}(t) = \begin{bmatrix} d_1(t) & \dots & d_N(t) \end{bmatrix}^T$ (similarly for the noise vector $\boldsymbol{v}(t)$), and the regressor matrix

$$\boldsymbol{U}(t) = \operatorname{diag}\{\boldsymbol{u}_1(t), \boldsymbol{u}_2(t), \dots, \boldsymbol{u}_N(t)\}, \tag{5}$$

where $\boldsymbol{0}_M$ is an $M \times 1$ null vector. Let $\mathbb{1}$ denote an N-dimensional vector of ones, that is, $\mathbb{1} = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}^T$. Then the convexity constraint (3) corresponds to requiring that

$$\boldsymbol{A}^T\mathbb{1} = \mathbb{1}, \tag{6}$$

i.e., $\mathbb{1}$ is an eigenvector of $\boldsymbol{A}^T$ relative to the eigenvalue 1.

In terms of the global weight vector $\boldsymbol{w}(t)$, the continuous-time diffusion LMS algorithm(2) and (4a)-(4b) is described by the following differential equations:

$$\dot{\boldsymbol{w}} = -\gamma_0\left(\boldsymbol{I}_{MN} - \boldsymbol{A}^T \otimes \boldsymbol{I}_M\right)\boldsymbol{w}(t) + \boldsymbol{U}(t)\boldsymbol{\Gamma}\boldsymbol{e}(t), \tag{7a}$$

$$\boldsymbol{e}(t) = \boldsymbol{d}(t) - \boldsymbol{U}^T(t)\left(\boldsymbol{A}^T \otimes \boldsymbol{I}_M\right)\boldsymbol{w}(t), \tag{7b}$$

where $\boldsymbol{\Gamma} = \operatorname{diag}\{\gamma_k\}$, $\boldsymbol{I}_M$ represents the $M \times M$ identity matrix, and $\otimes$ denotes the Kronecker product. With these definitions, the linear model (1) can be rewritten as

$$\begin{aligned} \boldsymbol{d}(t) &= \boldsymbol{U}^T(t)\left(\mathbb{1} \otimes \boldsymbol{w}_\text{o}\right) + \boldsymbol{v}(t) \\ &= \boldsymbol{U}^T(t)\left(\boldsymbol{A}^T \otimes \boldsymbol{I}_M\right)\left(\mathbb{1} \otimes \boldsymbol{w}_\text{o}\right) + \boldsymbol{v}(t), \end{aligned} \tag{8}$$

where we used (6) and the fact that, for any matrices of appropriate dimensions, $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{AC}) \otimes (\boldsymbol{BD})$ [12]. Define the weight error vector $\tilde{\boldsymbol{w}}(n) \triangleq \mathbb{1} \otimes \boldsymbol{w}_\text{o} - \boldsymbol{w}(t)$. We can write the error vector $\boldsymbol{e}(t)$ in terms of $\tilde{\boldsymbol{w}}(t)$ as below:

$$\boldsymbol{e}(t) = \boldsymbol{U}^T(t)\left(\boldsymbol{I}_{MN} - \boldsymbol{A}^T \otimes \boldsymbol{I}_M\right)\tilde{\boldsymbol{w}}(t) + \boldsymbol{v}(t). \tag{9}$$

From (6), it follows that $(\boldsymbol{I}_{MN} - \boldsymbol{A}^T \otimes \boldsymbol{I}_M)(\mathbb{1} \otimes \boldsymbol{w}_\text{o}) = \boldsymbol{0}_{MN}$. Noting also that $\dot{\tilde{\boldsymbol{w}}}(t) = -\dot{\boldsymbol{w}}(t)$, expression (7a) can be rewritten in terms of $\tilde{\boldsymbol{w}}(t)$:

$$\dot{\tilde{\boldsymbol{w}}}(t) = \boldsymbol{B}(t)\tilde{\boldsymbol{w}}(t) - \boldsymbol{U}(t)\boldsymbol{\Gamma}\boldsymbol{v}(t), \tag{10}$$

with

$$\begin{aligned} \boldsymbol{B}(t) \triangleq& -\gamma_0\left(\boldsymbol{I}_{MN} - \boldsymbol{A}^T \otimes \boldsymbol{I}_M\right) \\ &- \boldsymbol{U}(t)\boldsymbol{\Gamma}\boldsymbol{U}^T(t)\left(\boldsymbol{A}^T \otimes \boldsymbol{I}_M\right). \end{aligned} \tag{11}$$

In [9], we showed that $\tilde{\boldsymbol{w}} = \boldsymbol{0}_{MN}$ is a stable equilibrium point of (10) if the noise is identically zero and the mixing matrix $\boldsymbol{A}$ is symmetric and positive-definite, for any $\gamma_n > 0$ (stability under non-zero noise requires a persistence of excitation condition on $\boldsymbol{U}(t)$ — see Section 3).

The positive-definiteness restriction on $\boldsymbol{A}$ is necessary: if $\boldsymbol{A}$ is not positive-definite, a large $\boldsymbol{U}(t)$ might make $\tilde{\boldsymbol{w}}(t)$ diverge. For example, consider a network with $N = 2$, $M = 1$, $u_1(t) = u_2(t) \equiv 2$, and

$$\boldsymbol{A} = \begin{bmatrix} 0.2 & 0.8 \\ 0.8 & 0.2 \end{bmatrix}, \quad \text{for which} \quad \boldsymbol{B}(t) \equiv \boldsymbol{B} = \begin{bmatrix} -1.6 & -2.4 \\ -2.4 & -1.6 \end{bmatrix},$$

with eigenvalues -4 and $+0.8$. Since $\boldsymbol{B}$ is constant, the positive eigenvalue means (10) will be unstable. In the next section, we show how the update law (7) can be modified to avoid the restrictions of symmetry and positive-definiteness on $\boldsymbol{A}$.

3. LMS DIFFUSION WITH ASYMMETRIC MIXING

A close examination of the previous example suggests that removing the factor $\boldsymbol{A}^T \otimes \boldsymbol{I}_M$ multiplying the input term $\boldsymbol{U}(t)\boldsymbol{\Gamma}\boldsymbol{U}^T(t)$ in (11) should avoid instability. This amounts to modifying the error definition (4a) to

$$e_k(t) = d_k(t) - \boldsymbol{u}_k^T(t)\boldsymbol{w}_k(t), \tag{12}$$

using $\boldsymbol{w}_k(t)$ instead of $\boldsymbol{\psi}_k(t)$ to compute the local approximation to $d_k(t)$. It can be checked that this modification guarantees stability for any constant inputs in the previous example. In the remainder of this section, we study general conditions on $\boldsymbol{A}$ so that the modified diffusion algorithm will remain stable for any inputs $\boldsymbol{u}_k(t)$ and any value of M. With the modified error (12), the error equation is now given by

$$\begin{aligned} \dot{\tilde{\boldsymbol{w}}}(t) = &-\left[\gamma_0\left(\boldsymbol{I}_{MN} - \boldsymbol{A}^T \otimes \boldsymbol{I}_M\right) + \boldsymbol{U}(t)\boldsymbol{\Gamma}\boldsymbol{U}^T(t)\right]\tilde{\boldsymbol{w}}(t) \\ &- \boldsymbol{U}(t)\boldsymbol{\Gamma}\boldsymbol{v}(t). \end{aligned} \tag{13}$$

Equation (13) describes a linear time-varying system of differential equations, with system matrix:

$$\boldsymbol{D}(t) = \underbrace{-\gamma_0\left(\boldsymbol{I}_{MN} - \boldsymbol{A}^T \otimes \boldsymbol{I}_M\right)}_{\triangleq \boldsymbol{D}_1} \underbrace{-\boldsymbol{U}(t)\boldsymbol{\Gamma}\boldsymbol{U}^T(t)}_{\triangleq \boldsymbol{D}_2(t)}.$$

Requiring the eigenvalues of $\boldsymbol{D}(t)$ to have real parts less than a negative constant does not imply stability [13]. Thus we search for a Lyapunov function to study the stability of (13).

Since $\boldsymbol{A}$ is a left-stochastic matrix (i.e., all its terms are nonnegative and $\boldsymbol{A}^T\mathbb{1} = \mathbb{1}$), it has an eigenvalue at 1, and all other eigenvalues λ_i of $\boldsymbol{A}$ satisfy $|\lambda_i| \le 1$ [14]. This means that $\boldsymbol{D}_1$ is singular, and its eigenvalues have absolute values bounded by $2\gamma_0$. On the other hand, $\boldsymbol{D}_2(t)$ is positive semi-definite.

We shall use the following candidate Lyapunov function

$$V(t) = \tilde{\boldsymbol{w}}^T(t)\tilde{\boldsymbol{w}}(t). \qquad (14)$$

Its derivative (assuming $\boldsymbol{v}(t) \equiv \boldsymbol{0}_{MN}$ for now) is

$$\begin{aligned}\dot{V}(t) &= \tilde{\boldsymbol{w}}^T(t)\left[\boldsymbol{D}^T(t) + \boldsymbol{D}(t)\right]\tilde{\boldsymbol{w}}(t)\\ &= -\tilde{\boldsymbol{w}}^T(t)\left[\gamma_0\left(2\boldsymbol{I}_{MN} - \boldsymbol{A}\otimes\boldsymbol{I}_M - \boldsymbol{A}^T\otimes\boldsymbol{I}_M\right)\right.\\ &\quad \left.+2\boldsymbol{U}(t)\boldsymbol{\Gamma}\boldsymbol{U}^T(t)\right]\tilde{\boldsymbol{w}}(t) \triangleq -\tilde{\boldsymbol{w}}^T(t)\boldsymbol{P}(t)\tilde{\boldsymbol{w}}(t), \qquad (15)\end{aligned}$$

where we used the fact that $(\boldsymbol{A}\otimes\boldsymbol{B})^T = \boldsymbol{A}^T \otimes \boldsymbol{B}^T$, for any matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ [12].

The origin in (13) is stable if $\dot{V}(t)$ is nonpositive, or equivalently, if $\boldsymbol{P}(t)$ is positive semi-definite. We show next that this happens when $\boldsymbol{A}$ is doubly stochastic, that is, if in addition to (6), it also holds that $\boldsymbol{A}\mathbb{1} = \mathbb{1}$.

When $\boldsymbol{A}$ is doubly stochastic, the matrix $\bar{\boldsymbol{A}} \triangleq \boldsymbol{A}^T\otimes\boldsymbol{I}_M + \boldsymbol{A}\otimes\boldsymbol{I}_M$ satisfies $\bar{\boldsymbol{A}}(\mathbb{1}\otimes\mathbb{1}) = 2(\mathbb{1}\otimes\mathbb{1})$. Since all entries of $\bar{\boldsymbol{A}}$ are nonnegative, for any positive ϵ, $\boldsymbol{Q}(\epsilon) \triangleq (2+\epsilon)\boldsymbol{I}_{MN} - \bar{\boldsymbol{A}}$ is symmetric, has positive entries on the diagonal, and negative entries outside. Since $\boldsymbol{Q}(\epsilon)(\mathbb{1}\otimes\mathbb{1}) = \epsilon(\mathbb{1}\otimes\mathbb{1})$, we conclude that $\boldsymbol{Q}(\epsilon)$ is strictly diagonally dominant with positive diagonal entries, that is, its entries $q_{ij}(\epsilon)$ satisfy

$$q_{ii}(\epsilon) > \sum_{j=1, j\neq i}^{MN} |q_{ij}(\epsilon)|, \text{ for all } 1 \le i \le MN.$$

It follows that $\boldsymbol{Q}(\epsilon)$ is positive-definite for all $\epsilon > 0$ [14]. Since the eigenvalues of a matrix are continuous functions of its entries, this implies that $\boldsymbol{Q}(0) = 2\boldsymbol{I}_{MN} - \bar{\boldsymbol{A}}$ is positive semi-definite. Noting that $2\boldsymbol{U}(t)\boldsymbol{\Gamma}\boldsymbol{U}^T(t)$ is also positive semi-definite, we conclude that $\boldsymbol{P}(t)$ is positive semi-definite, and thus the origin in (13) is a stable equilibrium point in the absence of noise, irrespective of the inputs $\boldsymbol{u}_k(t)$.

If, in addition, there exist constants $0 < \alpha_1 \le \alpha_2 < \infty$ such that

$$\alpha_1\boldsymbol{I}_{MN} \le \int_{t_0}^{t_0+T} \boldsymbol{P}(t)dt \le \alpha_2\boldsymbol{I}_{MN} \qquad (16)$$

for all t_0, it can be shown that $\tilde{\boldsymbol{w}} = \boldsymbol{0}_{MN}$ in (13) is exponentially stable (this proof follows the same steps as in the analysis of the stand-alone continuous-time LMS algorithm [15]). Using the total stability theorem [16, Lemma 5.2], it then follows that $\tilde{\boldsymbol{w}}(t)$ will remain bounded in the presence of sufficiently small bounded noise.

An interesting property of the new continuous-time diffusion strategy(12)+(4b) proposed here is that it always improves stability, compared to a collection of stand-alone filters. Indeed, for stand-alone filters (no diffusion) we have $\boldsymbol{A} = \boldsymbol{I}_N$ and condition (16) reduces to

$$\alpha_1\boldsymbol{I}_M \le \int_{t_0}^{t_0+T} \gamma_k\boldsymbol{u}_k(t)\boldsymbol{u}_k^T(t)dt \le \alpha_2\boldsymbol{I}_M, \qquad (17)$$

for all k between 1 and N. This is equivalent to requiring that all nodes have persistently exciting inputs $\boldsymbol{u}_k(t)$ as defined in [15]. However, for any doubly stochastic $\boldsymbol{A}$ it holds that

$$\begin{aligned}2\boldsymbol{U}(t)\boldsymbol{\Gamma}\boldsymbol{U}^T(t) \le \gamma_0&\left(2\boldsymbol{I}_{MN} - \boldsymbol{A}\otimes\boldsymbol{I}_M - \boldsymbol{A}^T\otimes\boldsymbol{I}_M\right)\\ &+ 2\boldsymbol{U}(t)\boldsymbol{\Gamma}\boldsymbol{U}^T(t),\end{aligned}$$

and (16) may hold even if (17) does not. In the example of Section 2, if $u_1(t) \equiv 0.1$ and $u_2(t) \equiv 0$, we have

$$\boldsymbol{P}(t) = \begin{bmatrix} 1.61 & -1.60 \\ -1.60 & 1.60 \end{bmatrix} > 0.$$

In this case, (16) holds even though the persistence of excitation condition (17) does not hold for node 2. In the absence of noise, the diffusion algorithm guarantees that both nodes will converge to $\boldsymbol{w}_o$, while in the stand-alone case, node 2 would never leave its initial condition. In the next section we provide a similar example involving a 10-node network.

4. SIMULATIONS

In this section we provide an example of the advantages of the new diffusion learning strategy, using the 10-node network depicted in Figure 2 (weights a_{kk} are not drawn; their values are such that (3) is satisfied). We simulated our new

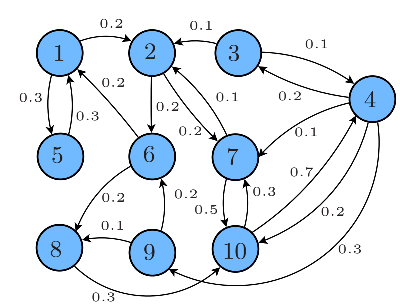

Fig. 2. 10-node network. Self-connections not drawn.

strategy (12)+(4b) with the mixing coefficients as in Figure 2 and in the stand-alone case ($\boldsymbol{A} = \boldsymbol{I}_N$). The regressors $\boldsymbol{u}_k(t)$ were either sinusoids or filtered white noise, chosen so that the persistence of excitation condition (17) is *not* satisfied for any node, although (16) is satisfied for the network (note that, since the vectors $\boldsymbol{u}_i(t)$ do not span all directions in $\mathbb{R}^2$, the stand-alone filters do not converge to a neighborhood of the true weights). For example, for nodes 1 and 6 we have

$$\boldsymbol{u}_1(t) = \begin{bmatrix}\sin(20\pi t) & 0\end{bmatrix}^T, \quad \boldsymbol{u}_6(t) = s_6(t)\begin{bmatrix}1 & 1\end{bmatrix}^T,$$

where $s_6(t)$ is the output of a linear filter with transfer function $H(s) = (s+1)/(s^2+s+2)$, with Simulink's band-limited white noise with power 0.2 and sampling time 0.001 as input. The optimum weight vector is $\boldsymbol{w}_{\rm o} = \begin{bmatrix}0.5 & -0.1\end{bmatrix}^T$.

Figure 3 shows the weight estimates obtained in Simulink with the new diffusion strategy and, for comparison, with stand-alone filters. It can be seen that the stand-alone filters do not converge to the optimum coefficients, while the diffusion strategy is able to correctly identify both weights. The plots for the other nodes are similar.

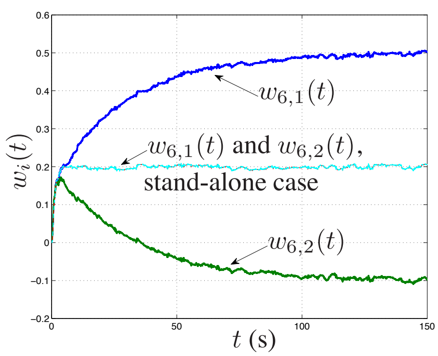

Fig. 3. Parameter estimates $\boldsymbol{w}_6(t)$ for node 6, for the new diffusion strategy and for stand-alone filters.

5. CONCLUSION

We proposed a new continuous-time diffusion LMS algorithm that allows for asymmetric doubly-stochastic mixing matrices. We proved stability and improved performance of the new strategy, compared to stand-alone learning.

6. REFERENCES

[1] W. Ren and R. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655 – 661, May 2005.

[2] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2508 – 2530, Jun. 2006.

[3] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3122–3136, 2008.

[4] T. Aysal, M. Yildiz, A. Sarwate, and A. Scaglione, "Broadcast gossip algorithms for consensus," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2748 –2761, Jul. 2009.

[5] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1035–1048, 2010.

[6] S. Chouvardas, K. Slavakis, and S. Theodoridis, "Adaptive robust distributed learning in diffusion sensor networks," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4692–4707, 2011.

[7] F. S. Cattivelli and A. H. Sayed, "Modeling bird flight formations using diffusion adaptation," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2038–2051, 2011.

[8] S.-Y. Tu and A. H. Sayed, "Mobile adaptive networks," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 4, pp. 649–664, 2011.

[9] V. H. Nascimento and A. H. Sayed, "Continuous-time distributed estimation," in *Proc. of the 45th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2011, pp. 1–5.

[10] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[11] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. IEEE Conference on Decision and Control*, Dec. 2007, pp. 5492 –5498.

[12] A. J. Laub, *Matrix Analysis for Scientists and Engineers*. Society for Industrial and Applied Mathematics, PA, Dec. 2005.

[13] A. Ilchmann, D. Owens, and D. Prätzel-Wolters, "Sufficient conditions for stability of linear time-varying systems," *Systems & Control Letters*, vol. 9, no. 2, pp. 157–163, Aug. 1987.

[14] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1987.

[15] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Prentice Hall, NJ, 1996.

[16] H. K. Khalil, *Nonlinear Systems*, 2nd ed. Prentice Hall, NJ, 1996.