Centralized Adaptation for Parameter Estimation over Wireless Sensor Networks

Reza Abdolee, Student Member, IEEE, and Benoit Champagne, Senior Member, IEEE

Abstract—We study the performance of centralized least-mean-squares (CLMS) algorithms in wireless sensor networks where nodes transmit their data over fading channels to a central processing unit (e.g., fusion center or cluster head), for parameter estimation. Wireless channel impairments, including fading and path loss, distort the transmitted data, cause link failure and degrade the performance of the adaptive solutions. To address this problem, we propose a novel CLMS algorithm that uses a refined version of the transmitted data and benefits from a link failure alarm strategy to discard severely distorted data. Furthermore, to remove the bias due to communication noise from the estimate, we introduce a bias-elimination scheme that also leads to a lower steady-state mean-square error. Our theoretical findings are supported by numerical simulation results.

Index Terms—centralized parameter estimation, LMS adaptive algorithms, wireless sensor networks, fading channels.

I. INTRODUCTION

CEN TERALIZED least-mean-squares (CLMS) algorithms were introduced in previous studies for applications in networks with multiple processing nodes [1], [2]. In the development of these algorithms, it is commonly assumed that the communication links between the nodes and the fusion center are error free, i.e., there is no channel distortion such as fading, path loss, and noise during the data transmission. Consequently, these algorithms may experience substantial performance degradation when used over wireless sensor networks (WSN), as evidenced in recent studies on distributed estimation in WSN operating over fading channels [3], [4].

Reference [3] and [4] propose novel distributed adaptive algorithms for networks with imperfect communication links. However, these algorithms do not include provisions to handle possible link failures as well as the estimation bias resulting from the link noise. Work in [5] studies the bias compensation problem for scenarios where only the regression data are noisy. Previous work [6] investigates the problem of space-varying parameter estimation over networks using distributed processing approaches. In these works, however, the communication links between the nodes are assumed to be ideal, i.e., no fading and communication noise are taken into consideration.

We, in this paper, bring all these factors into account and study the performance of CLMS algorithms for parameter estimation in WSN in the presence of radio channel impairments. Specifically, we propose a novel CLMS algorithm that uses a refined version of the transmitted information and benefits from a link failure alarm strategy to discard severely distorted data. We demonstrate the presence of a bias in the adaptive parameter estimates, consequence of the link noise, and introduce a bias-elimination scheme that also significantly decreases the steady-state mean-square error (MSE) of the network. Beside the application of our proposed bias-removal technique in centralized networks, it can be also used in networks with a distributed hierarchical structure, where the cluster heads act as fusion centers and distributed processing is performed at a cluster level.

The adaptive feature of the proposed algorithm is highly desirable in non-stationary signal environments where the underlying network parameters change over time. In summary, the contribution of this letter are as follow: a) development of a new CLMS algorithm for parameter estimation in WSN operating over fading channels; b) analysis of the developed algorithm and identification of the main technical issue under such condition, namely the estimation bias; c) derivation of the relationship between the bias estimates with that of the channel perturbations; d) development of a new bias-compensation technique to remove the bias and e) performance analysis of the developed bias-compensated algorithm, including the derivation of the stability conditions and the steady-state mean-square deviation (MSD) expressions.

II. PROBLEM FORMULATION

Consider a network of $N$ randomly distributed sensor nodes that are deployed over a geographical area to estimate an unknown parameter vector $w^o \in \mathbb{C}^{M \times 1}$. At time instant $i \in \mathbb{N}$, node $k \in \{1, 2, \ldots, N\}$ collects a measurement datum, $d_k(i) \in \mathbb{C}$, that is related to $w^o$ through a linear model\footnote{In this letter, the same mathematical notation as in [6] is used, where boldface fonts are reserved for random variables.}, i.e.: \[
d_k(i) = u_{k,i} w^o + v_{k,i} (1)\]
where the regression vectors $u_{k,i} \in \mathbb{C}^{1 \times M}$ are zero-mean i.i.d. in time and independent over space with positive definite covariance matrix $R_{u,k} = \mathbb{E}[u_{k,i}^* u_{k,i}]$, the noise terms $v_{k,i} \in \mathbb{C}$ are zero-mean i.i.d. in time and independent over space with variance $\sigma_v^2 k$. The regression data $u_{k,i}$ and the measurement noise $v_{k,i}$ are independent [1], [2], [7].

Let us assume that each node $k$ assembles its available data into a row vector as $z_{k,i} = [u_{k,i}, d_k(i)]$ and sends it over a wireless channel to the network fusion center. During its transmission, the data experience distortion of the form: \[
\tilde{z}_{k,i} = a_k(i) z_{k,i} + v_{k,i}^{(z)} (2)\]
where $\tilde{z}_{k,i} = [\hat{u}_{k,i}, \hat{d}_k(i)]$ denotes the received data at the fusion center and $v_{k,i}^{(z)} \in \mathbb{C}^{1 \times (M+1)}$ is the link noise with
\[ v^{(z)} = [v^{(u)}_{k,i}, v^{(d)}_{k,i}] \]. We consider \( a_k(i) = h_k(i) \sqrt{P} r_k \) for analog data transmission, where \( h_k(i) \in \mathbb{C} \) represents the fading channel coefficient, \( P \) is the transmit signal power, \( r_k \) is the distance between node \( k \) and the fusion center and \( \alpha \) is the path loss exponent. For digital-type data communication, the coefficient \( a_k(i) \) will be an indicator function that models the link-failure (packet loss) in baseband after decoding. The fading coefficients \( h_k(i) \) are zero-mean circular complex Gaussian, i.i.d. in time and independent over space with variance \( \sigma_{h,k}^2 \). The link noise vectors \( v^{(z)}_{k,i} \) are zero-mean, i.i.d. in time and independent over space with covariance matrix \( \sigma_{v,k}^2 \mathbf{I} \), while the components \( v^{(u)}_{k,i} \) and \( v^{(d)}_{k,i} \) are independent. The random variables \( h_k(i), u_{k,i}, v_{k,i} \) and \( v^{(z)}_{k,i} \) are all mutually independent. We use \( r^o_k \) to denote the free-space transmission range of node \( k \), as obtained using Friis formula for a given power budget, antenna gains and carrier frequency \( f \). Furthermore, we let \( \varsigma^o_k \) represent the threshold signal-to-noise ratio (SNR) of the received signal from node \( k \), defined as the SNR of the received signal over a non-fading link with communication range \( r^o_k \), i.e., \( \varsigma^o_k = P/\sigma_{v,k}^2 (r^o_k)^\alpha \). We use \( \varsigma_k(i) \) to denote the link instantaneous SNR, which is given by \( \varsigma_k(i) = |h_k(i)|^2 P/\sigma_{v,k}^2 r^o_k \). The transmission from node \( k \) to the fusion center will be successful if \( \varsigma_k(i) \) exceeds the threshold level \( \varsigma^o_k \), i.e., \( |h_k(i)|^2 \geq (r_k/r^o_k)^\alpha \); otherwise, the link fails. In this digital case, this amounts to a packet loss, while in the analog case this would correspond to a temporary loss of synchronization. Since the channel coefficients, \( h_k(i) \), are zero mean circular complex Gaussian random variables, their squared magnitudes, \( |h_k(i)|^2 \), have exponential distribution \[ \mathcal{E} \]. Let \( \lambda = 1/\sigma_{h,k}^2 \), then the probability of successful transmission will be:

\[
p_k = \Pr \left( |h_k(i)|^2 \geq (r_k/r^o_k)^\alpha \right) = e^{-\lambda (r_k/r^o_k)^\alpha} \quad (3)
\]

### III. CENTRALIZED LMS (CLMS) ALGORITHM

Let us first consider the analog-type data communication between nodes and the fusion center. In this case, the fusion center pre-process \{u_{k,i}, d_k(i)\}_{k=1}^N \) to partially recover \{u_{k,i}, d_k(i)\}_{k=1}^N \) before using them in the LMS iteration. This pre-processing can be of linear form, e.g.:

\[
\begin{align*}
\tilde{u}_{k,i} & \triangleq g_k(i) \hat{u}_{k,i} \\
\tilde{d}_k(i) & \triangleq g_k(i) \hat{d}_k(i)
\end{align*}
\]

where \( g_k(i) \) are scalar equalization coefficients. Assuming negligible channel estimation error, these coefficients can be obtained, in practice by, e.g., the least squares (LS) method:

\[
g_k(i) = \frac{h_k(i)}{|h_k(i)|^2} \sqrt{P} r_k
\]

Substituting \( \tilde{g}_k(i) \) into (4) and (5) leads to:

\[
\begin{align*}
\tilde{u}_{k,i} &= u_{k,i} + g_k(i) v_{k,i}^{(u)} \\
\tilde{d}_k(i) &= d_k(i) + g_k(i) v_{k,i}^{(d)}(i)
\end{align*}
\]

The network can now seek the unknown parameter vector \( \mathbf{w}^o \) using the pre-processed data by minimizing:

\[
J(w) = \sum_{k=1}^N \mathbb{E} \left[ |I_k(i)\tilde{d}_k(i) - \tilde{u}_{k,i}w|^2 \right] \quad (9)
\]

where \( I_k(i) \) is a random variable with Bernoulli distribution defined as:

\[
I_k(i) = \begin{cases} 
1, & \text{if } \varsigma_k(i) \geq \varsigma^o_k \\
0, & \text{otherwise}
\end{cases} \quad (10)
\]

That is, \( I_k(i) = 1 \) when the transmission from node \( k \) to the fusion center is successful and \( I_k(i) = 0 \) otherwise. We note that the probability of success, i.e., \( p_k = \mathbb{E} [I_k(i)] \), is given by (3). Since the cost function (9) is strictly convex, its optimal point will be its only stationary point. This leads to (10):

\[
u_{\text{w}^o} = \left( \sum_{k=1}^N (p_k R_{u,k} + R_{v,k}) \right)^{-1} \left( \sum_{k=1}^N p_k r_{du,k} \right)
\]

where \( r_{du,k} = \mathbb{E} [d_k(i) u_{k,i}^{*}] \), and

\[
R_{v,k} = \mathbb{E} [I_k(i)|g_k(i)|^2] R_{v,k}^{(a)}
\]

with \( R_{v,k}^{(a)} = \mathbb{E} [v_{k,i}^{(u)} v_{k,i}^{(d)}] \). The expectation term on the right hand side of (12) can be obtained as:

\[
\mathbb{E} [I_k(i)|g_k(i)|^2] = \int_{0}^{y_k^o} \frac{y_k^o^\alpha}{\Gamma(1+\alpha)} \lambda e^{-\lambda (y_k^o - x_k^o)^\alpha} dy_k
\]

where \( y_k^o \triangleq \left( \frac{r_k}{r^o_k} \right)^\alpha \) and \( x_k^o \triangleq \left( \frac{r_k}{r^o_k} \right)^\alpha \).

Since the covariance and cross-covariance of the data may not be available in practice, the estimate \( w^o \) given by (11) can alternatively be sought using the following CLMS algorithm:

\[
w_i = w_{i-1} + \mu \sum_{k=1}^N I_k(i) \tilde{u}_{k,i} (\tilde{d}_k(i) - \tilde{u}_{k,i} w_{i-1})
\]

where \( \mu > 0 \) is the step-size. The CLMS algorithm for digital data transmission model, i.e., when \( a_k(i) = I_k(i) \) will be the same except that \( \tilde{u}_{k,i} \) and \( \tilde{d}_k(i) \) are now replaced by \( \tilde{u}_{k,i} \) and \( \tilde{d}_k(i) \), respectively. In this case, \( I_k(i) \) will be an indicator function that models the link-failure or packet loss due to fading. We now proceed to investigate the convergence of the CLMS algorithm. To simplify the presentation, we first consider the case \( a_k(i) = 1 \) and assume that the links fail with probability \( p_k \). Later, we include fading and path loss effects into the analysis. Under this condition, from (2), we have:

\[
\hat{d}_k(i) = d_k(i) + v^{(d)}_{k,i}(i)
\]

\[
\hat{u}_{k,i} = u_{k,i} + v^{(u)}_{k,i}
\]

Using (11), (15) and (16), we then obtain:

\[
\hat{d}_k(i) = \hat{u}_{k,i} w^o + \tilde{v}_k(i)
\]

\[
\hat{v}_k(i) = v_k(i) + v^{(d)}_{k,i}(i) - v^{(u)}_{k,i} w^o
\]

The variance of \( \hat{v}_k(i) \) is then given by \( \sigma^2_{\hat{v},k} = \sigma^2_{v,k} + \sigma^2_{v,k} + w^o R_{v,k}^{(a)} w^o \). We now define the weight error-vector \( \tilde{w}_i = w^o - w_i \) and use (14) and (17) to arrive at:

\[
\tilde{w}_i = D_i \tilde{w}_{i-1} - \mu t_i
\]

where \( D_i \triangleq I - \mu \sum_{k=1}^N I_k(i) \hat{u}_{k,i} \hat{u}_{k,i}^* \) and \( t_i \triangleq \sum_{k=1}^N I_k(i) \hat{u}_{k,i} \hat{v}_k(i) \). Taking expectation from both sides of (19), we obtain:

\[
\mathbb{E}[\tilde{w}_i] = D \mathbb{E} [\tilde{w}_{i-1}] - \mu t_f
\]
where $D \triangleq \mathbb{E}[\mathbf{D}_t] = I - \mu \sum_{k=1}^{N} p_k (R_{u,k} + R_{v,k}^{(u)})$ and $t_f \triangleq \mathbb{E}[t_i] = - \sum_{k=1}^{N} p_k R_{o,k}^{(u)}$. From recursion (20), it can be verified that the CLMS algorithm will be mean-stable if

$$0 < \mu < \frac{2}{\lambda_{\text{max}} \left( \sum_{k=1}^{N} p_k (R_{u,k} + R_{v,k}^{(u)}) \right)}$$

(21)

If the step-size satisfy (21), from (20), we can also show that the mean estimate of the algorithm deviates from the optimal estimate by:

$$b \triangleq w^* - \lim_{t \to \infty} \mathbb{E}[\mathbf{w}_t] = \left( \sum_{k=1}^{N} p_k (R_{u,k} + R_{v,k}^{(u)}) \right)^{-1} t_f$$

(22)

We observe that the bias, $b$, does not depend on the step-size, $\mu$. Therefore, choosing a smaller step-size will not improve the accuracy of the estimated parameters.

To incorporate the effects of fading and path-loss into the above results, we can use expressions (7) and (8) and repeat the analysis. Doing so, we obtain similar results as before, except that in this case the power of the communication noise $\nu_k(i)$ and $\nu_v(i)$ increase by the factor of $\mathbb{E}[\mathcal{I}_k(i)|g_k(i)|^2]$. Consequently, (21) and (22) can still be used if, in these expressions, we replace $p_k R_{v,k}^{(u)}$ with $R_{v,k}$.}

IV. BIAS-COMPENSATED CLMS

As seen from (22), the mean of the weight error vector of the CLMS algorithm in (14), i.e., $\mathbb{E}[\mathbf{w}_t]$, converges to a non-zero vector. In what follows, we propose a bias-compensation scheme and develop a new form of centralized LMS algorithm whose mean estimation error converges to zero.

As shown in (22), the estimation bias $b$ is due to the term $t_f$, which in turn is caused by the regression noise component $\nu_k(i)$. Assuming zero regression noise in (9), the unbiased optimal estimate of the network will be:

$$w^o = \left( \sum_{k=1}^{N} p_k R_{u,k} \right)^{-1} \left( \sum_{k=1}^{N} p_k r_{du,k} \right)$$

(23)

We, in this paper, propose a CLMS algorithm that runs over the noisy received data and achieves the optimal estimate (23). The basic idea in our development is to construct an objective function whose gradient vector is identical to that of the cost function with zero regression noise (9). The following objective function satisfies this criterion:

$$J(w) = \sum_{k=1}^{N} \mathbb{E}[\mathcal{I}_k(i) (|\nu_k(i) - \nu_k(i)|^2 - |R_{v,k}^{(u)}|^2)]$$

(24)

It can be verified that the Hessian of $J(w)$ is positive definite and hence this cost function is strongly convex. From (24), we then arrive at the following bias-compensated CLMS (BC-CLMS) algorithm where $\tilde{\mu}$ is the new step-size:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \tilde{\mu} \sum_{k=1}^{N} \mathcal{I}_k(i) \left\{ \tilde{\nu}_{k,i} (\nu_k(i) - \nu_k(i)) - \tilde{g}_k(i) R_{v,k}^{(u)} \right\} + g_k(i) R_{v,k}^{(u)}$$

(25)

This algorithm requires the covariance matrix of the regression noise, which can be estimated off-line or during the operation of the algorithm in real-time as in e.g., [5], [12]. The bias-compensated CLMS algorithm for digital data transmission will be the same if we replace $\tilde{u}_{k,i}$ and $\tilde{d}_k(i)$ with $\hat{u}_{k,i}$ and $\hat{d}_k(i)$, respectively and set $g_k(i)$ to one.

We now proceed to analyze the proposed algorithm. Similar to previous section, we first consider the case with $a_k(i) = 1$. Under this condition, it can be verified that the weight error vector of the algorithm evolves with time according to:

$$\mathbf{w}_i = \tilde{D} \mathbf{w}_{i-1} - \mu \mathbf{t}_i$$

(26)

where, in this case, $D \triangleq I - \mu \sum_{k=1}^{N} \mathcal{I}_k(i) (\tilde{u}_{k,i} \tilde{u}_{k,i} - R_{v,k}^{(u)})$ and $\mathbf{t}_i \triangleq \sum_{k=1}^{N} \mathcal{I}_k(i) (\tilde{u}_{k,i} \hat{v}_k(i) + R_{v,k}^{(u)} \nu_v(i))$. By taking the expectation of (26), we arrive at:

$$\mathbb{E}[\mathbf{w}_i] = \tilde{D} \mathbb{E}[\mathbf{w}_{i-1}]$$

(27)

where $\tilde{D} \triangleq \mathbb{E}[\mathbf{D}_i] = I - \mu \sum_{k=1}^{N} p_k R_{u,k}$. It can be verified that $\lim_{i \to \infty} \mathbb{E}[\mathbf{w}_i] \to 0$ if

$$0 < \tilde{\mu} < \frac{2}{\lambda_{\text{max}} \left( \sum_{k=1}^{N} p_k R_{u,k} \right)}$$

(28)

The mean-square deviation (MSD) recursion of this algorithm can be expressed as:

$$\mathbb{E}[\|\mathbf{w}_i\|^2] = \mathbb{E}[\|\mathbf{w}_{i-1}\|^2] + \mathbb{E}[\|\nu_v(i)\|^2] \sum_{k=1}^{N} p_k R_{v,k}^{(u)}$$

(29)

where $\mathbb{E}[\|\mathbf{w}_i\|^2] = \mathbb{E}[\|\mathbf{w}_i - \mathbf{w}_i^o\|^2]$, $\Sigma \geq 0$, $\Sigma' \triangleq \mathbb{E}[\mathbf{D}_i \Sigma \mathbf{D}_i^T]$ and

$$\tilde{X} = \sum_{k=1}^{N} p_k \left( \sigma_v^2 \nu_{v,k} R_{u,k} + (\beta - 1) (R_{v,k}^{(u)} w^o R_{v,k}^{(u)})^T \right) + R_{v,k}^{(u)} \mathbb{E}[\nu_v(i) \nu_v(i)]$$

(30)

Using this recursion, the steady-state MSD of the algorithm converges to:

$$\tilde{\eta}_{\text{est}} = [\mathbb{E}[\nu_v^2(i)] (I - \tilde{F})^{-1} \mathbb{E}[\mathbf{I}_M]$$

(31)

where $\tilde{F} \approx \tilde{D}^T \otimes \tilde{D}^*$. As mentioned before, the steady-state MSD of the network under fading condition can be computed by replacing $p_k R_{v,k}^{(u)}$ with $R_{v,k}^{(u)}$ given by (12).

V. SIMULATION RESULTS

In this section, we present computer experiments to illustrate the performance of the CLMS algorithm (14) and the BC-CLMS algorithm (25) over a wireless sensor network. We consider a WSN consisting of $N = 5$ nodes, uniformly distributed over a squared area of $[x, y] = [1, 1]$ km with the fusion center located at $x = y = 0.5$ km. The unknown parameter vector to be estimated is $w^o = [-0.8006, -0.3203 + j0.1601, 0.4804]^T$ with $M = 3$. All nodes have equal transmit signal power $P = 10$ mW, the nominal transmission range is $r_i = 0.3$ km, and the path loss exponent is $\alpha = 2.5$. We initialize the algorithms with $\mathbf{w}_{-1} = [0, 0, 0]^T$ and set the step-size to $\mu = 0.003$. To generate communication noise vectors $\nu_v(i)$, we adopt complex normal distributions with zero mean and variances $\sigma_v^2 \in \{0.0617, 0.0560, 0.0923, 0.0831, 0.0476\}$. The measurement noise $\nu_v(i)$ is generated using a normal distribution with variances $\sigma_v^2 \in \{0.069, 0.090, 0.087, 0.092, 0.061, \}$. The fading coefficients of the radio links are obtained from a circular complex normal distribution with zero mean and unit variance. The regression data $\nu_v(i)$ are generated from a circular complex Gaussian distribution where the $(n, m)$-th
CLMS algorithm under imperfect channel information at about 8dB. This figure also shows the performance of the BC-CLMS algorithm proposed in [2]. As this performance of the proposed BC-CLMS algorithm with that of the conventional CLMS algorithm achieves a significantly lower steady-state error vector $E[w]$. As expected, the BC-CLMS algorithm decreases while its steady-state performance remains invariant. In contrast, as the channel estimation error increases the steady-state MSD of the algorithm increases. As this result indicates, the proposed algorithm is robust to small channel estimation error and temporal correlation.

VI. CONCLUSION

We studied the performance of CLMS algorithms for parameter estimation over wireless sensor networks where links between nodes and the fusion center are impaired by fading and noise. The analysis and the numerical simulations show that the proposed BC-CLMS algorithm is stable and converges to an unbiased estimate for small adaptation step-sizes.

REFERENCES

[1] C. G. Lopes and A. H. Sayed, “Diffusion least-mean squares over adaptive networks: Formulation and performance analysis,” IEEE Trans. on Signal Processing, vol. 56, no. 7, pp. 3122–3136, July 2008.
[2] F. S. Cattivelli and A. H. Sayed, “Diffusion LMS strategies for distributed estimation,” IEEE Trans. on Signal Processing, vol. 58, no. 3, pp. 1035–1048, Mar. 2010.
[3] R. Abdolee and B. Champagne, “Diffusion LMS algorithms for sensor networks over non-ideal inter-sensor wireless channels,” in Proc. of Int. Conf. on Dist. Computing in Sensor Systems and Workshops, Barcelona, Spain, June 2011, pp. 1–6.
[4] R. Abdolee, B. Champagne, and A. H. Sayed, “Diffusion LMS strategies for parameter estimation over fading wireless channels,” in Proc. of IEEE Int. Conf. on Communications (ICC), Budapest, Hungary, June 2013, pp. 1926–1930.
[5] A. Bertrand, M. Moonen, and A. H. Sayed, “Diffusion bias-compensated RLS estimation over adaptive networks,” IEEE Trans. on Signal Processing, pp. 5212–5224, Nov. 2011.
[6] R. Abdolee, B. Champagne, and A. Sayed, “Diffusion bias-compensated RLS estimation over adaptive networks,” IEEE Trans. on Signal Processing, vol. 62, no. 2, pp. 403–418, Jan. 2014.
[7] X. Zhao and A. H. Sayed, “Performance limits for distributed estimation over LMS adaptive networks,” IEEE Trans. on Signal Processing, vol. 60, no. 10, pp. 5107–5124, Oct. 2012.
[8] P. V. Nikitin and K. S. Rao, “Theory and measurement of backscattering from RFID tags,” IEEE Antennas and Propagation Magazine, vol. 48, no. 6, pp. 212–218, Dec. 2006.
[9] A. Leon-Garcia, Probability and Random Processes for Electrical Engineering, 2nd edition, Addison-Wesley, Reading, MA, 1994.
[10] A. H. Sayed, Adaptive Filters. Wiley, NJ, 2008.
[11] R. Abdolee, B. Champagne, and A. H. Sayed, “A diffusion LMS strategy for parameter estimation in noisy regressor applications,” in Proc. European Signal Processing Conference (EUSIPCO), Bucharest, Romania, Aug. 2012, pp. 749–753.
[12] W. X. Zheng, “A least-squares based algorithm for FIR filtering with noisy data,” in Proc. IEEE International Symposium on Circuits and Systems, May 2003, pp. IV-444–IV-447.