Abstract—We study the effect of fading in the communication channels between nodes on the performance of the incremental least mean square (ILMS) algorithm. We derive steady-state performance metrics, including the mean-square deviation (MSD), excess mean-square error (EMSE), and mean-square error (MSE). We obtain the sufficient conditions to ensure mean-square convergence, and verify our results through simulations. Simulation results show that our theoretical analysis closely matches the actual steady state performance.

Index Terms—Adaptive networks, distributed estimation, incremental least mean square

I. INTRODUCTION

There are already several useful strategies for estimation over distributed networks. An example of a distributed method is the consensus strategy [1]–[3] in which each node performs a local estimation and fuses its estimate with those of its neighbors so that all nodes converge to the same estimate as the number of iterations increases. The main problem of mentioned methods is that the framework does not allow the network to undertake a continuous learning and optimization which motivated the development of adaptive networks. An adaptive network is a collection of agents (nodes) that collaborate with each other through in-network local processing rules in order to estimate and track parameters of interest [5]. Two major classes of adaptive networks are incremental strategy [6]–[9] and diffusion strategy [7], [11], [12]. In the incremental mode, nodes passing updates to each other in a Hamiltonian cycle in the network, while in diffusion mode, each node can communicate with a subset of neighboring nodes.

In this paper, we investigate the performance of the ILMS algorithm in a wireless adaptive network with communication links between neighboring nodes modeled as fading channels. In the original ILMS strategy proposed by [13], it is assumed that nodes communicate with each other via ideal links which is typically not true in practice. In [14]–[18], the effect of additive link noise in the communication channels between nodes have been investigated. In [19], the performance of general adaptive diffusion algorithms in the presence of imperfect information exchanges, including quantization errors, and model non-stationarities has been considered. All these works however cannot be directly applied to a wireless adaptive network, whose communication links are modeled by fading channels [20]. The references [21], [22] propose diffusion LMS algorithms for wireless sensor networks with fading channels but under the assumption that channel state information is known so that channel equalization can be performed.

In this paper, our objective is to investigate how the performance of the ILMS algorithm is affected by the channel fading statistics of the communication channels between nodes. We show that the ILMS algorithm over fading channels, like the traditional ILMS algorithm, is stable in mean if the step size is chosen in an appropriate range. This range now not only depends on the regression correlation, but also on the mean of the channel fading gains. Moreover, our analysis reveals that in general, fading communication channels lead to biased estimates in steady state. We derive closed-form expressions for the steady state performance metrics, including the mean-square deviation (MSD), excess mean-square error (EMSE) and mean-square error (MSE), of the ILMS algorithm in the presence of fading channels, and under a Gaussian model. We show explicitly how these metrics are affected by the fading channel statistics. We derive sufficient conditions for the convergence of the MSD, EMSE and MSE, and show that for a fixed step size, mean-square stability is lost if the channel gain variances become large. We present simulations to verify that our theoretical analysis closely matches the actual steady state performance.

Notation: We adopt boldface letters for random quantities. The symbol * denotes conjugation for scalars and Hermitian transpose for matrices. The notation diag{·} will be used in two ways: $X = \text{diag}\{x\}$ is a diagonal matrix whose entries are those of the vector $x$, and $x = \text{diag}\{X\}$ is a vector containing the main diagonal of $X$. If $\Sigma$ is a matrix, we use the notation $\|x\|^2_{\Sigma} = x^*\Sigma x$ for the weighted square norm of $x$. If $\sigma$ is a vector, the notation $\|x\|_{\sigma}$ is used to represent $\|x\|_{\text{diag}\{\sigma\}}$.

II. ILMS ALGORITHM OVER FADING CHANNELS

Consider a network composed of $N$ nodes. At time $i$, node $k$ observes a scalar measurement $d_k(i)$ and a $1 \times M$ regression vector $u_{k,i}$, which are related via a linear regression model

$$d_k(i) = u_{k,i}w^o + v_k(i),$$

(1)

where $v_k(i)$ is the observation (or measurement) noise, and the $M \times 1$ vector $w^o$ is an unknown vector. The goal of the network is to estimate $w^o$, at every node $k$, using all observed data in the entire network. In the incremental LMS algorithm [6], [13],

A. Khalili and A. Rastegarnia are with Department of Electrical Engineering, Malayer University, Malayer 65719-95863, Iran e-mails: (rastegar, tinati, a-khalili@tabrizu.ac.ir).
each node $k$ receives a local estimate from the previous node $k-1$, updates it using its local data, and then sends it to the next node $k+1$. The update equations for the ILMS algorithm, at iteration $i$ is given by [6]

$$
\begin{align}
\begin{cases}
    w_{0,i} &\leftarrow w_{N,i-1} \\
    w_{k,i} &= w_{k-1,i} + \mu_k u_{k,i}^* (d_k(i) - u_{k,i} w_{k-1,i})
\end{cases}
\end{align}
$$

where $w_{k,i}$ is the local estimate of the node $k$ at time $i$. It is shown in [6] that $w_{k,i}$ converges to $w^o$ as $i \to \infty$ for every node $k$.

Now, we consider the case where nodes communicate over fading channels. By incorporating the impact of fading channels, the update equation for node $k$ in [3] becomes

$$
\begin{align}
    w_{k,i} &= h_k(i) w_{k-1,i} + q_{k,i} \\
    &\quad+ \mu_k u_{k,i}^* (d_k(i) - u_{k,i} (h_k(i) w_{k-1,i} + q_{k,i}))
\end{align}
$$

where $h_k(i)$ is the channel gain at time $i$ for the communication channel between node $k-1$ and $k$, and $q_{k,i}$ is the zero mean channel noise with covariance matrix $Q_k = \mathbb{E}[q_{k,i} q_{k,i}^*]$. We assume phase coherent reception at every node $k$ and model the channel gain as a non-negative random variable. We make the following assumptions regarding the fading channel statistics.

**Assumption 1:**
1) The channel gains $h_k(i)$ for all nodes $k = 1, \ldots, N$, and all observation times $i \geq 1$, are independent of each other. For each node $k$, the channel gains $\{h_k(i) : i \geq 1\}$ are identically distributed.
2) The channel gains $h_k(i)$ for all nodes $k = 1, \ldots, N$, and all observation times $i \geq 1$, are independent of $(d_j(i), u_{j,i})$ for all $l$ and $j$.

Let $m_k = \mathbb{E}[h_k(i)]$ and $s_k = \mathbb{E}[h_k^2(i)]$ be the mean and second order moment of the channel gain for node $k$, respectively. We note that in practice, even if the nodes perform channel state estimation, it is not possible to measure the channel gains with absolute certainty, especially if the channels experience fast fading. In this case, without loss of generality, the channel gain for node $k$ can still be modeled as a non-negative random variable with a non-trivial variance.

### III. Performance Analysis

In this section, we analyze the mean stability and steady state mean-square performance of the ILMS algorithm when communication channels between nodes are fading channels. Our analysis is based on the energy conservation approach of [6]. We make the following assumptions regarding the data model in [1]. These assumptions are commonly assumed in the literature [5], [13].

**Assumption 2:**
1) The regression vectors $u_{k,i}$ are independent over node indices $k$ and observation times $i$.
2) The measurement noises $v_k(i)$ are independent of each other and the regression vectors $u_{k,i}$.

In our analysis, we will use the deviation between an observed measurement and its prediction based on the current local estimate, which is defined as

$$ e_k(i) = d_k(i) - u_{k,i} w_{k-1,i}, $$

and the weight error vector, which is the deviation between the local estimate $w_{k,i}$ and its true value $w^o$, given by

$$ \bar{w}_{k,i} = w^o - w_{k,i}. $$

By subtracting $w^o$ from both sides of [3] and using the definition of $\bar{w}_{k,i}$, we obtain

$$
\begin{align}
    \bar{w}_{k,i} &= \bar{h}_k(i) \bar{w}_{k-1,i} \\
    &\quad+ (1 - \bar{h}_k(i)) w^o - \mu_k \bar{h}_k(i) u_{k,i}^* \bar{w}_{k-1,i} \\
    &\quad- \mu_k u_{k,i}^* v_k(i) - \mu_k (1 - \bar{h}_k(i)) u_{k,i}^* u_{k,i} w^o - q_{k,i} \\
    &\quad+ \mu_k u_{k,i}^* u_{k,i} q_{k,i}.
\end{align}
$$

In the steady-state analysis, we are interested to quantify the performance using the following metrics at every node $k$:

$$
\begin{align}
    \eta_k &\triangleq \lim_{i \to \infty} \mathbb{E}[||\bar{w}_{k-1,i}||^2] \quad \text{(MSD)} \\
    \zeta_k &\triangleq \lim_{i \to \infty} \mathbb{E}[||\bar{w}_{k-1,i}||^2 R_{u,k}] \quad \text{(EMSE)} \\
    \xi_k &\triangleq \lim_{i \to \infty} \mathbb{E}[||e_k(i)||^2] = \zeta_k + \sigma^2_{n,k} \quad \text{(MSE)}
\end{align}
$$

To derive the above steady state performance metrics, we need to evaluate quantities of the form $\mathbb{E}[||\bar{w}_{k,i}||^2_{\Sigma_k}]$ where $\Sigma_k$ is a positive semi-definite Hermitian matrix. To this end, we consider the weight vector update equation given by [4]. Let

$$ C_{k,i} = m_k J_k C_{k-1,i} + (1 - m_k) J_k, $$

where $J_k = I - \mu_k R_{u,k}$, $C_{0,i} = C_{N,i-1}$, and $C_{0,1} = I$. Note that the matrix $C_{k,i}$ is such that

$$
\mathbb{E}[\bar{w}_{k,i}] = C_{k,i} w^o.
$$

By equating the weighted norm of both sides of [4], taking expectations and using Assumptions 2 and 1 and 2, we obtain the following recursive relationship:

$$
\begin{align}
    \mathbb{E}[||\bar{w}_{k,i}||^2_{\Sigma_k}] &= \mathbb{E}[||\bar{w}_{k-1,i}||^2_{\Sigma_{k-1}} + \mu_k^2 \sigma^2_{n,k} \mathbb{E}[||u_{k,i}||^2_{\Sigma_k}]] \\
    &\quad+ \mathbb{E}[||q_{k,i}||^2_{G_k}] + ||w^o||^2_{T_k + H_{k,i}},
\end{align}
$$

where

$$
\begin{align}
    G_k &= \Sigma_k - \mu_k \mathbb{E}[\Sigma_k u_{k,i} u_{k,i}^* + u_{k,i}^* u_{k,i} \Sigma_k] \\
    &\quad+ \mu_k \mathbb{E}[||u_{k,i}||^2_{\Sigma_k} u_{k,i}^* u_{k,i}] \\
    \Sigma_k &= s_k G_k \\
    T_k &= (1 - 2m_k + s_k) G_k \\
    H_{k,i} &= (m_k - s_k) (C_{k-1,i} G_k + G_k C_{k-1,i}).
\end{align}
$$

In order to compute all the moments that appear in the recursive equation (10) and to obtain closed-form expressions, we now make the following assumption regarding the regression vectors $u_{k,i}$ for all nodes $k = 1, \ldots, N$.

**Assumption 3:** For each $k = 1, \ldots, N$, the distribution of $u_{k,i}$ is a Gaussian distribution with $R_{u,k} = U_k \Lambda_k U_k^*$, where $\Lambda_k$ is a diagonal matrix with diagonal elements being the eigenvalues of the correlation matrix $R_{u,k}$, and $U_k$ is unitary matrix.
Making use of Assumption 3, we further define the following transformed quantities:
\[
\begin{align*}
\mathbf{w}_{k,i}^* &= U_k^* \mathbf{w}_{k,i}, \\
\tilde{\Sigma}_k &= U_k^* \tilde{\gamma}_k U_k, \\
\tilde{T}_k &= U_k^* T_u U_k, \\
\tilde{H}_{k,i} &= U_k^* H_{k,i} U_k, \\
\tilde{\bar{v}}^o &= U_k^* \bar{v}^o, \\
\tilde{Q}_k &= U_k^* Q_k U_k, \\
D &= \tilde{\bar{v}}^o \tilde{\bar{v}}^{o*} = \bar{v}^o \bar{v}^{o*}.
\end{align*}
\]

From the above definitions, equation (10) can now be rewritten in the following equivalent form
\[
\begin{align*}
\mathbb{E} \left[ \left\| \mathbf{w}_{k,i} \right\|_{\Sigma_k}^2 \right] &= \mathbb{E} \left[ \left\| \mathbf{w}_{k-1,i} \right\|_{\Sigma_k}^2 \right] + \mu_k^2 \mathbb{E} \left[ \left\| \mathbf{u}_{k,i} \right\|_{\Sigma_k}^2 \right] \\
&\quad+ \mathbb{E} \left[ \left\| \mathbf{q}_{k,i} \right\|_{\tilde{\Sigma}_k}^2 \right] + \left\| \tilde{\bar{v}}^o \right\|^2_{\tilde{T}_k + \tilde{H}_{k,i}}.
\end{align*}
\] (21)

where in (21) we have
\[
\begin{align*}
\tilde{G}_k &= \tilde{\Sigma}_k - \mu_k \mathbb{E} \left[ \tilde{\Sigma}_k \tilde{a}_{k,i}^* \mathbf{u}_{k,i} + \tilde{\Sigma}_k \tilde{a}_{k,i}^* \tilde{a}_{k,i} \tilde{\Sigma}_k \right] \\
&\quad+ \mu_k^2 \mathbb{E} \left[ \left\| \mathbf{u}_{k,i} \right\|_{\tilde{\Sigma}_k}^2 \right] \tilde{a}_{k,i}^* \mathbf{u}_{k,i} \tilde{\Sigma}_k \\
&\quad= \tilde{\Sigma}_k - \mu_k (\tilde{\Sigma}_k \tilde{\Lambda} + \tilde{\Lambda} \tilde{\Sigma}_k) \\
&\quad+ \mu_k^2 (\tilde{\Lambda} \Sigma_k \tilde{\Lambda} + \gamma \Sigma_k \tilde{\Sigma}_k) \\
\tilde{\Sigma}_k &= \sigma_k \tilde{G}_k \\
\tilde{T}_k &= (1 - 2m_k + s_k) \tilde{G}_k \\
\tilde{H}_{k,i} &= (m_k - s_k) (\tilde{C}_k - \tilde{\gamma}_k \tilde{G}_k + \tilde{G}_k \tilde{C}_k - 1, 1) \\
&= 2(m_k - s_k) \tilde{C}_k - \tilde{\gamma}_k \tilde{G}_k. \\
\end{align*}
\] (22)

Further algebraic manipulations of (21) yields
\[
\begin{align*}
\mathbb{E} \left[ \left\| \mathbf{w}_{k,i} \right\|_{\Sigma_k}^2 \right] &= \mathbb{E} \left[ \left\| \mathbf{w}_{k-1,i} \right\|_{\Sigma_k}^2 \right] + \mu_k^2 \sigma_k^2 \tilde{a}_{k,i}^* \tilde{a}_{k,i} \Sigma_k \\
&\quad+ \mathbb{E} \left[ \left\| \mathbf{q}_{k,i} \right\|_{\tilde{\Sigma}_k}^2 \right] + \left\| \tilde{\bar{v}}^o \right\|^2_{\tilde{T}_k + \tilde{H}_{k,i}}.
\end{align*}
\] (23)

To derive (5), we only need to consider the case where \(\tilde{\Sigma}_k\) is a diagonal matrix. In this case, matrix \(\tilde{\Sigma}_k^r\) is also a diagonal matrix. We let
\[
\begin{align*}
\tilde{\sigma}_k^r &\equiv \text{diag} \{\tilde{\Sigma}_k\}, \\
\tilde{\sigma}_k^r &\equiv \text{diag} \{\tilde{\Sigma}_k^r\}, \\
\lambda_k &\equiv \text{diag} \{\tilde{\Lambda}_k\},
\end{align*}
\] (24)

and
\[
\tilde{F}_k = I - \mu_k X_k + \mu_k^2 Y_k, \quad (25)
\]

with \(X_k = 2\Lambda_k + Y_k = \Lambda_k^2 + \mu_k \lambda_k^2 \tilde{T}\). The \(M \times M\) matrix \(\tilde{F}_k\) contains the statistics of data local to node \(k\). We then have
\[
\begin{align*}
\mathbb{E} \left[ \left\| \mathbf{w}_{k,i} \right\|_{\Sigma_k}^2 \right] &= \mathbb{E} \left[ \left\| \mathbf{w}_{k-1,i} \right\|_{\Sigma_k}^2 \right] + g_{k,i} \tilde{\sigma}_k, \\
\end{align*}
\] (26)

where \(g_{k,i}\) and \(\tilde{\sigma}_k^r\) are given respectively by
\[
\begin{align*}
g_{k,i} &= \mu_k^2 \sigma_k^2 \tilde{a}_{k,i}^* \tilde{a}_{k,i} + \text{diag} \{\tilde{Q}_k\} \tilde{F}_k \\
&\quad+ (1 - 2m_k + s_k) \text{diag} \{\tilde{D}\} \tilde{F}_k \\
&\quad+ 2(m_k - s_k) \text{diag} \{\tilde{D}\} \tilde{C}_k - 1, i \tilde{F}_k, \\
\tilde{\sigma}_k &= s_k \tilde{F}_k \tilde{\sigma}_k. \\
\end{align*}
\] (27)

We can use (28) to derive conditions that guarantee convergence in the mean-square sense for the ILMS algorithm with fading channels. Under assumptions 2 and 3, the ILMS algorithm over fading channels converges in the mean-square sense if the step sizes \(\mu_k\) are chosen to be sufficiently small so that the
\[
s_k \rho(\tilde{F}_k) < 1. \quad (30)
\]
are generated from Gaussian distributions with covariance matrix $Q_k = \sigma_{v,k}^2 I$, for $k = 1, \ldots, 20$. The values of $\sigma_{v,k}^2$ are shown in Figure 2. We generate the channel gains $h_{c,k}(i)$ using a Rayleigh distribution with $m_k = \sqrt{2/\pi}$ for all values of $k$. To obtain the steady-state values of MSD, EMSE and MSE, we run the ILM algorithm with 2000 iterations and average the last 200 samples. Finally, each steady-state value is obtained by averaging over 100 independent runs.

In Figure 2 we show the steady-state performance metrics MSD, EMSE and MSE as functions of the node index $k$ when the step size $\mu = 0.02$. We can see that the simulated results closely match the theoretical results.

V. CONCLUSION

In this paper, we have investigated the steady state performance of the ILMS algorithm when the links between nodes are fading channels, and we do not have perfect channel state information. Our analysis reveals how the mean-square stability depends on the channel gain variances. We also derived steady state performance metrics, including the MSD, EMSE and MSE. We present simulation results to verify our theoretical analysis.