Computing (2023) 105:2443–2458
https://doi.org/10.1007/s00607-023-01187-5

REGULAR PAPER

Steady-state analysis of diffusion least-mean squares with deficient length over wireless sensor networks

Ghanbar Azarnia[1]

Received: 14 October 2022 / Accepted: 28 May 2023 / Published online: 19 June 2023

Abstract
In recent years, distributed adaptive processing has received much attention from both theoretical and practical aspects. One of the efficient cooperation structures in distributed adaptive processing is the diffusion strategy, which provides a platform for the cooperation of nodes that run an adaptive algorithm, such as the least mean-squares (LMS) algorithm. Despite the studies that have been done on the diffusion-based LMS algorithm, the effect of deficient length on such structures has been overlooked. Accordingly, in this paper, we study the steady-state performance of the deficient length diffusion LMS algorithm. The results of this study show, in particular, that setting the tap length below its actual value leads to drastic degradation of the steady-state excess mean-square error (EMSE) and mean-square deviation (MSD) in diffusion adaptive networks. Furthermore, unlike the full-length case, where the steady-state MSD and EMSE decrease significantly with the step size reduction, this study shows that in the deficient-length scenario, there are no significant improvements in the steady-state performance by reducing the step size. Therefore, according to this study, the tap length plays a key role in diffusion adaptive networks since the performance deterioration due to deficient selection of tap length could not be compensated by an adjustment in the step size. Experiments exhibit a very good match between simulations and theory.

Keywords Wireless sensor networks · Diffusion least-mean squares · Deficient length · Distributed adaptive estimation · Steady-state analysis

1 Introduction

Distributed parameter estimation is gathered more increased interest in the wireless sensor networks (WSNs) application areas. It just relies on the local information exchange between neighbouring sensors. It, hence, eliminates the necessity of

✉ Ghanbar Azarnia
g.azarnia@uut.ac.ir

1 Engineering Faculty of Khoy, Urmia University of Technology, Urmia, Iran

a powerful fusion center to collect and process the data from the entire sensor network and, as such, decreases the communications power and bandwidth of the conventional centralized approach while preserving similar performance. There are two widely considered topologies, diffusion and incremental, for distributed estimation over adaptive networks. [1]. However, amongst these structures, diffusion topologies show superior performance over incremental topologies in terms of robustness to the nodes and link failures and are manageable to distributed implementations [2]. As a result, diffusion-based adaptive networks are widely studied that can learn and adapt from consecutive data streaming and exhibit good tracking capability and fast convergence rates. Besides, based on the diffusion structure, different diffusion techniques have been presented, such as the diffusion recursive least squares (D-RLS) [3], diffusion least mean square (D-LMS) [4], diffusion Kalman filter [5], diffusion affine projection (D-APA) [6], and diffusion improved multiband-structured subband adaptive filter (D-IMSAF) [7]. The convex combination of two D-LMS algorithms is proposed in [8] to overcome the tradeoff between the convergence rate and steady-state error in the traditional D-LMS algorithm. Two D-LMS algorithms with different step sizes are combined in this algorithm, resulting in a lower steady-state error and faster convergence rate at the cost of increased computational complexity. Also, motivated by the idea of combining adaptive algorithms, [9] proposed an affine combination framework for diffusion topologies, where affine combination coefficients are adjusted based on the minimum mean-square-error (MSE) criterion. This combined diffusion structure enjoys the best characteristics of all component strategies. Since the conventional diffusion adaptive networks experience a significant performance degeneration in the presence of impulsive noise, robust D-RLS algorithms are developed in [10] to enhance the performance in such noisy scenarios. In another approach, to mitigate the performance deterioration experienced in diffusion adaptive networks in the presence of impulsive noise, [11] proposed a diffusion-based affine projection M-estimate algorithm, which employs a robust M-estimate based cost function. Paper [12] has studied the performance analysis of diffusion adaptive networks by considering the communication delays of the links between nodes. This paper presents the stability conditions in the mean and mean-square sense. This paper shows that the delayed D-LMS algorithm could converge under the same step-sizes limitation of the conventional D-LMS algorithm without considering delays. The transient performance of the D-LMS algorithm for the non-stationary systems has been studied in [13], where each node employs the different types of cyclostationary white non-Gaussian signals.

In [14], the authors have developed the zeroth-order (ZO)-diffusion algorithm by utilizing the gradient-free approach to the adaptation phase of the conventional diffusion structure. Furthermore, a modification of the stochastic variance reduced gradient (SVRG) named time-averaging SVRG (TA-SVRG) is provided in [14] for streaming data processing. Finally, the TA-SVRG algorithm has been applied in the ZO-diffusion to decrease the estimation variance and enhance the convergence rate.

Distributed diffusion adaptive techniques are implemented in two phases: the combination phase, in which each sensor combines its neighbourhood information, and the adaptation phase, in which each sensor updates its estimate based on an adaptive rule. The performance of diffusion networks could be enhanced by developing combination policies that set the combination coefficients based on the data quality. Accordingly, the design of combination coefficients has been considered in the literature [15–18].

As discussed, the above-mentioned papers have been developed to estimate the elements of an unknown parameter vector of interest in a distributed manner. These cooperative methods utilize a simple assumption that the dimension of the unknown vector or the tap-length is known a priori. So, they overlook one crucial point: in some contexts, the optimal length of the desired parameter is also unknown, similar to its elements. On this basis, the tap-length estimation problem has been considered in the distributed context based on diffusion [19–21] and incremental strategies [22, 23].

However, in some cases, for reasons such as power storage (since power consumption is a critical concern in WSNs), such variable tap length approaches are not applicable. So a suppositional length for the unknown vector is considered at each sensor. Usually, the exact length of the filter is unknown in advance, so it is possible to apply in each sensor node a deficient tap-length adaptive filter, i.e., an adaptive filter whose tap-length is smaller than that of the unknown desired vector length.

In distributed adaptive filter settings, the steady-state study for deficient length scenarios has been conducted only in incremental-based adaptive networks [24]. Despite the importance of studying the deficient length scenario for diffusion LMS algorithm, in almost all diffusion adaptive networks theoretical analysis, it is supposed that the length of the adaptive filter is the same as that of the unknown desired vector. On the other hand, theoretical findings on sufficient tap-length diffusion adaptive algorithms could not necessarily apply to the practical deficient length cases. Accordingly, from practical aspects, it becomes very critical to analyze and quantify the statistical behavior of the deficient length diffusion adaptive algorithms.

On this basis, in this paper, we study the deficient length diffusion LMS adaptive networks' performance in the steady state. More precisely, we derive closed-form expressions for the EMSE and MSD for each node to explain the efficiency of length deficiency on the steady-state performance of each node. This study shows that setting the tap length smaller than its actual value leads to drastic degradation of the steady-state EMSE and MSD in diffusion adaptive networks. Also, unlike the full-length case, where the steady-state MSD and EMSE decrease significantly with the step size reduction, this study shows that in the deficient-length scenario, there are no significant improvements in the steady-state performance by reducing the step size. Consequently, according to this study, the tap length plays a key role in diffusion adaptive networks since the performance deterioration due to deficient selection of tap length could not be compensated by an adjustment in the step size.

*Notation*As a convenience to the reader, Tables 1 and 2 list the main acronyms and symbols used in this paper.

2 Background

Consider a network of J sensors implemented in a diffusion structure to estimate the unknown vector $\boldsymbol{w}^o_{L_{opt}}$ of length L_{opt}. It is supposed that the network nodes have no prior information about the length L_{opt}, and each node is equipped with a filter of length

Table 1 List of main acronyms

Acronym	Description
CTA	Combine-Then-Adapt
D-RLS	Diffusion recursive least squares
EMSE	Excess mean-square error
LMS	Least mean-squares
MSD	Mean-square deviation
MSE	Mean-square-error
WSN	Wireless sensor network

Table 2 The main symbols

Symbol	Description
$\|x\|$	The absolute value of a scalar x
$\|\boldsymbol{x}\|_2$	The ℓ_2-norm of a vector $\boldsymbol{x}$
$\|\boldsymbol{x}\|^2_{\Sigma}$	The weighted norm for a column vector $\boldsymbol{x}$; $\boldsymbol{x}^* \sum \boldsymbol{x}$
$E\{X\}$	The statistical expectation of X
$\mathbf{0}_{N\times 1}$	The column vector of zeros of length N
$\mathbf{0}_N$	The $N \times N$ square-zero matrix
$\mathbf{0}_{N\times M}$	The $N \times M$ zero matrix
I_N	The $N \times N$ identity matrix
$()^T$	The transpose of a matrix or a vector
$()^*$	The adjoint (or Hermitian transpose)
tr(.)	The trace of a matrix
$\boldsymbol{A} \bigotimes \boldsymbol{B}$	The Kronecker product of two matrices
$\boldsymbol{A} \bigodot \boldsymbol{B}$	The block Kronecker product of two block matrices
vec	The standard vectorization operator
bvec	The block vectorization operator
diag{.}	The (block) diagonal matrix consisting of the specified arguments
col{.}	The column vector obtained by stacking the specified arguments

$N \leqslant L_{opt}$. Considering the Combine-Then-Adapt (CTA) strategy, the filter coefficients update equation in diffusion adaptive networks would be as [2]:

$$\begin{cases} \boldsymbol{\varphi}_k^{(i-1)} = \sum_{\ell \in \mathcal{J}_k} c_{k,\ell} \boldsymbol{\psi}_\ell^{(i-1)} \\ \boldsymbol{\psi}_k^{(i)} = \boldsymbol{\varphi}_k^{(i-1)} + \mu_k \boldsymbol{u}_{k,i}^* \left(d_k(i) - \boldsymbol{u}_{k,i} \boldsymbol{\varphi}_k^{(i-1)} \right) \end{cases} \tag{1}$$

where the $N \times 1$ column vector $\boldsymbol{\psi}_k^{(i)}$ represents the deficient-length local estimate of $\boldsymbol{w}^o_{L_{opt}}$ at sensor k and time i, and where $\mathcal{J}_k$ stands for the topologically connected nodes to node k (the neighborhood of sensor k, including itself). Note that, in (1), all vectors have length N. In (1), the positive constant μ_k is the step-size at sensor k, and the local combiners $\{c_{k,\ell} \geqslant 0\}$ meet the requirement $\sum_{\ell=1}^{J} c_{k,\ell} = 1$. The combiners $\{c_{k,\ell}\}$ are collected in a matrix $C = [c_{k,\ell}]$ that reflects the network topology. In (1), the data are supposed to follow the linear model [25]:

$$d_k(i) = \boldsymbol{u}_{L_{opt}k,i}\boldsymbol{w}^o_{L_{opt}} + v_k(i) \tag{2}$$

In this equation, $v_k(i)$ represents white background noise with zero-mean and variance $\sigma^2_{v,k}$, which is assumed to be independent of $\boldsymbol{u}_{L_{opt}\ell,j}$ for all ℓ, j.

In (2), $\boldsymbol{u}_{L_{opt}k,i}$ is a vector of length L_{opt}, and $\boldsymbol{u}_{k,i}$ consists of N first elements of $\boldsymbol{u}_{L_{opt}k,i}$. In this regard, $\boldsymbol{u}_{L_{opt}k,i}$ could be written as $\boldsymbol{u}_{L_{opt}k,i} = \left[\boldsymbol{u}_{k,i}, \bar{\boldsymbol{u}}_{k,i}\right]$, where $\boldsymbol{u}_{k,i} = \left[u_k(i), u_k(i-1), ..., u_k(i-N+1)\right]$ and $\bar{\boldsymbol{u}}_{k,i} = \left[u_k(i-N), ..., u_k\left(i - L_{opt} + 1\right)\right]$, with $u_k(i)$ representing the ith element of $\boldsymbol{u}_{k,i}$.

3 Steady-state analysis of deficient tap-length CTA diffusion LMS algorithm

Now, we intend to analyze the steady-state behavior of the CTA diffusion LMS algorithm when a conjectural deficient length is applied for the unknown parameter. In this regard, the unknown vector is treated as the composition of two vectors stacked on top of each other as $\begin{bmatrix} \boldsymbol{w}^{(1)}_N \\ \boldsymbol{w}^{(2)}_{L_{opt}-N} \end{bmatrix}$, where $\boldsymbol{w}^{(1)}_N$ consists of the first N entries of $\boldsymbol{w}^o_{L_{opt}}$, and $\boldsymbol{w}^{(2)}_{L_{opt}-N}$ consists of the $L_{opt} - N$ final elements.

To assess the steady-state performance, we analyze the MSD and EMSE measures for each sensor k, which are defined as [2]:

$$\mathrm{MSD}_k = E\left\{ \left\| \overline{\boldsymbol{\psi}}^{(i-1)}_{L_{opt},k} \right\|^2 \right\} \tag{3a}$$

$$\mathrm{EMSE}_k = E\left\{ \left| \boldsymbol{u}_{L_{opt},k,i}\overline{\boldsymbol{\psi}}^{(i-1)}_{L_{opt},k} \right|^2 \right\} \tag{3b}$$

where $\overline{\boldsymbol{\psi}}^{(i-1)}_{L_{opt},k}$ is defined as:

$$\begin{aligned} \overline{\boldsymbol{\psi}}^{(i-1)}_{L_{opt},k} &= -\boldsymbol{w}^o_{L_{opt}} + \begin{bmatrix} \boldsymbol{\psi}^{(i-1)}_k \\ \boldsymbol{0}_{(L_{opt}-N)\times 1} \end{bmatrix} \\ &= \begin{bmatrix} -\boldsymbol{w}^{(1)}_N \\ -\boldsymbol{w}^{(2)}_{L_{opt}-N} \end{bmatrix} \\ &+ \begin{bmatrix} \boldsymbol{\psi}^{(i-1)}_k \\ \boldsymbol{0}_{(L_{opt}-N)\times 1} \end{bmatrix} = \begin{bmatrix} \overline{\boldsymbol{\psi}}^{(i-1)}_{N,k} \\ -\boldsymbol{w}^{(2)}_{L_{opt}-N} \end{bmatrix} \end{aligned} \tag{4}$$

where $\overline{\boldsymbol{\psi}}^{(i-1)}_{N,k} = \boldsymbol{\psi}^{(i-1)}_k - \boldsymbol{w}^{(1)}_N$. With the definition of $\overline{\boldsymbol{\psi}}^{(i-1)}_{L_{opt},k}$, (3) could be rewritten as:

$$\mathrm{MSD}_k = E\left\{ \left\| \overline{\boldsymbol{\psi}}_{N,k}^{(i-1)} \right\|^2 \right\} + \left\| \boldsymbol{w}_{L_{opt}-N}^{(2)} \right\|^2 \tag{5a}$$

$$\begin{aligned}\mathrm{EMSE}_k &= E\left\{ \left| \left[\boldsymbol{u}_{k,i} \ \overline{\boldsymbol{u}}_{k,i} \right] \begin{bmatrix} \overline{\boldsymbol{\psi}}_{N,k}^{(i-1)} \\ -\boldsymbol{w}_{L_{opt}-N}^{(2)} \end{bmatrix} \right|^2 \right\} = E\left\{ \left| \boldsymbol{u}_{k,i} \overline{\boldsymbol{\psi}}_{N,k}^{(i-1)} - \overline{\boldsymbol{u}}_{k,i} \boldsymbol{w}_{L_{opt}-N}^{(2)} \right|^2 \right\} \\ &= E\left\{ \left\| \overline{\boldsymbol{\psi}}_{N,k}^{(i-1)} \right\|^2_{R_{\boldsymbol{u}_k}} \right\} + \left\| \boldsymbol{w}_{L_{opt}-N}^{(2)} \right\|^2_{R_{\overline{\boldsymbol{u}}_k}}\end{aligned} \tag{5b}$$

where $R_{\boldsymbol{u}_k} = E\left\{ \boldsymbol{u}_{k,i}^{*} \boldsymbol{u}_{k,i} \right\}$ and $R_{\overline{\boldsymbol{u}}_k} = E\left\{ \overline{\boldsymbol{u}}_{k,i}^{*} \overline{\boldsymbol{u}}_{k,i} \right\}$. The components of the regressors are assumed to be uncorrelated zero-mean Gaussian processes with variance $\sigma_{u,k}^2$. It is also assumed that $\boldsymbol{u}_{L_{opt}k,i}$ is independent of $\boldsymbol{u}_{L_{opt}\ell,j}$ for $k \neq \ell$ and $i \neq j$. This assumption will imply the independence of $\boldsymbol{u}_{k,i}$ from $\overline{\boldsymbol{\psi}}_{N,k}^{(i-1)}$. To proceed, we rewrite (1) as:

$$\begin{aligned}\boldsymbol{\varphi}_k^{(i-1)} &= \sum_{\ell \in J_k} c_{k,\ell} \boldsymbol{\psi}_\ell^{(i-1)} \\ \boldsymbol{\psi}_k^{(i)} &= \boldsymbol{\varphi}_k^{(i-1)} + \mu_k \boldsymbol{u}_{k,i}^{*} \left[\left(\boldsymbol{u}_{L_{opt},k,i} \boldsymbol{w}_{L_{opt}}^{o} + v_k(i) \right) - \boldsymbol{u}_{k,i} \boldsymbol{\varphi}_k^{(i-1)} \right] \\ &= \boldsymbol{\varphi}_k^{(i-1)} + \mu_k \boldsymbol{u}_{k,i}^{*} \left[\left[\boldsymbol{u}_{k,i} \ \overline{\boldsymbol{u}}_{k,i} \right] \begin{bmatrix} \boldsymbol{w}_N^{(1)} \\ \boldsymbol{w}_{L_{opt}-N}^{(2)} \end{bmatrix} + v_k(i) - \boldsymbol{u}_{k,i} \boldsymbol{\varphi}_k^{(i-1)} \right] \\ &= \boldsymbol{\varphi}_k^{(i-1)} + \mu_k \boldsymbol{u}_{k,i}^{*} \left[\boldsymbol{u}_{k,i} \boldsymbol{w}_N^{(1)} + \overline{\boldsymbol{u}}_{k,i} \boldsymbol{w}_{L_{opt}-N}^{(2)} - \boldsymbol{u}_{k,i} \boldsymbol{\varphi}_k^{(i-1)} + v_k(i) \right] \\ &= \sum_{\ell \in J_k} c_{k,\ell} \boldsymbol{\psi}_\ell^{(i-1)} + \mu_k \boldsymbol{u}_{k,i}^{*} \\ &\quad \left[-\boldsymbol{u}_{k,i} \left(\sum_{\ell \in J_k} c_{k,\ell} \left(\boldsymbol{\psi}_\ell^{(i-1)} - \boldsymbol{w}_N^{(1)} \right) \right) + \overline{\boldsymbol{u}}_{k,i} \boldsymbol{w}_{L_{opt}-N}^{(2)} + v_k(i) \right]\end{aligned} \tag{6}$$

where the condition $\sum_{\ell \in J_k} c_{k,\ell} = 1$ is imposed. Let us subtract $\boldsymbol{w}_N^{(1)}$ from both sides of (6):

$$\begin{aligned}\overline{\boldsymbol{\psi}}_{N,k}^{(i)} = \sum_{\ell \in J_k} c_{k,\ell} \overline{\boldsymbol{\psi}}_{N,\ell}^{(i-1)} - \mu_k \boldsymbol{u}_{k,i}^{*} \boldsymbol{u}_{k,i} \sum_{\ell \in J_k} c_{k,\ell} \overline{\boldsymbol{\psi}}_{N,\ell}^{(i-1)} \\ + \mu_k \boldsymbol{u}_{k,i}^{*} \overline{\boldsymbol{u}}_{k,i} \boldsymbol{w}_{L_{opt}-N}^{(2)} + \mu_k \boldsymbol{u}_{k,i}^{*} v_k(i)\end{aligned} \tag{7}$$

For ease of analysis, we define some quantities:

$$\boldsymbol{U}_i = \mathrm{diag}\{\boldsymbol{u}_{1,i}, \ldots, \boldsymbol{u}_{J,i}\} \qquad \overline{\boldsymbol{U}}_i = \mathrm{diag}\{\overline{\boldsymbol{u}}_{1,i}, \ldots, \overline{\boldsymbol{u}}_{J,i}\}$$

$$\boldsymbol{D} = \mathrm{diag}\{\mu_1 I_N, \ldots, \mu_J I_N\} \qquad \overline{\boldsymbol{D}} = \mathrm{diag}\left\{\mu_1 I_{L_{opt}-N}, \ldots, \mu_J I_{L_{opt}-N}\right\}$$

$$\overline{\boldsymbol{\Psi}}_N^{(i)} = \mathrm{col}\left\{\overline{\boldsymbol{\psi}}_{N,1}^{(i)}, \ldots, \overline{\boldsymbol{\psi}}_{N,J}^{(i)}\right\} \qquad \mathcal{W}_{L_{opt}-N}^{(O)} = \mathrm{col}\left\{\boldsymbol{w}_{L_{opt}-N}^{(2)}, \ldots, \boldsymbol{w}_{L_{opt}-N}^{(2)}\right\}$$

$$\boldsymbol{V}_i = \mathrm{col}\{\boldsymbol{v}_1(i), \ldots, \boldsymbol{v}_J(i)\}$$

Let $C = \{c_{k,\ell}\}$ represents the combination coefficient matrix; we also define $G = C \bigotimes I_N$, where $\bigotimes$ represents the standard Kronecker product.

With the expressed notation, we rewrite (7) as:

$$\overline{\boldsymbol{\Psi}}_N^{(i)} = G\overline{\boldsymbol{\Psi}}_N^{(i-1)} - \boldsymbol{D}\boldsymbol{U}_i^*\boldsymbol{U}_i G\overline{\boldsymbol{\Psi}}_N^{(i-1)} + \boldsymbol{D}\boldsymbol{U}_i^*\overline{\boldsymbol{U}}_i \mathcal{W}_{L_{opt}-N}^{(O)} + \boldsymbol{D}\boldsymbol{U}_i^*\boldsymbol{V}_i \tag{8}$$

Also, with the expressed notation, the local MSD and EMSE at sensor k can be rewritten as:

$$\mathrm{MSD}_k = E\left\{ \left\|\overline{\boldsymbol{\Psi}}_N^{(i-1)}\right\|^2_{\mathrm{diag}\{\mathbf{0}_{(k-1)N}, I_N, \mathbf{0}_{(J-k)N}\}} \right\} + \left\|\boldsymbol{w}_{L_{opt}-N}^{(2)}\right\|^2 \tag{9a}$$

$$\begin{aligned} \mathrm{EMSE}_k = {} & E\left\{ \left\|\overline{\boldsymbol{\Psi}}_N^{(i-1)}\right\|^2_{\mathrm{diag}\left\{\mathbf{0}_{(k-1)N}, R_{\boldsymbol{u}_k}, \mathbf{0}_{(J-k)N}\right\}} \right\} \\ & + \left\|\mathcal{W}_{L_{opt}-N}^{(O)}\right\|^2_{\mathrm{diag}\left\{\mathbf{0}_{(k-1)\left(L_{opt}-N\right)}, R_{\overline{\boldsymbol{u}}_k}, \mathbf{0}_{(J-k)\left(L_{opt}-N\right)}\right\}} \end{aligned} \tag{9b}$$

Also, to characterize the network performance, the network MSD and EMSE could be considered as:

$$\begin{aligned} \mathrm{MSD}^{Net} &= \frac{1}{J}\sum_{k=1}^{J} MSD_k = \frac{1}{J}\sum_{k=1}^{J} E\left\{ \left\|\overline{\boldsymbol{\psi}}_{N,k}^{(i-1)}\right\|^2 \right\} + \left\|\boldsymbol{w}_{L_{opt}-N}^{(2)}\right\|^2 \\ &= \frac{1}{J} E\left\{ \left\|\overline{\boldsymbol{\Psi}}_N^{(i-1)}\right\|^2_{I_{NJ}} \right\} + \left\|\boldsymbol{w}_{L_{opt}-N}^{(2)}\right\|^2 \end{aligned} \tag{10a}$$

$$\begin{aligned} \mathrm{EMSE}^{Net} &= \frac{1}{J}\sum_{k=1}^{J} EMSE_k = \frac{1}{J}\sum_{k=1}^{J} E\left\{ \left\|\overline{\boldsymbol{\psi}}_{N,k}^{(i-1)}\right\|^2_{R_{\boldsymbol{u}_k}} \right\} + \frac{1}{J}\sum_{k=1}^{J} \left\|\boldsymbol{w}_{L_{opt}-N}^{(2)}\right\|^2_{R_{\overline{\boldsymbol{u}}_k}} \\ &= \frac{1}{J} E\left\{ \left\|\overline{\boldsymbol{\Psi}}_N^{(i-1)}\right\|^2_{R_U} \right\} + \frac{1}{J}\left\|\mathcal{W}_{L_{opt}-N}^{(O)}\right\|^2_{R_{\overline{U}}} \end{aligned} \tag{10b}$$

where $R_U = E\{\boldsymbol{U}_i^*\boldsymbol{U}_i\} = \mathrm{diag}\{\mathrm{R}_{\boldsymbol{u}_1}, \ldots, \mathrm{R}_{\boldsymbol{u}_J}\}$ and $R_{\overline{U}} = E\left\{\overline{\boldsymbol{U}}_i^*\overline{\boldsymbol{U}}_i\right\} = \mathrm{diag}\left\{R_{\overline{\boldsymbol{u}}_1}, \ldots, R_{\overline{\boldsymbol{u}}_J}\right\}$.

What all these measures have in common is the weighted norm $\left\|\overline{\boldsymbol{\Psi}}_N^{(i-1)}\right\|_{\Sigma}^2$. On this basis, evaluating the weighted norm on both sides of (8) and then taking expectations yields the following relation:

$$\begin{aligned} E\left\{\left\|\overline{\boldsymbol{\Psi}}_N^{(i)}\right\|_{\Sigma}^2\right\} &= E\left\{\left\|\overline{\boldsymbol{\Psi}}_N^{(i-1)}\right\|_{\tilde{\Sigma}}^2\right\} \\ &+ \mathcal{W}_{L_{opt}-N}^{*(O)} E\left\{\overline{\boldsymbol{U}}_i^* \boldsymbol{U}_i D \sum D \boldsymbol{U}_i^* \overline{\boldsymbol{U}}_i\right\} \mathcal{W}_{L_{opt}-N}^{(O)} \\ &+ E\left\{\boldsymbol{V}_i^* \boldsymbol{U}_i D \sum D \boldsymbol{U}_i^* \boldsymbol{V}_i\right\} \end{aligned} \tag{11}$$

where

$$\begin{aligned} \widetilde{\sum} &= G^* \sum G - G^* \sum D E\{\boldsymbol{U}_i^* \boldsymbol{U}_i\} G \\ -G^* E\{\boldsymbol{U}_i^* \boldsymbol{U}_i\} D \sum G &+ G^* E\left\{\boldsymbol{U}_i^* \boldsymbol{U}_i D \sum D \boldsymbol{U}_i^* \boldsymbol{U}_i\right\} G \end{aligned} \tag{12}$$

The term $\mathcal{W}_{L_{opt}-N}^{*(O)} E\left\{\overline{\boldsymbol{U}}_i^* \boldsymbol{U}_i D \sum D \boldsymbol{U}_i^* \overline{\boldsymbol{U}}_i\right\} \mathcal{W}_{L_{opt}-N}^{(O)}$ on the right-hand side of (11) is due to the deficient length adaptive filter application at each node. This term, which contains all the omitted entries of the unknown vector, does not appear in the sufficient length scenario.

To proceed, we calculate the moments in the relations (11) and (12). First, $E\{\boldsymbol{V}_i^* \boldsymbol{U}_i D \sum D \boldsymbol{U}_i^* \boldsymbol{V}_i\}$ could be evaluated as [25]:

$$\begin{aligned} E\left\{\boldsymbol{V}_i^* \boldsymbol{U}_i D \sum D \boldsymbol{U}_i^* \boldsymbol{V}_i\right\} &= \mathrm{tr}\left\{D\, E\{\boldsymbol{U}_i^* \boldsymbol{V}_i \boldsymbol{V}_i^* \boldsymbol{U}_i\} D \sum\right\} \\ &= \mathrm{tr}\left\{D\, E\{\boldsymbol{U}_i^* R_{\boldsymbol{V}} \boldsymbol{U}_i\} D \sum\right\} \\ &= \mathrm{tr}\left\{D\left(R_{\boldsymbol{V}} \bigotimes I_N\right) E\{\boldsymbol{U}_i^* \boldsymbol{U}_i\} D \sum\right\} \\ &= \mathrm{tr}\left\{D^2\left(R_{\boldsymbol{V}} \bigotimes I_N\right) E\{\boldsymbol{U}_i^* \boldsymbol{U}_i\} \sum\right\} \\ &= \mathrm{tr}\left\{D^2\left(R_{\boldsymbol{V}} \bigotimes I_N\right) R_{\boldsymbol{U}} \sum\right\} \\ &= \left(\mathrm{bvec}\left\{D^2\left(R_{\boldsymbol{V}} \bigotimes I_N\right) R_{\boldsymbol{U}}\right\}\right)^T \mathrm{bvec}\{\sum\} \end{aligned} \tag{13}$$

where $R_{\boldsymbol{V}} = E\{\boldsymbol{V}_i \boldsymbol{V}_i^*\} = \mathrm{diag}\left\{\sigma_{v,1}^2, \ldots, \sigma_{v,J}^2\right\}$

To evaluate $\mathcal{W}_{L_{opt}-N}^{*(O)} E\left\{\overline{\boldsymbol{U}}_i^* \boldsymbol{U}_i D \sum D \boldsymbol{U}_i^* \overline{\boldsymbol{U}}_i\right\} \mathcal{W}_{L_{opt}-N}^{(O)}$, we start by assessing $\Upsilon = E\left\{\overline{\boldsymbol{U}}_i^* \boldsymbol{U}_i \sum \boldsymbol{U}_i^* \overline{\boldsymbol{U}}_i\right\}$ as:

$k \neq \ell$;

$$\Upsilon_{k\ell} = E\left\{\overline{\boldsymbol{u}}_{k,i}^* \boldsymbol{u}_{k,i} \sum\nolimits_{k\ell} \boldsymbol{u}_{\ell,i}^* \overline{\boldsymbol{u}}_{\ell,i}\right\} = E\left\{\overline{\boldsymbol{u}}_{k,i}^* \boldsymbol{u}_{k,i}\right\} \sum\nolimits_{k\ell} E\left\{\boldsymbol{u}_{\ell,i}^* \overline{\boldsymbol{u}}_{\ell,i}\right\} = 0 \tag{14}$$

$$
\begin{aligned}
k = \ell; Y_{kk} &= E\left\{\overline{\boldsymbol{u}}_{k,i}^{*}\boldsymbol{u}_{k,i}\sum\nolimits_{kk}\boldsymbol{u}_{k,i}^{*}\overline{\boldsymbol{u}}_{k,i}\right\} = E\left\{\overline{\boldsymbol{u}}_{k,i}^{*}\mathrm{tr}\left\{\boldsymbol{u}_{k,i}\sum\nolimits_{kk}\boldsymbol{u}_{k,i}^{*}\right\}\overline{\boldsymbol{u}}_{k,i}\right\} \\
&= E\left\{\overline{\boldsymbol{u}}_{k,i}^{*}\overline{\boldsymbol{u}}_{k,i}\right\}\mathrm{tr}\left\{E\left\{\boldsymbol{u}_{k,i}^{*}\boldsymbol{u}_{k,i}\right\}\sum\nolimits_{kk}\right\} = R_{\overline{\boldsymbol{u}}_k}\mathrm{tr}\left\{R_{\boldsymbol{u}_k}\sum\nolimits_{kk}\right\}
\end{aligned}
\tag{15}
$$

where $\Upsilon_{k\ell}$ is the $k\ell$-block of Υ. Therefore, $\mathcal{W}_{L_{opt}-N}^{*(O)}E\left\{\overline{\boldsymbol{U}}_i^{*}\boldsymbol{U}_iD\sum D\boldsymbol{U}_i^{*}\overline{\boldsymbol{U}}_i\right\}\mathcal{W}_{L_{opt}-N}^{(O)}$ could be calculated as:

$$
\begin{aligned}
&\mathcal{W}_{L_{opt}-N}^{*(O)}E\left\{\overline{\boldsymbol{U}}_i^{*}\boldsymbol{U}_iD\sum D\boldsymbol{U}_i^{*}\overline{\boldsymbol{U}}_i\right\}\mathcal{W}_{L_{opt}-N}^{(O)} = \mathcal{W}_{L_{opt}-N}^{*(O)}\overline{D}\Upsilon\overline{D}\mathcal{W}_{L_{opt}-N}^{(O)} \\
&= \mathrm{tr}\left\{\mathcal{W}_{L_{opt}-N}^{*(O)}\overline{D}\Upsilon\overline{D}\mathcal{W}_{L_{opt}-N}^{(O)}\right\} \\
&= \mathrm{tr}\left\{\mathcal{W}_{L_{opt}-N}^{(O)}\mathcal{W}_{L_{opt}-N}^{*(O)}\overline{D}\Upsilon\overline{D}\right\} \\
&= \left(\mathrm{bvec}\left\{\left(\mathcal{W}_{L_{opt}-N}^{(O)}\mathcal{W}_{L_{opt}-N}^{*(O)}\right)^{T}\right\}\right)^{T}\mathrm{bvec}\left\{\overline{D}\Upsilon\overline{D}\right\} \\
&= \left(\mathrm{bvec}\left\{\left(\mathcal{W}_{L_{opt}-N}^{(O)}\mathcal{W}_{L_{opt}-N}^{*(O)}\right)^{T}\right\}\right)^{T}\left(\overline{D}\odot\overline{D}\right)\mathrm{bvec}\{\Upsilon\}
\end{aligned}
\tag{16}
$$

We could express $\mathrm{bvec}\{\Upsilon\}$ as $bvec\{\Upsilon\} = \mathrm{col}\left\{\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_J\right\}$, where:

$$
\boldsymbol{\gamma}_k = \mathrm{col}\left\{\mathbf{0}_{(L_{opt}-N)^2(k-1)\times 1}, \overline{\boldsymbol{r}}_k\boldsymbol{r}_k^T\sigma_{kk}, \mathbf{0}_{(L_{opt}-N)^2(J-k)\times 1}\right\}
\tag{17}
$$

where $\sigma_{kk} = \mathrm{vec}\left\{\sum_{kk}\right\}$, $\boldsymbol{r}_k = \mathrm{vec}\left\{R_{\boldsymbol{u}_k}\right\}$, and $\overline{\boldsymbol{r}}_k = \mathrm{vec}\{R_{\overline{\boldsymbol{u}}_k}\}$. It follows that

$$
\begin{aligned}
\boldsymbol{\gamma}_k &= \Gamma_k\sigma_k \\
&= \mathrm{diag}\left\{\mathbf{0}_{(L_{opt}-N)^2(k-1)\times N^2(k-1)}, \overline{\boldsymbol{r}}_k\boldsymbol{r}_k^T, \mathbf{0}_{(L_{opt}-N)^2(J-k)\times N^2(J-k)}\right\}\sigma_k
\end{aligned}
\tag{18}
$$

where $\sigma_k = \mathrm{col}\left\{\sigma_{1k},\ \sigma_{2k}, \ldots, \sigma_{Jk}\right\}$. So, $\mathrm{bvec}\{\Upsilon\}$ could be rewritten as:

$$
b\mathrm{vec}\{\Upsilon\} = \mathrm{diag}\left\{\Gamma_1, \Gamma_2, \ldots, \Gamma_J\right\}\mathrm{bvec}\left\{\sum\right\} = \mathcal{F}\mathrm{bvec}\left\{\sum\right\}
\tag{19}
$$

Consequently, we find that:

$$
\begin{aligned}
&\mathcal{W}_{L_{opt}-N}^{*(O)}E\left\{\overline{\boldsymbol{U}}_i^{*}\boldsymbol{U}_iD\sum D\boldsymbol{U}_i^{*}\overline{\boldsymbol{U}}_i\right\}\mathcal{W}_{L_{opt}-N}^{(O)} \\
&= \left(\mathrm{bvec}\left\{\left(\mathcal{W}_{L_{opt}-N}^{(O)}\mathcal{W}_{L_{opt}-N}^{*(O)}\right)^{T}\right\}\right)^{T}\left(\overline{D}\odot\overline{D}\right)\mathcal{F}\ \mathrm{bvec}\left\{\sum\right\}
\end{aligned}
\tag{20}
$$

Also, applying the block vectorization operator bvec to $\widetilde{\sum}$ in (12) will result in [25]:

$$\begin{aligned} \mathrm{bvec}\{\widetilde{\textstyle\sum}\} = & \left(G^T \odot G^*\right)\mathrm{bvec}\{\textstyle\sum\} \\ & - \left(G^T \odot G^*\right)\left(DR_U \odot I_{JN}\right)\mathrm{bvec}\{\textstyle\sum\} \\ & - \left(G^T \odot G^*\right)\left(I_{JN} \odot R_U D\right)\mathrm{bvec}\{\textstyle\sum\} \\ & + \left(G^T \odot G^*\right)\left(D \odot D\right)\mathcal{B}\,\mathrm{bvec}\{\textstyle\sum\} \end{aligned} \tag{21}$$

where $\mathcal{B} = \mathrm{diag}\left\{\mathfrak{B}_1, \mathfrak{B}_2, \dots, \mathfrak{B}_J\right\}$, and

$$\mathfrak{B}_k = \mathrm{diag}\left\{\mathfrak{b}_k^{(1)}, \dots, \mathfrak{b}_k^{(J)}\right\} \tag{22}$$

where $\mathfrak{b}_k^{(\ell)}$ is given by

$$\mathfrak{b}_k^{(\ell)} = \begin{cases} \boldsymbol{r}_k \boldsymbol{r}_k^T + \beta R_{\boldsymbol{u}_k} \otimes R_{\boldsymbol{u}_k} & \ell = k \\ R_{\boldsymbol{u}_\ell} \otimes R_{\boldsymbol{u}_k} & \ell \neq k \end{cases} \tag{23}$$

where $\beta = 2$ for real regressors and $\beta = 1$ for complex data.

Substituting the resulting moments in (11) and (12) leads to the following relation for deficient length diffusion LMS adaptive networks.

$$E\left\{\left\|\overline{\boldsymbol{\Psi}}_N^{(i)}\right\|^2_{\mathrm{bvec}\{\sum\}}\right\} = E\left\{\left\|\overline{\boldsymbol{\Psi}}_N^{(i-1)}\right\|^2_{\mathcal{S}\ \mathrm{bvec}\{\sum\}}\right\} + \mathcal{Z}\,\mathrm{bvec}\{\textstyle\sum\} \tag{24}$$

where

$$\begin{aligned} \mathcal{S} = & \left(G^T \odot G^*\right) \\ & \times \left[I - \left(DR_U \odot I_{JN}\right) - \left(I_{JN} \odot R_U D\right) + \left(D \odot D\right)\mathcal{B}\right] \end{aligned} \tag{25}$$

and

$$\begin{aligned} \mathcal{Z} = & \left(\mathrm{bvec}\left\{\left(\mathcal{W}^{(O)}_{L_{opt}-N}\mathcal{W}^{*(O)}_{L_{opt}-N}\right)^T\right\}\right)^T \left(\overline{D} \odot \overline{D}\right)_F \\ & + \left(\mathrm{bvec}\left\{D^2\left(R_V \otimes I_N\right)\ R_U\right\}\right)^T \end{aligned} \tag{26}$$

In steady-state, (24) yields to

$$E\left\{\left\|\overline{\boldsymbol{\Psi}}_N^{(\infty)}\right\|^2_{(I-\mathcal{S})\ \mathrm{bvec}\{\sum\}}\right\} = \mathcal{Z}\,\mathrm{bvec}\{\textstyle\sum\} \tag{27}$$

Therefore, the network and local MSD and EMSE could be evaluated as:

$$\mathrm{MSD}^{Net} = \frac{1}{J}\mathcal{Z}(I-\mathcal{S})^{-1}\mathrm{bvec}\{I_{JN}\} + \left\|\boldsymbol{w}^{(2)}_{L_{opt}-N}\right\|^2 \tag{28}$$

$$\mathrm{EMSE}^{Net} = \frac{1}{J}\mathcal{Z}(I-\mathcal{S})^{-1}\mathrm{bvec}\{R_U\} + \frac{1}{J}\left\|\mathcal{W}^{(O)}_{L_{opt}-N}\right\|^2_{R_{\overline{U}}} \tag{29}$$

$$\mathrm{MSD}_k = \mathcal{Z}(I-\mathcal{S})^{-1}\mathrm{bvec}\left\{\mathrm{diag}\left\{\mathbf{0}_{(k-1)N}, I_N, \mathbf{0}_{(J-k)N}\right\}\right\} + \left\|\boldsymbol{w}^{(2)}_{L_{opt}-N}\right\|^2 \tag{30}$$

$$\begin{aligned}\mathrm{EMSE}_k &= \mathcal{Z}(I-\mathcal{S})^{-1}\mathrm{bvec}\left\{\mathrm{diag}\left\{\mathbf{0}_{(k-1)N}, R_{\boldsymbol{u}_k}, \mathbf{0}_{(J-k)N}\right\}\right\} \\ &+ \left\|\mathcal{W}^{(O)}_{L_{opt}-N}\right\|^2_{\mathrm{diag}\left\{\mathbf{0}_{(k-1)\left(L_{opt}-N\right)}, R_{\overline{\boldsymbol{u}}_k}, \mathbf{0}_{(J-k)\left(L_{opt}-N\right)}\right\}}\end{aligned} \tag{31}$$

4 Simulation results

This section compares the theoretical derivations with the computer experiments. All simulations are performed in MATLAB© software and examined in the HP laptop with Windows 10 Pro 64-bit, with Processor Intel®-core™ i5-3340M CPU @ 2.70GHz and 8GB of RAM. All expectations are resulted from averaging over 100 independent runs. We consider a network with $J = 10$ sensors where each sensor runs a local filter with N taps to estimate the unknown vector $w^o_{L_{opt}} = col\{1, 1, ..., 1\}/\sqrt{L_{opt}}$ with length $L_{opt} = 10$. The zero-mean Gaussian noise and regressors are independent in space, and i.i.d. in time. Their statistics profile and the network topology are illustrated in Fig. 1. The step size at each sensor is set to $\mu_k = 0.01$. Also, for the local combiners $\{c_{k,\ell} \geqslant 0\}$, the Metropolis rule is utilized as [25]:

$$c_{k,\ell} = \begin{cases} 1/\max(j_k, j_\ell) & \text{if } k \neq \ell \text{ are linked} \\ 0 & \text{if } k \text{ and } \ell \text{ are not linked} \\ 1 - \sum\limits_{m \in \mathcal{J}_k \setminus \{k\}} c_{k,m} & \text{if } k = \ell \end{cases} \tag{32}$$

where j_k Indicates the degree of node k, i.e., $j_k = |\mathcal{J}_k|$.

Figure 2 shows the steady-state MSD per node, where every node runs an adaptive filter of lengths $N = 5$, $N = 8$ (to model the deficient length scenario), and $N = 10$ (which implies the sufficient length scenario). As can be seen from this figure, there is a good match between simulation and theoretical results. Also, this figure confirms that by decreasing the tap-length from its actual value, the steady-state MSD deteriorates drastically, such that, for only two units of length reduction ($N = 8$), the steady-state MSD in node #1 decreases from −51.49 to −6.98 dB.

A similar achievement is observed for the EMSE evaluation criterion, illustrated in Fig. 3. This figure confirms the close match between simulation and theoretical results of the EMSE measure. According to Fig. 3, similar to MSD, by decreasing the tap length from its actual value, the performance from the steady-state EMSE point of view degrades considerably. For example, for node #7, as indicated in

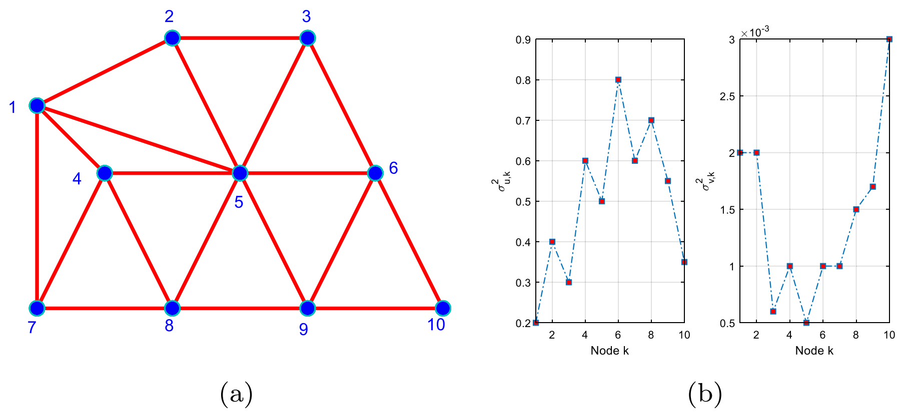


Fig. 1 Network topology **a**, and statistical settings **b** for $J = 10$ sensors

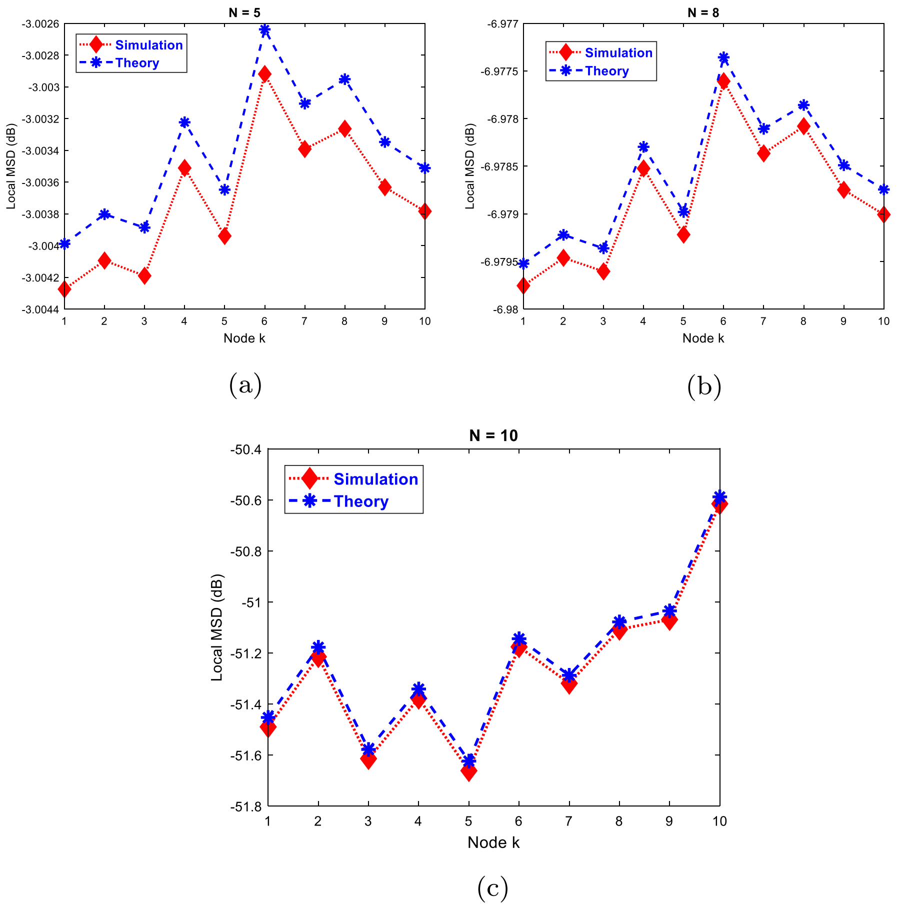


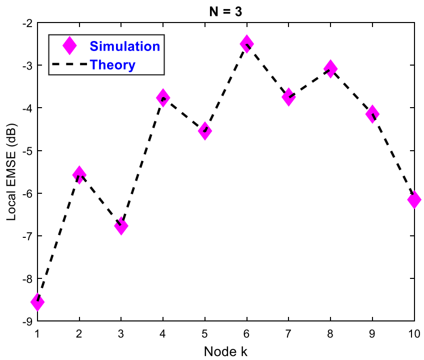
Fig. 2 The steady-state local MSD vs. node for $N = 5$ **a**, $N = 8$ **b**, and $N = 10$ **c**

Fig. 3, for tap lengths $N = 3$, $N = 7$, and $N = 10$, the steady-state EMSE is −3.75, −7.54, and −53.55 dB, respectively.

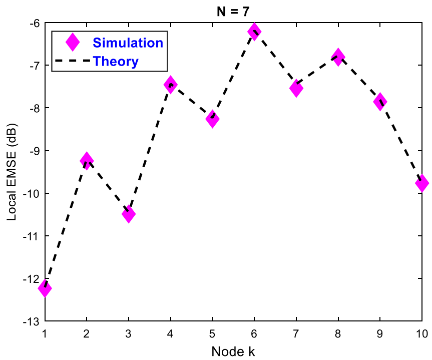
It is evident from Figs. 2 and 3 that the steady-state MSD and EMSE are sensitive to the sensor statistics in both sufficient and deficient length scenarios.

The dependency of the steady-state MSD and EMSE on the filter tap-length is seen clearly in Fig. 4. According to this figure, the steady-state measures are improved as the tap length tends to its actual value. However, this improvement occurs slowly for $N < L_{opt}$, changing the length from 9 to 10 improves the value of the steady-state measures by about 41 dB.

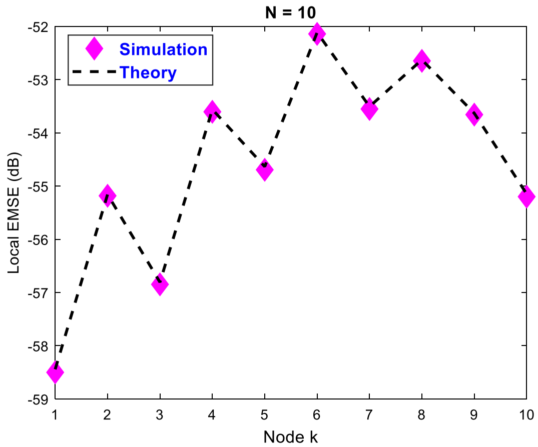
Figures 5 and 6 show another interesting finding concerning deficient length diffusion LMS adaptive networks: in the case of deficient length, the steady-state MSD and EMSE curves continue to grow at a slower rate as μ increases. These figures also reveal the acceptable matching between simulation and theoretical findings. Also,



(a)



(b)



(c)

Fig. 3 The steady-state local EMSE vs. node for $N = 3$ **a**, $N = 7$ **b**, and $N = 10$ **c**

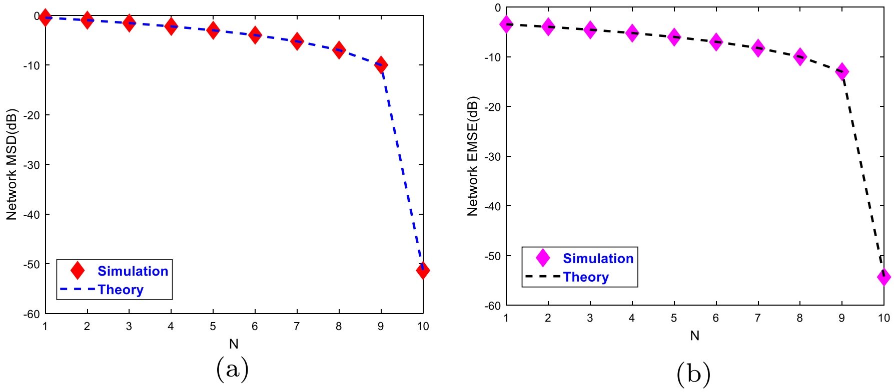


Fig. 4 Network MSD **a**, and Network EMSE **b** per N

we can see from these figures that, by tending the tap length to its optimum value, the performance increases as expected. As can be seen from Fig. 5, by decreasing the step size from 0.01 to 0.001, the network MSD is improved by about 10.483 dB in the case of sufficient length, but this improvement is about 0.006 dB in the case of deficient length $N = 6$.

5 Concluding remarks

In this paper, we studied the steady-state behavior of the deficient length diffusion LMS algorithm. We provided a closed-form expression for the EMSE and MSD of each node to clarify the efficiency of length deficiency on the steady-state performance of each sensor. The results verified the dependency of the steady-state MSD and EMSE on the filter tap length. It was concluded that the performance of

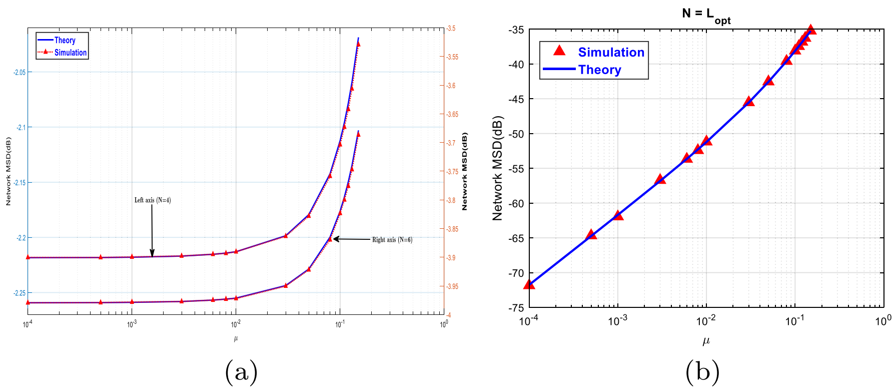
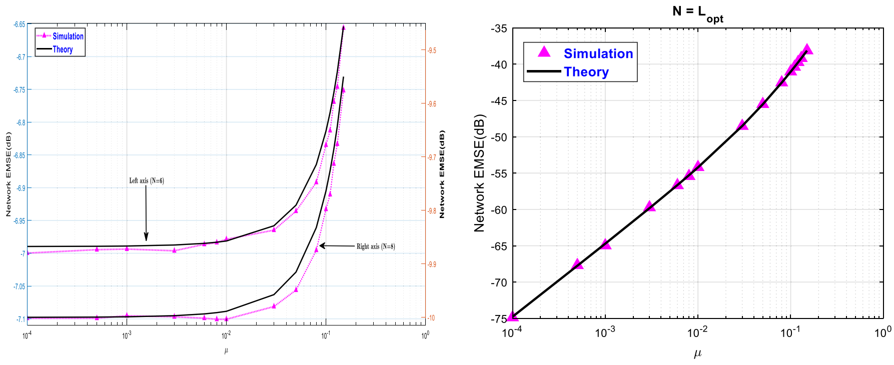


Fig. 5 Network MSD vs. μ for deficient length **a**and full length **b** scenarios



(a) (b)

Fig. 6 Network EMSE vs. μ for deficient length **a** and full length **b** scenarios

diffusion adaptive networks is considerably affected in the deficient length scenario, or equivalently, the steady-state measures are improved as the tap length tends to its actual value. Unlike the full-length case, where the steady-state MSD and EMSE decrease significantly with the step size reduction, this research showed that in the deficient-length scenario, there are no considerable improvements in the steady-state performance by reducing the step size, or equivalently, in the case of deficient length, the steady-state MSD and EMSE curves continue to grow at a slower rate as step-size increases. Therefore, it can be concluded that the tap length plays a critical role in diffusion adaptive networks since the performance deterioration due to the deficient selection of tap length could not be compensated by an adjustment in the step size.

References

1. Azarnia G, Tinati MA, Sharifi AA, Shiri H (2020) Incremental and diffusion compressive sensing strategies over distributed networks. Digital Signal Process 101:102732
2. Cattivelli FS, Sayed AH (2010) Diffusion LMS strategies for distributed estimation. IEEE Trans Signal Process 58(3):1035–1048
3. Gao W, Chen J, Richard C (2021) Transient theoretical analysis of diffusion RLS algorithm for cyclostationary colored inputs. IEEE Signal Process Lett 28:1160–1164
4. Zayyani H (2021) Robust minimum disturbance diffusion LMS for distributed estimation. IEEE Trans Circuits Syst II Express Briefs 68(1):521–525
5. Khalili A, Vahidpour V, Rastegarnia A, Bazzi WM, Sanei S (2021) Partial diffusion kalman filter with adaptive combiners. IEEE Trans Aerosp Electron Syst 57(3):1972–1980
6. Song P, Zhao H, Zhu Y (2021) Robust multitask diffusion affine projection algorithm for distributed estimation. IEEE Trans Circuits Syst II Express Briefs 69(3):1892–6
7. Abadi MSE, Ahmadi MJ (2018) Diffusion improved multiband-structured subband adaptive filter algorithm with dynamic selection of nodes over distributed networks. IEEE Trans Circuits Syst II Express Briefs 66(3):507–511
8. Abadi MSE, Adabi AP (2019) Convex combination of two diffusion LMS for distributed estimation. In: 2019 27th Iranian conference on electrical engineering (ICEE), , IEEE, pp 1715–1719
9. Jin D, Chen J, Richard C, Chen J, Sayed AH (2020) Affine combination of diffusion strategies over networks. IEEE Trans Signal Process 68:2087–2104

10. Yu Y, Zhao H, de Lamare RC, Zakharov Y, Lu L (2019) Robust distributed diffusion recursive least squares algorithms with side information for adaptive networks. IEEE Trans Signal Process 67(6):1566–1581
11. Song P, Zhao H, Zeng X (2019) Robust diffusion affine projection algorithm with variable step-size over distributed networks. IEEE Access 7:150484–150491
12. Hua F, Nassif R, Richard C, Wang H, Sayed AH (2020) Diffusion LMS with communication delays: stability and performance analysis. IEEE Signal Process Lett 27:730–734
13. Gao W, Chen J (2019) Performance analysis of diffusion LMS for cyclostationary white non-gaussian inputs. IEEE Access 7:91243–91252
14. Zhang M, Jin D, Chen J, Ni J (2021) Zeroth-order diffusion adaptive filter over networks. IEEE Trans Signal Process 69:589–602
15. Khalili A, Vahidpour V, Rastegarnia A, Bazzi WM, Sanei S (2020) Partial diffusion kalman filter with adaptive combiners. IEEE Trans Aerosp Electron Syst 57(3):1972–1980
16. Ijeoma CA, Hossin MA, Bemnet HA, Tesfaye AA, Hailu AH, Chiamaka CN (2020) "Exploration of diffusion LMS over static and adaptive combination policy," In: 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), pp. 424–427, IEEE
17. Merched R, Vlaski S, Sayed AH (2019) Enhanced diffusion learning over networks. In: 2019 27th European signal processing conference (EUSIPCO), IEEE, pp 1–5
18. Erginbas YE, Vlaski S, Sayed AH (2021) Gramian-based adaptive combination policies for diffusion learning over networks. In: ICASSP 2021-2021 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, pp 5215–5219
19. Li L, Feng J, He J, Chambers JA (2013) A distributed variable tap-length algorithm within diffusion adaptive networks. AASRI Procedia 5:77–84
20. Zhang Y, Wang C, Zhao L, Chambers JA (2015) A spatial diffusion strategy for tap-length estimation over adaptive networks. IEEE Trans Signal Process 63(17):4487–4501
21. Azarnia G (2021) Diffusion fractional tap-length algorithm with adaptive error width and step-size. Circuits, Syst Signal Process 41(1):321–45
22. Li L, Zhang Y, Chambers JA (2008) Variable length adaptive filtering within incremental learning algorithms for distributed networks. In: 2008 42nd Asilomar conference on signals, systems and computers, IEEE, pp 225–229
23. Azarnia G, Hassanlou M (2021) A variable tap-length dilms algorithm with variable parameters for wireless sensor networks. Int J Sensor Netw 36(2):97–107
24. Azarnia G, Tinati MA (2015) Steady-state analysis of the deficient length incremental LMS adaptive networks. Circuits Syst Signal Process 34(9):2893–2910
25. Lopes CG, Sayed AH (2008) Diffusion least-mean squares over adaptive networks: formulation and performance analysis. IEEE Trans Signal Process 56(7):3122–3136