A LOW-COMPLEXITY MAP DETECTOR FOR DISTRIBUTED NETWORKS

Allan E. Feitosa *Vítor H. Nascimento* *Cassio G. Lopes*

Dept. of Electronic Systems Engineering of University of São Paulo

ABSTRACT

This work describes a generalization of our previous maximum likelihood (ML) detector to a maximum a posteriori (MAP) detector in distributed networks using the diffusion LMS algorithm. Nodes in the network must decide between two concurrent hypotheses concerning their environment, using local measurements and shared estimates from neighbors. The generalization is provided by a new approximation concerning the network connectivity, whose accuracy is shown by simulations. The new MAP detector inherits from our ML formulation an exponential decay rate in probability of error independent of the LMS step size, if it is sufficiently small.

***Index Terms*—** adaptive networks, distributed detection, maximum a posteriori, network connectivity.

1. INTRODUCTION

Detection in distributed sensor network has been studied in many different contexts, such as event and security detection in IoT environments [1–3] and in one-bit messaging systems [4, 5]. Some works adopt data diffusion schemes [5–10] or consensus schemes [2, 3, 11–14], and concerning detection itself, a common approach is the classical *Neyman-Pearson* (NP) detector [4, 6, 7], which is historically used, for example, to detect the presence of a target in a radar. Another common approach is the Bayesian formulation [15, 16], when prior probabilities are set for the possible outcomes.

In our previous work [17], we developed a distributed maximum likelihood (ML) detector (i.e., when the prior probabilities are equal) which attains an exponential decay in the probability of error during the transient, due to a specific initialization of the diffusion LMS estimator. Furthermore, the transient performance of that detector is almost indifferent to the value of the LMS step size, if it is sufficiently small. This work proposes the more general *Maximum a Posteriori* (MAP) detector for when the prior probabilities are not equal, keeping the performance of the ML detector. For that we introduce new approximations assuming, only for derivations, that every node in the network is connected to every other node; this is reasonable due to the diffusion LMS equalization effect on connected networks [18]. We also assume, for an appropriate development, deterministic regressors with known covariance matrices, a sufficiently small LMS step size, and that we can determine the noise existent in measurements.

This work has been supported in part by the ELIOT, grant ANR-18-CE40-0030, by São Paulo Research Foundation (FAPESP), grants 2018/12579-7 and 2018/26040-2, and by CNPq, grant 304714/2018-6.

2. DETECTION OVER DISTRIBUTED NETWORKS

In a distributed network with K nodes, each node k at every time i has access to a scalar observation $d_k(i)$ of its environment and to a *deterministic* regressor vector $\boldsymbol{u}_k(i) \in \mathbb{R}^M$, i.e., we assume the $\boldsymbol{u}_k(j)$, $0 \leq j \leq i$, to be known at time i. Each node receives additional information only from its neighbors (without a fusion center), which can improve the quality of the decision process [6, 7, 17, 19]. The following data model:

$$d_k(i) = \boldsymbol{u}_k^T(i)\boldsymbol{w}_s + v_k(i), \quad k = 1, \ldots, K\,, \tag{1}$$

relates the observable variables $d_k(i)$ and $\boldsymbol{u}_k(i)$ to an unknown parameter vector $\boldsymbol{w}_s \in \mathbb{R}^M$, where $v_k(i) \sim \mathcal{N}(0, \sigma_{v,k}^2)$ is random noise. Nodes must decide between two hypotheses about the unknown parameter: $\boldsymbol{w}_s = \boldsymbol{w}_1$ if H_1 is true, or $\boldsymbol{w}_s = \boldsymbol{w}_2$ if H_2 is true. Let $p(H_1)$ and $p(H_2)$ be the respective prior probabilities. We assume that the two possibilities are known *a priori*, although it is not known which of the two is in place. As in [17], we wish to minimize the maximum probability of error of the distributed detector, using the following test structure, for each node k, as in [6, 7, 17, 19]:

$$t_k\big(\boldsymbol{w}_k(i)\big) \underset{H_1}{\overset{H_2}{\gtrless}} \gamma\,, \tag{2}$$

where t_k is a local function dependent on a local estimate $\boldsymbol{w}_k(i)$ of $\boldsymbol{w}_s$, and γ is an appropriate threshold. This structure exploits the connection between detection and estimation [20], which can be verified, for instance, when $\boldsymbol{w}_k(i)$ is the *Minimum Variance Unbiased Estimate* for linear models.

In order to describe a fully distributed MAP detector, we start by collecting all data of all nodes up to time i as follows:

$$\begin{aligned}
\boldsymbol{d}_{0:i} &= \mathrm{col}\{d_1(i), \ldots, d_K(i), \ldots, d_1(0), \ldots, d_K(0)\},\\
U_{0:i} &= \mathrm{col}\{\boldsymbol{u}_1^T(i), \ldots, \boldsymbol{u}_K^T(i), \ldots \boldsymbol{u}_1^T(0), \ldots, \boldsymbol{u}_K^T(0)\},\\
\boldsymbol{v}_{0:i} &= \mathrm{col}\{v_1(i), \ldots, v_K(i), \ldots v_1(0), \ldots, v_K(0)\},
\end{aligned}$$

where the col operator stacks its arguments column-wise. Therefore, equation (1) can be rewritten as

$$\boldsymbol{d}_{0:i} = U_{0:i}\boldsymbol{w}_s + \boldsymbol{v}_{0:i}\,, \tag{3}$$

which describes a global linear model for the network data under a global Gaussian perturbation vector $\boldsymbol{v}_{0:i} \sim \mathcal{N}(\mathbf{0}, \Sigma_{0:i})$, where $\mathbf{0}$ denotes the zero vector and

$$\Sigma_{0:i} = E\{\boldsymbol{v}_{0:i}\boldsymbol{v}_{0:i}^T\} = I_{i+1} \otimes \Sigma\,,\ \Sigma = \mathrm{diag}(\sigma_{v,1}^2, \ldots, \sigma_{v,K}^2),$$

where I_{i+1} is the identity matrix of dimension $i+1$ and the operator $\otimes$ denotes the Kronecker product of matrices.

In order to obtain the local estimate $\boldsymbol{w}_k$ in (2) in a distributed and cooperatively way, we use the diffusion LMS algorithm, an estimation technique given by the equations

$$\boldsymbol{\psi}_k(i) = \boldsymbol{w}_k(i-1) + \mu_k \boldsymbol{u}_k(i)[d_k(i) - \boldsymbol{u}_k^T(i)\boldsymbol{w}_k(i-1)]\,, \quad (4)$$

$$\boldsymbol{w}_k(i) = \sum_{l=1}^{K} a_{k,l}\boldsymbol{\psi}_l(i)\,, \quad \sum_{l=1}^{K} a_{k,l} = 1,\ a_{k,l} \geq 0,\ \forall k, l\,, \quad (5)$$

where $\boldsymbol{\psi}_k(i)$ are intermediate estimates, μ_k is a positive step size and the scalar $a_{k,l}$ represents the weighted contribution of $\boldsymbol{\psi}_l(i)$ in (4) from node l to estimate $\boldsymbol{w}_k(i)$ of node k. Furthermore, consider that the estimates $\boldsymbol{w}_k(i)$ in (5) are linearly related to $\boldsymbol{d}_{0:i}$ (assuming $\boldsymbol{w}_k(-1) = \mathbf{0}$); that is,

$$\boldsymbol{w}_k(i) = L_k(i)\boldsymbol{d}_{0:i}\,, \quad (6)$$

where $L_k(i) \in \mathbb{R}^{M\times(i+1)K}$ defines the local linear estimation procedure; thus, it is related to the diffusion LMS in our case.

Define $P_k(i) = L_k(i)U_{0:i} \in \mathbb{R}^{M\times M}$; we have then from (3) and (6) that

$$\boldsymbol{w}_k(i) = P_k(i)\boldsymbol{w}_s + L_k(i)\boldsymbol{v}_{0:i}\,, \quad (7)$$

which is also a linear model of the data under the Gaussian perturbation $L_k(i)\boldsymbol{v}_{0:i} \sim \mathcal{N}(\mathbf{0}, S_k(i))$, where we defined $S_k(i) = L_k(i)\Sigma_{0:i}L_k^T(i) \in \mathbb{R}^{M\times M}$, which is the covariance matrix of $\boldsymbol{w}_k(i)$. Considering the data model in (7) and $\boldsymbol{w}_k(-1) = \mathbf{0}$, an exact local test for the distributed MAP detector using diffusion LMS is given as [20, p. 115][1]

$$t_{k,\mathrm{ex}}(i) = \boldsymbol{w}_k^T(i)Q_k(i)(\boldsymbol{w}_2 - \boldsymbol{w}_1) - \frac{1}{2}(\mathcal{E}_{k,2} - \mathcal{E}_{k,1}) \underset{H_1}{\overset{H_2}{\gtrless}} \gamma, \quad (8)$$

in which we defined $Q_k(i) = S_k^{-1}(i)P_k(i) \in \mathbb{R}^{M\times M}$, $\mathcal{E}_{k,s} = \boldsymbol{w}_s^T P_k^T(i)Q_k(i)\boldsymbol{w}_s$, $s \in \{1, 2\}$, and $\gamma = \ln(p(H_1)/p(H_2))$.

The computation of quantities $Q_k(i)$ and $P_k(i)$ is quite demanding, since it would require their exchanging in the diffusion LMS update. In order to simplify the computations and make them amenable for distributed implementations, [6] proves, for diffusion LMS, that $Q_k(i)$ in (8) can be approximated by the identity matrix multiplied by a constant β_k, supposing that: 1) no diffusion update is used to compute $Q_k(i)$, that is, $a_{k,l} = 1$ for $l = k$ and $a_{k,l} = 0$ for $k \neq l$ in (5); and 2) the step sizes μ_k are sufficiently small.

[1] When $\boldsymbol{w}_k(-1) \neq \mathbf{0}$, $t_{k,\mathrm{ex}}(i)$ is slightly different; we must subtract from $\boldsymbol{w}_k(i)$ in (8) a term that accounts for the LMS adaptation of $\boldsymbol{w}_k(-1)$, which cancels the effect of initializing $\boldsymbol{w}_k(-1) \neq \mathbf{0}$. As a result, the performance of the exact distributed MAP detector does not depend on the initial estimate.

In [17], we followed these assumptions to propose our distributed ML detector, and since the performance of a detector is not affected when the test in (8) is multiplied by a constant when the threshold γ is set to zero (which is the case for the ML detector), we can simply approximate $Q_k(i)$ by the identity matrix. Nevertheless, in the more general case where $\gamma \neq 0$, the constants β_k cannot be ignored and must be estimated for each node k.

In light of the above, we take a different approach in this work; instead of supposing no diffusion when computing $Q_k(i)$, we suppose the complete opposite, that is, a totally connected network where each node communicates directly with every other node. This is reasonable, since the diffusion LMS nearly equalizes the overall estimation performance of nodes in a network, as long the network is connected [18]; thus, we expect the matrices $Q_k(i)$ to be equalized as well. In this sense, we exploit this equalization property to deduce a single approximation $Q_k(i)$ for every node.

3. DEVELOPING THE MAP DETECTOR

We show in this section how the new hypothesis concerning the diffusion learning process produces an implementable detector. We start developing an expression for $P_k(i)$, which was not present in our previous work. We first deduce an expression for $L_k(i)$ in (6) for diffusion LMS; Equations (4) and (5) can be rewritten in a single expression as follows:

$$\boldsymbol{w}_k(i) = \sum_{l=1}^{K} a_{k,l}\mu_l\mathcal{W}_l(i)\boldsymbol{d}(i) + \sum_{l=1}^{K} a_{k,l}Y_l(i)\boldsymbol{w}_l(i-1)\,, \quad (9)$$

where we defined

$$\begin{aligned} \mathcal{W}_k(i) &= [\mathbf{0}\ \mathbf{0} \ldots \boldsymbol{u}_k(i) \ldots \mathbf{0}], \quad (M \times K) \\ \boldsymbol{d}(i) &= \mathrm{col}\{d_1(i), \ldots, d_K(i)\}, \ (K \times 1) \\ Y_k(i) &= I_M - \mu_k\boldsymbol{u}_k(i)\boldsymbol{u}_k^T(i), \quad (M \times M) \end{aligned}$$

and $I_M \in \mathbb{R}^{M\times M}$ is the identity matrix. Using (6) to substitute $\boldsymbol{w}_l(i-1)$ by $L_l(i-1)\boldsymbol{d}_{0:i-1}$ in (9), we can rewrite it as

$$\boldsymbol{w}_k(i) = \left[\sum_{l=1}^{K} a_{k,l}\mu_l\mathcal{W}_l(i) \quad \sum_{l=1}^{K} a_{k,l}Y_l(i)L_l(i-1)\right]\begin{bmatrix}\boldsymbol{d}(i)\\ \boldsymbol{d}_{0:i-1}\end{bmatrix};$$

thus, by the definition in (6), we have

$$L_k(i) = \left[\sum_{l=1}^{K} a_{k,l}\mu_l\mathcal{W}_l(i) \quad \sum_{l=1}^{K} a_{k,l}Y_l(i)L_l(i-1)\right]. \quad (10)$$

Let us find a global description for $P_k(i)$. Define

$$\begin{aligned} \mathcal{P}(i) &= \mathrm{col}\{P_1(i), \ldots, P_K(i)\}, \quad (KM \times M) \\ \mathcal{L}(i) &= \mathrm{col}\{L_1(i), \ldots, L_K(i)\}, \quad (KM \times (i+1)K) \end{aligned}$$

and, by the definition of $P_k(i)$ right before (7), it holds that

$$\mathcal{P}(i) = \mathcal{L}(i)U_{0:i}\,. \quad (11)$$

Also define $A = (a_{k,l}) \in \mathbb{R}^{K\times K}$, such that $A\mathbb{1}_K = \mathbb{1}_K$, with $\mathbb{1}_K$ being a vector with entries equal to one in $\mathbb{R}^K$. We also need the following expressions to complete the $\mathcal{P}(i)$ description:

$$\begin{aligned}
\mathcal{H}(i) &= \mathrm{col}\{\boldsymbol{u}_1(i)\boldsymbol{u}_1^T(i), \ldots, \boldsymbol{u}_K(i)\boldsymbol{u}_K^T(i)\}, && (KM\times M)\\
\mathcal{A} &= A\otimes I_M \in \mathbb{R}^{KM\times KM}, && (KM\times KM)\\
\mathcal{M} &= \mathrm{diag}\{\mu_1 I_M, \ldots, \mu_K I_M\}, && (KM\times KM)\\
\mathcal{Y}(i) &= \mathrm{diag}\{Y_1(i), \ldots, Y_K(i)\}. && (KM\times KM)
\end{aligned}$$

It is then straightforward to establish from (10) and (11) that the following recursion holds:

$$\mathcal{P}(i) = \mathcal{A}\big(\mathcal{M}\mathcal{H}(i) + \mathcal{Y}(i)\mathcal{P}(i-1)\big). \tag{12}$$

Setting $\mathcal{P}(-1) = 0$, and also $\mu_k = \mu$ for every node for simplicity, a nonrecursive expression for (12) is given by

$$\mathcal{P}(i) = \mu\sum_{j=0}^{i}\mathcal{F}(i,j+1)\mathcal{A}\,\mathcal{H}(j)\,, \tag{13}$$

in which we defined the following matrix:

$$\mathcal{F}(i,j) = \begin{cases}\prod_{n=j}^{i}\mathcal{A}\,\mathcal{Y}(i+j-n) & \text{if } i\geq j\,,\\ I_{KM} & \text{if } i<j\,.\end{cases} \tag{14}$$

We showed in [17] that (14) can be expanded as a polynomial on μ, and that the zero-order coefficient is given by $\mathcal{C}_0(i,j) = \mathcal{A}^{i-j+1}$. Substituting $\mathcal{F}(i,j)$ by $\mathcal{C}_0(i,j)$ in (13) results in a first-order polynomial approximation for $\mathcal{P}(i)$, given as

$$\mathcal{P}_{1\text{st}}(i) = \mu\sum_{j=0}^{i}\mathcal{A}^{i-j+1}\,\mathcal{H}(j)\,, \tag{15}$$

which is adequate for sufficient small step size μ.

Our next step is finding an approximation for $S_k(i)$. Defining $\boldsymbol{w}(i) = \mathrm{col}\{\boldsymbol{w}_1(i), \ldots, \boldsymbol{w}_K(i)\}$, a nonrecursive expression for the covariance matrix $\mathcal{S}(i)$ of $\boldsymbol{w}(i)$ was developed in [17] and is given as

$$\mathcal{S}(i) = \mu^2\sum_{j=0}^{i}\mathcal{F}(i,j+1)\,\mathcal{A}\,\mathcal{D}_v(j)\mathcal{A}^T\mathcal{F}^T(i,j+1), \tag{16}$$

where $\mathcal{D}_v(i) = \mathrm{diag}\{\sigma_{v,1}^2\boldsymbol{u}_1(i)\boldsymbol{u}_1^T(i), \ldots, \sigma_{v,K}^2\boldsymbol{u}_K(i)\boldsymbol{u}_K^T(i)\}$ (recall that $\boldsymbol{u}_k(i)$ is deterministic). Also note that $S_k(i)$ is the kth $M\times M$ diagonal block of $\mathcal{S}(i)$. Likewise, substituting $\mathcal{F}(i,j)$ by $\mathcal{C}_0(i,j)$ in (16) results in a second-order polynomial approximation for $\mathcal{S}(i)$, given as

$$\mathcal{S}_{2\text{nd}}(i) = \mu^2\sum_{j=0}^{i}\mathcal{A}^{i-j+1}\mathcal{D}_v(j)\big(\mathcal{A}^T\big)^{i-j+1}. \tag{17}$$

We now apply the totally connected network hypothesis in order to complete our formulation. As a result of doing this, the matrix A reduces to $(1/K)\,\mathbb{1}_K\,\mathbb{1}_K^T$ and, by consequence, $\mathcal{A} = \mathcal{A}^T = (1/K)\mathbb{I}\,\mathbb{I}^T$, where $\mathbb{I} = \mathrm{col}\{I_M, \ldots, I_M\} \in \mathbb{R}^{MK\times M}$. Under this condition, it is straightforward to note that $\mathcal{A}^n = \mathcal{A},\ \forall n\in\mathbb{N}^*$. Therefore, substituting the resultant $\mathcal{A}$ in (15) and (17), we have the following approximations:

$$\begin{aligned}
P_k(i) &\approx \frac{\mu}{K}\sum_{l=1}^{K}\left(\sum_{j=0}^{i}\boldsymbol{u}_l(j)\boldsymbol{u}_l^T(j)\right),\\
S_k(i) &\approx \frac{\mu^2}{K^2}\sum_{l=1}^{K}\sigma_{v,l}^2\left(\sum_{j=0}^{i}\boldsymbol{u}_l(j)\boldsymbol{u}_l^T(j)\right),
\end{aligned} \tag{18}$$

valid for a sufficient small step size μ and $\forall k$. In order to derive a simple expression for the innermost sums in (18), we assume that the deterministic regressors $\boldsymbol{u}_k(i)$ are a stationary process with covariance matrix $R_{u,k}$ to approximate $\sum_{j=0}^{i}\boldsymbol{u}_l(j)\boldsymbol{u}_l^T(j) \approx (i+1)R_{u,l}$; then we can rewrite (18) as

$$\begin{aligned}
P_k(i) &\approx \frac{\mu}{K}(i+1)\sum_{l=1}^{K}R_{u,l},\\
S_k(i) &\approx \frac{\mu^2}{K^2}(i+1)\sum_{l=1}^{K}\sigma_{v,l}^2R_{u,l}.
\end{aligned} \tag{19}$$

Let us make the approximation $R_{u,k} \approx \sigma_{u,k}^2 I_M$ in (19), which is reasonable, provided the correlations between different entries in $\boldsymbol{u}_k(i)$ are small. Then, we can approximate $Q_k(i) = S_k^{-1}(i)P_k(i)$ as

$$Q_k(i) \approx \frac{K}{\mu}\frac{\sum_{l=1}^{K}\sigma_{u,l}^2}{\sum_{l=1}^{K}\sigma_{v,l}^2\sigma_{u,l}^2}I_M = \beta I_M\,, \tag{20}$$

provided the $\sigma_{v,l}^2$ are possible to be determined.[2]

Therefore, under all the assumptions above, we approximate $Q_k(i)$ in (8) by (20) at every node. As in [17], we also approximate $P_k(i)$ by I_M in $\mathcal{E}_{k,s}$ of (8), since $P_k(i)\approx I$ after convergence of diffusion LMS. Thereby, we propose the following local approximate and low-complexity MAP detector:

$$\boxed{t_k(i) = \boldsymbol{w}_k^T(i)(\boldsymbol{w}_2-\boldsymbol{w}_1) - \frac{1}{2}\big(||\boldsymbol{w}_2||_2^2 - ||\boldsymbol{w}_1||_2^2\big)\underset{H_1}{\overset{H_2}{\gtrless}}\gamma',} \tag{21}$$

where $\gamma' = \gamma/\beta$ and $\gamma = \ln(p(H_1)/p(H_2))$. Since $t_k(i)$ in (21) is the same as our ML detector, it has the same performance features, namely an exponential decay in the probability of error during the transient, which is almost indifferent to the value of the step size μ, if it is sufficiently small. These features are obtained when our algorithm is initialized at

$$\boldsymbol{w}_k(-1) = \frac{\boldsymbol{w}_1+\boldsymbol{w}_2}{2} = \bar{\boldsymbol{w}}\,, \tag{22}$$

which also avoids the initial time in which the probability of error does not decrease, as showed in [17, Fig. 1]. As a matter

[2]For instance, in some applications, $\sigma_{v,l}^2$ can be estimated shutting down the filters running in the nodes, i.e., $\boldsymbol{u}_k(i) = \mathbf{0}$ in (1), so that $d_k(i) = v_k(i)$.

Table 1. Network Statistics

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$\alpha_k \times 10$	7.7	2.8	7.6	4.9	4.2	6.1	4.2	8.9	6.2	8.5	0.1	6.7	2.3	0.5	6.2	7.7	3.9	4.8	7.9	8.5
$\sigma^2_{u,k} \times 10$	5.3	1.8	6.5	6.0	2.0	2.4	8.1	1.1	5.3	4.7	10	6.5	5.4	0.9	4.7	5.5	1.8	9.3	9.0	2.3
$\sigma^2_{v,k} \times 100$	1.5	6.7	1.1	2.2	1.0	3.6	1.1	4.0	2.6	4.8	1.7	8.9	2.6	0.1	5.4	8.1	8.4	1.5	7.9	9.5

of fact, the same applies to our MAP detector, even though γ' is different from zero. To see that, let us investigate the probability of error of the MAP detector, which is given as

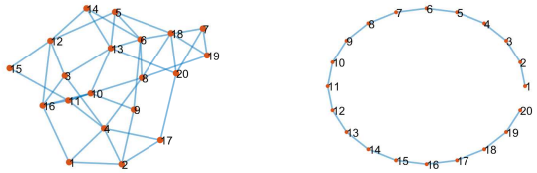
$$p_{\mathrm{e},k}(i) = p(H_1)\mathcal{Q}\left(\frac{\gamma' - E\{t_k(i)|H_1\}}{\sigma_k(i)}\right) + p(H_2)\mathcal{Q}\left(\frac{E\{t_k(i)|H_2\} - \gamma'}{\sigma_k(i)}\right), \quad (23)$$

where $\sigma_k(i)$ is the standard deviation of $t_k(i)$ and $\mathcal{Q}(z) = \mathrm{Prob}(Z > z)$ for a standard Gaussian variable Z. We showed in [17] that when $\boldsymbol{w}_k(-1)$ is as (22), the ratio between $E\{t_k(i)|H_s\}$ and $\sigma_k(i)$ does not depend on μ, if μ is sufficient small and during the transient. We also showed that $\sigma_k(i)$ can be approximated by a linear function of μ [17, Eq. 28]; therefore, the ratio between the threshold γ' and $\sigma_k(i)$ also does not depend on μ, since $1/\beta$ is a linear function of μ as well (20). Consequently, under our assumptions, the probability of error is approximately independent of μ during the transient of the filters.

4. SIMULATIONS

We consider two network topologies, A and B, with $K = 20$ nodes in Fig. 1. Topology A is the same used in [6], and B tests the simplest topology in which the network is still connected. The prior probabilities are $p(H_1) = 1/3$ and $p(H_2) = 2/3$. In order to obtain regressors $\boldsymbol{u}_k(i)$ with full covariance matrices, we generate 20 signals using first order Markov processes (which are used, among other applications, to model certain communication channels), with power $\sigma^2_{u,k}$ and correlation α_k. Recall that at time instant i the regressors $\boldsymbol{u}_k(i)$, $0 \leq j \leq i$ are known, and our detector approximates the optimal detector given $\boldsymbol{u}_k(j)$, $0 \leq j \leq i$. The noise in the measurements has power $\sigma^2_{v,k}$. The values of $\sigma^2_{u,k}$, α_k and $\sigma^2_{v,k}$ were obtained randomly and can be seen in Table 1. The weights $a_{k,l}$ in (5) are obtained using the Metropolis rule [21], resulting $A = A^T$. We choose $\boldsymbol{w}_1 = [1, 1, \ldots, 1]^T$ and $\boldsymbol{w}_2 = [0.97, 1, \ldots, 1]^T$ very close to each other in order to test an unfavorable detection situation.

The plots in Fig. 2 show maximum probability of error $p_{\mathrm{e,max}}(i) = \max_k p_{\mathrm{e},k}(i)$ of our simulated MAP detector from (21) in transient, for topologies A and B and using 10^5 realizations for estimating the probabilities. We compare the results with the theoretical MAP detector computing the exact values of $Q_k(i)$ and $P_k(i)$ [see (8)] and considering the same topologies. Our detector is initialized as (22), whereas the exact one is accordingly initialized as $\boldsymbol{w}(-1) = \boldsymbol{0}$. The LMS



(a) Topology A (b) Topology B

Fig. 1. Topologies used for simulations.

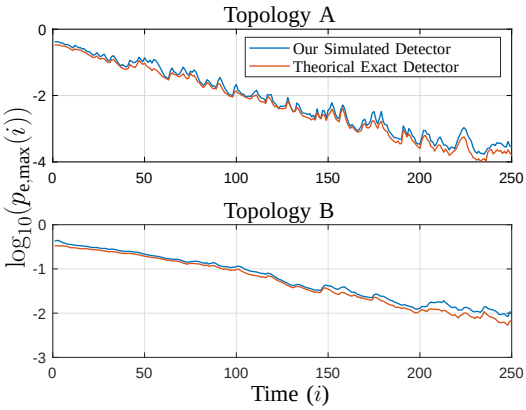


Fig. 2. Logarithm of the maximum probability of error in the network, for topologies A and B.

step size is $\mu = 1.10^{-6}$, which is sufficiently small for our purposes; as we showed in [17, Fig. 2], choosing smaller step sizes maintains the same decay rate while lowering the steady state probability of error (not shown in Fig. 2). As we can see, our detector approximates very closely the exact one for the two topologies, confirming that the totally connected network hypothesis works very well, even for topology B, and the cost of using this approximation is just the small difference between our detector from (21) and the exact in both cases.

5. CONCLUSION

We developed a low-complexity distributed MAP detector that is a generalization of our previous ML detector, retaining the exponential convergence of the probability of error in stationary environments and the insensibility to the step size. We assumed a totally connected network hypothesis to approximate the local tests, and the simulated results showed that even when the network connectivity is very low our method presents very good performance.

6. REFERENCES

[1] A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for internet of things," *Future Generation Computer Systems*, vol. 82, pp. 761–768, 2018.

[2] H. Wu and W. Wang, "A game theory based collaborative security detection method for internet of things systems," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 6, pp. 1432–1445, June 2018.

[3] S. Li, S. Zhao, P. Yang, P. Andriotis, L. Xu, and Q. Sun, "Distributed consensus algorithm for events detection in cyber-physical systems," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2299–2308, Apr. 2019.

[4] X. Wang, G. Li, C. Quan, and P. K. Varshney, "Distributed detection of sparse stochastic signals with quantized measurements: The generalized Gaussian case," *IEEE Transactions on Signal Processing*, vol. 67, no. 18, pp. 4886–4898, Sept. 2019.

[5] S. Marano and A. H. Sayed, "Detection under one-bit messaging over adaptive networks," *IEEE Transactions on Information Theory*, vol. 65, no. 10, pp. 6519–6538, Oct. 2019.

[6] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks using diffusion adaptation," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 1917–1932, May 2011.

[7] S. Al-Sayed, J. Plata-Chaves, M. Muma, M. Moonen, and A. M. Zoubir, "Node-specific diffusion LMS-based distributed detection over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 66, no. 3, pp. 682–697, Feb. 2018.

[8] V. Matta, P. Braca, S. Marano, and A. H. Sayed, "Diffusion-based adaptive distributed detection: Steady-state performance in the slow adaptation regime," *IEEE Transactions on Information Theory*, vol. 62, no. 8, pp. 4710–4732, Aug. 2016.

[9] V. Matta, P. Braca, S. Marano, and A. H. Sayed, "Distributed detection over adaptive networks: Refined asymptotics and the role of connectivity," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 4, pp. 442–460, Dec. 2016.

[10] A. Ainomäe, M. Bengtsson, and T. Trump, "Distributed largest eigenvalue-based spectrum sensing using diffusion LMS," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 2, pp. 362–377, June 2018.

[11] D. Bajovic, D. Jakovetic, J. M. F. Moura, J. Xavier, and B. Sinopoli, "Large deviations performance of consensus+ innovations distributed detection with non-Gaussian observations," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5987–6002, Nov. 2012.

[12] A. K. Sahu and S. Kar, "Recursive distributed detection for composite hypothesis testing: Nonlinear observation models in additive gaussian noise," *IEEE Transactions on Information Theory*, vol. 63, no. 8, pp. 4797–4828, Aug. 2017.

[13] C. A. M. Sotomayor, R. P. David, and R. Sampaio-Neto, "Adaptive nonassisted distributed detection in sensor networks," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 6, pp. 3165–3174, Dec. 2017.

[14] M. R. Leonard and A. M. Zoubir, "Robust sequential detection in distributed sensor networks," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5648–5662, Nov. 2018.

[15] H. R. Ahmadi and A. Vosoughi, "Distributed detection with adaptive topology and nonideal communication channels," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2857–2874, June 2011.

[16] D. Li, S. Kar, and S. Cui, "Distributed quickest detection in sensor networks via two-layer large deviation analysis," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 930–942, Apr. 2018.

[17] A. E. Feitosa, V. H. Nascimento, and C. G. Lopes, "Adaptive detection in distributed networks using maximum likelihood detector," *IEEE Signal Processing Letters*, vol. 25, no. 7, pp. 974–978, July 2018.

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

[19] J.-W. Lee, S.-E. Kim, W.-J. Song, and A. H. Sayed, "Spatio-temporal diffusion strategies for estimation and detection over networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4017–4034, Aug. 2012.

[20] S. M. Kay, *Fundamentals of statistical signal processing: Detection theory, vol. 2*, Prentice Hall Upper Saddle River, NJ, USA, 1998.

[21] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. IPSN 2005. Fourth Int. Symp. Information Processing in Sensor Networks*, Apr. 2005, pp. 63–70.