DISTRIBUTED DETECTION OVER ADAPTIVE NETWORKS BASED ON DIFFUSION ESTIMATION SCHEMES

Federico S. Cattivelli *Ali H. Sayed*

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

ABSTRACT

We study the problem of distributed detection, where a set of nodes are required to decide between two hypotheses based on their measurements. We seek fully distributed implementations, where all nodes make individual decisions by communicating with their immediate neighbors, and no fusion center is necessary. This scheme provides the network with more flexibility, saves energy for communication and networking resources. Our distributed detection algorithm is based on a previously proposed distributed estimation algorithm. We establish the connection between the detection and estimation problems, propose a distributed detection algorithm, and analyze the performance of the algorithm in terms of its probabilities of detection and false alarm. We also provide simulation results comparing with other cooperation schemes.

1. 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 *fully distributed* implementations, 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 and is more efficient in terms of communication power and networking resources [1]. Thus, 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 unknown parameter.

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.

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

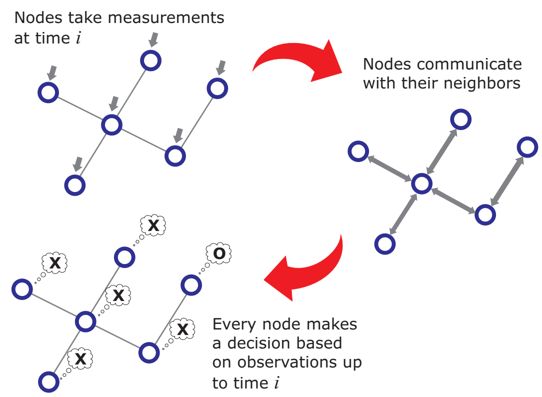

Fig. 1. Distributed detection scheme.

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] and diffusion RLS [11]. In this work, we employ the connection between 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. Then, we use our previously proposed diffusion RLS algorithm [11] to implement the diffusion-based detection algorithm. 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.

2. THE DETECTION PROBLEM

2.1. 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:

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

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 $\mathcal{H}_1$, $\mathbf{d}_i \sim \mathcal{CN}(\mathrm{U}_i w_1, R_{v,i})$.

2.2. 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 [12]:

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

where

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

and α_i is any real, positive constant (the value of γ_i will typically depend on the choice of α_i). Noticing that

$$\alpha_i w_1^* \mathrm{U}_i^* R_{v,i}^{-1}\mathbf{d}_i \sim \mathcal{CN}(\alpha_i w_1^* \mathrm{U}_i^* R_{v,i}^{-1}\mathrm{U}_i w^o, \alpha_i^2 w_1^* \mathrm{U}_i^* R_{v,i}^{-1}\mathrm{U}_i w_1)$$

we have

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

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.

2.3. Relation between NP detector and MVU estimator

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 [13]:

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

Note that the MVU estimator is also the solution to the following weighted least-squares problem:

$$\hat{\boldsymbol{w}}_i^{\text{mvu}} = \arg\min_w \|\mathbf{d}_i - \mathrm{U}_i w\|^2_{R_{v,i}^{-1}} \tag{6}$$

The error covariance matrix of the MVU estimator is

$$R_{\tilde{w}_i^{\text{mvu}}} = \mathrm{E}\,\tilde{\boldsymbol{w}}_i^{\text{mvu}}\tilde{\boldsymbol{w}}_i^{\text{mvu}*} = (\mathrm{U}_i^* R_{v,i}^{-1}\mathrm{U}_i)^{-1}$$

where $\tilde{\boldsymbol{w}}_i^{\text{mvu}} = \hat{\boldsymbol{w}}_i^{\text{mvu}} - w^o$. Now, the optimal test statistic (3) can be rewritten in terms of (5) as follows

$$\boxed{\boldsymbol{T}_i(\mathbf{d}_i) = \alpha_i \mathrm{Re}\{w_1^* \mathrm{U}_i^* R_{v,i}^{-1}\mathrm{U}_i \hat{\boldsymbol{w}}_i^{\text{mvu}}\}} \tag{7}$$

3. DISTRIBUTED DETECTION

3.1. Detection with incomplete data

Equation (7) 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 (7), 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 a subset of the data $\mathbf{d}_i$, and therefore will obtain an estimate of w^o which is not necessarily the global estimate. The question is how to define a test statistic based on this new estimator, and what will be the probabilities of detection and false alarm given this new statistic.

Thus, assume that every node k in the network has access to data

$$\bar{\mathbf{d}}_{k,i} = W_{k,i}\mathbf{d}_i \qquad \bar{\mathrm{U}}_{k,i} = W_{k,i}\mathrm{U}_i \qquad \bar{R}_{v,k,i} = W_{k,i}R_{v,i}W_{k,i}^*$$

where $W_{k,i}$ is some weighting matrix that determines what data is obtained by node k at time i.

Assume now that at time i, node k has knowledge of an *unbiased*, *linear* estimator of w^o, and let us denote this estimator by $\hat{\boldsymbol{w}}_{k,i}$. Based on (7) we can define the following test statistic:

$$\boxed{\boldsymbol{T}_{k,i}(\bar{\mathbf{d}}_{k,i}) = \alpha_{k,i}\mathrm{Re}\{w_1^* \bar{\mathrm{U}}_{k,i}^* \bar{R}_{v,k,i}^{-1}\bar{\mathrm{U}}_{k,i}\hat{\boldsymbol{w}}_{k,i}\}} \tag{8}$$

Thus, $\boldsymbol{T}_{k,i}(\bar{\mathbf{d}}_{k,i})$ is also Gaussian, and distributed according to:

$$\boldsymbol{T}_{k,i}(\bar{\mathbf{d}}_{k,i}) \sim \begin{cases} \mathcal{N}(0, \sigma_{k,i}^2) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(\mu_{k,i}, \sigma_{k,i}^2) & \text{under } \mathcal{H}_1 \end{cases} \tag{9}$$

with

$$\boxed{\begin{aligned}
\mu_{k,i} &= \alpha_{k,i} w_1^* \bar{\mathrm{U}}_{k,i}^* \bar{R}_{v,k,i}^{-1}\bar{\mathrm{U}}_{k,i} w_1 \\
\sigma_{k,i}^2 &= s\alpha_{k,i}^2 w_1^* (\bar{\mathrm{U}}_{k,i}^* \bar{R}_{k,v,i}^{-1}\bar{\mathrm{U}}_{k,i}) R_{\tilde{w}_{k,i}} (\bar{\mathrm{U}}_{k,i}^* R_{v,k,i}^{-1}\bar{\mathrm{U}}_{k,i}) w_1
\end{aligned}}$$

The probabilities of false alarm and detection for node k at time i can be computed from (4) replacing $\{\mu_i, \sigma_i^2\}$ with the above $\{\mu_{k,i}, \sigma_{k,i}^2\}$.

3.2. The diffusion RLS algorithm

The diffusion RLS algorithm from [11] allows every node in the network to estimate the parameter w^o from a linear observation model as in (1) by attempting to solve (6) in a distributed manner. The nodes only need to communicate with their neighbors.

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

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

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

Diffusion RLS Algorithm **[11]**
Start with $\hat{w}_{k,-1} = 0$ and $P_{k,-1} = \Pi^{-1}$ for every node k.
For every time instant i, repeat
Incremental update: for every node k, repeat
$\psi_{k,i} = \hat{w}_{k,i-1}$
$P_{k,i} = \lambda^{-1} P_{k,i-1}$
for all $l \in \mathcal{N}_k$

$$\begin{aligned} \psi_{k,i} &\leftarrow \psi_{k,i} + \tfrac{c_{l,k} P_{k,i} u_{l,i}^* [d_l(i) - u_{l,i}\psi_{k,i}]}{\sigma_{v_l}^2 + c_{l,k} u_{l,i} P_{k,i} u_{l,i}^*} \\ P_{k,i} &\leftarrow P_{k,i} - \tfrac{c_{l,k} P_{k,i} u_{l,i}^* u_{l,i} P_{k,i}}{\sigma_{v_l}^2 + c_{l,k} u_{l,i} P_{k,i} u_{l,i}^*} \end{aligned} \tag{10}$$

end
Spatial update: for every node k, repeat
$\hat{w}_{k,i} = \sum_{l \in \mathcal{N}_k} a_{l,k} \psi_{l,i}$

The algorithm assumes knowledge of the observation noise variances, $\sigma_{v_k}^2$. When these variances are equal across all nodes, we can set $\sigma_{v_k}^2 = 1$ in the diffusion RLS algorithm. The algorithm also uses a forgetting factor λ, which enhances its tracking capabilities. If no forgetting factor is needed, we can select $\lambda = 1$.

3.3. Diffusion-based detection algorithm

Based on (8), we can formulate a distributed detection algorithm that uses the diffusion RLS algorithm (10) to compute $\hat{w}_{k,i}$. Let $\lambda = 1$ and $c_{l,k} = 1/n_k$ if $l \in \mathcal{N}_k$. Since a node has access to data $\{d_l(i), u_{l,i}\}$ from its neighbors $l \in \mathcal{N}_k$, we have

$$P_{k,i}^{-1} = \Pi + \sum_{j=0}^{i} \sum_{l \in \mathcal{N}_k} \frac{u_{l,j}^* u_{l,j}}{n_k \sigma_{v_l}^2} = \Pi + \frac{1}{n_k} \bar{\mathrm{U}}_{k,i}^* \bar{R}_{k,v,i}^{-1} \bar{\mathrm{U}}_{k,i} \triangleq Q_{k,i}$$

Also notice that

$$Q_{k,i} = Q_{k,i-1} + \sum_{l \in \mathcal{N}_k} \frac{u_{l,i}^* u_{l,i}}{n_k \sigma_{v_l}^2}$$

Then, for the choice $\Pi = 0$ and $\alpha_i = 1/[(i+1)n_k]$, we can rewrite (8) as

$$\boldsymbol{T}_{k,i}(\bar{\mathbf{d}}_{k,i}) = \frac{1}{i+1} \mathrm{Re}\{w_1^* Q_{k,i} \hat{\boldsymbol{w}}_{k,i}\}$$

The proposed algorithm is shown in (12). It uses the diffusion RLS algorithm (10) to compute $\hat{w}_{k,i}$, and then uses this estimate to compute the test statistic $T_{k,i}$. It also computes the matrix $Q_{k,i}$ without requiring matrix inversion, except initially to obtain P_{k,i_0-1}.

Alg. (12) can be specialized to different cooperation schemes of interest. In this work we consider four different cooperation schemes based on the available data at every node: *global*, *individual*, *local* 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. This is the best possible scenario. The individual solution corresponds to the case where nodes do not communicate with each other, and are isolated. This corresponds to the worst possible scenario. In a local solution, nodes exchange measurements with their neighbors, but do not perform a diffusion step. Finally, in a diffusion solution (our proposed scheme), nodes exchange measurements with their neighbors as in the local case, but also exchange estimates and perform diffusion. Alg. (12) 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 1. One possible choice for A in the diffusion case is the relative-degree rule [11]:

$$a_{l,k} = \begin{cases} \frac{n_l}{\sum_{l \in \mathcal{N}_k} n_l} & \text{if } l \in \mathcal{N}_k \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Diffusion-based Detection Algorithm
Compute, for every node k,

$$Q_{k,i_0-1} = \sum_{j=0}^{i_0-1} \sum_{l \in \mathcal{N}_k} \frac{u_{l,j}^* u_{l,j}}{n_k \sigma_{v_l}^2} \qquad q_{k,i_0-1} = \sum_{j=0}^{i_0-1} \sum_{l \in \mathcal{N}_k} \frac{u_{l,j}^* d_l(j)}{n_k \sigma_{v_l}^2}$$

until a time instant $i_0 - 1$ such that Q_{k,i_0-1} becomes non-singular. Then, start with $\hat{w}_{k,i_0-1} = Q_{k,i_0-1}^{-1} q_{k,i_0-1}$, $P_{k,i_0-1} = Q_{k,i_0-1}^{-1}$ and for every time instant $i \geq i_0$, repeat
Incremental update: for every node k, repeat
$\psi_{k,i} = \hat{w}_{k,i-1}$
$P_{k,i} = \lambda^{-1} P_{k,i-1}$
$Q_{k,i} = \lambda Q_{k,i-1}$
for all $l \in \mathcal{N}_k$

$$\begin{aligned} \psi_{k,i} &\leftarrow \psi_{k,i} + \tfrac{P_{k,i} u_{l,i}^* [d_l(i) - u_{l,i}\psi_{k,i}]}{n_k \sigma_{v_l}^2 + u_{l,i} P_{k,i} u_{l,i}^*} \\ P_{k,i} &\leftarrow P_{k,i} - \tfrac{P_{k,i} u_{l,i}^* u_{l,i} P_{k,i}}{n_k \sigma_{v_l}^2 + u_{l,i} P_{k,i} u_{l,i}^*} \\ Q_{k,i} &\leftarrow Q_{k,i} + \tfrac{u_{l,i}^* u_{l,i}}{n_k \sigma_{v_l}^2} \end{aligned} \tag{12}$$

end
Spatial update: for every node k, repeat
$\hat{w}_{k,i} = \sum_{l \in \mathcal{N}_k} a_{l,k} \psi_{l,i}$
$T_{k,i} = \frac{1}{i+1} \mathrm{Re}\{w_1^* Q_{k,i} \hat{w}_{k,i}\}$
Decision: for every node k, repeat
$T_{k,i} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \gamma_{k,i}$

Scheme	Choice of $\mathcal{N}_k$ in (12)	Choice of A in (12)
Global	$\{1, \ldots, N\}$	I
Individual	$\{k\}$	I
Local	$\mathcal{N}_k$	I
Diffusion	$\mathcal{N}_k$	A

Table 1. Choices of A and $\mathcal{N}_k$ for different cooperation schemes.

In the following section we analyze the performance of Alg. (12) in terms of its probabilities of detection and false alarm. The diffusion RLS algorithm (10) was analyzed in [11] for the case where the regressors $u_{k,i}$ are random. In this work we provide analysis for the case where regressors are deterministic.

4. PERFORMANCE ANALYSIS

4.1. Mean-performance

In this section we show that Alg. (12) is unbiased, i.e., $\mathrm{E}\,\hat{\boldsymbol{w}}_{k,i} = w^o$ for all k and i. We start by rewriting (see [11]) for $i \geq i_0$:

$$\boldsymbol{\psi}_{k,i} = P_{k,i}\left[P_{k,i-1}^{-1}\hat{\boldsymbol{w}}_{k,i-1} + \sum_{l\in\mathcal{N}_k}\frac{1}{n_k\sigma_{v_l}^2}u_{l,i}^*\boldsymbol{d}_l(i)\right]$$

Taking expectation in the above expression when $i = i_0$, we obtain

$$\mathrm{E}\,\boldsymbol{\psi}_{k,i_0} = P_{k,i_0}\sum_{l\in\mathcal{N}_k}\frac{1}{n_k\sigma_{v_l}^2}u_{l,i_0}^*u_{l,i_0}w^o$$

But since $P_{k,i_0}^{-1} = Q_{k,i_0} = \sum_{l\in\mathcal{N}_k}\frac{1}{n_k\sigma_{v_l}^2}u_{l,i_0}^*u_{l,i_0}$, we have

$$\mathrm{E}\,\boldsymbol{\psi}_{k,i_0} = w^o$$

and therefore $\boldsymbol{\psi}_{k,i_0}$ is unbiased, and so is $\hat{\boldsymbol{w}}_{k,i_0}$ because of the convexity of the coefficients $a_{l,k}$. For $i > i_0$, assume $\hat{\boldsymbol{w}}_{k,i-1}$ is unbiased. Then, we have

$$\mathrm{E}\,\boldsymbol{\psi}_{k,i} = P_{k,i}\left[P_{k,i-1}^{-1}w^o + \sum_{l\in\mathcal{N}_k}\frac{1}{n_k\sigma_{v_l}^2}u_{l,i}^*u_{l,i}w^o\right] = w^o$$

and $\hat{\boldsymbol{w}}_{k,i}$ is also unbiased. By induction, we conclude that $\hat{\boldsymbol{w}}_{k,i}$ is unbiased for all $i \geq i_0$.

4.2. Mean-square performance

Now define the error quantities

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

Then, for $i \geq i_0$ we have:

$$\tilde{\boldsymbol{\psi}}_{k,i} = P_{k,i}\left[P_{k,i-1}^{-1}\tilde{\boldsymbol{w}}_{k,i-1} + \sum_{l\in\mathcal{N}_k}\frac{1}{n_k\sigma_{v_l}^2}u_{l,i}^*\boldsymbol{v}_l(i)\right]$$

and

$$\tilde{\boldsymbol{w}}_{k,i} = \sum_{l\in\mathcal{N}_k}a_{l,k}\tilde{\boldsymbol{\psi}}_{k,i}$$

The above expression can be written more compactly as

$$\tilde{\boldsymbol{w}}_i = \mathcal{A}^T\mathcal{P}_i[\mathcal{P}_{i-1}^{-1}\tilde{\boldsymbol{w}}_{i-1} + \mathcal{C}^T\mathcal{U}_i^*\mathcal{R}_{v,i}^{-1}\boldsymbol{v}_i] \tag{13}$$

where we defined the extended quantities

$$\begin{aligned}
\tilde{\boldsymbol{w}}_i &= \mathrm{col}\{\tilde{\boldsymbol{w}}_{1,i},\ldots,\tilde{\boldsymbol{w}}_{N,i}\}\\
\boldsymbol{v}_i &= \mathrm{col}\{\boldsymbol{v}_1(i),\ldots,\boldsymbol{v}_N(i)\}\\
\mathcal{A} &= A\otimes I_M\\
\mathcal{C} &= C\otimes I_M\\
\mathcal{P}_i &= \mathrm{diag}\{P_{1,i},\ldots,P_{N,i}\}\\
\mathcal{U}_i &= \mathrm{diag}\{u_{1,i},\ldots,u_{N,i}\}\\
\mathcal{R}_{v,i} &= \mathrm{diag}\{\sigma_{v_1}^2,\ldots,\sigma_{v_N}^2\}
\end{aligned}$$

From (13), we find that the covariance matrix of the error vector $\tilde{\boldsymbol{w}}_i$ is given by:

$$\boxed{R_{\tilde{w}_i} = \mathcal{A}^T\mathcal{P}_i\left[\mathcal{P}_{i-1}^{-1}R_{\tilde{w}_{i-1}}\mathcal{P}_{i-1}^{-1} + \mathcal{C}^T\mathcal{U}_i^*\mathcal{R}_{v,i}^{-1}\mathcal{U}_i\mathcal{C}\right]\mathcal{P}_i\mathcal{A}} \tag{14}$$

Note that $R_{\tilde{w}_{k,i}}$ is given by the k-th $M\times M$ diagonal block of $R_{\tilde{w}_i}$. The initial condition is

$$R_{\tilde{w}_{i_0-1}} = \mathcal{P}_{i_0-1}\mathcal{C}^T\left(\sum_{j=0}^{i_0-1}\mathcal{U}_j^*\mathcal{R}_{v,j}^{-1}\mathcal{U}_j\right)\mathcal{C}\mathcal{P}_{i_0-1}$$

4.3. Detection performance

We conclude that the test statistic of node k at time i, given by $\boldsymbol{T}_{k,i} = \frac{1}{i+1}\mathrm{Re}\{w_1^*Q_{k,i}\hat{w}_{k,i}\}$ is distributed according to (9), with

$$\boxed{\begin{aligned}\mu_{k,i} &= \tfrac{1}{i+1}w_1^*Q_{k,i}w_1\\ \sigma_{k,i}^2 &= \tfrac{s}{(i+1)^2}w_1^*Q_{k,i}R_{\tilde{w}_{k,i}}Q_{k,i}w_1\end{aligned}} \tag{15}$$

The probabilities of false alarm and detection are given by:

$$\boxed{\begin{aligned}P_f &= Q\left(\tfrac{\gamma_{k,i}}{\sigma_{k,i}}\right)\\ P_d &= Q\left(\tfrac{\gamma_{k,i}-\mu_{k,i}}{\sigma_{k,i}}\right) = Q\left(Q^{-1}(P_f) - \tfrac{\mu_{k,i}}{\sigma_{k,i}}\right)\end{aligned}} \tag{16}$$

Notice that given P_f, we can compute $\gamma_{k,i} = \sigma_{k,i}Q^{-1}(P_f)$. We summarize our results in the following lemma:

Lemma 1 *For every node k and every time instant i, $\hat{\boldsymbol{w}}_{k,i}$ in Alg. (12) is an unbiased estimator of w_o, with covariance matrix given by the k-th $M\times M$ diagonal block of (14). The test statistic $T_{k,i}$ in Alg. (12) is distributed according to (9) with $\mu_{k,i}$ and $\sigma_{k,i}$ given by (15), and probabilities of false alarm and detection given by (16).*

5. SIMULATIONS

In this section we provide simulation results for Alg. (12) and compare with the global, individual and local 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.

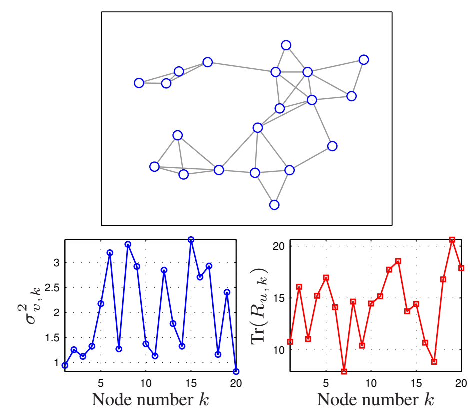

Fig. 2. Network topology (top), noise variances $\sigma_{v,k}^2$ (bottom, left) and trace of regressor covariances $\mathrm{Tr}(R_{u,k})$ (bottom, right).

Figures 3 and 4 show the expectation and variance, respectively, of the test statistic $\boldsymbol{T}_{k,i}$ for node 1 and for different cooperation schemes. We show the theoretical expressions from (15), and a simulation averaging over 200 experiments. It can be observed that the theoretical curves agree with the simulation results. Also notice that the individual algorithm starts at a later time than the rest. This is because it requires a few more samples before the matrix Q_{k,i_0-1} becomes invertible.

Fig. 5 shows the probability of mis-detection $P_e = 1 - P_d$ for different cooperation schemes, where for every node the threshold

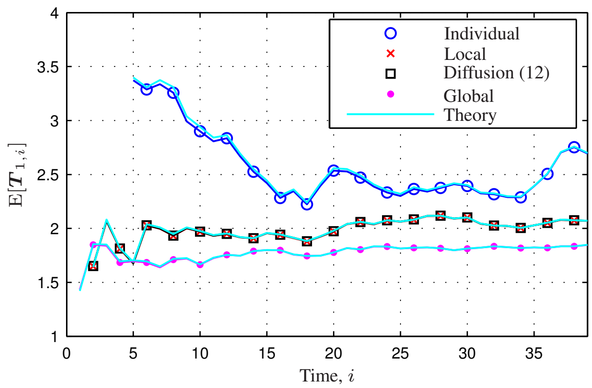

Fig. 3. Expectation of T for node 1, simulation and theory.

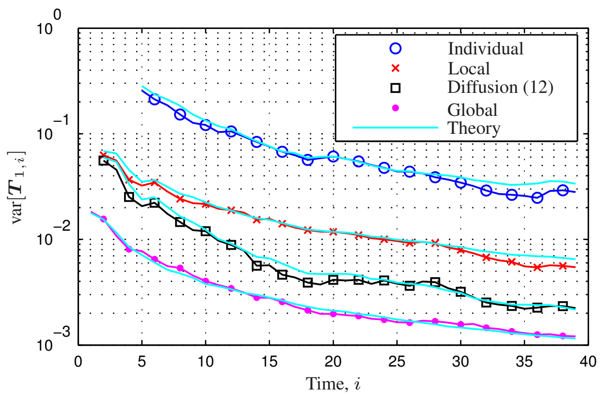

Fig. 4. Variance of T for node 1, simulation and theory.

$\gamma_{k,i}$ is determined in such a way that its probability of false alarm is $P_f = 10^{-9}$. The probabilities were computed using the theoretical expressions for the mean and variance of the test statistic as in (15), 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, and it also outperforms the local solution, showing that the diffusion step improves over using local information only. As expected, the global scheme has better performance than diffusion.

6. 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 theory and uses the diffusion RLS distributed estimation algorithm. We provided performance analysis and simulations showing that the diffusion algorithm outperforms the cases where there is no cooperation and where only local data is used.

7. 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.

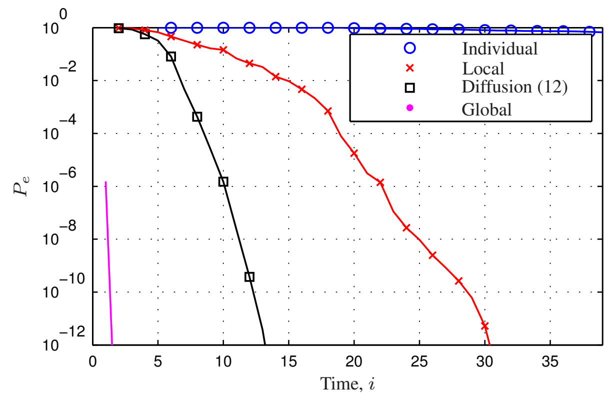

Fig. 5. Probability of error ($P_e = 1 - P_d$) for $P_f = 10^{-9}$.

[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, 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.

[12] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. 1, Prentice Hall, NJ, 1998.

[13] A. H. Sayed, *Adaptive Filters*, Wiley, NJ, 2008.