Partial-Diffusion Least Mean-Square Estimation Over Networks Under Noisy Information Exchange

Wael M. Bazzi
Electrical Engineering Department
American University in Dubai
Dubai, UAE
Email: wbazzi@aud.edu

Vahid Vadidpour, Amir Rastegarnia, Azam Khalili
Department of Electrical Engineering
Malayer University
Malayer, Iran, 65719-95863
Email: {vahidpour, rastegarnia, khalili}@malayeru.ac.ir

Abstract— **Partial diffusion scheme is an effective method for reducing computational load and power consumption in adaptive network implementation. The Information is exchanged among the nodes, usually over noisy links. In this paper, we consider a general version of Partial-Diffusion Least-Mean-Square (PDLMS) algorithm in the presence of various sources of imperfect information exchanges. Like the established PDLMS, we consider two different schemes to select the entries, sequential and stochastic, for transmission at each iteration. Our objective is to analyze the aggregate effect of these perturbations on general PDLMS strategies. Simulation results demonstrate that considering noisy link assumption adds a new complexity to the related optimization problem and the trade-off between communication cost and estimation performance in comparison to ideal case becomes unbalanced. Our simulation results substantiate the effect of noisy links on PDLMS algorithm and verify the theoretical analysis.**

Keywords- adaptive networks; diffusion adaptation; noisy informatin exchange; partial diffusion; sequential, stochastic.

I. INTRODUCTION

Due to limited electrical power and bandwidth resources for inter-node communication over a practical Wireless Sensor Networks (WSN) or ad hoc networks, data transmission through radio communication links can become prohibitively expensive for realizing a collaborative task. Generally speaking, although benefits of diffusion strategies achieved by increasing inter-node communications, they are compromised by the communication cost. As a result, since various nodes can have various numbers of neighbors, they may require disparate hardware or consume power dissimilarity [1]–[4]. Therefore, reducing the communication cost while maintaining the benefits of cooperation is of practical importance.

There have been several attempts to reduce the communication cost without considerable degradation of the estimation and compromising the cooperation benefits in diffusion algorithms. Among them diffusion least mean-square (LMS), such as reducing the dimension of the estimates [5]–[7], selecting a subset of the entries of the estimates [1] [2] [8] [9], set-membership filtering [10][11], or partial updating [12] have been reported in [13]–[15]. Among these methods, we focus on [1] which LMS algorithm for adaptive distribute estimation has been formulated and analyzed by utilizing partial-diffusion. In [1],

an adapt-then-combine (ATC) Partial-Diffusion Least-Mean-Square (PDLMS) algorithm is proposed for distributed estimation over adaptive networks in which, at each iteration, each node transmits a subsets of the entries of intermediate estimate vector to its neighbors.

In the PDLMS strategy proposed in [1], the weight estimates that are exchanged among the nodes can be subject to perturbations over communication links. The effect of link noise during the exchange of weight estimates, already appear for the diffusion algorithm in the works [16]–[20]. It should be noted that since our objective is to minimize the inter-node communications, the nodes only exchange their intermediate estimates with their neighbors. Therefore, we allow for noisy exchange just during the two combination steps. We subsequently study the performance of this general case utilizing the energy conservation argument. We established its stability and convergence in the mean and mean-square senses. We also derive a theoretical expression for the steady-state Mean-Square-Deviation (MSD) and verify its accuracy via numerical simulations.

The main contributions in this paper include:

- Focusing on [1] which involves transmission of a subset of entries of the inter-node estimate vectors named partial diffusion, we provide a more general algorithmic structure of which [1] is just a special case. To achieve this, we consider the fact that the weight estimates exchanged among the nodes can be subject to quantization errors and additive noise over communication links. We also consider two different schemes for selecting the weight vector entries for transmission at each iteration. We allow for noisy exchange during the two combination steps only. It should be noted that since our objective is to minimize the inter-node communication, the nodes only exchange their intermediate estimates with their neighbors;

- Using the energy conservation argument [21] we analyze the stability of algorithms in mean and mean square senses under certain statistical conditions.

- We illustrate the comparable convergence performance of PDLMS algorithm with noisy links using different numerical examples.

This paper is organized as follows. In Section II, we formulate the PDLMS under noisy information exchange. The performance analyses are examined in Section III. We

provide simulation results in Section IV and draw conclusions in Section V.

A. Notation

We adopt the lowercase letters to denote vectors, uppercase letter for matrices, normal font for nonrandom (deterministic) quantities, and the boldface letters for random quantities. The notation $(.)^*$ refers to conjugate transposition, $Tr(.)$ refer to the trace of its matrix argument, $\otimes$ for the Kronecker product, and $vec(.)$ for a vector formed by stacking the columns of its matrix argument. We shall also use $col(...)$ to denote a column vector formed by stacking its arguments on top of each other and $diag(...)$ to denote a (block) diagonal matrix formed from its argument. All vectors in our treatment are column vectors, with the exception of regression vectors, $\boldsymbol{u}_{k,i}$.

II. Partial Diffusion Algorithms with Imperfect Information Exchange

Consider a connected network consisting of N nodes. At time instant $i \geq 0$, each node k has access to scalar measurements $\boldsymbol{d}_k(i)$ and $1 \times M$ regression data vectors $\boldsymbol{u}_{k,i}$. The data across all nodes are assumed to be related to an unknown $M \times 1$ vector w^o via linear regression model of the form [17]:

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

where $\boldsymbol{v}_k(i)$ denotes the measurement noise with zero mean and variance $\sigma_{v,k}^2$ and the vector w^o refers to the parameter of interest.

We are now interested in solving optimization problems of the type:

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

The nodes in the network would like to estimate w^o by solving the equation above in adaptive and collaborative manners. We review the diffusion adaptation strategies with imperfect information exchange below.

A. Diffusion Adaptation with Imperfect Information Exchange

Consider the following general adaptive diffusion strategies with $C = I_N$ corresponding to the case in which the nodes only share the weight estimates for $i \geq 0$ [21]:

$$\boldsymbol{\phi}_{k,i} = \sum_{l \in \mathcal{N}_k} a_{1,lk}\boldsymbol{w}_{l,i-1} \tag{3}$$

$$\boldsymbol{\psi}_{k,i} = \boldsymbol{\phi}_{k,i-1} + \mu_k \boldsymbol{u}_{k,i}^*\left[\boldsymbol{d}_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{\phi}_{k,i-1}\right] \tag{4}$$

$$\boldsymbol{w}_{k,i} = \sum_{l \in \mathcal{N}_k} a_{2,lk}\boldsymbol{\psi}_{l,i} \tag{5}$$

The scalars $\{a_{1,lk}, a_{2,lk}\}$ are non-negative real coefficients corresponding to the (l,k) entries of $N \times N$ combination matrices $\{A_1, A_2\}$, respectively. The role of

these combination matrices is in convergence behavior of the diffusion strategy (3)-(5). These coefficients are zero whenever node $l \notin \mathcal{N}_k$, where $\mathcal{N}_k$ denotes the neighborhood of node k. These matrices are assumed to satisfy the conditions:

$$A_1^T \mathbb{1}_N = \mathbb{1}_N, \qquad A_2^T \mathbb{1}_N = \mathbb{1}_N \tag{6}$$

where the notation $\mathbb{1}$ denotes an $N \times 1$ column vector with all its entries equal to one.

We model the noisy data received by node k from its neighbor l as follows (see Figure 1):

$$\boldsymbol{w}_{lk,i-1} = \boldsymbol{w}_{l,i-1} + \boldsymbol{v}_{lk,i-1}^{(w)} \tag{7}$$

$$\boldsymbol{\psi}_{lk,i} = \boldsymbol{\psi}_{l,i} + \boldsymbol{v}_{lk,i}^{(\psi)} \tag{8}$$

where $\boldsymbol{v}_{lk,i-1}^{(w)}$ $(M \times 1)$ and $\boldsymbol{v}_{lk,i}^{(\psi)}$ $(M \times 1)$ are vector noise signal. They are temporally white and spatially independent random process with zero mean and covariance given by $\left\{R_{v,lk}^{(w)}, R_{v,lk}^{(\psi)}\right\}$. The quantities $\left\{R_{v,lk}^{(w)}, R_{v,lk}^{(\psi)}\right\}$ are all zero if $l \notin \mathcal{N}_k$ or when $l = k$. It should be noted that the subscript lk indicates that l is the source and k the sink, the flow of information is from l to k.

Using the perturbed data (7) and (8), the adaptive strategy (3)-(5) becomes

$$\boldsymbol{\phi}_{k,i} = \sum_{l \in \mathcal{N}_k} a_{1,lk}\boldsymbol{w}_{lk,i-1} \tag{9}$$

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

$$\boldsymbol{w}_{k,i} = \sum_{l \in \mathcal{N}_k} a_{2,lk}\boldsymbol{\psi}_{lk,i} \tag{11}$$

B. Partial-Diffusion with Impeferct Information Exchange

In order to lower the level of inter-node communication required among the nodes, we utilize partial-diffusion strategy proposed in [1], to transmit L out of M entries of the intermediate estimates at each time instant where the integer L is fixed and pre-specified. Again, we develop a more general class of PDLMS of which [1] is a special case. The selection of to-be-transmitted entries at node k and time instant i can be portrayed by an $M \times M$ diagonal entry-selection matrix, denoted by $\boldsymbol{\Lambda}_{k,i}$, that has L ones and $M - L$ zeros on its diagonal. The position of ones states the selected entries. Multiplication of an intermediate estimate vector by this matrix replaces its non-selected entries with zero.

According to (9) and (11) that can also be expressed as:

$$\boldsymbol{\phi}_{k,i} = a_{1,kk}\boldsymbol{w}_{k,i-1} + \sum_{l \in \mathcal{N}_k \setminus \{k\}} a_{1,lk}\left[\boldsymbol{\Lambda}_{l,i-1}\boldsymbol{w}_{lk,i-1} + \left(I_M - \boldsymbol{\Lambda}_{l,i-1}\right)\boldsymbol{w}_{lk,i-1}\right] \tag{12}$$

$$\boldsymbol{\psi}_{k,i} = \boldsymbol{\phi}_{k,i-1} + \mu_k \boldsymbol{u}_{k,i}^*\left[\boldsymbol{d}_k(i) - \boldsymbol{u}_{k,i}\boldsymbol{\phi}_{k,i-1}\right] \tag{13}$$

$$\boldsymbol{w}_{k,i} = a_{2,kk}\boldsymbol{\psi}_{k,i} + \sum_{l \in \mathcal{N}_k \setminus \{k\}} a_{2,lk}\left[\boldsymbol{\Lambda}_{l,i}\boldsymbol{\psi}_{lk,i} + \left(I_M - \boldsymbol{\Lambda}_{l,i}\right)\boldsymbol{\psi}_{lk,i}\right]$$
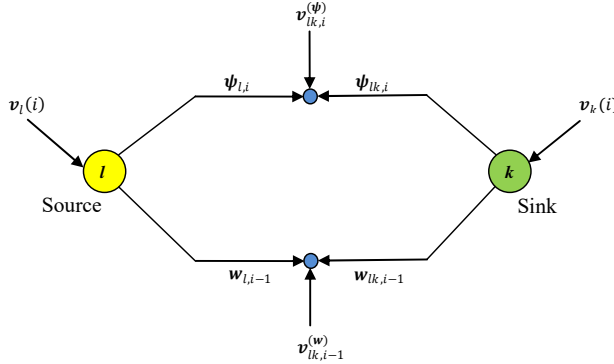
Figure 1. Several additive nise sources perturb the exchange of information from node l to node k.

Each node needs the information of all entries of its neighbors' intermediate estimate vectors for the consultation phase. However, when the intermediate estimates are broadcast partially ($0 < L < M$), nodes have no access to the non-communicated entries. To resolve this indistinctness, we allow the nodes utilize their own intermediate estimates entries instead of ones from the neighbors that have not been communicated, i.e., at node k, substitute

$$(I_M - \Lambda_{l,i-1})w_{k,i-1} \qquad \forall l \in \mathcal{N}_k \setminus \{k\} \qquad (15)$$

for

$$(I_M - \Lambda_{l,i-1})w_{lk,i-1} \qquad \forall l \in \mathcal{N}_k \setminus \{k\} \qquad (16)$$

and

$$(I_M - \Lambda_{l,i})\psi_{k,i} \qquad \forall l \in \mathcal{N}_k \setminus \{k\} \qquad (17)$$

for

$$(I_M - \Lambda_{l,i})\psi_{lk,i} \qquad \forall l \in \mathcal{N}_k \setminus \{k\} \qquad (18)$$

Based on this approach together with using perturbed data as introduced in (7) and (8), we formulate general PDLMS under noisy exchange as follows:

$$\phi_{k,i} = a_{1,kk}w_{k,i-1} + \sum_{l\in\mathcal{N}_k\setminus\{k\}} a_{1,lk}[\Lambda_{l,i-1}w_{l,i-1} + (I_M - \Lambda_{l,i-1})w_{k,i-1}] + \sum_{l\in\mathcal{N}_k\setminus\{k\}} a_{1,lk}\Lambda_{l,i-1}v_{lk,i-1}^{(w)} \qquad (19)$$

$$\psi_{k,i} = \phi_{k,i-1} + \mu_k u_{k,i}^*[d_k(i) - u_{k,i}\phi_{k,i-1}] \qquad (20)$$

$$w_{k,i} = a_{2,kk}\psi_{k,i} + \sum_{l\in\mathcal{N}_k\setminus\{k\}} a_{2,lk}[\Lambda_{l,i}\psi_{l,i} + (I_M - \Lambda_{l,i})\psi_{k,i}] + \sum_{l\in\mathcal{N}_k\setminus\{k\}} a_{2,lk}\Lambda_{l,i}v_{lk,i}^{(\psi)} \qquad (21)$$

Remark: The probability of transmission for all the entries at each node is equal and state as [1], [9]

$$p = {}^L/_M \qquad (22)$$

Moreover, the entry selection matrices, $\Lambda_{k,i}$, do not rely on any data/parameter with the exception of L and M.

Introduce the following aggregate $M \times 1$ zero mean noise signals:

$$v_{k,i-1}^{(w)} \triangleq \sum_{l\in\mathcal{N}_k\setminus\{k\}} a_{1,lk}\Lambda_{l,i-1}v_{lk,i-1}^{(w)} \qquad (23)$$

$$v_{k,i}^{(\psi)} \triangleq \sum_{l\in\mathcal{N}_k\setminus\{k\}} a_{2,lk}\Lambda_{l,i}v_{lk,i-1}^{(\psi)} \qquad (24)$$

These noises correspond to the cumulative effect on node k of all selected exchange noises from the neighbors of node k while exchanging the estimates $\{w_{l,i-1}, \psi_{l,i}\}$ in the course of the two consultation steps. The $M \times M$ covariance matrices of these noises are given by:

$$R_{v,k}^{(w)} \triangleq \sum_{l\in\mathcal{N}_k\setminus\{k\}} a_{1,lk}^2\Lambda_{l,i-1}R_{v,lk}^{(w)} \qquad (25)$$

$$R_{v,k}^{(\psi)} \triangleq \sum_{l\in\mathcal{N}_k\setminus\{k\}} a_{2,lk}^2\Lambda_{l,i}R_{v,lk}^{(\psi)} \qquad (26)$$

Thus, the PDLMS algorithm, Adapt Then Combine (ATC) approach, under noisy information exchange takes the following form:

TABLE I. PDLMS ALGORITHM UNDER NOISY INFORMATION EXCHANGE

Initialization:
Start with $w_{k,-1} = 0$ and given non-negative real coefficient $\{a_{lk}\}$, $A_1 = I_N$ and $A_2 = A$, satifying (6),
for $i \geq 0$, every node k computes
Step 1: Incremental Phase
$$\psi_{k,i} = w_{k,i-1} + \mu_k u_{k,i}^*[d_k(i) - u_{k,i}\phi_{k,i-1}]$$
Step 2: Diffusion Phase
$$w_{k,i} = a_{kk}\psi_{k,i} + \sum_{l\in\mathcal{N}_k\setminus\{k\}} a_{lk}[\Lambda_{l,i}\psi_{lk,i} + (I_M - \Lambda_{l,i})\psi_{lk,i}]$$

C. Entry Selection Method

In order to select L out of M entries of the intermediate estimates of each node at each iteration, the methods we utilized are comparable to the selection processes in stochastic and sequential *partial-update* schemes [12], [22], [23]. Here, we just review these methods namely *sequential* and *stochastic* partial-diffusion.

In sequential partial-diffusion the entry selection matrices, $\Lambda_{k,i}$, is a diagonal matrix:

$$\Lambda_{k,i} = \begin{bmatrix} r_{1,i} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & r_{M,i} \end{bmatrix}, \quad r_{\ell,i} = \begin{cases} 1 & if \ \ell \in \mathcal{I}_{(i \bmod \bar{B})+1} \\ 0 & otherwise \end{cases} \quad (27)$$

with $\bar{B} = \lceil M/L \rceil$. The number of selection entries at each iteration is limited by L. The coefficient subsets $\mathcal{I}_i$ are not unique as long as they meet the following requirements [12]:

1. Cardinality of $\mathcal{I}_i$ is between 1 and L;
2. $\bigcup_{r=1}^{\bar{B}} \mathcal{I}_r = \mathcal{S}$ where $\mathcal{S} = \{1,2,...,M\}$;
3. $\mathcal{I}_r \cap \mathcal{I}_p = \emptyset, \forall r, p \in \{1,...,\bar{B}\}$ and $r \neq p$.

The description of the entry selection matrices, $\mathbf{\Lambda}_{k,i}$, in stochastic partial-diffusion is similar to that of sequential one. The only difference is as follows. At a given iteration, $\boldsymbol{i}$, sequential case one of the set $\mathcal{J}_r, r = 1, \dots, \bar{B}$ is chosen in a predetermined fashion, whereas for stochastic case, one of the sets $\mathcal{J}_r$ is sampled at random from $\{\mathcal{J}_1, \mathcal{J}_2, \dots, \mathcal{J}_{\bar{B}}\}$. One might ask why these methods are considered to organize these selection matrices. To answer this question, it is worth mentioning that the nodes need to recognize which entries of their neighbors' intermediate estimates have been propagated at each iteration. These schemes bypass the need for addressing (position in the vector) [1], [9].

III. STEADY-STATE PERFORMANCE ANALYSIS

We now move on to examine the behavior of the general PDLMS implementations (19)-(21), and the influence of the mentioned perturbations on its convergence and steady-state performance. For this reason, we shall study the convergence of the weight estimates both in the mean and mean-square senses. In order to make the analysis tractable, we introduce the following assumptions on statistical properties of the measurement data and noise signals.

Assumptions:

1. The regression data $\boldsymbol{u}_{k,i}$ are temporally white and spatially independent random variables with zero mean and covariance matrix $R_{u,k} \triangleq \mathrm{E}\boldsymbol{u}_{k,i}^*\boldsymbol{u}_{k,i} \geq 0$.

2. The noise signal $\boldsymbol{v}_k(i)$, $\boldsymbol{v}_{k,i-1}^{(w)}$ and $\boldsymbol{v}_{k,i}^{(\psi)}$ are temporally white and spatially independent random variable with zero mean and covariance $\sigma_{v,k}^2$, $R_{v,k}^{(w)}$ and $R_{v,k}^{(\psi)}$, respectively. In addition, the quantities $\{R_{v,lk}^{(w)}, R_{v,lk}^{(\psi)}\}$ are all zero if $l \notin N_k$ or when $l = k$.

3. The regression data $\{\boldsymbol{u}_{m,i_1}\}$, the model noise signals $\boldsymbol{v}_n(i_2)$, and the link noise signals $\boldsymbol{v}_{l_1 k_1, j_1}^{(w)}$ and $\boldsymbol{v}_{l_2 k_2, j_2}^{(\psi)}$ are mutually independent random variables for all indexes $\{i_1, i_2, j_1, j_2, k_1, k_2, l_1, l_2, m, n\}$.

4. The step-sizes, $\mu_k \ \forall k$, are small enough such that their squared values are negligible.

We are interested in examining the evolution of the weight-error vectors. To do so, we let:

$$\widetilde{\boldsymbol{w}}_{k,i} \triangleq w^o - \boldsymbol{w}_{k,i} \qquad (28)$$

We import the information from across the network into block vectors and matrices as follows:

$$\mathcal{R}_{u,i} \triangleq diag\{\boldsymbol{u}_{1,i}^*\boldsymbol{u}_{1,i}, \boldsymbol{u}_{2,i}^*\boldsymbol{u}_{2,i}, \dots, \boldsymbol{u}_{N,i}^*\boldsymbol{u}_{N,i}\} \qquad (29)$$

$$\boldsymbol{s}_i \triangleq diag\{\boldsymbol{u}_{1,i}^*\boldsymbol{v}_1(i), \boldsymbol{u}_{2,i}^*\boldsymbol{v}_2(i), \dots, \boldsymbol{u}_{N,i}^*\boldsymbol{v}_N(i)\} \qquad (30)$$

$$\boldsymbol{v}_i^{(w)} \triangleq col\{\boldsymbol{v}_{1,i}^{(w)}, \boldsymbol{v}_{2,i}^{(w)}, \dots, \boldsymbol{v}_{N,i}^{(w)}\} \qquad (31)$$

$$\boldsymbol{v}_i^{(\psi)} \triangleq col\{\boldsymbol{v}_{1,i}^{(\psi)}, \boldsymbol{v}_{2,i}^{(\psi)}, \dots, \boldsymbol{v}_{N,i}^{(\psi)}\} \qquad (32)$$

$$\mathcal{M} \triangleq diag\{\mu_1 I_M, \dots, \mu_N I_M\} \qquad (33)$$

$$\widetilde{\boldsymbol{w}}_i \triangleq col\{\widetilde{\boldsymbol{w}}_{1,i}, \dots, \widetilde{\boldsymbol{w}}_{N,i}\} \qquad (34)$$

Subsequently, some algebra demonstrates that

$$\boxed{\begin{aligned}\widetilde{\boldsymbol{w}}_i &= \mathcal{A}_{2,i}(I_{NM} - \mathcal{M}\mathcal{R}_{u,i})\mathcal{A}_{1,i-1}\widetilde{\boldsymbol{w}}_{i-1} \\ &\quad -\mathcal{A}_{2,i}(I_{NM} - \mathcal{M}\mathcal{R}_{u,i})\boldsymbol{v}_{i-1}^{(w)} - \mathcal{A}_{2,i}\mathcal{M}\boldsymbol{s}_i - \boldsymbol{v}_i^{(\psi)}\end{aligned}} \qquad (35)$$

where

$$\mathcal{A}_{r,i} = \begin{bmatrix} A_{1,1,i} & \cdots & A_{1,N,i} \\ \vdots & \ddots & \vdots \\ A_{N,1,i} & \cdots & A_{N,N,i} \end{bmatrix}, \forall r \in \{1,2\} \qquad (36)$$

$$A_{p,q,i} = \begin{cases} I_M - \sum_{l \in N_p \backslash \{p\}} a_{r,lp}\mathbf{\Lambda}_{l,i} & if \ q = p \\ a_{r,qp}\mathbf{\Lambda}_{q,i} & if \ q \in N_p \backslash \{p\} \\ O_M & other \ wise \end{cases} \qquad (37)$$

A. Mean Performance

Taking expectation of both sides of (35) under *Remark* and *Assumptions*, we find that the mean error vector evolves according to the following recursion:

$$\mathbb{E}\widetilde{\boldsymbol{w}}_i = \mathcal{Q}_2 \mathrm{E}(I_{NM} - \mathcal{M}\mathcal{R}_u)\mathcal{Q}_1 \mathbb{E}\widetilde{\boldsymbol{w}}_{i-1} \qquad (38)$$

where

$$\mathcal{Q}_1 = \mathbb{E}[\mathcal{A}_{1,i}], \quad \mathcal{Q}_2 = \mathbb{E}[\mathcal{A}_{2,i-1}] \qquad (39)$$

From (38), we observe that in order for the recursion to be stable in the mean sense, the matrix $\mathcal{Q}_2 \mathrm{E}(I_{NM} - \mathcal{M}\mathcal{R}_u)\mathcal{Q}_1$ should be stable [24]. Picking $\mathcal{Q}_1$ and $\mathcal{Q}_2$ which all their entries are real non-negative and all their rows add up to unity [1]. Therefore, in the light of lemma 1 of [25], mean stability and asymptotic unbiasedness of the algorithm is guaranteed if the matrix $I_{NM} - \mathcal{M}\mathcal{R}_u$ is stable or equivalently if

$$|\lambda_{max}\{I_{NM} - \mathcal{M}\mathcal{R}_u\}| < 1 \qquad (40)$$

where $\lambda_{max}\{.\}$ refers to the largest eigenvalue of a matrix. The set of the eigenvalue of $I_{NM} - \mathcal{M}\mathcal{R}_u$ is the union of the set of the eigenvalue of $I_M - \mu_k R_{u,k} \ \forall k$ [25]. Thus, (40) is satisfied when $|\lambda_{max}\{I_M - \mu_k R_{u,k}\}| < 1, \forall k$ or $|1 - \mu_k \lambda_{max}\{R_{u,k}\}| < 1, \ \forall k$. These inequalities determine the stability bounds for step-sizes as

$$0 < \mu_k < \frac{2}{\lambda_{max}\{R_{u,k}\}} \quad \forall k \qquad (41)$$

B. Mean-Square Performance

The weighted variance relation for the error vector $\widetilde{w}_i$ can be obtained from the error recursion (35) as:

$$\mathbb{E}\|\widetilde{w}_i\|_\Sigma^2 = \mathbb{E}\|\widetilde{w}_{i-1}\|_{\Sigma'}^2 + \mathbb{E}\left(s_i^* \mathcal{M} \mathcal{A}_{2,i}^T \Sigma \mathcal{A}_{2,i} \mathcal{M} s_i\right)$$
$$+ \mathbb{E}\left(v_{i-1}^{*(w)} \mathcal{H}_i^* \Sigma \mathcal{H}_i v_{i-1}^{(w)}\right) + \mathbb{E}\left(v_i^{*(\psi)} \Sigma v_i^{(\psi)}\right) \tag{42}$$

where Σ is an arbitrary positive semi-definite Hermitian matrix of size $NM \times NM$ and

$$\mathcal{H}_i \triangleq \mathcal{A}_{2,i}\left(I_{NM} - \mathcal{M} \mathcal{R}_{u,i}\right) \tag{43}$$

Furthermore, the matrix Σ' can be expressed as

$$\Sigma' \triangleq \mathbb{E} \mathcal{B}_i^* \Sigma \mathcal{B}_i \tag{44}$$

with

$$\mathcal{B}_i \triangleq \mathcal{A}_{2,i}\left(I_{NM} - \mathcal{M} \mathcal{R}_{u,i}\right) \mathcal{A}_{1,i-1} \tag{45}$$

The variance relation becomes

$$\mathbb{E}\|\widetilde{w}_i\|_\Sigma^2 = \mathbb{E}\|\widetilde{w}_{i-1}\|_{\Sigma'}^2$$
$$+ \left(vec^T\{\mathcal{G}\}\mathcal{D}_2 + vec^T(\mathcal{H} R_v^{(w)} \mathcal{H}^*)\mathcal{D}_2\right.$$
$$\left. + vec^T\left(R_v^{(\psi)}\right)\right) vec(\Sigma) \tag{46}$$

where

$$\mathcal{D}_1 = \mathbb{E}\left(\mathcal{A}_{1,i-1}^T \otimes \mathcal{A}_{1,i-1}^T\right) \tag{47}$$

$$\mathcal{D}_2 = \mathbb{E}\left(\mathcal{A}_{2,i}^T \otimes \mathcal{A}_{2,i}^T\right) \tag{48}$$

$$\mathcal{G} = \mathcal{M} \mathbb{E}[s_i s_i^*] \mathcal{M} = diag\{\mu_1 \sigma_{v,1}^2 R_{u,1}, \ldots, \mu_N \sigma_{v,N}^2 R_{u,N}\} \tag{49}$$

$$\mathcal{H} \triangleq \mathbb{E}[I_{NM} - \mathcal{M} \mathcal{R}_{u,i}] = I_{NM} - \mathcal{M} \mathcal{R}_u \tag{50}$$

$$R_v^{(w)} \triangleq \mathbb{E} v_{i-1}^{(w)} v_{i-1}^{*(w)} = diag\{R_{v,1}^{(w)}, \ldots, R_{v,N}^{(w)}\} \tag{51}$$

$$R_v^{(\psi)} \triangleq \mathbb{E} v_i^{(\psi)} v_i^{*(\psi)} = diag\{R_{v,1}^{(\psi)}, \ldots, R_{v,N}^{(\psi)}\} \tag{52}$$

Therefore, the steady-state weighted variance relation (46) becomes

$$\lim_{i\to\infty} \mathbb{E}\|\widetilde{w}_i\|_{unvec\left[\left(I_{N^2M^2} - \mathcal{D}_1\left[\mathbb{E}(I_{NM} - \mathcal{R}_{u,i}\mathcal{M}) \otimes (I_{NM} - \mathcal{R}_{u,i}\mathcal{M})\right]\mathcal{D}_2\right) vec(\Sigma)\right]}^2$$

$$= \left(vec^T\{\mathcal{G}\}\mathcal{D}_2 + vec^T\left(\mathcal{H} R_v^{(w)} \mathcal{H}^*\right)\mathcal{D}_2 + vec^T\left(R_v^{(\psi)}\right)\right) vec(\Sigma) \tag{53}$$

It is known that a recursion of type (53) is stable and convergent if the matrix $\mathcal{D}_1\left[\mathbb{E}(I_{NM} - \mathcal{R}_{u,i}\mathcal{M}) \otimes (I_{NM} - \mathcal{R}_{u,i}\mathcal{M})\right]\mathcal{D}_2$ is stable [24], [26]. All the entries of $\mathcal{D}_1$ and $\mathcal{D}_2$ are real non-negative and all its columns sum up to one. Moreover, the eigenvalue of $\mathcal{T} \otimes \mathcal{T}$ are square of the eigenvalue of $\mathcal{T}$. Therefore, stability of this matrix has the same conditions as the stability of $I_{NM} - \mathcal{R}_u \mathcal{M}$. This means that choosing the step-sizes is accordance with (41) makes the algorithm stable likewise in the mean-square sense hence convergent to steady state.

The network MSD is defined as:

$$\text{MSD}^{network} \triangleq \lim_{i\to\infty} \frac{1}{N} \sum_{k=1}^N \mathbb{E}\|\widetilde{w}_{k,i}\|^2 \tag{54}$$

Since we are free to choose Σ, we select it as $I_{N^2M^2} - \mathcal{D}_1\left[\mathbb{E}(I_{NM} - \mathcal{R}_{u,i}\mathcal{M}) \otimes (I_{NM} - \mathcal{R}_{u,i}\mathcal{M})\right]\mathcal{D}_2 = vec(I_{NM}/N)$ then the expression (53) gives

$$\text{MSD}_{imperfect}^{network} = \frac{1}{N}\left[vec(\mathcal{G})\mathcal{D}_2 + vec(\mathcal{H} R_v^w \mathcal{H}^*)\mathcal{D}_2 + vec\left(R_v^{(\psi)}\right)\right]^T (I_{N^2M^2} - \mathcal{F})^{-1} vec(I_{NM}) \tag{55}$$

IV. NUMERICAL STUDIES

In order to illustrate the PLDMS strategies performance under noisy information exchange, we present some simulation results in this section.

A. Simulation

We consider an adaptive network with a random topology and $N = 10$ where each node is, in average, connected to two other nodes. The unknown parameter w^o of length $M = 8$ is randomly generated. We adopt a uniform step-size, $\mu_k = 0.01$, The measurements were generated according to model (1), and regressors, $u_{k,i}$, were chosen Gaussian i.i.d with randomly generated different diagonal covariance matrices, $R_{u,k}$. The additive noises at nodes are zero mean Gaussian with variances $\sigma_{v,k}^2$ and independent of the regression data. The traces of the covariance matrix regressors and the noise variances at all nodes, $Tr(R_{u,k})$ and $\sigma_{v,k}^2$, are shown in Figure 2.

We also use white Gaussian link noise signals such that $R_{v,lk}^{(w)} = \sigma_{w,lk}^2 I_M$ and $R_{v,lk}^{(\psi)} = \sigma_{\psi,lk}^2 I_M$. All link noise variances $\{\sigma_{w,lk}^2, \sigma_{\psi,lk}^2\}$ are randomly generated. The average power of each type of link noise across the network is 35 dB less than that of the model noise. In Figure 3, we plot the experimental network MSD curves for ATC case $(A_1 = I_N)$ of PDLMS algorithm using both sequential and stochastic partial diffusion schemes under noisy information exchange for different numbers of entries at each iteration, M. We use uniform weights for $\{a_{1,lk}, a_{2,lk}\}$ at combination phase at this stage. In Figure 4, we compare network MSD learning curve of PDLMS for both ATC and CTA, Combine Then Adapt, cases $(A_2 = I_N)$ under ideal and noisy links.
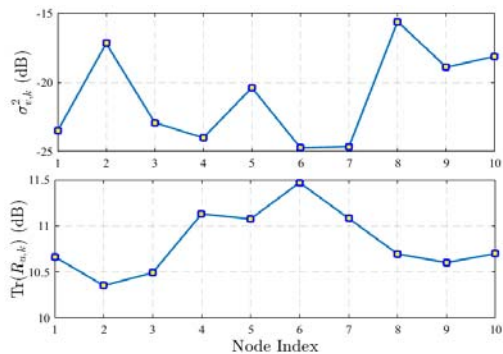
Figure 2. Covariance matrix trace of the input signal and the variance of the noise at each node.
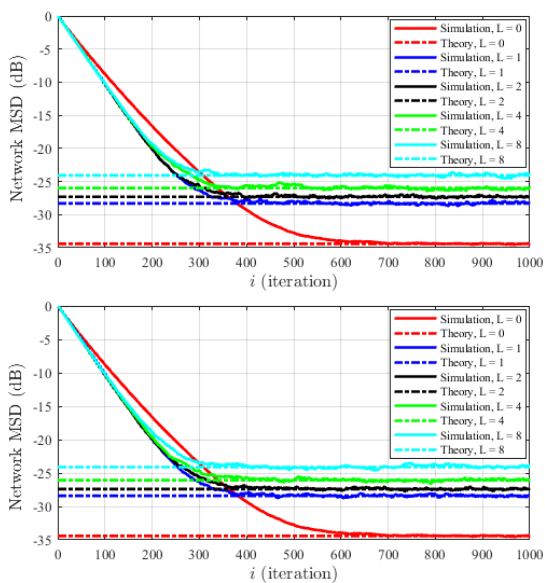

Figure 3. Network MSD learning curve of PDLMS for ATC under noisy links (top) sequential and (bottom) stochastic.

B. Discussion

From the simulation results, we make the following observations:

In [1], authors emphasized that the PDLMS algorithm delivers a trade-off between communications cost and estimation performance. However, noisy links add a new complexity to the network optimization problem. In addition, in the presence of noisy links, the trade-off between communication cost and estimation performance becomes unbalanced. This is because as more entries are broadcast at each iteration, more perturbed weight estimates are interred in consultant phase.

The sequential partial-diffusion schemes outperform the stochastic partial-diffusion. Finally, the adaptive ATC strategy outperforms the adaptive CTA strategy for both perfect and imperfect cases.
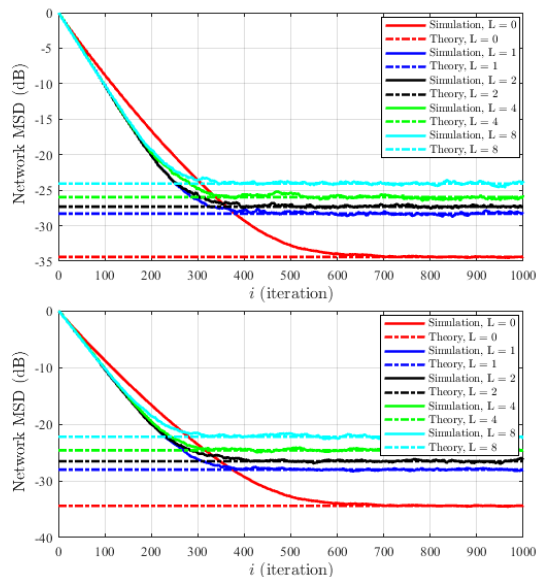

Figure 4. Comparison of network MSDs of sequential PDLMS algorithm for (top) ATC and (bottom) CTA under noisy information exchange

V. CONCLUSION

In this work, we presented a general form of PDLMS algorithms, formulated the ATC and CTA versions of PDLMS under noisy Links, and investigated the performance of partial-diffusion algorithms under several sources of noise during information exchange for both sequential and stochastic schemes. We also illustrated that the PDLMS strategy could still stabilize the mean and mean-square convergence of the network with noisy information exchange. We derived analytical expressions for network learning curve MSD. The important result is that the noisy links are the main factor in performance degradation of a diffusion LMS strategy running in a network with imperfect communication. Furthermore, there is no direct relation between the MSD performance and number of selected entries under imperfect information exchange. The performance degradation incurred by noisy links depends not only on the link noise variances, $\sigma_{w,lk}^2$ and $\sigma_{\psi,lk}^2$, but also on the other parameters of the network, i.e., the measurement and state noise variances, the network topology, and combination weights. These topics will be addressed in future work.

REFERENCE

[1] R. Arablouei, S. Werner, Y.-F. Huang, and K. Dogancay, "Distributed least mean-square estimation with partial diffusion," *Signal Process. IEEE Trans.*, vol. 62, no. 2, pp. 472–484, 2014.

[2] V. Vahidpour, A. Rastegarnia, A. Khalili, and S. Sanei, "Analysis of partial diffusion recursive least squares adaptation over noisy links," *IET Signal Process.*, 2017.

[3] V. Vahidpour, A. Rastegarnia, A. Khalili, and S. Sanei, "Partial

Diffusion Kalman Filtering for Distributed State Estimation in Multiagent Networks," *Neural Networks Learn. Syst. IEEE Trans.*, 2019.

[4] E. E. Tsiropoulou, S. T. Paruchuri, and J. S. Baras, "Interest, energy and physical-aware coalition formation and resource allocation in smart IoT applications," in *2017 51st Annual Conference on Information Sciences and Systems (CISS)*, 2017, pp. 1–6.

[5] M. O. Sayin and S. S. Kozat, "Compressive diffusion strategies over distributed networks for reduced communication load," *Signal Process. IEEE Trans.*, vol. 62, no. 20, pp. 5308–5323, 2014.

[6] M. O. Sayin and S. S. Kozat, "Single bit and reduced dimension diffusion strategies over distributed networks," *Signal Process. Lett. IEEE*, vol. 20, no. 10, pp. 976–979, 2013.

[7] S. Chouvardas, K. Slavakis, and S. Theodoridis, "Trading off complexity with communication costs in distributed adaptive learning via Krylov subspaces for dimensionality reduction," *Sel. Top. Signal Process. IEEE J.*, vol. 7, no. 2, pp. 257–273, 2013.

[8] R. Arablouei, K. Dogancay, S. Werner, and Y.-F. Huang, "Adaptive distributed estimation based on recursive least-squares and partial diffusion," *Signal Process. IEEE Trans.*, vol. 62, no. 14, pp. 3510–3522, 2014.

[9] V. Vahidpour, A. Rastegarnia, A. Khalili, W. M. Bazzi, and S. Sanei, "Analysis of Partial Diffusion LMS for Adaptive Estimation Over Networks with Noisy Links," *IEEE Trans. Netw. Sci. Eng.*, 2017.

[10] J. R. Deller Jr and Y. F. Huang, "Set-membership identification and filtering for signal processing applications," *Circuits, Syst. Signal Process.*, vol. 21, no. 1, pp. 69–82, 2002.

[11] S. Gollamudi, S. Nagaraj, S. Kapoor, and Y.-F. Huang, "Set-membership filtering and a set-membership normalized LMS algorithm with an adaptive step size," *Signal Process. Lett. IEEE*, vol. 5, no. 5, pp. 111–114, 1998.

[12] K. Dogancay, *Partial-update adaptive signal processing: Design Analysis and Implementation*. Academic Press, 2008.

[13] S. Werner, T. Riihonen, and Y.-F. Huang, "Energy-efficient distributed parameter estimation with partial updates," in *Green Circuits and Systems (ICGCS), 2010 International Conference on*, 2010, pp. 36–40.

[14] S. Werner and Y.-F. Huang, "Time-and coefficient-selective diffusion strategies for distributed parameter estimation," in *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*, 2010, pp. 696–697.

[15] A. Malipatil, Y.-F. Huang, and S. Werner, "An SMF approach to distributed average consensus in clustered sensor networks," in *Signal Processing Advances in Wireless Communications, 2009. SPAWC'09. IEEE 10th Workshop on*, 2009, pp. 81–85.

[16] R. Abdolee and B. Champagne, "Diffusion LMS algorithms for sensor networks over non-ideal inter-sensor wireless channels," in *Distributed Computing in Sensor Systems and Workshops (DCOSS), 2011 International Conference on*, 2011, pp. 1–6.

[17] A. Khalili, M. A. Tinati, and A. Rastegarnia, "Performance analysis of distributed incremental LMS algorithm with noisy links," *Int. J. Distrib. Sens. Networks*, vol. 2011, 2011.

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

[19] A. Khalili, M. A. Tinati, A. Rastegarnia, and J. A. Chambers, "Transient analysis of diffusion least-mean squares adaptive networks with noisy channels," *Int. J. Adapt. Control Signal Process.*, vol. 26, no. 2, pp. 171–180, 2012.

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

[21] A. H. Sayed, "Adaptive Filters. Hoboken." NJ: John Wiley & Sons, 2008.

[22] A. H. Sayed, "Diffusion adaptation over networks," *Acad. Press Libr. Signal Process.*, vol. 3, pp. 323–454, 2013.

[23] M. Godavarti and A. O. Hero III, "Partial update LMS algorithms," *Signal Process. IEEE Trans.*, vol. 53, no. 7, pp. 2382–2399, 2005.

[24] C. D. Meyer, *Matrix analysis and applied linear algebra*, vol. 2. Siam, 2000.

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

[26] K. M. Abadir and J. R. Magnus, *Matrix algebra*, vol. 1. Cambridge University Press, 2005.