



Steady-state analysis of distributed incremental variable fractional tap-length LMS adaptive networks

Ghanbar Azarnia¹ · Abbas Ali Sharif²

Accepted: 10 August 2021 / Published online: 21 August 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Recently, cooperative and distributed processing has been attracted a lot of attention, especially in wireless sensor networks, to prolong the network's lifetime. So, distributed adaptive filtering, which operates in a distributed and adaptive manner, has been established. In the distributed adaptive networks, in addition to filter coefficients, the length of the adaptive filter is also unknown in general, and it should be estimated. The distributed incremental fractional tap-length (FT) algorithm is an approach to determine the adaptive filter length in a distributed scheme. In the current study, we analyze the steady-state behavior of the distributed incremental variable FT LMS algorithm. According to the analysis, we derive the mathematical expression for the steady-state tap-length at each particular sensor. The obtained results indicate that this algorithm overestimates optimal tap-length. Numerical simulations are provided to confirm the theoretical analyses.

Keywords Adaptive networks · DILMS algorithm · Distributed estimation · Fractional tap-length (FT)

1 Introduction

Wireless sensor network (WSN) has been adopted as one of the most potential applications in many research fields, such as environmental monitoring, factory instrumentation, power transmission and distribution systems, and military surveillance [1]. In many applications, the objective is to estimate a desired unknown parameter employing measurements collected from the sensor nodes and the distributed solution is usually used to estimate the parameter. The distributed estimation extracts the information collected from the sensors distributed in an area. The distribution of the sensors, in addition to the temporal dimension, provides spatial diversity to improve the robustness of the processing tasks [2]. In the distributed solution, each sensor node relies on its local measurements, communicates with its immediate neighbors, and the processing is distributed between all sensors. In this solution,

the processing and communications measurements are significantly decreased.

Recently, distributed adaptive estimation approaches (or adaptive networks) have been proposed to solve the linear estimation problem in a distributed and adaptive manner. The distributed adaptive approach extends the adaptive filters to the network domain, which is performed without prior knowledge of data statistics. The adaptive networks' performance strongly depends on the cooperation modes of incremental and diffusion [3]. In the incremental scheme, a cyclic path is needed, and the sensors communicate with their neighbors inside the path. The incremental LMS algorithm [4], incremental RLS algorithm [5], incremental affine projection-based adaptive (APA) algorithm [6] are distributed adaptive estimations that use incremental cooperation. When more communications and energy resources are accessible, a cooperative diffusion mode is utilized. In this mode, each node communicates with all its neighbors as managed by the network topology, and it is not required any cyclic path. The diffusion LMS [7], diffusion RLS algorithm [8], and diffusion APA [9] are distributed adaptive estimation algorithms that use diffusion cooperation.

In most of the previous studies of distributed adaptive estimation algorithms [4–9], the tap-length of the adaptive

✉ Ghanbar Azarnia
g.azarnia@uut.ac.ir

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

² Department of Electrical Engineering, University of Bonab, Bonab, Iran

filters is constant for each node. This issue is not proper in general because the optimal value of tap-length is variable or unknown. The tap-length of a linear filter significantly influences its performance. The deficient tap-length likely increases the mean-squared error (MSE), whereas the excess mean-squared error (EMSE) and computational cost will be high when the tap-length is too large. Hence, there is an optimal value for tap-length to balance the conflicting requirements of performance and complexity. Furthermore, in many applications, the optimal tap-length may be time-variant. So, a variable tap-length algorithm is required to determine the optimal length. In the single filter domain, many algorithms have been introduced for this purpose [10–21]. The most efficient algorithm among them is the FT algorithm [21]. This algorithm, like the LMS and its variants [22–24] is simple and has an excellent performance. Thus, it is considered a popular algorithm, and in the adaptive networks domain, it is more suitable than all variable tap-length algorithms. In [25], the FT algorithm has developed in the context of incremental learning for distributed networks. Two recursive algorithms run concurrently, one for estimating the tap-length and the other for estimating the tap-weights. The authors in [25] only provide a steady-state analysis to examine the steady-state performance of the tap-weight vector. This issue is performed by deriving the theoretical expressions for the mean-squared deviation (MSD), MSE, and EMSE for each node within the network. For this purpose, the authors assume a fixed tap-length at the steady-state scenario, but they have not provided any theoretical expression for this steady-state tap-length. In this paper, we investigate a theoretical expression for steady-state tap-length achieved in the incremental structure tap-length adaption for each node. The achieved results indicate that this algorithm overestimates the optimal tap-length. Simulation results validate the derived theoretical expressions.

The remains of the paper are organized as follows. In Sect. 2, we provide the background that includes the estimation problem and its incremental-based solution. In part 3, we analyze the performance of distributed incremental variable FT LMS algorithm. In this part, we also calculate the mathematical expression for the steady-state tap-length at each sensor. The comparison of the simulation results and theoretical analyses is given in Sect. 4, while conclusion remarks are presented in Sect. 5.

Notation: For ease of reference, the main symbols used in this paper are listed in Table 1.

2 Incremental FT-LMS algorithm

We consider N sensor nodes spatially distributed in a region. The objective is to estimate the coefficients and the length L_{opt} of an unknown vector $\mathbf{w}_{L_{opt}}^o$ from multiple measurements obtained from N sensors of the network. The adaption rules for the coefficients and the length of $\mathbf{w}_{L_{opt}}^o$ are decoupled, where the selection of each node does not depend on the other. As we will present later, the tap-length L is calculated using the tap-length estimation algorithm. We also suppose that each node k has access to time realizations $\{d_k(i), \mathbf{u}_{k,i}\}$ of zero-mean spatial data $\{d_k, \mathbf{u}_k\}$, where d_k and \mathbf{u}_k are respectively the scalar and $1 \times L$ row regression vector.

By collecting the measurement data and regression into global matrices, the following definitions are made:

$$\begin{aligned} \mathbf{d} &\triangleq \text{col}\{d_1, d_2, \dots, d_N\} & (N \times 1) \\ \mathbf{U} &\triangleq \text{col}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N\} & (N \times L) \end{aligned} \quad (1)$$

First, we consider the unknown parameter coefficients estimation. The aim is the estimation of $L \times 1$ vector \mathbf{w} that solves $\arg \min_{\mathbf{w}} J_L(\mathbf{w})$, where

$$J_L(\mathbf{w}) = E\{\|\mathbf{d} - \mathbf{U}\mathbf{w}\|^2\} \quad (2)$$

Let $\psi_k^{(i)}$ be the local estimation of the unknown parameter at node k in time instant i . In [4], a distributed incremental LMS (DILMS) approach is reported to solve the above optimization problem as follows:

$$\begin{aligned} \psi_k^{(i)} &= \psi_{k-1}^{(i)} + \mu_k \mathbf{u}_{k,i}^* (d_k(i) - \mathbf{u}_{k,i} \psi_{k-1}^{(i)}) \\ k &\in \mathcal{N} \end{aligned} \quad (3)$$

where the parameter μ_k is the proper selected positive local step-size. Each sensor k employs the local data $d_k(i)$, $\mathbf{u}_{k,i}$ and $\psi_{k-1}^{(i)}$ received from the sensor $k - 1$. Then, $\psi_N^{(i)}$ is used as the initial condition for the next time at sensor node $k = 1$.

Considering the tap-length L , the segmented cost function is defined as [25]:

$$J_M^{(L)}(\mathbf{w}) \triangleq E\{\|\mathbf{d} - \mathbf{U}_M \mathbf{w}_M\|^2\} \quad (4)$$

where $1 \leq M \leq L$, \mathbf{U}_M and \mathbf{w}_M respectively contain the initial M column vectors of \mathbf{U} and M elements of \mathbf{w} such as:

$$\begin{aligned} \mathbf{w}_M &\triangleq \text{col}\{w(1), w(2), \dots, w(M)\} & (1 \times M) \\ \mathbf{U}_M &\triangleq \text{col}\{\mathbf{u}_1(1:M), \mathbf{u}_2(1:M), \dots, \mathbf{u}_N(1:M)\} & (N \times M) \end{aligned} \quad (5)$$

where $w(j)$ is the j 'th element of \mathbf{w} and $\mathbf{u}_k(1:M)$ consists of the M initial elements of \mathbf{u}_k . In [25], to explore the optimal length L_{opt} , the smallest difference of MSE estimation is proposed as:

Table 1 The main symbols

Symbol	Description
$ \cdot $	Absolute operation
$\ \cdot\ ^2$	Squared Euclidean norm operation
$E\{\cdot\}$	Statistical expectation
$col\{\cdot\}$	Column vector
$\mathbf{0}_{M \times 1}$	$M \times 1$ zero vector
$\mathbf{0}_{1 \times M}$	$1 \times M$ zero vector
\mathbf{I}_L	$L \times L$ identity matrix
$tr(\mathbf{A})$	Trace of matrix \mathbf{A}
$(\cdot)^T$	Transposition
$(\cdot)^*$	Conjugation for scalars and Hermitian transpose for matrices

$$\min\{L \mid J_{L-\Delta}^{(L)}(\mathbf{w}) - J_L^{(L)}(\mathbf{w}) \leq \varepsilon\} \tag{6}$$

where ε is a small positive value and it is predetermined by the system requirements. The integer value Δ prevents the tap-length to be suboptimal.

In [25], to solve the optimization problem (6) a solution by developing the FT algorithm within the context of incremental cooperation has been proposed. With the assumption that $\ell_{k,f}(i)$ denotes the local estimation of the FT at sensor k in time instant i , in the specific topological circle, sensor k receives the calculated FT $\ell_{k-1,f}(i)$ from sensor $k - 1$ and updates its local estimation as:

$$\ell_{k,f}(i) = (\ell_{k-1,f}(i) - \alpha_k) - \gamma_k \left[(e_{k,L_k(i)}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}))^2 - (e_{k,L_k(i)-\Delta}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}))^2 \right] \tag{7}$$

where

$$e_{k,L_k(i)}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}) = d_k(i) - \mathbf{u}_{k,i} \boldsymbol{\psi}_{k-1}^{(i)} \tag{8}$$

and

$$e_{k,L_k(i)-\Delta}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}) = d_k(i) - \mathbf{u}_{k,i}(1 : L_k(i) - \Delta) \boldsymbol{\psi}_{k-1}^{(i)}(1 : L_k(i) - \Delta) \tag{9}$$

In which $\mathbf{u}_{k,i}(1 : L_k(i) - \Delta)$ and $\boldsymbol{\psi}_{k-1}^{(i)}(1 : L_k(i) - \Delta)$ consist of the initial $L_k(i) - \Delta$ elements of $\mathbf{u}_{k,i}$ and $\boldsymbol{\psi}_{k-1}^{(i)}$ respectively. In (7) α_k denotes the local leakage factor and γ_k indicates the local step-size for $\ell_{k,f}(i)$ adaption at sensor k . $\ell_{k,f}(i)$ is no longer forced to be an integer value, and the local estimation of integer tap-length $L_k(i)$ is achieved from the FT as follows:

$$L_{k+1}(i) = \begin{cases} \lfloor \ell_{k,f}(i) \rfloor & \text{if } |L_k(i) - \ell_{k,f}(i)| \geq \delta_k \\ L_k(i) & \text{o.w} \end{cases} \tag{10}$$

where $\lfloor \cdot \rfloor$ indicates the floor operation, which rounds down to the nearest integer, and δ_k is a small integer. Also the

tap-length during its evolution is conditioned to be not less than a lower floor value L_{min} , where $L_{min} > \Delta$ (usually $L_{min} = \Delta + \delta_k + 1$), i.e., when the tap-length fluctuates under L_{min} , it is set to L_{min} . The necessity of this operation is due to using $L_k(i) - \Delta$ as a tap-length in the incremental FT algorithm.

With $L_k(i)$, the tap coefficients are then updated by (3) in which the length of vectors $\mathbf{u}_{k,i}$ and $\boldsymbol{\psi}_{k-1}^{(i)}$ are adjusted by $L_k(i)$.

3 Steady-state analysis of incremental variable FT LMS networks

In this part, we intend to obtain the mathematical expression for the steady-state tap-length of incremental FT algorithm for each sensor node. First, we assume a linear model as:

$$d_k(i) = \mathbf{u}_{L_{opt},k,i} \mathbf{w}_{L_{opt}}^o + v_k(i) \tag{11}$$

where $v_k(i)$ is temporally and spatially white noise element with zero mean and variance $\sigma_{v,k}^2$ and independent of $\mathbf{u}_{L_{opt},\ell,j}$ and $d_\ell(j)$ for all ℓ, j . In (11), $\mathbf{u}_{L_{opt},k,i}$ is a vector of length L_{opt} . Since in this scenario, the length of the vectors varies in each iteration, so it is convenient to define an upper bound for the length and pad the vectors by zeros to achieve this length. So, we obtain vectors with the same length that simplifies the calculations. Thus, we define the upper bound for length L_{ub} as:

$$L_{ub} > \max\{L_{opt}, L_k(i)\} \quad \text{for all } i \tag{12}$$

By this definition, using $\mathbf{w}_{L_{ub}}$ we denote the unknown parameter $\mathbf{w}_{L_{opt}}^o$, where $\mathbf{w}_{L_{ub}}$ is obtained by padding $\mathbf{w}_{L_{opt}}^o$ with $L_{ub} - L_{opt}$ zeros. In this way, we define the vector $\boldsymbol{\psi}_{L_{ub},k-1}^{(i)}$ that is obtained by padding $\boldsymbol{\psi}_{k-1}^{(i)}$ with $L_{ub} - L_k(i)$

zeros. Now we partition the unknown parameter \mathbf{w}_{Lub} into three parts as:

$$\begin{bmatrix} \dot{\mathbf{w}} \\ \ddot{\mathbf{w}} \\ \ddot{\mathbf{w}} \end{bmatrix} \tag{13}$$

where $\dot{\mathbf{w}}$ is the part modeled by $\dot{\boldsymbol{\psi}}_{k-1}^{(i)}$, and $\ddot{\mathbf{w}}$ is the part modeled by $\ddot{\boldsymbol{\psi}}_{k-1}^{(i)}$ such that:

$$\begin{aligned} \dot{\boldsymbol{\psi}}_{k-1}^{(i)} &= \boldsymbol{\psi}_{k-1}^{(i)}(1 : L_k(i) - \Delta) \\ \ddot{\boldsymbol{\psi}}_{k-1}^{(i)} &= \boldsymbol{\psi}_{k-1}^{(i)}(L_k(i) - \Delta + 1 : L_k(i)) \end{aligned} \tag{14}$$

where $\boldsymbol{\psi}_{k-1}^{(i)}(1 : L_k(i) - \Delta)$ consists of the initial $L_k(i) - \Delta$ elements of $\boldsymbol{\psi}_{k-1}^{(i)}$ and $\boldsymbol{\psi}_{k-1}^{(i)}(L_k(i) - \Delta + 1 : L_k(i))$ consists of the last Δ elements of $\boldsymbol{\psi}_{k-1}^{(i)}$ and $\ddot{\mathbf{w}}$ is the undermodeled part. Now we define the vector $\overline{\boldsymbol{\psi}}_{Lub,k-1}^{(i)}$, that computes the difference among the estimated weight at sensor $k - 1$ and the desired solution \mathbf{w}_{Lub} as:

$$\overline{\boldsymbol{\psi}}_{Lub,k-1}^{(i)} = \mathbf{w}_{Lub} - \boldsymbol{\psi}_{Lub,k-1}^{(i)} \tag{15}$$

and it is partitioned into three parts as:

$$\begin{aligned} \overline{\boldsymbol{\psi}}_{Lub,k-1}^{(i)} &= \mathbf{w}_{Lub} - \boldsymbol{\psi}_{Lub,k-1}^{(i)} = \begin{bmatrix} \dot{\mathbf{w}} \\ \ddot{\mathbf{w}} \\ \ddot{\mathbf{w}} \end{bmatrix} - \begin{bmatrix} \dot{\boldsymbol{\psi}}_{k-1}^{(i)} \\ \ddot{\boldsymbol{\psi}}_{k-1}^{(i)} \\ \mathbf{0}_{(Lub-L_k(i)) \times 1} \end{bmatrix} \\ &= \begin{bmatrix} \dot{\mathbf{w}} - \dot{\boldsymbol{\psi}}_{k-1}^{(i)} \\ \ddot{\mathbf{w}} - \ddot{\boldsymbol{\psi}}_{k-1}^{(i)} \\ \ddot{\mathbf{w}} \end{bmatrix} = \begin{bmatrix} \dot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} \\ \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} \\ \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} \end{bmatrix} \end{aligned} \tag{16}$$

For convenience of analysis, similar to \mathbf{w}_{Lub} and $\overline{\boldsymbol{\psi}}_{Lub,k-1}^{(i)}$, we partition the regression vector $\mathbf{u}_{Lub,k,i}$ with length Lub into three parts $\dot{\mathbf{u}}_{k,i}$, $\ddot{\mathbf{u}}_{k,i}$ and $\ddot{\mathbf{u}}_{k,i}$. Considering the above notations and replacing (11) and (15) into (8) and (9) and padding all the vectors in (8), (9) and (11) with zeros to make their length equal to Lub , we have:

$$\begin{aligned} e_{k,L_k(i)}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}) &= \mathbf{u}_{Lopt,k,i} \mathbf{w}_{Lopt}^o - \mathbf{u}_{k,i} \boldsymbol{\psi}_{k-1}^{(i)} + v_k(i) \\ &= \mathbf{u}_{Lub,k,i} \begin{bmatrix} \mathbf{w}_{Lopt}^o \\ \mathbf{0}_{(Lub-Lopt) \times 1} \end{bmatrix} - \mathbf{u}_{Lub,k,i} \begin{bmatrix} \boldsymbol{\psi}_{k-1}^{(i)} \\ \mathbf{0}_{(Lub-L_k(i)) \times 1} \end{bmatrix} \\ &\quad + v_k(i) \\ &= \mathbf{u}_{Lub,k,i} \mathbf{w}_{Lub} - \mathbf{u}_{Lub,k,i} \boldsymbol{\psi}_{Lub,k-1}^{(i)} + v_k(i) \\ &= \mathbf{u}_{Lub,k,i} (\mathbf{w}_{Lub} - \boldsymbol{\psi}_{Lub,k-1}^{(i)}) + v_k(i) \\ &= \mathbf{u}_{Lub,k,i} \overline{\boldsymbol{\psi}}_{Lub,k-1}^{(i)} + v_k(i) \\ &= [\dot{\mathbf{u}}_{k,i} \quad \ddot{\mathbf{u}}_{k,i} \quad \ddot{\mathbf{u}}_{k,i}] \begin{bmatrix} \dot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} \\ \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} \\ \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} \end{bmatrix} + v_k(i) \\ &= \dot{\mathbf{u}}_{k,i} \dot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} + \ddot{\mathbf{u}}_{k,i} \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} + \ddot{\mathbf{u}}_{k,i} \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} + v_k(i) \end{aligned} \tag{17}$$

and

$$\begin{aligned} e_{k,L_k(i)-\Delta}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}) &= \mathbf{u}_{Lopt,k,i} \mathbf{w}_{Lopt}^o \\ &\quad - \mathbf{u}_{k,i}(1 : L_k(i) - \Delta) \boldsymbol{\psi}_{k-1}^{(i)}(1 : L_k(i) - \Delta) + v_k(i) \\ &= \mathbf{u}_{Lub,k,i} \begin{bmatrix} \mathbf{w}_{Lopt}^o \\ \mathbf{0}_{(Lub-Lopt) \times 1} \end{bmatrix} - \mathbf{u}_{Lub,k,i} \begin{bmatrix} \dot{\boldsymbol{\psi}}_{k-1}^{(i)} \\ \mathbf{0}_{(Lub-L_k(i)+\Delta) \times 1} \end{bmatrix} \\ &\quad + v_k(i) \\ &= \mathbf{u}_{Lub,k,i} \begin{bmatrix} \dot{\mathbf{w}} - \dot{\boldsymbol{\psi}}_{k-1}^{(i)} \\ \ddot{\mathbf{w}} \\ \ddot{\mathbf{w}} \end{bmatrix} + v_k(i) \\ &= [\dot{\mathbf{u}}_{k,i} \quad \ddot{\mathbf{u}}_{k,i} \quad \ddot{\mathbf{u}}_{k,i}] \begin{bmatrix} \dot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} \\ \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} \\ \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} \end{bmatrix} + v_k(i) \\ &= \dot{\mathbf{u}}_{k,i} \dot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} + \ddot{\mathbf{u}}_{k,i} \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} + \ddot{\mathbf{u}}_{k,i} \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} + v_k(i) \end{aligned} \tag{18}$$

The main term in the FT updating (7), is written as:

$$\begin{aligned} &(e_{k,L_k(i)}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}))^2 - (e_{k,L_k(i)-\Delta}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}))^2 \\ &= (e_{k,L_k(i)}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}) - e_{k,L_k(i)-\Delta}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)})) \\ &\quad \times (e_{k,L_k(i)}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}) + e_{k,L_k(i)-\Delta}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)})) \end{aligned} \tag{19}$$

Substituting (17) and (18) in (19) results:

$$\begin{aligned} &(e_{k,L_k(i)}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}))^2 - (e_{k,L_k(i)-\Delta}^{(L_k(i))}(\boldsymbol{\psi}_{k-1}^{(i)}))^2 \\ &= (\dot{\mathbf{u}}_{k,i} \dot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} - \ddot{\mathbf{u}}_{k,i} \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)}) \\ &\quad \times (2\dot{\mathbf{u}}_{k,i} \dot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} + \ddot{\mathbf{u}}_{k,i} \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} \\ &\quad + \ddot{\mathbf{u}}_{k,i} \ddot{\overline{\boldsymbol{\psi}}}_{k-1}^{(i)} + 2v_k(i)) \end{aligned} \tag{20}$$

By multiplying the parentheses in (20), we get:

$$\begin{aligned}
 & (e_{k,L_k(i)}^{(L_k(i))}(\psi_{k-1}^{(i)}))^2 - (e_{k,L_k(i)-\Delta}^{(L_k(i))}(\psi_{k-1}^{(i)}))^2 \\
 &= 2v_k(i)\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)} - 2v_k(i)\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}} \\
 &+ 2\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)} - 2\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}} \\
 &+ \left[\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)} \right]^2 - \left[\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}} \right]^2 \\
 &+ 2\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}}\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)} - 2\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}}\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}}
 \end{aligned} \tag{21}$$

For the ease of presentation, we define the following parameters:

$$\begin{aligned}
 A &= 2v_k(i)\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}, & B &= 2v_k(i)\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}} \\
 C &= 2\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}, & D &= 2\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}} \\
 E &= \left[\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)} \right]^2, & F &= \left[\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}} \right]^2 \\
 G &= 2\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}}\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}, & H &= 2\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}}\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}}
 \end{aligned} \tag{22}$$

From (7), (21) and (22) we have:

$$\begin{aligned}
 \ell_{k,f}(i) &= (\ell_{k-1,f}(i) - \alpha_k) \\
 &- \gamma_k[A - B + C - D + E - F + G - H]
 \end{aligned} \tag{23}$$

Now, according to the update equation, we analyze the steady-state performance. For simplicity of analysis, the following assumptions are made.

1. The regressors $\mathbf{u}_{L_{ub},k,i}$ are spatially and temporally independent.
2. The components of $\mathbf{u}_{L_{ub},k,i}$ are obtained from white Gaussian random process with zero mean and variance $\sigma_{u,k}^2$. The covariance matrix of $\mathbf{u}_{L_{ub},k,i}$ is $\mathbf{R}_{u,k} = \sigma_{u,k}^2 \mathbf{I}$.
3. In the steady-state, the tap-length in node k approximately converges to the constant value $L_k(\infty)$. Also we assume that at steady-state, $L_k(\infty) > L_{opt}$. This overestimate phenomenon will be justified by the results of the analysis.
4. The elements of the unknown parameter are obtained from the zero-mean random sequence with variance σ_o^2 .
5. We assume that in the steady-state

$$\left\{ \begin{bmatrix} \dot{\psi}_{k-1}^{(i)} \\ \ddot{\psi}_{k-1}^{(i)} \end{bmatrix} \begin{bmatrix} \dot{\psi}_{k-1}^{(i)} \\ \ddot{\psi}_{k-1}^{(i)} \end{bmatrix}^T \right\} = \sigma_{\psi,k-1}^2 \mathbf{I}, \text{ where } \sigma_{\psi,k-1}^2 \text{ is the variance of the elements of } \dot{\psi}_{k-1}^{(i)} \text{ and } \ddot{\psi}_{k-1}^{(i)}.$$

As a result of these assumptions, we can say that $\psi_{k-1}^{(i)}$ and consequently $\ddot{\psi}_{L_{ub},k-1}^{(i)}$ (or $\dot{\psi}_{k-1}^{(i)}$ and $\ddot{\psi}_{k-1}^{(i)}$) are independent from $\mathbf{u}_{L_{ub},k,i}$ and $v_k(i)$.

Considering the assumptions 1-5 and performing the mathematical expectations from both sides of (23) in steady-state, we have:

$$\begin{aligned}
 E\{\ell_{k,f}(i)\} &= (E\{\ell_{k-1,f}(i)\} - \alpha_k) \\
 &- \gamma_k E\{A - B + C - D + E - F + G - H\}, i \rightarrow \infty
 \end{aligned} \tag{24}$$

With the assumption of $L_k(\infty)$ and $L_{k-1}(\infty)$ as steady-state tap-length of nodes k and $k - 1$ we have:

$$\begin{aligned}
 L_k(\infty) &= (L_{k-1}(\infty) - \alpha_k) \\
 &- \gamma_k E\{A - B + C - D + E - F + G - H\}, i \rightarrow \infty
 \end{aligned} \tag{25}$$

To continue, we evaluate the moments in (25). From the assumptions 1–5 and their results we have:

$$\begin{aligned}
 E\{A\} &= E\{2v_k(i)\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\} = 2E\{v_k(i)\}E\{\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\} = 0 \\
 E\{B\} &= E\{2v_k(i)\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}}\} = 2E\{v_k(i)\}E\{\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}}\} = 0 \\
 E\{C\} &= E\{2\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\} \\
 &= 2E\{\ddot{\mathbf{u}}_{k,i}\}E\{\ddot{\psi}_{k-1}^{(i)}\}E\{\ddot{\mathbf{u}}_{k,i}\}E\{\ddot{\psi}_{k-1}^{(i)}\} = 0 \\
 E\{D\} &= E\{2\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}}\} = 2E\{\ddot{\mathbf{u}}_{k,i}\}E\{\ddot{\psi}_{k-1}^{(i)}\}E\{\ddot{\mathbf{u}}_{k,i}\}E\{\ddot{\mathbf{w}}\} = 0 \\
 E\{E\} &= E\left\{ \left[\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)} \right]^2 \right\} = \text{tr}\left(E\left\{ \left[\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)} \right]^2 \right\} \right) \\
 &= E(\text{tr}\{\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\}) \\
 &= E(\text{tr}\{\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\ddot{\psi}_{k-1}^{(i)T}\ddot{\mathbf{u}}_{k,i}^T\}) \\
 &= E(\text{tr}\{\ddot{\mathbf{u}}_{k,i}^T\ddot{\mathbf{u}}_{k,i}\ddot{\psi}_{k-1}^{(i)}\ddot{\psi}_{k-1}^{(i)T}\}) \\
 &= \text{tr}(E\{\ddot{\mathbf{u}}_{k,i}^T\ddot{\mathbf{u}}_{k,i}\}E\{\ddot{\psi}_{k-1}^{(i)}\ddot{\psi}_{k-1}^{(i)T}\}) \\
 &= \text{tr}(\sigma_{u,k}^2 \mathbf{I}_{\Delta} E\{\ddot{\psi}_{k-1}^{(i)}\ddot{\psi}_{k-1}^{(i)T}\}) \\
 &= \sigma_{u,k}^2 E(\text{tr}\{\ddot{\psi}_{k-1}^{(i)}\ddot{\psi}_{k-1}^{(i)T}\}) = \sigma_{u,k}^2 E\{\|\ddot{\psi}_{k-1}^{(i)}\|^2\} \\
 E\{F\} &= E\{[\ddot{\mathbf{u}}_{k,i}\ddot{\mathbf{w}}]^2\} = \sigma_{u,k}^2 E\{\|\ddot{\mathbf{w}}\|^2\}
 \end{aligned} \tag{26}$$

According to assumption in the steady-state, $L_k(\infty) > L_{opt}$, so, in steady-state $\ddot{\mathbf{w}} = 0$ and it is resulted that in the steady-state the terms G and H will disappear. From this and substituting(26) in (25) result:

$$L_k(\infty) = (L_{k-1}(\infty) - \alpha_k) - \gamma_k \sigma_{u,k}^2 \{E\{\|\ddot{\psi}_{k-1}^{(i)}\|^2\} - E\{\|\ddot{\mathbf{w}}\|^2\}\}, i \rightarrow \infty \tag{27}$$

If $(L_{k-1}(\infty) - \Delta) > L_{opt}$, then in the steady-state $\ddot{\mathbf{w}} = 0$, but, if $(L_{k-1}(\infty) - \Delta) \leq L_{opt}$, then $\ddot{\mathbf{w}}$ consists of $L_{opt} - L_{k-1}(\infty) + \Delta$ nonzero elements of $\mathbf{w}_{L_{opt}}^o$, so we have:

$$E\{\|\ddot{\mathbf{w}}\|^2\} = \begin{cases} (L_{opt} - L_{k-1}(\infty) + \Delta)\sigma_o^2 & \text{if } L_{opt} \leq L_{k-1}(\infty) \leq L_{opt} + \Delta \\ 0 & \text{if } L_{k-1}(\infty) > L_{opt} + \Delta \end{cases} \tag{28}$$

In fact, if $L_k(\infty)$ and $L_{k-1}(\infty)$ are not exactly equal, but are almost equal. We can temporarily assume $L_k(\infty) = L_{k-1}(\infty)$, and rewrite (27) as:

$$\{E\{\|\ddot{\psi}_{k-1}^{(i)}\|^2\} - E\{\|\ddot{\mathbf{w}}\|^2\}\} = -\alpha_k / \gamma_k \sigma_{u,k}^2 \quad i \rightarrow \infty \tag{29}$$

Now if we assume that $L_{k-1}(\infty) > (L_{opt} + \Delta)$, then according to (28) discrepancy is seen between the parties of (29). Because of this assumption, the left-hand side of (29) is higher than zero, but the right-hand side has a negative value. Thus, we conclude that $L_{opt} \leq L_{k-1}(\infty) \leq (L_{opt} + \Delta)$ will hold true. According to this and by defining $\mathbf{P}_{L_{ub},k-1} = \overline{\psi}_{L_{ub},k-1}^{(\infty)}$ and partitioning of $\mathbf{P}_{L_{ub},k-1}$ similar to $\overline{\psi}_{L_{ub},k-1}^{(i)}$ in (16) into three parts $\dot{\mathbf{P}}_{k-1}$, $\ddot{\mathbf{P}}_{k-1}$ and $\ddot{\mathbf{P}}_{k-1}$ we can rewrite (27) as:

$$L_k(\infty) = (L_{k-1}(\infty) - \alpha_k) - \gamma_k \sigma_{u,k}^2 \{E\{\|\dot{\mathbf{P}}_{k-1}\|^2\} - (L_{opt} - L_{k-1}(\infty) + \Delta)\sigma_o^2\} \tag{30}$$

In order to evaluate the term $E\{\|\dot{\mathbf{P}}_{k-1}\|^2\}$ we need the steady-state MSD of the deficient length DILMS algorithm. In [26], we have introduced the concept of deficient length DILMS algorithm and derived an expression for its steady-state MSD. The main results are included in appendix A to make the paper self-contained. Here we rewrite the resulting MSD that appears in appendix A, (71), as:

$$E\{\|\mathbf{P}_{L_{ub},k-1}\|^2\} = (1 - \Pi_{k,1})^{-1} s_k \tag{31}$$

In comparison with (71) here the subscript L_{opt} in $E\{\|\mathbf{P}_{L_{opt},k-1}\|^2\}$ is modified as L_{ub} , since here we are assuming that the length of the unknown parameter equals to L_{ub} . In (31), $\Pi_{k,1}$ and s_k are defined as:

$$\begin{aligned} \Pi_{k,\ell} &\triangleq \beta_{k-1} \beta_{k-2} \dots \beta_1 \beta_N \beta_{N-1} \dots \\ &\beta_{k+\ell} \beta_{k+\ell-1} \quad \ell = 1, \dots, N \\ s_k &\triangleq \Pi_{k,2} f_k + \Pi_{k,3} f_{k+1} + \dots + \Pi_{k,N-1} f_{k-3} \\ &+ \Pi_{k,N} f_{k-2} + f_{k-1} \end{aligned} \tag{32}$$

with

$$\beta_k = 1 - 2\mu_k \sigma_{u,k}^2 + \mu_k^2 \sigma_{u,k}^4 (L_{k-1}(\infty) + 2) \tag{33}$$

and

$$\begin{aligned} f_k &= 2\mu_k \sigma_{u,k}^2 (1 - \mu_k \sigma_{u,k}^2) \|\ddot{\mathbf{w}}\|^2 \\ &+ \mu_k^2 \sigma_{v,k}^2 \sigma_{u,k}^2 L_{k-1}(\infty) \\ &= \mu_k^2 \sigma_{v,k}^2 \sigma_{u,k}^2 L_{k-1}(\infty) \end{aligned} \tag{34}$$

By comparing (33) and (34) with (66) and (69), we observe that in (33) and (34), M has changed to $L_{k-1}(\infty)$, since here in the steady-state the length of adaptive filter in sensor $k - 1$ is equal to $L_{k-1}(\infty)$, while in appendix A, length M is considered for the adaptive filter of each sensor. Also $\|\ddot{\mathbf{w}}_{L_{opt}-M}\|^2$ in (69) has changed to $\|\ddot{\mathbf{w}}\|^2$ in (34), since according to partitioning in (13), here $\ddot{\mathbf{w}}_{L_{opt}-M}$ is equivalent to $\ddot{\mathbf{w}}$, which is also consistent with $\ddot{\mathbf{w}} = 0$ in the steady-state. Now using (31)-(34) we want to compute $E\{\|\dot{\mathbf{P}}_{k-1}\|^2\}$. In order to do so, first we divide the steady-state MSD into three parts as:

$$E\{\|\mathbf{P}_{L_{ub},k-1}\|^2\} = E\{\|\dot{\mathbf{P}}_{k-1}\|^2\} + E\{\|\ddot{\mathbf{P}}_{k-1}\|^2\} + E\{\|\ddot{\mathbf{P}}_{k-1}\|^2\} \tag{35}$$

Using the assumption 5, and since $\ddot{\mathbf{P}}_{k-1} = \overline{\psi}_{k-1}^{(\infty)} = \ddot{\mathbf{w}} = 0$, we have:

$$\begin{aligned} E\{\|\dot{\mathbf{P}}_{k-1}\|^2\} &= (L_{k-1}(\infty) - \Delta)\sigma_{\psi,k-1}^2 \\ E\{\|\ddot{\mathbf{P}}_{k-1}\|^2\} &= \Delta\sigma_{\psi,k-1}^2 \\ E\{\|\ddot{\mathbf{P}}_{k-1}\|^2\} &= 0 \end{aligned} \tag{36}$$

Substituting (36) into (35), we get:

$$E\{\|\mathbf{P}_{L_{ub},k-1}\|^2\} = L_{k-1}(\infty)\sigma_{\psi,k-1}^2 \tag{37}$$

From (37) and (36) we have:

$$E\{\|\ddot{\mathbf{P}}_{k-1}\|^2\} = \frac{\Delta}{L_{k-1}(\infty)} E\{\|\mathbf{P}_{L_{ub},k-1}\|^2\} \tag{38}$$

According to (31), the equation (38) is rewritten as:

$$E\{\|\ddot{\mathbf{P}}_{k-1}\|^2\} = \frac{\Delta}{L_{k-1}(\infty)} (1 - \Pi_{k,1})^{-1} s_k \tag{39}$$

or

$$E\{\|\ddot{\mathbf{P}}_{k-1}\|^2\} = \Delta(1 - \Pi_{k,1})^{-1} \dot{s}_k \tag{40}$$

where

$$\begin{aligned} \dot{s}_k \triangleq \frac{s_k}{L_{k-1}(\infty)} &= \Pi_{k,2} \dot{f}_k + \Pi_{k,3} \dot{f}_{k+1} + \dots \\ &+ \Pi_{k,N-1} \dot{f}_{k-3} \\ &+ \Pi_{k,N} \dot{f}_{k-2} + \dot{f}_{k-1} \end{aligned} \tag{41}$$

with

$$\dot{f}_k = \frac{f_k}{L_{k-1}(\infty)} = \mu_k^2 \sigma_{v,k}^2 \sigma_{u,k}^2 \tag{42}$$

Substituting (40) into (30) results:

$$\begin{aligned} L_k(\infty) &= L_{k-1}(\infty) - \alpha_k \\ &- \gamma_k \sigma_{u,k}^2 \{ \Delta(1 - \Pi_{k,1})^{-1} \dot{s}_k \\ &- (L_{opt} - L_{k-1}(\infty) + \Delta) \sigma_o^2 \} \end{aligned} \tag{43}$$

Equation (43) can be arranged as the following:

$$\begin{aligned} L_k(\infty) &= (1 - \gamma_k \sigma_{u,k}^2 \sigma_o^2) L_{k-1}(\infty) \\ &+ \gamma_k \sigma_{u,k}^2 \sigma_o^2 \{ L_{opt} + \Delta - \frac{\Delta}{\sigma_o^2} (1 - \Pi_{k,1})^{-1} \dot{s}_k \} - \alpha_k \end{aligned} \tag{44}$$

In the resulting equation, $\Pi_{k,1}$ depends on β_k and according to (33) depends on $L_{k-1}(\infty)$. But, on the other hand, as previously explained, we know $L_{opt} \leq L_{k-1}(\infty) \leq (L_{opt} + \Delta)$, and in practice $\Delta \ll L_{opt}$; thus, $L_{k-1}(\infty)$ is very close to L_{opt} . The equation (33) can be rewritten as:

$$\beta_k = 1 - 2\mu_k \sigma_{u,k}^2 + \mu_k^2 \sigma_{u,k}^4 (L_{opt} + 2) \tag{45}$$

By defining h_k and g_k as:

$$\begin{aligned} h_k &= 1 - \gamma_k \sigma_{u,k}^2 \sigma_o^2 \\ g_k &= \gamma_k \\ \sigma_{u,k}^2 \sigma_o^2 \{ L_{opt} + \Delta - \frac{\Delta}{\sigma_o^2} (1 - \Pi_{k,1})^{-1} \dot{s}_k \} - \alpha_k \end{aligned} \tag{46}$$

(44) can be written as:

$$L_k(\infty) = h_k L_{k-1}(\infty) + g_k \tag{47}$$

It should be noted that (47) is a coupled equation: it contains $L_k(\infty)$ and $L_{k-1}(\infty)$, i.e., information of two different locations. To mitigate this issue, we use ring topology [4]-[6]. Thus, by iterating (47), we have:

$$\begin{aligned} L_1(\infty) &= h_1 L_N(\infty) + g_1 \\ L_2(\infty) &= h_2 L_1(\infty) + g_2 \\ &\vdots \\ L_{k-2}(\infty) &= h_{k-2} L_{k-3}(\infty) + g_{k-2} \\ L_{k-1}(\infty) &= h_{k-1} L_{k-2}(\infty) + g_{k-1} \\ &\vdots \\ L_N(\infty) &= h_N L_{N-1}(\infty) + g_N \end{aligned} \tag{48}$$

Observe that according to (48), $L_{k-1}(\infty)$ can be express in term of $L_{k-3}(\infty)$ as:

$$L_{k-1}(\infty) = h_{k-1} h_{k-2} L_{k-3}(\infty) + h_{k-1} g_{k-2} + g_{k-1} \tag{49}$$

By iterating in this manner, we have:

$$\begin{aligned} L_{k-1}(\infty) &= h_{k-1} h_{k-2} \dots h_1 h_N h_{N-1} \dots h_k L_{k-1}(\infty) \\ &+ h_{k-1} h_{k-2} \dots h_1 h_N h_{N-1} \dots h_{k+1} g_k \\ &+ h_{k-1} h_{k-2} \dots h_1 h_N h_{N-1} \dots h_{k+2} g_{k+1} \\ &+ \dots + h_{k-1} h_{k-2} \dots h_1 h_N g_{N-1} \\ &+ h_{k-1} h_{k-2} \dots h_1 g_N + h_{k-1} h_{k-2} \dots h_2 g_1 \\ &+ h_{k-1} h_{k-2} \dots h_4 h_3 g_2 + \dots \\ &+ h_{k-1} h_{k-2} g_{k-3} + h_{k-1} g_{k-2} + g_{k-1} \end{aligned} \tag{50}$$

We use N quantities for each sensor k as:

$$\begin{aligned} G_{k,\ell} \triangleq &h_{k-1} h_{k-2} \dots h_1 h_N h_{N-1} \dots \\ &h_{k+\ell} h_{k+\ell-1}, \quad \ell = 1, \dots, N \end{aligned} \tag{51}$$

and m_k is defined as:

$$\begin{aligned} m_k \triangleq &G_{k,2} g_k + G_{k,3} g_{k+1} + \dots \\ &+ G_{k,N-1} g_{k-3} + G_{k,N} g_{k-2} + g_{k-1} \end{aligned} \tag{52}$$

From (50)-(52) we have:

$$L_{k-1}(\infty) = G_{k,1} L_{k-1}(\infty) + m_k \tag{53}$$

From (53), we illustrate the desired steady-state tap-length as:

$$L_{k-1}(\infty) = (1 - G_{k,1})^{-1} m_k \tag{54}$$

The equation (54) renders a mathematical expression of the steady-state tap-length. Because of the complicated form of the obtained results for the steady-state tap-length, it is almost difficult to make beneficial information about the steady-state tap-length of each sensor node. To achieve a clear view of the steady-state tap-length, we simplify the

equation (54). For this issue, we assume that all sensors employ the same step size for the tap-length adaption, i.e., $\gamma_k = \gamma, \forall k \in \mathcal{N}$, then, we assume that $\mathbf{R}_{u,k} = \sigma_u^2 \mathbf{I}$. With these assumptions, we have:

$$h_k = 1 - \gamma \sigma_u^2 \sigma_o^2 \tag{55}$$

Furthermore, we assume $\gamma \sigma_u^2 \sigma_o^2 \ll 1$ such that:

$$\begin{aligned} G_{k,1} &= h_1 h_2 \dots h_N \\ &= (1 - \gamma \sigma_u^2 \sigma_o^2)^N \approx 1 - N \gamma \sigma_u^2 \sigma_o^2 \end{aligned} \tag{56}$$

Also, m_k can be approximate as:

$$\begin{aligned} m_k &\approx \sum_{k=1}^N g_k = N \gamma \sigma_u^2 \sigma_o^2 (L_{opt} + \Delta) \\ &- \gamma \sigma_u^2 \Delta \sum_{k=1}^N (1 - \Pi_{k,1})^{-1} \dot{s}_k - \sum_{k=1}^N \alpha_k \end{aligned} \tag{57}$$

Substituting (56) and (57) in (54) result:

$$\begin{aligned} L_{k-1}(\infty) &= (L_{opt} + \Delta) \\ &- \frac{\Delta}{N \sigma_o^2} \sum_{k=1}^N (1 - \Pi_{k,1})^{-1} \dot{s}_k \\ &- \frac{\sum_{k=1}^N \alpha_k}{N \gamma \sigma_u^2 \sigma_o^2} \end{aligned} \tag{58}$$

Since the value of steady-state tap-length in (58) rarely is an integer, the steady-state tap-length usually varies nearby this value. The last two terms of the right-hand side of (58) are small and can be ignored. So, the steady-state tap-length is close to the value $L_{opt} + \Delta$, as will be presented in the simulations. This gives the result that the incremental FT algorithm overestimates the steady-state tap-length, and the third assumption is confirmed.

4 Simulation results

In this part, simulations are presented to verify the theoretical results of the paper. Numerical simulations are made under three low noise, high noise, and large step sizes for tap-length adaption conditions.

4.1 Low-noise scenario (SNR=20dB)

We assume a network with $N=12$ nodes exploring an unknown parameter with length $L_{opt}=20$. The elements of the unknown parameter are obtained from a white and zero-mean random sequence with variance $\sigma_o^2=0.01$. The regressors are independent zero-mean Gaussian variables, and their covariance matrix is $\mathbf{R}_{u,k} = \mathbf{I}$. The observation noise is zero-mean white Gaussian and scaled to have the $SNR=20\text{dB}$. The step size for the DILMS algorithm is set to $\mu_k=0.05$ for all nodes. The parameters of the incremental FT algorithm are chosen as $\delta_k=1, \alpha_k=0.001, \gamma_k=6, \Delta=6$ with

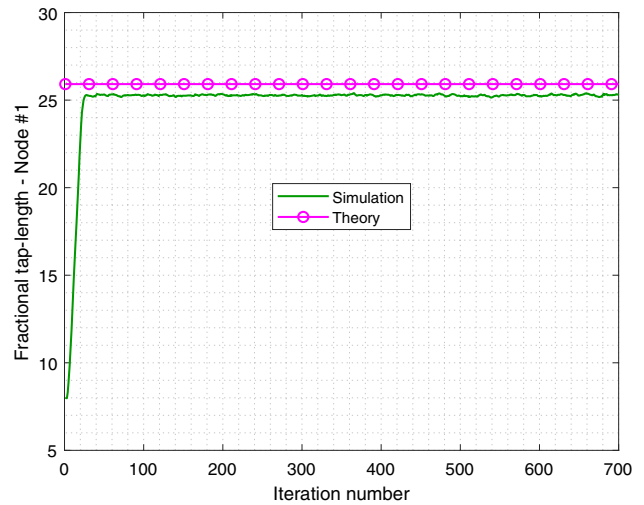


Fig. 1 Evolution of the fractional tap-length in low-noise scenario

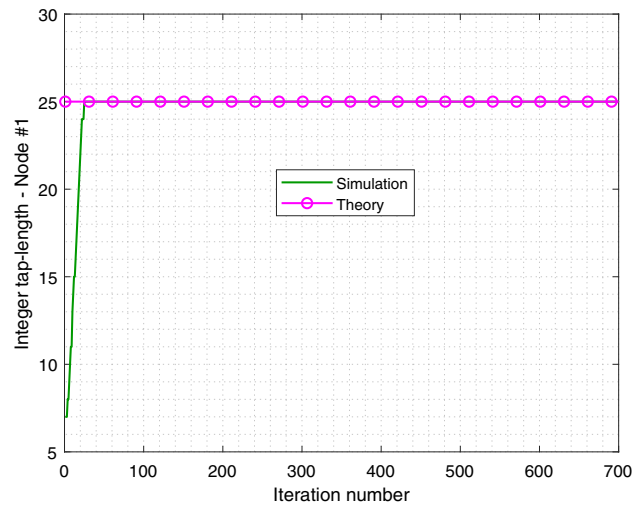


Fig. 2 Evolution of the integer tap-length in low-noise scenario

initial tap-length that are set in the minimum value $L_{min}=8$. Figures 1 and 2 respectively show the evolutions of the fractional and integer tap-length, under a low noise condition for node $k=1$. To obtain the integer values for the theoretical tap-lengths, we applied the floor operator $\lfloor \cdot \rfloor$ to (54). As shown in the figures, this node converges to a steady-state tap-length that corresponds well with the value obtained for the theoretical steady-state tap-length. We observe through the simulations that this is true for all nodes.

4.2 High-noise scenario (SNR=0dB)

In this case, a high-noise environment is considered for simulations. The parameters used in the simulations are the same as that of the previous case, but the observation noise is scaled to reach the $SNR=0\text{dB}$. The step size of the

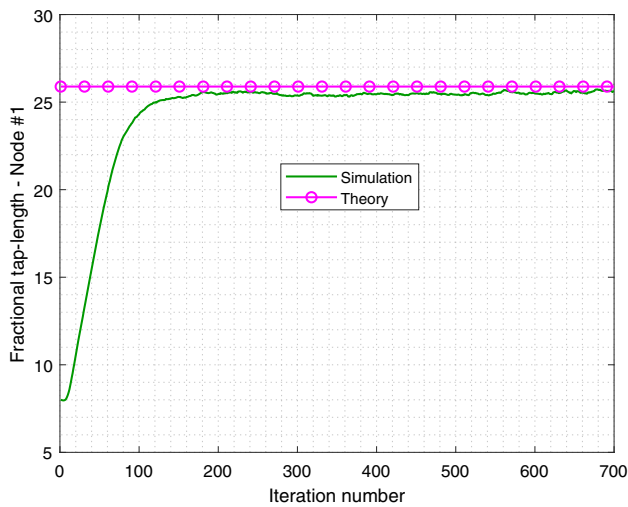


Fig. 3 Evolution of the fractional tap-length in high-noise scenario

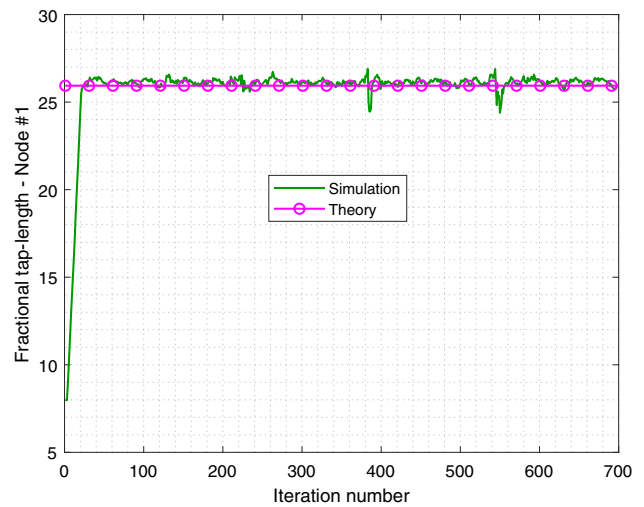


Fig. 5 Evolution of the fractional tap-length for large tap-length adaption step size, $\gamma_k=30$

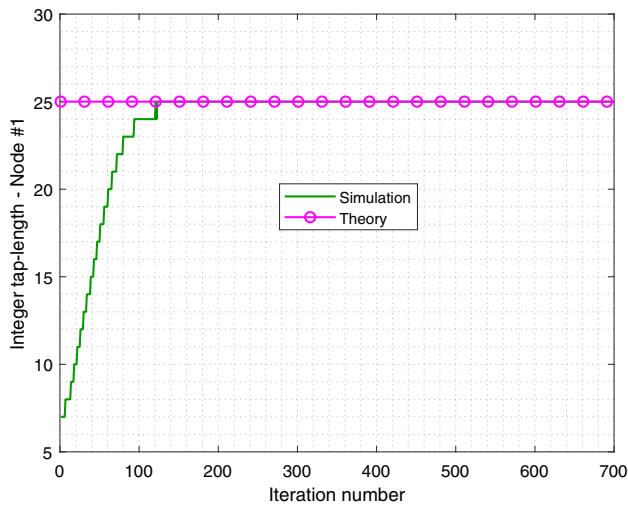


Fig. 4 Evolution of the integer tap-length in high-noise scenario

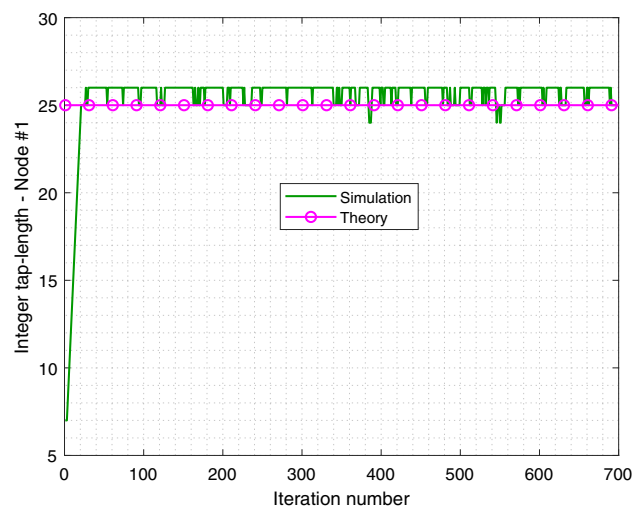


Fig. 6 Evolution of the integer tap-length for large tap-length adaption step size, $\gamma_k=30$

DILMS and tap-length adaption algorithm are set to $\mu_k=0.001$, and $\gamma_k=2$, respectively. The evolutions of the fractional and integer tap-length, under this condition, are shown in Figs. 3 and 4 for node $k=1$. It is observed that the simulation results support the theoretical analyses.

4.3 Large step size for tap-length adaption

The simulation setup of this scenario is the same as that of the first scenario, except that, in this case, we use a significant value for the tap-length adaption step size $\gamma_k=30$. However, increasing the step size γ_k increases the convergence rate, but a large γ_k causes a significant fluctuation of the evolution curve. The evolution curves of the fractional and integer tap-length, under this condition, are presented in Figs. 5 and 6 for node $k=1$. It is easily seen

from the figures that the steady-state tap-length fluctuates are about the theoretical values.

Table 2 shows the simulation results for different SNR values and compares them with the theoretical results. All parameters are following the previous experiments except for $\mu_k=0.001$ and $\gamma_k=2$. As the table shows, for different values of SNRs, the theoretical results are well compatible with the simulation results.

Table 3 also shows the results for different L_{opt} values and compares them with the theoretical results. In this simulation, $\mu_k=0.001$ and $\gamma_k=10$. As shown, first, the simulation results are compatible with the theoretical results. Secondly, all these results show that the length of the steady-state is overestimated almost as much as Δ . This overestimation of length is important because it shows that

Table 2 The steady-state tap-length performance in different SNRs

Length (Theoretical)	Length (Simulations)	SNR (dB)
25.7374	26.0317	−4
25.8715	25.7825	−2
25.8979	25.5523	0
25.9145	25.3627	2
25.9152	25.3960	4
25.9320	25.3807	6
25.9452	25.3675	8
25.9403	25.3026	10
25.9460	25.2738	12
25.9483	25.3649	14
25.9485	25.2994	16
25.9493	25.2700	18
25.9491	25.2947	20
25.9496	25.2510	22
25.9497	25.3360	24

Table 3 The steady-state tap-length performance in different values of L_{opt}

Length (Theoretical)	Length (Simulations)	L_{opt}
25.9896	25.5805	20
55.9891	55.6295	50
105.9875	105.6483	100
133.9862	133.5513	128
261.9831	261.9681	256
517.9772	518.2319	512

length is not estimated deficiently. As shown in [26], deficient length estimation increases the steady-state error. On the other hand, a length much larger than the optimal value increases the computational complexity, which is not proper for a network of low-power sensors. However, the overestimation of the length is not significant and almost equal to Δ , which is much smaller than the optimal length.

5 Conclusions

In this paper, we investigated the performance of the distributed incremental variable fractional tap-length (FT) LMS algorithm. Since in this scenario, the length of the vectors varies in each iteration, so we defined an upper bound for the size and padded the vectors by zeros to achieve this length. We obtained the vectors with the same sizes that simplified the calculations. Based on these

findings, we analyzed the incremental variable FT LMS algorithm and derived the mathematical expression for the steady-state tap-length at each sensor node. The most important result is that this algorithm overestimates the optimal value of the tap-length. Computer simulation results confirmed the mathematical analysis and the derived closed-form expression of tap-length.

Appendix A.

In [26], we have presented the concept of a short-length DILMS algorithm, and we have provided an expression for its steady-state MSD. This steady-state MSD is needed for the evaluation of the term $E\{\|\ddot{\mathbf{P}}_{k-1}\|^2\}$ in (30), so the main results are included here. For this aim, we assume a set of N sensors in order to find an $L_{opt} \times 1$ unknown vector $\mathbf{w}_{L_{opt}}^o$ with unknown length L_{opt} from multiple measurements collected at N sensor nodes in the network. As previously mentioned, to find the length of the unknown parameter, a variable tap-length algorithm is needed. But, we assume that such an algorithm is not applicable for reasons such as energy storage (since energy consumption is an essential issue in WSNs). So, at each sensor, a conjectural length M for the unknown parameter is considered. More clearly, each sensor is equipped with an adaptive filter with M coefficients ($M < L_{opt}$). Now at each sensor node, only the algorithm (3) is applicable in which all vectors have length M .

Now, we present a steady-state analysis for this deficient length case. To perform this analysis, first, we partition the unknown parameter $\mathbf{w}_{L_{opt}}^o$ as follow:

$$\begin{bmatrix} \dot{\mathbf{w}}_M \\ \ddot{\mathbf{w}}_{L_{opt}-M} \end{bmatrix} \tag{59}$$

where $\dot{\mathbf{w}}_M$ is the partition of $\mathbf{w}_{L_{opt}}^o$ that is modeled by $\psi_k^{(i)}$ in each sensor, and $\ddot{\mathbf{w}}_{L_{opt}-M}$ is the partition of $\mathbf{w}_{L_{opt}}^o$ that excluded in the estimation of $\mathbf{w}_{L_{opt}}^o$. The partitioning simplifies the work with the variable-length vectors. Regarding the partitioning and using the data model (11), the update equation (3) is expressed as:

$$\begin{aligned} \psi_k^{(i)} &= \psi_{k-1}^{(i)} - \mu_k \mathbf{u}_{k,i}^* \mathbf{u}_{L_{opt},k,i} \overline{\psi}_{L_{opt},k-1}^{(i)} \\ &\quad + \mu_k \nu_k(i) \mathbf{u}_{k,i}^* \end{aligned} \tag{60}$$

where the vector $\overline{\psi}_{L_{opt},k-1}^{(i)}$ with length L_{opt} , computes the difference among the weight at sensor $k - 1$ and $\mathbf{w}_{L_{opt}}^o$ as:

$$\bar{\Psi}_{L_{opt},k-1}^{(i)} = \begin{bmatrix} \Psi_{k-1}^{(i)} \\ \mathbf{O}_{(L_{opt}-M) \times 1} \end{bmatrix} - \begin{bmatrix} \dot{\mathbf{w}}_M \\ \ddot{\mathbf{w}}_{L_{opt}-M} \end{bmatrix} \quad (61)$$

By padding the $L_{opt} - M$ zeros in vectors with length M in (60), and subtracting the unknown parameter $\mathbf{w}_{L_{opt}}^o$ from both of its sides results in:

$$\bar{\Psi}_{L_{opt},k}^{(i)} = \Lambda_k(i) \bar{\Psi}_{L_{opt},k-1}^{(i)} + \mu_k v_k(i) \begin{bmatrix} \mathbf{u}_{k,i}^* \\ \mathbf{O}_{(L_{opt}-M) \times 1} \end{bmatrix} \quad (62)$$

where

$$\Lambda_k(i) = I_{L_{opt}} - \mu_k \begin{bmatrix} \mathbf{u}_{k,i}^* \\ \mathbf{O}_{(L_{opt}-M) \times 1} \end{bmatrix} \mathbf{u}_{L_{opt},k,i} \quad (63)$$

To demonstrate the steady-state MSD, first we write the $\|\bar{\Psi}_{L_{opt},k}^{(i)}\|^2$ as:

$$\begin{aligned} \|\bar{\Psi}_{L_{opt},k}^{(i)}\|^2 &= \bar{\Psi}_{L_{opt},k-1}^{*(i)} \Lambda_k^*(i) \Lambda_k(i) \bar{\Psi}_{L_{opt},k-1}^{(i)} \\ &+ \mu_k v_k(i) \bar{\Psi}_{L_{opt},k-1}^{*(i)} \Lambda_k^*(i) \begin{bmatrix} \mathbf{u}_{k,i}^* \\ \mathbf{O}_{(L_{opt}-M) \times 1} \end{bmatrix} \\ &+ \mu_k v_k(i) \begin{bmatrix} \mathbf{u}_{k,i} & \mathbf{O}_{1 \times (L_{opt}-M)} \end{bmatrix} \Lambda_k(i) \bar{\Psi}_{L_{opt},k-1}^{(i)} \\ &+ \mu_k^2 v_k^2(i) \|\mathbf{u}_{k,i}\|^2 \end{aligned} \quad (64)$$

Considering the assumptions 1, 2, and employing the mathematical expectations from both sides of (64), after some tedious algebra leads to:

$$E\{\|\bar{\Psi}_{L_{opt},k}^{(i)}\|^2\} = \beta_k E\{\|\bar{\Psi}_{L_{opt},k-1}^{(i)}\|^2\} + (\eta_k - \beta_k) \|\ddot{\mathbf{w}}_{L_{opt}-M}\|^2 + \tau_k \quad (65)$$

where

$$\begin{aligned} \beta_k &= 1 - 2\mu_k \sigma_{u,k}^2 + \mu_k^2 \sigma_{u,k}^4 (M + 2) \\ \eta_k &= 1 + \mu_k^2 \sigma_{u,k}^4 M \end{aligned} \quad (66)$$

and

$$\tau_k = \mu_k^2 \sigma_{v,k}^2 \sigma_{u,k}^2 M \quad (67)$$

Now we write (65) in the brief form of below:

$$E\{\|\bar{\Psi}_{L_{opt},k}^{(i)}\|^2\} = \beta_k E\{\|\bar{\Psi}_{L_{opt},k-1}^{(i)}\|^2\} + f_k \quad (68)$$

where

$$f_k = (\eta_k - \beta_k) \|\ddot{\mathbf{w}}_{L_{opt}-M}\|^2 + \tau_k \quad (69)$$

According to the steady-state analysis, as $(i \rightarrow \infty)$ in (68), and assuming $\mathbf{P}_{L_{opt},k} = \bar{\Psi}_{L_{opt},k}^{(\infty)}$, (68) is rewritten as:

$$E\{\|\mathbf{P}_{L_{opt},k}\|^2\} = \beta_k E\{\|\mathbf{P}_{L_{opt},k-1}\|^2\} + f_k \quad (70)$$

This equation has a similar structure with (47), therefore in

the same manner that was used for the solving of the recursive equation (47) we have:

$$E\{\|\mathbf{P}_{L_{opt},k-1}\|^2\} = (1 - \Pi_{k,1})^{-1} s_k \quad (71)$$

where

$$\begin{aligned} \Pi_{k,\ell} &\triangleq \beta_{k-1} \beta_{k-2} \dots \beta_1 \beta_N \beta_{N-1} \dots \\ &\beta_{k+\ell} \beta_{k+\ell-1}, \quad \ell = 1, \dots, N \end{aligned} \quad (72)$$

and

$$\begin{aligned} s_k &\triangleq \Pi_{k,2} f_k + \Pi_{k,3} f_{k+1} + \dots + \Pi_{k,N-1} f_{k-3} + \Pi_{k,N} f_{k-2} \\ &\quad + f_{k-1} \end{aligned} \quad (73)$$

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

References

- Liu, J., Zhao, Z., Ji, J., & Miaolong, H. (2020). Research and application of wireless sensor network technology in power transmission and distribution system. *Intelligent and Converged Networks, 1*(2), 199–220.
- Zhou, Y., Yang, L., Yang, L., & Ni, M. (2019). Novel energy-efficient data gathering scheme exploiting spatial-temporal correlation for wireless sensor networks. *Wireless Communications and Mobile Computing, 2019*.
- Azarnia, G., Tinati, M. A., Sharifi, A. A., & Shiri, H. (2020). Incremental and diffusion compressive sensing strategies over distributed networks. *Digital Signal Processing, 101*, 102732.
- Saeed, M. O. B., & Zerguine, A. (2019). An incremental variable step-size lms algorithm for adaptive networks. *IEEE Transactions on Circuits and Systems II: Express Briefs, 67*(10), 2264–2268.
- Dimple, K., Kotary, D. K., & Nanda, S. J. (2017). An incremental rls for distributed parameter estimation of iir systems present in computing nodes of a wireless sensor network. *Procedia Computer Science, 115*, 699–706.
- Li, L., Chambers, J. A., Lopes, C. G., & Sayed, A. H. (2009). Distributed estimation over an adaptive incremental network based on the affine projection algorithm. *IEEE Transactions on Signal Processing, 58*(1), 151–164.
- Huang, W., Li, L., Li, Q., & Yao, X. (2018). Diffusion robust variable step-size lms algorithm over distributed networks. *IEEE Access, 6*, 47511–47520.
- Rastegarnia, A. (2019). Reduced-communication diffusion rls for distributed estimation over multi-agent networks. *IEEE Transactions on Circuits and Systems II: Express Briefs, 67*(1), 177–181.
- Li, L., & Chambers, J. A. (2009). Distributed adaptive estimation based on the apa algorithm over diffusion networks with changing topology. In *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*, (pp. 757–760). IEEE.
- Riera-Palou, F., Noras, J. M., & Cruickshank, D. G. M. (2001). Linear equalisers, with dynamic and automatic length selection. *Electronics Letters, 37*(25), 1553–1554.

11. Shaozhong, F., Jianhua, G., & Yong, W. Fast adaptive algorithms for updating variable length equalizer based on exponential. In *2008 4th International Conference on Wireless Communications, Networking and Mobile Computing*.
12. Rusu, C., & Cowan, C. F. N. (2001). Novel stochastic gradient adaptive algorithm with variable length. In *European Conference on Circuit Theory and Design (ECCTD'01)*.
13. Bilcu, R. C., Kuosmanen, P., & Egiazarian, K. (2002). A new variable length lms algorithm: Theoretical analysis and implementations. In *9th International conference on Electronics, Circuits and Systems*, volume 3, (pp. 1031–1034). IEEE.
14. Bilcu, R. C., Kuosmanen, P., & Egiazarian, K. (2006). On length adaptation for the least mean square adaptive filters. *Signal Processing*, 86(10), 3089–3094.
15. Yuantao, G., Tang, K., Cui, H., & Wen, D. (2003). Convergence analysis of a deficient-length lms filter and optimal-length sequence to model exponential decay impulse response. *IEEE Signal Processing Letters*, 10(1), 4–7.
16. Zhang, Y., Chambers, J. A., Sanei, S., Kendrick, P., & Cox, T. J. (2007). A new variable tap-length lms algorithm to model an exponential decay impulse response. *IEEE Signal Processing Letters*, 14(4), 263–266.
17. Shi, K., Ma, X., Tong, G., & Zhou. (2009). A variable step size and variable tap length lms algorithm for impulse responses with exponential power profile. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 3105–3108). IEEE.
18. Mayyas, K. (2005). Performance analysis of the deficient length lms adaptive algorithm. *IEEE Transactions on Signal Processing*, 53(8), 2727–2734.
19. Yuantao, G., Tang, K., & Cui, H. (2004). Lms algorithm with gradient descent filter length. *IEEE Signal Processing Letters*, 11(3), 305–307.
20. Gong, Yu., & Cowan, C. F. N. (2004). Structure adaptation of linear mmse adaptive filters. *IEE Proceedings-Vision, Image and Signal Processing*, 151(4), 271–277.
21. Gong, Yu., & Cowan, C. F. N. (2005). An lms style variable tap-length algorithm for structure adaptation. *IEEE Transactions on Signal Processing*, 53(7), 2400–2407.
22. Ali, A., Moinuddin, M., & Al-Naffouri, T. Y. (2021). The nlms is steady-state schur-convex. *IEEE Signal Processing Letters*, 28, 389–393.
23. Lu, L., Chen, L., Zheng, Z., Yi, Yu., & Yang, X. (2020). Behavior of the lms algorithm with hyperbolic secant cost. *Journal of the Franklin Institute*, 357(3), 1943–1960.
24. Luo, L., & Xie, A. (2020). Steady-state mean-square deviation analysis of improved l0-norm-constraint lms algorithm for sparse system identification. *Signal Processing*, 175, 107658.
25. Li, L., Zhang, Y., & Chambers, J. A. (2008). Variable length adaptive filtering within incremental learning algorithms for distributed networks. In *2008 42nd Asilomar Conference on Signals, Systems and Computers* (pp. 225–229). IEEE.
26. Azarnia, G., & Tinati, M. A. (2015). Steady-state analysis of the deficient length incremental lms adaptive networks. *Circuits, Systems, and Signal Processing*, 34(9), 2893–2910.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Ghanbar Azarnia was born in 1988 in Iran. He received B.Sc., M.Sc. and Ph.D. degrees in Electrical Engineering (Communication) from University of Tabriz, Tabriz, Iran, in 2011, 2013 and 2017 respectively. He is currently with the Engineering Faculty of Khoy, Urmia University of Technology, Urmia, Iran. His current research interests include compressed sensing, statistical signal processing, biomedical signal processing, adaptive filtering, wireless sensor networks and distributed processing.



Abbas Ali Sharifi received the Ph.D. degree in telecommunication engineering from University of Tabriz, Iran, in 2015. Now he is an associate professor in the department of electrical engineering at the University of Bonab. His current research interests include wireless communication, signal processing, and cognitive radio networks.