Diffusion LMS-Based Distributed Detection over Adaptive Networks

Federico Cattivelli Ali H. Sayed

Department of Electrical Engineering
University of California, Los Angeles, CA 90095
Emails: {fcattiv, sayed}@ee.ucla.edu

Abstract— **We consider the problem of distributed detection, where a set of nodes are required to decide between two hypotheses based on their measurements. In diffusion implementations, nodes communicate with their neighbors and no fusion center is needed. In previous work we proposed a distributed detection scheme which was based on diffusion least-squares techniques. In this work we consider the case where nodes utilize diffusion LMS techniques instead. The proposed detector is capable of tracking changes in the active hypothesis. We analyze the performance of the detector, and provide simulation results comparing with other cooperation schemes.**

I. Introduction

We study the problem of distributed detection, where a set of nodes are required to decide between two hypotheses based on their measurements of some physical process. We seek a *fully distributed* implementation, where all nodes make individual decisions by communicating with their immediate neighbors only, and no fusion center is necessary. This scheme provides the network with more flexibility in comparison to a centralized solution, and can be more efficient in terms of communication power and networking resources [1]. Every node in the network will reach a decision. Moreover, our proposed detection algorithm is *adaptive*, in the sense that at every time instant, every node obtains a new measurement, and uses it to obtain a new decision based on the measurements up to that time instant. This makes our algorithm attractive for distributed real-time implementations, since there is no need to wait until a number of measurements are obtained, and more importantly, the algorithm allows tracking of changes in the active hypothesis.

Distributed detection schemes have been proposed before in the literature. The so-called "decentralized" detection schemes require communicating the measurements to a fusion center for processing [2]-[4]. More recently, detection schemes based on average consensus have been proposed, which avoid the use of a fusion center, and where every node in the network makes an individual decision [5]-[8]. Consensus-based schemes assume that all the nodes take a set of measurements, and subsequently run an iterative algorithm to reach consensus. Thus, these algorithms employ two time-scales: one to take the measurements and another to run the consensus algorithm, making them different from our proposed approach.

The proposed distributed detection algorithm is based on our prior work on distributed estimation. Diffusion-based estimation solutions, where nodes communicate with their neighbors in an isotropic manner have been proposed in the context of distributed adaptive filtering, including diffusion LMS [9], [10], [11] and diffusion RLS [12]. In [13] we proposed a distributed detection scheme based on diffusion RLS, which takes advantage of the connection between

This material was based on work supported in part by the National Science Foundation under awards ECS-0601266 and ECS-0725441.

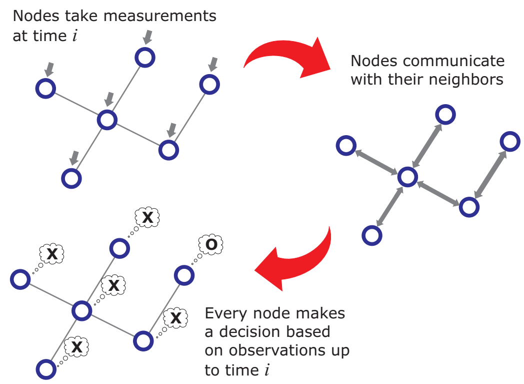

Fig. 1. Distributed detection scheme.

Neyman-Pearson detection and minimum-variance estimation for linear systems in Gaussian noise in order to formulate the detection problem in terms of an estimation problem. In this work, we propose a detector based on diffusion LMS instead. Thus, the solution has lower computational complexity than its RLS counterpart. Moreover, the solution inherits the tracking abilities of the LMS algorithm as we shall see. We provide performance analysis in terms of probabilities of detection and false alarm, and provide simulation results comparing with other techniques, such as the centralized solution and the case where nodes do not cooperate.

II. The Detection Problem

A. Data model

Consider a set of N nodes distributed over some region. We say that two nodes are connected if they can communicate directly with each other. Every node is always connected to itself. The set of nodes connected to node k is called the *neighborhood* of node k, and is denoted by $\mathcal{N}_k$. The number of neighbors of node k including itself is called the *degree* of node k and is denoted by n_k. At every time instant i, every node k takes a scalar measurement $d_k(i)$ of some random process $\boldsymbol{d}_k(i)$, which is related to an unknown vector w^o of size M as follows:

$$\boldsymbol{d}_k(i) = u_{k,i} w^o + \boldsymbol{v}_k(i) \tag{1}$$

where $u_{k,i}$ is a *known* deterministic row vector of size M, and $\boldsymbol{v}_k(i)$ is a scalar zero-mean WSS complex circular Gaussian random process, uncorrelated in time and space, i.e.,

$$\mathrm{E}\,\boldsymbol{v}_k(i)\boldsymbol{v}_l(j) = \delta_{kl}\delta_{ij}\sigma^2_{v_k}$$

The operator E denotes expectation, and δ_{kl} is the Kronecker delta.

The objective is for every node in the network to distinguish between two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, where:

$$w^o = \begin{cases} 0 & \text{under } \mathcal{H}_0 \\ w_1 & \text{under } \mathcal{H}_1 \end{cases}$$

Thus, under $\mathcal{H}_0$, the observations only contain noise, whereas under $\mathcal{H}_1$, the observations contain a signal component.

We collect the data for all nodes $k = 1, \ldots, N$ and for all time instants $j = 0, \ldots, i$ up to time i as follows:

$$\begin{aligned} \mathbf{d}_i &= \text{col}\{\boldsymbol{d}_1(i), \ldots, \boldsymbol{d}_N(i), \boldsymbol{d}_1(i-1), \ldots, \boldsymbol{d}_N(i-1), \\ &\quad \ldots, \boldsymbol{d}_1(0), \ldots, \boldsymbol{d}_N(0)\} \qquad ((i+1)N \times 1) \\ \mathrm{U}_i &= \text{col}\{u_{1,i}, \ldots, u_{N,i}, u_{1,i-1}, \ldots, u_{N,i-1}, \\ &\quad \ldots, u_{1,0}, \ldots, u_{N,0}\} \qquad ((i+1)N \times M) \\ \mathbf{v}_i &= \text{col}\{\boldsymbol{v}_1(i), \ldots, \boldsymbol{v}_N(i), \boldsymbol{v}_1(i-1), \ldots, \boldsymbol{v}_N(i-1), \\ &\quad \ldots, \boldsymbol{v}_1(0), \ldots, \boldsymbol{v}_N(0)\} \qquad ((i+1)N \times 1) \end{aligned}$$

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

$$\mathbf{d}_i = \mathrm{U}_i w^o + \mathbf{v}_i \tag{2}$$

Notice that $\mathbf{v}_i \sim \mathcal{CN}(0, R_{v,i})$, where $R_{v,i} = \mathrm{E}\,\mathbf{v}_i\mathbf{v}_i^*$ is a diagonal matrix, and * denotes complex conjugate transposition. Thus, under $\mathcal{H}_0$, $\mathbf{d}_i \sim \mathcal{CN}(0, R_{v,i})$, whereas under $\mathcal{H}_1$, $\mathbf{d}_i \sim \mathcal{CN}(\mathrm{U}_i w_1, R_{v,i})$.

B. Neyman-Pearson detection

According to the Neyman-Pearson (NP) criterion, the detector that maximizes the probability of detection P_d (i.e., the probability of selecting $\mathcal{H}_1$ when $\mathcal{H}_1$ is true) given a probability of false alarm P_f (i.e., the probability of selecting $\mathcal{H}_1$ when $\mathcal{H}_0$ is true) is [14]:

$$\boldsymbol{T}_i(\mathbf{d}_i) \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \gamma_i$$

where

$$\boxed{\boldsymbol{T}_i(\mathbf{d}_i) \triangleq \alpha_i \text{Re}\{w_1^* \mathrm{U}_i^* R_{v,i}^{-1} \mathbf{d}_i\}} \tag{3}$$

and α_i is any real, positive constant (the value of γ_i will typically depend on the choice of α_i). Then we have (see [13] for details):

$$\boldsymbol{T}_i(\mathbf{d}_i) \sim \begin{cases} \mathcal{N}(0, \sigma_i^2) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(\mu_i, \sigma_i^2) & \text{under } \mathcal{H}_1 \end{cases}$$

where

$$\mu_i = \alpha_i w_1^* \mathrm{U}_i^* R_{v,i}^{-1} \mathrm{U}_i w_1 \qquad \sigma_i^2 = s\alpha_i^2 w_1^* \mathrm{U}_i^* R_{v,i}^{-1} \mathrm{U}_i w_1$$

and $s = 1$ if the vector $\mathbf{d}_i$ is real, and $s = 1/2$ if it is complex. The probabilities of false alarm and detection at time i are given, respectively, by

$$\begin{aligned} P_f &= Q\left(\tfrac{\gamma_i}{\sigma_i}\right) \\ P_d &= Q\left(\tfrac{\gamma_i - \mu_i}{\sigma_i}\right) = Q\left(Q^{-1}(P_f) - \tfrac{\mu_i}{\sigma_i}\right) \end{aligned} \tag{4}$$

Note that given P_f, we can determine $\gamma_i = \sigma_i Q^{-1}(P_f)$, and also, that P_d does not depend on the choice of α_i, and therefore we are free to choose this constant to our convenience.

Under the linear model assumption (2), and the statistical assumptions on the observation noise $\mathbf{v}_i$, we have that the minimum-variance-unbiased (MVU) estimator of w^o given $\mathbf{d}_i$ is given by [15]:

$$\hat{\boldsymbol{w}}_i^{\text{mvu}} = (\mathrm{U}_i^* R_{v,i}^{-1} \mathrm{U}_i)^{-1} \mathrm{U}_i^* R_{v,i}^{-1} \mathbf{d}_i \tag{5}$$

Then, the optimal test statistic (3) can be rewritten in terms of (5) as follows

$$\boxed{\boldsymbol{T}_i(\mathbf{d}_i) = \alpha_i \text{Re}\{w_1^* \mathrm{U}_i^* R_{v,i}^{-1} \mathrm{U}_i \hat{\boldsymbol{w}}_i^{\text{mvu}}\}} \tag{6}$$

III. Distributed Detection

A. Detection with incomplete data

Equation (6) is key for our development, since it indicates how we can calculate the optimal NP test statistic $\boldsymbol{T}_i$ from the MVU estimator $\hat{\boldsymbol{w}}_i^{\text{mvu}}$. Notice that in order to calculate (3) or (6), we need knowledge of the data $\{\boldsymbol{d}_k(j), u_{k,j}\}$ for all nodes k and for all instants j up to time i. Thus, a fusion center would collect all these measurements coming from the different nodes, and compute the optimal NP test statistic. This is the global solution to the problem.

In practice, it may be the case that a certain node only has access to an estimate $\hat{\boldsymbol{w}}_{k,i}$ of w^o which is not necessarily the optimal MVU estimate. The question is how to define a test statistic based on this new estimator, and what will be the resulting probabilities of detection and false alarm. In this work we consider *linear* estimates $\hat{\boldsymbol{w}}_{k,i}$ obtained through the diffusion LMS algorithm.

B. The diffusion LMS algorithm

The diffusion LMS algorithm [9], [11] allows every node in the network to obtain a linear estimate of the unknown parameter w^o from a linear observation model as in (1).

Consider $N \times N$ matrices A and C with non-negative real entries $a_{l,k}$ and $c_{l,k}$, respectively, satisfying

$$a_{l,k} = c_{l,k} = 0 \text{ if } l \notin \mathcal{N}_k \qquad \mathbb{1}^T A = \mathbb{1}^T \quad \mathbb{1}^T C = \mathbb{1}^T$$

The diffusion LMS algorithm obtains for every node k, and for every instant i, an estimate $\hat{w}_{k,i}$ of w^o. The Adapt-Then-Combine (ATC) version of the algorithm is shown below for convenience. Notice that nodes only need to communicate with their neighbors, at every time i, their data $\{d_k(i), u_{k,i}\}$ and vectors $\psi_{k,i}$ of size M.

***ATC Diffusion LMS Algorithm* [10]**

Start with $\hat{w}_{k,-1} = 0$ for every node k. For every time instant $i \geq 0$, repeat

Incremental update: for every node k, repeat

$$\psi_{k,i} = \hat{w}_{k,i-1} + \mu_k \sum_{l \in \mathcal{N}_k} c_{l,k} u_{l,i}^* [d_l(i) - u_{l,i}\hat{w}_{k,i-1}]$$

Diffusion update: for every node k, repeat

$$\hat{w}_{k,i} = \sum_{l \in \mathcal{N}_k} a_{l,k} \psi_{l,i} \tag{7}$$

C. Diffusion LMS-based detection algorithm

Based on (6), we can formulate a distributed detection algorithm that uses the diffusion LMS algorithm (7) to compute $\hat{w}_{k,i}$. We will be interested in test statistics of the form:

$$\boldsymbol{T}_{k,i}(\hat{\boldsymbol{w}}_{k,i}) = \alpha_i \text{Re}\{w_1^* Q_{k,i} \hat{\boldsymbol{w}}_{k,i}\} \tag{8}$$

Notice that (6) is a special case of (8) when $\hat{\boldsymbol{w}}_{k,i}$ is the MVU estimator, and we choose $Q_{k,i} = \mathrm{U}_i^* R_{v,i}^{-1} \mathrm{U}_i$. We will now consider linear estimators of the form:

$$\hat{\boldsymbol{w}}_{k,i} = K_{k,i} \mathbf{d}_i$$

where $K_{k,i}$ is a $M \times (i+1)N$ matrix assumed full rank. The following proposition establishes the optimal choice of $Q_{k,i}$ in (8).

Proposition 1: Consider the observation model (2), and assume that every node k, at time i, obtains a *linear* estimator of w^o, denoted by $\hat{\boldsymbol{w}}_{k,i} = K_{k,i}\mathbf{d}_i$, where $K_{k,i}$ is an $M \times (i+1)N$ full rank matrix. Then, the optimal test statistic for NP detection is given by (8) with:

$$Q_{k,i} = Q_{k,i}^{\text{opt}} \triangleq \mathrm{U}_i^* K_{k,i}^* (K_{k,i} R_{v,i} K_{k,i}^*)^{-1} \tag{9}$$

Proof: The result is obtained by considering now the linear observation model:

$$\hat{\boldsymbol{w}}_{k,i} = (K_{k,i}\mathrm{U}_i)w^o + K_{k,i}\mathbf{v}_i$$

and applying the NP detection theorem for linear models as discussed in Section II-B. ■

Equation (9) gives us an expression for the optimal value of $Q_{k,i}$ for a given linear estimation scheme with estimation matrix $K_{k,i}$. Even though this optimal value of $Q_{k,i}$ will give us the best performance, it may be inefficient in practice to compute the quantities $K_{k,i}\mathrm{U}_i$ and $(K_{k,i}R_{v,i}K_{k,i}^*)^{-1}$ in a distributed manner. Though possible, this will require exchanging matrices between neighboring nodes at every iteration. Thus, we now propose a simpler choice of $Q_{k,i}$ that yields good results as shown in Sec. V.

In order to derive $Q_{k,i}$, and in order to avoid further communications, we will assume for the moment that there is no diffusion process, i.e., $A = I$. In this case, we have:

$$\hat{w}_{k,i} = \hat{w}_{k,i-1} + \mu_k \sum_{l\in\mathcal{N}_k} c_{l,k}u_{l,i}^*[d_l(i) - u_{l,i}\hat{w}_{k,i-1}]$$

Assuming $\hat{w}_{k,-1} = 0$, the above estimator is of the form $\hat{w}_{k,i} = K_{k,i}\mathrm{d}_i$. As shown in Appendix , a good and simple approximation for $Q_{k,i}$ (up to a constant) when the step-size μ_k is small is given by

$$Q_{k,i} \approx I. \tag{10}$$

Combining (8) and (10), we obtain the proposed algorithm shown in (11). It uses the diffusion LMS algorithm (7) to compute $\hat{w}_{k,i}$, and then uses this estimate to compute the test statistic $T_{k,i}$.

Diffusion LMS Detection Algorithm

Start $\hat{w}_{k,-1} = 0$ for all k. For every node k, and for every time instant $i > i_0$, repeat:

Incremental update: for every node k, repeat

$$\psi_{k,i} = \hat{w}_{k,i-1} + \mu_k \sum_{l\in\mathcal{N}_k} c_{l,k}u_{l,i}^*[d_l(i) - u_{l,i}\hat{w}_{k,i-1}]$$

Diffusion update: for every node k, repeat

$$\hat{w}_{k,i} = \sum_{l\in\mathcal{N}_k} a_{l,k}\psi_{l,i}$$

Decision: for every node k, repeat

$$T_{k,i} = \alpha_i \mathrm{Re}\{w_1^*\hat{w}_{k,i}\}$$

$$T_{k,i} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \gamma_{k,i}$$

(11)

Alg. (11) can be specialized to different cooperation schemes of interest. In this work we consider three different cooperation schemes based on the available data at every node: *global*, *no-cooperation* and *diffusion*. The global solution corresponds to the case where all nodes have access to all the data from the network, as in a a fully connected or centralized solution. The no-cooperation solution corresponds to the case where nodes do not communicate with each other, and are isolated. In a diffusion solution (our proposed scheme), nodes exchange measurements and estimates with their neighbors. Alg. (11) can be specialized to each of these cases by appropriately selecting the matrix A and the neighborhood of node k. The choices are summarized in Table I.

Scheme	Choice of $\mathcal{N}_k$ in (11)	Choice of A in (11)
Global	$\{1,\ldots,N\}$	I
No cooperation	$\{k\}$	I
Diffusion	$\mathcal{N}_k$	A

TABLE I

CHOICES OF A AND $\mathcal{N}_k$ FOR DIFFERENT COOPERATION SCHEMES.

IV. PERFORMANCE ANALYSIS

In this section we study the performance of the diffusion LMS-based detector (11). We will now consider the estimates $\hat{\boldsymbol{w}}_{k,i}$ and $\boldsymbol{\psi}_{k,i}$ to be random quantities. We define the error quantities:

$$\tilde{\boldsymbol{\psi}}_{k,i} = \boldsymbol{\psi}_{k,i} - w^o \qquad \tilde{\boldsymbol{w}}_{k,i} = \hat{\boldsymbol{w}}_{k,i} - w^o$$

We also introduce the following extended vectors, obtained by stacking the vectors of every node:

$$\tilde{\boldsymbol{\psi}}_i = \mathrm{col}\{\tilde{\boldsymbol{\psi}}_{1,i},\ldots,\tilde{\boldsymbol{\psi}}_{N,i}\} \qquad \tilde{\boldsymbol{w}}_i = \mathrm{col}\{\tilde{\boldsymbol{w}}_{1,i},\ldots,\tilde{\boldsymbol{w}}_{N,i}\}$$

From the incremental update of Alg. (11), we have:

$$\boldsymbol{\psi}_{k,i} = \hat{\boldsymbol{w}}_{k,i-1} + \mu_k \sum_{l\in\mathcal{N}_k} c_{l,k}u_{l,i}^*[\boldsymbol{d}_l(i) - u_{l,i}\hat{\boldsymbol{w}}_{k,i-1}]$$

and from model (1) we obtain:

$$\tilde{\boldsymbol{\psi}}_{k,i} = \tilde{\boldsymbol{w}}_{k,i-1} + \mu_k \sum_{l\in\mathcal{N}_k} c_{l,k}u_{l,i}^*[\boldsymbol{v}_l(i) - u_{l,i}\tilde{\boldsymbol{w}}_{k,i-1}] \tag{12}$$

We define the block-diagonal matrices

$$\mathcal{D}_i = \mathrm{diag}\left\{\sum_{l\in\mathcal{N}_1} c_{l,k}u_{l,i}^*u_{l,i},\ldots,\sum_{l\in\mathcal{N}_N} c_{l,k}u_{l,i}^*u_{l,i}\right\} \tag{13}$$

$$\mathcal{E}_i = \mathrm{diag}\left\{\sum_{l\in\mathcal{N}_1} b_{l,k}u_{l,i}^*u_{l,i},\ldots,\sum_{l\in\mathcal{N}_N} b_{l,k}u_{l,i}^*u_{l,i}\right\} \tag{14}$$

and the diagonal matrix

$$\mathcal{M} = \mathrm{diag}\{\mu_1 I_M,\ldots,\mu_N I_M\} \tag{15}$$

and the extended weighting matrices

$$\mathcal{A} = A\otimes I_M \qquad \mathcal{C} = C\otimes I_M$$

where $\otimes$ denotes the Kronecker product. Then we can rewrite (12) as:

$$\tilde{\boldsymbol{\psi}}_i = [I - \mathcal{M}\mathcal{D}_i]\tilde{\boldsymbol{w}}_{i-1} + \mathcal{M}\mathcal{C}^T\mathcal{U}_i^*\boldsymbol{v}_i$$

where

$$\mathcal{U}_i = \mathrm{col}\{u_{1,i},\ldots,u_{N,i}\}$$

and

$$\boldsymbol{v}_i = \mathrm{col}\{\boldsymbol{v}_{1,i},\ldots,\boldsymbol{v}_{N,i}\} \qquad \mathrm{E}\,\boldsymbol{v}_i\boldsymbol{v}_i^* = \mathcal{R}_{v,i}$$

From the diffusion update of (11), we have:

$$\tilde{\boldsymbol{w}}_i = \mathcal{A}^T\tilde{\boldsymbol{\psi}}_i$$

and therefore

$$\tilde{\boldsymbol{w}}_i = \mathcal{A}^T[I - \mathcal{M}\mathcal{D}_i]\tilde{\boldsymbol{w}}_{i-1} + \mathcal{A}^T\mathcal{M}\mathcal{C}^T\mathcal{U}_i^*\boldsymbol{v}_i \tag{16}$$

A. Mean performance

Taking expectation of (16) we get

$$\mathrm{E}\,\tilde{\boldsymbol{w}}_i = \mathcal{A}^T[I - \mathcal{M}\mathcal{D}_i]\,\mathrm{E}\,\tilde{\boldsymbol{w}}_{i-1} \tag{17}$$

At time $i=-1$, the initial condition is $\mathrm{E}\,\tilde{\boldsymbol{w}}_{-1} = -\mathbb{1}\otimes w^o$.

Notice that if we start at some time i_0 from an unbiased initial condition (i.e., $\mathrm{E}\,\boldsymbol{w}_{k,i_0} = w^o$ for all k), then Alg. (11) will be unbiased for $i \geq i_0$. Asymptotic unbiased can be obtained in some cases for any initial condition. For instance, Alg. (11) will be asymptotically unbiased if we can find a sub-multiplicative matrix norm $\|\cdot\|$ and a time instant i_0 such that the matrices $\mathcal{A}^T[I-\mathcal{M}\mathcal{D}_i]$ satisfy $\|\mathcal{A}^T[I-\mathcal{M}\mathcal{D}_i]\| \leq \beta < 1$ for all $i > i_0$.

B. Mean-square performance

From (16), and noticing that $\boldsymbol{v}_k(i)$ is uncorrelated with $\hat{\boldsymbol{w}}_{l,i-1}$ for all l, we find that the covariance matrix of the error vector $\tilde{\boldsymbol{w}}_i$ is given by:

$$\boxed{\begin{array}{rcl} R_{\tilde{w}_i} & = & \mathcal{A}^T[I-\mathcal{M}\mathcal{D}_i]R_{\tilde{w}_{i-1}}[I-\mathcal{D}_i^*\mathcal{M}]\mathcal{A} \\ & & +\mathcal{A}^T\mathcal{M}\mathcal{C}^T\mathcal{U}_i^*\mathcal{R}_{v,i}\mathcal{U}_i\mathcal{C}\mathcal{M}\mathcal{A} \end{array}} \tag{18}$$

Note that $R_{\tilde{w}_{k,i}}$, the covariance matrix of $\tilde{\boldsymbol{w}}_{k,i}$, is given by the k-th $M\times M$ diagonal block of $R_{\tilde{w}_i}$. At time $i=-1$, the initial condition for (18) is $R_{\tilde{w}_{-1}} = (\mathbb{1}\otimes w^o)(\mathbb{1}\otimes w^o)^*$.

C. Detection performance

From model (1), we have that the test statistic (8) with $Q_{k,i}=I$ is Gaussian, and distributed according to:

$$\boxed{\boldsymbol{T}_{k,i}(\hat{\boldsymbol{w}}_{k,i}) \sim \mathcal{N}(\mu_{k,i}, \sigma_{k,i}^2)} \tag{19}$$

with mean and variance given, respectively, by:

$$\mu_{k,i} = \alpha_i \mathrm{Re}\{w_1^*\,\mathrm{E}\,\hat{\boldsymbol{w}}_{k,i}\} \tag{20}$$

$$\sigma_{k,i}^2 = s\alpha_i^2 w_1^* R_{\tilde{w}_{k,i}} w_1) \tag{21}$$

where the parameter s is given by $s=1$ if the vector $\mathbf{d}_i$ is real, and $s=1/2$ if it is complex.

Since $\mathrm{E}\,\hat{\boldsymbol{w}}_{k,i} = K_{k,i}\mathrm{U}_i w^o$, it clearly depends on the active hypothesis. For the case $w^o = 0$, we also have $\mathrm{E}\,\hat{\boldsymbol{w}}_{k,i} = 0$. Therefore, the probability of detection at node k and time i is given by

$$P_d = Q\left(\frac{\gamma_{k,i} - \alpha_i\mathrm{Re}\{w_1^*\,\mathrm{E}[\hat{\boldsymbol{w}}_{k,i}|\,\mathcal{H}_1]\}}{\sigma_{k,i}}\right) \tag{22}$$

The probability of false alarm is given by:

$$P_f = Q\left(\frac{\gamma_{k,i}}{\sigma_{k,i}}\right) \tag{23}$$

and the threshold can be computed from $\gamma_{k,i} = \sigma_{k,i}Q^{-1}(P_f)$.

V. Simulations

We now provide simulation results for Alg. (11) and compare with other cooperation schemes. We use a network with $N=20$ nodes and unknown complex vector of size $M=5$. The regressors were drawn according to a complex Gaussian distribution, independent in time and space. The network topology, noise variances and trace of regressor covariances are shown in Fig. 2. For the diffusion algorithm, we used relative-degree weights for A [12] and for C we used weights of the form $c_{l,k} = \beta_k/\sigma_{v_l}^2$, where β_k is a normalizing constant such that $\sum_l c_{l,k} = 1$. The step-size is $\mu_k = 0.05$ in all cases and for all k.

Fig. 3 shows the probability of mis-detection $P_e = 1 - P_d$ for different cooperation schemes, where for every node the threshold

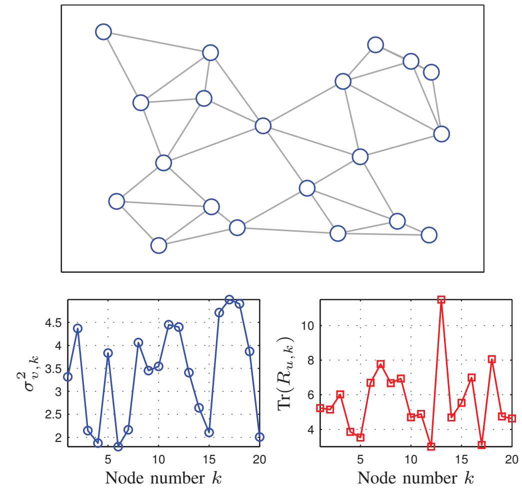

Fig. 2. Network topology (top), noise variances $\sigma_{v,k}^2$ (bottom, left) and trace of regressor covariances $\mathrm{Tr}(R_{u,k})$ (bottom, right).

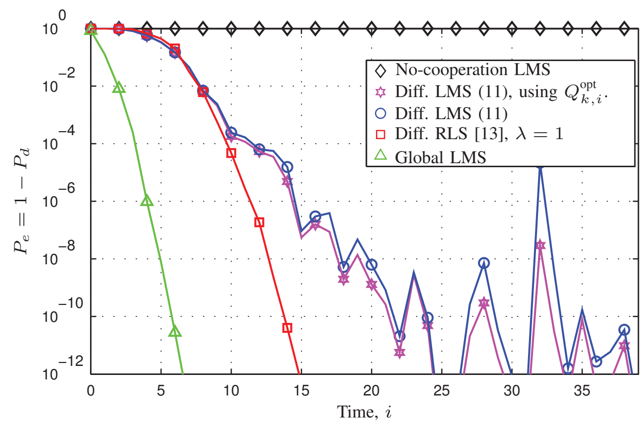

Fig. 3. Probability of error ($P_e = 1 - P_d$) for $P_f = 10^{-9}$.

$\gamma_{k,i}$ is determined in such a way that its probability of false alarm is $P_f = 10^{-9}$. The probability P_e was computed using the provided theoretical expressions, and taking the *maximum* over all nodes at each time instant. We observe that the diffusion-based solution considerably outperforms the case where there is no cooperation. As expected, the global LMS scheme has better performance than diffusion. Also, the proposed algorithm (11) with $Q_{k,i}=I$ has a performance which is very close to the optimal choice of $Q^{\mathrm{opt}}_{k,i}$ in (9). The algorithm also outperforms the diffusion RLS algorithm [13] during the initial stages, though as more measurements are taken into account, the converse is true.

The tracking performance of the proposed algorithm is illustrated in Fig. 4. We show the probability of mis-detection for $P_f = 10^{-9}$, for the case where the active hypothesis changes from $\mathcal{H}_1$ to $\mathcal{H}_0$ at time $i=15$. We observe that the proposed LMS algorithm has better tracking capabilities than the diffusion RLS detection algorithm [13], and also outperforms it when we set a forgetting factor $\lambda = 0.95$.

VI. Conclusions

We proposed a distributed detection algorithm for a binary hypothesis testing problem in Gaussian noise. Our algorithm exploits the connection between detection and estimation and uses the diffusion

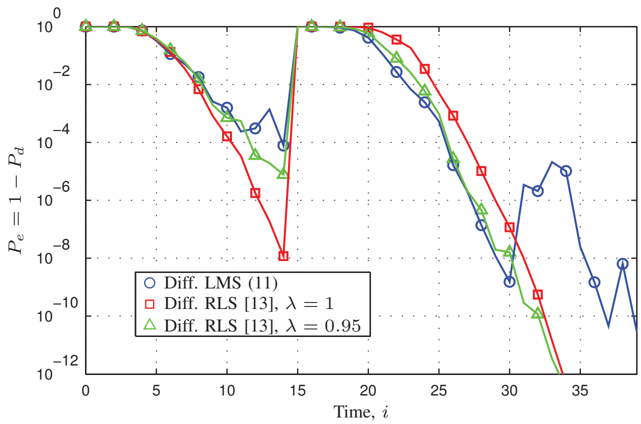

Fig. 4. Probability of error ($P_e = 1 - P_d$) for $P_f = 10^{-9}$, where the active hypothesis changes at time 15.

LMS distributed estimation algorithm. We provided performance analysis and simulations showing that the diffusion algorithm outperforms the case where there is no cooperation and has enhanced tracking capabilities.

Appendix

We proceed to show that for small step-size, we have $Q_{k,i} \approx I$. For the diffusion LMS algorithm with $A = I$ and $\hat{w}_{k,-1} = 0$, we have $\hat{w}_{k,i} = K_{k,i}\mathrm{d}_i$, where

$$K_{k,i} = \mu_k[\mathcal{U}_i^* W_k \quad H_i\mathcal{U}_{i-1}^* W_k \quad H_i H_{i-1}\mathcal{U}_{i-2}^* W_k \quad \ldots \quad H_i H_{i-1}\ldots H_1 \mathcal{U}_0^* W_k]$$

and we defined

$$\mathcal{U}_i = \mathrm{col}\{u_{1,i}, \ldots, u_{N,i}\}$$
$$H_i = I - \mu_k \mathcal{U}_i^* W_k \mathcal{U}_i$$
$$W_k = \mathrm{diag}(Ce_k)$$

For the choice $c_{l,k} = \beta_k/\sigma_{v_l}^2$ we have $W_k = \beta_k R_v^{-1}$ and

$$K_{k,i}\mathrm{U}_i = \mu_k \sum_{j=0}^{i} (H_i \ldots H_{j+1})\mathcal{U}_j^* W_k \mathcal{U}_j$$

$$K_{k,i}R_{v,i}K_{k,i} = \beta_k\mu_k^2 \sum_{j=0}^{i} (H_i \ldots H_{j+1})\mathcal{U}_j^* W_k \mathcal{U}_j (H_i \ldots H_{j+1})^*$$

For small step-size, we can approximate $H_i \ldots H_{j+1} \approx I - \mu_k \sum_{m=j+1}^{i} \mathcal{U}_m^* W_k \mathcal{U}_m$ and

$$K_{k,i}R_{v,i}K_{k,i} = \beta_k\mu_k^2 \Big[K_{k,i}\mathrm{U}_i - \mu_k \sum_{j=0}^{i} (H_i \ldots H_{j+1})\mathcal{U}_j^* W_k \mathcal{U}_j \big(\sum_{m=j+1}^{i} \mathcal{U}_m^* W_k \mathcal{U}_m \big) \Big] \approx \beta_k \mu_k^2 K_{k,i}\mathrm{U}_i$$

Finally,

$$Q_{k,i} = (K_{k,i}\mathrm{U}_i)^* (K_{k,i}R_{v,i}K_{k,i}^*)^{-1} \approx \beta_k^{-1}\mu_k^{-2} I$$

References

[1] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," in *Proc. ICASSP*, Salt Lake City, UT, May 2001, pp. 2033–2036 vol.4.
[2] M. Longo, T. D. Lookabaugh, and R. M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. on Inf. Theory*, vol. 36, no. 2, pp. 241–255, March 1990.
[3] N. Patwari and A. O. Hero, "Hierarchical censoring for distributed detection in wireless sensor networks," in *Proc. IEEE ICASSP*, Hong Kong, April 2003, vol. 4, pp. 848–851.
[4] J. F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. on Signal Processing*, vol. 51, no. 2, pp. 407–416, February 2003.
[5] M. Alanyali, S. Venkatesh, O. Savas, and S. Aeron, "Distributed bayesian hypothesis testing in sensor networks," in *Proc. American Control Conference*, Boston, MA, July 2004, vol. 6, pp. 5369–5374.
[6] R. Olfati-Saber, E. Franco, E. Frazzoli, and J. S. Shamma, "Belief consensus and distributed hypothesis testing in sensor networks," in *Proc. Workshop on Network Embedded Sensing and Control*, South Bend, IN, October 2005.
[7] V. Saligrama, M. Alanyali, and O. Savas, "Distributed detection in sensor networks with packet losses and finite capacity links," *IEEE Trans. on Signal Processing*, vol. 54, no. 11, pp. 4118–4132, November 2006.
[8] S. Kar and J. M. F. Moura, "Consensus based detection in sensor networks: Topology optimization under practical," in *Proc. Workshop Inf. Theory in Sensor Networks*, Santa Fe, NM, June 2007.
[9] 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.
[10] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS algorithms with information exchange," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, October 2008.
[11] F. S. Cattivelli and A. H. Sayed, "Diffusion lms strategies for distributed estimation," *to appear, IEEE Trans. on Signal Processing*, 2009.
[12] F. S. Cattivelli, C. G. Lopes, and A. H. Sayed, "Diffusion recursive least-squares for distributed estimation over adaptive networks," *IEEE Trans. on Signal Processing*, vol. 56, no. 5, pp. 1865–1877, May 2008.
[13] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks based on diffusion estimation schemes," in *Proc. IEEE SPAWC*, Perugia, Italy, June 2009.
[14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. 1, Prentice Hall, NJ, 1998.
[15] A. H. Sayed, *Adaptive Filters*, Wiley, NJ, 2008.