19th Mediterranean Conference on Control and Automation
Aquis Corfu Holiday Palace, Corfu, Greece
June 20-23, 2011

Consensus Based Distributed Change Detection Using Generalized Likelihood Ratio Methodology

Nemanja Ilić, Srdjan S. Stanković, Miloš S. Stanković and Karl Henrik Johansson

***Abstract*— In this paper a novel distributed recursive algorithm based on the Generalized Likelihood Ratio methodology is proposed for real time change detection using sensor networks. The algorithm is based on a combination of recursively generated local statistics and a global consensus strategy, and does not require any fusion center, so that the state of any node can be tested w.r.t. a given common threshold. Two different problems are discussed: detection of an unknown change in the mean and in the variance of an observed random process. Performance of the algorithm for change detection in the mean is analyzed in the sense of a measure of the error with respect to the corresponding centralized algorithm. The analysis encompasses constant and randomly time varying matrices describing communications in the network. Simulation results illustrate characteristic properties of the algorithms.**

***Index Terms*— Sensor networks, Distributed change detection, Generalized Likelihood Ratio, Consensus, Convergence.**

I. INTRODUCTION

Decision making in large scale systems consisting of multiple decision makers is a very challenging problem in many real world situations, such as air-traffic control, oil exploration, military command and control, electric power networks, control of complex industrial systems, fault detection and isolation, etc. One of typical tasks of *sensor networks*, which is in the focus of many researchers, is *distributed detection* (*e.g.*, [1], [2], [3]). The classical multi-sensor detection schemes require the existence of a *fusion center*, where the final decision is made. Such a topology has been found to be too restrictive for many applications. Distribution of functions has been found to have, in principle, many advantages, consisting of increased reliability, reduced communication bandwidth requirements and reduced overall cost, leading, at the same time, to a certain loss of performance with respect to the optimal centralized system.

There have been some recent attempts to apply *consensus techniques* to the distributed detection problem [4] but they introduce the dynamic agreement process *after all data had been collected*, implying inapplicability to real time change detection problems. In [5], [6], [7] algorithms for distributed real time state and parameter estimation have been proposed by combining local estimation schemes with a dynamic consensus algorithm. An analogous combination of recursive geometric moving average control charts with a consensus algorithm has lead to a novel distributed consensus-based change detection algorithm proposed in [8], [9]. Similar algorithms based on "running consensus" have been proposed and discussed in [10], [11].

In this paper two new algorithms are proposed for *distributed detection* of unknown changes in: a) the mean and b) the variance of a piecewise stationary random process. Both algorithms are derived using the Generalized Likelihood Ratio (GLR) methodology, so that all the nodes in the network can generate local decision variables using the corresponding recursions. By applying a dynamic consensus scheme, one obtains an algorithm which asymptotically provides nearly equal behavior of all the nodes, *i.e.*, *any node* can be selected for testing the decision variable w.r.t. a pre-specified threshold. The algorithm for change detection in the mean is analyzed for both constant and randomly time varying asymmetric consensus matrices characterizing the network. It is shown that the ratio of the mean square error between the proposed decision variables and the centralized decision variable and the mean square value of the centralized decision variable is bounded in the case of constant consensus matrices by $K(1-\alpha)^2$, where $0 < \alpha < 1$ is the forgetting factor of the algorithm, while in the case of random consensus gains it is bounded by $K(1-\alpha)$. A number of simulation results are given as an illustration of the characteristic properties of the proposed algorithm. They also show that the results of the analysis connected to the change in the mean hold also for the detection of the change in the variance.

The outline of the paper is as follows. In Section II local recursive algorithms are derived using the GLR methodology for detecting changes in the mean and in the variance of piecewise stationary random processes. In Section III a novel distributed change detection scheme based on a consensus algorithm is presented. In Section IV analysis of the error with respect to the centralized algorithm is given, while Section V finalizes the paper with some illustrative simulation results.

II. RECURSIVE DETECTION ALGORITHMS

A. Change in the Mean

Assume that we have a sensor network containing n nodes, in which the measurement signal of the i-th node is given by

$$y_i(t) = \theta_i + \epsilon_i(t), \tag{1}$$

N. Ilić, who is the corresponding author, and S. S. Stanković are with the Faculty of Electrical Engineering, University of Belgrade, 11000 Belgrade, Serbia, Phone: 381-11-337-0150, Fax: 381-11-324-8681; E-mail: in085049p@student.etf.rs, stankovic@etf.rs

M. S. Stanković and K. H. Johansson are with the School of Electrical Engineering, Royal Institute of Technology, 100-44 Stockholm, Sweden; E-mail: milsta@kth.se, kallej@kth.se

where $\epsilon_i(t) \sim N(0, \sigma_i^2)$, $i = 1, \ldots, n$, are mutually independent iid processes. We shall assume further that until an unknown time $t = t_0$ we have $\theta_i = \theta_i^0 = 0$ (hypothesis H_0^i), and from $t = t_0$, $\theta_i = \theta_i^1 \neq 0$ (hypothesis H_1^i). In the case when θ_i^1, $i = 1, \ldots, n$, is not a priori known, it is possible to apply the GLR methodology and to obtain the following local statistics based on N successive measurements [12]

$$s_i^l(N) = \max_{\theta_i^1} \sum_{t=1}^{N} \log \frac{p_{\theta_i^1}(y_i(t))}{p_{\theta_i^0}(y_i(t))} = \frac{N}{2} \bar{y}_i(N)^2 \sigma_i^{-2}, \quad (2)$$

where $\bar{y}_i(N) = \frac{1}{N} \sum_{t=1}^{N} y_i(t)$.

Calculation of $s_i^l(N)$ can be performed on-line, recursively. Introducing t for current time, we obtain, using [12], the following basic local recursion for the local decision function

$$s_i^l(t+1) = \frac{t}{t+1} s_i^l(t) + \frac{\sigma_i^{-2}}{t+1} [(t+1)\bar{y}_i(t+1) - \\ -\frac{1}{2} y_i(t+1)] y_i(t+1). \quad (3)$$

After replacing $\frac{t}{t+1}$ by a constant α close to one, which acts as a forgetting factor, and after neglecting the second term in the brackets at the right hand side, the following recursive algorithm is obtained [12]

$$s_i^l(t+1) = \alpha s_i^l(t) + \sigma_i^{-2} \bar{y}_i(t+1) y_i(t+1), \; s_i^l(0) = 0, \quad (4)$$

where $\bar{y}_i$ is also generated recursively by

$$\bar{y}_i(t+1) = \alpha \bar{y}_i(t) + (1-\alpha) y_i(t+1), \quad \bar{y}_i(0) = 0. \quad (5)$$

B. Change in the Variance

Assume, without loss of generality, that we have the following zero-mean system model

$$y_i(t) = \epsilon_i(t), \quad (6)$$

where the hypothesis H_0^i is that $\epsilon_i(t) \sim N(0, (\sigma_i^0)^2)$ and the hypothesis H_1^i that $\epsilon_i(t) \sim N(0, (\sigma_i^1)^2)$; $\{\epsilon_i(t)\}$ under each hypothesis are supposed to be mutually independent iid processes. We shall assume further that until an unknown time $t = t_0$ we have hypothesis H_0^i, and from $t = t_0$ hypothesis H_1^i. In the case when $(\sigma_i^1)^2$ is not a priori known, the application of the GLR methodology leads to the following statistics based on N successive measurements [12]

$$s_i^l(N) = \max_{\sigma_i^1} \sum_{t=1}^{N} \log \frac{p_{\sigma_i^1}(y_i(t))}{p_{\sigma_i^0}(y_i(t))} = \\ = N \log \frac{\sigma_i^0}{\bar{\sigma}_i(N)} + \frac{1}{2(\sigma_i^0)^2} \sum_{t=1}^{N} y_i(t)^2 - \frac{N}{2}, \quad (7)$$

where $\bar{\sigma}_i(N)^2 = \frac{1}{N} \sum_{t=1}^{N} y_i(t)^2$.

Introducing t for current time, we derive, similarly as in (3), the following basic local recursions for calculating $s_i^l(t)$:

$$s_i^l(t+1) = \frac{t}{t+1} s_i^l(t) + \frac{2t+1}{2(t+1)} \log \frac{(\sigma_i^0)^2}{\bar{\sigma}_i(t+1)^2} + \\ + \frac{1}{2} (\frac{t}{t+1} \frac{1}{(\sigma_i^0)^2} - (\frac{t}{t+1})^2 \frac{1}{\bar{\sigma}_i(t+1)^2}) y_i(t+1)^2 + \\ + \frac{1}{2(\sigma_i^0)^2} (\bar{\sigma}_i(t+1)^2 - (\sigma_i^0)^2). \quad (8)$$

For t sufficiently large, we introduce the approximations $\frac{1}{t+1} \ll 1$ and $\frac{t}{t+1} \approx 1$ connected to all the terms at the right hand side except the first (which includes $s_i^l(t)$), and, after replacing $\frac{t}{t+1}$ by α close to 1, we finally obtain the following recursion analogous to (4)

$$s_i^l(t+1) = \alpha s_i^l(t) + \log \frac{(\sigma_i^0)^2}{\bar{\sigma}_i(t+1)^2} + \\ + \frac{1}{2} (\frac{1}{(\sigma_i^0)^2} - \frac{1}{\bar{\sigma}_i(t+1)^2}) y_i(t+1)^2 + \\ + \frac{1}{2(\sigma_i^0)^2} (\bar{\sigma}_i(t+1)^2 - (\sigma_i^0)^2), \quad (9)$$

where $\bar{\sigma}_i(t+1)^2$ is generated recursively by

$$\bar{\sigma}_i(t+1)^2 = \alpha \bar{\sigma}_i(t)^2 + (1-\alpha) y_i(t+1)^2. \quad (10)$$

Denoting the terms at the right hand side of the recursions not including $s_i^l(t)$ in both derived detection algorithms (4) and (9) as $x_i(t+1)$, we obtain the following general form of local recursive algorithms based on the GLR methodology which can be used on-line for change detection purposes

$$s_i^l(t+1) = \alpha s_i^l(t) + x_i(t+1), \; s_i^l(0) = 0. \quad (11)$$

These algorithms obviously belong to the class of geometric moving average control charts [13], [14].

Complexity of the expression for $x_i(t+1)$ in the case of detecting change in the variance (recursively generated $\bar{\sigma}_i(t+1)^2$ in the denominator, correlated with $y_i(t+1)^2$, plus the logarithmic term) makes any analysis regarding statistical properties of $x_i(t)$ very difficult. An analysis connected to the corresponding recursively generated statistics is even more difficult so that properties of the change in the variance detection algorithm will be analyzed in Section V by means of simulation.

One can simplify calculation in the recursions by replacing $x_i(t)$ with $x_i^*(t) = \log \frac{\sigma_i^0}{\bar{\sigma}_i(t)} + \frac{1}{2} (\frac{1}{(\sigma_i^0)^2} - \frac{1}{\bar{\sigma}_i(t)^2}) y_i(t)^2$. As it can be seen from Fig. 1, where mathematical expectations of the two terms are represented w.r.t. σ_i^1 (assuming that α is sufficiently close to 1, so that $\bar{\sigma}_i(t)^2$ has converged to σ_i^1), $x_i^*(t)$ has the same sign as $x_i(t)$, but with smaller ordinates. It will be shown in Section V that the resulting detection scheme is also efficient, provided an adequate threshold is selected.

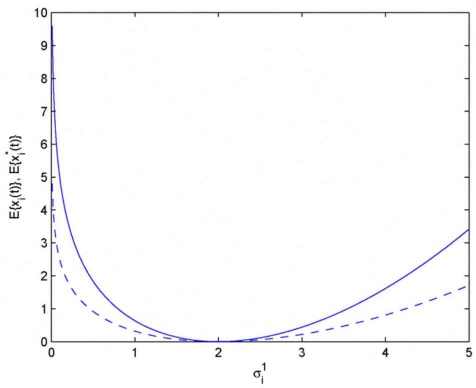

Fig. 1. Mean values of the terms $x_i(t)$ (solid line) and $x_i^*(t)$ (dashed line); value of σ_i^0 is 2.

III. Distributed Change Detection Algorithm

Assuming independence of the local measurements, the global statistics for the whole sensor network are defined as a sum of the statistics given in either (2) or (7). According to the previous section, the centralized decision variable is generated by

$$s_c(t+1) = \alpha s_c(t) + \sum_{i=1}^{n} w_i x_i(t+1), \qquad s_c(0) = 0 \quad (12)$$

where w_i are the weight vector components all equal to 1. For the sake of convenience, we shall adopt that the weights are normalized, leading to $w_i = \frac{1}{n}$. Moreover, we shall generalize this setting by assuming that the weight vector $w = [w_1 \cdots w_n]^T$ in (12) is arbitrarily selected, but satisfying $w_i \geq 0$ and $\sum_{i=1}^{n} w_i = 1$. For example, in the case of (4) we can have $x_i(t) = \bar{y}_i(t)y_i(t)$ instead of $x_i(t) = \sigma_i^{-2}\bar{y}_i(t)y_i(t)$, so that the normalized weights become $w_i = \sigma_i^{-2}(\sum_{i=1}^{n}\sigma_i^{-2})^{-1}$. The global detection procedure is based on testing the decision function $s_c(t)$ with respect to an appropriately chosen threshold $\lambda_c > 0$, so that a change is detected when, *e.g.*, $|s_c(t)| > \lambda_c$. Notice that the algorithm requires a *fusion center*.

The aim of this paper is to propose a distributed change detection algorithm which does not require a fusion center and in which the output of any preselected node can be used as a representative of the whole network and tested w.r.t. a prespecified common threshold. The basic assumption for this algorithm is that the nodes of the network are connected in accordance with an $n \times n$ time varying matrix $C(t) = [c_{ij}(t)]$ satisfying $c_{ij}(t) \geq 0$, $i \neq j$ and $c_{ii}(t) > 0$, $i, j = 1, \ldots, n$, which formally represents the weighted adjacency matrix for the underlying time varying graph representing the network ($c_{ij}(t)$ represents the communication gain from the node j to the node i), and that $C(t)$ is row stochastic for each t. We shall assume, additionally, that matrices $C(t)$ are random, iid and statistically independent from the sequences $x_i(t)$, $i = 1, \ldots, n$. We propose the following algorithm for generating the vector decision function $s(t) = [s_1(t) \cdots s_n(t)]^T$:

$$s(t+1) = \alpha C(t)s(t) + C(t)x(t+1), \qquad s(0) = 0 \quad (13)$$

where $x(t) = [x_1(t) \cdots x_n(t)]^T$. Notice that the consensus matrix $C(t)$ performs for each node "convexification" of the neighboring states and enforces in such a way consensus between the nodes, according to the general principle of consensus schemes. After achieving $s_i(t) \approx s_j(t)$, $i, j = 1, \ldots, n$, change detection can be done by testing $|s_i(t)|$ for any i with respect to the same λ_c as in the case of (12), provided (13) achieves a good approximation of $s_c(t)$ generated by (12).

In order to implement the proposed algorithm it is necessary to set the network communication gains $C(t)$ in accordance with the communication structure constraints resulting from the availability of communication links. We shall assume, in general, that the sequence $\{C(t)\}$, $t = 0, 1, \ldots$, is a sequence of mutually independent identically distributed random matrices independent from the sequence $\{x(t)\}$, such that $C(t)$ is realized at each discrete time instant t as $C^{(k)}$ with probability p_k, $k = 1, \ldots, N$, $N < \infty$, $\sum_{k=1}^{N} p_k = 1$ (the case of constant gains simply follows as a special case). The realization matrices $C^{(k)} = [c_{ij}^{(k)}]$, $k = 1, \ldots, N$, $i, j = 1, \ldots, n$, will be assumed to be constant nonnegative row stochastic matrices, satisfying $c_{ii}^{(k)} > 0$, $i = 1, \ldots, n$, so that we have $\bar{C} = E\{C(t)\} = \sum_{k=1}^{N} C^{(k)}p_k$. This formal setting obviously encompasses the asynchronous asymmetric gossip algorithm with one message at a time, various types of synchronous asymmetric gossip algorithms, as well as communication faults.

We shall assume further that:

A1) $\bar{C}$ has the eigenvalue 1 with algebraic multiplicity 1;

A2) $\lim_{i\to\infty} \bar{C}^i = \mathbf{1}w^T$.

The first assumption is related to the *a priori* given topology of the underlying multi-agent network, implying that the graph associated with $\bar{C}$ has a spanning tree and that $\bar{C}^i$ converges to a nonnegative row stochastic matrix with equal rows when i tends to infinity, *e.g.*, [15], [16]. Assumption A2) establishes a formal connection between the algorithm (13) and the centralized scheme (12), implying that the realization matrices $C^{(k)}$, the corresponding probabilities p_k and the weight vector w are connected by the relation

$$w^T\bar{C} = w^T \sum_{k=1}^{N} C^{(k)} p_k = w^T. \quad (14)$$

For an a priori given vector w, according to the requirements resulting from the selected centralized detector (12), equation (14) should be solved for $C^{(k)}$ and p_k. It is a nonlinear equation, which can be solved in practice by adopting one set of parameters (probabilities p_k, for example) and solving the linear programming problem for the remaining set of parameters (parameters in $C^{(k)}$), or *vice versa* [17]. It is to be emphasized that solving (14) in the case when all $w_i = n^{-1}$ results in symmetric average consensus matrices $\bar{C}$ when the communication links allow such a structure; otherwise, we have an asymmetric $\bar{C}$, satisfying (14). The related literature covers only the symmetric case [10], [18]; the asymmetric case has been treated in [8], [9], [17].

IV. Analysis of the Algorithm

The theoretical analysis given in this section will be concerned with the relationship between the proposed consensus based algorithm (13) and the centralized scheme (12) taken as a reference, assuming that assumption A2) holds. The error vector between the corresponding states is defined as

$$e(t) = s(t) - \mathbf{1}s_c(t), \quad (15)$$

where $\mathbf{1} = [1\cdots 1]^T$. Iterating (13) and (12) back to the zero initial conditions, we get $s(t) = \sum_{i=0}^{t-1}\alpha^i\varphi(t-1, t-i-1)x(t-i)$, where $\varphi(i,j) = C(i)\cdots C(j)$, $i \geq j$, and $s_c(t) = \sum_{i=0}^{t-1}\alpha^i w^T x(t-i)$, wherefrom

$$e(t) = \sum_{i=0}^{t-1}\alpha^i[\varphi(t-1, t-i-1) - \mathbf{1}w^T]x(t-i). \quad (16)$$

From (16) we obtain directly

$$E\{e(t)\} = \sum_{i=0}^{t-1}\alpha^i(\bar{C} - \mathbf{1}w^T)^{i+1}m = \sum_{i=0}^{t-1}\alpha^i\tilde{C}^{i+1}m, \quad (17)$$

where $m = E\{x(t)\}$ and $\tilde{C} = \bar{C} - \mathbf{1}w^T$, having in mind that, under A2), we have $(\bar{C} - \mathbf{1}w^T)^i = \bar{C}^i - \mathbf{1}w^T$. Obviously, $s(t)$ is a biased estimator of $\mathbf{1}s_c(t)$ when $m \neq \mu\mathbf{1}$, where μ is a given scalar.

By assumptions A1) and A2), it follows that $\bar{C}$ and $\mathbf{1}w^T$ have the same eigenvectors. Therefore, $\bar{C}$ has the same eigenvalues as $\tilde{C}$, except for the eigenvalue 1 of $\bar{C}$ which is replaced by the eigenvalue 0 of $\tilde{C}$. Having in mind that $c_{ii} > 0$, $i = 1, \ldots, n$, it follows that the modules of all the eigenvalues of $\tilde{C}$ are strictly less than 1 [15]. We denote $|\lambda(\tilde{C})|_{max} = \lambda_M < 1$. Now we can see that

$$\|E\{e(t)\}\| \leq \sum_{i=0}^{t-1}\alpha^i\|\tilde{C}^{i+1}\|\|m\| \leq \frac{k\lambda_M\|m\|}{1-\alpha\lambda_M} < \frac{k\lambda_M\|m\|}{1-\lambda_M}, \quad (18)$$

having in mind that $\|\tilde{C}^i\| \leq k\lambda_M^t$ for any matrix norm, where k is an appropriately chosen constant, and that $\lambda_M < 1$. A comparison with the properties of an analogous algorithm presented in [9] shows that in the case of (13) the upper limit of $\|E\{e(t)\}\|$ does not depend on α.

However, the obtained quality of approximating the centralized solution can be more adequately expressed by a normalized criterion, defined as the ratio of the norm of mathematical expectation of the error (18) and mathematical expectation of the centralized decision variable. In this case we readily obtain that

$$\frac{\|E\{e(t)\}\|}{E\{s_c(t)\}} \leq K(1-\alpha). \quad (19)$$

where $K < \infty$, having in mind that $E\{s_c(t)\} = w^T m\frac{1-\alpha^t}{1-\alpha}$.

A more complete insight into the quality of approximation can be obtained from an analysis of the mean square error matrix

$$Q(t) = E\{e(t)e(t)^T\}. \quad (20)$$

The following lemma serves as a prerequisite.

Lemma 1. The covariance function $r_i(\tau) = E\{(x_i(t) - m_i)(x_i(t+\tau) - m_i)\}$ for algorithm (4) satisfies

$$\sum_{\tau=0}^{\infty}|r_i(\tau)| \leq K_1; \quad i = 1, \ldots, n, \ \ 0 < K_1 < \infty. \quad (21)$$

Proof: For (4) we have $x_i(t) = \sigma_i^{-2}\bar{y}_i(t)y_i(t)$, $\bar{y}_i(t) = (1-\alpha)\sum_{j=0}^{t-1}\alpha^j y(t-i)$ and $y_i(t) = \theta_i + \epsilon_i(t)$), so that

$$r_i(\tau) = E\{(\tfrac{\bar{y}_i(t)y_i(t)}{\sigma_i^2} - m_i)(\tfrac{\bar{y}_i(t+\tau)y_i(t+\tau)}{\sigma_i^2} - m_i)\} =$$
$$= E\{\tfrac{1-\alpha}{\sigma_i^2}\textstyle\sum_{j=0}^{t-1}\alpha^j\theta_i(\epsilon_i(t) + \epsilon_i(t-j))$$
$$\tfrac{1-\alpha}{\sigma_i^2}\textstyle\sum_{k=0}^{t+\tau-1}\alpha^k\theta_i(\epsilon_i(t+\tau) + \epsilon_i(t+\tau-k))\}. \quad (22)$$

Since $r_i(\tau) = r_i(-\tau)$ we can see that for $\tau > 0$ we have non-zero terms in (22) in the cases when $k = \tau$ and $k = \tau + j$, so that we obtain the following expression for $r_i(\tau)$

$$r_i(\tau) = \tfrac{(1-\alpha)^2}{\sigma_i^4}E\{\textstyle\sum_{j=0}^{t-1}\alpha^j\theta_i^2\alpha^\tau(\epsilon_i^2(t) +$$
$$+\alpha^j\epsilon_i^2(t-j))\} = (1-\alpha)^2\tfrac{\theta_i^2}{\sigma_i^2}(\tfrac{1-\alpha^t}{1-\alpha} + \tfrac{1-\alpha^{2t}}{1-\alpha^2})\alpha^\tau. \quad (23)$$

Therefore, (21) is satisfied, since $r_i(\tau) \sim (1-\alpha)\alpha^\tau$. Q.E.D.

Theorem 1. Let assumptions A1) and A2) hold. Then, under hypothesis H_1, in the case of constant consensus matrices

$$\frac{\|Q(t)\|_\infty}{E\{s_c(t)^2\}} \leq K(1-\alpha)^2,$$

while in the case of random consensus matrices

$$\frac{\|Q(t)\|_\infty}{E\{s_c(t)^2\}} \leq K(1-\alpha),$$

where $K < \infty$ is a constant that does not depend on α and $\|A\|_\infty = \max_i\sum_j|a_{ij}|$, where $A = [a_{ij}]$ is a given matrix.

Proof: First we obtain a lower bound for the variance of the centralized statistics:

$$\mathrm{var}\{s_c(t)\} =$$
$$= E\{(\textstyle\sum_{j=0}^{t-1}\alpha^j w^T x(t-j) - \tfrac{w^T m}{1-\alpha})^2\} =$$
$$= E\{(\textstyle\sum_{j=0}^{t-1}\alpha^j w^T(x(t-j) - m))^2\} \geq$$
$$\geq E\{\textstyle\sum_{j=0}^{t-1}(\alpha^j w^T(x(t-j) - m))^2\} =$$
$$= \textstyle\sum_{j=0}^{t-1}\alpha^{2j}w^T\mathrm{diag}\{\mathrm{var}(x_i)\}w =$$
$$= \tfrac{1-\alpha^{2t}}{1-\alpha^2}w^T\mathrm{diag}\{\mathrm{var}(x_i)\}w \geq m_1(1-\alpha)^{-1}, \quad (24)$$

where $m_1 < \infty$ does not depend on α. Having in mind that $E\{s_c(t)\} = w^T m\frac{1-\alpha^t}{1-\alpha}$ we obtain

$$E\{s_c(t)^2\} = \mathrm{var}\{s_c(t)\} + E\{s_c(t)\}^2 \geq m_2(1-\alpha)^{-2}, \quad (25)$$

where $m_2 < \infty$ does not depend on α.

Further, consider an arbitrary deterministic n-vector y and analyze the quadratic form $y^T Q(t)y$.

In the case of constant consensus gains we have that $Q(t) = Q_1(t) + Q_2(t)$, in which

$$Q_1(t) = \Phi(t)^T\tilde{R}(t)\Phi(t) \quad (26)$$

and

$$Q_2(t) = \Phi(t)^T m_X(t)m_X(t)^T\Phi(t), \quad (27)$$

978-1-4577-0123-8/11/$26.00 ©2011 IEEE

where $\Phi(t) = [\alpha^{t-1}\tilde{C}^t \vdots \alpha^{t-2}\tilde{C}^{t-1} \vdots \cdots \vdots \alpha^0\tilde{C}]^T$, $\tilde{R}(t) = R(t) - m_X(t)m_X(t)^T$, $R(t) = E\{X(t)X(t)^T\}$, $X(t) = [x(1)^T \cdots x(t)^T]^T$ and $m_X(t) = E\{X(t)\}$. Following the results in [8] it can be easily shown that

$$y^T Q_1(t) y \leq k_1 \|y\|^2, \tag{28}$$

where $k_1 < \infty$ does not depend on α, while analyzing $Q_2(t)$ we find that

$$\begin{aligned} y^T Q_2(t) y \leq (\textstyle\sum_{i=0}^{t-1} \alpha^i \|\tilde{C}^{i+1}\| \|m\|)^2 \|y\|^2 \leq \\ \leq k' (\tfrac{\lambda_m}{1-\lambda_m})^2 \|y\|^2 \leq k_2 \|y\|^2, \end{aligned} \tag{29}$$

where $k_2 < \infty$ does not depend on α, having in mind that $\|\tilde{C}^i\| \leq k\lambda_m^i$, where $k < \infty$ and that modules of all the eigenvalues of $\tilde{C}$ are strictly less than 1.

In the case of random consensus gains the mean square error matrix is decomposed as $Q(t) = Q_3(t) + Q_4(t)$, where

$$Q_3(t) = E\{E_x\{e(t)e(t)^T\} - E_x\{e(t)\}E_x\{e(t)\}^T\} \tag{30}$$

and

$$Q_4(t) = E\{E_x\{e(t)\}E_x\{e(t)\}^T\}\}, \tag{31}$$

$E_x\{.\}$ denoting the conditional expectation given the σ-algebra generated by $\{C(t)\}$.

We obtain, in analogy with (26) and (27), that

$$Q_3(t) = E\{\tilde{\Phi}(t)^T \tilde{R}(t) \tilde{\Phi}(t)\}, \tag{32}$$

where $\tilde{\Phi}(t) = [\alpha^{t-1}(\varphi(t-1,0) - \mathbf{1}w^T) \vdots\ \alpha^{t-2}(\varphi(t-1,1) - \mathbf{1}w^T) \vdots \cdots \vdots\ \alpha^0(\varphi(t-1,t-1) - \mathbf{1}w^T)]^T$ and

$$Q_4(t) = E\{\tilde{\Phi}(t)^T m_X(t) m_X(t)^T \tilde{\Phi}(t)\}\}. \tag{33}$$

Following the results in [9] it can be shown that

$$y^T Q_3(t) y \leq k_3 \|y\|^2 (1-\alpha)^{-1}, \tag{34}$$

where $k_3 < \infty$ does not depend on α, while the term $y^T Q_4(t) y$ can be analyzed analogously. We use the fact that $E\{\tilde{\Phi}(t)^T m_X(t) m_X(t)^T \tilde{\Phi}(t)\} \leq 2\alpha^{2(t-1)} E\{(\varphi(t-1,0) - \mathbf{1}w^T) m m^T (\varphi(t-1,0) - \mathbf{1}w^T)^T\} + \ldots + 2\alpha^{2 \cdot 0} E\{(\varphi(t-1,t-1) - \mathbf{1}w^T) m m^T (\varphi(t-1,t-1) - \mathbf{1}w^T)^T\}$ and obtain that

$$y^T Q_4(t) y \leq m' \|y\|^2 \sum_{i=0}^{t-1} \alpha^{2i} \leq k_4 \|y\|^2 (1-\alpha)^{-1}, \tag{35}$$

where $k_4 < \infty$ does not depend on α, based on the result from [9] that norm of the matrices $D(t-1,j) = E\{(\varphi(t-1,j) - \mathbf{1}w^T)(\varphi(t-1,j) - \mathbf{1}w^T)^T\}$, where $j = 0, \ldots, t-1$ has a finite upper bound that does not depend on α.

Consequently, by choosing $y = e_i$, where e_i denotes the n-vector of zeros with only the i-th entry equal to one, one obtains that in the case of constant consensus gains $Q_{ii}(t) \leq k_{12}$, where $k_{12} < \infty$, $i = 1, \ldots, n$. Furthermore, $|Q_{ij}(t)| \leq \max_i Q_{ii}(t)$, having in mind elementary properties of positive semidefinite matrices. In the case of random consensus matrices, we have that $\max_{i,j} Q_{ij}(t) \leq k_{34} \frac{1}{1-\alpha}$, where $k_{34} < \infty$. Thus, dividing the mean square error matrices with the mean square value of the centralized decision variable (25) gives the result. Q.E.D.

V. Simulation Results

Change in the mean. Let us consider a sensor network with $n = 10$ nodes, where the means θ_i^1 (unknown to the designer of the detection scheme) are randomly taken from the interval $(0, 1]$ and the variances σ_i^2 randomly taken from the interval $[0.5, 1.5]$; it is assumed that $\theta_i^0 = 0$ in the case of no change, $i = 1, \ldots, n$. Communication gains are obtained by solving the equation (14) for both constant and time varying cases under the constraints that the consensus matrices are row stochastic and possess a predefined structure (places of zeros). The assumed network topology corresponds to the Geometric Random Graphs in which the nodes represent, *e.g.*, randomly spatially distributed agents within a square area and they are connected if their distance is less than some predetermined threshold (in this case half of the side of the square, see, *e.g.*, [18]). The weight vector components are chosen as $w_i = \sigma_i^{-2}(\sum_{i=1}^n \sigma_i^{-2})^{-1}$ (see Section III). In the case of random consensus gains the asymmetric asynchronous "gossip" algorithm with one communication at a time is assumed; values of the elements of the realizations of the consensus matrices corresponding to communicating nodes are taken to be 0.5, and (14) is solved for the probabilities of individual realizations, see [17]. Fig. 2 contains, for comparison, one typical realization of the centralized decision function (12) for $\alpha = 0.9$ and $\alpha = 0.99$, together with the corresponding realizations obtained at one randomly selected node in the network for constant and random consensus gains (one component of (13)). The moment of change is chosen to be $t = 200$. Fig. 3 illustrates the dependence of the error on the forgetting factor α, according to Section IV. For the above network with 10 nodes, the ratio of the mean square error for one randomly selected node and the mean square value of the centralized statistics at t=1000 is calculated using 1000 Monte Carlo runs as a function of $(1-\alpha)^2$ in the case of constant consensus gains and of $(1-\alpha)$ in the case of random consensus gains. The results of Theorem 1 are clearly justified.

Change in the variance. A similar network as the one described above is used, where $(\sigma_i^1)^2$ (unknown to the designer of the detection scheme) are randomly taken from the interval $(0.5, 1]$ and $(\sigma_i^0)^2$ randomly taken from the interval $(0, 0.5]$. Communication gains are obtained by solving the equation (14) similarly as above, with the weight vector components $w_i = \frac{1}{n}$. Fig. 4 contains, for comparison, the mean $\pm$ one standard deviation of the centralized decision function for $\alpha = 0.99$, together with the corresponding decentralized statistics obtained at one randomly selected node in the network for constant and random consensus gains and for two different choices of the terms used in recursive schemes ($x_i(t)$ and $x_i^*(t)$, see Section II). As can be seen, both mentioned choices can be used, since they result in statistics with similar behavior. In addition, similarly as above, analysis of the error is given in Fig. 3 (dashed lines), confirming that all the theoretical results from Section IV connected to the change in the mean hold also for the detection algorithm of the change in the variance. Fig. 2

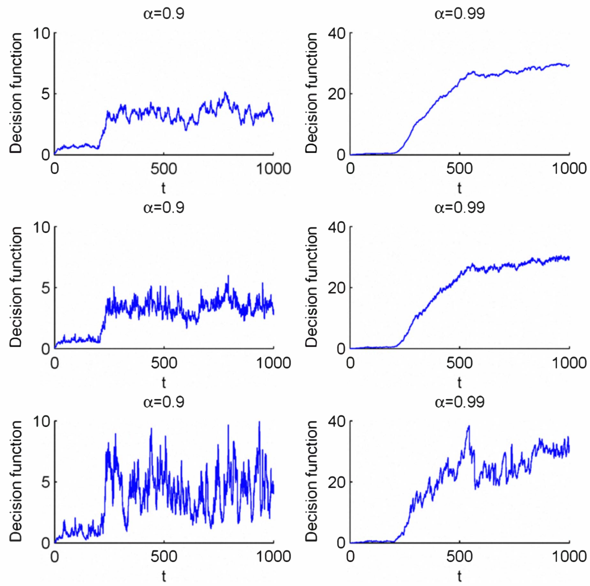

Fig. 2. Realizations of decision functions: centralized strategy (top), constant consensus gains (middle), random consensus gains (bottom)

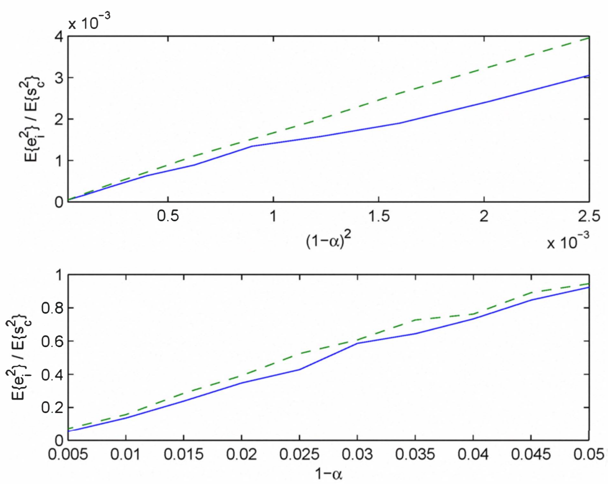

Fig. 3. Ratio of the mean square error and mean square value of the centralized statistics: constant consensus gains (top), random C (bottom); change in the mean (solid line), change in the variance (dashed line)

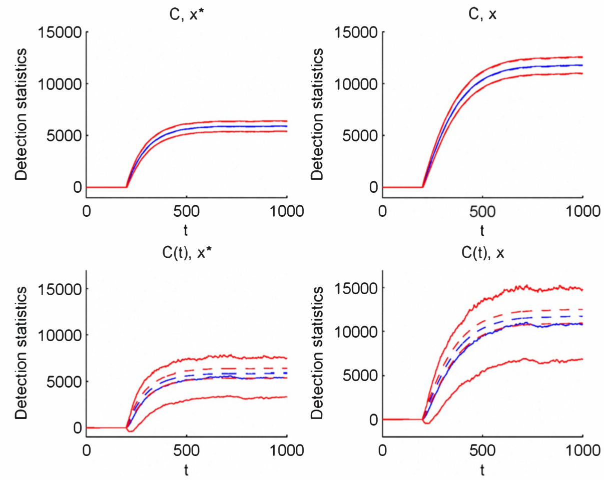

Fig. 4. Means $\pm$ one standard deviation for decision functions: centralized strategy (dashed lines), proposed algorithm (solid lines); constant consensus gains (top), random C (bottom); used term $x_i^*(t)$ (left), $x_i(t)$ (right)

and Fig. 4 also show that the detection efficiency of our distributed detection scheme is satisfactory even in the case of random consensus gains since the mean square value of the statistics is very small under the hypothesis H_0 (it can be shown that it is $\frac{K}{1-\alpha}$ times smaller than corresponding ones under H_1).

References

[1] J. F. Chamberland and V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Proc.*, vol. 50, pp. 407–416, 2003.

[2] R. Vishwanathan and P. Varshney, "Distributed detection with multiple sensors: Part i - fundamentals," *Proc. of the IEEE*, vol. 85, pp. 54–63, 1997.

[3] A. Speranzon, C. Fischione, and K. H. Johansson, "Distributed and collaborative estimation over wireless sensor networks," in *Proc. IEEE Conf. on Decision and Control*, 2006.

[4] E. Franco, R. Olfati-Saber, T. Parisini, and M. M. Polycarpou, "Distributed fault diagnosis using sensor networks and consensus based filters," in *Proc. 45th IEEE CDC Conf.*, 2006.

[5] S. S. Stanković, M. S. Stanković, and D. M. Stipanović, "Consensus based overlapping decentralized estimator," *IEEE Trans. Autom. Control*, vol. 54, pp. 410–415, 2009.

[6] ——, "Consensus based overlapping decentralized estimation with missing observations and communication faults," *Automatica*, vol. 45, pp. 1397–1406, 2009.

[7] ——, "Decentralized parameter estimation by consensus based stochastic approximation," in *Proc. 46th IEEE Conference on Decision and Control*, 2007, pp. 1535–1540.

[8] S. S. Stanković, N. Ilić, M. S. Stanković, and K. H. Johansson, "Distributed change detection based on a consensus algorithm," in *Proc. 2nd IFAC Workshop on Distr. Estim. Contr. Networked Systems*, 2010.

[9] ——, "Distributed change detection based on a randomized consensus algorithm," in *Proc. IEEE Conf. Circuits and Systems for Communications*, 2010.

[10] P. Braca, S. Marano, and V. Matta, "Enforcing consensus while monitoring the environment in wireless sensor networks," *IEEE Trans. Signal Processing*, vol. 56, pp. 3375–3380, 2008.

[11] P. Braca, S. Marano, V. Matta, and P. Willett, "Asymptotic optimality of running consensus in testing binary hypotheses," *IEEE Trans. Signal Processing*, vol. 58, pp. 814–825, 2010.

[12] S. X. Ding, *Model Based Fault Diagnosis Techniques - Design Schemes, Algorithms and Tools*. Springer Verlag, 2008.

[13] M. Basseville and L. V. Nikiforov, *Detection of Abrupt Changes: Theory and Applications*. Prentice Hall, 1993.

[14] F. Gustafsson, *Adaptive Filtering and Change Detection*. John Wiley & Sons Ltd, 2000.

[15] W. Ren and R. Beard, "Consensus seeking in multi-agent systems using dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, pp. 655–661, 2005.

[16] R. Olfati-Saber, A. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, pp. 215–233, 2007.

[17] N. Ilić and S. S. Stanković, "Communication gains design in a consensus based distributed change detection algorithm," in *Proc. 8th European Workshop on Advanced Control and Diagnosis*, 2010.

[18] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inf. Theory*, vol. 52, pp. 2508–2530, 2006.