Diffusion LMS Strategies for Parameter Estimation over Fading Wireless Channels

Reza Abdolee and Benoit Champagne
McGill University
Dept. of Elec. and Comp. Engineering
Montreal, QC, H3A 2A7, Canada
Email: reza.ab@ieee.org, benoit.champagne@mcgill.ca

Ali H. Sayed
University of California
Dept. of Electrical Engineering
Los Angeles, CA, 90095, USA
Email: sayed@ee.ucla.edu

***Abstract*—We propose a modified diffusion strategy for parameter estimation in sensor networks where nodes exchange information over fading wireless channels. We show that the effect of fading can be mitigated by incorporating local equalization coefficients into the diffusion process. We explain how the equalization coefficients are chosen and show that the (mean) stability of the network continues to be insensitive to the choice of the combination weights and to the network topology. Our computer experiments demonstrate that the performance of the modified diffusion algorithm in fading scenario is nearly identical to that of centralized least-mean square (LMS) with equalized input data.**

***Index Terms*—distributed estimation, diffusion adaptation, wireless sensor networks, link-failure, fading channels.**

I. Introduction

Diffusion strategies are powerful schemes that enable adaptation and learning in real-time over networks in response to streaming data [1]–[3]. Previous works on distributed estimation and adaptation over networks examined the effect of noisy communication links on the performance of the diffusion learning schemes [4]–[6]. The main conclusion is that performance degradation occurs unless the combination weights are adjusted accordingly in order to counter the effect of noise over the links.

In this work, we take a step further and consider the effect of fading and path loss, in addition to additive noise. Previously, reference [6] proposed diffusion-type strategies for parameter estimation over non-ideal communication links in sensor networks. This reference, however, overlooked the ramifications of exchanging information over fading wireless channels such as random link failure and time-varying network topology. We, in this paper, consider these phenomena and provide brief convergence analysis for Adapt-then-Combine (ATC) diffusion LMS strategy in such conditions. Our analysis shows that the effect of fading can be overcome by incorporating local equalization coefficients into the diffusion strategy. We explain how the equalization coefficients are chosen and show that the (mean) stability of the network continues to be insensitive to the choice of the combination weights and to the network topology. Simulation results are used to illustrate the findings.

Notation: We use boldface letters to represent random variables, and normal font to represent deterministic quantities. For complex vectors and matrices, $(\cdot)^*$ denotes complex conjugate transposition. I_M denotes the identity matrix of size $M \times M$, and $\mathbb{E}[\cdot]$ is the expectation operator.

The work of R. Abdolee and B. Champagne was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada.
The work of A. H. Sayed was supported in part by NSF grant CCF-1011918.

II. Diffusion Estimation over Fading Channels

Consider a set of N interconnected nodes that are distributed over some spatial domain. At every time instant i, node k collects data $\{\boldsymbol{d}_k(i), \boldsymbol{u}_{k,i}\}$ that are related to some unknown $M \times 1$ vector w^o:

$$\boldsymbol{d}_k(i) = \boldsymbol{u}_{k,i} w^o + \boldsymbol{v}_k(i) \tag{1}$$

where $\boldsymbol{d}_k(i) \in \mathbb{C}$ is a scalar measurement, $\boldsymbol{u}_{k,i} \in \mathbb{C}^{1\times M}$ is a regression vector, and $\boldsymbol{v}_k(i)$ is measurement noise.

Assumption 1. *The data in model (1) are assumed to satisfy the following conditions:*

a) *The regressors $\{\boldsymbol{u}_{k,i}\}$ are zero-mean, i.i.d in time and spatially independent with covariance matrix $R_{u,k} = \mathbb{E}[\boldsymbol{u}_{k,i}^* \boldsymbol{u}_{k,i}] > 0$.*
b) *The noise process $\boldsymbol{v}_k(i)$ is zero-mean, temporally white and independent over space with variance $\sigma_{v,k}^2$. The measurement noise is also independent of the regressors $\{\boldsymbol{u}_{m,j}\}$ for all k, m, i, j.*

In this network, two nodes are said to be neighbors if they can share information. The set of neighbors of node k, including node k itself, is denoted by $\mathcal{N}_k$. The objective of the network is to estimate the unknown parameter vector w^o, in a distributed manner, over fading wireless channels where nodes are allowed to communicate with their neighbors only. The network seeks this objective by minimizing the following global cost function:

$$J^{\text{glob}}(w) = \sum_{k=1}^{N} \mathbb{E}|\boldsymbol{d}_k(i) - \boldsymbol{u}_{k,i}w|^2 \tag{2}$$

Several adaptive diffusion strategies have been proposed to solve (2) in a distributed manner over ideal communication links [1], [2]. One such strategy is the following ATC algorithm:

$$\boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu_k \boldsymbol{u}^*_{k,i}\big[\boldsymbol{d}_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{w}_{k,i-1}\big] \tag{3a}$$

$$\boldsymbol{w}_{k,i} = \sum_{\ell \in \mathcal{N}_k} a_{\ell,k}\boldsymbol{\psi}_{\ell,i} \tag{3b}$$

where $\mu_k > 0$ is the step-size used by node k, and the $\{a_{\ell,k}\}$ denote nonnegative entries of a left-stochastic matrix A that satisfy:

$$a_{\ell,k} = 0 \;\text{ if }\; \ell \notin \mathcal{N}_k \quad \text{and} \quad \sum_{\ell \in \mathcal{N}_k} a_{\ell,k} = 1 \tag{4}$$

In this implementation, step (3a) is an adaptation step where node k updates its intermediate estimate $\boldsymbol{w}_{k,i-1}$ to $\boldsymbol{\psi}_{k,i}$ using its measured data $\{\boldsymbol{u}_{k,i}, \boldsymbol{d}_k(i)\}$. The second step, (3b), is a combination step in which each node k combines its intermediate estimate $\boldsymbol{\psi}_{k,i}$ with that of its neighbors to obtain $\boldsymbol{w}_{k,i}$. The above algorithm works well over ideal communication channels. When the exchange of information between neighboring nodes is subject to noise, some degradation in performance occurs [4]–[8]; reference [5] shows how the combination weights $\{a_{\ell,k}\}$ can be adjusted to counter the effect of noisy links in the presence of additive noise.

In this work, we study the case in which the exchange of information is subject to fading, path loss, and link failure. Specifically, we assume the transmitted signal from node ℓ to node k experiences channel distortion of the following form [9]:

$$\boldsymbol{\psi}_{\ell k,i} = \boldsymbol{h}_{\ell,k}(i)\sqrt{\frac{P_o}{r^{\alpha}_{\ell,k}}}\boldsymbol{\psi}_{\ell,i} + \boldsymbol{v}^{(\psi)}_{\ell k,i} \tag{5}$$

where $\boldsymbol{\psi}_{\ell k,i} \in \mathbb{C}^{M\times 1}$ is the distorted estimate, $\boldsymbol{h}_{\ell,k}(i) \in \mathbb{C}$ denotes the fading channel coefficient, $P_o \in \mathbb{R}^+$ is the transmit signal power, $r_{\ell,k} = r_{k,\ell} \in \mathbb{R}^+$ is the distance between node k and ℓ, and $\alpha \in \mathbb{R}^+$ is the path loss exponent. We let the quantity $\boldsymbol{h}_{\ell,k}(i)\sqrt{\frac{P_o}{r^{\alpha}_{\ell,k}}}$ denote the (ℓ,k)-th entry of the network channel matrix, $\boldsymbol{H}_i$. Moreover, the link noise $\boldsymbol{v}^{(\psi)}_{\ell k,i}$ is assumed to be a zero mean white random process with covariance matrix $\sigma^{2(\psi)}_{v,\ell k}I$. Transmission from node ℓ to node k is considered successful if the SNR between node ℓ and k, denoted by $\boldsymbol{\varsigma}_{\ell k}(i)$, exceeds some threshold level $\varsigma^o_{\ell k}$. The threshold level is chosen as the SNR in the non-fading link scenario and is defined as:

$$\varsigma^o_{\ell k} = \frac{P_o}{\sigma^{2(\psi)}_{v,\ell k} r^{\alpha}_o} \tag{6}$$

where r_o is the node's transmission range. When fading is present, the instantaneous SNR is:

$$\boldsymbol{\varsigma}_{\ell k}(i) = \frac{|\boldsymbol{h}_{\ell,k}(i)|^2 P_o}{\sigma^{2(\psi)}_{v,\ell k} r^{\alpha}_{\ell,k}} \tag{7}$$

The transmission is successful if $\boldsymbol{\varsigma}_{\ell k}(i) \geq \varsigma^o_{\ell k}$, i.e.,

$$|\boldsymbol{h}_{\ell,k}(i)|^2 \geq \frac{r^{\alpha}_{\ell,k}}{r^{\alpha}_o} \tag{8}$$

Let us assume $\boldsymbol{h}_{\ell,k}(i)$ is a complex Gaussian random variable with zero-mean and unit variance. Then, its envelop is Rayleigh distributed and its squared magnitude, $|\boldsymbol{h}_{\ell,k}(i)|^2$, has exponential distribution with parameter one [10]. In this case, the probability of successful transmission can be computed as:

$$p_{\ell,k} = p\Big(|\boldsymbol{h}_{\ell,k}(i)|^2 \geq \frac{r^{\alpha}_{\ell,k}}{r^{\alpha}_o}\Big) = e^{-(r_{\ell k}/r_o)^{\alpha}} \tag{9}$$

This expression shows that the probability of successful transmission decreases as the distance between two nodes increases. Motivated by this result, we redefine the neighborhood of node k as consisting of all nodes ℓ for which $\boldsymbol{\varsigma}_{\ell k}(i)$ exceeds $\varsigma^o_{\ell k}$. In this way, the neighborhood becomes a random process and we denote it by $\boldsymbol{\mathcal{N}}_{k,i}$.

In view of (5), we modify the ATC strategy (3) as follows:

$$\boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu_k \boldsymbol{u}^*_{k,i}\big[\boldsymbol{d}_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{w}_{k,i-1}\big] \tag{10a}$$

$$\boldsymbol{w}_{k,i} = \sum_{\ell \in \boldsymbol{\mathcal{N}}_{k,i}} \boldsymbol{a}_{\ell,k}(i)\boldsymbol{g}_{\ell,k}(i)\boldsymbol{\psi}_{\ell k,i} \tag{10b}$$

where the combination weights $\{\boldsymbol{a}_{\ell,k}(i)\}$ are now allowed to vary with time and each scalar gain $\boldsymbol{g}_{\ell,k}(i)$ is the (ℓ,k)-th entry of some network equalization matrix, $\boldsymbol{G}_i$, used by the algorithm to counter the effect of fading. In (10b), $\boldsymbol{\psi}_{\ell k,i}$ is the received distorted estimate from node ℓ expressed by (5). We set $\boldsymbol{\psi}_{kk,i} = \boldsymbol{\psi}_{k,i}$ to maintain consistency in the notation.

The randomness of the coefficients $\{\boldsymbol{a}_{\ell,k}(i)\}$ can be explained using equation (8). The communication between node ℓ and k is successful if (8) is satisfied. Otherwise, the link fails. When this happens, the associated combination weight $\boldsymbol{a}_{\ell,k}(i)$ must be set to zero, meaning that the remaining combination coefficients for node k need to be adjusted to preserve their sum to one. This suggests that the neighborhood set $\boldsymbol{\mathcal{N}}_{k,i}$ has to be updated whenever the SNR of one of the received signals crosses the threshold in either direction:

$$\boldsymbol{\mathcal{N}}_{k,i} = \Big\{\ell \in \{1,\cdots,N\} \Big|\, \boldsymbol{\varsigma}_{\ell k}(i) \geq \varsigma^o_{\ell k}\Big\} \tag{11}$$

Let us assume the combination weights are chosen according to the following structure:

$$\boldsymbol{a}_{\ell,k}(i) = \gamma \boldsymbol{\mathcal{I}}_{\ell,k}(i), \quad \boldsymbol{a}_{k,k}(i) = 1 - \sum_{\ell \in \boldsymbol{\mathcal{N}}_{k,i}\setminus\{k\}} \boldsymbol{a}_{\ell,k}(i) \tag{12}$$

where $\gamma \le 1/N$, and $\boldsymbol{\mathcal{I}}_{\ell,k}(i)$ is defined as:

$$\boldsymbol{\mathcal{I}}_{\ell,k}(i) = \begin{cases} 1, & \text{if } \ell \in \boldsymbol{\mathcal{N}}_{k,i} \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

In other words, when reception at node k from node ℓ is successful $\boldsymbol{\mathcal{I}}_{\ell,k}(i) = 1$. Otherwise, $\boldsymbol{\mathcal{I}}_{\ell,k}(i) = 0$. This implies that the indicator operator, $\boldsymbol{\mathcal{I}}_{\ell,k}(i)$, is a random variable with binomial distribution for which the probability of success, $p_{\ell,k}$, is given by the exponential function (9). The mean and variance of this distribution are respectively $p_{\ell,k}$ and $p_{\ell,k}(1-p_{\ell,k})$ [10].

One way to compute the equalization coefficients that appear in (10b) is by using a zero-forcing type construction as follows:

$$\boldsymbol{g}_{\ell,k}(i) = \frac{\boldsymbol{h}^*_{\ell,k}(i)}{|\boldsymbol{h}_{\ell,k}(i)|^2}\sqrt{\frac{r^{\alpha}_{\ell,k}}{P_o}} \tag{14}$$

This approach assumes that the nodes are able to estimate the fading coefficients. Alternatively, if the noise variance $\sigma^{2(\psi)}_{v,\ell k}$ is known, then one can use a minimum-mean-square-error construction for the equalization coefficients; we continue with (14) in this paper to illustrate the main ideas. Using (5) and (14) in (10b) gives:

$$\boldsymbol{w}_{k,i} = \sum_{\ell \in \boldsymbol{\mathcal{N}}_{k,i}} \boldsymbol{a}_{\ell,k}(i)\boldsymbol{\psi}_{\ell,i} + \boldsymbol{v}^{(\psi)}_{k,i} \tag{15}$$

where

$$\boldsymbol{v}^{(\psi)}_{k,i} = \sum_{\ell \in \boldsymbol{\mathcal{N}}_{k,i}} \boldsymbol{a}_{\ell,k}(i)\boldsymbol{g}_{\ell,k}(i)\boldsymbol{v}^{(\psi)}_{\ell k,i} \tag{16}$$

The combination step (15) appears to be similar to that of diffusion LMS with imperfect information exchange analyzed in [5]. However, there are two main differences. First, the combination weights are random. And, second, the noise power over the fading links are generally larger.

III. Performance Analysis

In this section, we examine the performance of algorithm (10) under Assumptions 1 and 2. Derivations are omitted for brevity.

Assumption 2. *It is assumed that the fading channel satisfies the following conditions:*

a) *All channel coefficients $\{\boldsymbol{h}_{\ell,k}(i)\}$ are zero mean i.i.d in time and independent over space with unit variance. These coefficients are assumed to be independent from the regressors $\{\boldsymbol{u}_{m,j}\}$, measurement noise, $\{\boldsymbol{v}_m(j)\}$, and the communication noise $\{\boldsymbol{v}^{(\psi)}_{nm}(j)\}$ for all ℓ, k, n, m, i, j.*
b) *The noise vector $\boldsymbol{v}^{(\psi)}_{\ell k,i}$ is zero-mean Gaussian, i.i.d in time and independent over space with a covariance matrix $\sigma^{2(\psi)}_{v,\ell k} I$. This noise is also independent from $\boldsymbol{v}_m(j)$, $\boldsymbol{d}_m(j)$, and $\boldsymbol{u}_{m,j}$ for all k, m, i, j.*

A. Convergence in the Mean

Let $\tilde{\boldsymbol{w}}_{k,i} \triangleq w^o - \boldsymbol{w}_{k,i}$ and $\tilde{\boldsymbol{\psi}}_{k,i} \triangleq w^o - \boldsymbol{\psi}_{k,i}$ and define:

$$\tilde{\boldsymbol{\psi}}_i \triangleq \text{col}\{\tilde{\boldsymbol{\psi}}_{1,i}, \tilde{\boldsymbol{\psi}}_{2,i}, \ldots, \tilde{\boldsymbol{\psi}}_{N,i}\} \tag{17}$$

$$\tilde{\boldsymbol{w}}_i \triangleq \text{col}\{\tilde{\boldsymbol{w}}_{1,i}, \tilde{\boldsymbol{w}}_{2,i}, \ldots, \tilde{\boldsymbol{w}}_{N,i}\} \tag{18}$$

We collect the $\{\boldsymbol{a}_{\ell,k}(i)\}$ into a left-stochastic matrix $\boldsymbol{A}_i$ and let $\boldsymbol{\mathcal{A}}_i \triangleq \boldsymbol{A}_i \otimes I_M$, where $\otimes$ is the Kronecker product. We further introduce the variables:

$$\boldsymbol{\mathcal{R}}_i \triangleq \text{diag}\Big\{\boldsymbol{u}^*_{1,i}\boldsymbol{u}_{1,i}, \cdots, \boldsymbol{u}^*_{N,i}\boldsymbol{u}_{N,i}\Big\} \tag{19}$$

$$\mathcal{M} \triangleq \text{diag}\Big\{\mu_1 I_M, \cdots, \mu_N I_M\Big\} \tag{20}$$

$$\boldsymbol{p}_i \triangleq \text{col}\{\boldsymbol{u}^*_{1,i}\boldsymbol{v}_1(i), \cdots, \boldsymbol{u}^*_{N,i}\boldsymbol{v}_N(i)\} \tag{21}$$

$$\boldsymbol{v}^{(\psi)}_i \triangleq \text{col}\{\boldsymbol{v}^{(\psi)}_{1,i}, \cdots, \boldsymbol{v}^{(\psi)}_{N,i}\} \tag{22}$$

A recursion for the network error vector, $\tilde{\boldsymbol{w}}_i$, can be obtained by subtracting w^o from both sides of (10):

$$\tilde{\boldsymbol{w}}_i = \boldsymbol{\mathcal{A}}^T_i(I - \mathcal{M}\boldsymbol{\mathcal{R}}_i)\tilde{\boldsymbol{w}}_{i-1} - \boldsymbol{\mathcal{A}}^T_i\mathcal{M}\boldsymbol{p}_i + \boldsymbol{v}^{(\psi)}_i \tag{23}$$

Under Assumptions 1 and 2, we obtain:

$$\mathbb{E}[\tilde{\boldsymbol{w}}_i] = \mathcal{A}^T(I - \mathcal{M}\mathcal{R})\mathbb{E}[\tilde{\boldsymbol{w}}_{i-1}] \tag{24}$$

where $\mathcal{A}^T \triangleq \mathbb{E}[\boldsymbol{\mathcal{A}}^T_i]$ and $\mathcal{R} \triangleq \mathbb{E}[\boldsymbol{\mathcal{R}}_i]$ are given by:

$$\mathcal{A}^T = A^T \otimes I_M \tag{25}$$

$$\mathcal{R} = \text{diag}\{R_{u,1}, \ldots, R_{u,N}\} \tag{26}$$

with the elements of A computed as:

$$a_{\ell,k} = \gamma p_{\ell,k}, \quad a_{k,k} = 1 - \gamma\sum_{\ell=1\backslash k}^{N} p_{\ell,k} \tag{27}$$

Observe that $A^T\mathbb{1} = \mathbb{1}$, where $\mathbb{1}$ denotes a vector with all entries equal to one. From (24), it can be verified that $\lim_{i\to\infty}\mathbb{E}[\tilde{\boldsymbol{w}}_i]$ vanishes if the step-sizes are chosen as [3], [5]:

$$0 < \mu_k < \frac{2}{\lambda_{\max}(R_{u,k})} \tag{28}$$

where $\lambda_{\max}(\cdot)$ denotes the maximum eigenvalue of its matrix argument. Note that the mean-convergence of the algorithm is not restricted by the particular choice of $\boldsymbol{A}_i$ made in (12). Theoretically, other choices also lead to the same conclusion as long as the average of $\boldsymbol{A}_i$ yields a left-stochastic matrix.

B. Optimal Combination Weights

Based on the above result, several of the earlier combination rules proposed for diffusion LMS can be used in networks where link failures are possible due to fading. For example, the optimized combination rule given in [5] becomes:

$$\boldsymbol{a}_{\ell,k}(i) = \begin{cases} \frac{\gamma^{-2}_{\ell,k}(i)}{\sum_{m\in\boldsymbol{\mathcal{N}}_{k,i}}\gamma^{-2}_{m,k}(i)}, & \text{if } \ell \in \boldsymbol{\mathcal{N}}_{k,i} \\ 0, & \text{otherwise} \end{cases} \tag{29}$$

where

$$\gamma^2_{\ell,k}(i) = \begin{cases} \mu_\ell^2 \sigma^2_{v,\ell} \mathrm{Tr}(R_{u,\ell}) + M\sigma^{2(\psi)}_{v,\ell k}, & \text{if } \ell \in \boldsymbol{\mathcal{N}}_{k,i} \backslash \{k\} \\ \mu_k^2 \sigma^2_{v,k} \mathrm{Tr}(R_{u,k}), & \text{if } \ell = k \end{cases} \tag{30}$$

In this formulation, $\gamma^2_{\ell,k}(i)$ varies with time because the neighborhood, $\boldsymbol{\mathcal{N}}_{k,i}$, changes due to fading. This combination rule can be improved if we exploit the channel state information (CSI) while optimizing the algorithm performance over $\boldsymbol{A}_i$. Assuming known CSI at each node and following the optimization strategy in [5], we arrive at:

$$\gamma^2_{\ell,k}(i) = \begin{cases} \mu_\ell^2 \sigma^2_{v,\ell} \mathrm{Tr}(R_{u,\ell}) + M|\boldsymbol{g}_{\ell,k}(i)|^2\sigma^{2(\psi)}_{v,\ell k}, & \text{if } \ell \in \boldsymbol{\mathcal{N}}_{k,i} \backslash \{k\} \\ \mu_k^2 \sigma^2_{v,k} \mathrm{Tr}(R_{u,k}), & \text{if } \ell = k \end{cases} \tag{31}$$

where we assume the channel changes slowly enough to enable optimization. We then substitute (31) into (29) to obtain $\boldsymbol{A}_i$.

IV. Simulation Results

In this section, we present computer experiments to illustrate the performance of the modified ATC diffusion LMS algorithm over wireless channels. We consider a network with $N = 10$ nodes that are randomly spread over a unit square area with the nominal transmission range of $r^o = 0.4$ unit length. We choose $\mu_k = \mu_{\text{diff}} = 0.01$ and $\boldsymbol{w}_{k,-1} = 0$ for all k. The objective of the network is to cooperatively estimate the vector $w^o = [-2, -1, 1, 2]^T$.

We compare the performance of the distributed solution against a centralized LMS approach. In the centralized set-up, we assume nodes transmit their data, $\{\boldsymbol{d}_k(i), \boldsymbol{u}_{k,i}\}$, over wireless channels to the fusion center located at the network center, say, at location $x = y = 0.5$. The links between the nodes and the fusion center have random behavior and may fail if condition (8) does not hold. The corrupted version of the data received by the fusion center is expressed as:

$$\boldsymbol{u}_{kf,i} = \boldsymbol{h}_{kf}(i)\sqrt{\frac{P_o}{r^\alpha_{kf}}}\boldsymbol{u}_{k,i} + \boldsymbol{v}^{(u)}_{kf,i} \tag{32}$$

$$\boldsymbol{d}_{kf}(i) = \boldsymbol{h}_{kf}(i)\sqrt{\frac{P_o}{r^\alpha_{kf}}}\boldsymbol{d}_k(i) + \boldsymbol{v}^{(d)}_{kf}(i) \tag{33}$$

where the link noise $\boldsymbol{v}^{(u)}_{kf,i}$ is a zero mean white random process with covariance matrix $\sigma^{2(u)}_{v,kf} I$, $\boldsymbol{h}_{kf}(i)$ is a random fading coefficient with zero mean and power $\sigma^2_{h,kf} = 1$, and $\boldsymbol{v}^{(d)}_{kf}(i)$ is zero mean white noise with variance $\sigma^{2(d)}_{v,kf}$. The random variables $\boldsymbol{v}^{(u)}_{kf,i}$, $\boldsymbol{h}_{kf}(i)$, and $\boldsymbol{v}^{(d)}_{kf}(i)$ are assumed to be mutually independent over time and space. At time instant i, the centralized LMS algorithm starts with $\boldsymbol{\psi}_{0,i} = \boldsymbol{w}_{i-1}$ and iterates incrementally over data from node $k = 1$ to node N:

$$\boldsymbol{\psi}_{k,i} = \boldsymbol{\psi}_{k-1,i} + \mu\hat{\boldsymbol{u}}^*_{kf,i}[\hat{\boldsymbol{d}}_{kf}(i) - \hat{\boldsymbol{u}}_{kf,i}\boldsymbol{\psi}_{k-1,i}] \tag{34}$$

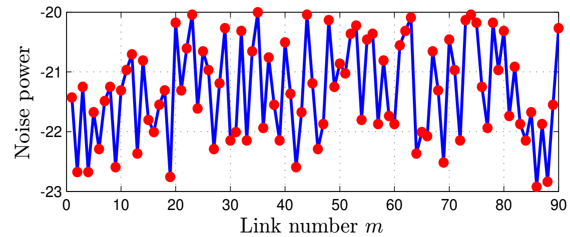

Fig. 1. Noise variances, $\sigma^{2(\psi)}_{v,\ell k}$, across the links in the network.

where $\mu > 0$ is the step-size, and

$$\hat{\boldsymbol{u}}_{kf,i} = \boldsymbol{g}_{kf}(i)\boldsymbol{u}_{kf,i} \tag{35}$$

$$\hat{\boldsymbol{d}}_{kf,i} = \boldsymbol{g}_{kf}(i)\boldsymbol{d}_{kf,i} \tag{36}$$

are the equalized data at the fusion center. Once the algorithm finishes iterating over k, it updates $\boldsymbol{w}_{i-1}$ to $\boldsymbol{w}_i = \boldsymbol{\psi}_{N,i}$.

The parameters $\mathrm{Tr}(R_{u,k})$, $\sigma^2_{v,k}$ and $\sigma^2_{v,kf}$ are presented in Table I. For the distributed set-up, the communication noise power, $\sigma^{2(\psi)}_{v,\ell k}$, is plotted in Fig. 1 in terms of the link number defined as follows. We denote the link from node ℓ to node k as $l_{\ell,k}$ where $\ell \neq k$. We arrange the links $\{l_{\ell,k},\ \ell \in \{1, 2, \cdots, N\}\backslash k\}$ in an ascending order of ℓ in the list L_k for each node k. The list L_k is a set with ordered elements. We collect $\{L_k\}$ in an ascending order of k to get the overall list $L \triangleq \{L_1, L_2 \cdots, L_N\}$. In this representation, the m-th link in the network is given by the m-th element in the list L.

To have a fair comparison between centralized and diffusion LMS algorithms, the communication noise power, $\sigma^{2(d)}_{v,kf}$, $\sigma^{2(u)}_{v,kf}$ and $\sigma^{2(\psi)}_{v,\ell k}$, are chosen uniformly from the range $[50, 100] \times 10^{-3}$. The step-size in the centralized LMS algorithm is chosen as $\mu = 2.5\mu_{\text{diff}}/N$ to ensure the same convergence rate as that of diffusion LMS. Note that in networks with static topologies and no link failure, the step-size for the centralized LMS is chosen as $\mu = \mu_{\text{diff}}/N$. The reason that here we have an extra 2.5 factor is that in fading condition not all the nodes are always able to transmit their data to the fusion center. Therefore, the centralized algorithm needs to take a larger step-size to compensate the absence of the contribution from these nodes in enhancing its convergence speed.

Figure 2 illustrates the convergence of the network mean error vector $\mathbb{E}[\tilde{\boldsymbol{w}}_i]$ for the equalized diffusion LMS algorithm (10). Figure 3 shows the network transient mean-square deviation (MSD) for diffusion and centralized LMS defined respectively by:

$$\eta_{\text{diff}}(i) = \frac{1}{N}\sum_{k=1}^{N}\mathbb{E}\|\tilde{\boldsymbol{w}}_{k,i}\|^2, \qquad \eta_{\text{ctrl}}(i) = \mathbb{E}\|\tilde{\boldsymbol{w}}_i\|^2 \tag{37}$$

TABLE I
NETWORK SIGNAL AND NOISE POWER PROFILE

	Node number k									
Parameters	1	2	3	4	5	6	7	8	9	10
$\sigma^2_{v,k}$	0.0580	0.0840	0.0730	0.0880	0.0990	0.0880	0.0730	0.0690	0.0530	0.0680
$\sigma^{2(d)}_{v,kf}, \sigma^{2(u)}_{v,kf}$	0.0058	0.0074	0.0093	0.0093	0.0066	0.0095	0.0067	0.0063	0.0063	0.0100
$\mathrm{Tr}(R_{u,k})$	5.100	4.4000	3.4000	3.8000	4.2000	5.9000	3.0000	4.2000	5.9000	2.0000

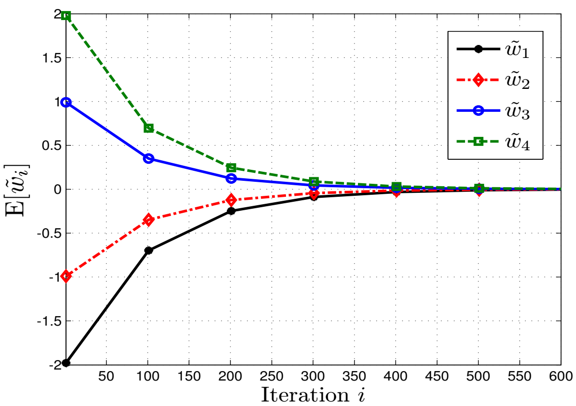

Fig. 2. Network error vector $\mathbb{E}[\tilde{\boldsymbol{w}}_i]$.

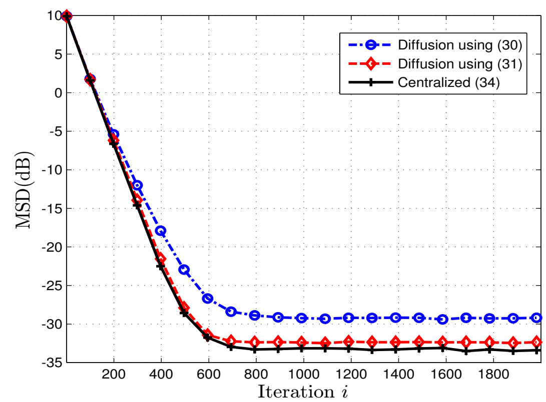

Fig. 3. Network MSD learning curve.

The results are drawn from the average of 200 independent runs. The MSD curves suggest that the performance of diffusion LMS depends on the combination weights $\boldsymbol{A}_i$. As expected, diffusion LMS with $\boldsymbol{A}_i$ computed using (31) outperforms diffusion LMS with (30). The figure also shows that the MSD of diffusion LMS with combination rule (31) approaches that of centralized LMS. Note that in fading channel scenarios, where the communication range, path loss and random fading are taken into account, the performance of the centralized network is adversely affected as well and often more severely than in a distributed set-up. The reason is that in diffusion LMS, the probability of successful transmission is higher than that of the centralized algorithm. Indeed according to (9), in the centralized set-up, nodes locating at the network edge have higher probability of link-failure in transmitting their data to the fusion center, whereas in diffusion set-up, the link-failure probability is lower since only nearby neighbors need to exchange information.

V. CONCLUSION

We modified the diffusion LMS algorithm for parameter estimation applications in wireless networks and investigated its mean convergence behavior under fading conditions. The results show that under a given step-size condition, the modified strategy will converge in the mean-sense if the network combination matrix preserves its left-stochastic property.

REFERENCES

[1] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Trans. on Signal Processing*, vol. 56, pp. 3122–3136, July 2008.

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

[3] A. H. Sayed, "Diffusion adaptation over networks," to appear in *E-Reference Signal Processing*, R. Chellapa and S. Theodoridis, *Eds.*, Elsevier, 2013. Also available as manuscript arXiv:1205.4220v1, May 2012.

[4] A. Khalili, M. Tinati, A. Rastegarnia, and J. Chambers, "Steady-state analysis of diffusion LMS adaptive networks with noisy links," *IEEE Trans. on Signal Processing*, vol. 60, no. 2, pp. 974–979, Feb. 2012.

[5] X. Zhao, S. Y. Tu, and A. H. Sayed, "Diffusion adaptation over networks under imperfect information exchange and non-stationary data," *IEEE Trans. on Signal Processing*, vol. 6, no. 7, pp. 3460–3475, July 2012.

[6] R. Abdolee and B. Champagne, "Diffusion LMS algorithms for sensor networks over non-ideal inter-sensor wireless channels," in *Proc. Int. Conf. on Dist. Computing in Sensor Systems and Workshops*, Barcelona, Spain, June 2011, pp. 1–6.

[7] A. Khalili, M. Tinati, A. Rastegarnia, and J. Chambers, "Transient analysis of diffusion least-mean squares adaptive networks with noisy channels," *Int. J. of Adaptive Cont. and Signal Proc.*, Feb. 2012.

[8] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks: Link failures and channel noise," *IEEE Trans. on Signal Processing*, vol. 57, no. 1, pp. 355–369, Jan. 2009.

[9] K. Chan, A. Swami, Q. Zhao, and A. Scaglione, "Consensus algorithms over fading channels," in *Proc. of Military Communication Conference (MILCOM)*, San Jose, California, Oct. 2010, pp. 549–554.

[10] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*. 2nd edition, Addison-Wesley, Reading, MA, 1994.