Distributed Kalman Filtering over Sensor Networks with Unknown Random Link Failures

Stefano Battilotti, *Senior Member, IEEE,* Filippo Cacace, *Member, IEEE,* Massimiliano d'Angelo,
and Alfredo Germani, *Senior Member, IEEE*

Abstract—In this paper we consider the distributed consensus-based filtering problem for linear time-invariant systems over sensor networks subject to random link failures when the failure sequence is not known at the receiving side. We assume that the information exchanged, traveling along the channel, is corrupted by a noise and hence, it is no more possible to discriminate with certainty if a link failure has occurred. Therefore, in order to process the only significant information, we endow each sensor with detectors which decide on the presence of link failures. At each sensor the proposed approach consists of three steps: failure detection, local data aggregation and Kalman consensus filtering. Numerical examples show the effectiveness of this method.

Index Terms—Sensor Networks, Kalman filtering, Fault Detection.

I. INTRODUCTION

IN recent years the fast technological developments in the area of sensor networks have attracted a considerable amount of research on the problem of distributed estimation. In the context of the ongoing Internet of Things era, distributed estimation and control, by successfully exploiting the communication among the nodes of the underlying network, are currently being adopted to deal with several applications, *e.g.* in the domains of telecommunications [4] and intelligent transportation systems [24]. In particular, the typical scenario of distributed estimation consists of a group of autonomous sensors that are deployed in a monitored region and cooperate in some monitoring task by sharing their local information via wireless communication links. In this framework the estimation process is distributed in the sense that no central elaboration is assumed and each node of the sensor network utilizes both the local information and the messages from the neighbors to generate an estimate [5], [12], [13], [15], [16], [17], [21]. Even if the use of shared information across neighbors improves the local estimate, there is no guarantee for the consensus of the estimates across the network. To this aim, several consensus procedures have been proposed [1], [2], [3], [16]. In practical applications temporary link failures are an important issue, due to power constraints, multipath fading, background noise or external attacks. This problem has been widely investigated for centralized estimation algorithms [22], [6] and in the consensus problem of random networks [8], [10], [11], [23]. In the context of sensor networks and distributed consensus algorithms the topology design of the network in presence of link failures and other communication constraints is investigated in [9]. The convergence rate of the consensus in presence of failure is studied in [19], [20]. The recent paper [14] proposes a two-stage Kalman-consensus filtering approach over unreliable channels when the link failures are known at the receiving side. In these conditions sufficient conditions for the boundedness of the estimation error covariance are provided.

In this paper we consider the distributed consensus-based filtering problem for linear time-invariant systems over sensor networks subject to random link failures when the failure sequence is not known at the receiving side. Our approach extends the method proposed in [17]-[18] to the case of unreliable channels by introducing a failure detection strategy proposed for the case of intermittent observations in [6]-[7]. At each sensor the proposed approach consists of three main steps: failure detection provided by some local nonlinear optimal detectors, local aggregation of data and covariance matrices and Kalman-consensus filtering fed by the decisions made by the detectors.

The paper is organized as follows. The problem setting is formalized in Section II. Section III describes the detection procedure, which is part of the distributed filtering algorithm described in Section IV. The performance of the algorithm is investigated in Section V and conclusions follow.

Notation If $A \in \mathbb{R}^{n \times n}$ then $A^\top$ denotes its transpose and $|A|$ denotes its determinant. If $v_1, \ldots, v_n$ are vectors in $\mathbb{R}^n$, then $v = \mathrm{col}(v_1, \ldots, v_n)$ denotes the vector $v = [v_1^\top, \ldots, v_n^\top]^\top$. Moreover, if $v \in \mathbb{R}^n$, then we denote with $\mathrm{diag}(v) \in \mathbb{R}^{n \times n}$ the diagonal matrix with entries the components of v. If A and B are two matrices in $\mathbb{R}^{n \times n}$, then the direct sum of matrices is denoted by $A \oplus B$ and the Kroneker product with $A \otimes B$. We indicate with I the identity matrix of appropriate dimension. The euclidean norm in $\mathbb{R}^{n \times n}$ is denoted with $\| \cdot \|$. Finally, given a random variable (r.v.) X in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, implicit in the rest of the paper, we denote with $\mathbb{E}[X]$ its expectation. We denote with $X \sim \mathcal{N}(\mu, \sigma^2)$ a gaussian r.v. X with mean μ and variance σ^2.

Stefano Battilotti and Massimiliano d'Angelo {battilotti,mdangelo}@diag.uniroma1.it are with DIAG, Sapienza Università di Roma, Italy.

Filippo Cacace f.cacace@unicampus.it is with Università Campus Bio-medico di Roma, Italy.

Alfredo Germani germani@ing.univaq.it is with Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica, Università dell'Aquila, Italy.

II. PROBLEM FORMULATION

A. Topology Structure

The topology structure is characterized by an undirected graph with N nodes $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with $\mathcal{V} = \{1, \ldots, N\}$ the set of the nodes of the graph, namely the sensors, and the edges $(i, j) \in \mathcal{E}$ correspond to the links among such nodes. We denote by $\mathcal{E}_i = \{k : (i, k) \in \mathcal{E}\}$ the set of neighbors of node i. The number of measurements available at the sensor i from the neighbors is $d_i = |\mathcal{E}_i|$ and we define $J_i = \mathcal{E}_i \cup \{i\}$.

B. Target Plant and measurement models

We consider the class of linear time-invariant systems described by the following state-space model

$$x_{k+1} = Ax_k + f_k, \quad k \geq 0, \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state of the process and $f_k \in \mathbb{R}^n$ is a stochastic sequence affecting the process at time k. For $i \in \mathcal{V}$ and $j \in \mathcal{E}_i$, the sensing model of the i-th node is described by

$$y_k^i = C_i x_k + g_k^i, \tag{2}$$
$$y_k^{ij} = \gamma_k^{ij} y_k^j + v_k^{ij}, \tag{3}$$

where, at time k, the vector $y_k^i \in \mathbb{R}^m$ is the measurement of node i, while the measurement $y_k^{ij} \in \mathbb{R}^m$ is the information available at node i delivered from node j. The sequences $g_k^i \in \mathbb{R}^m$ and $v_k^{ij} \in \mathbb{R}^m$ are stochastic noise terms affecting the measurements whilst $\gamma_k^{ij} \in \{0, 1\}$ is a random variable modeling the link failure between nodes i and j at time k.

We note that if $\gamma_k^{ij} = 1$, then a link failure has not occurred and node i receives from node j the measurement y_k^j corrupted by an additive noise v_k^{ij} because of the transmission channel. If $\gamma_k^{ij} = 0$, then the information delivered from node j to node i consists of the pure noise signal v_k^{ij}. The initial state x_0 and the stochastic sequences $\{f_k\}$, $\{g_k^i\}$ and $\{\gamma_k^{ij}\}$ satisfy the following conditions for $k \geq 0$:

1) x_0 is a zero-mean Gaussian random variable with covariance matrix $\mathbb{E}[x_0 x_0^\top] = \Sigma_0$,
2) $\{f_k\}$ is a white sequence of zero-mean Gaussian random vectors with covariance matrix $\mathbb{E}\{f_k f_k^\top\} = Q$,
3) $\{g_k^i\}$ is a white sequence of zero-mean Gaussian random vectors with covariance matrix $\mathbb{E}\{g_k^i g_k^{i\top}\} = R_i > 0$ with $i \in \mathcal{V}$,
4) $\{v_k^{ij}\}$ is a white sequence of zero-mean Gaussian random vectors with covariance matrix $\mathbb{E}\{v_k^{ij} v_k^{ij\top}\} = V_{ij} > 0$ with $i \in \mathcal{V}$ and $j \in \mathcal{E}_i$,
5) For a fixed integer $L \geq 0$ the following joint probability mass functions are known[1]

$$\begin{cases} \overline{\mathrm{P}}_k(\gamma_k^{ij}, \ldots, \gamma_1^{ij}), & \text{for } 1 \leq k \leq L+1 \\ \overline{\mathrm{P}}_k(\gamma_k^{ij}, \ldots, \gamma_{k-L}^{ij}), & \text{for } k > L+1 \end{cases}$$

6) For any $i \in \mathcal{V}$, the random quantities x_0, $\{f_k\}$, $\{g_k^i\}$, $\{v_k^{ij}\}$ and $\{\gamma_k^{ij}\}$ are statistically independent $\forall j \in \mathcal{E}_i$.

We can assume without loss of generality that $V_{ij} = V_{ji}$ and $\gamma_k^{ij} = \gamma_k^{ji}$, i.e. a link failure occurring in the channel from

---

[1]With a slight abuse of notation, we omit the superscript ij for the probability mass functions.

---

node i to node j affects the transmission from node j to node i too. This is a reasonable but not restrictive assumption for the methodology we develop in this paper. Moreover, we note that the joint probability mass function of assumption (5) modeling the link failures could be different for different links.

In the recent paper [14] as well as in other contributions, it is assumed that the link failures are known to the receiver. Consequently, the knowledge of the sequence $\{\gamma_k^{ij}\}$ is needed to implement the filter.

In this paper we assume that the sequence of failures is not available at the receiving nodes. In the case of a link failure the observation y_k^{ij} of (3) consists of a pure noise signal v_k^{ij}. Therefore, in order to only process significant information, we shall endow each sensor with detectors which decide on the presence of link failures.

C. The ideal distributed filter with known failure sequences

In this paragraph we recall the case of known link failures, i.e. known $\{\gamma_k^{ij}\}$ sequences. It is clear that the information y_k^{ij} would be no longer beneficial for estimation purpose if a failure occurs (namely $\gamma_k^{ij} = 0$). Hence, if the sequence $\{\gamma_k^{ij}\}$ were available, the only reasonable choice when $\gamma_k^{ij} = 0$ is to reject the measurement y_k^{ij} since it does not convey any information on the state process. Consequently, if the sequences $\{\gamma_k^{ij}\}$ are assumed to be known for all $i \in \mathcal{V}$ and $j \in \mathcal{E}_i$, we can frame the problem as a distributed filtering problem on a switching (or dynamic) topology with $\mathcal{G}(k) = (\mathcal{V}, \mathcal{E}(k))$ where the edge $(i, j) \in \mathcal{E}(k)$ and the node $j \in \mathcal{E}_i(k)$ if and only if $\gamma_k^{ij} = 1$. In other words, there is a link between nodes i and j when $\gamma_k^{ij} = 1$, whilst the link is absent if $\gamma_k^{ij} = 0$. In that case, recast with our notation, the author of [17] and [18] has proposed the following consensus-based Kalman filtering algorithm for any sensor $i \in \mathcal{V}$

Initialization:

$$\bar{x}_0^i = 0, \qquad P_0^i = \Sigma_0. \tag{4}$$

Locally aggregate data and covariance matrices $\forall j \in J_i(k)$:

$$u_k^{ij} = C_j^\top W_{ij}^{-1} y_k^{ij}, \quad U_{ij} = C_j^\top W_{ij}^{-1} C_j, \tag{5}$$

where $y_k^{ii} \doteq y_k^i$, $W_{ij} \doteq \gamma_k^{ij} R_j + V_{ij}$, $\gamma_k^{ii} \doteq 1$ and $V_{ii} \doteq 0$,

$$z_k^i = \sum_{j \in J_i(k)} u_k^{ij}, \qquad Z_k^i = \sum_{j \in J_i(k)} U_{ij}. \tag{6}$$

Compute the Kalman-consensus estimate:

$$M_k^i = \left(P_k^{i\,-1} + Z_k^i\right)^{-1}, \tag{7}$$

$$\hat{x}_k^i = \bar{x}_k^i + M_k^i \left(z_k^i - Z_k^i \bar{x}_k^i\right) + \epsilon M_k^i \sum_{j \in \mathcal{E}_i(k)} \left(\bar{x}_k^j - \bar{x}_k^i\right). \tag{8}$$

Update the state of the Kalman-consensus filter:

$$P_{k+1}^i = A M_k^i A^\top + Q, \tag{9}$$
$$\bar{x}_{k+1}^i = A \hat{x}_k^i. \tag{10}$$

In (8) the scalar $\epsilon > 0$ is a relative small constant (e.g. chosen to be of the order of the integration time-step of the discretization of a continuous-time process [18]). We refer

to this algorithm as the *ideal* distributed filter (DF) of [18] with known failure sequences, since it makes use of the sequences $\{\gamma_k^{ij}\}$ which are not available in reality. In fact, the time-varying nature of the set $\mathcal{E}_i(k)$, and consequently $J_i(k)$, descends from the fact that the sequence $\{\gamma_k^{ij}\}$ is known. Moreover, since each sensor can reconstruct by itself the values W_{ij}, and thus U_{ij}, we see that the message broadcasted at time k to all neighbors by node i is $(y_k^i, \bar{x}_k^i)$.

In this paper, a consensus-based Kalman filtering algorithm is derived from the one reported above when the information delivered by the nodes is subjected to link failures. For simplicity, in what follows, we assume that for each node i and at any time $k \geq 0$, the information $\bar{x}_k^i$ is available to the neighbors $j \in \mathcal{E}_i$ with certainty (e.g. travel along a reliable channel). We stress that in order to detect the link failures, this assumption is not necessary. In fact, even though the information $\bar{x}_k^i$ travel along a faulty and noisy channel, the proposed methodology can be applied following exactly the same procedure.

III. OPTIMAL DETECTION OF LINK FAILURES

In this section we consider the problem of detecting the link failures and we adopt the decision strategy used in [6] for the problem of intermittent observations. It is clear that each sensor should be endowed with a number of detectors equal to the number of links. We note that, even though by assumption $\gamma_k^{ij} = \gamma_k^{ji}$ for all $k \geq 0$, the decision $\hat{\gamma}_k^{ij}$ given by the detector (i, j), i.e. the one deciding for the sequence $\{\gamma_k^{ij}\}$, is different of the decision $\hat{\gamma}_k^{ji}$ given by the detector (j, i). Clearly, a low probability of error of the detection stage is beneficial for the subsequent state estimation stage, thus improving the overall performance. Hence, for any $i \in \mathcal{V}$, the sensor i has d_i detectors, one for each link $j \in \mathcal{E}_i$. In order to detect the possible link failures with its neighbors, sensor i chooses as optimal detector for the link with j, the one that guarantees the minimum probability of error given the available observations delivered by node j, namely $y_1^{ij}, \ldots, y_k^{ij}$. This is equivalent to the maximum a posteriori (MAP) probability decision rule and thus for each $i \in \mathcal{V}$ and $j \in \mathcal{E}_i$ we have

$$\hat{\gamma}_k^{ij} = \arg \max_{\gamma_k^{ij} \in \{0,1\}} \mathrm{P}_k \left(\gamma_k^{ij} \mid y_k^{ij}, \ldots, y_1^{ij} \right) \tag{11}$$

where $\mathrm{P}_k \left(\gamma_k^{ij} \mid y_k^{ij}, \ldots, y_1^{ij} \right)$ is the probability mass function of γ_k^{ij} at time k conditional on the available observations delivered by node j.

A direct consequence is that the memory and the complexity of the MAP detector (11) increase with time. For, in order to obtain detectors with finite memory, it is sufficient to carry out the decision on the last $L+1$ measurements, with $L \geq 0$, namely

$$\hat{\gamma}_k^{ij} = \arg \max_{\gamma_k^{ij} \in \{0,1\}} \mathrm{P}_k(\gamma_k^{ij} | y_k^{ij}, \ldots, y_{k-L}^{ij}), \tag{12}$$

which is the optimal decision rule used in what follows.

To derive an explicit expression of (12) we first note that by equation (2), we can rewrite equation (3) as

$$\begin{aligned} y_k^{ij} &= \gamma_k^{ij} C_j x_k + \gamma_k^{ij} g_k^j + v_k^{ij} \\ &= \gamma_k^{ij} C_j x_k + w_k^{ij}(\gamma_k^{ij}) \end{aligned} \tag{13}$$

where $w_k^{ij}(\gamma_k^{ij}) = \gamma_k^{ij} g_k^j + v_k^{ij}$ and thus $w_k^{ij}(\gamma_k^{ij}) | \gamma_k^{ij} \sim \mathcal{N}\left(0, W_{ij}(\gamma_k^{ij})\right)$ with $W_{ij}(\gamma_k^{ij}) = \gamma_k^{ij} R_j + V_{ij}$ for $i \in \mathcal{V}$ and $j \in \mathcal{E}_i$. Proceeding similarly to [6], we set

$$\boldsymbol{y}_k^{ij} \doteq \mathrm{col}(y_k^{ij}, y_{k-1}^{ij}, \ldots, y_{k-L}^{ij}), \tag{14}$$

$$\boldsymbol{x}_k \doteq \mathrm{col}(x_k, x_{k-1}, \ldots, x_{k-L}), \tag{15}$$

$$\boldsymbol{\gamma}_k^{ij} \doteq \mathrm{col}(\gamma_k^{ij}, \gamma_{k-1}^{ij}, \ldots, \gamma_{k-L}^{ij}), \tag{16}$$

$$\boldsymbol{w}_k^{ij}(\boldsymbol{\gamma}_k^{ij}) \doteq \mathrm{col}(w_k^{ij}(\gamma_k^{ij}), \ldots, w_{k-L}^{ij}(\gamma_{k-L}^{ij})) \tag{17}$$

and the diagonal matrix

$$\Gamma_k^{ij} = \mathrm{diag}(\boldsymbol{\gamma}_k^{ij}). \tag{18}$$

With the positions above, the decision rule (12) becomes

$$\hat{\gamma}_k^{ij} = \arg \max_{\hat{\gamma}_k^{ij} \in \{0,1\}} \mathrm{P}_k \left(\hat{\gamma}_k^{ij} | \boldsymbol{y}_k^{ij} \right)$$

$$= \arg \max_{\hat{\gamma}_k^{ij} \in \{0,1\}} \sum_{\hat{\gamma}_{k-1}^{ij} \in \{0,1\}} \sum_{\hat{\gamma}_{k-2}^{ij} \in \{0,1\}} \cdots$$

$$\cdots \sum_{\hat{\gamma}_{k-L}^{ij} \in \{0,1\}} f\left(\boldsymbol{y}_k^{ij} | \boldsymbol{\gamma}_k^{ij} \right) \bar{\mathrm{P}}_k \left(\boldsymbol{\gamma}_k^{ij} \right), \tag{19}$$

where $f(\boldsymbol{y}_k^{ij} | \boldsymbol{\gamma}_k^{ij})$ is the probability density function of $\boldsymbol{y}_k^{ij}$ conditional on $\boldsymbol{\gamma}_k^{ij}$ and $\bar{\mathrm{P}}_k(\boldsymbol{\gamma}_k^{ij})$ is the probability mass function of $\boldsymbol{\gamma}_k^{ij}$, namely the joint probability mass function of $\gamma_k^{ij}, \gamma_{k-1}^{ij}, \ldots, \gamma_{k-L}^{ij}$ which is known by assumption 5 of Section II-B. To derive $f(\boldsymbol{y}_k^{ij} | \boldsymbol{\gamma}_k^{ij})$ note that we have the following relation

$$\boldsymbol{y}_k^{ij} = (\Gamma_k^{ij} \otimes C_j) \boldsymbol{x}_k + \boldsymbol{w}_k^{ij}(\boldsymbol{\gamma}_k^{ij}),$$

with $\boldsymbol{w}_k^{ij}(\boldsymbol{\gamma}_k^{ij}) | \boldsymbol{\gamma}_k^{ij} \sim \mathcal{N}\left(0, \boldsymbol{W}_{ij}(\boldsymbol{\gamma}_k^{ij})\right)$, where

$$\begin{aligned} \boldsymbol{W}_{ij}(\boldsymbol{\gamma}_k^{ij}) &= W_{ij}(\gamma_k^{ij}) \oplus W_{ij}(\gamma_{k-1}^{ij}) \oplus \cdots \oplus W_{ij}(\gamma_{k-L}^{ij}) \\ &= \bigoplus_{i=0}^{L} W_{ij}(\gamma_{k-i}^{ij}). \end{aligned} \tag{20}$$

From the equation above, it follows that $\boldsymbol{y}_k^{ij}$ conditional on $\boldsymbol{\gamma}_k^{ij}$ is Gaussian with

$$\boldsymbol{y}_k^{ij} | \boldsymbol{\gamma}_k^{ij} \sim \mathcal{N}\left(0, (\Gamma_k^{ij} \otimes C_j) \boldsymbol{\Xi}_k (\Gamma_k^{ij} \otimes C_j^\top) + \boldsymbol{W}_{ij}(\boldsymbol{\gamma}_k^{ij})\right) \tag{21}$$

where $\boldsymbol{\Xi}_k = \mathbb{E}\left\{\boldsymbol{x}_k \boldsymbol{x}_k^\top\right\}$. It is easy to see that the matrix $\boldsymbol{\Xi}_k$ has the following structure

$$\boldsymbol{\Xi}_k = \begin{bmatrix} \Sigma_k & A\Sigma_{k-1} & \cdots & A^L \Sigma_{k-L} \\ \Sigma_{k-1}A^\top & \Sigma_{k-1} & \cdots & A^{L-1}\Sigma_{k-L} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{k-L}(A^L)^\top & \Sigma_{k-L}(A^{L-1})^\top & \cdots & \Sigma_{k-L} \end{bmatrix} \tag{22}$$

where $\Sigma_h = \mathbb{E}\{x_h x_h^\top\}$, with $h = k-L, \ldots, k$, can be computed iteratively as follows

$$\Sigma_{k+1} = A\Sigma_k A^\top + Q.$$

Finally, by using (21) and simplifying terms, the decision rule (19) becomes

$$\hat{\gamma}_k^{ij} = \arg \max_{\gamma_k^{ij} \in \{0,1\}} \sum_{\gamma_{k-1}^{ij} \in \{0,1\}} \sum_{\gamma_{k-2}^{ij} \in \{0,1\}} \cdots$$
$$\cdots \sum_{\gamma_{k-L}^{ij} \in \{0,1\}} \frac{\exp\left[-\frac{1}{2} \boldsymbol{y}_k^{ij\top} \boldsymbol{Z}_k^{ij^{-1}} \boldsymbol{y}_k^{ij}\right]}{\sqrt{\left|\boldsymbol{Z}_k^{ij}\right|}} \overline{P}_k(\boldsymbol{\gamma}_k^{ij}) \quad (23)$$

where

$$\boldsymbol{Z}_k^{ij} = (\Gamma_k^{ij} \otimes C_j) \boldsymbol{\Xi}_k (\Gamma_k^{ij} \otimes C_j^\top) + \boldsymbol{W}_{ij}(\boldsymbol{\gamma}_k^{ij}).$$

A relevant case is when $\{\gamma_k^{ij}\}$ is a Markov chain with initial probability $\pi_0^{ij} = [P_0^{ij}(0), P_0^{ij}(1)]$ and transition matrix

$$\Pi_k^{ij} = \begin{bmatrix} P_k(0|0) & P_k(1|0) \\ P_k(0|1) & P_k(1|1) \end{bmatrix}, \quad (24)$$

where $P_k(\alpha|\beta) = \mathbb{P}\{\gamma_k^{ij} = \alpha | \gamma_{k-1}^{ij} = \beta\}^2$. The probability mass function $P_k(\gamma_k^{ij})$ is given by

$$P_k(\gamma_k^{ij}) = \pi_0^{ij} \prod_{\ell=1}^{k} \Pi_\ell^{ij} \begin{bmatrix} 1 - \gamma_k^{ij} \\ \gamma_k^{ij} \end{bmatrix}, \quad (25)$$

which simplifies in the case of homogeneous chain as

$$P_k(\gamma_k^{ij}) = \pi_0^{ij} \Pi^{ij^k} \begin{bmatrix} 1 - \gamma_k^{ij} \\ \gamma_k^{ij} \end{bmatrix}.$$

With these positions, the probability mass function $\overline{P}_k(\boldsymbol{\gamma}_k^{ij})$ in (23) can be written as

$$\overline{P}_k(\boldsymbol{\gamma}_k^{ij}) = \prod_{i=0}^{L-1} P_{k-i}(\gamma_{k-i}^{ij}|\gamma_{k-i-1}^{ij}) P_{k-L}(\gamma_{k-L}^{ij})$$
$$= \prod_{i=0}^{L-1} \Pi_{k-i}^{ij}(\gamma_{k-i-1}^{ij} + 1, \gamma_{k-i}^{ij} + 1) P_{k-L}(\gamma_{k-L}^{ij})$$
$$(26)$$

where $\Pi_{k-i}^{ij}(\gamma_{k-i-1}^{ij}+1, \gamma_{k-i}^{ij}+1)$ is the element in the position $(\gamma_{k-i-1}^{ij} + 1, \gamma_{k-i}^{ij} + 1)$ of the matrix Π_{k-i}^{ij} and $P_{k-L}(\gamma_{k-L}^{ij})$ is defined as in (25).

The block scheme in Figure 1 represents an example of information processing towards a node. Note that the proposed optimal detector applies to both stable and unstable systems.

IV. DISTRIBUTED FILTERING ALGORITHM

We state in this Section the complete algorithm of the proposed distributed filter with *unknown* random link failures. Firstly, we note that the decision rule (12) requires $L + 1$ measurements and thus, the optimal detectors when $k \leq L$ use the available observations until time k, namely

$$\hat{\gamma}_k^{ij} = \arg \max_{\gamma_k^{ij} \in \{0,1\}} P_k\left(\gamma_k^{ij} | y_k^{ij}, \ldots, y_1^{ij}\right)$$

whereas, when $k > L$ the decision rule (12) should be used. Hence, we modify the consensus-based Kalman filtering

---

^{2}As for $P_k(\cdot)$, with a slight abuse of notation, we use $P_k(\alpha|\beta)$ instead of $P_k^{ij}(\alpha|\beta)$.
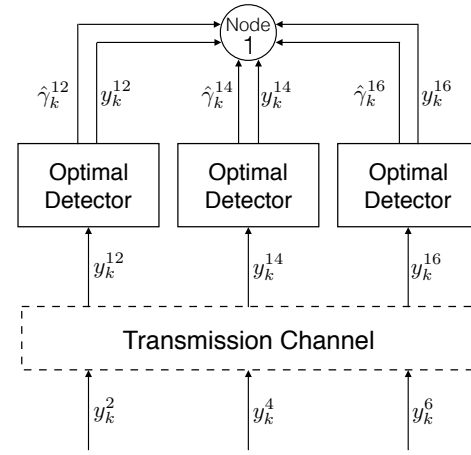

Fig. 1. Block scheme of the exchanged information towards node 1 in the example of Section V.

algorithm (4)–(10) of [18] for any $i \in \mathcal{V}$ as follows
Initialization:

$$\bar{x}_0^i = 0, \qquad P_0^i = \Sigma_0. \quad (27)$$

Optimal decisions of link failures $\forall j \in \mathcal{E}_i$:

$$\hat{\gamma}_k^{ij} = \begin{cases} \arg \max_{\gamma_k^{ij} \in \{0,1\}} P_k(\gamma_k^{ij} | y_k^{ij}, \ldots, y_1^{ij}), & \text{if } k \leq L \\ \\ \arg \max_{\gamma_k^{ij} \in \{0,1\}} P_k(\gamma_k^{ij} | y_k^{ij}, \ldots, y_{k-L}^{ij}), & \text{if } k > L \end{cases}$$
$$(28)$$

Locally aggregate data and covariance matrices $\forall j \in J_i$:

$$u_k^{ij} = C_j^\top \widehat{W}_{ij}^{-1} y_k^{ij}, \quad U_{ij} = C_j^\top \widehat{W}_{ij}^{-1} C_j, \quad (29)$$

where $y_k^{ii} \doteq y_k^i$, $\widehat{W}_{ij} \doteq \hat{\gamma}_k^{ij} R_j + V_{ij}$, $\hat{\gamma}_k^{ii} \doteq 1$ and $V_{ii} \doteq 0$,

$$z_k^i = \sum_{j \in J_i} \hat{\gamma}_k^{ij} u_k^{ij}, \qquad \widehat{Z}_k^i = \sum_{j \in J_i} \hat{\gamma}_k^{ij} U_{ij}. \quad (30)$$

Compute the Kalman-consensus estimate:

$$\widehat{M}_k^i = \left(P_k^{i^{-1}} + \widehat{Z}_k^i\right)^{-1}, \quad (31)$$

$$\hat{x}_k^i = \bar{x}_k^i + \widehat{M}_k^i \left(z_k^i - \widehat{Z}_k^i \bar{x}_k^i\right) + \epsilon \widehat{M}_k^i \sum_{j \in \mathcal{E}_i} \left(\bar{x}_k^j - \bar{x}_k^i\right). \quad (32)$$

Update the state of the Kalman-consensus filter:

$$P_{k+1}^i = A \widehat{M}_k^i A^\top + Q, \quad (33)$$
$$\bar{x}_{k+1}^i = A \hat{x}_k^i. \quad (34)$$

As before, in (32) the scalar $\epsilon > 0$ is a relative small constant (*e.g.* chosen to be of the order of the integration time-step of the discretization of a continuous-time process [18]). We note that the sets $\mathcal{E}_i$ and consequently J_i are constant for $k \geq 0$, however the fact that two nodes could be disconnected at some times is captured by $\hat{\gamma}_k^{ij}$ in (29)–(30). We remark that each node should run the algorithm (27)–(34) which is scalable in the number of nodes N. Moreover, each node i should solve the optimization problem (28) for each of its d_i neighbors, thus the complexity of the MAP decision strategy (28) is $O(L2^L d_i)$. We note that, according to the local computational power of the nodes, the value of L could be node-dependent and adjusted according to the number of the d_i neighbors.

V. SIMULATION EXAMPLE

In this section we provide a simulation example, with two different scenario of probability distributions of the link failures, to show the effectiveness of the proposed approach. The undirected graph is characterized by $\mathcal{V} = \{1, 2, \ldots, 6\}$ and it contains the edges $\mathcal{E} = \{(1, 2), (1, 4), (1, 6), (2, 3), (3, 5), (4, 5), (5, 6)\}$, thus the nodes $\{1, 5\}$ have three neighbors and the nodes $\{2, 3, 4, 6\}$ have two neighbors. We note that $\sum_{i=1}^{6} d_i = 14$. We consider the same continuous-time target dynamics as in [17] (*i.e.* a point moving on noisy circular trajectories) given by

$$\dot{x}_t = A_0 x_t + n_t, \quad A_0 = 2 \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad (35)$$

where $n_t \sim \mathcal{N}(0, Q_0)$, with $Q_0 = 5 \cdot 10^{-2}$. The associated discrete-time model has the form (1) where $A = I + \epsilon A_0 + \frac{\epsilon^2}{2} A_0^2 + \frac{\epsilon^3}{6} A_0^3$ and $x_0 \sim (0, I)$, $f_k \sim (0, Q)$ with $Q = \epsilon Q_0$ and $\epsilon = 0.015$ (≈ 70 Hz). The sensor i performs noisy measurements given by (2) where $g_k^i \sim (0, R)$ with $R = 2 \cdot 10^{-2}$ for all $i \in \mathcal{V}$, $C_i = [1, 0]$ for $i = 1, 3, 5$ and $C_i = [0, 1]$ for $i = 2, 4, 6$. Moreover, at each $k \geq 0$, the sensor i receives the measurements y_k^{ij} with $j \in \mathcal{E}_i$ modeled by (3) where $v_k^{ij} \sim (0, V)$ with $V = 2 \cdot 10^{-3}$ for all $i \in \mathcal{V}$ and $j \in \mathcal{E}_i$. For all the links, the random sequences $\{\gamma_k^{ij}\}$, representing the link failures, is modeled as a Markov chain characterized by the same probability transition matrix and we consider two scenarios, i.e. for all $i \in \mathcal{V}$ and for all $j \in \mathcal{E}_i$ we have

$$\Pi_1 = \begin{bmatrix} 0.05 & 0.95 \\ 0.1 & 0.9 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} 0.15 & 0.85 \\ 0.2 & 0.8 \end{bmatrix}. \qquad (36)$$

We stress that, the matrix Π_1 characterizes the event of having a link failure with a small probability. Nonetheless, we shall see that these packet losses among nodes, even if small, are responsible of a not negligible degradation of the performance. To this end, we compare the proposed approach, *i.e.* algorithm (27)–(34), with the algorithm of [18] which does not consider any packet loss in the transmission, *i.e.* algorithm (27)–(34) with $\hat{\gamma}_k^{ij} = 1$ for all $i \in \mathcal{V}$, $j \in \mathcal{E}$ and $k \geq 0$ and the algorithm of [18] with known sequence of link failures, *i.e.* algorithm (4)–(10) which represents an ideal (not implementable) filter since it makes use of the real values of the sequence $\{\gamma_k^{ij}\}$ (that are not available in reality). We consider a time-horizon of 150 and 300 independent realizations. We shall compute the mean square error (MSE), evaluated averaging the arithmetic mean of the MSE of the nodes over independent realizations of state and output noise and mode sequences, namely

$$\text{MSE}(k) = \frac{1}{300} \frac{1}{6} \sum_{r=1}^{300} \sum_{i=1}^{6} \left\| \hat{x}_k^{r,i} - x_k^r \right\|^2,$$

where x_k^r and $\hat{x}_k^{r,i}$ denote the sample path of the state and of the estimated state of the sensor i, relative to the r-th realization of noise and mode sequences at time k.
We consider also an index which measures the disagreement of estimates (as in [18]) defined as

$$\delta(k) = \frac{1}{300} \sum_{r=1}^{300} \left(\sum_{i=1}^{6} \left\| \hat{x}_k^{r,i} - \mu_k^r \right\|^2 \right)^{1/2},$$

where $\mu_k^r = \frac{1}{6} \sum_{i=1}^{6} \hat{x}_k^{r,i}$. Figure 2 shows the scenario with Π_1, namely the average MSE and the average disagreement of estimates defined above over the first 25 steps of the Distributed filter (DF) of [18], the proposed DF where the detector has memory $L = 0$ and $L = 1$, and the ideal DF of [18] with known sequences of link failures $\{\gamma_k^{ij}\}$. We stress the fact that the last algorithm, *i.e.* the ideal DF of [17] with known $\{\gamma_k^{ij}\}$, is just a theoretical limit of the performance: it is not implementable in reality since it makes use of the not available values γ_k^{ij}. We see that, even though the probability of the link failures described by the matrix Π_1 is small, the proposed approach outperforms the DF algorithm of [18] in terms of estimation error. Moreover, it provides even much cohesive estimates. Table I summarizes the results for both scenarios that run with different probability transition matrices Π_1 and Π_2 (36). It shows the averaged MSE and the averaged disagreement of estimates, *i.e.* $\overline{\text{MSE}} \doteq \frac{1}{150} \sum_{k=0}^{150} \text{MSE}(k)$ and $\Delta \doteq \frac{1}{150} \sum_{k=0}^{150} \delta(k)$, a performance index and the probability of error, namely

$$P_{\text{err}} \doteq \frac{1}{300} \frac{1}{150} \frac{1}{6} \frac{1}{14} \sum_{r=1}^{300} \sum_{k=0}^{150} \sum_{i=1}^{6} \sum_{j \in \mathcal{E}_i} \mathbb{P}\{\hat{\gamma}_k^{r,i,j} \neq \gamma_k^{r,i,j}\}$$

and

$$\alpha_{\overline{\text{MSE}}} \doteq 10^2 \cdot \frac{(\overline{\text{MSE}}^{\text{DF}} - \overline{\text{MSE}})}{\overline{\text{MSE}}^{\text{DF}}} \%,$$

where $\hat{\gamma}_k^{r,i,j}$ and $\gamma_k^{r,i,j}$ are the detected and real value of γ_k^{ij} for the r-th realization of noise and mode sequences at time k, $\overline{\text{MSE}}$ is the averaged MSE of the used approach and $\overline{\text{MSE}}^{\text{DF}}$ is the averaged MSE of the DF of [18] with unknown sequences of link failures. As the intuition suggests, all the algorithms performs better in terms of estimation error and disagreement of estimates when the probability of receiving the message ($\gamma_k^{ij} = 1$) is high: in the scenario with Π_1 the probability of receiving the message is higher than the scenario with Π_2, indeed the MSE and the disagreement of estimates Δ are consistent with the intuition above. Moreover, the more Π is characterized by a high probability of receiving the message from the neighbors the more all the curves collapse towards the ideal filter performance, *i.e.* the ideal DF of [17] with known $\{\gamma_k^{ij}\}$ sequences. We see as, in the scenario with Π_1, the proposed approach with the detectors with memory $L = 0$ outperforms the DF of [17], it improves the $\overline{\text{MSE}}$ of 44% and Δ of 34%.

Scenario Π_1	$\overline{\text{MSE}}\, 10^{-2}$	$\Delta\, 10^{-2}$	α_{MSE}	$P_{\text{err}}\, 10^{-2}$
DF of [18]	2.98	31.8	-	9.5
Proposed DF ($L = 0$)	1.65	20.9	44.7%	6.3
Proposed DF ($L = 1$)	1.53	19.1	48.7%	3.6
Ideal DF of [18]	1.47	17.7	50.9%	0

Scenario Π_2	$\overline{\text{MSE}}\, 10^{-2}$	$\Delta\, 10^{-2}$	α_{MSE}	$P_{\text{err}}\, 10^{-2}$
DF of [18]	6.56	45.9	-	19
Proposed DF ($L = 0$)	1.67	21.2	74.5%	7.4
Proposed DF ($L = 1$)	1.56	19.9	75.8%	5.2
Ideal DF of [18]	1.50	18.4	77.1%	0

TABLE I
COMPARISON OF THE DISTRIBUTED FILTER (DF) OF [18], THE PROPOSED DF WITH $L = 0$ AND $L = 1$, AND THE IDEAL DF OF [18].
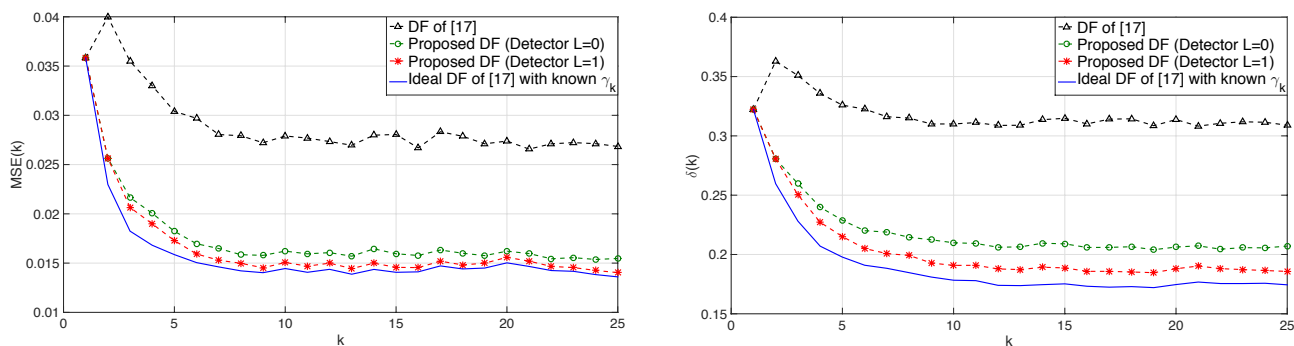
Fig. 2. Comparison of the Distributed filter (DF) of [18], the proposed DF where the detector has memory $L = 0$ and $L = 1$, and the ideal DF of [18] with known sequences of link failures $\{\gamma_k^{ij}\}$. It shows the MSE(k)(left) and the disagreement of estimates $\delta(k)$ (right) over the first 25 steps. The time-horizon is 150 and the number of realizations is 300.

The example demonstrates the importance of having a decision strategy even though the probability of failures among links is small.

VI. CONCLUSIONS

In this paper we consider the distributed consensus-based filtering problem for linear time-invariant systems over sensor networks subject to random link failures when the failure sequence is not known at the receiving side. Our approach extends the method proposed in [17]-[18] to the case of unreliable channels by introducing a failure detection strategy proposed for the case of intermittent observations in [6]-[7]. For each sensor, the proposed approach consists of three stage: some local nonlinear optimal detectors, the locally aggregation of data and covariance matrices and a Kalman-consensus filter fed by the decisions of the first step. The numerical example demonstrates the importance of having a decision strategy even though the probability of failures among links is small. The performance of the proposed method gets closer to the filter with known sequences of link failures when the memory L of the detectors increases. Moreover, a compromise between computational cost of the detectors and the performance of the overall distributed filter can be easily achieved by changing the value L of the carried observations.

REFERENCES

[1] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano. Consensus CPHD filter for distributed multitarget tracking. *IEEE Journal of Selected Topics in Signal Processing*, 7(3), 508–520, 2013.

[2] G. Battistelli, L. Chisci, G. Mugnai, A. Farina, and A. Graziano. Consensus-based linear and nonlinear filtering. *IEEE Trans on Autom. Contr.*, 60(5), 1410–1415, 2015.

[3] S. Das and J. Moura. Distributed Kalman filtering with dynamic observations consensus,. *IEEE Trans. Signal Processing*, vol. 63, no. 17, 4458–4473, 2015.

[4] F. Delli Priscoli, A. Di Giorgio, F. Lisi, S. Monaco, A. Pietrabissa, L. Ricciardi Celsi, V. Suraci, Multi-Agent Quality of Experience Control. *International Journal of Control, Automation, and Systems, 2017, vol. 15, no. 2, pp. 892-904*

[5] D. Ding, Z. Wang, D. W. C. Ho and G. Wei, Distributed recursive filtering for stochastic systems under uniform quantizations and deception attacks through sensor networks. *Automatica*, 78, 231-.240, 2017.

[6] A. Fasano, A. Germani, and A. Monteriú. A detection-estimation approach to filtering for Gaussian systems with intermittent observations. In *Decision and Control, 53rd IEEE Conference on*, 4172–4177, 2014.

[7] A. Fasano, A. Monteriú, and V. Villani. A detection-estimation approach to filtering with intermittent observations with generally correlated packet dropouts. In *Decision and Control, 54th IEEE Conference on*, 4356–4361, 2015.

[8] Y. Hatano and M. Mesbahi. Agreement over random networks. *IEEE Trans on Autom. Contr.*, 50(11), 1867–1872, 2005.

[9] S. Kar, and J.M. Moura. Sensor networks with random links: topology design for distributed consensus. *IEEE Trans. Signal Processing*, 56(7), 3315–3326, 2008.

[10] S. Kar, and J.M. Moura. Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise. *IEEE Trans. Signal Processing*, 57(1), 355-369, 2009.

[11] K. Lee, H.W. Lee, and E. Modiano. Reliability in layered networks with random link failures. *IEEE/ACM Transactions on Networking*, 19(6), 1835–1848, 2011.

[12] Q. Liu, Z. Wang, X. He, and D. H. Zhou. Event-based recursive distributed filtering over wireless sensor networks. *IEEE Trans on Autom. Contr.*, 60(9), 2470–2475, 2015.

[13] Q. Liu, Z. Wang, X. He, G. Ghinea, and F.E. Alsaadi. A Resilient Approach to Distributed Filter Design for Time-Varying Systems Under Stochastic Nonlinearities and Sensor Degradation. *IEEE Trans. Signal Processing*, 65(5), 1300–1309, 2017.

[14] Q. Liu, Z. Wang, X. He, and D. Zhou. On Kalman-Consensus Filtering with Random Link Failures over Sensor Networks. *IEEE Trans on Autom. Contr.*, (in press), 2017.

[15] L. Ma, Z. Wang, Q.L. Han, and H.K. Lam. Variance-constrained distributed filtering for time-varying systems with multiplicative noises and deception attacks over sensor networks. *IEEE Sensors Journal*, 17(7), 2279–2288, 2017.

[16] I. Matei, and J.S. Baras. Consensus-based linear distributed filtering. *Automatica*, 48(8), 1776–1782, 2012.

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

[18] R. Olfati-Saber. Kalman-consensus filter: Optimality, stability, and performance. In *Decision and Control, 48th IEEE Conference on*, 7036–7042, 2009.

[19] S. Patterson, B. Bamieh, and A. El Abbadi. Convergence tates of distributed average consensus with stochastic link failures. *IEEE Trans on Autom. Contr.*, 55(4), 880–892, 2010.

[20] S. Silva Pereira, A. Pagès-Zamora. Mean Square Convergence of Consensus Algorithms in Random WSNs. *IEEE Trans. Signal Processing*, 58(5), 2866–2774, 2010.

[21] A. Simonetto, T. Keviczky, and R. Babuška. Distributed nonlinear estimation for robot localization using weighted consensus. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, 3026–3031, 2010.

[22] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan and S. S. Sastry. Kalman filtering with intermittent observations. *IEEE Trans on Autom. Contr.*, 49(9), 1453..1464, 2004.

[23] A. Tahbaz-Salehi and A. Jadbabaie. Consensus over ergodic stationary graph processes. *IEEE Trans on Autom. Contr.*, 55(1), 225–230, 2010.

[24] B. Zhou, J. Cao, X. Zeng, and H. Wu. Adaptive traffic light control in wireless sensor network-based intelligent transportation system. *In Vehicular technology conference fall (VTC 2010-Fall), 2010 IEEE 72nd (pp. 1-5). IEEE.*