Performance trade-offs for networked jump observer-based fault diagnosis

Daniel Dolz, Ignacio Peñarrocha, Roberto Sanchis

***Abstract*—In this paper, we address the fault diagnosis problem for discrete-time multi-sensor systems over communication networks with measurement dropouts. We use the measurement outcomes to model the measurement reception scenarios. Based on this, we propose the use of a jump observer to diagnose multiple faults. We model the faults as slow time-varying signals and introduce this dynamic in the observer to estimate the faults and to generate a residual. The fault detection is assured by comparing the residual signal with a prescribed threshold. We design the jump observer, the residual and the threshold to attain disturbance attenuation, fault tracking and detection conditions and a given false alarm rate. The false alarm rate is upper bounded by means of Markov's inequality. We explore the trade-offs between the minimum detectable faults, the false alarm rate and the response time to faults of the fault diagnoser. By imposing the disturbances and measurement noises to be Gaussian, we tighten the false alarm rate bound which improves the time needed to detect a fault. A numerical example is provided to illustrate the effectiveness of the theory developed in the paper.**

***Index Terms*—Fault diagnosis, false alarm rate, time to detect faults, jump linear system, dropouts.**

I. Introduction

Networked control systems have been extended to many industrial applications due to the diverse offered advantages, as the reduction on the installation cost or the increase on the flexibility, provided by the communication network [1]. In these kinds of systems, the controller unit, the sensors and the actuator are not collocated and the exchange of information is done through a shared network, leading to some network-induced issues as time delays and dropouts [2], [3]. Owing to the need for reliability, safety and efficient operation of these networked systems, model-based fault diagnosis methods [4] have been recently introduced to operate over networks [5].

Fault detection over communication networks when using an observer-based fault detection scheme is addressed by the comparison between a residual signal generated with the estimated system outputs and a threshold. The residual is conceived to balance the robustness against network effects and disturbances, and the fault sensitivity [6]–[9].

Assuring a predefined false alarm rate (FAR) is a key problem. In the majority of the networked fault detection proposals, the threshold is chosen to reduce the FAR to the minimum [10], [11], but without quantifying it. Some works as [7], [8], [12] characterize the mean and variance of the residual and use Markov's inequality to impose a desired FAR bound. However, Markov's inequality is known to be conservative [13]. The main problem to get a proper FAR bound is to obtain the probability distribution of the residual signal. In [14] the residual was computed as a quadratic form of the outputs estimation error by means of the inverse of the outputs estimation error matrix covariance given by a Kalman filter. With that, their residual signal follows a chi-squared distribution and an exact FAR can be fixed. But, to the best of the authors' knowledge, the extension to observers with predefined gains (which have less implementation cost) for networked systems with dropouts has not been addressed.

Regarding the fault estimation problem, the most common approach is to make the residual track the fault or a weighted fault signal by guaranteeing some performances of the fault estimation error under disturbances and the network issues [15]–[19]. Recently, to improve the fault estimation performances, the authors in [20] introduced a dynamic of the fault signal on the fault estimator. Fault detection and estimation can be combined to attain fault diagnosis.

According to [21], the performance of a fault detection algorithm is defined by means of the trade-offs between the time to detect a fault and the FAR. This definition can be extended to the fault diagnosis case by considering also the convergence speed of a norm of the fault estimation error. The authors in [22] show that there exists a trade-off between the fault detection rate and the FAR. More recently, the existence of a compromise between the time to detect a fault and the fault sensitivity has been demonstrated in [23]. Nevertheless, none of them explores the compromises between the minimum detectable faults, the FAR and the fault diagnosis (detection and estimation) speed.

The dropouts in the fault diagnosis problem over communication networks have been mainly studied in the packetized case [7], [8], [15]. The multi-sensor case was studied in [24] with an invariant observer gain approach, however the use of jump observers that adapt their gains to the network scenario has been proved to enhance the estimation performances [25], [26]. Networked jump observer-based fault estimators have recently started to receive attention [6], [8].

Motivated by the previous analysis, in this paper we face the fault diagnosis problem for multi-sensor systems with dropouts through the combination of fault detection and fault estimation. The faults are characterized as slow-time varying signals and the network dropouts are modeled with the combination of

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

Daniel Dolz, Ignacio Peñarrocha, Roberto Sanchis are with Department of Industrial System Engineering and Design, University Jaume I of Castellón , Spain {ddolz,ipenarro,rsanchis}@uji.es.

This work has been funded by MICINN project DPI2011-27845-C02-02 from the Spanish government and by project P1·1B2013-51 and grant PREDOC/2011/37 from Universitat Jaume I.

available measurements at the fault diagnoser. We introduce a jump observer to estimate the faults and define the residual signal as a quadratic form of the estimated fault vector. The design of the jump observer and residual is addressed through an iterative linear matrix inequalities (LMIs) procedure that allows obtaining the predefined set of observer gains and the fault detector parameters. The design is carried out to achieve disturbance and measurement noise attenuation, and fault diagnosis performances under a prescribed FAR. We propose two design strategies: the first one consists of fixing the response speed to faults and minimizing the minimum detectable fault, and the second one consists of fixing the minimum detectable fault and minimizing the response time. The trade-offs between the minimum detectable faults, the FAR and the delay between fault occurrence and detection (response time of the fault estimator) are highlighted. Furthermore, we derive two ways of bounding the FAR depending on whether the residual signal probability distribution is unknown (Markov's inequality approach) or known as a result of assuming Gaussian disturbances and measurement noises (chi-squared approach).

Notation : Let A and B be some matrices. $A(i,i)$ defines the i-th diagonal element of A. The maximum and minimum eigenvalues of A are denoted by $\overline{\lambda}(A)$ and $\underline{\lambda}(A)$ respectively. $A \preceq B$ means that matrix $A - B$ is negative semidefinite. Similar applies to $\succeq$. The direct sum is represented by $\bigoplus$, where $A \bigoplus B$ is a block diagonal matrix with A and B on its diagonal. Operator $\mathrm{vec}(A)$ generates a vector by stacking the columns of matrix A. Let $x[t] \in \mathbb{R}^n$ be a stochastic process. Expected value and probability are denoted as $\mathbf{E}\{\cdot\}$ and $\mathbf{Pr}\{\cdot\}$. We write $\|x[t]\|_2^2 \triangleq x[t]^T x[t]$ for the ℓ_2 norm of $x[t]$, $\|x\|_\infty \triangleq \max_t \max_i |x_i[t]|$ for the ℓ_∞ norm of x and $\|x\|_{\mathrm{RMS}}^2 \triangleq \lim_{K\to\infty} \sum_{t=0}^{K-1} \frac{1}{K}\|x[t]\|_2^2$ for its RMS norm.

II. Problem Formulation

Let us consider linear time invariant discrete-time systems defined by equations

$$x[t+1] = A\,x[t] + B_u\,u[t] + B_w\,w[t] + B_f\,f[t], \qquad (1)$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^{n_u}$ is the vector of known inputs, $w \in \mathbb{R}^{n_w}$ is the state disturbance assumed as a random signal, uncorrelated in time, with zero mean and known covariance matrix $\mathbf{E}\{w[t]^T w[t]\} = W$ for all t, and $f \in \mathbb{R}^{n_f}$ is the fault vector. Throughout this work we assume that the known input u is causally available at all times, see Fig. 1. This general model includes as a particular case a system without known inputs, by simply taking $B_u = 0$.

The measurable outputs of the system are modeled by equation

$$y[t] = C\,x[t] \qquad (2)$$

where $y \in \mathbb{R}^{n_y}$ is the output vector.

Different sensors with different characteristics on sampling rate or noise, that may have faults, can be connected to one single measurable output, but at least each measurable output is measured by one sensor, having $n_m \geq n_y$ sensors. We define the measurement value as

$$m_j[t] = c_j\,x[t] + h_j\,f[t] + v_j[t], \quad j = 1,\ldots,n_m \qquad (3)$$

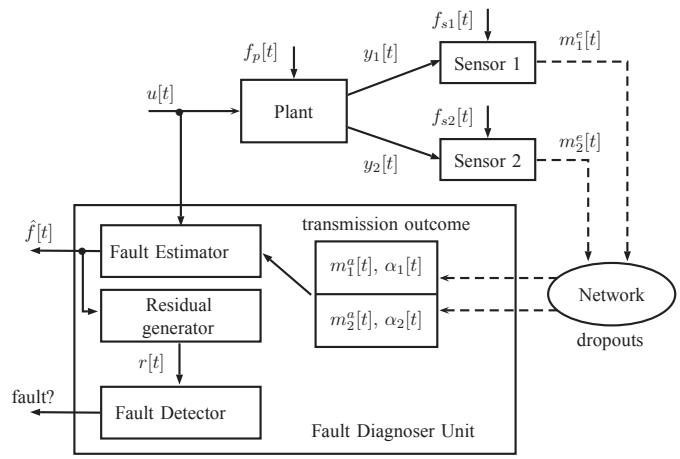

Fig. 1. Networked fault diagnosis problem for two sensors with possible faults in the plant (f_p for actuators and other faulty components) and in the sensors (f_{s1}, f_{s2}).

where $m_j[t] \in \mathbb{R}$ represents the t-th measurement of the j-th sensor and $v_j[t] \in \mathbb{R}$ the j-th sensor noise assumed as a zero mean random signal with known variance $\mathbf{E}\{v_j[t]^2\} = \sigma_j^2$ for all t, that is uncorrelated with respect to the time index t. We also consider that v_i is mutually uncorrelated with $v_{j\neq i}$. c_j denotes one row of matrix C (several c_j could be equal and correspond to the same row of C) and h_j denotes each one of the rows of matrix H.

In the current work, we model the fault signal as a slow time-varying one (cf. [20], [27]), i.e.,

$$f[t+1] = f[t] + \Delta f[t] \qquad (4)$$

where $\Delta f[t]$ is the variation of the fault from instant t to $t+1$. Equation (4) allows modeling, for instance, step signals ($\Delta f[t]$ only takes a nonzero value at the time the fault appears) or ramp signals ($\Delta f[t]$ takes a constant value), that have been widely used in the literature to analyze the behavior of fault detection algorithms [4]. Along this paper, we consider that $w[t]$, $v_j[t]$ for all $j = 1,\ldots,n_m$ and $\Delta f[t]$ are mutually uncorrelated for all t.

We introduce an extended order model to include the fault dynamic as

$$z[t+1] = \bar{A}z[t] + \bar{B}_u u[t] + \bar{B}_w w[t] + \bar{B}_f \Delta f[t] \qquad (5)$$

with $z[t] = \begin{bmatrix} x[t]^T & f[t]^T \end{bmatrix}^T$ and

$$\bar{A} = \begin{bmatrix} A & B_f \\ 0 & I \end{bmatrix},\ \bar{B}_u = \begin{bmatrix} B_u \\ 0 \end{bmatrix},\ \bar{B}_w = \begin{bmatrix} B_w \\ 0 \end{bmatrix},\ \bar{B}_f = \begin{bmatrix} 0 \\ I \end{bmatrix}$$

where $z \in \mathbb{R}^{\bar{n}}$ with $\bar{n} = n + n_f$.

In this work we intend to detect and estimate (diagnose) the possible system faults (represented by vector $f[t]$) when the measurements are transmitted through a communication network that may induce dropouts. In this case, the system output measurements are not available at every discrete time instant. When the dropout rate is high, the fault estimation problem becomes more difficult and the importance of a fast response to faults and a low FAR becomes more evident.

The measured and transmitted value from sensor j at instant t is

$$m_j^e[t] = \bar{c}_j z[t] + v_j[t], \tag{6}$$

with $\bar{c}_j = [c_j \ h_j]$ and $j = 1, \ldots, n_m$. We assume that the pair $(\bar{A}, \bar{C})$ is detectable (being $\bar{C}$ the matrix whose rows are $\bar{c}_j$).

Remark 1. *If the pair $(\bar{A}, \bar{C})$ is not detectable (i.e., $n_f > n_m$), only a combination of the faults can be detected. Then, a previous transformation of the system, as proposed in [28], must be done (leading to new $\bar{n}_f$ faults such that $\bar{n}_f \leq n_m$) before the proposed technique becomes applicable.*

A. Network transmissions characterization

Each sensor samples its output synchronously with the known input update and sends independently a time-tagged packet with the measurement $m_j^e[t]$ to the fault diagnoser station, through an unreliable communication network with packet dropouts.

We define the binary variable $\alpha_j[t]$ that indicates the availability of the j-th sensor measurement ($j = 1, \ldots, n_m$) at each instant t, as.

$$\alpha_j[t] = \begin{cases} 0 & \text{if } m_j^e[t] \text{ is not received at } t, \\ 1 & \text{if } m_j^e[t] \text{ is received at } t. \end{cases} \tag{7}$$

Then, the availability matrix $\alpha[t] = \bigoplus_{j=1}^{n_m} \alpha_j[t]$ is a binary diagonal matrix that can only have ones in its diagonal. Thus, using $\alpha[t]$ we can redefine the available measurements at instant t as

$$m^a[t] = \alpha[t] \left(\bar{C} z[t] + v[t] \right). \tag{8}$$

Note that a component of vector $m^a[t]$ is null when the corresponding measurement is not available. $v[t] = [v_1[t] \cdots v_{n_m}[t]]^T$ is the measurement noise vector with covariance $\mathbf{E}\{v[t]v[t]^T\} = V = \bigoplus_{j=1}^{n_m} \sigma_j^2$ (for all t).

The possible values of $\alpha[t]$ at each instant t belong to a known finite set

$$\alpha[t] \in \Xi = \{\eta_0, \eta_1, \ldots, \eta_q\}, \quad q = 2^{n_m} - 1, \tag{9}$$

where η_i denotes each possible combination of the available measurements at the fault diagnoser station (measurement reception scenario). Matrix η_0 denotes the scenario in which there is no measurement available, i.e., $\eta_0 = 0$. We characterize the network behavior using the total probability of each scenario in Ξ. We denote by $p_i = \mathbf{Pr}\{\alpha[t] = \eta_i\}$ the probability of having the measurement reception scenario η_i at instant t. p_0 denotes the probability of having no measurements.

In the current paper, we assume that the arrival probability from each sensor is governed by an independent and identically distributed process [29]. We denote by β_j the probability of having available the measurement from sensor j at instant t, i.e., $\beta_j = \mathbf{Pr}\{\alpha_j[t] = 1\}$. Then, the probability of having a given combination of available measurements $\eta_i \in \Xi$ is

$$p_i = \mathbf{Pr}\{\alpha[t] = \eta_i\} = \prod_{j \in \mathcal{I}(\eta_i)} (1 - \beta_j) \prod_{j \notin \mathcal{I}(\eta_i)} \beta_j \tag{10}$$

for all $i = 0, \ldots, q$ where $\mathcal{I}(\eta_i) \triangleq \{j | \eta_i(j,j) = 0\}$.

B. Fault diagnosis method

We propose the following fault estimation algorithm for system (5)-(6). At each instant t, the model is run in open loop leading to

$$\hat{z}[t^-] = \bar{A}\, \hat{z}[t-1] + \bar{B}_u u[t-1]. \tag{11}$$

If no measurement is received, we keep the open loop estimation, i.e., $\hat{z}[t] = \hat{z}[t^-]$. If a measurement arrives at instant $t = t_k$, the state is updated as

$$\hat{z}[t_k] = \hat{z}[t_k^-] + L[t_k] \left(m^a[t_k] - \alpha[t_k]\, \bar{C}\, \hat{z}[t_k^-] \right), \tag{12}$$

where $L[t_k]$ is the updating gain matrix and $m^a[t_k]$ is defined in (8).

Remark 2. *While $t \in \mathbb{N}$ refers to each time instant, t_k (with $k \in \mathbb{N}$) enumerates only the instants where some measurements are received. For instance, if we receive some measurements only at instants $t_k = 8$ and $t_{k+1} = 10$, but not at $t = 9$, then instant $t_k + 1 = 9$ (or $t_{k+1} - 1 = 9$) refers to instant 9, when no measurement is received.*

Let us denote $z[t_k]$ by z_k. Defining the extended state estimation error at updating instants as $\tilde{z}_k = z_k - \hat{z}_k$, the estimation error dynamics is given by

$$\begin{aligned} \tilde{z}_k =& (I - L_k \alpha_k \bar{C}) \bar{A}^{N_k} \tilde{z}_{k-1} - L_k \alpha_k v_k \\ &+ \sum_{l=1}^{N_k} (I - L_k \alpha_k \bar{C}) \bar{A}^{l-1} B_{\mathbb{W}} \mathbb{W}[t_{k-1} + l - 1] \end{aligned} \tag{13}$$

being $B_{\mathbb{W}} = [\bar{B}_w \ \ \bar{B}_f]$ and $\mathbb{W}[t_{k-1} + l - 1] = [w[t_{k-1} + l - 1]^T \ \ \Delta f[t_{k-1} + l - 1]^T]^T$. N_k denotes the number of consecutive instants without measurements (which is unbounded), i.e., $N_k = t_k - t_{k-1}$.

The fault detection algorithm uses the estimated faults to compute a residual signal at instants $t = t_k$ as

$$r_k = \hat{f}_k^T F^{-1} \hat{f}_k, \tag{14}$$

where the common fault detection decision is given by

$$\begin{cases} \text{if } r_k \leq r_{\text{th}} & \text{No fault} \\ \text{if } r_k > r_{\text{th}} & \text{Fault} \end{cases}$$

being $r_{\text{th}} > 0$ a threshold to be defined. Then, fault isolation is achieved by means of the combination of fault detection and fault estimation, allowing us to identify which is the origin of the fault.

Remark 3. *According to [4], the minimum detectable fault is a fault that drives the residual to its threshold, provided no other faults, disturbances and measurement noises are present. Then, assuming a zero fault estimation error (i.e. $\hat{f} = f$), each diagonal element of F in* (14) *multiplied by r_{th} defines the minimum detectable fault as $f_{\min,l} = r_{\text{th}} F(l,l)$ for the corresponding channel ($l = 1, \ldots, n_f$).*

Considering the fault detection logic, the FAR is defined as the average probability of rising false alarms over an infinite-time window, i.e.

$$\Psi = \lim_{K \to \infty} \sum_{k=0}^{K-1} \mathbf{Pr}\{r_k > r_{\text{th}} \,|\, f_k = 0\}. \tag{15}$$

The aim of this work is to compute the gain matrices L_k, the matrix F, and the threshold $r_{\rm th}$ such that the fault diagnoser attains disturbance and measurement noise attenuation, and fault diagnosis performances for a given FAR. These objectives can be reached with an invariant observer gain (as in the majority of reviewed works), or with a jump one (e.g. [6], [8]). In this work, we relate the gain L_k to the sampling scenario α_k, as $L_k = L(\alpha_k)$, with the law $L_k = L_i$ when $\alpha_k = \eta_i$ for $\alpha_k = \eta_1, \ldots, \eta_q$. Then, the matrices are computed off-line leading to the finite set

$$L_k \in \mathcal{L} = \{L_1, \ldots, L_q\}. \tag{16}$$

III. Fault diagnoser design: dropout-free

Let us first consider the case without measurement dropouts, i.e., $\beta_j = 1$ for all $j = 1, \ldots, n_m$. In this case, $\alpha[t]$ is always the identity, which implies that each instant t is a measurement instant ($t_k = t$) leading to $L_k = L$ and $N_k = 1$, for all k. The following theorem presents how to design the observer gain L and the matrix F that defines the residual (14) based on the H_2 norm of system (13).

Theorem 1. *Consider the estimation algorithm* (11)-(12) *applied to system* (1)-(4) *with standard sampling. If there exist symmetric matrices P, F, Γ_w, Γ_v, Γ_f, and full matrix X fulfilling*

$$\begin{bmatrix} P & \bar{\mathcal{A}} & 0 \\ \bar{\mathcal{A}}^T & P & \bar{B}_f \\ 0 & \bar{B}_f^T & F \end{bmatrix} \succeq 0, \tag{17a}$$

$$\begin{bmatrix} P & \bar{\mathcal{B}}_\mu \\ \bar{\mathcal{B}}_\mu^T & \Gamma_\mu \end{bmatrix} \succeq 0, \ \mu = \{w, v, f\} \tag{17b}$$

with

$$\bar{\mathcal{A}} = (P - X\bar{C})\bar{A}, \ \bar{\mathcal{B}}_w = (P - X\bar{C})\bar{B}_w,$$
$$\bar{\mathcal{B}}_v = -X, \ \bar{\mathcal{B}}_f = (P - X\bar{C})\bar{B}_f$$

then, defining the observer gain matrices as $L = P^{-1}X$, the following statements hold:

i) In the absence of disturbances, faults, and measurement noises the extended state estimation error (13) *converges to zero.*

ii) Under zero initial conditions (i.e., $\tilde{z}_0 = 0$), the fault estimation error is bounded by

$$\mathbf{E}\{\|\tilde{f}\|^2_{\rm RMS}\} \leq \overline{\lambda}(F)\left({\rm tr}(\bar{\Gamma}) + n_f\overline{\lambda}(\Gamma_f)\Delta f^2_{\max}\right), \tag{18}$$

where $\bar{\Gamma} = \Gamma_w W + \Gamma_v V$ and $\|\Delta f\|_\infty \leq \Delta f_{\max}$.

Proof. See Appendix A. □

The above theorem states that F is related to the expected value of the squared RMS norm of the fault estimation error. We can extract from (18) that the fault estimation (and therefore the residual signal) is more sensitive to disturbances and measurement noises when the maximum of the minimum detectable faults (by means of $\overline{\lambda}(F)$) is higher. Furthermore, the lower the value $\overline{\lambda}(\Gamma_f)$, the lower the effect of the faults on the estimation error. The next theorem extends the results of the previous one to bound the FAR.

Theorem 2. *For a given threshold $r_{\rm th} > 0$ and $0 \leq \phi \leq 1$, and under the premisses of Theorem 1, if*

$${\rm tr}(\Gamma_w W) + {\rm tr}(\Gamma_v V) = \phi\, r_{\rm th}, \tag{19}$$

and constraints (17) *are fulfilled, then, the following additional statement holds:*

iii) In the absence of faults and under zero initial conditions, the fault detection algorithm assures a FAR (15) *bounded by ϕ.*

Proof. See Appendix B. □

The next theorem extends the previous one showing how the fault estimation error decays at each measurement instant.

Theorem 3. *For a given threshold $r_{\rm th} > 0$ and $0 \leq \phi \leq 1$, and under the premisses of Theorem 2, if*

$$\Gamma_f - \bar{B}_f^T P \bar{B}_f \succeq 0, \tag{20}$$

and constraints (17), (19) *are fulfilled, then, the following additional statement holds:*

iv) The fault estimation error given by $\mathbf{E}\{\|\tilde{f}_k\|_2^2\}$ decays with

$$\rho = 1 - \frac{1}{\overline{\lambda}(\Gamma_f F)}. \tag{21}$$

Proof. See Appendix C. □

The above theorem shows that $\mathbf{E}\{\|\tilde{f}_k\|_2^2\}$ decays with ρ, from the initial conditions to the steady state region (see (41)). ρ depends on the maximum eigenvalue of the product $\Gamma_f F$. If F is fixed to assure the detection of some given minimum faults, Γ_f determines the response time of the fault estimator (by means of ρ) and therefore the time to detect a fault (as the residual is defined with the estimated faults).

Remark 4. *Under a step-like fault, the number of instants with measurement reception, denoted by $\mathcal{K}$, until the initial value of the fault estimation error is decreased below $\xi(\%)$, characterizes the settling time of the fault estimation vector (time to achieve the $100 - \xi$ of the final value). $\mathcal{K}$ can be obtained approximately by solving equation $\rho^{k+1} = \xi/100$, (see* (41)*) leading to*

$$\mathcal{K} = \left\lceil \frac{\log(\xi/100)}{\log(\rho)} - 1 \right\rceil \tag{22}$$

where $\lceil\cdot\rceil$ is the operator that rounds its argument to the nearest integer towards infinity. One of the most used values for ξ in system theory is $\xi = 2\%$. Thus, $\xi = 2\%$ refers to the number of time instants until reaching the 98% of the fault estimation final value.

Remark 5. *For a fixed value of F, increasing the FAR by means of ϕ leads to an increase in the values of Γ_w and Γ_v, see* (19)*. Higher values on these variables alleviate the constraints over P in* (17b)*, increasing the solution space in the search for a feasible matrix P. This would allow, for instance, structure constraints over matrix P. Matrix Γ_f in* (20) *constrains the last diagonal block on P. Then, increasing ϕ can enlarge the solution space to find lower values on Γ_f, which, in turn, lead to lower values of ρ (faster*

fault diagnosers). These ideas are analyzed in the examples section.

We used Markov's inequality in Theorem 2 to bound the FAR. However, it is well known that the bound yielded by Markov's inequality may be very conservative (see [13]) because it does not consider the probability distribution of the residual r_k. This may result in a real FAR that is some orders of magnitude lower than the desired one, which, as shown in the examples, may lead to a very slow response of the fault diagnoser (characterized by ρ in Theorem 3). Most of the works in the literature share this important drawback. In order to overcome this, a more accurate bound on the FAR would be desirable. Assuming that the disturbances w_k and the measurement noises v_k are Gaussian, we show in the next theorem how to impose an appropriate value to matrix F to force the residual r_k follow a chi-squared distribution, which allows us to tighten the FAR bound.

Theorem 4. *For the fixed threshold $r_{\rm th} = n_f$ and for a given $0 \leq \phi \leq 1$, under the premisses of Theorem 3, if*

$$F = \phi^{-1}\Sigma_f \tag{23}$$

and constraints (17), (19), (20) *are fulfilled, with* [1]

$$\begin{aligned} \mathrm{vec}(\Sigma_f) &= \left(I - G\bar{A} \otimes G\bar{A}\right)^{-1} \mathrm{vec}\left(Y_1\right), \\ Y_1 &= G\bar{B}_w W \bar{B}_w^T G^T + P^{-1} X\, V\, (P^{-1}X)^T, \\ G &= (I - P^{-1}X\bar{C}), \end{aligned} \tag{24}$$

then, in the absence of faults, under zero initial conditions and Gaussian disturbances and measurement noises, if the fault diagnoser gain is defined as $L = P^{-1}X$, then the FAR is given by

$$\Psi = 1 - \mathrm{CDF}_{\mathcal{X}^2_{n_f}}\left(\frac{r_{\rm th}}{\phi}\right) \tag{25}$$

where $\mathrm{CDF}_{\mathcal{X}^2_{n_f}}(\frac{r_{\rm th}}{\phi}) = \mathbf{Pr}\{\frac{r_k}{\phi} \leq \frac{r_{\rm th}}{\phi}\}$ denotes the cumulative distribution function (CDF) of a chi-squared random variable with n_f degrees of freedom, $\mathcal{X}^2_{n_f}$.

Proof. See Appendix D. □

Remark 6. *Following the definition of the CDF of a chi-squared random variable, the value of ϕ needed to obtain a desired FAR ψ with the chi-squared approach is always higher (for any value of n_f) that the one required with the Markov's inequality approach. For instance, if $n_f = 2$ and $\psi = 10^{-3}$ using Theorem 2 requires $\phi = 10^{-3}$ while Theorem 4 requires $\phi = 0.145$. Following Remark 5, this implies that using the chi-squared approach could lead to fault diagnosers with a faster response to faults than employing the result on Theorem 2. However, Markov's inequality approach (from Theorem 2) has wider applications because it does not require Gaussian disturbances and noises, as the chi-square approach (from Theorem 4) does.*

[1]The Kronecker product between $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{p\times q}$ is a block matrix such as $A \otimes B \begin{bmatrix} a_{11}A & \cdots & a_{1m}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{bmatrix} \in \mathbb{R}^{np\times mq}$. The vectorization of matrix A is $\mathrm{vec}(A) = \begin{bmatrix} a_{11} & \cdots & a_{n1} & a_{12} & \cdots & a_{nm} \end{bmatrix}^T$.

Theorem 4 has shown how to reduce the conservatism of the approach when assuming Gaussian disturbances, but at the cost of including new nonlinear equality constraints that are hard to handle. We will show in the design strategies section how to overcome this issue.

IV. Application to networked transmission

In the previous section we presented how to design the fault diagnoser and to characterize the obtained FAR and response time to faults for measurement transmission without dropouts. In this section we extend the previous results to a more interesting case where measurement information is not always available due to network dropouts. This will stress the need of fast fault detection with a low FAR. The following theorem extends Theorem 3 and shows how to find the set of observer gain matrices (16) and the matrix that defines the residual (14).

Theorem 5. *For a given threshold $r_{\rm th} > 0$ and $0 \leq \phi \leq 1$, consider the estimation algorithm* (11)-(12) *applied to system* (1)-(4)*. Assume that there can be q different measurement reception scenarios η_i ($i = 0, \ldots, q$) with a probability p_i. If there exist symmetric matrices P, Q, F, Γ_v, Γ_w, Γ_f, and full matrices X_i fulfilling*

$$\begin{bmatrix} P - M_1 & \bar{B}_f \\ \bar{B}_f^T & F \end{bmatrix} \succeq 0, \tag{26a}$$

$$\Gamma_w - \bar{B}_w^T M_2 \bar{B}_w \succeq 0, \tag{26b}$$

$$\begin{bmatrix} \bigoplus_{l=1}^q P & M_3 \\ M_3^T & \Gamma_v \end{bmatrix} \succeq 0, \tag{26c}$$

$$\Gamma_f - \bar{B}_f^T (M_5 + M_6) \bar{B}_f \succeq 0, \tag{26d}$$

$$\begin{bmatrix} \bigoplus_{l=1}^q P & M_4 \\ M_4^T & Q \end{bmatrix} \succeq 0, \tag{26e}$$

$$\mathrm{tr}(\Gamma_w W) + \mathrm{tr}(\Gamma_v V) = \phi\, r_{\rm th}, \tag{26f}$$

$$\Gamma_f - \bar{B}_f^T P \bar{B}_f \succeq 0, \tag{26g}$$

with

$$\begin{aligned} &\mathrm{vec}(M_1) = \varphi(\bar{A})^{-1}\mathrm{vec}(\bar{A}^T Q \bar{A}), \\ &M_2 = (1-p_0)M_5 + \frac{p_0}{1-p_0}M_1, \quad M_5 = \frac{1}{(1-p_0)^2}Q, \\ &M_3 = \begin{bmatrix} \sqrt{\frac{1}{1-p_0}p_1}\, X_1\eta_1 \\ \vdots \\ \sqrt{\frac{1}{1-p_0}p_q}\, X_q\eta_q \end{bmatrix}, \quad M_4 = \begin{bmatrix} \sqrt{p_1}(P - X_1\eta_1\bar{C}) \\ \vdots \\ \sqrt{p_q}(P - X_q\eta_q\bar{C}) \end{bmatrix}, \\ &\mathrm{vec}(M_6) = \varphi(\bar{A})^{-1}\left(\mathrm{vec}(\bar{A}^T M_5 \bar{A}) + \frac{p_0}{1-p_0}\mathrm{vec}(M_1)\right), \end{aligned}$$

and $\varphi(\bar{A}) = I - p_0\bar{A}^T \otimes \bar{A}^T$, then, defining the observer gain matrices as $L_i = P^{-1}X_i$, the following statements hold:

i) In the absence of disturbances, faults and measurement noises, (13) *converges to zero in average.*

ii) Under zero initial conditions, the fault estimation error is bounded by

$$\mathbf{E}\{\|\tilde{f}\|^2_{\rm RMS}\} \leq \overline{\lambda}(F) \cdot \left(\phi\, r_{\rm th} + n_f \overline{\lambda}(\Gamma_f)\overline{\Delta f}^2_{\max}\right), \tag{27}$$

where $\overline{\Delta f}_{\max}$ is a constant that depends on the fault magnitude that bounds vector $\overline{\Delta f}_k \in \mathbb{R}^{n_f}$ as $\|\overline{\Delta f}\|_\infty \leq \overline{\Delta f}_{\max}$, being $\overline{\Delta f}_k$ a vector that fulfills for all k that

$$\sum_{N=1}^{\infty} N p_0^{N-1} \sum_{l=0}^{N-1} (\star)^T Q \underbrace{\left(\bar{A}^l \bar{B}_f \Delta f[t_k+l]\right)}_{\star} = \overline{\Delta f}_k{}^T \left(\sum_{N=1}^{\infty} N p_0^{N-1} \sum_{l=0}^{N-1} \bar{B}_f^T (\bar{A}^l)^T Q \bar{A}^l \bar{B}_f \right) \overline{\Delta f}_k. \tag{28}$$

iii) Under zero initial conditions and in the absence of faults, the residual evaluation assures a FAR (15) *bounded by ϕ.*

iv) The fault estimation error given by $\mathbf{E}\{\|\tilde{f}_k\|_2^2\}$ decays with

$$\rho = 1 - \frac{1}{\overline{\lambda}(\Gamma_f F)}. \tag{29}$$

Proof. See Appendix E. □

Remark 7. *The existence of vector $\overline{\Delta f}_k$ defined in* (28) *is assured because it represents an equality constrained problem with one equation and n_f degrees of freedom. For instance, under ramp-like faults ($\Delta f[t_k+l]$ is constant), $\Delta f[t_k+l] = \overline{\Delta f}_k$ (for all $l = 0, 1, \ldots$) and $\overline{\Delta f}_{\max} = \|\Delta f\|_\infty$. Furthermore, the exact value of $\overline{\Delta f}_{\max}$ is not relevant for the analysis.*

In the aim of reducing the conservativeness introduced by Markov's inequality to bound the FAR, the next theorem extends Theorem 4 by forcing r_k to follow a chi-squared distribution when measurements are subject to dropouts.

Theorem 6. *If the threshold is set as $r_{\text{th}} = n_f$ and for a given $0 \leq \phi \leq 1$, under the premisses of Theorem 5, if*

$$F = \phi^{-1} \Sigma_f \tag{30}$$

and constraints (26) *are fulfilled for $i = 1, \ldots, q$, with*

$$\begin{aligned}
&\Sigma_f = \bar{B}_f^T (R - p_0 \bar{A} R \bar{A}^T) \bar{B}_f, \quad \text{vec}(R) = Y_1^{-1} \text{vec}(Y_2), \\
&Y_1 = \varphi(\bar{A}) - (\sum_{i=1}^{q} p_i (G_i \bar{A}) \otimes (G_i \bar{A})), \\
&\varphi(\bar{A}) = I - p_0 \bar{A}^T \otimes \bar{A}^T, \\
&Y_2 = \frac{1}{1-p_0} \sum_{i=1}^{q} p_i L_i \, \eta_i V \eta_i^T L_i^T + \sum_{i=1}^{q} p_i G_i (S_W) G_i^T, \\
&S_W = \frac{1}{1-p_0} \left(\bar{B}_w W \bar{B}_w^T + p_0 \bar{A} S_{W,\infty} \bar{A}^T \right), \\
&\text{vec}(S_{W,\infty}) = \varphi(\bar{A})^{-1} \text{vec}(\bar{B}_w W \bar{B}_w^T), \\
&L_i = P^{-1} X_i, \;\; G_i = I - L_i \eta_i \bar{C}
\end{aligned}$$

then, in the absence of faults, under zero initial conditions and Gaussian disturbances and measurement noises, the FAR is given by (25).

Proof. See Appendix F. □

V. Fault diagnosis strategies

Based on the derived results on Theorem 5, we propose the following two strategies to address the design of a fault diagnoser depending on the needs of the application.

First, let us consider that we desire to detect faults over a certain value, i.e to fix the minimum detectable fault on each channel $f_{\min,l}$ (for $l = 1, \ldots, n_f$), with a guaranteed FAR, and to detect as fast as possible the appearance of faults (i.e., with the lowest ρ). The next optimization problem deals with this design problem.

Strategy 1. *For a given threshold $r_{\text{th}} > 0$, let ψ be the desired FAR, fix ϕ to be $\phi = \psi$, and let $\mathcal{F}$ be a diagonal matrix such that $\mathcal{F} = \bigoplus_{l=1}^{n_f} f_{\min,l}^2 / r_{\text{th}}$. Then, the minimization problem*

$$\begin{aligned}
&\text{minimize} \quad \gamma \\
&\text{subject to} \quad \mathcal{X}_1 = \{(26),\ F \preceq \mathcal{F},\ \Gamma_f \mathcal{F} \preceq \gamma I\}
\end{aligned} \tag{31}$$

along variables γ, P, Q, F, Γ_v, Γ_w, Γ_f, and X_i (with $i = 1, \ldots, q$), leads to the fault diagnoser with the fastest response under faults, able to detect faults over $f_{\min,l}$ (with $l = 1, \ldots, n_f$) with a FAR below ψ.

Remark 8. *The computational complexity of Strategy 1 can be described as follows. The size of the full involved LMI constraint is*

$$(2^{n_m} + 1)(n + n_f) + 5n_f + n_w + n_m + 1.$$

Symmetric matrices as $P \in \mathbb{R}^{(n+n_f)\times(n+n_f)}$ have $(n + n_f)(1 + n + n_f)/2$ decision variables, while full matrices as $X_i \in \mathbb{R}^{(n+n_f)\times n_m}$ have $(n+n_f)n_m$. Furthermore, Strategy 1 is based on semidefinite programming and therefore does not require a high computing capacity. This kind of problems can be solved using MATLAB toolboxes as Yalmip [30] (which can handle large scale problems).

Second, let us assume that we desire to impose the response speed under the appearance of faults (by means of ρ) with a guaranteed FAR. Then, the minimum detectable faults can be minimized through the next optimization problem.

Strategy 2. *For a given threshold $r_{\text{th}} > 0$, let ψ be the desired FAR, fix ϕ to be $\phi = \psi$, and let $\bar{\rho}$ be the given upper bound on how the fault estimation error decays, i.e., $\rho \leq \bar{\rho}$. Then, the minimization problem*

$$\begin{aligned}
&\text{minimize} \quad \gamma \\
&\text{subject to} \quad \mathcal{X}_2 = \left\{ \begin{array}{l} (26),\ \text{tr}(F) \leq \gamma, \\ \Gamma_f F \leq (1-\bar{\rho})^{-1} I \end{array} \right\}
\end{aligned} \tag{32}$$

along variables γ, P, Q, F, Γ_v, Γ_w, Γ_f and X_i (with $i = 1, \ldots, q$), leads to the fault diagnoser with the minimum value of the sum of the squared minimum detectable faults (defined by matrix F) with $\rho \leq \bar{\rho}$ and a FAR below ψ.

Remark 9. *Optimization problem* (32) *is nonlinear because of the bilinear matrix inequality (BMI) that affects the product $\Gamma_f F$. This can be solved with the following rank constrained problem*

$$\begin{bmatrix} (1-\bar{\rho})^{-1} F & F \\ F & \Lambda \end{bmatrix} \succeq 0, \ \text{rank} \left(\begin{bmatrix} \Gamma_f & I \\ I & \Lambda \end{bmatrix} \right) \leq n_f$$

where a new symmetric decision matrix Λ has been added. This problem can be iteratively handled with the well known cone complementarity linearization (CCL) algorithm [31]

(which only addresses feasibility by relaxing the rank constraint with a positive semidefinite constraint on the involved matrix) over a bisection algorithm. Solving Strategy 2 is more time consuming than Strategy 1 because of the iterations introduced by the CCL and the bisection algorithm. Nevertheless, it only introduces $n_f(1+n_f)$ decision variables (due to Λ) and only increases the full LMI size in $2n_f+1$ over a semidefinite programming problem, and therefore, the computational complexity is not really an issue.

Both design strategies are still valid when including nonlinear equality constraints (30) but need more computational effort. The next strategy extends the previous ones to consider the chi-squared approach presented in Theorem 6.

Strategy 3. *The minimization problem*

$$\begin{aligned} &\text{minimize} && \gamma \\ &\text{subject to} && \mathcal{X}_j,\ (30), \\ & && \psi = 1 - \mathrm{CDF}_{\mathcal{X}^2_{n_f}}\left(\frac{r_{\mathrm{th}}}{\phi}\right) \end{aligned} \tag{33}$$

along variables γ, P, Q, F, Γ_v, Γ_w, Γ_f, and X_i (with $i = 1,\dots,q$) with $r_{\mathrm{th}} = n_f$, extends the design made in Strategy 1, if $j=1$, or in Strategy 2, if $j=2$, to tighten the FAR bound with the chi-squared approach.

Remark 10. *Optimization problem* (33) *is nonlinear due to constraint* (30)*. This optimization problem can be solved iteratively with LMI constraints by forcing matrix F at each step k to be as $F \preceq \phi^{-1}\Sigma_f(\mathcal{L}^{k-1})$, until $\Sigma_f(\mathcal{L}^{k-1})$ converges to a constant value, where $\Sigma_f(\mathcal{L}^{k-1})$ is the covariance matrix in* (30) *evaluated with the observer gains at step $k-1$. The computational burden of each of the iterations is nearly the same as in Strategy 1 (or Strategy 2), but the total computing time is multiplied at most by the number of iterations. However, we are again dealing with a semidefinite programming problem, therefore the computational load is not a problem.*

Remark 11. *Strategy 3 will lead, in general, to minimum detectable faults under $f_{\min,l}$ (for $l = 1,\dots,n_f$). If we do not intend to detect faults under $f_{\min,l}$, we can first solve the optimization problem involved in Strategy 3 and then use $r_k = \hat{f}_k^T \mathcal{F}^{-1} \hat{f}_k$ in the real-time implementation (where $\mathcal{F}$ includes the original prescribed minimum detectable faults, $f_{\min,l}$). In this case, as we impose in the design that $\phi^{-1}\Sigma_f \preceq \mathcal{F}$, the obtained FAR will be upper-bounded by* (43)*.*

VI. Example

Let us consider an industrial continuous-stirred tank reactor process (borrowed from [32]) where the discretized state-space model is

$$A = \begin{bmatrix} 0.972 & -0.001 \\ -0.034 & 0.863 \end{bmatrix},\ B_u = \begin{bmatrix} -0.084 & 0.023 \\ 0.076 & 0.414 \end{bmatrix},$$
$$B_w = B_u,\ C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

We desire to detect faults from the second actuator and the first sensor, i.e.

$$B_f = \begin{bmatrix} 0.023 & 0 \\ 0.414 & 0 \end{bmatrix},\ H = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

The state disturbances and measurement noises are Gaussian with covariance matrices

$$W = \begin{bmatrix} 0.11 & 0.03 \\ 0.03 & 0.13 \end{bmatrix},\ V = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}.$$

We consider that the measurements are independently acquired through a communication network where the probabilities of having available the measurements from each sensor are $\beta = [0.58\ 0.46]$.

For ease of analysis, in this example we will only explore the case when we impose that the minimum detectable faults are below some given values and we try to obtain the fastest response to faults of the fault diagnoser, i.e. we will only analyze Strategies 1 and 3. For ease of notation, let us assume that the requirement over the minimum detectable faults is such that $F \preceq f_{\min} I$. In the next, we impose the threshold to be $r_{\mathrm{th}} = n_f$.

First, let us study the compromises between the minimum detectable faults f_{min}, the desired FAR ψ and the speed of the fault diagnoser by means of ρ in the design procedure. Fig. 2 illustrates these trade-offs for five different desired FARs with $\psi = [10^{-1}\ 10^{-2}\ 10^{-3}\ 10^{-4}\ 10^{-5}]$ and for the two presented approaches to assure them: through Markov's inequality (left hand side figure, Strategy 1) and through characterizing the probability distribution of the residual signal (right hand side figure, Strategy 3). We note that imposing smaller minimum detectable faults or lower FARs results in a slower response time to faults (ρ higher). We also find that forcing F to be as defined in (30) (chi-squared approach) results in a faster response under faults (ρ smaller) for the same minimum detectable faults than using Markov's inequality approach. Furthermore, Fig. 2 shows an asymptotic behavior of ρ with respect to $f_{\min}$, leading to a minimum achievable value.

Second, let us study the behaviour of some fault diagnosers in simulation, where $u[t] = 0$ for all t. Table I compares the fault diagnosis performances for the case when F is unconstrained, case C1 (where Markov's inequality approach is used, Strategy 1) and when F is constrained to be as in (30), case C2 (where the chi-squared approach is used, Strategy 3). For both cases we impose $\psi = 10^{-3}$ and $f_{\min} = 0.6$. We also include in Table I a case C3 where we reduced the $f_{\min}$ from case C2 to the half. The matrices F obtained for the three cases are:

$$F_{\mathrm{C1}} = \begin{bmatrix} 0.18 & 0 \\ 0 & 0.18 \end{bmatrix},\quad F_{\mathrm{C2}} = \begin{bmatrix} 0.161 & -0.025 \\ -0.025 & 0.107 \end{bmatrix},$$
$$F_{\mathrm{C3}} = \begin{bmatrix} 0.022 & -0.008 \\ -0.008 & 0.041 \end{bmatrix}.$$

As illustrated in Table I, for case C3, we can detect smaller faults than in case C2 at the expense of being slower than in case C2. However, we still are much faster than in case C1 where the guaranteed detectable faults were higher. Moreover, as stated in Remark 11 cases C2 and C3 can detect faults below the imposed $f_{\min}$ ($f_{\min,1}$ for the actuator fault and $f_{\min,2}$ for the sensor fault). Concerning the computational burden, obtaining C1 takes 0.4sec (using Yalmip with SeDuMi solver [33] in a i7-3770 processor at 3.40 GHz) while C2

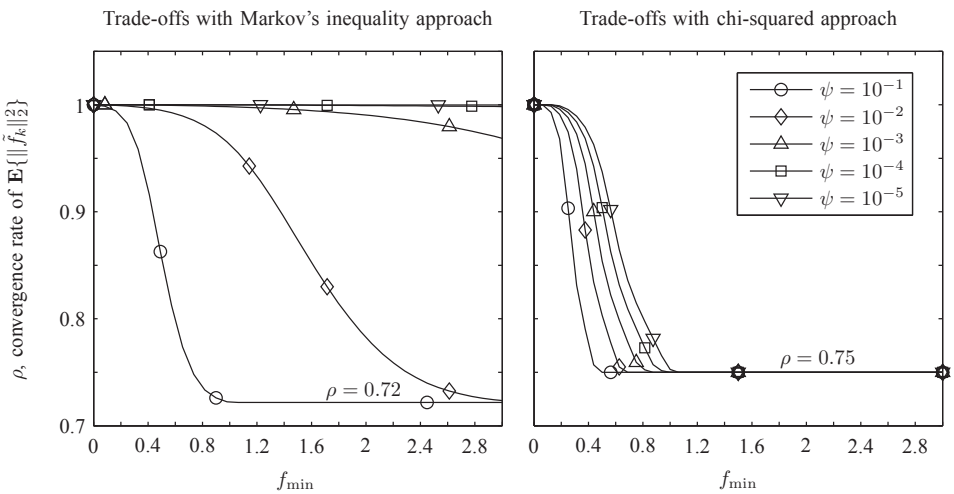

Fig. 2. Trade-offs on the observer-based fault diagnoser design.

requires 2.5sec (with 10 iterations). Note that as previously stated in Remark 8 and 10, the computational cost is not an issue.

After a simulation of 10^6 instants with no faults, we verify that the FAR obtained in simulation (by dividing the number of risen alarms by the total number of simulation time instants) for case C2 and case C3 is the same as forecasted in the design, but for case C1 is much lower (several orders of magnitude) than the imposed bound. This conservativeness of the Markov's approach results in an extremely slow residual dynamics (as seen in Fig. 3), and a huge time to detect the fault (characterized by 6101 measurement instants, see (22)), that is useless in practice. To alleviate this conservativeness, we add to the analysis a fourth case C4 (with $F_{C4} = F_{C1}$) where, as a difference from case C1, we impose $\phi = 0.1$ ($\psi \leq 0.1$). Then, we obtain a fault diagnoser similar to C2 with a FAR in simulation of 10^{-4} (see Table I), which is under the desired one of 10^{-3}. This shows that we can compensate the conservativeness of the Markov's approach by increasing the value of ϕ and then verifying in simulation if the prescribed bound is fulfilled, but we cannot guarantee a priori a given tight false alarm rate or minimum detectable faults.

Fig. 3 and Fig. 4 show the fault estimation and fault detection performances resulting from simulating the fault diagnosers from Table I under the appearance of two step faults, one for each channel, of an amplitude of 0.7 at time $t = 100$ (disappearing at $t = 400$) for f_1, and at $t = 200$ (disappearing at $t = 500$) for f_2.

The fault diagnosers for case C2 and C4 are the fastest ones to detect the faults and their estimation of the faults have the lowest settling time. However they are the most sensitive under state disturbances and measurement noises (as they have the highest $\phi\overline{\lambda}(F)$ product, see (27)). For case C1, the fault detector cannot detect the faults on time because it has a too slow dynamic due to the conservativeness introduced by the Markov's inequality. Case C3, is an intermediate case between C1 and C2. Even if for case C3 the estimated faults converge slower to the faults than for cases C2 and C4, the detection mechanism only takes 6 more instants to detect the fault. This is due to the fact that C3 can detect lower faults than C2 and C4 (note that the diagonal of F_{C3}^{-1} are higher than the ones of F_{C2}^{-1} and F_{C4}^{-1}). Finally, note that the settling time at the 98% ($\xi = 2\%$) for the fault estimation, measured in terms of the number of measurement instants, is in the order of $\mathcal{K}$ (defined in (22)). For example, for case C3, the settling time is of 60 measurement instants for $\hat{f}_1$ and of 130 for $\hat{f}_2$, while it was characterized by $\mathcal{K} = 167$ from (22).

VII. Conclusion

In the current work, we designed a jump observer-based fault diagnoser to detect and estimate faults under measurements dropouts. We constructed the residual signal using a quadratic form of the estimated faults. A finite set of observer gains is used to estimate the faults and each gain is applied depending on the measurement outcomes. We employed the measurement successful reception probabilities from each sensor to describe the possible measurement reception scenarios.

The proposed design method allows finding a trade off between the achievable minimum detectable faults and the response time to faults, while guaranteeing a prescribed false alarm rate. Two design strategies can be used: fixing the minimum detectable faults and then minimizing the response time, or fixing the response time and then minimizing the minumum detectable faults.

We developed two ways of imposing a desired false alarm rate depending on the assumed knowledge about the probability distribution of the residual signal. If no information is assumed to be known, the Markov's inequality leads to a very conservative bound on the false alarm rate. If the disturbances and noise are assumed to be Gaussian, a certain condition imposed on matrix F leads to a chi-squared residual distribution. In this case a very precise bound on the false alarm rate is attained, improving the fault diagnosis performance.

Further research may include extensions to delayed measurements with Markovian models for the missing measurements and analytical characterization of the missing fault rate.

TABLE I
FAULT DIAGNOSERS COMPARISON.

Case	Design							Simulation
	$f_{\min}$	$f_{\min,1}$	$f_{\min,2}$	ϕ	ψ	ρ	$\mathcal{K}$	FAR
C1	0.6	0.6	0.6	10^{-3}	10^{-3}	0.999	6101	0
C2	0.56	0.46	0.52	0.145	10^{-3}	0.808	18	10^{-3}
C3	0.21	0.29	0.29	0.145	10^{-3}	0.977	167	10^{-3}
C4	0.6	0.6	0.6	0.1	0.1	0.798	17	10^{-4}

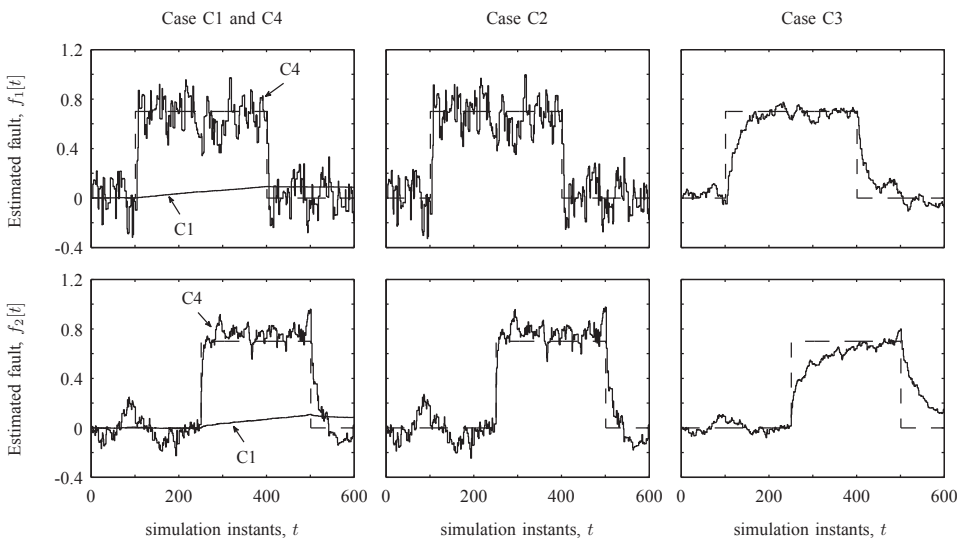

Fig. 3. Fault estimation performances for the analyzed cases on Table I.

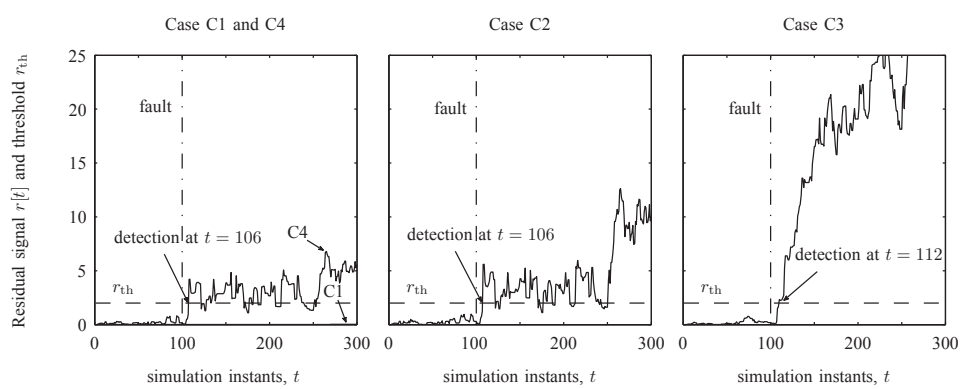

Fig. 4. Fault detection performances for the analyzed cases on Table I.

APPENDIX

Let us first introduce the following lemmas.

Lemma 1 ([34]). *Let ω be a stochastic vector with mean μ and a covariance matrix W, and P a symmetric matrix. Then*

$$\mathbf{E}\{\omega^T P \omega\} = \mu^T P \mu + \mathrm{tr}(PW).$$

Lemma 2 ([35]). *Let P be a positive semidefinite matrix, x_i a vector with appropriate dimensions and $\mu_i \geq 0$ scalar constants (with $i = 1, 2, \ldots$). If the series concerned is convergent, then we have*

$$\left(\sum_{i=1}^{\infty} \mu_i x_i\right)^T P \left(\sum_{i=1}^{\infty} \mu_i x_i\right) \leq \left(\sum_{i=1}^{\infty} \mu_i\right) \sum_{i=1}^{\infty} \mu_i x_i^T P x_i.$$

A. Proof of Theorem 1

Let us define the Lyapunov function at instant $t = t_k$ as $V_k = \tilde{z}_k^T P \tilde{z}_k$.

i) In the absence of disturbances, faults and measurement noises, after taking Schur's complements on (17a) and premultiplying the result by $\tilde{z}_k^T$ and postmultipliying by its transpose, we obtain that $V_{k+1} - V_k \leq 0$ that assures that the extended

state estimation error (13) converges to zero under standard sampling.

ii) Performing similar steps on (17b) (Schur's complements and operations with w_k^T, v_k^T and Δf_k), taking expected value on the results and adding the obtained constraints with the one from (17a) we get

$$\begin{aligned}&\mathbf{E}\{V_{k+1}\} - \mathbf{E}\{V_k\} + \mathbf{E}\{\tilde{f}_k^T F^{-1} \tilde{f}_k\} - \mathbf{E}\{w_k^T \Gamma_w w_k\} \\ &- \mathbf{E}\{v_{k+1}^T \Gamma_v v_{k+1}\} - \Delta f_k{}^T \Gamma_f \Delta f_k \leq 0 \end{aligned} \tag{34}$$

where we have considered the uncorrelation between $\tilde{z}_k$, w_k, v_{k+1} and Δf_k. Applying Lemma 1 over w_k^T and v_k^T, considering zero initial conditions ($V(0) = 0$) and adding the result from $k = 0$ to $K - 1$ we get

$$\sum_{k=0}^{K-1} \mathbf{E}\{\tilde{f}_k^T F^{-1} \tilde{f}_k\} \leq K \operatorname{tr}(\bar{\Gamma}) + \sum_{k=0}^{K-1} \Delta f_k^T \Gamma_f \Delta f_k^T \tag{35}$$

where we have taken into account that $P \succeq 0$ and that $\bar{\Gamma} = \Gamma_w W + \Gamma_v V$. Dividing the above expressions by K, taking the limit when $K \to \infty$ and considering that

$$\begin{aligned}&\mathbf{E}\{\tilde{f}_k^T F^{-1} \tilde{f}_k\} \geq \underline{\lambda}(F^{-1}) \mathbf{E}\{\tilde{f}_k^T \tilde{f}_k\}, \\ &\Delta f_k{}^T \Gamma_f \Delta f_k \leq n_f \overline{\lambda}(\Gamma_f) \|\Delta f\|_\infty^2 \leq n_f \overline{\lambda}(\Gamma_f) \Delta f_{\max}^2,\end{aligned}$$

and that $\underline{\lambda}(F^{-1}) = 1/\overline{\lambda}(F)$ (as F is a positive definite matrix), it leads to (18), which concludes this proof.

B. Proof of Theorem 2

If there is no fault on the system (i.e. $\tilde{f}_k = -\hat{f}_k$ and $\Delta f_k = 0$ for all k), we have that $\mathbf{E}\{\tilde{f}_k^T F^{-1} \tilde{f}_k\} = \mathbf{E}\{\hat{f}_k^T F^{-1} \hat{f}_k\} = \mathbf{E}\{r_k\}$. Then, following the proof of Theorem 1, dividing expression (35) by K, taking the limit when K tends to infinity and considering constraint (19), we obtain

$$\lim_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathbf{E}\{r_k\} \leq \phi\, r_{\text{th}}. \tag{36}$$

Considering the above result and the FAR definition given in (15), we can employ Markov's inequality[2] to obtain

$$\Psi \leq \lim_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K-1} \frac{\mathbf{E}\{r_k\}}{r_{\text{th}}} \leq \phi,$$

proving that ϕ bounds the FAR.

C. Proof of Theorem 3

Let us define vector $\tilde{f}_k'$ by $\tilde{f}_k' = F^{-\frac{1}{2}} \tilde{f}_k$. With that, (34) can be rewritten as

$$\mathbf{E}\{V_{k+1}\} - \mathbf{E}\{V_k\} \leq -\mathbf{E}\{\|\tilde{f}_k'\|_2^2\} + r_{\text{th}} + n_f \overline{\lambda}(\Gamma_f) \Delta f_{\max}^2. \tag{37}$$

Inequality (20) implies that Γ_f minus the diagonal block of P corresponding to the fault estimation error is positive semidefinite. Then, there exists a finite real constant $d_1 \geq 0$ that fulfills

$$\begin{aligned}\mathbf{E}\{V_k\} &\leq \mathbf{E}\{\tilde{f}_k^T \Gamma_f \tilde{f}_k\} + d_1 = \mathbf{E}\{\tilde{f}_k'^T F^{\frac{1}{2}^T} \Gamma_f F^{\frac{1}{2}} \tilde{f}_k'\} + d_1 \\ &\leq \overline{\lambda}(\Gamma_f F) \mathbf{E}\{\|\tilde{f}_k'\|_2^2\} + d_1 \end{aligned} \tag{38}$$

[2]If x is a positive random variable and $a > 0$, then $\mathbf{Pr}\{x > a\} \leq \frac{\mathbf{E}\{x\}}{a}$.

for all k, considering the fact that $\Gamma_f F$ and $F^{\frac{1}{2}^T} \Gamma_f F^{\frac{1}{2}}$ are similar matrices[3]. From this expression we can upper bound $-\mathbf{E}\{\|\tilde{f}_k'\|_2^2\}$ allowing us to rewrite expression (37) as

$$\mathbf{E}\{V_{k+1}\} \leq \rho \mathbf{E}\{V_k\} + \varepsilon + (1 - \rho) d_1, \tag{39}$$

for all k with ρ as defined in (21) and

$$\varepsilon = r_{\text{th}} + n_f \overline{\lambda}(\Gamma_f) \Delta f_{\max}^2.$$

Expressions (17a) imposes that $\bar{B}_f^T P \bar{B}_f \succeq F^{-1}$ which combined with (20) leads to $\Gamma_f F \succeq I$ guaranteing that $0 \leq \rho \leq 1$. Going backwards from k to $k = 0$, expression (39) becomes

$$\mathbf{E}\{V_{k+1}\} \leq \rho^{k+1} \mathbf{E}\{V_0\} + \sum_{l=0}^{k} \rho^l \left(\varepsilon + (1 - \rho) d_1\right).$$

Taking into account that $\sum_{l=0}^{k} \rho^l = \frac{1 - \rho^{k+1}}{1 - \rho} \leq \frac{1}{1 - \rho}$, then

$$\mathbf{E}\{V_{k+1}\} \leq \rho^{k+1} \mathbf{E}\{V_0\} + \frac{1}{1 - \rho} \varepsilon + d_1. \tag{40}$$

Constraint (17a) implies also that $\mathbf{E}\{V_k\} \geq \mathbf{E}\{\|\tilde{f}_k'\|_2^2\}$. Considering this, inequality (38) and the fact that

$$\underline{\lambda}(F^{-1}) \|\tilde{f}_k\|_2^2 \leq \|\tilde{f}_k'\|_2^2 \leq \overline{\lambda}(F^{-1}) \|\tilde{f}_k\|_2^2,$$

expression (40) leads to

$$\mathbf{E}\{\|\tilde{f}_{k+1}\|_2^2\} \leq \rho^{k+1} \frac{\kappa(F)}{(1 - \rho)} \mathbf{E}\{\|\tilde{f}_0\|_2^2\} + \frac{\overline{\lambda}(F)}{(1 - \rho)} \varepsilon + \overline{\lambda}(F) d_1, \tag{41}$$

where $\kappa(F) = \overline{\lambda}(F)/\underline{\lambda}(F)$ is the condition number of matrix F and where we have considered that $\underline{\lambda}(F^{-1}) = 1/\overline{\lambda}(F)$ because F is positive definite. Expression (41) proves that $\mathbf{E}\{\|\tilde{f}_k\|_2^2\}$ decays with ρ.

D. Proof of Theorem 4

First, in the absence of faults and under zero initial conditions, $\tilde{f}_k$ is normally distributed and has zero mean because the disturbances and measurement noises are normally distributed with zero mean. Second, let $Z_{k-1} = \mathbf{E}\{\tilde{z}_{k-1} \tilde{z}_{k-1}^T\}$ be the covariance matrix for the state estimation error updated at instants t_{k-1} (which is also the covariance at instants t, since we are dealing with standard sampling). Then, its expected value at instant t_k is given by

$$\mathbf{E}\{Z_k\} = G(\bar{A} Z_{k-1} \bar{A}^T + \bar{B}_w W \bar{B}_w^T) G^T + L\, V\, L^T. \tag{42}$$

As the observer gain L assures the stability of (13) (by Theorem 1), the series in (42) converges to a symmetric positive definite matrix $\Sigma_f = \mathbf{E}\{Z_k\} = Z_{k-1}$ when $k \to \infty$ (see [26]) given in (24). Then we have that $\tilde{f}_k^T \Sigma_f^{-1} \tilde{f}_k$ is distributed as $\mathcal{X}_{n_f}^2$ (see [36]). Considering (23), the signal $r_k/\phi = \tilde{f}_k^T F^{-1} \tilde{f}_k$ is then distributed as $\mathcal{X}_{n_f}^2$. From Theorem 2 we know that $\mathbf{E}\{r_k\}/\phi \leq r_{\text{th}}$, see (36). As the expected value of random variable that follows a $\mathcal{X}_{n_f}^2$ is n_f, if we fix the threshold to be $r_{\text{th}} = n_f$, then we have that the FAR is

$$\Psi = \mathbf{Pr}\left\{ \frac{r_k}{\phi} > \frac{r_{\text{th}}}{\phi} \middle| f_k = 0 \right\}, \tag{43}$$

and using the definition of the CDF, we obtain (25).

[3]Matrices A and B are similar if $B = C^{-1} A C$. Similar matrices share the same eigenvalues.

E. Proof of Theorem 5

Let us define the Lyapunov function at instant $t = t_k$ as $V_k = \tilde{z}_k^T P \tilde{z}_k$. Let us first study the evolution of the Lyapunov function. The expected value of the Lyapunov function at the next update instant $t = t_{k+1}$ given that a measurement was obtained at t_k, denoted by $\mathbf{E}\{V_{k+1}\}$, is

$$\begin{aligned}
&\sum_{N=1}^{\infty} p_0^{N-1} \sum_{i=1}^{q} p_i \, \mathbf{E}\{\tilde{z}_{k+1}^T P \tilde{z}_{k+1} | N_{k+1} = N, \alpha_{k+1} = \eta_i\} \\
&\quad = \mathbf{E}\left\{ \tilde{z}_k^T \left(\sum_{N=1}^{\infty} p_0^{N-1} (\bar{A}^N)^T \mathcal{Q} \bar{A}^N \right) \tilde{z}_k \right\} \\
&\quad + \mathbf{E}\left\{ w_k^T \left(\sum_{N=1}^{\infty} p_0^{N-1} \left(\sum_{l=0}^{N-1} \bar{B}_w^T \left(\bar{A}^l\right)^T \mathcal{Q} \bar{A}^l \bar{B}_w \right) \right) w_k \right\} \\
&\quad + \mathbf{E}\left\{ v_k^T \left(\sum_{N=1}^{\infty} p_0^{N-1} \sum_{i=1}^{r} \eta_i^T L_i^T P L_i \eta_i \right) v_k \right\} \\
&\quad + \sum_{N=1}^{\infty} p_0^{N-1} (\star)^T \mathcal{Q} \underbrace{\left(\sum_{l=0}^{N-1} \bar{A}^l \bar{B}_f \Delta f[t_k + l] \right)}_{\star}
\end{aligned} \tag{44}$$

considering the uncorrelation between $\tilde{z}[t_k]$, $w[t_k + l - 1]$, $v[t_{k+1}]$ and $\Delta f[t_k + l - 1]$ for $l = 1, \ldots, N_k - 1$ and the uncorrelation in time of $w[t]$. Matrix $\mathcal{Q}$ is defined by $\mathcal{Q} = \sum_{i=1}^{q} p_i G_i^T P G_i$, where $G_i = I - L_i \eta_i \bar{C}$ and $L_i = P^{-1} X_i$.

Le us denote by $\mathcal{V}_{k+1}$ the result of replacing in (44) $\mathcal{Q}$ by Q where $\mathcal{Q} \preceq Q$. We rewrite $\mathcal{V}_{k+1}$ as

$$\mathcal{V}_{k+1} = \mathcal{V}_{k+1}^z + \mathcal{V}_{k+1}^w + \mathcal{V}_{k+1}^v + \mathcal{V}_{k+1}^f.$$

Since $\mathcal{Q} \preceq Q$, we have that $\mathbf{E}\{V_{k+1}\} \leq \mathcal{V}_{k+1}$. If $p_0 \overline{\lambda}(\bar{A})^2 < 1$, the series involve in (44), and therefore in $\mathcal{V}_{k+1}$, are convergent. Then, the summatory in $\mathcal{V}_{k+1}^f$, which implies dealing with cross products between the different $\Delta f[t_k + l]$, can be bounded with Lemma 2 as $\mathcal{V}_{k+1}^f \leq \overline{\mathcal{V}}_{k+1}^f$ with $\overline{\mathcal{V}}_{k+1}^f$ given by (28). Therefore, defining $\overline{\mathcal{V}}_{k+1}$ as

$$\overline{\mathcal{V}}_{k+1} = \mathcal{V}_{k+1}^z + \mathcal{V}_{k+1}^w + \mathcal{V}_{k+1}^v + \overline{\mathcal{V}}_{k+1}^f, \tag{45}$$

we have that $\mathbf{E}\{V_{k+1}\} \leq \mathcal{V}_{k+1} \leq \overline{\mathcal{V}}_{k+1}$. Let us now analyze constraints (26a)-(26e). If (26e) holds, then matrix Q is such as $Q \succeq \mathcal{Q}$. Matrices M_1, M_2, M_5, M_6 can be rewritten as

$$\begin{aligned}
M_1 &= \sum_{N=1}^{\infty} p_0^{N-1} (\bar{A}^N)^T Q \bar{A}^N, \\
M_2 &= \sum_{N=1}^{\infty} p_0^{N-1} \left(\sum_{l=0}^{N-1} \bar{B}_w^T \left(\bar{A}^l\right)^T Q \bar{A}^l \bar{B}_w \right), \\
M_5 + M_6 &= \sum_{N=1}^{\infty} N p_0^{N-1} \sum_{j=0}^{N-1} (\bar{A}^j)^T Q \bar{A}^j.
\end{aligned}$$

Then taking Schur's complement from (26a) to (26d); premultiplying the result by $\tilde{z}_k^T$, w_k^T, v_k^T and $\overline{\Delta f}_k^{\,T}$ and postmultiplying by its transpose respectively; and taking expected values in both sides, we obtain

$$\begin{aligned}
&\mathbf{E}\left\{\tilde{z}_k^T M_1 \tilde{z}_k\right\} \leq \mathbf{E}\left\{V_k\right\} - \mathbf{E}\left\{\tilde{f}_k^T F^{-1} \tilde{f}_k\right\}, \\
&\mathbf{E}\left\{w_k^T M_2 w_k\right\} \leq \mathbf{E}\left\{w_k^T \Gamma_w w_k\right\}, \\
&\mathbf{E}\left\{ v_k^T \left(\sum_{N=1}^{\infty} p_0^{N-1} \sum_{i=1}^{r} \eta_i^T L_i^T P L_i \eta_i \right) v_k \right\} \leq \mathbf{E}\left\{v_k^T \Gamma_v v_k\right\}, \\
&\overline{\Delta f}_k^{\,T} \left(\bar{B}_f^T (M_5 + M_6) \bar{B}_f\right) \overline{\Delta f}_k \leq \overline{\Delta f}_k^{\,T} \Gamma_f \overline{\Delta f}_k,
\end{aligned}$$

Adding all the above expressions leads to

$$\begin{aligned}
\Upsilon = &\overline{\mathcal{V}}_{k+1} - \mathbf{E}\{V_k\} + \mathbf{E}\{\tilde{f}_k^T F^{-1} \tilde{f}_k\} \\
&- \mathbf{E}\{w_k^T \Gamma_w w_k\} - \mathbf{E}\{v_k^T \Gamma_v v_k\} - \overline{\Delta f}_k^{\,T} \Gamma_f \overline{\Delta f}_k \leq 0
\end{aligned} \tag{46}$$

where $\overline{\mathcal{V}}_{k+1}$ is as defined in (45). Let us define Θ as

$$\Theta = \Upsilon - \overline{\mathcal{V}}_{k+1} + \mathbf{E}\{V_{k+1}\} \leq 0.$$

Therefore, as $\mathbf{E}\{V_{k+1}\} \leq \overline{\mathcal{V}}_{k+1}$, if (46) holds, then we have that $\Theta \leq 0$ (analogous to (34)), since $\Theta \leq \Upsilon \leq 0$.

Using the fact that $\Theta \leq 0$ and following similar steps than in the proofs of Theorems 1, 2 and 3 we can prove with not much effort that the statements of Theorem 5 hold.

F. Proof of Theorem 6

Let $Z_{k-1} = \mathbf{E}\{\tilde{z}_{k-1} \tilde{z}_{k-1}^T\}$ be the covariance matrix for the state estimation error updated at the measurement instant t_{k-1}. Then, its expected value at instant t_k is given by

$$\begin{aligned}
\mathbf{E}\{Z_k\} = &\sum_{i=1}^{q} p_i G_i (\bar{A} R_{k-1} \bar{A}^T + S_W) G_i^T \\
&+ \frac{1}{1 - p_0} \sum_{i=1}^{q} p_i L_i \, \eta_i V \eta_i^T L_i^T
\end{aligned} \tag{47}$$

where $\sum_{N=1}^{\infty} p_0^{N-1} = 1/(1 - p_0)$ and $R_{k-1} = \sum_{i=0}^{\infty} p_0^i \bar{A}^i Z_{k-1} (\bar{A}^i)^T$ expressed as

$$\mathrm{vec}(R_{k-1}) = (I - p_0 \bar{A} \otimes \bar{A})^{-1} \mathrm{vec}(Z_{k-1}).$$

Following similar arguments than in the proof of Theorem 4, the series in (47) converges to a symmetric positive definite matrix Σ_f and $r_k/\phi = \tilde{f}_k^T \Sigma_f^{-1} \tilde{f}_k$ is distributed as $\mathcal{X}_{n_f}^2$, leading to a FAR given by (25).

Acknowledgement

The authors would like to thank the associate editor and the anonymous reviewers for their valuable comments and suggestions which have helped improving the presentation of the paper.

References

[1] R. A. Gupta and M.-Y. Chow, "Networked control system: overview and research trends," *IEEE Trans. Ind. Electron.*, vol. 57, no. 7, pp. 2527–2535, 2010.

[2] J. P. Hespanha and P. Naghshtabrizi, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, 2007.

[3] Z. Wang, F. Yang, D. W. C. Ho, and X. Liu, "Robust H_∞ filtering for stochastic time-delay systems with missing measurements," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2579–2587, 2006.

[4] J. Chen and R. J. Patton, *Robust model-based fault diagnosis for dynamic systems*. Springer Publishing Company, Incorporated, 1999, vol. 3.

[5] H. Fang, H. Ye, and M. Zhong, "Fault diagnosis of networked control systems," *Annual Reviews in Control*, vol. 31, no. 1, pp. 55–68, 2007.

[6] Z. Mao, B. Jiang, and P. Shi, "H_∞ fault detection filter design for networked control systems modelled by discrete Markovian jump systems," *IET Control Theory Appl.*, vol. 1, no. 5, p. 1336, 2007.

[7] Y. Wang, H. Ye, S. X. Ding, G. Wang, and D. Zhou, "Residual generation and evaluation of networked control systems subject to random packet dropout," *Automatica*, vol. 45, no. 10, pp. 2427–2434, 2009.

[8] W. Li, Z. Zhu, and S. Ding, "Fault detection design of networked control systems," *IET Control Theory Appl.*, vol. 5, no. 12, p. 1439, 2011.

[9] Y. Wang, S. X. Ding, H. Ye, and G. Wang, "A new fault detection scheme for networked control systems subject to uncertain time-varying delay," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5258–5268, 2008.

[10] C. Peng, D. Yue, E. Tian, and Z. Gu, "Observer-based fault detection for networked control systems with network Quality of Services," *Appl. Math. Model.*, vol. 34, no. 6, pp. 1653–1661, 2010.

[11] X. Wan, H. Fang, and S. Fu, "Observer-based fault detection for networked discrete-time infinite-distributed delay systems with packet dropouts," *Appl. Math. Model.*, vol. 36, no. 1, pp. 270–278, 2012.

[12] W. Li and S. X. Ding, "Remote fault detection system design with online channel reliability information," *Int. J. Syst. Sci.*, vol. 41, no. 8, pp. 957–970, 2010.

[13] B. A. Ogunnaike, *Random phenomena: fundamentals of probability and statistics for engineers*. Taylor & Francis US, 2011.

[14] F. N. Chowdhury, "Ordinary and neural Chi-squared tests for fault detection in multi-output stochastic systems," *IEEE Trans. Control Syst. Technol.*, vol. 8, no. 2, pp. 372–379, 2000.

[15] D. Huang and S. K. Nguang, "Robust fault estimator design for uncertain networked control systems with random time delays: An ILMI approach," *Information Sciences*, vol. 180, no. 3, pp. 465–480, 2010.

[16] J. Li and G.-Y. Tang, "Fault diagnosis for networked control systems with delayed measurements and inputs," *IET Control Theory Appl.*, vol. 4, no. 6, p. 1047, 2010.

[17] X. Yao, L. Wu, and W. X. Zheng, "Fault detection filter design for markovian jump singular systems with intermittent measurements," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3099–3109, 2011.

[18] J. Yu, M. Liu, W. Yang, P. Shi, and S. Tong, "Robust fault detection for Markovian jump systems with unreliable communication links," *Int. J. Syst. Sci.*, vol. 44, no. 11, pp. 2015–2026, 2013.

[19] B. Shen, S. X. Ding, and Z. Wang, "Finite-horizon H_∞ fault estimation for linear discrete time-varying systems with delayed measurements," *Automatica*, vol. 49, no. 1, pp. 293–296, 2013.

[20] K. Zhang, B. Jiang, and P. Shi, "Fault estimation observer design for discrete-time takagisugeno fuzzy systems based on piecewise lyapunov functions," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 1, pp. 192–200, 2012.

[21] I. Hwang, S. Kim, Y. Kim, and C. E. Seah, "A survey of fault detection, isolation, and reconfiguration methods," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 3, pp. 636–653, 2010.

[22] P. Zhang and S. X. Ding, "An integrated trade-off design of observer based fault detection systems," *Automatica*, vol. 44, no. 7, pp. 1886–1894, 2008.

[23] Y. Wang, H. Ye, S. X. Ding, Y. Cheng, P. Zhang, and G. Wang, "Fault detection of networked control systems with limited communication," *Int. J. Control*, vol. 82, no. 7, pp. 1344–1356, 2009.

[24] X. He, Z. Wang, Y. D. Ji, and D. H. Zhou, "Robust Fault Detection for Networked Systems with Distributed Sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 1, pp. 166 – 177, 2011.

[25] S. C. Smith and P. Seiler, "Estimation with lossy measurements: jump estimators for jump systems," *IEEE Trans. Autom. Control*, vol. 48, no. 12, pp. 2163–2171, 2003.

[26] D. Dolz, I. Peñarrocha, and R. Sanchis, "Jump state estimation with multiple sensors with packet dropping and delaying channels," *Int. J. Syst. Sci.*, 2014.

[27] J.-L. Chang, "Applying discrete-time proportional integral bbservers for state and disturbance estimations," *IEEE Trans. Autom. Control*, vol. 51, no. 5, pp. 814–818, 2006.

[28] Z. Li, E. Mazars, Z. Zhang, and I. M. Jaimoukha, "State-space solution to the $\mathcal{H}_-/\mathcal{H}_\infty$ fault-detection problem," *Int. J. Robust Nonlinear Control*, vol. 22, no. 3, pp. 282–299, 2012.

[29] Z. Wang, B. Shen, and X. Liu, "$\mathcal{H}_\infty$ filtering with randomly occurring sensor saturations and missing measurements," *Automatica*, vol. 48, no. 3, pp. 556–562, 2012.

[30] J. Löfberg, "Yalmip: A toolbox for modeling and optimization in matlab," in *IEEE International Symposium on Computer Aided Control Systems Design*, 2004, pp. 284–289.

[31] L. El Ghaoui, F. Oustry, and M. AitRami, "A cone complementarity linearization algorithm for static output-feedback and related problems," *IEEE Trans. Autom. Control*, vol. 42, no. 8, pp. 1171–1176, 1997.

[32] H. Gao, T. Chen, and L. Wang, "Robust fault detection with missing measurements," *Int. J. Control*, vol. 81, no. 5, pp. 804–819, 2008.

[33] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization methods and software*, vol. 11, no. 1-4, pp. 625–653, 1999.

[34] T. Söderström, "Discrete-time stochastic systems. estimation and control," 1994.

[35] Y. Liu, Z. Wang, J. Liang, and X. Liu, "Synchronization and state estimation for discrete-time complex networks with distributed delays," *IEEE Trans. Syst., Man, Cybern., Part B: Cybern.*, vol. 38, no. 5, pp. 1314–1325, 2008.

[36] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis (6th Edition)*. Pearson, 2007.

Daniel Dolz was born in Castellón, Spain in 1988. He received his M. Sc. Degree in Industrial Engineering in 2011 and his Ph.D. in Industrial Technologies in 2014 from the Universitat Jaume I of Castellón, Spain. He also holds from 2011 a M. Sc. Degree in Automatic and Electronic Engineering from INSA Toulose, France. Currently he is a postdoctoral fellow at the Department of Industrial Systems Engineering and Design at the Universitat Jaume I of Castellón, Spain. His research interests include estimation, fault diagnosis and control over networks, and wireless sensor networks.

Ignacio Peñarrocha was born in Castellón, Spain in 1978. He received the M. Sc. Degree in Industrial Engineering from the Universitat Jaume I de Castelló (UJI), Spain, in 2002, and his Ph.D. in Computering and Control Engineering from Universitat Politècnica de València (UPV), Spain, in 2006. He has been working since 2004 at the Universitat Jaume I de Castellón. His current position is as associate professor at the department of Industrial Systems Engineering and Design. He has participated in several local and national research projects. His research interests include identification, estimation, fault diagnosis and control over networks, and fault tolerant control of wind turbines and wind farms.

Roberto Sanchis was born in Genovés, València, Spain in 1968. He received the M.Sc Degree in Electrical Engineering in 1993 and his Ph.D. in Control Engineering in 1999 from the Polytechnic University of València (UPV), Spain. He was awarded the first national prize for university graduation in 1993. During 1994 and 1995 he was a teaching assistant at the Systems and Control Engineering Department of the UPV. He has been working since 1996 at the University Jaume I of Castellón, Spain. His current position is as Associate Professor (tenure lecturer) at the department of Industrial Systems Engineering and Design, where he is the leader of the Control Systems and Automation group. His research interests include missing-data estimation, identification and control, networked control systems, tuning and auto-tuning of PID controllers, and applications (ceramic industry and waste water treatment).