AN ADAPTIVE COMBINATION RULE FOR DIFFUSION LMS BASED ON CONSENSUS PROPAGATION

Ayano Nakai

Graduate School of Informatics,
Kyoto University
Sakyo-ku, Kyoto 606-8501, Japan

Kazunori Hayashi

Graduate School of Engineering,
Osaka City University
Sumiyoshi-ku, Osaka 558-8585, Japan

ABSTRACT

Diffusion least-mean-square (LMS) algorithm is a method that estimates an unknown global vector from its linear measurements obtained at multiple nodes in a network in a distributed manner. This paper proposes a novel combination rule in the algorithm used to integrate the local estimates at each node by using the idea of consensus propagation, which is known to be a fast algorithm to achieve the average consensus. Moreover, we optimize constants involved in the proposed combination rule in terms of the steady state mean-square-deviation (MSD) and show an adaptive combination rule, along with an adaptive implementation. Simulation results demonstrate that the proposed combination scheme achieves better MSD performance than conventional combination schemes.

***Index Terms*—** Diffusion LMS, in-network signal processing, consensus propagation, average consensus, combination weights

1. INTRODUCTION AND RELATED WORK

In-network signal processing, which is a framework of distributed signal processing in networks, has gained much attention recently [1]. When each node in the network needs to track an unknown parameter in real-time, one of the effective methods is diffusion least-mean-square (LMS) algorithm [2]–[6], where each node in the network iteratively updates the estimate by LMS algorithm using its noisy measurements and also by the weighted average of its neighbors' estimates obtained via communications, and finally all nodes in the network obtain the same estimate.

It is known that the choice of the weights used in the combine step of neighbors' estimates has a great impact on the convergence performance of the diffusion LMS. Thus, several combination rules have been proposed in the literature, such as uniform rule [7], maximum degree rule [8], Metropolis rule [9], and relative degree rule [5]. More sophisticated static rules are considered in [5], [10]–[12], which are derived by solving some optimization problems. In particular, a closed-form solution that minimizes the steady state error has been derived in [10] and [11], which is referred to as relative-variance rule. Since the relative-variance rule requires parameters including network statistics such as noise variances at all nodes, which are not available at each node in general, adaptive estimation methods of the parameters have been also proposed in [10] and [11].

In this paper, we propose a novel combination rule by using the idea of consensus propagation [13], which can be viewed as a special case of belief propagation [14]. If the network has a tree structure, all nodes in the network can obtain exact average only with the same number of iterations as the diameter of the tree by using consensus propagation. Based on this fact, we have proposed a diffusion LMS using consensus propagation for the network with the tree structure in [15] in our previous work. In this paper, we extend the method to general networks, which possibly have loops. Since the update rule of consensus propagation in the network with loops involves constants to be determined that control the convergence property, which are known to be difficult to optimize in general, we select values of constants in consensus propagation by minimizing steady state mean-squared-deviation (MSD) of the diffusion LMS as in [10], [11]. Moreover, we further extend the proposed combination rule to an adaptive version, which can be implemented in a fully distributed manner. Simulation results show that the proposed combination scheme achieves better performance with the lower sensitivity to the initial values of adaptation than the conventional combination rules.

2. DIFFUSION STRATEGY

2.1. System Model

Consider a network with N nodes, where each node k can communicate with its neighbors and aims to estimate an unknown deterministic vector of interest $\boldsymbol{w}^{o} \in \mathbb{C}^{M \times 1}$ through

This work was supported in part by the Grants-in-Aid for Scientific Research no. 15K06064 and 15H02252 from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

linear measurements of the form [2]–[6]:

$$d_k^{(i)} = \boldsymbol{u}_k^{(i)\mathrm{H}}\boldsymbol{w}^{\mathrm{o}} + v_k^{(i)}, \quad (1)$$

where $(\cdot)^{\mathrm{H}}$ denotes the Hermitian transpose, i (≥ 0) is a time index, $d_k^{(i)}$ is a scalar measurement, $\boldsymbol{u}_k^{(i)} \in \mathbb{C}^{M\times 1}$ is a random measurement vector, and $v_k^{(i)}$ is a zero-mean additive white Gaussian noise with variance of σ_k^2. The stochastic processes $\{d_k^{(i)}, \boldsymbol{u}_k^{(i)}\}$ are assumed to be jointly wide-sense stationary and zero-mean. For simplicity, we assume that all communications between neighbor nodes are perfect, i.e., we do not consider any communication error.

2.2. Diffusion LMS Algorithm

All nodes in the network estimate $\boldsymbol{w}^{\mathrm{o}}$ by solving the following optimization problem [2]–[6]:

$$\hat{\boldsymbol{w}}^{\mathrm{o}} = \underset{\boldsymbol{w}}{\operatorname{argmin}} \sum_{k=1}^{N} \mathrm{E}[|d_k^{(i)} - \boldsymbol{u}_k^{(i)\mathrm{H}}\boldsymbol{w}|^2], \quad (2)$$

where $\mathrm{E}[\cdot]$ stands for the expectation operator. Diffusion LMS algorithm [3]–[5] is the iterative method to solve this global problem in a distributed manner. Especially, by using $\boldsymbol{\psi}_k^{(i)}$ and $\boldsymbol{\phi}_k^{(i)}$ that denote the estimates of $\boldsymbol{w}^{\mathrm{o}}$ at node k at time i, the Adapt-then-Combine (ATC) version of diffusion LMS update can be described as [5]

$$\boldsymbol{\psi}_k^{(i)} = \boldsymbol{\phi}_k^{(i-1)} + \mu_k \boldsymbol{u}_k^{(i)}(d_k^{(i)} - \boldsymbol{u}_k^{(i)\mathrm{H}}\boldsymbol{\phi}_k^{(i-1)}), \quad (3)$$

$$\boldsymbol{\phi}_k^{(i)} = \sum_{l\in\mathcal{N}_k} a_{lk}\boldsymbol{\psi}_l^{(i)}, \quad (4)$$

where $\boldsymbol{\psi}_k^{(i)}$ is an immediate estimate obtained by LMS update, $\boldsymbol{\phi}_k^{(i)}$ is obtained by the weighted average of its neighbors' estimates with $\boldsymbol{\phi}_k^{(-1)} = 0$, μ_k is the step-size parameter, $\mathcal{N}_k$ is the set of neighbors of node k including k itself, and a_{lk} is a nonnegative combination weight, which is the (l, k) element of an $N \times N$ matrix $\boldsymbol{A}$ that satisfies $\mathbf{1}^{\mathrm{T}}\boldsymbol{A} = \mathbf{1}^{\mathrm{T}}$, where $\mathbf{1}$ denotes a vector whose elements are all 1. Possible choices of the combination weights a_{lk} will be uniform rule [7]

$$a_{lk} = \frac{1}{|\mathcal{N}_k|}, \quad (5)$$

Metropolis rule [9], relative degree rule [5], and so on. Note that the choice of the combination rule has a great impact on the convergence performance of the diffusion LMS [12].

3. PROPOSED DIFFUSION STRATEGY

Here, we propose a novel combination rule inspired by the fast message passing algorithm to achieve average consensus, namely consensus propagation (CP) [13], in order to improve the convergence performance of the diffusion LMS.

3.1. Consensus Propagation

Assume that, in a network composed of N nodes, each node k has an initial state value $x_k \in \mathbb{C}$. The goal of CP is that each node obtains the average $\frac{1}{N}\sum_{k=1}^{N} x_k$, which is called average consensus. CP consists of two types updates, message update between neighbor nodes and state update at each node, to calculate the approximate average using locally available information only. The updates of CP at the j-th iteration are given as follows:

$$K_{(k\to l)}^{[j]} = \frac{1 + \sum_{u\in\mathcal{N}_k\setminus l,k} K_{(u\to k)}^{[j-1]}}{1 + \frac{1}{\beta_k}(1 + \sum_{u\in\mathcal{N}_k\setminus l,k} K_{(u\to k)}^{[j-1]})}, \quad (6)$$

$$\theta_{(k\to l)}^{[j]} = \frac{x_k + \sum_{u\in\mathcal{N}_k\setminus l,k} K_{(u\to k)}^{[j-1]}\theta_{(u\to k)}^{[j-1]}}{1 + \sum_{u\in\mathcal{N}_k\setminus l,k} K_{(u\to k)}^{[j-1]}}, \quad (7)$$

$$x_k^{[j]} = \frac{x_k + \sum_{u\in\mathcal{N}_k\setminus k} K_{(u\to k)}^{[j]}\theta_{(u\to k)}^{[j]}}{1 + \sum_{u\in\mathcal{N}_k\setminus k} K_{(u\to k)}^{[j]}}, \quad (8)$$

where $K_{(k\to l)}^{[0]} = 0$ and β_k is a positive constant. By iterating (6)–(8), $x_k^{[j]}$ will converge to $\frac{1}{N}\sum_{k=1}^{N} x_k$.

Note that CP achieves exact average consensus with the same number of iterations as the diameter of the graph when the network has tree structure and $\beta_k = +\infty$. If the network has loops, however, the convergence behavior has not been fully understood yet because the messages and state values at any iterations cannot be described. Although it is known that β_k in (6) plays an important role to ensure the convergence for the case with loops, to the best of our knowledge, the optimal value of β_k has not been derived.

3.2. Diffusion LMS based on CP

We propose a novel combination rule by applying the first iteration of CP to the combine step of the diffusion LMS algorithm (4). By substituting $\boldsymbol{\psi}_k^{(i)}$ in (4) to x_k, we have

$$K_{(u\to k)}^{[1]} = \frac{\beta_k}{1+\beta_k}, \quad \theta_{(u\to k)}^{[1]} = x_k = \boldsymbol{\psi}_k^{(i)}, \quad (9)$$

and substituting $\boldsymbol{\phi}_k^{(i)}$ in (4) to $x_k^{[1]}$ and (9) to (8), then we have

$$\boldsymbol{\phi}_k^{(i)} = \frac{1+\beta_k}{1+|\mathcal{N}_k|\beta_k}\boldsymbol{\psi}_k^{(i)} + \frac{\beta_k}{1+|\mathcal{N}_k|\beta_k}\sum_{u\in\mathcal{N}_k\setminus k}\boldsymbol{\psi}_u^{(i)}. \quad (10)$$

This can be regarded as a novel combination rule summarized as

$$a_{lk} = \begin{cases} \frac{\beta_k}{1+|\mathcal{N}_k|\beta_k} & \text{if } l \in \mathcal{N}_k \text{ and } l \neq k \\ \frac{1+\beta_k}{1+|\mathcal{N}_k|\beta_k} & \text{if } k = l \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

which satisfies $\mathbf{1}^{\mathrm{T}}\boldsymbol{A} = \mathbf{1}^{\mathrm{T}}$.

Note that the proposed combination rule approaches to the uniform rule (5) as β_k increases.

4. WEIGHT OPTIMIZATION

4.1. Optimal Combination Weight

As mentioned in Sect. 3.1, how to select β_k has been an open issue [13]. In this section, we consider to choose β_k that minimizes mean-square-deviation (MSD) in the steady state of diffusion LMS algorithm, i.e., $\lim_{i\to\infty}\frac{1}{N}\sum_{k=1}^{N}\mathrm{E}[\|\boldsymbol{w}^{\mathrm{o}}-\boldsymbol{\phi}_k^{(i)}\|^2]$. Assuming that step-sizes $\{\mu_k\}$ are sufficiently small and that all measurement vectors $\{\boldsymbol{u}_k^{(i)}\}$ are temporally and spatially independent, it can be shown that the upper bound of the steady state MSD is proportional to [6], [10], [11]

$$\sum_{k=1}^{N}\sum_{l=1}^{N}\gamma_l^2 a_{lk}^2, \tag{12}$$

where $\gamma_l^2=\mu_l^2\sigma_l^2\mathrm{Tr}(\boldsymbol{R}_{u_l})$, $\mathrm{Tr}(\cdot)$ denotes the trace operator, and $\boldsymbol{R}_{u_l}=E[\boldsymbol{u}_l^{(i)}\boldsymbol{u}_l^{(i)\mathrm{H}}]$.

The steady state MSD in (12) has been used to determine $\{a_{lk}\}$ in (4) in the existing works [6], [10], [11] as

$$\begin{aligned}\{a_{lk}^{\mathrm{opt}}\}_{k=1}^{N} \quad &= \quad \underset{\{a_{lk}\}_{k=1}^{N}}{\mathrm{argmin}}\sum_{l=1}^{N}\gamma_l^2 a_{lk}^2, \\ \text{s.t.} \quad & \quad \sum_{l=1}^{N}a_{lk}=1,\ a_{lk}=0 \text{ if } l\notin\mathcal{N}_k.\end{aligned} \tag{13}$$

The solution results in the relative-variance rule [11] given by

$$a_{lk}^{\mathrm{opt}}=\begin{cases}\frac{[\gamma_l^2]^{-1}}{\sum_{m\in\mathcal{N}_k}[\gamma_m^2]^{-1}} & \text{if } l\in\mathcal{N}_k\\ 0 & \text{otherwise.}\end{cases} \tag{14}$$

For the proposed combination rule, the optimization problem to decide $\{\beta_k\}$ is written as

$$\{\beta_k^{\mathrm{opt}}\}_{k=1}^{N}=\quad\underset{\{\beta_k\}_{k=1}^{N}}{\mathrm{argmin}}\sum_{l=1}^{N}\gamma_l^2 a_{lk}^2,\quad \text{s.t. (11).} \tag{15}$$

The shape of the cost function in (15) largely depends on $A_k=\sum_{l\in\mathcal{N}_k}\gamma_l^2-|\mathcal{N}_k|\gamma_k^2\neq 0$. When $A_k>0$, the function has a global minimum in $\beta_k>0$ and the optimal value is obtained as

$$\beta_k^{\min}=\frac{(|\mathcal{N}_k|-1)\gamma_k^2}{A_k}. \tag{16}$$

On the other hand, it becomes monotonically decreasing function when $A_k<0$. Thus, in summary, the optimum β_k is given by

$$\beta_k^{\mathrm{opt}}=\begin{cases}\beta_k^{\min} & \text{if } A_k>0\\ +\infty & \text{otherwise.}\end{cases} \tag{17}$$

Algorithm 1 Diffusion LMS with weight adaptation in [10]

Initialization: $\boldsymbol{\phi}_k^{(-1)}=0$
for each time $i\geq 0$ and each node k **do**
$\quad\boldsymbol{\psi}_k^{(i)}=\boldsymbol{\phi}_k^{(i-1)}+\mu_k\boldsymbol{u}_k^{(i)}(d_k^{(i)}-\boldsymbol{u}_k^{(i)\mathrm{H}}\boldsymbol{\phi}_k^{(i-1)})$
$\quad\gamma_{lk}^{2,(i)}=(1-\nu_k)\gamma_{lk}^{2,(i-1)}+\nu_k\|\boldsymbol{\psi}_l^{(i)}-\boldsymbol{\phi}_k^{(i-1)}\|^2$
$\quad a_{lk}^{(i)}=\frac{[\gamma_{lk}^{2,(i)}]^{-1}}{\sum_{m\in\mathcal{N}_k}[\gamma_{mk}^{2,(i)}]^{-1}}$
$\quad\boldsymbol{\phi}_k^{(i)}=\sum_{l\in\mathcal{N}_k}a_{lk}^{(i)}\boldsymbol{\psi}_l^{(i)}$
end for

Algorithm 2 Diffusion LMS with proposed weight adaptation

Initialization: $\boldsymbol{\phi}_k^{(-1)}=0$
for each time $i\geq 0$ and each node k **do**
$\quad\boldsymbol{\psi}_k^{(i)}=\boldsymbol{\phi}_k^{(i-1)}+\mu_k\boldsymbol{u}_k^{(i)}(d_k^{(i)}-\boldsymbol{u}_k^{(i)\mathrm{H}}\boldsymbol{\phi}_k^{(i-1)})$
$\quad\gamma_{lk}^{2,(i)}=(1-\nu_k)\gamma_{lk}^{2,(i-1)}+\nu_k\|\boldsymbol{\psi}_l^{(i)}-\boldsymbol{\phi}_k^{(i-1)}\|^2$
$\quad$**if** $\sum_{l\in\mathcal{N}_k}\gamma_{lk}^{2,(i)}-\gamma_{kk}^{2,(i)}|\mathcal{N}_k|>0$ **then**
$\quad\quad\beta_k^{(i)}=\frac{(|\mathcal{N}_k|-1)\gamma_{kk}^{2,(i)}}{\sum_{l\in\mathcal{N}_k}\gamma_{lk}^{2,(i)}-\gamma_{kk}^{2,(i)}|\mathcal{N}_k|}$
$\quad$**else**
$\quad\quad\beta_k^{(i)}=+\infty$ (large positive constant)
$\quad$**end if**
$\quad a_{lk}^{(i)}=\frac{\beta_k^{(i)}}{1+\beta_k^{(i)}|\mathcal{N}_k|}\ (l\in\mathcal{N}_k\setminus k),\quad \frac{1+\beta_k^{(i)}}{1+\beta_k^{(i)}|\mathcal{N}_k|}\ (l=k)$
$\quad\boldsymbol{\phi}_k^{(i)}=\sum_{l\in\mathcal{N}_k}a_{lk}^{(i)}\boldsymbol{\psi}_l^{(i)}$
end for

4.2. Adaptive Combination Weight

The optimal values of $\{a_{lk}^{\mathrm{opt}}\}$ for the conventional methods in (14) and $\{\beta_k^{\mathrm{opt}}\}$ for the proposed scheme in (17) require the knowledge of γ_l^2, which depends on locally unavailable network statistics such as the measurement vectors and the measurement noise profile. Thus, in [6], [10], [11], the estimation method of γ_l^2 at each node is proposed as

$$\gamma_{lk}^{2,(i)}=(1-\nu_k)\gamma_{lk}^{2,(i-1)}+\nu_k\|\boldsymbol{\psi}_l^{(i)}-\boldsymbol{\phi}_k^{(i-1)}\|^2, \tag{18}$$

where ν_k is a forgetting factor ($0<\nu_k<1$) and $\gamma_{lk}^{2,(i)}$ is the estimate of γ_l^2 at node k and time i. By using this estimate, the adaptive version of (14) is proposed in [10] as

$$a_{lk}^{(i)}=\begin{cases}\frac{[\gamma_{lk}^{2,(i)}]^{-1}}{\sum_{m\in\mathcal{N}_k}[\gamma_{mk}^{2,(i)}]^{-1}} & \text{if } l\in\mathcal{N}_k\\ 0 & \text{otherwise.}\end{cases} \tag{19}$$

In the same way, the adaptive version of the proposed combination rule is given by

$$\beta_k^{(i)} = \begin{cases} \frac{(|\mathcal{N}_k|-1)\gamma_{kk}^{2,(i)}}{A_k^{(i)}} & \text{if } A_k^{(i)} > 0 \\ +\infty & \text{otherwise,} \end{cases} \tag{20}$$

$$a_{lk}^{(i)} = \begin{cases} \frac{\beta_k^{(i)}}{1+|\mathcal{N}_k|\beta_k^{(i)}} & \text{if } l \in \mathcal{N}_k \text{ and } l \neq k \\ \frac{1+\beta_k^{(i)}}{1+|\mathcal{N}_k|\beta_k^{(i)}} & \text{if } k = l \\ 0 & \text{otherwise,} \end{cases} \tag{21}$$

where $A_k^{(i)} = \sum_{l\in\mathcal{N}_k} \gamma_{lk}^{2,(i)} - |\mathcal{N}_k|\gamma_{kk}^{2,(i)}$. The algorithms of diffusion LMS at node k using the conventional adaptive combination rule in (19) and the proposed rule in (21) are shown in *Algorithm 1* and *Algorithm 2*, respectively. The computational complexity of *Algorithm 1* and *Algorithm 2* are almost the same because the complexity of the proposed rule (21) becomes comparable to that of the conventional rule (19) by directly substituting (20) to (21).

5. SIMULATION RESULTS

In this section, we compare the MSD performance of the proposed scheme with that of the conventional methods by computer simulations. We assume that the measurement vectors $\boldsymbol{u}_k^{(i)}$ are zero-mean white circular Gaussian random vectors according to [3] and [5]. The background noise power at each node is randomly determined by uniform distribution of $[10^{-3}, 10^{-1}]$. The unknown vector is set to be $\boldsymbol{w}^{\text{o}} = \frac{1}{\sqrt{M}}\mathbf{1}$ with $M = 5$. The step-sizes at all node are set to be $\mu_k = 0.03$ and $\nu_k = 0.05$. Simulation results are obtained by averaging over 100 independent trials.

We have generated a random network for the diffusion LMS with $N = 50$ and the average degree of 6 as shown in Fig. 1. Fig. 2 shows the learning curves for the diffusion LMS using the uniform rule (5), the conventional adaptive combination rule (19), and the proposed rule (21) in terms of the network MSD, which is defined as the average squared errors of estimates at all nodes, $\frac{1}{N}\sum_{k=1}^{N}\|\boldsymbol{w}^{\text{o}} - \boldsymbol{\phi}_k^{(i)}\|^2$. In the conventional adaptive rule and the proposed rule, we show the cases where the initial values are $\gamma_{lk}^{(0)} = 10^4$ and 0.1. In the figure, we see that the proposed adaptive combination rule outperforms the other rules. We also observe that, in the conventional scheme (19), the performance largely depends on the initial value $\gamma_{lk}^{(0)}$, while the performance of the proposed rule is less sensitive to $\gamma_{lk}^{(0)}$.

6. CONCLUSIONS

In this paper, we have proposed a novel combination rule for diffusion LMS algorithm by using the idea of consensus propagation and developed it to an adaptive form. Simulation results show that the proposed adaptive rule outperforms the conventional rules in terms of network MSD performance with the lower sensitivity to the initial values of adaptation.

Future work includes the extension to more flexible weight control using asymmetric updates in consensus propagation, i.e., using different constants at each node depending on the direction of the messages.

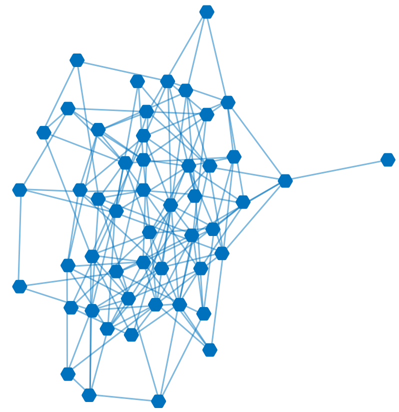

Fig. 1: Network with $N = 50$

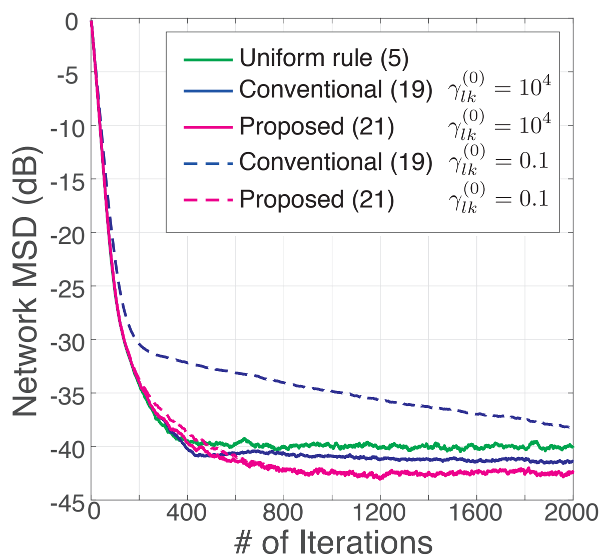

Fig. 2: Network MSD learning curves

7. REFERENCES

[1] S. Hara, H. Yomo, P. Popovski, and K. Hayashi, “New Paradigms in Wireless Communication Systems,” *Wirel. Pers. Commun.*, vol. 37, no. 3–4, pp. 233–241, May 2006.

[2] A. H. Sayed, S. Tu, J. Chen, X. Zhao, and Z. J. Tpwfic,

“Diffusion strategies for adaptation and learning over networks,” *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 155–171, May 2013.

[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, Jul. 2008.

[4] I. D. Schizas, G. Mateos, and G. B. Giannakis, “Distributed LMS for consensus-based in-network adaptive processing,” *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2365–2382, Jun. 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, Mar. 2010.

[6] A. H. Sayed, “Diffusion adaptation over networks,” *Academic Press Library in Signal Processing,* vol. 3, R. Chellapa and S. Theodoridis, *Eds.,* pp. 323–454, 2014.

[7] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, “Convergence in multiagent coordination, consensus, and flocking,” in *Proc. IEEE. CDC-ECC,* Seville, Spain, Dec. 2005, pp. 2996–3000.

[8] D. S. Scherber and H. C. Papadopoulos, “Locally constructed algorithms for distributed computations in ad hoc networks,” in *Proc. IPSN,* Berkeley, CA, Apr. 2014, pp. 11-19.

[9] L. Xiao and S. Boyd, “Fast linear iterations for distributed averaging,” *Syst. Control Lett.,* vol. 53, no. 1, pp. 65–78, Sep. 2004.

[10] X. Zhao, S.-Y. Tu, and A. H. Sayed, “Diffusion adaptation over networks under imperfect information exchange and non-stationary data,” *IEEE Trans. Signal Process.,* vol. 60, no. 7, pp. 3460–3475, Jul. 2012.

[11] S.-Y. Tu and A. H. Sayed, “Optimal combination rules for adaptation and learning over networks,” in *Proc. IEEE CAMSAP,* San Juan, Puerto Rico, Dec. 2011, pp. 317–320.

[12] N. Takahashi, I. Yamada, and A. H. Sayed, “Diffusion least-mean squares with adaptive combiners: Formulation and performance analysis,” *IEEE Trans. Signal Process.,* vol. 58, no. 9, pp. 4795–4810, Sept. 2010.

[13] C. C. Moallemi and B. V. Roy, “Consensus propagation,” *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4753–4766, Nov. 2006.

[14] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann Publishers Inc., 1988.

[15] A. Nakai and K. Hayashi, “Diffusion LMS using consensus propagation,” in *Proc. APSIPA ASC,* Kuala Lumpur, Malaysia, Dec. 2017.