ScienceDirect
IFAC Papers Online
CONFERENCE PAPER ARCHIVE
IFAC-PapersOnLine 48-21 (2015) 060–065

Optimal Topology for Distributed Fault Detection of Large-scale Systems ⋆

Francesca Boem * **Riccardo M. G. Ferrari** **
Thomas Parisini *** **Marios M. Polycarpou** ****

* *Imperial College London, UK,* (f.boem@imperial.ac.uk).
** *Danieli Automation S.p.A., Italy* (r.ferrari@dca.it)
*** *Imperial College London, UK, and University of Trieste, Italy*
(t.parisini@gmail.com)
**** *University of Cyprus, Cyprus.* (mpolycar@ucy.ac.cy)

Abstract: The paper deals with the problem of defining the optimal topology for a distributed fault detection architecture for non-linear large-scale systems. A stochastic model-based framework for diagnosis is formulated. The system structural graph is decomposed into subsystems and each subsystem is monitored by one local diagnoser. It is shown that overlapping of subsystems allows to improve the detectability properties of the monitoring architecture. Based on this theoretical result, an optimal decomposition design method is proposed, able to define the minimum number of detection units needed to guarantee the detectability of certain faults while minimizing the communication costs subject to some computation cost constraints. An algorithmic procedure is presented to solve the proposed optimal decomposition problem. Preliminary simulation results show the potential of the proposed approach.

Keywords: Fault detection, Distributed systems, Optimal decomposition, Topology

1. INTRODUCTION

Recently there has been a growing interest towards distributed architectures for fault diagnosis of large-scale and networked systems (see, for instance Boem et al. (2011); Ferrari et al. (2012); Zhang and Zhang (2012); Boem et al. (2013a,b); Keliris et al. (2015); Reppa et al. (2015)). As it is well known, the drawbacks of a centralized fault diagnosis architecture are scalability and robustness. A common solution is to decompose the overall system into subsystems that are monitored by some local agents, which we call Local Fault Diagnosers (LFDs). According to (Šiljak (1978)), the term *system decomposition* refers to the clustering of the states, inputs, and outputs system variables into subsets, which make up the subsystems. Since each LFD is devoted to monitor a subsystem, the decomposition of the overall system defines the topology of the diagnosis architecture. Given a large-scale interconnected system and its structural graph (Šiljak (1978)) whose nodes are the state and the input variables of the system, the goal is to identify: i) the number of local subsystems (and therefore the number of LFDs) needed to monitor the system, and ii) how to assign the system variables of the monitored system to each subsystem. The latter considers also which variables need to be shared among more than one subsystem. In this respect, overlapping decompositions are considered, that is, some state variables may be monitored by more than one LFD. An additional objective is to make the Fault Detection (FD) problem computationally tractable and to guarantee some performances related to given monitoring goals.

The problem of system decomposition is well-known problem in decentralized and distributed control and there are some recent papers presenting algorithms for non-overlapping (Ocampo-Martinez et al. (2011)) and possibly overlapping decompositions (Anderson and Papachristodoulou (2012))). On the other hand, the works proposing distributed monitoring schemes for discrete-time or continuous-time systems, like Boem et al. (2011); Ferrari et al. (2012); Zhang and Zhang (2012), assume that the decomposition of the system into subsystems is given a priori. The aim of this work is to study the decomposition problem specifically for the fault detection task. The goal is to understand how the decomposition and the adoption of distributed approaches can influence the detectability performances. In Bregon et al. (2014), a decentralized fault diagnosis task using structural model decomposition is considered, but an event-based method is implemented in a qualitative approach. In Staroswiecki and Amani (2014), the topology of the information pattern is studied in order to allow fault-tolerant control reconfiguration. In Grbovic et al. (2012), the decomposition is designed using the Sparse Principal Component Analysis algorithm, but the proposed decentralized fault detection architecture is a data-driven approach (see Yin et al. (2014) for a recent survey), while our method is a model-based one (see Venkatasubramanian et al. (2003)).

The main contributions of the paper are: i) a methodology to find the minimum number of LFDs needed to detect a certain set of faults is addressed. Once obtained the minimum required number of LFDs and the variables to be shared, an optimal topology is determined that minimizes the communication costs and satisfies some computational constraints. We show that the decomposition allows to improve detectability; ii) a novel stochastic formulation for the problem of distributed fault detection is proposed, while previous works by the authors presented deterministic approaches [1]; iii) the effectiveness of a consensus approach for diagnosis purposes is demonstrated: the proposed consensus protocol is used as a tool for shared

⋆ This paper has been partially supported by the EPSRC STABLE-NET grant EP/L014343/1.

[1] The proposed approach is based on the model of the system; statistics is introduced to manage the stochastic characterization of the uncertainty and to derive some stochastic thresholds.

10.1016/j.ifacol.2015.09.505

variables estimation and it allows to reduce the uncertainty and to improve detectability.

It is worth noting that, to the best of the authors' knowledge, it is the first time that the system decomposition problem is analyzed specifically for the distributed fault diagnosis purposes. Furthermore, it is the first time that the problem of graph decomposition design is considered in the overlapping case where the nodes to be shared are selected before the decomposition process. In this connection, we remark that in computer science, the problem of graph decomposition has been widely investigated. For instance, a multilevel graph partitioning method is proposed in Karypis and Kumar (1998) and Schloegel et al. (2000), where a non–overlapping decomposition is obtained.

Notation. Given a stochastic variable x, we represent as $\mathbb{E}[x]$ its expected value, and as $\mathrm{Var}[x]$ its variance. Given a vector a, we denote with $a^{(k)}$ its k-th component. Finally, let us denote by present $A^{(k)}$ the k-th row of matrix A.

2. PROBLEM FORMULATION

We introduce in qualitative terms the multi-objective optimization problem designed to obtain an optimal system decomposition for fault detection purposes. The proposed method consists of two phases. Let n^+ be a design parameter, representing the maximum number of state variables that each subsystem may contain due to computational cost limits. Then, we first define the optimization problem $\mathcal{P}_1(n^+)$. Let N denote the number of subsystems, n the number of state variables, and n_I be the number of state variables contained in the I-th subsystem [2], $I = 1, \dots, N$. Let $\mathbb{S}$ be the set of the variables shared among more than one subsystem and d_k be the overlap degree (Ferrari et al. (2012)) of the k-th state variable, $k = 1, \dots, n$, that is, the number of subsystems the variable belongs to. The objective is to find the minimum number of subsystems N^* and the overlap degrees d_k^*, needed to guarantee some detectability conditions $\mathcal{D}(d_k)$ that will be defined in the following, subject to the computational cost constraint. Given n^+, we formulate

$$\begin{aligned} \mathcal{P}_1(n^+) \ : \ N^* = \min_{N, d_k} \ & N \\ \text{s.t.} \ & \\ & \mathcal{D}(d_k) \text{ satisfied} \\ & n_I \leq n^+, \ \ \forall I = 1, \dots, N. \end{aligned} \tag{1}$$

The optimal d_k^* characterizes the optimal set $\mathbb{S}^*$ of the variables to be shared. Once the optimal values N^* and d_k^* are obtained from $\mathcal{P}_1(n^+)$, we then formulate a second optimization problem $\mathcal{P}_2(N^*, d_k^*, n^+)$. The objective is to minimize the communication cost C (defined in the following), given the number of subsystems, the variables to be shared and the computational cost constraint:

$$\begin{aligned} \mathcal{P}_2(N^*, d_k^*, n^+) \ : \ \min \ & C \\ \text{s.t.} \ & \\ & N = N^* \\ & \text{satisfy } d_k^*, \ \forall \ k = 1, \dots, n \\ & n_I \leq n^+, \ \ \forall I = 1, \dots, N. \end{aligned} \tag{2}$$

The outputs of the second optimization problem are the minimum communication cost C^* and the optimal decomposition Ω^* of the system graph.

Consider a large-scale system described by the possibly non-linear model equations:

$$x(t+1) = f(x(t), u(t)) + \eta(t) + \phi(x(t), u(t), t) \tag{3}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ the control input, $f : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$ represent the nominal healthy dynamics, while η describes modeling uncertainties in the state equation. Finally, ϕ is a fault function which is null for $t < T_0$ (T_0 is the time of fault occurrence). Each LFD uses the measurements obtained by some sensors. We assume that the state vector is completely measurable, that is we assume that at least n sensors are available, one for each state component. The possibility of sharing some variables among different subsystems implies the addition of some sensors, so that the LFDs sharing the involved state variable can determine that component value in different ways [3]. In the case that each state variable can be measured by more than one sensor, the output equation is

$$y(t) = Hx(t) + \xi(t),$$

where $y \in \mathbb{R}^p$, with $p \geq n$, represents the measurements vector affected by the measurement noise $\xi \in \mathbb{R}^p$ and H is a $p \times n$ matrix having for each row all the elements null but one, equal to 1. The matrix H describes the relationship between sensors and measured variables: the (i, j)-th element is equal to 1 if the i-th sensor measures the j-th variable. One of the tasks of the design problem is to define the matrix H, that is, to decide how to use the available sensors: we choose the number of sensors measuring each variable and so which variables can take advantage from having redundant measurements. This includes the possibility to add more sensors in order to improve the fault detection performances.

The following assumptions are needed:

Assumption 1. The modeling uncertainty η is an unknown function, modeled as a stochastic process of unknown distribution. We assume to know at each time instant t mean and variance of the stochastic variables $\eta(t)$:

$$\eta(t) \approx (\mu_\eta(t), \sigma_\eta(t)),$$

Assumption 2. The measurement noise ξ is a stochastic process of known distribution. We assume to know at each t the mean and variance of the stochastic variables $\xi(t)$:

$$\xi(t) \approx (\mu_\xi(t), \sigma_\xi(t)).$$

Once the system decomposition is chosen, it is possible to define some local models:

$$\mathscr{S}_I : \begin{cases} x_I(t+1) = f_I(x_I(t), z_I, u_I(t)) + \eta_I(t) \\ \qquad\qquad + \phi_I(x_I(t), u_I(t), t) \\ y_I(t) = x_I(t) + \xi_I(t), \end{cases} \tag{4}$$

where $x_I \in \mathbb{R}^{n_I}$ is the local state vector, $u_I \in \mathbb{R}^{m_I}$ the local input and $y \in \mathbb{R}^{n_I}$ the local output affected by the measurement noise ξ_I, $z_I \in \mathbb{R}^{n_I}$ collects the neighboring state variables affecting local state variables dynamics; f_I is the local nominal function, while η_I describes local uncertainties. Finally, ϕ_I is the local fault function.

Next, we address the model-based FD problem.

3. DISTRIBUTED FAULT DETECTION

We first consider the case in which the k-th state component, with $k = 1, \dots, n$, is monitored by a single LFD. The local estimation model is based on the local system model:

$$\hat{x}_I^{(k)}(t+1) = f_I^{(k)}(y_I(t), v_I(t), u_I(t)) + \lambda(\hat{x}_I^{(k)}(t) - y_I^{(k)}(t)),$$

where v_I is the communicated measurement of z_I, so that $v_I = z_I + \xi_{z,I}$. In the distributed FD architecture, at time $t+1$, the state estimate $\hat{x}_I(t+1)$, computed at time t, is

[2] N and n_I are variables of the optimization problem.

[3] There are many application examples where this fact can be applicable, especially in the more recent times thanks to the low cost of sensor and wireless communication capabilities.

compared component-by-component to the measurement $y_I(t+1)$, which yields a residual signal $\epsilon_I(t+1)$. A threshold $\bar{\epsilon}_I(t+1)$ for the residual $\epsilon_I(t+1)$ is derived at time t so that if the residual exceeds the threshold, then an alarm signal is triggered and the system is considered to be faulty. Let us analyze the residual signal:

$$\begin{aligned}\epsilon_I^{(k)}(t+1) &= y_I^{(k)}(t+1) - \hat{x}_I^{(k)}(t+1)\\ &= \lambda\epsilon_I^{(k)}(t) + \Delta f_I^{(k)}(t) + \eta_I^{(k)}(t) + \xi_I^{(k)}(t+1), \quad (5)\end{aligned}$$

where

$\Delta f_I^{(k)}(t) = f_I^{(k)}(x_I(t), z_I(t), u_I(t)) - f_I^{(k)}(y_I(t), v_I(t), u_I(t))$

and λ is a filtering parameter chosen so that $0 < \lambda < 1$ to guarantee the convergence of the estimator. The following further assumptions are needed.

Assumption 3. The measurement noise $\xi_I(t)$ and the modeling uncertainty $\eta_I(t)$ are not correlated.

Assumption 4. The measurement noise and the modeling uncertainty are zero-mean: $\mu_{\xi_I}(t) = 0$, $\mu_{\eta_I}(t) = 0$, $\forall t$.

We analyze the stochastic part of the residual [4]:

$$\chi_I^{(k)}(t+1) = \Delta f_I^{(k)}(t) + \eta_I^{(k)}(t) + \xi_I^{(k)}(t+1).$$

We can then compute

$$\mathbb{E}[\chi_I^{(k)}(t+1)] = \mathbb{E}[\Delta f_I^{(k)}(t)] \quad (6)$$

$$\begin{aligned}\mathrm{Var}[\chi_I^{(k)}(t+1)] = \mathrm{Var}[\Delta f_I^{(k)}(t)] + \sigma^2_{\eta_I^{(k)}}(t) + \sigma^2_{\xi_I^{(k)}}(t+1)\\ + 2\mathrm{Cov}[\Delta f_I^{(k)}(t), \xi_I^{(k)}(t+1)] \quad (7)\end{aligned}$$

We now derive some time-varying stochastic thresholds. Chebyshev inequalities can be used, without any assumption on the distribution of the residual. For a stochastic variable X, with mean $\mu(X)$ and variance $\sigma^2(X)$, it holds:

$$Pr\big(\mu(X) - \alpha\sigma(X) \le X \le \mu(X) + \alpha\sigma(X)\big) \ge 1 - 1/\alpha^2 \quad (8)$$

where α is a tunable, real positive-valued scalar. Therefore, it is possible to obtain a lower and a upper stochastic thresholds for the residual signal, so that in healthy conditions

$$\bar{\epsilon}_I^{(k)}(t)_{low} \le \epsilon_I^{(k)}(t) \le \bar{\epsilon}_I^{(k)}(t)^{upp} \quad (9)$$

with probability greater than $1 - 1/\alpha^2$. If the residual crosses one of the thresholds, then we can say that a fault has occurred with false-alarm probability lower than $1/\alpha^2$. The thresholds can be computed at each step t as

$$\begin{aligned}\bar{\epsilon}_I^{(k)}(t+1)^{upp}_{low} &= \lambda\bar{\epsilon}_I^{(k)}(t)^{upp}_{low} + \mathbb{E}[\chi_I^{(k)}(t+1)]\\ &\pm \alpha\left[\mathrm{Var}[\chi_I^{(k)}(t+1)]\right]^{\frac{1}{2}} = \lambda\bar{\epsilon}_I^{(k)}(t)^{upp}_{low} + \mathbb{E}[\Delta f_I^{(k)}(t)]\\ &\pm \alpha\Big[\mathrm{Var}[\Delta f_I^{(k)}(t)] + \sigma^2_{\eta_I^{(k)}}(t) + \sigma^2_{\xi_I^{(k)}}(t+1)\\ &\quad + 2\mathrm{Cov}[\Delta f_I^{(k)}(t), \xi_I^{(k)}(t+1)]\Big]^{\frac{1}{2}}. \quad (10)\end{aligned}$$

The value of α is a tuning parameter by which different values of guaranteed false-alarms rate can be set.

4. CONSENSUS-BASED FD METHODOLOGY

In the distributed FD architecture, considering possibly overlapping decomposition, certain state variables may be measured, estimated and monitored by more than one LFD. In this section, we address the case in which the k-th state component is shared by a set of LFDs $\mathcal{O}_k$. Each I-th LFD in this set communicates with the LFDs $J \in \mathcal{O}_k$ sharing the state variable $x^{(k)}$. In this section, the extension of the deterministic consensus mechanism illustrated in Ferrari et al. (2012) to a stochastic context is illustrated. In this case, the I-th estimation model is

$$\begin{aligned}\hat{x}_I^{(k)}(t+1) &= \sum_{J\in\mathcal{O}_k} w_k^{(I,J)}[f_J^{(k)}(y_J(t), v_J(t), u_J(t))\\ &\quad + \lambda(\hat{x}_J^{(k)}(t) - y_J^{(k)}(t))]\\ &= \sum_{J\in\mathcal{O}_k} w_k^{(I,J)}[-\lambda\epsilon_J^{(k)}(t) + f_J^{(k)}(y_J(t), v_J(t), u_J(t))],\end{aligned}$$

where $w_k^{(I,J)}$ is the (I,J)-th element of a stochastic matrix collecting the weights of the consensus for the k-th shared state component, being $\sum_{J\in\mathcal{O}_k} w_k^{(I,J)} = 1$. In this shared-variable case, the residual is

$$\begin{aligned}\epsilon_I^{(k)}(t+1) = \sum_{J\in\mathcal{O}_k} w_k^{(I,J)}[\lambda\epsilon_J^{(k)}(t) + \Delta f_J^{(k)}(t) + \eta_J^{(k)}(t)\\ + \xi_I^{(k)}(t+1)]. \quad (11)\end{aligned}$$

Similarly as before, we obtain the following expressions for the lower and upper thresholds:

$$\begin{aligned}\bar{\epsilon}_I^{(k)}(t+1)^{upp}_{low} = \sum_{J\in\mathcal{O}_k} w_k^{(I,J)}\left[\lambda\bar{\epsilon}_J^{(k)}(t)^{upp}_{low} + \mathbb{E}[\Delta f_J^{(k)}(t)]\right]\\ \pm \alpha\left[\mathrm{Var}[\chi^{(k)}(t+1)]\right]^{\frac{1}{2}} \quad (12)\end{aligned}$$

where

$$\begin{aligned}\left[\mathrm{Var}[\chi^{(k)}(t+1)]\right]^{\frac{1}{2}} = \Bigg[\sum_{J\in\mathcal{O}_k}(w_k^{(I,J)})^2[\mathrm{Var}[\Delta f_J^{(k)}(t)]\\ + \sigma^2_{\eta_J^{(k)}}(t) + \sigma^2_{\xi_I^{(k)}}(t+1) + 2\mathrm{Cov}[\Delta f_J^{(k)}(t), \xi_I^{(k)}(t+1)]]\Bigg]^{\frac{1}{2}}, \quad (13)\end{aligned}$$

It is worth noting that, since $0 \le w_k^{(I,J)} \le 1$ for every (I,J), then $\sum_{J\in\mathcal{O}_k}(w_k^{(I,J)})^2 \le 1$. Therefore, the variance component of the threshold for the shared case in (12) is lower that in the non-shared case in (10). Then, we showed that by using the proposed consensus approach and sharing some state variables among more than one LFD implies the reduction of the variance of the residual signal, thus leading to less conservative detection thresholds.

5. DETECTABILITY ANALYSIS

In this section we define some detectability indicators to be used in the decomposition optimization problem. Let us first consider the case with one single diagnoser I. In this case, the residual in (5) can be written as:

$$\epsilon^{(k)}(t) = U^{(k)}(t) + \sum_{h=T_0}^{t-1}\lambda^{t-1-h}\phi^{(k)}(h) \quad (14)$$

as $\phi(t) = \phi(x(t), u(t), t) = 0$ (with some abuse of notation) for $t < T_0$, and $U^{(k)}(t)$ represents the part of the residual collecting all the uncertainty terms:

$$U^{(k)}(t) = \sum_{h=0}^{t-1}\lambda^{t-1-h}(\chi^{(k)}(h+1)),$$

since $\hat{x}^{(k)}(0) = y^{(k)}(0)$ and $\epsilon^{(k)}(0) = 0$. A fault is detected at a certain time instant $t = T_d > T_0$ with a certain probability depending on α if

$$\epsilon^{(k)}(t) \notin \left(\bar{\epsilon}^{(k)}(t)_{low}, \bar{\epsilon}^{(k)}(t)^{upp}\right) \quad (15)$$

[4] At time t, $\epsilon_I(t)$ is not a random variable, since it can be computed as $\epsilon_I(t) = y_I(t) - \hat{x}_I(t)$.

for at least one state component $k \in \{1, \dots, n\}$. Following (14), condition (15) is equivalent to:

$$\sum_{h=T_0}^{t-1} \lambda^{t-1-h} \phi^{(k)}(h) \notin \left(\bar{\epsilon}^{(k)}(t)_{low} - U^{(k)}(t), \bar{\epsilon}^{(k)}(t)^{upp} - U^{(k)}(t) \right).$$

The uncertainty term can be expressed as $\chi^{(k)}(t) = \mathbb{E}[\chi^{(k)}(t)] + \Delta\chi^{(k)}(t)$, where $\Delta\chi^{(k)}$ is the deviation from its mean, and the thresholds (10) can be rewritten as

$$\bar{\epsilon}^{(k)}(t)^{upp}_{low} = \sum_{h=0}^{t-1} \lambda^{t-1-h} (\mathbb{E}[\chi^{(k)}(h+1)] \pm \alpha \mathrm{Var}[\chi^{(k)}(h+1)]^{\frac{1}{2}})$$

The detectability conditions (15) become:

$$\sum_{h=T_0}^{t-1} \lambda^{t-1-h} \phi^{(k)}(h) \notin \sum_{h=T_0}^{t-1} \lambda^{t-1-h} \Big(-\alpha \left[\mathrm{Var}[\chi^{(k)}(h+1)] \right]^{\frac{1}{2}} - \Delta\chi^{(k)}(h+1), +\alpha \left[\mathrm{Var}[\chi^{(k)}(h+1)] \right]^{\frac{1}{2}} - \Delta\chi^{(k)}(h+1) \Big).$$

Since $\Delta\chi^{(k)}(t)$ is zero–mean, we have that

$$-\alpha \left[\mathrm{Var}[\chi^{(k)}(t)] \right]^{\frac{1}{2}} \leq \Delta\chi^{(k)}(t) \leq \alpha \left[\mathrm{Var}[\chi^{(k)}(t)] \right]^{\frac{1}{2}}$$

with a certain probability using Chebishev inequalities. Therefore, the fault detection is guaranteed at time T_d with a certain false–alarms rate depending on α, when the following detectability condition is satisfied:

$$\left| \sum_{h=T_0}^{T_d-1} \lambda^{T_d-1-h} \phi^{(k)}(h) \right| > 2\alpha \sum_{h=T_0}^{T_d-1} \lambda^{T_d-1-h} \left[\mathrm{Var}[\chi^{(k)}(h+1)] \right]^{\frac{1}{2}}. \tag{16}$$

Consider now the problem of detecting a fault ϕ occurring at some time $t = T_0$ in the nonlinear system (3).

Assumption 5. Suppose that a value $\bar{\phi}^{(k)}$ is given for $k \in \{1, \dots, n\}$ such that for at least one time instant $T > T_0$

$$\left| \sum_{h=T_0}^{T-1} \lambda^{T-1-h} \phi^{(k)}(h) \right| \geq \bar{\phi}^{(k)}.$$

Note that $\bar{\phi}^{(k)}$ represents the minimum magnitude reached by the fault function component that we are interested in detecting. It is a given element of the problem, characterizing the fault that we want to detect. In order to detect the fault at a certain time T_d, from (16) it follows that

$$2\alpha \sum_{h=T_0}^{T_d-1} \lambda^{T_d-1-h} \left[\mathrm{Var}[\chi^{(k)}(h+1)] \right]^{\frac{1}{2}} < \bar{\phi}^{(k)}.$$

We assume that in the initial centralized condition without consensus, the uncertainty $\chi^{(k)}$, and so also the absolute value of thresholds $\bar{\epsilon}^{(k)upp}$ and $\bar{\epsilon}^{(k)}_{low}$, are too high and so the uncertainty can hide the presence of the fault and the detection may not be possible. Therefore we use the consensus scheme explained in Section 4 to reduce the uncertainty. The goal of the optimal decomposition is to obtain the threshold $\bar{\epsilon}^{(k)upp}$ and $\bar{\epsilon}^{(k)}_{low}$ so that in the faulty case at a certain time T condition (15) is satisfied. By remembering that in the shared case, we have (13), we observe that with the consensus it is possible to reduce the uncertainty terms on the right in (16). The decomposition is designed a-priori and it is fixed during the monitoring process. In applications it is typically possible to analyze off-line a number of typical modes of behaviour, thus identifying in which conditions the fault we want to detect commonly happens and its features in terms of magnitude. We assume to be able to define the following quantities

$$\bar{U}^{(k)}(w) = 2\alpha \sum_{h=T_0}^{T_d-1} \lambda^{T_d-1-h} \left[\mathrm{Var}[\chi^{(k)}(h+1)] \right]^{\frac{1}{2}}$$

and

$$\bar{\phi}^{(k)} \leq \left| \sum_{h=T_0}^{T_d-1} \lambda^{T_d-1-h} \phi^{(k)}(h) \right|.$$

The first one represents the integral of the uncertainty standard deviation at the desired detection time T_d. The second quantity represents the minimum fault magnitude we aim to detect [5]. The decomposition is designed to detect all the faults acting on the k-th component having a magnitude greater or equal to $\bar{\phi}^{(k)}$. In this subsection, we have defined the detectability conditions in the shared and non-shared case.

6. THE DECOMPOSITION PROBLEM

In this section, we address the solution of the optimal decomposition problem. This implies the optimal choice of which state variables have to be shared between suitable LFDs in order to guarantee the detectability of a given class of faults. At the same time, we take into account the inherent trade-off between sharing and the communication costs and hardware costs due to the addition of sensors. The goal is to define the minimum number of needed LFDs and assign each variable to one or more LFDs.

We refer to the optimization problems defined in (1) and (2). We are now able to characterize all the terms in a quantitative way. The computational complexity can be quantified by the number of nodes in each subsystem I denoted as n_I. We formulate the decomposition problem, that is the covering Ω of the set of the variables of the system, as a two-phases optimization problem. Given the maximum number of nodes for each subsystem n^+, (1) becomes

$$\begin{aligned} \mathcal{P}_1(n^+) \; : \; N^* = \min_{N,w,d_k} \; & N \\ \text{s.t.} \\ & \bar{U}^{(k)}(w) < \bar{\phi}^{(k)} \quad \forall \; k = 1, \dots, n \\ & n_I \leq n^+, \quad \forall I = 1, \dots, N. \end{aligned} \tag{17}$$

We define the objective function of problem (2): the communication cost of a decomposition can be computed as

C = number of cut edges + number of shared variables.

Given the optimal values obtained from (17), (2) can then be rewritten as

$$\begin{aligned} \mathcal{P}_2(N^*, d_k^*, n^+) \; : \; \min \; & C \\ \text{s.t.} \\ & N = N^* \\ & \text{satisfy } d_k^*, \; \forall \; k = 1, \dots, n \\ & n_I \leq n^+ \forall I = 1, \dots, N. \end{aligned} \tag{18}$$

6.1 Optimal Decomposition Algorithm

We here explain how to solve in an analytical way the problem (17): to find the minimum number of subsystems

[5] Let us for example consider a leakage fault: $\bar{\phi}^{(k)}$ could represent the minimum quantity of water loss in the interval time between T_0 and T_d that we are able to detect.

and the variables to be shared. Let us assume for the sake of simplicity that all the sensors related to different LFDs and measuring the same shared component s of the state are affected by the same level of uncertainty, that is, $\mu_{\xi I}^{(s)} = \mu_{\xi J}^{(s)}$ and $\sigma_{\xi I}^{(s)} = \sigma_{\xi J}^{(s)}$ $\forall I$ and J in $\mathcal{O}_s$. Moreover, we assume that in the consensus protocol each LFD weights equally all the terms of the consensus:

$$w_s^{(I,J)} = \frac{1}{d_s}, \quad \forall J \in \mathcal{O}_s,$$

being the overlap degree d_s equal to the cardinality of the set $\mathcal{O}_s$. Therefore, the variance term of the threshold $\sigma_{\bar{\epsilon}}^2$ of the non-shared case decreases linearly with the number of LFDs sharing that component in the shared case:

$$\sum_{J \in \mathcal{O}_s} (w_s^{(I,J)})^2 \sigma_{\bar{\epsilon}}^2 = d_s \cdot \left(\frac{1}{d_s}\right)^2 \sigma_{\bar{\epsilon}}^2 = \frac{1}{d_s}\sigma_{\bar{\epsilon}}^2$$

being $\sigma_{\bar{\epsilon}}^2 = \mathrm{Var}[\chi^{(k)}(t+1)] = \mathrm{Var}[\Delta f^{(k)}(t)] + \sigma_{\eta^{(k)}}^2(t) + \sigma_{\xi^{(k)}}^2(t+1) + 2\mathrm{Cov}[\Delta f^{(k)}(t), \xi^{(k)}(t+1)]$.

We implement the following steps:

(1) We identify the set $\mathbb{S}$ collecting the nodes s for which the detectability constraint is not satisfied, that is, $\bar{U}^{(s)}(w=1) > \bar{\phi}^{(s)}$ in the non-shared case.

(2) By using the rule we expressed about the linear decreasing of the threshold variance term, we compute, for each $s \in \mathbb{S}$, the minimum number N_{min}^s of LFDs sharing s we need to satisfy the detectability constraint. In fact, we want that

$$2\alpha \left[\frac{1}{N_{min}^s}\sigma_{\bar{\epsilon}}^2\right]^{\frac{1}{2}} \sum_{h=T_0}^{T_d-1} \lambda^{T_d-1-h} < \bar{\phi}^{(s)}.$$

We have $d_k^* = 1$ for $k \notin \mathbb{S}$, while for $s \in \mathbb{S}$:

$$d_s^* = N_{min}^s > \left(\frac{2\alpha\sigma_{\bar{\epsilon}} \sum_{h=T_0}^{T_d-1} \lambda^{T_d-1-h}}{\bar{\phi}^{(s)}}\right)^2.$$

(3) The minimum number of LFDs needed is

$$N^* = \max \left[\left\lceil \frac{n}{n^+} \right\rceil, \max_{s \in \mathbb{S}} N_{min}^s\right]$$

where the first argument of the max operator guarantees the fulfillment of the computational complexity constraint, and the second the detectability constraint.

(4) We can then solve the communication cost minimization problem (18) to assign each variable to a subsystem. The optimal solution of this problem requires the exploration of all the possible solutions. Due to obvious scalability problems of the last step of this algorithm, in the following simulations section we propose an heuristic algorithm to solve this second optimization problem for large-scale systems.

7. SIMULATION RESULTS

In this section we provide some preliminary simulation results. We tested the proposed decomposition and FD approach on a 12 tanks system (Fig. 1 left, where each node represents a state variable). Three pumps are present, feeding the fourth, sixth and tenth tanks with the following flows: $u_i = 1 + 0.5 \cdot \sin(0.1 \cdot t)$, $i = 1, 2, 3$. The nominal tank sections, interconnecting pipe cross-sections and drain pipes cross sections are set to 1 m^2, 0.2 m^2 and 0.02 m^2 respectively. All the pipes outflow coefficients are unitary. The actual cross-sections are affected by random

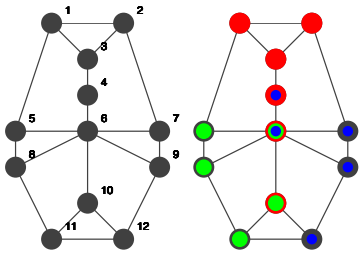

Fig. 1. The 12 tanks system graph and its optimal decomposition. Each node colour represents a different subsystem. Shared variables have more than one colour.

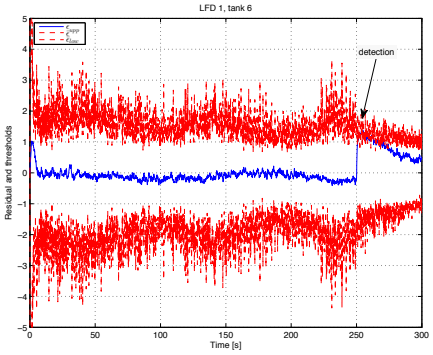

Fig. 2. Thresholds and residual signal for LFD 1 tank 6.

uncertainties with variance of 2.5% of the nominal values. Moreover, the tank levels measurements are affected by measurement noises with zero-mean and 0.05 standard deviation. The considered fault is modelled as an incipient leakage (with time profile parameter 8) affecting the fourth, sixth and tenth tanks, occurring at time 250 s. It is represented for each tank as a circular hole of 0.13 m radius in tank bottom, so that the outflow coefficient due to the leak is 1. We implemented an off-line Monte Carlo method for numerically estimating the mean and variance of the non linear function Δf. The procedure in Section 6.1 identifies nodes 4, 6 and 10 to be shared, two, three and two times, respectively: it is interesting to note how, being the model and uncertainty parameters equal for all the nodes, the detectability condition led to sharing more node 6, which has the highest degree and so is more influenced by the uncertainty of the neighboring nodes. Finally we assume that the constraint on the maximum number of nodes for each subsystem is 6: this implies at least 2 subsystems, but the detectability analysis eventually leads to a minimum number of 3 subsystems. In order to find the minimum communication cost decomposition, Point 4 of the algorithm in Section 6.1 is addressed by a modification of the multilevel partitioning scheme of Karypis and Kumar (1998) where the coarsening, initial partitioning and successive refinement stages are constrained so that the sharing of the nodes in $\mathbb{S}$, as well as the bound on the maximum number of nodes per subsystem, are automatically enforced. We obtain the optimal decomposition in Fig.1 right, having $x_1 = [1, 2, 3, 4, 6, 10]$, $x_2 = [5, 6, 8, 10, 11]$ and $x_3 = [4, 6, 7, 9, 12]$. Once the system decomposition was obtained by this off-line procedure, we tested the proposed distributed stochastic FD scheme, with $\alpha = 3$. We see in Fig. 2 that the designed optimal system decomposition allows the detection of the fault at time 251.2 s. Moreover, all the LFDs monitoring faulty variables are able to detect the fault [6].

[6] The residual decreases after fault detection because the tank is getting empty.

Table 1. Performance comparison

Decomposition	DD [s]	TRR before	TRR after	COMM	COMP
Optimal	1.20	357.33	1.61	9.33	5.33
Other 1	2.20	322.95	1.95	7.67	5.00
Other 2	1.90	863.04	6.51	9.33	5.66
Non overlapping	∞	740.00	12.55	5.33	4.00

We compared the performances obtained by the same FD stochastic approach with different decompositions. We chose the following decompositions (Fig.3): the first

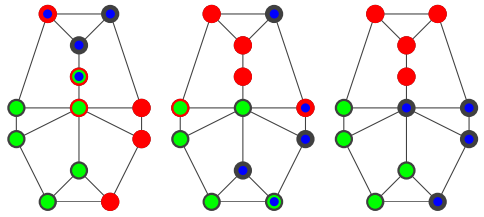

Fig. 3. The non-optimal decompositions considered for the comparison (Other 1, Other 2 and Non overlapping).

is an overlapping decomposition sharing the same variables as the optimal one, but with different cardinality of the overlap sets; the second is an overlapping decomposition as well, but sharing different variables, while the third is a non-overlapping decomposition minimizing the communication and computation costs, irrespective of the detectability condition. In order to compare the performances obtained by the different FD topologies, we introduce in Table 1 the following performance indicators: detection delay (DD); threshold/residual ratio before (TRR before) and after fault (TRR after) averaged on the variables affected by the fault function; communication (denoted in the table as COMM) and computation cost (average number of nodes of each subsystem, denoted as COMP). Table 1 shows that the designed optimal decomposition detects the fault with the lowest delay. Furthermore, it has the lowest threshold/residual ratio after fault, meaning that the thresholds are less conservative than the other cases. It is important to note that the non-overlapping decomposition is not able to detect the considered fault, even if it is the best decomposition in the sense of minimum communication and computation costs. In decompositions Other 1 and Other 2, the LFDs are not able to detect the fault for all the involved variables. Finally, we observe that the proposed optimal decomposition has communication and computation costs similar to those of the other cases with bigger detection delay.

8. CONCLUDING REMARKS

In this preliminary paper, a novel algorithm for finding an optimal decomposition of large-scale nonlinear systems is proposed in order to design a distributed fault detection architecture able to guarantee fault detectability in some defined standard conditions under communication and computation constraints. A novel stochastic framework is proposed for the problem of distributed fault detection and the effectiveness of the adoption of an overlapping decomposition is demonstrated.

REFERENCES

Anderson, J. and Papachristodoulou, A. (2012). A decomposition technique for nonlinear dynamical system analysis. *Automatic Control, IEEE Trans. on*, 57(6), 1516–1521.

Boem, F., Ferrari, R.M.G., and Parisini, T. (2011). Distributed fault detection and isolation of continuous-time nonlinear systems. *Europ. J. of Control*, (5-6), 603–620.

Boem, F., Ferrari, R.M., Parisini, T., and Polycarpou, M.M. (2013a). Distributed fault diagnosis for continuous-time nonlinear systems: The input–output case. *Annual Reviews in Control*, 37(1), 163–169.

Boem, F., Xu, Y., Fischione, C., and Parisini, T. (2013b). Distributed fault detection using sensor networks and pareto estimation. In *Proc. of European Control Conf.*, 932–937. IEEE.

Bregon, A., Daigle, M., Roychoudhury, I., Biswas, G., Koutsoukos, X., and Pulido, B. (2014). An event-based distributed diagnosis framework using structural model decomposition. *Artificial Intelligence*, 210, 1–35.

Ferrari, R., Parisini, T., and Polycarpou, M. (2012). Distributed fault detection and isolation of large-scale discrete-time nonlinear systems: An adaptive approximation approach. *Automatic Control, IEEE Trans. on*, 57(2), 275 –290.

Grbovic, M., Li, W., Xu, P., Usadi, A.K., Song, L., and Vucetic, S. (2012). Decentralized fault detection and diagnosis via sparse PCA based decomposition and maximum entropy decision fusion. *J. of Process Control*, 22(4), 738 – 750.

Karypis, G. and Kumar, V. (1998). A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on scientific Computing*, 20(1), 359–392.

Keliris, C., Polycarpou, M.M., and Parisini, T. (2015). A robust nonlinear observer-based approach for distributed fault detection of input–output interconnected systems. *Automatica*, 53, 408–415.

Ocampo-Martinez, C., Bovo, S., and Puig, V. (2011). Partitioning approach oriented to the decentralised predictive control of large-scale systems. *J. of process control*, 21(5), 775–786.

Reppa, V., Polycarpou, M., and Panayiotou, C. (2015). Distributed sensor fault diagnosis for a network of interconnected cyberphysical systems. *Control of Network Systems, IEEE Trans. on*, 2(1), 11–23.

Schloegel, K., Karypis, G., and Kumar, V. (2000). Parallel multilevel algorithms for multi-constraint graph partitioning. In *Euro-Par 2000 Parallel Processing*, 296–310. Springer.

Šiljak, D. (1978). *Large-Scale Dynamic Systems: Stability and Structure*. North Holland, New York.

Staroswiecki, M. and Amani, A.M. (2014). Fault-tolerant control of distributed systems by information pattern reconfiguration. *Int. J. of Adaptive Control and Signal Processing*.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N., and Yin, K. (2003). A review of process fault detection and diagnosis. Part I: Quantitative model-based methods. *Computers & Chem. Eng.*, 27, 293–311.

Yin, S., Ding, S.X., Xie, X., and Luo, H. (2014). A review on basic data-driven approaches for industrial process monitoring. *Industrial Electronics, IEEE Trans. on*, 61(11), 6418–6428.

Zhang, Q. and Zhang, X. (2012). A distributed detection scheme for process faults and sensor faults in a class of interconnected nonlinear uncertain systems. In *IEEE 51st Conf. on Decision and Control*, 586–591.