# Scalable Monitoring of Interconnected Stochastic Systems

Francesca Boem, Ruggero Carli, Marcello Farina, Giancarlo Ferrari-Trecate, Thomas Parisini

*Abstract*— In this paper, we propose a novel distributed fault detection method to monitor the state of a linear system, partitioned into interconnected subsystems. The approach hinges on the definition of a partition-based distributed Luenberger estimator, based on the local model of the subsystems and that takes into account the dynamic coupling terms between the subsystems. The proposed methodology computes – *in a distributed way* – a bound on the variance of a properly defined residual signal, considering the uncertainty related to the state estimates performed by the neighboring subsystems. This bound allows the computation of suitable local thresholds with guaranteed maximum false-alarms rate. The implementation of the proposed estimation and fault detection method is scalable, allowing *Plug & Play* operations and the possibility to disconnect the faulty subsystem after fault detection. Theoretical conditions guaranteeing the convergence of the estimates and of the bounds are provided. Simulation results show the effectiveness of the proposed method.

## I. INTRODUCTION

In recent years, the research dealing with design of systems which are reliable and robust with respect to uncertainties, changing environment and communication failures has grown in importance, especially in relation with distributed control and monitoring of large-scale and networked systems [1]. In this respect, the distributed state estimation problem is certainly central.

The problem dealt with in the paper consists in estimating the state of a large-scale system, characterized by interconnected subsystems, and taking decisions about the health status of the system, using a network of local diagnosers equipped with sensing, communication and computation capabilities. Differently from many contributions in the literature, where the full state of the system is estimated by all subsystems, (e.g., techniques based on consensus and diffusion strategies [2], [3]), here each local diagnosis unit only estimates a part of the global state vector. More specifically, by using a *partition-based* estimation technique and exchanging information with the diagnosers of neighboring

F. Boem is with the Dept. of Electrical and Electronic Engineering at the Imperial College London, UK. (f.boem@imperial.ac.uk)

R. Carli is with the Dip. di Ing. dell'Informazione, Università di Padova, Italy. (carlirug@dei.unipd.it)

M. Farina is with the Dip. di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Italy. (marcello.farina@polimi.it)

G. Ferrari-Trecate is with the Automatic Control Laboratory, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. (giancarlo.ferraritrecate@epfl.ch)

T. Parisini is with the Dept. of Electrical and Electronic Engineering at the Imperial College London, UK, and also with the Dept. of Engineering and Architecture at University of Trieste, Italy. (t.parisini@gmail.com)

subsystems, each local diagnoser monitors in a distributed way the state of the associated local subsystem only.

Recently, several different partition-based approaches have been proposed: for example, [4], [5], [6] propose Kalman-filter-based estimation schemes for discrete-time systems affected by stochastic noise, while [7], [8] assume that the system is affected by bounded noise. In this paper, we consider linear discrete-time systems affected by stochastic noises. We adapt the partition-based distributed estimation method introduced in [9] in order to compute an estimate of the local state, proposing a different definition of the correction gains. We then locally compute an upper bound for the covariance matrix of the estimation error and we use this bound to derive a suitable fault detection threshold for a local residual signal, *aiming at guaranteeing a maximum false-alarms rate*.

In the past few years, quite a few distributed model-based fault detection schemes have been proposed based on observers (see, for instance, [10], [11], [12], [13], [14], where process and sensor faults can be detected by means of a group of local detection agents). These methods usually consider deterministic bounds for noises and uncertainties in order to suitably determine detection thresholds. On the other hand, here, similarly to [15], [16], we consider a stochastic characterization of the noises and the definition of time-varying bounds that guarantee probabilistic performance. While [16] proposes a sensor network to monitor a system characterized by stochastic uncertainties, where each sensor takes noisy measurements of the entire state, in this paper only a part of the state is considered by each diagnoser.

Moreover, the proposed method is scalable and allows the unplugging of faulty subsystems in order to avoid or reduce the propagation of faults in the interconnected large-scale system. Once the issue has been solved, the disconnected subsystem can be re-connected to the network of subsystems. The reconfiguration process involves only communication with neighboring subsystems, in a scalable architecture. With respect to [15], where a Plug & Play (PnP) fault detection architecture is proposed dealing with nonlinear systems and possibly overlapping decompositions, here we do not assume to know the mean and variance of the coupling uncertainty as a given element of the problem, but we are able to compute a bound for the influence on the uncertainty of the neighboring estimates. Furthermore, in the present paper we remove the assumption, used in [15], that the state is fully measurable.

To sum up, the main contributions of the paper are (i) the design of a distributed estimation and fault detection scheme able to consider the dynamics of coupling terms between subsystems; (ii) a recursive equation for computing, in a

distributed fashion, an upper bound on the true covariance matrix of the estimation error, allowing the design of a distributed detection threshold that guarantees a maximum probability of false alarms at each time step and (iii) a proof that this upper bound converges.

The design and the online implementation of the proposed estimation scheme involve only transmission of a limited amount of data among neighboring subsystems. This enables PnP operations, meaning that, when a new subsystems issues a plug-in request, (i) the possibility of adding it without spoiling convergence of the estimation scheme is automatically checked and (ii) only subsystems that are at most two-hops-away from the new unit need to update local estimators and fault detectors.

The paper is structured as follows. In Section II, we introduce the distributed estimation and fault detection method. In Section III we propose a bound for the estimation error covariance matrix and we provide some convergence conditions. The scalability features are analyzed in Section IV. Finally, simulation results are presented in Section V using a multi-area power network.

**Notation**. Given a stochastic variable $x$, we represent as $\mathbb{E}[x]$ its expected value. The symbols $\geq$ and $>$ are used to denote positive semi-definite matrices and positive definite matrices, respectively. The cardinality of a set $\mathcal{N}$ is denoted with $|\mathcal{N}|$ and the spectral radius of a square matrix $A$ is $\sigma(A)$. A square matrix is Schur stable if $\sigma(A) < 1$. The Kronecker delta is $\delta_{ij}$.

## II. FAULT DETECTION PROBLEM FORMULATION

The monitored large-scale system is composed of (or can be decomposed in) $M$ interconnected subsystems. Each subsystem $\Sigma_i$, with $i = 1, \ldots, M$, is described by the following equations:

$$\Sigma_i \; : \; x_i(k+1) = A_{ii}x_i(k) + \sum_{j \neq i} A_{ij}x_j(k) + w_i(k),$$
$$y_i(k) = C_i x_i(k) + v_i(k), \tag{1}$$

where $x_i(k), w_i(k) \in \mathbb{R}^{n_i}$ and $y_i(k), v_i(k) \in \mathbb{R}^{p_i}$. We assume that $w_i(k)$ and $v_i(k)$ are zero-mean white noises, for all $i = 1, \ldots, M$, and $\mathbb{E}\{w_i(k)w_j^\top(k)\} = Q_i\delta_{ij}$, $\mathbb{E}\{v_i(k)v_j(k)\} = R_i\delta_{ij}$ (with $R_i > 0$ for all $i = 1, \ldots, M$), and that $\mathbb{E}\{w_i(k)v_j^\top(h)\} = 0$ for all $i, j = 1, \ldots, M$ and $h, k \geq 0$. For $i \in \{1, \ldots, M\}$, $\mathcal{N}_i$ denotes the set of neighbors (also called predecessors in [17]) of subsystem $i$ defined as $\mathcal{N}_i = \{j \mid A_{ij} \neq 0\}$ while $\mathcal{S}_i$ is the set of successors of subsystem $i$ defined as $\mathcal{S}_i = \{j \mid i \in \mathcal{N}_j\}$. In our setup we assume that subsystem $i$ can receive information from its neighbors. Note that $i$ is in general included in $\mathcal{S}_i$ and $\mathcal{N}_i$. For later use, we also define the set of strict neighbors and successors $\tilde{\mathcal{N}}_i = \mathcal{N}_i \backslash \{i\}$ and $\tilde{\mathcal{S}}_i = \mathcal{S}_i \backslash \{i\}$, respectively.

Each subsystem is monitored by a local diagnoser that can communicate with neighboring subsystems. Each diagnoser

locally implements a Luenberger observer to estimate the local state vector:

$$\hat{x}_i(k+1) = \sum_{j \in \mathcal{N}_i} \{A_{ij}\hat{x}_j(k) + L_{ij}[y_j(k) - C_j\hat{x}_j(k)]\}$$
$$\hat{y}_i(k) = C_i\hat{x}_i(k) \tag{2}$$

Then, each local diagnoser computes a *local residual signal*

$$r_i(k) := y_i(k) - \hat{y}_i(k)$$

and uses it, together with a properly designed threshold, to monitor the corresponding subsystem. Given $\alpha > 1$ and taking advantage of the Chebishev inequalities, for each $l$-th component $r_{i,l}$ of the residual $r_i$ we can write

$$\Pr(\mathbb{E}[r_{i,l}] - \alpha\sqrt{\mathtt{Var}[r_{i,l}]} \leq r_{i,l} \leq \mathbb{E}[r_{i,l}] + \alpha\sqrt{\mathtt{Var}[r_{i,l}]})$$
$$\geq 1 - \frac{1}{\alpha^2}.$$

We define component-wise the time-varying threshold

$$\bar{r}_{i,l}(k) = \alpha\sqrt{\mathtt{Var}[r_{i,l}(k)]}. \tag{3}$$

Therefore, since $\mathbb{E}[r_i(k)] = 0$ for all $k$, in healthy conditions

$$|r_i(k)| \leq \bar{r}_i(k),$$

with a probability greater than $1 - \frac{1}{\alpha^2}$.

It is now of interest to compute $\bar{r}_i(k)$. As it will be clearer in the following, this is possible in a distributed and scalable fashion at the price of using a suitable upper bound for the variance of the estimation error.

In order to compute the local threshold in an appropriate way, we note that the local residual can be written as

$$r_i(k) = C_i e_i(k) + v_i(k),$$

where $e_i(k) = x_i(k) - \hat{x}_i(k)$ is the local estimation error, whose dynamics is given by

$$e_i(k+1) = \sum_{j \in \mathcal{N}_i} \{(A_{ij} - L_{ij}C_j)e_j(k) - L_{ij}v_j(k)\} + w_i(k). \tag{4}$$

We introduce the extended vectors $e, v, w$, as column vectors collecting $e_i$, $v_i$ and $w_i$, respectively, for all $i = 1, \ldots, M$. Moreover, we define the extended matrices $A$, $L$, as block matrices having the $(i, j)$-th element equal to $A_{ij}$ and $L_{ij}$, respectively, $i = 1, \ldots, M$, $j = 1, \ldots, M$. Finally, $C$ is a block matrix collecting on the diagonal the matrices $C_i$, $i = 1, \ldots, M$. We can therefore describe the dynamics of the extended estimation error as

$$e(k+1) = (A - LC)e(k) - Lv(k) + w(k). \tag{5}$$

The covariance matrix of the extended estimation error $\Pi(k+1) := \mathbb{E}[e(k+1)e^\top(k+1)]$ obeys the recursive equation:

$$\Pi(k+1) = (A - LC)\Pi(k)(A - LC)^\top + LRL^\top + Q. \tag{6}$$

Note that, since the target residual for the diagnoser is $r_i(k) = C_i e_i(k) + v_i(k)$, its covariance matrix (in healthy

conditions) is given by

$$\mathbb{E}[r_i(k)r_i(k)^\top] = C_i\Pi_i(k)C_i^\top + R_i, \qquad (7)$$

where $\Pi_i(k) \in \mathbb{R}^{n_i \times n_i}$ is the $i$-th diagonal block of matrix $\Pi(k)$. However, equation (6) does not allow for a recursive distributed update, i.e., where only local computations are performed and where communication is required only among neighboring diagnosers. In the following, we define an upper bound $B_i(k)$ to the local estimation error covariance $\Pi_i(k)$ that can be computed in a distributed way and that can be used for the computation of the local thresholds. This, as a byproduct, leads to a scalable design procedure for the estimation gains $L_{ij}$ allowing for PnP operations.

## III. Upper bound to the error covariance matrix and convergence properties

Setting $k = 1$ as the initial time instant, we define the time-varying matrix $B_i(k)$, $i = 1, \ldots, M$ for all $k > 1$, using the following distributed update scheme

$$B_i(k+1) = \sum_{j \in \mathcal{N}_i} \left[ (\tilde{A}_{ij} - L_{ij}\tilde{C}_j)B_j(k)(\tilde{A}_{ij} - L_{ij}\tilde{C}_j)^\top \right.$$
$$\left. + L_{ij}\tilde{R}_j L_{ij}^\top \right] + Q_i, \qquad (8)$$

where, for all $i, j = 1, \ldots, M$, $\tilde{A}_{ij} = \sqrt{\varsigma_j}A_{ij}$, $\tilde{C}_i = \sqrt{\varsigma_i}C_i$, and $\tilde{R}_i = \varsigma_i R_i$, and $\varsigma_i = |\mathcal{S}_i|$.

We have the following result, namely that $B_i(k)$ can be used as an upper bound to $\Pi_i(k)$, for all $i = 1, \ldots, M$ and for all $k \geq 1$. The proof is omitted due to space constraints.

*Theorem 1:* If we set, for all $i = 1, \ldots, M$, $\mathrm{diag}(B_1(1), \ldots, B_M(1)) \geq \Pi(1)$ then, for all $k \geq 1$, it holds that $B_i(k) \geq \Pi_i(k)$. $\qquad \square$

Next, we give a centralized condition guaranteeing that, at the same time, the error dynamics (5) is asymptotically stable and $B_i(k)$ is bounded for all $k$.

Before to continue, some definitions are in place. We define, for all $i, j$, $\tilde{F}_{ij} = (\tilde{A}_{ij} - L_{ij}\tilde{C}_j)$ and the matrix $\tilde{F}$ as the matrix whose blocks are $\tilde{F}_{ij}$. Also, we define the following further matrix.

$$\mathbb{F} = \tilde{F} \odot \tilde{F} = \begin{bmatrix} \tilde{F}_{11} \otimes \tilde{F}_{11} & \ldots & \tilde{F}_{1M} \otimes \tilde{F}_{1M} \\ \vdots & \ddots & \vdots \\ \tilde{F}_{M1} \otimes \tilde{F}_{M1} & \ldots & \tilde{F}_{MM} \otimes \tilde{F}_{MM} \end{bmatrix} \qquad (9)$$

where $\odot$ denotes the Khatri-Rao product, while $\otimes$ denotes the Kronecker product [18].

Now we are in the position to state the second main result, whose proof is omitted due to length constraints.

*Theorem 2:* If matrix $\mathbb{F}$ is Schur stable, then
(i) There exists, for all $i = 1, \ldots, M$, a matrix $\bar{B}_i \geq 0$ such that $B_i(k) \to \bar{B}_i$ as $k \to +\infty$;
(ii) $A - LC$ is Schur stable. $\qquad \square$

## IV. Scalable design of local estimators

As it is evident from the previous section, the key condition guaranteeing the effectiveness of the proposed estimation/fault detection scheme is the Schur stability of the matrix $\mathbb{F}$. This condition can be checked *in a scalable way* via the following result.

*Proposition 1:* For matrices $L_{ii}$ such that $\tilde{F}_{ii}$ is Schur stable, if the following conditions are fulfilled

$$\alpha_i = \sum_{j \in \tilde{\mathcal{N}}_i} \sum_{k=0}^{\infty} ||\tilde{F}_{ii}^k \tilde{F}_{ij}||_\infty^2 < 1, \ \forall i = 1, \ldots, M \qquad (10)$$

then $\mathbb{F}$ is Schur stable. $\qquad \square$

*Proof:* Using some properties of the Kronecker products (see [18]), one has

$$||\tilde{F}_{ii}^k \tilde{F}_{ij}||_\infty^2 = ||(\tilde{F}_{ii}^k \tilde{F}_{ij}) \otimes (\tilde{F}_{ii}^k \tilde{F}_{ij})||_\infty =$$
$$= ||(\tilde{F}_{ii}^k \otimes \tilde{F}_{ii}^k)(\tilde{F}_{ij} \otimes \tilde{F}_{ij})||_\infty =$$
$$= ||(\tilde{F}_{ii} \otimes \tilde{F}_{ii})^k (\tilde{F}_{ij} \otimes \tilde{F}_{ij})||_\infty = ||\mathbb{F}_{ii}^k \mathbb{F}_{ij}||_\infty$$

where $\mathbb{F}_{ii} = \tilde{F}_{ii} \otimes \tilde{F}_{ii}$ and $\mathbb{F}_{ij} = \tilde{F}_{ij} \otimes \tilde{F}_{ij}$ denote the blocks composing the matrix $\mathbb{F}$ in (9). Then, conditions (10) can be written as

$$\sum_{j \in \tilde{\mathcal{N}}_i} \sum_{k=0}^{\infty} ||\mathbb{F}_{ii}^k \mathbb{F}_{ij}||_\infty < 1, \ \forall i = 1, \ldots, M. \qquad (11)$$

Since $\tilde{F}_{ii}$ is Schur stable, $\mathbb{F}_{ii}$ has the same property, because each eigenvalue of $\mathbb{F}_{ii}$ is the product of two eigenvalues of $\tilde{F}_{ii}$ [18]. Proposition 1 in [19] then shows that conditions (11) imply the Schur stability of $\mathbb{F}$. $\qquad \blacksquare$

The quantity $\alpha_i$ in (10) depends only upon pieces of information about subsystem $\Sigma_i$ (matrices $\tilde{A}_{ii}$, $\tilde{C}_i$ and $A_{ij}$, $j \in \mathcal{N}_i$), estimator $i$ (matrices $L_{ii}$ and $L_{ij}$, $j \in \mathcal{N}_i$), and matrices $\tilde{C}_j$ and parameters $\varsigma_j$ of strict neighbors $j \in \mathcal{N}_i$ (the latter are needed for computing matrices $\tilde{A}_{ij}$). In particular, no information about $\Sigma_j$ $j \neq \mathcal{N}_i$ is required. Therefore, $\alpha_i$ can be computed locally by subsystem $i$, after having exchanged information with its strict neighbors. Similarly, the following design problem can be solved locally and independently of $\Sigma_j$, $j \notin \mathcal{N}_i$:

*Problem 1:* Compute matrices $L_{ij}$, $j \in \mathcal{N}_i$ such that $\tilde{F}_{ii}$ is Schur stable and (10) holds.

As proposed in [19] for PnP control design, instead of computing $L_{ii}$ and $L_{ij}$, $j \in \tilde{\mathcal{N}}_i$ simultaneously, one can follow the more conservative (but simplified) procedure described in Algorithm 1.

---

**Algorithm 1** Design of a local estimator for subsystem $\Sigma_i$ (Problem 1)

**Input**: $\tilde{C}_i$, $\mathcal{N}_i$, $\{\tilde{A}_{ij}\}_{j \in \mathcal{N}_i}$

(i) $\forall j \in \tilde{\mathcal{N}}_i$, compute the matrix $L_{ij}$ solving the linear programming problem

$$\min_{L_{ij}} ||\tilde{F}_{ij}||_\infty. \qquad (12)$$

(ii) Compute $L_{ii}$ such that $\tilde{F}_{ii}$ is Schur stable and $\alpha_i < 1$. If it does not exist **stop**.
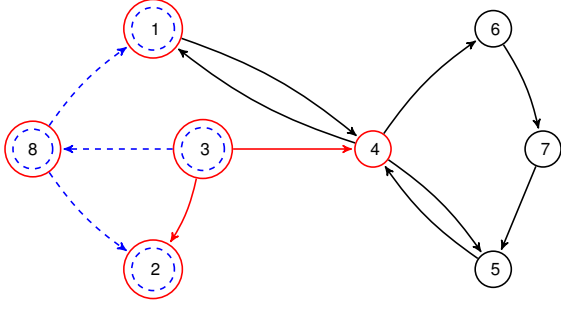
---

Fig. 1. Plug-in of $\Sigma_8$ at time $T_{PnP}$ in a network composed by 7 subsystems. Solid edges: coupling graph of the original network induced by sets $\mathcal{N}_i$ and $\mathcal{S}_i$, $i = 1, ..., 7$. Dashed blue edges: new coupling links given by $\tilde{\mathcal{N}}_8 = \{3\}$ and $\tilde{\mathcal{S}}_8 = \{1, 2\}$. Dashed blue nodes: subsystems that must succesfully run Algorithm 1 (in order to allow the plug-in of $\Sigma_8$) and use new estimators from time $T_{PnP}$. Red nodes: subsystems that must update the dynamics (8) from time $T_{PnP}$.

This approach is justified by the fact that, $\alpha_i$ can be bounded from above as

$$\alpha_i = \sum_{j \in \tilde{\mathcal{N}}_i} \sum_{k=0}^{\infty} ||\tilde{F}_{ii}^k \tilde{F}_{ij}||_\infty^2 \leq \sum_{k=0}^{\infty} ||\tilde{F}_{ii}||_\infty^{2k} \sum_{j \in \tilde{\mathcal{N}}_i} ||\tilde{F}_{ij}||_\infty^2. \tag{13}$$

Therefore, matrices $L_{ij}$ in step (i) of Algorithm 1 minimize the upper bound. According to step (ii) of Algorithm 1, the computation of $L_{ii}$ can be carried out by solving a nonlinear optimization problem. We defer the reader to [20, Chapter 4] for a discussion about some numerical aspects.

### A. PnP operations

After the addition or the removal of a subsystem, the update of the local state estimators and dynamics (8) might be needed for some subsystems. Next, we detail these changes, showing that they may impact at most on subsystems that are two-hops away from the entering/leaving unit. In both cases, the starting point is a network of subsystems equipped with observers produced by Algorithm 1. We denote with $T_{PnP}$ the planned plug-in/out time and use "+" for quantities that must be used after the plug-in/out event (if it takes place). For each subsystem $i$, we define $\rho_i^+ = \frac{\varsigma_i^+}{\varsigma_i}$ where $\varsigma_i^+ = |\mathcal{S}_i^+|$. We start noting that once a matrix $L_{ij}$ has been computed using (12), it never changes. Indeed, $L_{ij}$ minimizes $||\tilde{A}_{ij} + L_{ij}\tilde{C}_j||_\infty$ and, since $\tilde{A}_{ij}^+ = \sqrt{\rho_j^+}\tilde{A}_{ij}$ and $\tilde{C}_j^+ = \sqrt{\rho_j^+}\tilde{C}_j$, it also minimizes $||\tilde{A}_{ij}^+ + L_{ij}\tilde{C}_j^+||_\infty$, irrespectively of $\rho_i^+$.

Suppose, for example, that subsystem $\Sigma_{M+1}$ needs to be plugged-in and be connected with predecessors $\tilde{\mathcal{N}}_{M+1}$ and successors $\tilde{\mathcal{S}}_{M+1}$ (Figure 1 provides an example with $M + 1 = 8$, $\tilde{\mathcal{N}}_8 = \{3\}$ and $\tilde{\mathcal{S}}_8 = \{1, 2\}$).

First, each subsystem $j \in \mathcal{N}_{M+1}$ sends $\varsigma_j^+$ to its successors. In order to preserve properties (i) and (ii) of Theorem 2, one must design new estimators through Algorithm 1 for subsystem $M + 1$ and for

- all subsystems in $\tilde{\mathcal{S}}_{M+1}$, as they will be affected by new coupling terms (see the dashed blue edges in Figure 1);
- all subsystems $j \in \mathcal{N}_{M+1}$ because, without changing the gain $L_{jj}$, one would have

$$\tilde{F}_{jj}^+ = \sqrt{\rho_j^+}\tilde{F}_{jj} \tag{14}$$

and, since $\rho_j^+ > 1$, neither Schur stabiity of $\tilde{F}_{jj}^+$ nor $\alpha_j^+ < 1$ is guaranteed.

If a single instance of the optimization problem in step (ii) of Algorithm 1 is infeasible, the plug-in of $\Sigma_{M+1}$ is denied. Otherwise it is allowed and new estimators are activated at time $T_{PnP}$. Subsystems that must update dynamics (8) and use it from time $T_{PnP}$, are, besides $\Sigma_{M+1}$,

- subsystems in $\tilde{\mathcal{S}}_{M+1}$, as they must include new coupling terms;
- subsystems $j \in \tilde{\mathcal{N}}_{M+1}$ as $\rho_j^+ > 1$ and (8) must use the matrix $\tilde{F}_{jj}^+$ in (14);
- subsystems in $\mathcal{U}_{M+1} = \cup_{j \in \tilde{\mathcal{N}}_{M+1}} \tilde{\mathcal{S}}_j$. Indeed, for each $j \in \tilde{\mathcal{N}}_{M+1}$, the quantity $\rho_j^+ > 1$ has been sent to all subsystems $\ell \in \tilde{\mathcal{S}}_j$ and matrices $\tilde{A}_{\ell j}$, $\tilde{C}_j$ and $\tilde{R}_j$, used in (8) by subsystems $\ell \in \tilde{\mathcal{S}}_j$, must be updated by multiplying them by $\sqrt{\rho_j^+}$.

Summarizing the above points, all subsystems in the set $\mathcal{N}_{M+1} \cup \tilde{\mathcal{S}}_{M+1} \cup \mathcal{U}_{M+1}$ must update dynamics (8) (see the red nodes in Figure 1). We highlight that no other subsystem in the network need to change the corresponding local estimators or dynamics (8). This motivates the scalability of the plug-in operation.

We discuss now the unplugging of a subsystem (say $\Sigma_u$) at time $T_{PnP}$. First, each subsystem $j \in \tilde{\mathcal{N}}_u$, having a successor less, sends $\varsigma_j^+$ to its successors. Then,

- subsystems $i \in \tilde{\mathcal{S}}_u$ can update the local estimator (2) by just using the new set $\tilde{\mathcal{N}}_i^+$ and without changing gains $L_{ii}$ and $L_{ij}$ Indeed, $\tilde{\mathcal{N}}_i^+$ will have one element less and so will the sum in (10). Moreover, in (10), matrices $\tilde{F}_{ij}^+ = \tilde{A}_{ij}^+ - L_{ij}\tilde{C}_j^+$ verify

$$\tilde{F}_{ij}^+ = \sqrt{\rho_j^+}\tilde{F}_{ij} \tag{15}$$

with $\sqrt{\rho_j^+} \leq 1$. The above conditions guarantee that $\alpha_i^+ < 1$. Finally, by using the matrix $L_{ii}$ previously designed, matrices $\tilde{F}_{ii}^+$ and $\tilde{F}_{ii}$ are related by (14) with $\rho_i^+ \leq 1$. This shows that Schur stability of $\tilde{F}_{ii}^+$ is guaranteed.
- subsystems $j \in \tilde{\mathcal{N}}_u$ will have one successor less. Therefore, as shown in the previous point, without changing gains $L_{jj}$ and $L_{\ell j}$, $\ell \in \mathcal{S}_j^+$ one has that matrices $\tilde{F}_{jj}^+$ are Schur stable and $\alpha_\ell^+ < 1$.

It follows that the unplugging of $\Sigma_u$ can be always performed without spoiling properties (i) and (ii) of Theorem 2, and, similarly to the plug-in operation, subsystems in $\tilde{\mathcal{N}}_u \cup \tilde{\mathcal{S}}_u \cup \mathcal{U}_u$ will have to update dynamics (8) from time $T_{PnP}$.

## V. SIMULATION RESULTS

In this section, we provide some simulation results illustrating the effectiveness of the proposed distributed fault detection technique.

As a quite significant case-study, we consider a power network system including a number of power generation areas coupled through tie-lines. The dynamics of each power generation area, equipped with primary control and linearized around the equilibrium value for all variables, is described by the following continuous time LTI model [21]:

$$\dot{x}_i(t) = A_{ii}^c x_i(t) + B_i^c u_i + L_i^c \Delta P_{L_i} + \sum_{j \in \mathcal{N}_i} A_{ij}^c x_j, \quad (16)$$

where $x_i = (\Delta\theta_i, \Delta\omega_i, \Delta P_{m_i}, \Delta P_{v_i})$ is the state, $u_i = \Delta P_{ref_i}$ is the control input of each area, and $\Delta P_{L_i}$ is the local power load. Note that the letter $\Delta$ is used to denote the deviation from steady-state. The matrices of system (16) are

$$A_{ii}^c = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{\sum_{j \in \mathcal{N}_i} P_{ij}}{2H_i} & -\frac{D_i}{2H_i} & \frac{1}{2H_i} & 0 \\ 0 & 0 & -\frac{1}{T_{t_i}} & \frac{1}{T_{t_i}} \\ 0 & -\frac{1}{R_i T_{g_i}} & 0 & -\frac{1}{T_{g_i}} \end{bmatrix},$$

$$B_i^c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{T_{g_i}} \end{bmatrix}, A_{ij}^c = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \frac{P_{ij}}{2H_i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, L_i^c = \begin{bmatrix} 0 \\ -\frac{1}{2H_i} \\ 0 \\ 0 \end{bmatrix}$$

where the parameters and their numerical values are defined in [21]. Since both $\Delta P_{ref_i}$ and $\Delta P_{L_i}$ are assumed to be constant and known, for the sake of simplicity, we neglect them in our analysis.

We discretize the process (16) with a sampling interval $T = 1$ sec according to the technique proposed in [22], leading to the discrete-time model (1) where the matrices $A_{ii}$, $A_{ij}$ can be easily constructed from (16). The matrix $C_i$ is

$$C_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

For $i = 1, \ldots, M$, $\mathbb{E}\left[ w_i w_i^\top \right] = Q_i = 0.0001\,I_4$ and $\mathbb{E}\left[ v_i v_i^\top \right] = R_i = 0.0001 I_2$ where $I_k$ is the identity matrix of order $k$.

In this section we consider the scenario 1 in [21], where $M = 4$ and where the adjacency matrix $Ad$, defining the neighboring relationships between areas[1], is

$$Ad = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

namely, $Ad_{ij} \neq 0$ if and only if $\frac{P_{ij}}{2H_i} \neq 0$.

The gains $L_{ij}$ of the Luenberger observer proposed in (2), have been computed based on the steady state solutions of suitable local Riccati equations introduced in [23], where a distributed Kalman filtering scheme with theoretical guarantees has been proposed. We have verified that the corresponding matrices $\tilde{F}_{ij} = (\tilde{A}_{ij} - L_{ij}\tilde{C}_j)$ satisfy (10).

---

[1]In this example, neighboring relations are induced by electric lines and they are symmetric since electric power flows in both directions.

At time instant $t = 100$, the following fault occurs in area 1: the inertia constant $H_1$ is reduced from 12 to 1, which means, from an electrical point of view, that there is a fault in a local generator and hence the faulty area must be isolated for safety reasons, not to propagate faults in the PNS. In order to define the threshold, we set $\alpha = 2.7$. In Figure 2, where the experiment is repeated 100 times using random initial conditions and noise realizations, we can see residuals and thresholds signals for each area of the PNS. We can see that the local diagnosers are able to detect the fault in Area 1 and that the false-alarms are rare events. In Figure 3, we can see the detection by the local diagnoser in Area 1 for component 2 of the residual (i.e. $r_{1,2}(k)$), for a single experiment.
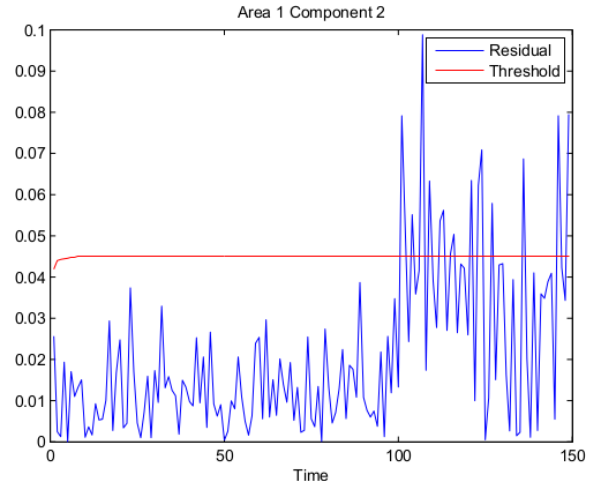


Fig. 3. Residual signal $r_{1,2}(k)$ (in blue) and threshold $\bar{r}_{1,2}(k)$ (in red) for Area 1 component 2 for a single experiment. The inertia constant $H_1$ is reduced from 12 to 1 at time $k = 100$.

Finally, we tested the fault detection scheme on different scenarios with different faults and with faults in different areas and we obtained similar results.

## VI. CONCLUDING REMARKS

In this paper, we propose a novel distributed fault detection method for interconnected linear systems, allowing the computation of suitable local thresholds with guaranteed maximum false-alarms rate. This is achieved by a partition-based distributed estimation method that takes into account the dynamic coupling terms between subsystems. Moreover, a bound on the variance of the estimation error that is computed in a distributed way. Notably, the proposed estimation and fault detection method enjoys scalability features, allowing PnP operations. As a future work, we are going to analyse the conservativeness of the proposed bound on the variance of the estimation error and we will provide extensive simulation analysis.

### REFERENCES

[1] T. Samad and T. Parisini, "Systems of systems," in *The Impact of Control Technology*, T. Samad and A. M. Annaswamy, Eds. IEEE Control Systems Society, 2011, pp. 175–183.
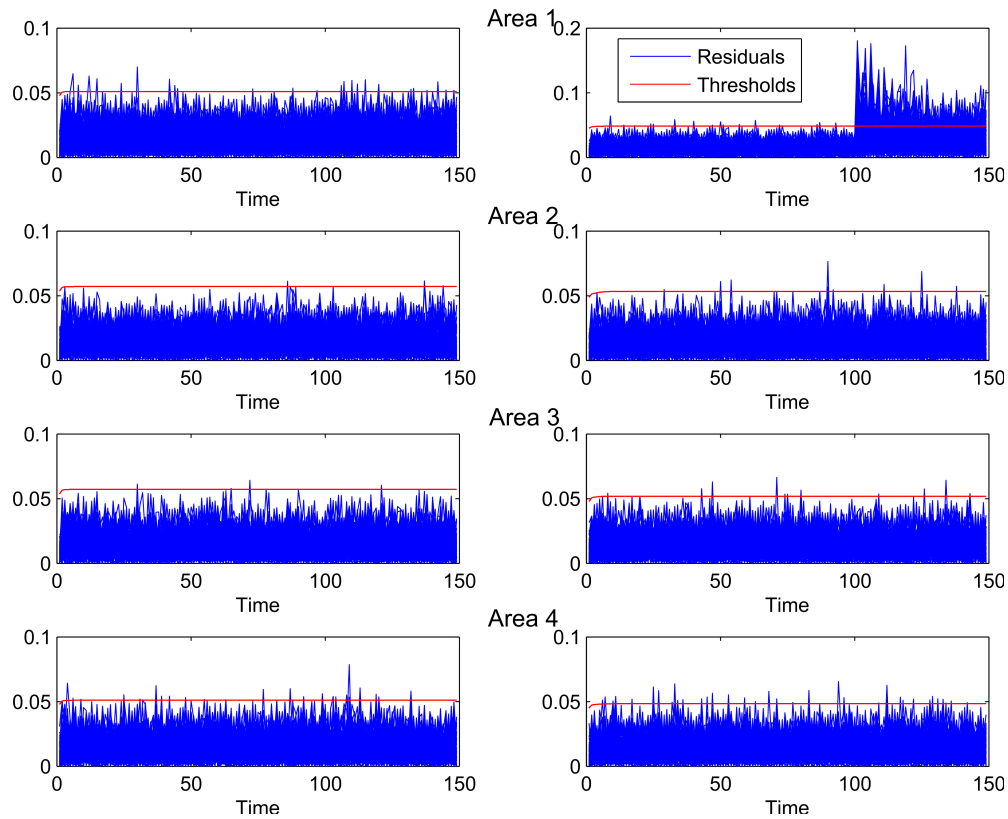
Fig. 2. For each area, residual signals (in blue) and thresholds (in red) for 100 experiments. Left and right panels correspond to the first and second scalar outputs, respectively. The inertia constant $H_1$ is reduced from 12 to 1 at time $k = 100$.

[2] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *46th IEEE Conference on Decision and Control*, 2007, pp. 5492–5498.

[3] S.-Y. Tu and A. H. Sayed, "Diffusion strategies outperform consensus strategies for distributed estimation over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6217–6234, 2012.

[4] R. Vadigepalli and F. Doyle III, "A distributed state estimation and control algorithm for plantwide processes," *IEEE Transactions on Control Systems Technology*, vol. 11, no. 1, pp. 119 – 127, 2003.

[5] U. A. Khan and J. M. F. Moura, "Distributing the Kalman Filters for Large-Scale Systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4919–4935, 2008.

[6] S. Roshany-Yamchi, M. Cychowski, R. R. Negenborn, B. De Schutter, K. Delaney, and J. Connell, "Kalman filter-based distributed predictive control of large-scale multi-rate systems: Application to power networks," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 1, pp. 27–39, 2013.

[7] M. Farina, G. Ferrari-Trecate, and R. Scattolini, "Moving-horizon partition-based state estimation of large-scale systems," *Automatica*, vol. 46, no. 5, pp. 910–918, 2010.

[8] S. Riverso, M. Farina, R. Scattolini, and G. Ferrari-Trecate, "Plug-and-play distributed state estimation for linear systems," in *IEEE 52nd Annual Conference on Decision and Control*, 2013, pp. 4889–4894.

[9] M. Farina and R. Carli, "Plug and play partition-based state estimation based on Kalman filter," in *IEEE 54th Annual Conference on Decision and Control*, 2015, pp. 3155–3160.

[10] F. Boem, R. Ferrari, T. Parisini, and M. Polycarpou, "Distributed fault detection and isolation of continuous-time nonlinear systems," *European Journal of Control*, vol. 5-6, pp. 603–620, 2011.

[11] I. Shames, A. M. Teixeira, H. Sandberg, and K. H. Johansson, "Distributed fault detection for interconnected second-order systems," *Automatica*, vol. 47, no. 12, pp. 2757–2764, 2011.

[12] X. Zhang and Q. Zhang, "Distributed fault diagnosis in a class of interconnected nonlinear uncertain systems," *International Journal of Control*, vol. 85, no. 11, pp. 1644–1662, 2012.

[13] V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, "Decentralized isolation of multiple sensor faults in large-scale interconnected nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 60, no. 6, pp. 1582–1596, 2015.

[14] S. Riverso, F. Boem, G. Ferrari-Trecate, and T. Parisini, "Plug-and-play fault detection and control-reconfiguration for a class of nonlinear large-scale constrained systems," *IEEE Transactions on Automatic Control*, vol. PP, no. 99, pp. 1–1, 2016.

[15] F. Boem, S. Riverso, G. Ferrari-Trecate, and T. Parisini, "Stochastic fault detection in a plug-and-play scenario," in *IEEE 54th Annual Conference on Decision and Control*, 2015, pp. 3137–3142.

[16] Y. Zhou, F. Boem, C. Fischione, and T. Parisini, "Distributed fault detection with sensor networks using pareto-optimal dynamic estimation method," in *European Control Conference (accepted)*, 2016.

[17] S. Riverso, M. Farina, and G. Ferrari-Trecate, "Plug-and-play decentralized model predictive control for linear systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 10, pp. 2608–2614, 2013.

[18] R. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 2012.

[19] S. Riverso and G. Ferrari-Trecate, "Plug-and-play distributed model predictive control with coupling attenuation," *Optim. Control Appl. Meth., DOI: 10.1002/oca.2142*, vol. 36, no. 3, pp. 292–305, 2015.

[20] S. Riverso, "Distributed and plug-and-play control for constrained systems," Ph.D. dissertation, Università degli Studi di Pavia, 2014. [Online]. Available: http://sisdin.unipv.it/pnpmpc/phpinclude/papers/phd_thesis_riverso.pdf

[21] S. Riverso and G. Ferrari-Trecate, "Hycon2 benchmark: Power network system," *arXiv:1207.2000v1*, 2012.

[22] M. Farina, P. Colaneri, and R. Scattolini, "Block-wise discretization accounting for structural constraints," *Automatica*, vol. 49, no. 11, pp. 3411–3417, 2013.

[23] M. Farina and R. Carli, "Partition-based distributed Kalman filter with plug and play features," *arXiv:1507.06820*.