Received 25 February 2024, accepted 7 March 2024, date of publication 11 March 2024, date of current version 20 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3375757

RESEARCH ARTICLE

Hidden Markov Model Decorrelated Diffusion Leaky LMS With Optimized Leaky Factor

HADI ZAYYANI[1], (Member, IEEE), MOHAMMAD SALMAN[2], (Senior Member, IEEE), MEHDI KORKI[3], (Member, IEEE), AND AHMED A. F. YOUSSEF[2], (Member, IEEE)

[1]Department of Electrical and Computer Engineering, Qom University of Technology (QUT), Qom 37181-46645, Iran

[2]College of Engineering and Technology, American University of the Middle East, Egaila 54200, Kuwait

[3]School of Science, Computing, and Engineering Technologies, Swinburne University of Technology, Melbourne, VIC 3122, Australia

Corresponding author: Hadi Zayyani (zayyani@qut.ac.ir)

ABSTRACT In this paper, two basic novel modifications of Diffusion Leaky Least Mean Square (DL-LMS) algorithm are proposed. First, the optimized leaky factor is obtained in closed form using the minimum disturbance principle, thanks to quadratic form of disturbance. Then, a low complexity decorrelated version of DL-LMS is presented based on the statistical Hidden Markov Modeling (HMM) of time correlations of input signals. The decorrelation is performed using a simple one-tap filter. Simulation results demonstrate the effectiveness of both modifications in comparison to other methods.

INDEX TERMS Distributed estimation, leaky LMS, hidden Markov model, diffusion.

I. INTRODUCTION

Distributed estimation problem in a network is a well known field in signal processing which has applications in channel estimation, spectrum sensing, target tracking and etc [1]. The nodes of the network collaborate to each other to estimate an unknown parameter vector. The collaboration strategies are incremental, consensus, and diffusion [1]. The diffusion strategy reported to have more numerical stability, scalability, ease of computations and outperforms other strategies [2].

Among diffusion algorithms [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], diffusion LMS is the most versatile diffusion algorithm. In LMS adaptive filters, the implementation of LMS algorithm reported to be problematic [19], [20]. Two problems are numerical problem and stagnation behavior [20]. Numerical problem is due to inadequacy of excitation in the input data while the stagnation behavior is due to low input signal [20]. So, as such leaky LMS is suggested in adaptive filter literature [19], [21] to solve these problems, the diffusion leaky LMS algorithm is proposed for distributed estimation as well [20]. In [20], a variable Leaky LMS algorithm is also presented in which the leaky factor is updated recursively based on minimizing the square of the instantaneous error. Moreover, [22] proposed a leaky zero attracting DLMS algorithm in which some sparsity regularizers are exploited to enhance the performance in sparse parameter vector settings.

In different direction of research, some papers deal with the correlated input cases for distributed estimation [23], [24]. The correlated input signal makes the algorithm converge slowly. In [23], to increase the convergence rate in presence of correlated input, the decorrelated method is presented for the distributed estimation problem. In the decorrelation method for system identification in [23], both the input signal to the adaptive filter and the output of the system are decorrelated by a transversal Finite Impulse Response (FIR) filter. These decorrelation filters makes the algorithm complex especially when it would be implemented in a sensor network. In addition, [23] uses two convex-combined decorrelated transversal filter which doubly complicates the scheme. Moreover, in a recent work [24], a class of diffusion Bayesian decorrelation least mean squares algorithms are presented based on decorrelated observation models. Again, the Bayesian nature of decorrelation method utilized in this paper is complex to implement in a sensor network in which the size and limited computational complexity is a requisite.

In this paper, due to the property of limited convergence rate of leaky LMS in presence of correlated input signal, the focus will be on improving the convergence rate of diffusion leaky LMS using the decorrelation method. The

The associate editor coordinating the review of this manuscript and approving it for publication was Jingen Ni.

real-life application of this work is highlighted when there are implications such as acoustic echo cancellers when there is some correlated input signal. The contribution of our paper is twofold. First, since the performance of the diffusion leaky LMS is dependent on the leaky factor, an optimized closed-form leaky factor is derived which is calculated during the iterations of the algorithm. Unlike the variable diffusion leaky LMS presented in [20] which is recursive and is just based on the instantaneous error, it has a closed-form formula and is derived based on minimizing the disturbance term. Second contribution is to devise a simple decorrelation method based on Hiden Markov Model (HMM) rather than Auto-Regressive (AR) model used in [23]. HMM modeling has a well-established history in the statistical signal processing community [25], [26], in speech processing [26], [27], and in other applications such as in power systems [28]. In this paper, a two-state first-order HMM model to ensure that the decorrelation algorithm is of low complexity, both in its derivation and implementation is employed. The HMM decorrelator is a one-tap FIR filter (which is shown in Fig. 1) which is derived mathematically based on decorrelation conditions. Simulation results show the effectiveness of the proposed algorithm in both using optimized leaky factor and also HMM decorrelator.

The rest of the paper is organized as follows. Section II introduces the problem and the model employed in the paper. Section III introduces the diffusion leaky LMS. In Section IV, the proposed improved diffusion leaky LMS algorithm is derived and explained. The simulation results are presented in Section V. Finally, conclusions are given in Section VI.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A network with N nodes is assumed. Each node observes an scalar measurement $d_{k,i}$ which is equal to

$$d_{k,i} = \mathbf{u}_{k,i}^T \boldsymbol{\omega}^o + v_{k,i}, \tag{1}$$

where k is the index of the node, i is the time index, $\boldsymbol{\omega}^o$ is a $L \times 1$ unknown vector where L is the size of unknown vector, $\mathbf{u}_{k,i}$ is a $L \times 1$ known regression vector, and $v_{k,i}$ is the measurement noise. All the N sensors of the network collect the measurements and try to collaboratively estimate the unknown vector $\boldsymbol{\omega}^o$. In Adapt-then-combine (ATC) diffusion algorithms, each node firstly updates its local estimate based on a local cost function. In diffusion LMS algorithms, the cost function is a MSE function as $\mathrm{J}_{\mathrm{Loc}}(\boldsymbol{\omega}) = \sum_{l\in\mathcal{N}_k} a_{l,k}\mathrm{E}\{||d_{l,i} - \mathbf{u}_{l,i}^T\boldsymbol{\omega}^o||^2\}$, where $\mathcal{N}_k$ is the neighborhood set of node k, and $a_{l,k}$ is the combination coefficients in the adaptation step [1]. The adaptation step of ATC diffusion LMS is [1]

$$\boldsymbol{\phi}_{k,i} = \boldsymbol{\omega}_{k,i} + \mu \sum_{l\in\mathcal{N}_k} a_{l,k}\mathbf{u}_{l,i}^T(d_{l,i} - \mathbf{u}_{l,i}^T\boldsymbol{\omega}_{k,i}), \tag{2}$$

where $\boldsymbol{\omega}_{k,i}$ is the estimation of node k at the end of index i, $\boldsymbol{\phi}_{k,i}$ is the intermediate estimation of node k at time index i and μ is the step size. After adaptation, each node sends its local updated estimate to its neighborhoods. Then, each node combines the received local estimates from its neighborhoods

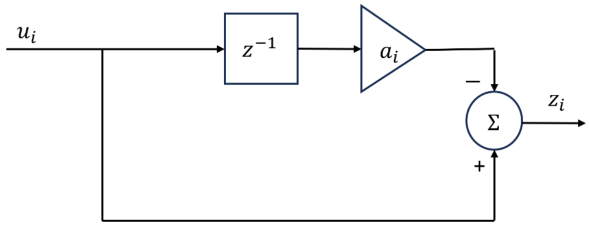

FIGURE 1. Block diagram of 1-tap decorrelator.

which are $\tilde{\boldsymbol{\phi}}_{l,i}$ and may be different in comparison to true intermediate estimations $\boldsymbol{\phi}_{l,i}$. So, in the combination step, we have

$$\boldsymbol{\omega}_{k,i+1} = \sum_{l\in\mathcal{N}_k} c_{l,k}\tilde{\boldsymbol{\phi}}_{l,i}, \tag{3}$$

where c_{lk} is the combination coefficients in the combination step [1]. After Combination step, by receiving new measurements, the diffusion LMS algorithm go ahead until final convergence. In the next section, the diffusion leaky LMS algorithm is introduced.

III. THE DIFFUSION LEAKY LMS ALGORITHM

Leaky LMS is a well known adaptive filter algorithm which has advantages of numerical stability over classical LMS algorithm. It is generalized to diffusion Leaky LMS to estimate the unknown vector in a collaborative network. The only difference between Diffusion Leaky LMS (DL-LMS) and DLMS is in the adaptation step which is

$$\boldsymbol{\phi}_{k,i} = (1 - \mu\gamma_{k,i})\boldsymbol{\omega}_{k,i} + \mu \sum_{l\in\mathcal{N}_k} a_{l,k}\mathbf{u}_{l,i}^T(d_{l,i} - \mathbf{u}_{l,i}^T\boldsymbol{\omega}_{k,i}), \tag{4}$$

where $\gamma_{k,i}$ is the leaky factor which has the role of tradeoff between the final MSE and convergence rate. In the next section, the proposed algorithm is introduced.

IV. THE PROPOSED HMM DECORRELATION DIFFUSION LEAKY LMS WITH OPTIMIZED LEAKY FACTOR

In this paper, two improvements to diffusion leaky LMS is suggested. Firstly, it is suggested to use an optimized value for the leaky factor based on minimum disturbance principle. Minmum disturbance principle is a well known principle in adaptive filtering theory which means that the adaptive filter design should obey a minimum disturbance principle to gradually find the optimum Wiener filter solution of adaptive filter [29]. This principle is used for robust distributed estimation against impulsive noise in [11]. Secondly, an HMM model for correlation of input signals of local leaky LMS adaptive filters is used. Then, a low complexity decorrelation diffusion leaky LMS algorithm is presented. These two modifications are introduced in the two next subsections.

A. OPTIMIZED LEAKY FACTOR

To optimally design the leaky factor, minimum disturbance principle is used. Nominating the difference of estimation in successive iterations as $\Delta\Omega \triangleq \boldsymbol{\omega}_{k,i+1} - \boldsymbol{\omega}_{k,i}$, then we have:

$$\begin{aligned}\Delta\Omega &= \boldsymbol{\omega}_{k,i+1} - \boldsymbol{\omega}_{k,i} = \sum_{l\in\mathcal{N}_k} c_{l,k}\tilde{\boldsymbol{\phi}}_{l,i} - \boldsymbol{\omega}_{k,i} \\ &= \sum_{l\in\mathcal{N}_k} c_{l,k}\Big[(1-\mu\gamma)\boldsymbol{\omega}_{l,i} + \mu\mathbf{p}_l\Big] - \boldsymbol{\omega}_{k,i} \\ &= (1-\mu\gamma)\sum_{l\in\mathcal{N}_k^-} c_{l,k}\boldsymbol{\omega}_{l,i} \\ &\quad + \mu\sum_{l\in\mathcal{N}_k} c_{l,k}\mathbf{p}_l + c_{kk}(1-\mu\gamma)\boldsymbol{\omega}_{k,i} - \boldsymbol{\omega}_{k,i}, \end{aligned} \tag{5}$$

where $\mathcal{N}_k^-$ is the neighborhood of node k excluding k'th node, and $\mathbf{b}_k \triangleq \sum_{l\in\mathcal{N}_k} c_{l,k}\mathbf{u}_{l,i}^T(d_{l,i} - \mathbf{u}_{l,i}^T\boldsymbol{\omega}_{k,i})$. So, following (5), we have

$$\Delta\Omega = (1-\mu\gamma)\mathbf{d}_k + \mathbf{h}_k + g_k\boldsymbol{\omega}_{k,i}, \tag{6}$$

where

$$g_k \triangleq c_{kk}(1-\mu\gamma) - 1, \tag{7}$$

and

$$\mathbf{d}_k \triangleq \sum_{l\in\mathcal{N}_k^-} c_{l,k}\boldsymbol{\omega}_{l,i}, \tag{8}$$

and

$$\mathbf{h}_k \triangleq \mu\sum_{l\in\mathcal{N}_k} c_{l,k}\mathbf{p}_l. \tag{9}$$

Hence, to find the optimum leaky factor γ_{opt} based on minimization of disturbance term which is defined as to $||\Delta\Omega||^2$[1] [11], [29], the quadratic disturbance cost function can be written as

$$\begin{aligned}\mathrm{f}(\gamma) &= ||\Delta\Omega||^2 = ||(1-\mu\gamma)\mathbf{d}_k + \mathbf{h}_k + g_k\boldsymbol{\omega}_{k,i}||^2 \\ &= (1-\mu\gamma)^2||\mathbf{d}_k||^2 + g_k^2||\boldsymbol{\omega}_{k,i}||^2 + ||\mathbf{h}_k||^2 \\ &\quad + (1-\mu\gamma)\mathbf{d}_k^T\mathbf{h}_k + g_k(1-\mu\gamma)\mathbf{d}_k^T\boldsymbol{\omega}_{k,i} + g_k\mathbf{h}_k^T\boldsymbol{\omega}_{k,i} \\ &\quad + (1-\mu\gamma)\mathbf{h}_k^T\mathbf{d}_k + g_k\boldsymbol{\omega}_{k,i}^T\mathbf{h}_k \\ &\quad + g_k(1-\mu\gamma)\boldsymbol{\omega}_{k,i}^T\mathbf{d}_k. \end{aligned} \tag{10}$$

The quadratic function $\mathrm{f}(\gamma)$ is reduced to

$$\mathrm{f}(\gamma) = \bar{a}_k\gamma^2 + \bar{b}_k\gamma + \bar{c}_k, \tag{11}$$

where we have

$$\bar{a}_k = \mu^2||\mathbf{d}_k||^2 + c_{kk}^2\mu^2||\boldsymbol{\omega}_{k,i}||^2 + 2c_{kk}\mu^2\mathbf{d}_k^T\boldsymbol{\omega}_{k,i}, \tag{12}$$

and

$$\begin{aligned}\bar{b}_k &= -2\mu||\mathbf{d}_k||^2 + 2\mu c_{kk}||\boldsymbol{\omega}_{k,i}||^2 - \mu\mathbf{h}_k^T\mathbf{d}_k \\ &\quad - 2\mu c_{kk}^2||\boldsymbol{\omega}_{k,i}||^2 + 2\mu\mathbf{d}_k^T\boldsymbol{\omega}_{k,i} - 2\mu c_{kk}\mathbf{d}_k^T\boldsymbol{\omega}_{k,i} \\ &\quad - 2\mu c_{kk}\mathbf{h}_k^T\boldsymbol{\omega}_{k,i} - 2\mu c_{kk}\boldsymbol{\omega}_{k,i}^T\mathbf{d}_k, \end{aligned} \tag{13}$$

[1] The well known quadratic form of disturbance makes the mathematical derivations tractable.

and

$$\begin{aligned}\bar{c}_k &= ||\mathbf{d}_k||^2 + (c_{kk}-1)^2||\boldsymbol{\omega}_{k,i}||^2 + ||\mathbf{h}_k||^2 \\ &\quad + 2\mathbf{d}_k^T\mathbf{h}_k + (c_{kk}-1)[\mathbf{d}_k^T\boldsymbol{\omega}_{k,i} \\ &\quad + \mathbf{h}_k^T\boldsymbol{\omega}_{k,i} + \boldsymbol{\omega}_{k,i}^T(\mathbf{d}_k + \mathbf{h}_k)]. \end{aligned} \tag{14}$$

Therefore, if $\bar{a}_k \geq 0$, the minimum of a scalar quadratic function of (11) is equal to

$$\gamma_{\text{opt}} = -\frac{\bar{b}_k}{2\bar{a}_k}. \tag{15}$$

Since the leaky factor is a positive relatively small constant, in practice, a fixed γ_0 is used if $\gamma_{\text{opt}} < 0$ and a maximum $\gamma_{\max}$ if $\gamma_{\text{opt}} > \gamma_{\max}$.

For the minimum disturbance, replace (15) in (11). Then, the minimum disturbance equals to

$$\mathrm{D}_{\min} = \frac{\bar{b}_k^2}{2\bar{a}_k} + \bar{c}_k. \tag{16}$$

B. HMM DECORRELATION

In [23], an AR model is used for correlation model of input signal which is actually the regression vector with respect to time index i. What meant of correlation of the input signal is that successive samples of the input signal are correlated random variables. At least, two consecutive samples are correlated. The process of decorrelation is relatively complex and need a two decorrelated filters for decorrelating the input signal and decorrelating the output of the measurements. So, in this paper, a first order HMM model for modeling the correlation of input signal is used and then use a low complex one tap decorrelating filter instead. In fact, only correlation between two consecutive samples of the input signal is considered.

At first, the adaptation step of decorrelated diffusion leaky LMS is written as

$$\boldsymbol{\phi}_{k,i} = (1-\mu\gamma_{k,i})\boldsymbol{\omega}_{k,i} + \mu\sum_{l\in\mathcal{N}_k} a_{l,k}\bar{\mathbf{u}}_{l,i}^T(\bar{d}_{l,i} - \bar{\mathbf{u}}_{l,i}^T\boldsymbol{\omega}_{k,i}), \tag{17}$$

where $\bar{d}_{k,i}$ is the decorrelated version of $d_{k,i}$, and $\bar{\mathbf{u}}_{k,i}$ is the decorrelated version of $\mathbf{u}_{k,i}$. The block diagram of decorrelated leaky LMS used in the diffusion algorithm is shown in Fig. 1. The type of decorrelation used in this scheme is decorrelation using HMM model which is denoted by HMM-decorrelation and explained in the sequel.

Secondly the first order two state HMM model [30], [31] is introduced. The trellis diagram and state diagram of this HMM model are shown in Fig. 2 and Fig. 3. In fact, there may be used higher order HMM models or HMM models with more than two states which is quite versatile in real-world applications such as in speech processing. But, the objective of the presented work is to suggest a decorrelation step as simple as possible. So, a two-state first order HMM model is used for correlation in which only correlation between successive samples is considered. If the index k in $\mathbf{u}_{k,i}$ is

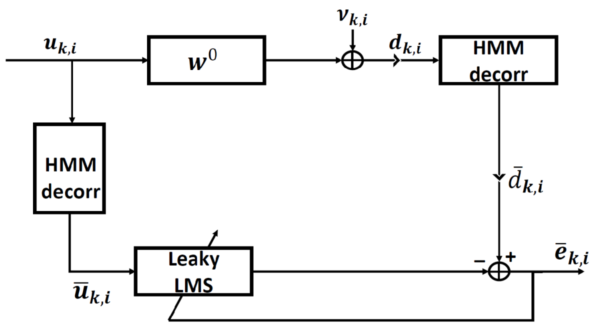

FIGURE 2. **Block diagram of HMM-decorrelated leaky LMS used in the diffusion algorithm.**

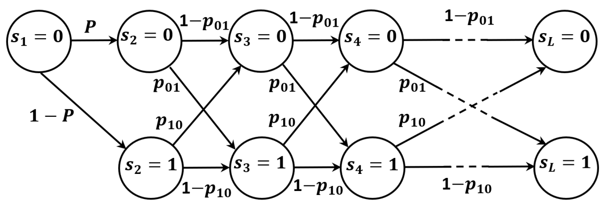

FIGURE 3. **Trellis diagram of the two state Marvov model.**

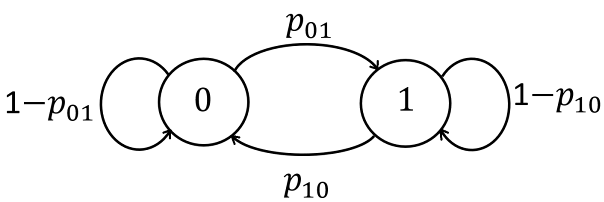

FIGURE 4. **State diagram of the two state Markov model.**

removed for simplicity and assuming the stream of input signal as u_i, the input signal occurs at two states. Then, for distribution of u_i, we have

$$\begin{cases} u_i \sim N(0, \sigma_0^2) & s_i = 0, \\ u_i \sim N(0, \sigma_1^2) & s_i = 1, \end{cases} \tag{18}$$

where σ_1^2 is the variance of state 1, and σ_0^2 is the variance in state 0. Also, for the correlation coefficient between successive samples in each state, assume a fixed correlation coefficient $r = \frac{\mathrm{E}(u_i u_{i+1})}{\sigma_i \sigma_{i+1}}$, where σ_i is the standard deviation of sample u_i. The transition probabilities between two states are defined as $p_{rk} = p\{s_{i+1} = 1 | s_i = r\}$ for $r, k = 0, 1$. The probability of a state is defined as $p_r = p\{s_i = r\}$. Then, the state probability vector is defined as $\Pi^i = [p_0, 1 - p_0]^T$ and is obtained from $\Pi^i = \Pi^i \mathbf{P}$, where $\mathbf{P} = [p_{kr}]$ is the transition probability matrix. To decorrelate an input signal which has a first order two state Markov model, a variable one tap filter is used. So, if the decorrelated output of the filter is $z_i = \bar{u}_i$, then, we have $z_i = u_i - a_i u_{i-1}$ and $z_{i+1} = u_{i+1} - a_{i+1} u_i$. To have uncorrelatedness between z_i and z_{i+1}, then $\mathrm{E}\{z_i z_{i+1}\} = 0$. We can write:

$$\begin{aligned} \mathrm{E}\{z_i z_{i+1}\} &= \mathrm{E}\{(u_i - a_i u_{i-1})(u_{i+1} - a_{i+1} u_i)\} \\ &= \mathrm{E}(u_i u_{i+1}) - \mathrm{E}\{a_{i+1} u_i^2\} - \mathrm{E}\{a_i u_{i-1} u_{i+1}\} \\ &\quad + \mathrm{E}\{a_i a_{i+1} u_i u_{i-1}\} \end{aligned} \tag{19}$$

Since the uncorrelatedness between u_{i-1} and u_{i+1} is assumed, from (19), the following formula is achieved by assuming the decorrelation condition:

$$\mathrm{E}(u_i u_{i+1}) - a_{i+1} \mathrm{E}(u_i^2) + a_i a_{i+1} \mathrm{E}(u_i u_{i-1}). \tag{20}$$

Now, $\mathrm{E}(u_i u_{i+1})$ is calculated as

$$\begin{aligned} &p\{s_i = 0\} \mathrm{E}\{u_i u_{i+1} | s_i = 0\} + p\{s_i = 1\} \mathrm{E}\{u_i u_{i+1} | s_i = 1\} \\ &\quad = p_0 \mathrm{E}\{u_i u_{i+1} | s_i = 0\} + (1 - p_0) \mathrm{E}\{u_i u_{i+1} | s_i = 1\}. \end{aligned} \tag{21}$$

To calculate $\mathrm{E}\{u_i u_{i+1} | s_i = 0\}$, it can be written as

$$p_0 \mathrm{E}(u_i u_{i+1} | s_i s_{i+1} = 00) + (1 - p_0) \mathrm{E}(u_i u_{i+1} | s_i s_{i+1} = 01). \tag{22}$$

Then, from (21) and (22), we can reach to

$$\begin{aligned} \mathrm{E}(u_i u_{i+1}) = p_0 \Big[&p_0 \mathrm{E}(u_i u_{i+1} | s_i s_{i+1} = 00) \\ &+ \bar{p}_0 \mathrm{E}(u_i u_{i+1} | s_i s_{i+1} = 01) \Big] \\ &+ \bar{p}_0 \Big[p_0 \mathrm{E}(u_i u_{i+1} | s_i s_{i+1} = 10) \\ &+ \bar{p}_0 \mathrm{E}(u_i u_{i+1} | s_i s_{i+1} = 11) \Big], \end{aligned} \tag{23}$$

where $\bar{p}_0 = 1 - p_0$. Since $\mathrm{E}(u_i u_{i+1} | s_i s_{i+1} = 00) = r\sigma_0^2$, $\mathrm{E}(u_i u_{i+1} | s_i s_{i+1} = 01) = r\sigma_1\sigma_0$, $\mathrm{E}(u_i u_{i+1} | s_i s_{i+1} = 10) = r\sigma_1\sigma_0$, and $\mathrm{E}(u_i u_{i+1} | s_i s_{i+1} = 11) = r\sigma_1^2$, then, (23) can be written as:

$$\mathrm{E}(u_i u_{i+1}) = r p_0^2 \sigma_0^2 + 2 p_0 \bar{p}_0 r \sigma_1 \sigma_0 + r \bar{p}_o^2 \sigma_1^2 \triangleq A. \tag{24}$$

Also, in (20), we have

$$\mathrm{E}(u_i^2) = p_0 \sigma_0^2 + \bar{p}_0 \sigma_1^2 \triangleq B. \tag{25}$$

Also, similarly, following the stationarity of the input signal, we have $\mathrm{E}(u_i u_{i-1}) = A$. So, the decorrelation condition in (20) reduces to

$$A - a_{i+1} B + a_i a_{i+1} A = 0, \tag{26}$$

which results to a recursive equation for updating the time varying filter coefficient a_i as

$$a_{i+1} = \frac{A}{B - A a_i}. \tag{27}$$

Therefore, the final low complexity one tap decorrelation filter is as

$$z_i = \bar{u}_i = u_i - a_i u_{i-1}, \tag{28}$$

where a_i is obtained via (27).

The one-tap decorrelator used in simulations is depicted in Fig. 1.

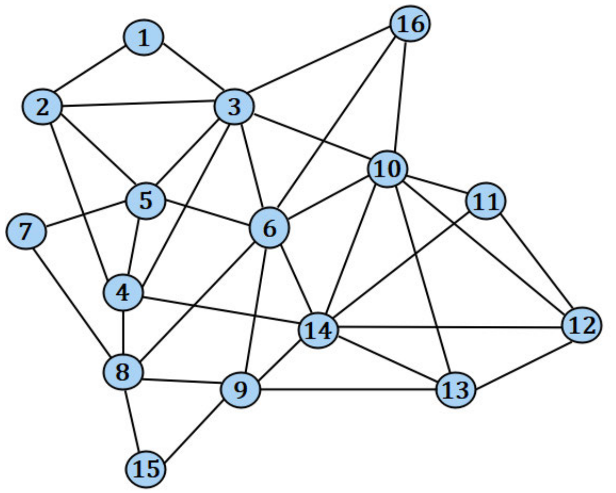

FIGURE 5. **The network used in the simulations.**

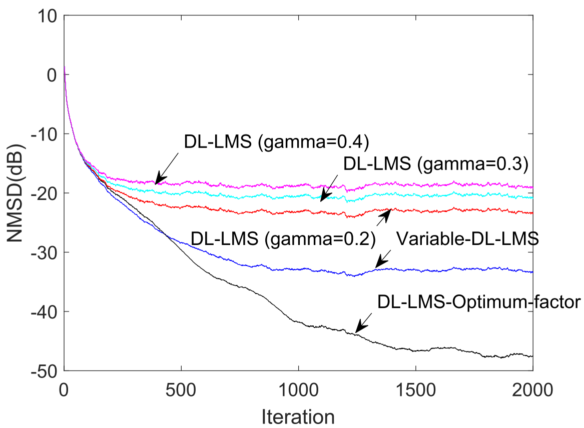

FIGURE 6. **NMSD versus iteration number for conventional and optimized DL-LMS algorithm in comparison to variable DL-LMS (without decorrelation).**

V. SIMULATION RESULTS

In this section, the performance of the proposed HMM-Decorrelated Diffusion Leaky LMS (HMM-Decorrelated-DL-LMS) is investigated. The two versions of the proposed algorithm which are conventional (with constant leaky factor) and optimized (with optimized leaky factor) are implemented and discussed. Three experiments are performed. In the first experiment, the DL-LMS with optimized leaky factor is implemented and compared to conventional DL-LMS with constant leaky factor and variable DL-LMS [20]. In the second experiment, the HMM-Decorrelated DL-LMS is examined and compared with other algorithms. In the third experiment, the proposed HMM-decorrelated DL-LMS algorithm is compared with AR-decorrelated DNLMS.

The network has $N = 16$ sensors and the topology is the same as was used in [32] and depicted in Fig. 4. In all experiments, the input signals which in fact are the regression vectors are derived from the two state first order HMM model described in Section. IV-B. The parameters are selected as

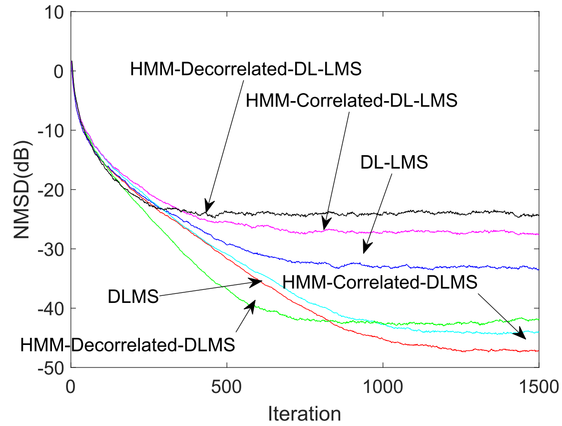

FIGURE 7. **MSD versus iteration number for performance comparison of proposed HMM-decorrelated DL-LMS algorithm in comparison to other algorithms.**

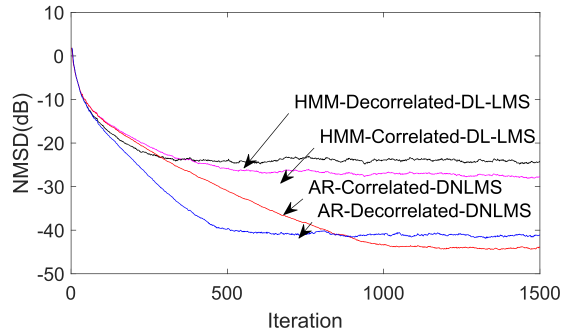

FIGURE 8. **MSD versus iteration number for performance comparison of proposed HMM decorrelation without optimized leaky factor and an AR decorrelation.**

$p_{01} = p_{10} = 0.2$, $p_{00} = p_{11} = 0.8$, $r = 0.5$, $\sigma_1 = 1$, and $\sigma_0 = 0.2$. The measurement noise is assumed to be White Gaussian Noise (WGN) with zero mean and variance equal to $\sigma_v^2 = 0.05$. The step sizes are all selected as $\mu = 0.02$. The performance metric is defined as Normalized Mean Square Deviation (NMSD), which is defined as

$$\text{NMSD(dB)} = 20 \log_{10}\left(\frac{||\boldsymbol{\omega} - \boldsymbol{\omega}_o||_2}{||\boldsymbol{\omega}^o||_2}\right), \tag{29}$$

which is averaged over $R = 50$ Monte Carlo simulation runs with random noise, random input signals, and random parameter vector. The size of parameter vector is set as $L = 50$. The unknown parameter vector $\boldsymbol{\omega}^o$ is drawn from a WGN with zero mean and unit variance. For simulating the optimized leaky factor, It is assumed that $\gamma_0 = 0.2$ and $\gamma_{max} = 0.5$.

In the first experiment, the DL-LMS with uncorrelated input signal and with different leaky factors equal to $\gamma = 0.4$, $\gamma = 0.3$, and $\gamma = 0.2$ is simulated. Also, the DL-LMS with optimized leaky factor and the variable DL-LMS [20] is simulated as the competing algorithm. The NMSD versus iteration number is depicted in Fig. 5. It demonstrates the superiority of the proposed DL-LMS with optimized leaky factor over other algorithms in terms of achieving less final

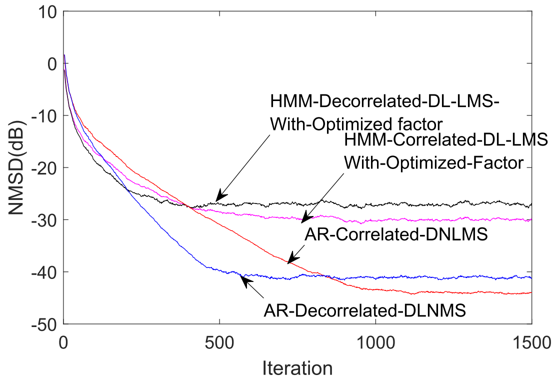

FIGURE 9. MSD versus iteration number for performance comparison of proposed HMM decorrelation with optimized leaky factor and an AR decorrelation.

MSD. In fact, the optimized leaky factor has about 10dB less final NMSD in comparison to the case of $\gamma = 0.2$.

In the second experiment, the performance of HMM-Decorrelated algorithms in presence of correlated input signal are investigated. Hence, in this experiment, the decorrelation modification to the algorithm is only considered. In this experiment, the simulated algorithms are the proposed HMM-Decorrelated DL-LMS, HMM-Decorrelated DLMS, classical DLMS with HMM correlation (nominated as HMM-correlated DLMS), classical DL-LMS with HMM correlation with $\gamma = 0.2$ (nominated as HMM-correlated DL-LMS), classical DLMS with no correlation (nominated as DLMS), and classical DL-DLMS with no correlation with $\gamma = 0.2$. The NMSD versus iteration number is shown in Fig. 6. Some points are deduced. First, the DLMS has better performance than DL-LMS in terms of final NMSD. Second, The HMM-decorrelation speeds up the convergence both in DLMS and DL-LMS with price of higher final NMSD. Third, HMM decorrelated DL-DLMS is better than HMM correlated DL-DLMS around 5dB.

In the third experiment, the HMM-decorrelated DL-LMS is compared with AR-decorrelated NLMS and also by versions that do not use decorrelations in the case of not using the optimized leaky factor. The NMSD versus iteration number is depicted in Fig. 7. It demonstrates that AR-decorrelation NLMS algorithm has better performance of 10dB, but with higher complexity of using a decorrelation filter of length $M_{dec} = 5$ instead of $M_{dec} = 1$ used in the proposed algorithm.

In Fig. 5, only modification of using the optimized leaky factor is used. In Figs. 6 and 7, only modification of decorrelation in the proposed algorithm is used. Now, in the final sketch of Fig. 8, both modifications of decorrelation and using optimized leaky factor are considered. The Fig. 8 shows a low improvement when both using decorrelation and optimized leaky factor in comparison to when using decorrelation without optimizing leaky factor. This improvement is around 2-3dB less final MSD. It shows that using the optimized leaky factor has low influence on the performance when there are correlation in the input signal.

VI. CONCLUSION AND FUTURE WORK

In this paper, the main concentration was on diffusion leaky LMS algorithm as a distributed estimation algorithm with good numerical stability. This algorithm has a leaky factor that controls the speed of convergence and final achieved MSD. The leakage factor is optimally calculated based on minimizing the disturbance of the algorithm in one iteration. Thanks to quadratic nature of disturbance the optimized leaky factor is calculated in closed form. In addition, the performance of this algorithm is deteriorated in terms of convergence speed when the input signal is correlated in time. The existing decorrelation methods are computationally complex specially when intended to be used in a sensor network. So, a statistical HMM model for correlation in time is presented and it is used in a low complex one tap filter for decorrelation. Again, the coefficient of the one-tap filter is calculated in closed form. Moreover, simulation results show the effectiveness of optimized DL-LMS algorithm in comparison to variable DL-LMS algorithm and also show the effectiveness of speeding up the convergence when using the HMM-decorrelated DL-LMS algorithm in comparison to not using the decorrelation step when there exists HMM correlation. For the future work, to speed up the convergence rate further, one can combine the leaky LMS with proportionate algorithms. Also, a prediction filter introduced in [33], could be used for the decorrelation step.

REFERENCES

[1] A. H. Sayed, "Adaptation, learning and optimization over networks," *Found. Trends Mach. Learn.*, vol. 7, nos. 4–5, pp. 311–801, 2014.

[2] S.-Y. Tu and A. H. Sayed, "Diffusion strategies outperform consensus strategies for distributed estimation over adaptive networks," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6217–6234, Dec. 2012.

[3] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3122–3136, Jul. 2008.

[4] Y. Liu, C. Li, and Z. Zhang, "Diffusion sparse least-mean squares over networks," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4480–4485, Aug. 2012.

[5] F. Wen, "Diffusion least-mean P-power algorithms for distributed estimation in alpha-stable noise environments," *Electron. Lett.*, vol. 49, no. 21, pp. 1355–1356, Oct. 2013.

[6] L. Lu, H. Zhao, W. Wang, and Y. Yu, "Performance analysis of the robust diffusion normalized least mean p-power algorithm," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 12, pp. 2047–2051, Dec. 2018.

[7] M. Korki and H. Zayyani, "Weighted diffusion continuous mixed p-norm algorithm for distributed estimation in non-uniform noise environment," *Signal Process.*, vol. 164, pp. 225–233, Nov. 2019.

[8] W. Ma, B. Chen, J. Duan, and H. Zhao, "Diffusion maximum correntropy criterion algorithms for robust distributed estimation," *Digit. Signal Process.*, vol. 58, pp. 10–19, Nov. 2016.

[9] Y. He, F. Wang, S. Wang, P. Ren, and B. Chen, "Maximum total correntropy diffusion adaptation over networks with noisy links," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 2, pp. 307–311, Feb. 2019.

[10] V. C. Gogineni, S. P. Talebi, S. Werner, and D. P. Mandic, "Fractional-order correntropy adaptive filters for distributed processing of α-stable signals," *IEEE Signal Process. Lett.*, vol. 27, pp. 1884–1888, 2020.

[11] H. Zayyani, "Robust minimum disturbance diffusion LMS for distributed estimation," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 68, no. 1, pp. 521–525, Jan. 2021.

[12] H. Zayyani and A. Javaheri, "A robust generalized proportionate diffusion LMS algorithm for distributed estimation," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 68, no. 4, pp. 1552–1556, Apr. 2021.

[13] A. Nakai-Kasai and K. Hayashi, "Diffusion LMS based on message passing algorithm," *IEEE Access*, vol. 7, pp. 47022–47033, 2019.
[14] J.-W. Lee, J.-T. Kong, W.-J. Song, and S.-E. Kim, "Data-reserved periodic diffusion LMS with low communication cost over networks," *IEEE Access*, vol. 6, pp. 54636–54650, 2018.
[15] W. Gao and J. Chen, "Performance analysis of diffusion LMS for cyclostationary white non-Gaussian inputs," *IEEE Access*, vol. 7, pp. 91243–91252, 2019.
[16] W. Huang, L. Li, Q. Li, and X. Yao, "Diffusion robust variable step-size LMS algorithm over distributed networks," *IEEE Access*, vol. 6, pp. 47511–47520, 2018.
[17] A. M. Wilson, T. Panigrahi, B. P. Mishra, and S. L. Sabat, "Adaptive geman-McClure estimator for robust distributed channel estimation," *IEEE Access*, vol. 9, pp. 93691–93702, 2021.
[18] A. N. Sadigh and H. Zayyani, "A proportionate robust diffusion recursive least exponential hyperbolic cosine algorithm for distributed estimation," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 4, pp. 2381–2385, Apr. 2022.
[19] K. Mayyas and T. Aboulnasr, "Leaky LMS algorithm: MSE analysis for Gaussian data," *IEEE Trans. Signal Process.*, vol. 45, no. 4, pp. 927–934, Apr. 1997.
[20] L. Lu and H. Zhao, "Diffusion leaky LMS algorithm: Analysis and implementation," *Signal Process.*, vol. 140, pp. 77–86, Nov. 2017.
[21] M. S. Salman, "Sparse leaky-LMS algorithm for system identification and its convergence analysis," *Int. J. Adapt. Control Signal Process.*, vol. 28, no. 10, pp. 1065–1072, Oct. 2014.
[22] L. Shi and H. Zhao, "Diffusion leaky zero attracting least mean square algorithm and its performance analysis," *IEEE Access*, vol. 6, pp. 56911–56923, 2018.
[23] S. Zhang, H. C. So, W. Mi, and H. Han, "A family of adaptive decorrelation NLMS algorithms and its diffusion version over adaptive networks," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 2, pp. 638–649, Feb. 2018.
[24] F. Huang, S. Zhang, J. Zhang, H. Chen, and A. H. Sayed, "Diffusion Bayesian decorrelation algorithms over networks," *IEEE Trans. Signal Process.*, vol. 71, pp. 571–586, 2023.
[25] B.-J. Yoon and P. P. Vaidyanathan, "Context-sensitive hidden Markov models for modeling long-range dependencies in symbol sequences," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4169–4184, Nov. 2006.
[26] M. Ostendorf, V. V. Digalakis, and O. A. Kimball, "From HMM's to segment models: A unified view of stochastic modeling for speech recognition," *IEEE Trans. Speech Audio Process.*, vol. 4, no. 5, pp. 360–378, 1996.
[27] H. Sameti, H. Sheikhzadeh, L. Deng, and R. L. Brennan, "HMM-based strategies for enhancement of speech signals embedded in nonstationary noise," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 5, pp. 445–455, Sep. 1998.
[28] Y. Li, B. Hu, T. Niu, S. Gao, J. Yan, K. Xie, and Z. Ren, "GMM-HMM-based medium- and long-term multi-wind farm correlated power output time series generation method," *IEEE Access*, vol. 9, pp. 90255–90267, 2021.
[29] S. O. Haykin, *Adaptive Filter Theory*, 5th ed. Pearson, 2013.
[30] Z. Habibi and H. Zayyani, "Markovian adaptive filtering algorithm for block-sparse system identification," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 68, no. 8, pp. 3032–3036, Aug. 2021.
[31] Z. Habibi, H. Zayyani, and M. Korki, "A robust Markovian block sparse adaptive algorithm with its convergence analysis," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 71, no. 3, pp. 1546–1550, 2024.
[32] H. Zayyani, F. Oruji, and I. Fijalkow, "An adversary-resilient doubly compressed diffusion LMS algorithm for distributed estimation," *Circuits, Syst., Signal Process.*, vol. 41, no. 11, pp. 6182–6205, Nov. 2022.
[33] K. Vasudevan, *Digital Communications and Signal Processing*, 2nd ed. India: Universities Press (India) Private Limited, 2010.

HADI ZAYYANI (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering and in the communication field from the Sharif University of Technology, in 2000, 2003, and 2010, respectively. Since 2012, he has been an Assistant Professor with Qom University of Technology (QUT), where he has been an Associate Professor, since 2019. He is selected as top 2 percent scientists from Stanford University, in 2021 and 2022, respectively. He has published around 90 articles which around half of them are in IEEE venue. He is also an Editor of *Iranian Journal of Science and Technology* and a reviewer of high impact IEEE and Elsevier journals. His research interests include statistical signal processing, sparse signal processing, adaptive filter theory, radar signal processing, graph signal processing, and distributed signal processing.

MOHAMMAD SALMAN (Senior Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electrical and electronic engineering from Eastern Mediterranean University (EMU), in 2006, 2007, and 2011, respectively. He is currently an Associate Professor with the Electrical Engineering Department, American University of the Middle East, Kuwait. He has over 70 peer reviewed journals and conference publications. He has served as a guest editor, the general chair, the program chair, and a TPC member for many international journals and conferences. His research interests include signal processing, adaptive filters, image processing, and sparse signal representation.

MEHDI KORKI (Member, IEEE) received the Ph.D. degree from the School of Software and Electrical Engineering, Swinburne University of Technology, Melbourne, Australia, in 2016. He is currently a Lecturer with the Swinburne University of Technology. His main research interests include statistical signal processing, machine learning, power line communication (PLC), source localization, and smart grid communications.

AHMED A. F. YOUSSEF (Member, IEEE) received the Ph.D. degree in electrical engineering in Canada, in 2010. He is currently an Associate Professor. His research interests include wireless communications, digital signal processing, optimization, and statistical signal processing.