

Distributed Anomaly Detection and Estimation over Sensor Networks: Observational-Equivalence and Q -Redundant Observer Design

Mohammadreza Doostmohammadian, Themistoklis Charalambous, *Senior Member, IEEE*

Abstract—In this paper, we study stateless and stateful physics-based anomaly detection scenarios via *distributed estimation* over sensor networks. In the stateful case, the detector keeps track of the sensor residuals (i.e., the difference of estimated and true outputs) and reports an alarm if certain statistics of the recorded residuals deviate over a predefined threshold, e.g., χ^2 (Chi-square) detector. Instead, only instantaneous deviation of the residuals raises the alarm in the stateless case without considering the history of the sensor outputs and estimation data. Given (approximate) false-alarm rate for both cases, we propose a probabilistic threshold design based on the noise statistics. We show by simulation that increasing the window length in the stateful case may not necessarily reduce the false-alarm rate. On the other hand, it adds unwanted delay to raise the alarm. The distributed aspect of the proposed detection algorithm enables local isolation of the faulty sensors with possible recovery solutions by adding redundant observationally-equivalent sensors. We, then, offer a mechanism to design Q -redundant distributed observers, robust to failure (or removal) of up to Q sensors over the network.

Index Terms—Anomaly detection, networked estimation, observational-equivalence, q -redundant observability

I. INTRODUCTION

The recent advancements in wireless communication, high-performance networking, sensing, and processing devices, along with cloud-computing and Internet-of-Thing (IoT) [1], [2] have motivated localized (or *distributed*) estimation over sensor networks. Such localized setup, further, mandates *distributed detection* mechanisms to find possible anomalies (faults, failures, or malicious attacks) *locally*, with potential large-scale applications from social networks [3] to Cyber-Physical Energy Systems (CPES) [4], [5]. Such distributed methods are privileged with *no single node of failure* and outperform existing centralized detection [6]–[9].

The existing distributed methods either require (i) *local* system observability in the neighborhood of every sensor via high-traffic communication network with large quantity of data-transfer [2], or (ii) fast data-sharing and processing units to perform many iterations of consensus and communication between every two samples of system dynamics (*double time-scale*) [10], [11]. In this work, similar to [12]–[14], we consider *single time-scale* distributed estimation (i.e., iterations at the same time scale of system dynamics) with no assumption on local system observability at any sensor [4], [15] to relax the networking, communication,

and computational needs. A challenge is to build distributed estimation networks robust to faults, anomalies, and even cyber-attacks. Such detection and mitigation techniques are wide-spread in *centralized* setup [16]. Some works propose preventive mechanisms as in privacy-preserving [17], and resilient estimation [18], [19], while many other use detection mechanisms, e.g., observer-based (or predictive) fault detection and isolation (FDI) [8], [9]. Machine learning and binary classification algorithms (e.g., support-vector-machine [20]) are also used to generate decision boundaries separating *normal space* of system states from *abnormal space* (operating with fault or anomaly) [7]. Graph-theoretic methods via *structured systems theory* are also adopted for *generic* design of detection/mitigation protocols irrespective of numerical parameter values and only based on system *structure* [21], [22]. What missing in the literature is a *distributed* algorithm to *localize* the detection, in both stateful and stateless setups, along with network design for *redundant distributed observability* (distributed fault-tolerant models).

Main contributions: We propose a local mechanism to detect possible bias/anomaly in system output while performing single time-scale distributed estimation. We design probabilistic detection thresholds in both stateless and stateful scenarios, define the false-alarm (false-positive) rate (FAR) and false-negative rate for each case, and compare the performance of both solutions in terms of FAR (and delay in raising the alarm). Further, in contrast to the full-rank model in [3], [12], this work considers distributed detection on (possibly) rank-deficient systems, known to require more information-sharing over the distributed estimation networks [15]. We specifically use constrained LMI gain design to *isolate* (faulty) sensor residuals. Another contribution is to design Q -redundant distributed observers, robust to failure or removal of any Q (failed or faulty) sensors. Using graph-theory (structural methods), we first define the set of *observationally-equivalent* sensors/state-outputs using *strongly-connected-components* (SCCs) and *contractions* in the *system digraph*. Then, using k -vertex-connected (or k -connected) graphs [23], we propose sensor-network structures resilient to Q sensor removals such that the remaining sensor-network successfully tracks the underlying system (which is not necessarily observable in the neighborhood of any sensor). The proposed algorithms are of polynomial-order complexity.

Paper organization: the problem is set up in Section II. The distributed detection mechanism and Q -redundant observer are proposed in Section III and IV. The simulations and conclusion are given in Section V and VI.

Authors are with the School of Electrical Engineering at Aalto University, Espoo, Finland, Email: firstname.surname@aalto.fi. M. Doostmohammadian is also with the Faculty of Mechanical Engineering at Semnan University, Semnan, Iran, Email: doost@semnan.ac.ir. T. Charalambous is also with the Electrical Engineering Department at the University of Cyprus, surname.name@ucy.ac.cy.

II. THE FRAMEWORK

A. System model

In this paper, we consider LTI systems in the form,

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + \boldsymbol{\nu}(k), \quad (1)$$

with k as the time index, A as the system matrix (possibly rank-deficient), $\mathbf{x} \in \mathbb{R}^n$ as the column-vector of states, i.e., $\mathbf{x} = [x_1; \dots; x_n]$, and $\boldsymbol{\nu} \sim \mathcal{N}(0, \Sigma_\nu)$ as the Gaussian noise. The output vector at time k is in the form,

$$\mathbf{y}(k) = C\mathbf{x}(k) + \boldsymbol{\zeta}(k) + \mathbf{f}(k), \quad (2)$$

with $\mathbf{y} = [y_1, \dots, y_N] \in \mathbb{R}^N$, $\boldsymbol{\zeta} \sim \mathcal{N}(0, \Sigma_\zeta)$ as the output noise (with diagonal Σ_ζ), and $\mathbf{f} \in \mathbb{R}^N$ as the *additive bias* to the outputs (due to faults or anomalies) defined as,

$$\begin{cases} f_i(k) = 0 & \text{no-fault at output } i \\ f_i(k) \neq 0 & \text{faulty output } i \end{cases} \quad (3)$$

Without loss of generality, in this paper, we assign every sensor i with one state-output y_i . Given the structural representation of A, C matrices (i.e., the *system digraph* \mathcal{G}_A), sufficient conditions for structural (or *generic*) (A, C) -observability are given in [15], [24]. Such structural methods imply observability for *almost all* numerical values of non-zero system parameters. This is referred to as the *linear-structure-invariant* (LSI) model, where the structure is fixed and the non-zero entries in A and C can take almost any value (with non-admissible parameter values lying on an algebraic subspace with zero Lebesgue measure) [15].

B. Distributed consensus-based estimation

We consider a distributed estimation framework, where the (group of) sensors estimate the system A locally via information-sharing over a sensor network to gain *distributed observability* (as defined later in Lemma 1). The following distributed estimator in *single time-scale*, proposed in [15], is considered at sensor i .

$$\hat{\mathbf{x}}_i(k|k-1) = \sum_{j \in \mathcal{N}_\beta(i)} W_{ij} A \hat{\mathbf{x}}_j(k-1|k-1), \quad (4)$$

$$\begin{aligned} \hat{\mathbf{x}}_i(k|k) = & \hat{\mathbf{x}}_i(k|k-1) \\ & + K_i \sum_{j \in \mathcal{N}_\alpha(i)} C_j \left(y_j(k) - C_j^\top \hat{\mathbf{x}}_i(k|k-1) \right), \end{aligned} \quad (5)$$

where $\hat{\mathbf{x}}_i(k|k-1)$ and $\hat{\mathbf{x}}_i(k|k)$ denote the priori and posteriori estimates via the received data up to time $k-1$ and k , respectively. K_i is the local gain (to be designed), and $\mathcal{N}_\alpha(i)$ and $\mathcal{N}_\beta(i)$ denote the in-neighborhood respectively over networks \mathcal{G}_α and \mathcal{G}_β (defined in Lemma 1), with 0-1 matrix $U = [U_{ij}]$ and row-stochastic matrix $W = [W_{ij}]$. Recall that $\mathcal{N}_\alpha(i)$ includes the so-called α -sensors with outputs of rank-deficient part of the system [15], [21], where for $j \in \mathcal{N}_\alpha(i)$ the entry $U_{ij} = 1$. The following assumption distinguishes this work from many literature on single time-scale distributed estimation and observer-based detection.

Assumption 1: The pair (A, C) is observable and the pair $(A, C_{j \in \mathcal{N}_\alpha(i)})$ (with $C_{j \in \mathcal{N}_\alpha(i)}$ representing outputs in α -neighborhood of i) is *not necessarily observable* at any

sensor i . In other words, we assume *global observability* at the group of *all sensors* and *no local observability* in the *neighborhood of any sensor*.

Remark 1: The proposed single time-scale protocol, similar to [12]–[14], is more suitable for large-scale, as compared to double time-scale methods [10], [11], in terms of needed communication traffic and computation loads on sensors. This is because, in the latter, sensors perform a large number of consensus/data-sharing iterations between steps k and $k+1$ of system dynamics, in contrast to only 1 iteration in single time-scale methods. See details in [4], [12], [15].

Define the estimation error as $\mathbf{e}_i(k) = \hat{\mathbf{x}}_i(k|k) - \mathbf{x}(k)$ and $\mathbf{e} = [\mathbf{e}_1, \dots, \mathbf{e}_N]$. The error dynamics under (4)–(5) is,

$$\mathbf{e}(k+1) = (W \otimes A) \mathbf{e}(k) - K D_C (W \otimes A) \mathbf{e}(k) + \boldsymbol{\eta}(k+1), \quad (6)$$

$$\boldsymbol{\eta}(k+1) = (I - K D_C) \mathbf{1}_N \otimes \boldsymbol{\nu}(k) - K \bar{D}_C (\boldsymbol{\zeta}(k+1) + \mathbf{f}(k+1)) \quad (7)$$

with $\boldsymbol{\eta}(k+1)$ containing the noise and fault terms, $nN \times nN$ matrices $D_C = \text{diag}[U_{ij} C_j C_j^\top]$ and $\bar{D}_C \triangleq (U \otimes \mathbf{1}_n) \circ (\mathbf{1}_N \otimes C^\top)$ ("o" and " \otimes " respectively as the entry-wise and Kronecker product). Let $\hat{A} := W \otimes A - K D_C (W \otimes A)$. Eq. (6) represents a cumulative LTI dynamics with system matrix \hat{A} , where $\rho(\hat{A})$ determines the error stability (with $\rho(\cdot)$ as the spectral radius). Following the Kalman stability theorem, one can design block-diagonal gain matrix $K = K \circ \mathcal{K} = \text{diag}[K_i]$ (with $\mathcal{K} = I_N \otimes \mathbf{1}_{n \times n}$, I_N and $\mathbf{1}_{n \times n}$ respectively as identity and ones matrix of size N , n) to stabilize (6) if the pair $(W \otimes A, D_C)$ is observable. This is referred to as the distributed observability, with the sufficient conditions discussed in the next lemma.

Lemma 1 ([15]): The pair $(W \otimes A, D_C)$ is observable if (i) \mathcal{G}_β is strongly-connected, and (ii) every α -sensor is a *hub* of \mathcal{G}_α , where every sensor i with a system-output recovering *structural rank-deficiency*¹ is regarded as an α -sensor.

Then, the block-diagonal gain matrix K can be designed, e.g., via the LMI in [15], [26], such that the error dynamics is Schur stable, i.e., $\rho(\hat{A}) < 1$ for general (possibly unstable) A . Define the steady state error variance as $\Sigma_e = \lim_{k \rightarrow \infty} \mathbb{E}(\mathbf{e}_i(k) \mathbf{e}_i(k)^\top)$. It is shown in [15], [26] that $\lim_{k \rightarrow \infty} \mathbb{E}(\mathbf{e}_i(k)) = \mathbf{0}$, and,

$$\|\Sigma_e\|_2 = \frac{a_1 N \|\Sigma_\nu\|_2 + a_2 \|\bar{\Sigma}_\zeta\|_2}{1 - b^2}, \quad (8)$$

with $b := \|(W \otimes A) - K D_C (W \otimes A)\|_2 < 1$, $a_1 := \|I_{Nn} - K D_C\|_2^2$, $a_2 := \|K\|_2^2$, $\bar{\Sigma}_\zeta := \text{diag}[\sum_{j \in \mathcal{N}_\alpha(i)} C_j^\top \Sigma_\zeta^j C_j]$, where Σ_ζ^j denotes the j th diagonal entry of Σ_ζ .

III. MAIN RESULTS: LOCAL DETECTION OF ANOMALIES

A. Stateless Detector

Given the error dynamics (6)–(7), define the residual of sensor i (including possible anomaly/bias f_i) at time k as,

$$\begin{aligned} r_i(k) &= y_i(k) - C_i \hat{x}_i(k|k) = C_i \mathbf{e}_i(k) + \zeta_i(k) + f_i(k) \\ &= C_i \hat{A}_i \mathbf{e}_i(k-1) + C_i \boldsymbol{\eta}_i(k) + \zeta_i(k) + f_i(k) \end{aligned} \quad (9)$$

¹Later in Section IV, using the system graph representation, an α -sensor is defined as the sensor with output of an state node in a contraction component. We refer interested readers to [25] for more details.

Algorithm 1: Constrained iterative LMI gain design

1 Input: matrices A, W, U, C , scale factor ϵ
2 Calculate $D_C = \text{diag}[U_{ij}C_jC_j^\top]$;
3 if $(W \otimes A, D_C)$ is not observable **then**
4 Terminate;
5 else
6 Iteratively solve for $\hat{A} = W \otimes A - KD_C(W \otimes A)$
 min $\text{trace}(XY)$
 s.t. $X, Y \succ 0$, $K \leftarrow K \circ \mathcal{K}$
 $\begin{pmatrix} X & \hat{A}^\top \\ \hat{A} & Y \end{pmatrix} \succ 0$, $\begin{pmatrix} X & I \\ I & Y \end{pmatrix} \succ 0$,
 $\frac{|C_i K_i C_j^\top|}{|1 - C_j K_j C_j^\top|} < \epsilon, \forall j \in \mathcal{N}_\alpha(i), j \neq i$
7 Output Block-diagonal gain matrix $K = K \circ \mathcal{K}$;

with

$$\begin{aligned} \zeta_i(k) + f_i(k) + C_i \eta_i(k) &= \zeta_i(k) + f_i(k) + C_i \nu(k-1) \\ &- C_i K_i \sum_{j \in \mathcal{N}_\alpha(i)} (C_j^\top \zeta_j(k) + C_j^\top f_i(k) + C_j C_j^\top \nu(k-1)) \end{aligned} \quad (10)$$

In steady-state (large enough k), the noise and (non-zero) fault terms in (10) mainly value the residual of sensor i . From (10), other than bias f_i , any fault in the α -neighborhood of sensor i , say $f_j, j \in \mathcal{N}_\alpha(i)$, also appears in $r_i(k)$. This makes the fault isolation more challenging as compared to the full-rank system models [3], [12] in which $\mathcal{N}_\alpha(i) = \{i\}$. To overcome this and isolate the faults on $j \in \mathcal{N}_\alpha(i), j \neq i$, from (10), we constrain the gain matrix K_i such that,

$$|C_i^\top K_i C_j| \leq \epsilon |1 - C_j^\top K_j C_j|, \forall j \in \mathcal{N}_\alpha(i), j \neq i \quad (11)$$

with $\epsilon < 1$ as a design constant, implying that the fault-related term $C_i^\top K_i C_j f_j$ in the residual $r_i(k)$ due to fault f_j is down-scaled by ϵ as compared to the fault-related term in $r_j(k)$ (of α -sensor j itself). The constraint (11) helps to isolate possible faults at α -sensors in rank-deficient systems via Algorithm 1 as the modified version of the LMI gain design in [12]. This algorithm can be run either once centralized offline with K_i s given to the nodes after termination or iteratively over k . From (8) and results in [12], [27], one can define the *confidence intervals* for $r_i(k)$. First, note that from (9) for the case of no anomaly/fault $f_i(k) = 0, \forall i$ and,

$$\Sigma_r^i \leq \frac{\|C_i\|_2 \|\Sigma_e\|_2}{N} + \Sigma_\zeta^i \quad (12)$$

where, assuming one state measurement by every sensor i , $\|C_i\|_2 = |C_i|$. Then, the probability that the residual lies in the confidence interval $|r_i(k)| \leq \kappa \Sigma_r^i$ is $\text{erf}(\frac{\kappa}{\sqrt{2}})$ with $\text{erf}(\cdot)$ as the *Gauss error function*. One can similarly claim that the probability of $|r_i(k)| \geq \kappa \Sigma_r^i$ is $1 - \text{erf}(\frac{\kappa}{\sqrt{2}})$ for zero additive bias $f_i(k) = 0$. Define the *probabilistic threshold* $\theta_p = \kappa \Sigma_r^i$. Then, the detection logic in this paper is as

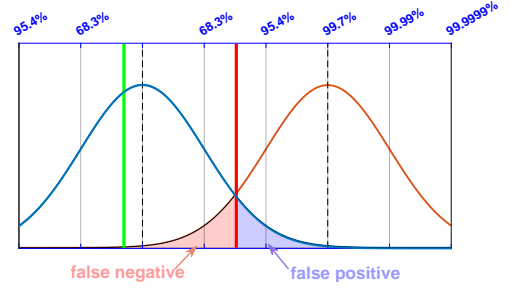


Fig. 1. This figure shows example distributions for unbiased residual with $f_i(k) = 0$ (blue curve) versus the biased residual with $f_i(k) \neq 0$ (red curve). The confidence intervals (of the unbiased PDF) are shown with the associated probabilities. Two vertical green and red lines represent two example residual values r_k^i , where the green one (most likely) belongs to the unbiased curve as it is close to the expected value (zero) of the unbiased PDF and, thus, most likely is due to system/measurement noise; the red sensor residual, however, is far from the expected value (zero) and, based on the shown confidence-intervals, represents anomalous output with FAR $4.6\% < p < 31.7\%$ (more accurately $p = 1 - \text{erf}(\frac{\kappa=1.5}{\sqrt{2}}) = 13.3\%$ with $\theta_p = 1.5\Sigma_r^i$). Considering the absolute residual $|r_k^i|$ and folding-over the red PDF to the right-hand-side (RHS) of 0 axis, the blue shaded area equals to $\frac{p}{2}$, representing half of the probability of false-alarm, where the other half on the LHS of 0 is not shown for simplicity. Similarly, the red shaded area, equal to $\frac{p}{2}$, approximately represents the false-negative rate.

follows: for given FAR (or false-positive probability) p define the threshold θ_p with $\kappa = \sqrt{2}\text{erf}^{-1}(1-p)$. If $|r_i(k)| \geq \theta_p$ raise the anomaly detection alarm (associated with FAR p). Recall that we use *absolute residual* $|r_i(k)|$, and thus, *folded* Gaussian distribution, i.e., to fold-over the probability mass to the RHS of 0 axis by taking the absolute value. Similarly, the probability of *false negative* can be defined as,

$$\bar{p} = \frac{1 - \text{erf}(\kappa/\sqrt{2})}{2} - \frac{1 - \text{erf}(3\kappa/\sqrt{2})}{2} = \frac{\text{erf}(3\kappa/\sqrt{2}) - \text{erf}(\kappa/\sqrt{2})}{2} \quad (13)$$

For $\kappa \geq 2$, we have $1 - \text{erf}(3\kappa/\sqrt{2}) \approx 2e-10$ and one can approximate \bar{p} by $\frac{p}{2}$. This is better explained in Fig. 1. This detection logic uses the instant value of the $r_i(k)$ at every time k with no use of residual history, known as the *stateless* mechanism. Next, we consider residual history for anomaly detection, which is known as the *stateful* mechanism.

B. Stateful Detector

Define *distance measure* $\iota_i(k)$ as,

$$\iota_i^T(k) = \sum_{m=k-T+1}^k \frac{r_i(m)^2}{\Sigma_r^i}. \quad (14)$$

over a *sliding time-window* of length T . It is known that the summation of squared random variables from normal distribution (i.e., $\frac{r_i(k)^2}{\Sigma_r^i}$) follows the so-called *Chi-squared* distribution with T degrees of freedom (denoted by χ_T^2) [3], [16], [28], where $\mathbb{E}(\iota_i^T) = T$. Following the same line of reasoning as in the stateless case, the probabilistic threshold on the variable ι_i^T (for given FAR p) can be defined as,

$$\theta_p^T = 2\Gamma^{-1}(1-p, \frac{T}{2}) \quad (15)$$

where $\Gamma^{-1}(\cdot, \cdot)$ denotes the *inverse regularized lower incomplete gamma function*. Weighted distance measure [6]

Algorithm 2: Localized Detection of Output Bias

- 1 **Input:** $A, W_{ij}, U_{ij}, C_i, y_i$ at sensor i , FAR p , K_i via Algorithm 1, time-window T , weight-factor μ
 - 2 Find $r_i(k)$ via (9), $\iota_i^T(k)$ via (14), or $\bar{\iota}_i(k)$ via (16) at step k ;
 - 3 Define $\theta_p = \kappa \Sigma_r^i$, θ_p^T via (15), or θ_p^μ via (17);
 - 4 **begin** Stateless detection:
 - 5 **if** $r_i(k) \geq \theta_p$ **then**
 - 6 Declare: alarm at sensor i ;
 - 7 **begin** Stateful detection:
 - 8 **if** $\iota_i^T(k) \geq \theta_p^T$ or $\bar{\iota}_i(k) \geq \theta_p^\mu$ **then**
 - 9 Declare: alarm at sensor i ;
 - 10 **Output** Binary decision (Alarm or no-Alarm) associated with pre-specified FAR p ;
-

further can be considered to put more weights on the recent (normalized) residuals and less on the far past residuals as,

$$\bar{\iota}_i(k) = \sum_{m=k-T+1}^k \mu^{k-m} \frac{r_i(m)^2}{\Sigma_r^i}. \quad (16)$$

with $0 < \mu \leq 1$ as the weight factor. $\bar{\iota}_i$ is referred to as the weighted sum of Chi-squared distributions [28]. Although, χ^2 -distribution is typically defined for positive integer T , one can find similar results for the weighted χ^2 given by (16) [28], with the expected value of $\bar{\iota}_i$ and the probabilistic thresholds (for a pre-specified FAR p) similar to (15),

$$\mathbb{E}(\bar{\iota}_i) = \frac{1 - \mu^T}{1 - \mu}, \quad \theta_p^\mu = 2\Gamma^{-1} \left(1 - p, \frac{1 - \mu^T}{2 - 2\mu} \right) \quad (17)$$

For the stateful detectors (15) and (17), the FAR is a function of T and μ using the CDF of the χ^2 -distribution,

$$p = 1 - \frac{\gamma(\frac{T}{2}, \frac{T}{2})}{\Gamma(\frac{T}{2})}, \quad p = 1 - \frac{\gamma(\frac{\bar{\iota}_i}{2}, \frac{1 - \mu^T}{2 - 2\mu})}{\Gamma(\frac{1 - \mu^T}{2 - 2\mu})} \quad (18)$$

with $\gamma(\cdot, \cdot)$ as the *lower incomplete gamma function*. Our fault detection logics are summarized in Algorithm 2 with either stateless (first **if**) or stateful detection (second **if**).

Remark 2: For the stateful case, longer time-window T results in less FAR; however, it also increases the false alarm delay [16]. A similar statement holds for the weight factor μ ; greater μ results in lower FAR and more-delayed alarm. In general, there is a trade-off between detection accuracy (lower FAR via increasing T and μ) and detection delay in raising the alarm; see examples in Section V.

Remark 3: The proposed detection logic in Algorithm 2 is *localized* at every sensor i with no need of centralized decision making. This is more feasible in large-scale applications as compared to the centralized detection methods [7]–[9].

IV. Q-REDUNDANT DISTRIBUTED OBSERVER DESIGN BASED ON STRUCTURAL OBSERVATIONAL EQUIVALENCE

In this section, we provide Q -redundant version of the distributed estimator (4)–(5) such that it tolerates removal (or

isolation) of any Q (faulty) sensors while holding distributed observability over the remaining sensor-network. Thus, the other sensors can locally estimate the system and detect possible output bias/anomaly. This problem is twofold: (i) it provides sufficient outputs such that after removal of up to Q rows of matrix C , the pair (A, C) remains structurally observable (known as Q -redundant observability [19]), and (ii) designs Q -redundant communication network such that the conditions in Lemma 1 hold after removing Q sensors and cutting their linking over the network. We address these via the notion of observational-equivalence [25], [29].

A. Q -redundant observability

Given an observable pair (A, C) , two outputs $y_i = C_i \mathbf{x}$ and $y_j = C_j \mathbf{x}$ are *observationally-equivalent* if losing either of the two does not affect system observability, while removing both makes the system *unobservable*. Let \bar{C}_i denote the output matrix after removing the row C_i from C (i.e., removing sensor i). Then, for observationally equivalent sensors/outputs i, j , (A, \bar{C}_i) and (A, \bar{C}_j) are observable, but $(A, \bar{C}_{i,j})$ is not observable. Consider the system digraph $\mathcal{G}_A = \{\mathcal{V}, \mathcal{E}\}$ with set $\mathcal{V} = \{1, \dots, n\}$ denoting the state nodes and link set \mathcal{E} denoting the state interactions. Define a *contraction*, \mathcal{C}_l , as the set of nodes such that $|\mathcal{N}(\mathcal{C}_l)| < |\mathcal{C}_l|$, where $\mathcal{N}(\mathcal{C}_l) = \{j | (i, j) \in \mathcal{E}, i \in \mathcal{C}_l\}$ [25], [30]. Define an SCC, \mathcal{S}_l , as the set of nodes $i, j \in \mathcal{V}$ such that $i \xrightarrow{\text{path}} j$ and $j \xrightarrow{\text{path}} i$. A *parent SCC* is, then, defined as an SCC with no outgoing link to any other SCC [25]. Output of parent SCCs are known to recover the *output-connectivity* of \mathcal{G}_A , and output of contractions are known to recover the *cyclicity* of \mathcal{G}_A and rank of the system matrix A [25].

Lemma 2 ([15], [24]): Given system digraph \mathcal{G}_A , outputs of one state node in every parent SCC \mathcal{S}_l and every contraction \mathcal{C}_l are sufficient for structural (A, C) -observability, i.e., to satisfy Assumption 1.

The above lemma implies that the set of state nodes in a contraction \mathcal{C}_l and a parent SCC \mathcal{S}_l in \mathcal{G}_A are observationally equivalent. In this direction, $Q + 1$ different state-outputs from every contraction and parent-SCC are sufficient for Q -redundant observability. We assign these outputs to $Q + 1$ sensors. Then, removing any Q sensors, the remaining ones include (at least) one output from every \mathcal{S}_l and \mathcal{C}_l . Thus, from Lemma 2, the pair (A, \bar{C}_Q) (with \bar{C}_Q as the output matrix after removal of any Q rows) remains observable.

B. Q -redundant distributed observer

Given the outputs and sensors for Q -redundant observability, this section provides the Q -redundant distributed observer. Following Lemma 1, we improve the network-connectivity of \mathcal{G}_β and \mathcal{G}_α to gain Q -redundant $(W \otimes A, D_C)$ -observability as follows,

- 1) Design \mathcal{G}_β to be Q -vertex-connected, e.g., via computationally-efficient algorithms in [31]. Recall that a Q -vertex-connected graph *remains strongly-connected after cutting any Q (sensor) nodes from the network*.
- 2) For the *hub-network* \mathcal{G}_α , following Lemma 1, every α -sensor i (say α_1) as a network hub, directly shares its

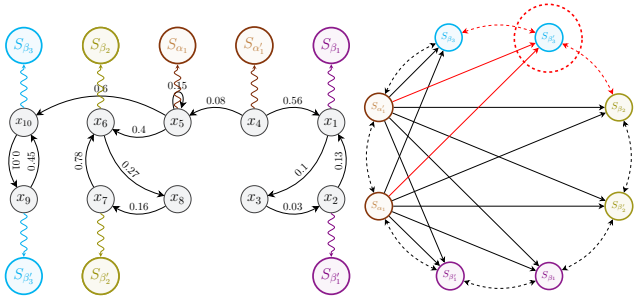


Fig. 2. (Left) This figure shows an example system digraph \mathcal{G}_A (gray nodes) and sensors with outputs of state nodes in 4 parent SCCs and contraction. A link from, e.g., node x_1 to x_3 with weight 0.1 in \mathcal{G}_A implies that $A_{31} = 0.1$. Sensors of the same color (e.g., blue colored β_3 and β_1) are observationally equivalent with outputs of the same component (SCC or contraction). This implies that by isolating/removing one faulty output, the system remains observable via the other ones. For example, removing β_3 , \mathcal{G}_A is globally observable to the remaining 7 sensors. (Right) This figure shows an example sensor network with solid links as network \mathcal{G}_α (α sensors as the hubs) and dashed-links as an example 1-connected SC network \mathcal{G}_β . Both \mathcal{G}_β and \mathcal{G}_α include self-loops (not shown for simplicity). By this setup, $(W \otimes A, D_C)$ is 1-redundant observable, implying that by isolating any one sensor, e.g., β_3 (shown via red dashed circle) and removing its incoming/outgoing links (red-colored), $(W \otimes A, D_C)$ remains observable and sufficient conditions in Lemmas 1 and 2 hold.

output with every other sensor j not observationally equivalent with i , (all other sensors except α_2).

The above connectivity ensures that every sensor directly links from $Q + 1$ set of α -sensors and $Q + 1$ paths to and from other sensors. Then, protocol (4)-(5) over these Q -redundant networks \mathcal{G}_β and \mathcal{G}_α can tolerate failure/isolation of any Q sensors without losing distributed observability over the network (from Lemma 1), i.e., a Q -redundant distributed estimator. In similar setup, Q -edge-connectivity can be used for survivable network design [32] resilient to link removal.

C. Illustrative Example

To illustrate the Q -redundant results, consider the system digraph \mathcal{G}_A in Fig. 2(Left), containing three parent SCCs $\mathcal{S}_1 = \{x_1, x_2, x_3\}$, $\mathcal{S}_2 = \{x_6, x_7, x_8\}$, $\mathcal{S}_3 = \{x_9, x_{10}\}$ and one contraction $\mathcal{C}_1 = \{x_4, x_5, x_2, x_7, x_9\}$. We aim to design a 1-redundant distributed observer, i.e., an observer robust to removal/isolation of any one output (sensor). Following Section IV-A, taking 4 outputs from $\{x_1, x_5, x_6, x_9\}$ the system digraph is structurally observable. By adding another set of 4 outputs from states $\{x_2, x_4, x_7, x_{10}\}$, the system is (structurally) 1-redundant observable. Following Section IV-B, we design \mathcal{G}_β as the 2-vertex-connected graph and \mathcal{G}_α as a hub-network shown in Fig. 2(Right) to gain 1-redundant distributed observability. Then, in protocol (4)-(5), matrix W (as the adjacency of \mathcal{G}_β) is designed row-stochastic; for example, by considering random positive non-zero entries, and then, dividing each row by the row-sum. Matrix U is the 0-1 adjacency matrix of \mathcal{G}_α . Then, block-diagonal K can be designed via Algorithm 1. This distributed observer is robust to failure of 1 sensor, or, more precisely, to failure of 1 output from every parent SCC and contraction.

V. SIMULATION

For simulation, we consider the same example in Section IV-C representing a rank deficient system of $n =$

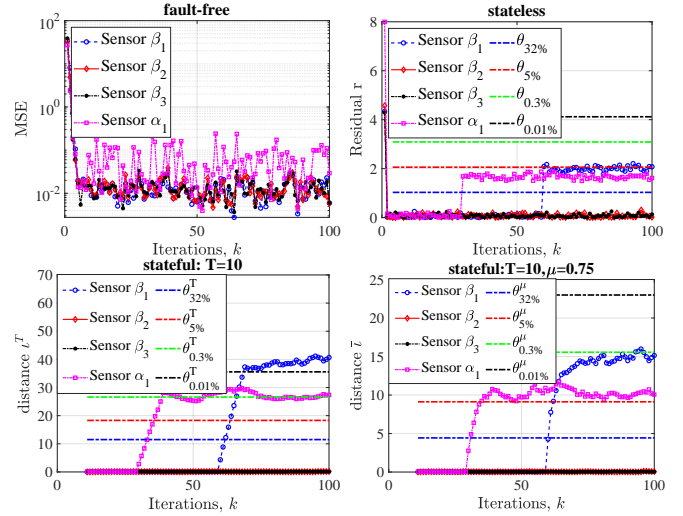


Fig. 3. This figure shows (TopLeft) the bounded-MSE performance of the proposed protocol (4)-(5) in the *absence* of output bias, i.e., $f_i = 0, \forall i$, and *localized detection* in the *presence* of output bias, $f_{\beta_1}(k \geq 60) = 2$ and $f_{\alpha_1}(k \geq 30) \sim \mathcal{N}(2, 0.5)$: (TopRight) stateless case considering no history of the residuals, (BelowLeft) stateful case over time-window $T = 10$ and equally-weighting all residual history, (BelowRight) stateful case over the same time-window and weighting the residual history by $\mu = 0.75$.

10 nodes with non-zero entries as the given link weights in Fig. 2(Left). We consider $N = 4$ outputs from $\{x_1, x_6, x_9, x_5\}$ associated with sensors $\{\beta_1, \beta_2, \beta_3, \alpha_1\}$, satisfying Lemma 2 and Assumption 1. Network \mathcal{G}_α includes direct links (with weight 1) from α_1 to $\beta_1, \beta_2, \beta_3$, and \mathcal{G}_β is considered as a directed cycle $1 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow 1$ with random row-stochastic link weights. The non-zero entries of the output matrix C are set equal to 1. The gain matrix K is designed via Algorithm 1 with $\epsilon = 0.14$. The mean-square error (MSE) under the proposed protocol (4)-(5) is bounded steady-state as shown in Fig. 3(TopLeft) with $\nu \sim \mathcal{N}(0, 0.01)$ and $\zeta \sim \mathcal{N}(0, 0.01)$. To check the performance in the presence of faults or anomalies, we consider two bias $f_{\beta_1}(k \geq 60) = 2$ and $f_{\alpha_1}(k \geq 30) \sim \mathcal{N}(2, 0.5)$. Our localized (or distributed) detection logic follows algorithm 2 for FAR $p = 32\%, 5\%, 0.3\%, 0.01\%$. For the stateless case, the residual r_i at the faulty sensors β_1 and α_1 are between the thresholds $\theta_{32\%}$ and $\theta_{5\%}$. Therefore, the local detector declares the fault probability of more than 68% and less than 95% at these sensors and no fault at the other two. For the stateful case, considering the distance measure \bar{r}_i^T in (14) with $T = 10$ -steps and thresholds (15), the local detector declares higher probabilities of 99.7% and 99.99% respectively at sensors α_1 and β_1 , however *with certain delay* (about 10-steps delay as in Fig. 3(BelowLeft)). One can reduce this alarm delay by weighting the residual history by μ as in (16), however via thresholds of higher FAR. For $\mu = 0.75$, the weighted distance measure \bar{r}_i and the thresholds (17) are shown in Fig. 3(BelowRight). The detection probability with this logic is approximately 99.7% at sensor α_1 and 95% at sensor β_1 , with alarm delay reduced to 7 time-steps.

Next, we compare the performance of the stateful detectors for different values of T and μ . Consider a constant fault resulting in biased residual $r_i^2 = 2\Sigma_r^i$. For the stateless

detector, the associated FAR is equal to 5%. For the stateful case, the threshold's FAR can be defined via (18), shown in Fig. 4(Left) for different T and μ values. Clearly, for greater μ , the FAR is lower; however, as discussed in Remark 2 and Fig. 3, large μ values result in longer delays to raise the alarm. Further, for smaller values, e.g., $\mu = 0.5, 0.6$, the FAR is almost constant for $T \geq 8$ and, thus, longer time windows do not improve the FAR. Fig. 4(Right) shows the FAR versus $\frac{r_i^2}{\sum r_i^2}$ for different μ values (for $T = 8$). It is clear that as $\mu \rightarrow 1$ the FAR decreases (with $\mu = 1$ giving (15)).

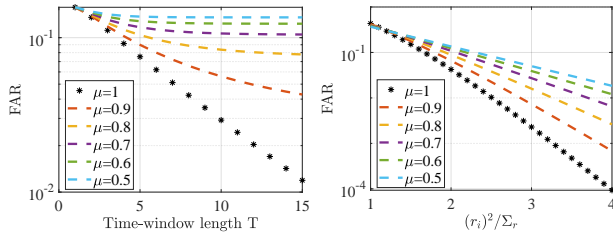


Fig. 4. (Left) FAR increases for some shorter time-windows T and smaller μ (for normalized residual $\frac{r_i^2}{\sum r_i^2} = 2$). (Right) FAR increases for larger values of normalized residual and smaller values of μ in (16) (for fixed $T = 8$).

VI. CONCLUSION AND FUTURE WORKS

Stateful and stateless local detection mechanisms over distributed estimation networks are considered with a trade-off between alarm delay and FAR. The stateful case shows lower FAR with possibly delayed alarm over long sliding time windows. The solutions are of polynomial-order complexity; for example, system digraph decomposition into parent SCCs and contractions is of complexity $\mathcal{O}(n^{2.5})$ [30]. As future research, one may consider possible time-delay in the communication network [33], [34].

REFERENCES

- [1] M. Doostmohammadian and H. R. Rabiee, "On the observability and controllability of large-scale IoT networks: Reducing number of unmatched nodes via link addition," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1747–1752, 2021.
- [2] Y. Chen, S. Kar, and J. M. F. Moura, "The internet of things: Secure distributed inference," *IEEE Signal Proc. Mag.*, vol. 35, no. 5, pp. 64–75, 2018.
- [3] M. Doostmohammadian, T. Charalambous, M. Shafie-khah, N. Meskin, and U. A. Khan, "Simultaneous distributed estimation and attack detection/isolation in social networks: Structural observability, kronecker-product network, and chi-square detector," in *IEEE International Conference on Autonomous Systems*, 2021, pp. 344–348.
- [4] U. A. Khan and M. Doostmohammadian, "A sensor placement and network design paradigm for future smart grids," in *IEEE Workshop on Computational Adv. in Multi-Sensor Adap. Proc.*, 2011, pp. 137–140.
- [5] M. D. Ilić, L. Xie, U. A. Khan, and J. M. F. Moura, "Modeling of future cyber-physical energy systems for distributed sensing and control," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 40, no. 4, pp. 825–838, 2010.
- [6] D. Umsonst, *Tuning of anomaly detectors in the presence of sensor attacks*, Ph.D. thesis, KTH Royal Institute of Technology, 2019.
- [7] M. Abbaszadeh, "System and method for anomaly and cyber-threat detection in a wind turbine," 2019, US Patent App. 15/988,515.
- [8] M. Davoodi, N. Meskin, and K. Khorasani, "Event-triggered multi-objective control and fault diagnosis: A unified framework," *IEEE Trans. on Industrial Informatics*, vol. 13, no. 1, pp. 298–311, 2016.
- [9] M. Navi, N. Meskin, and M. Davoodi, "Sensor fault detection and isolation of an industrial gas turbine using partial adaptive KPCA," *Journal of Process Control*, vol. 64, pp. 37–48, 2018.

- [10] X. He, X. Ren, H. Sandberg, and K. H. Johansson, "How to secure distributed filters under sensor attacks?," *IEEE Transactions on Automatic Control*, *arXiv preprint arXiv:2004.05409*, 2021.
- [11] S. Battilotti, F. Cacace, and M. d'Angelo, "A stability with optimality analysis of consensus-based distributed filters for discrete-time linear systems," *Automatica*, vol. 129, pp. 109589, 2021.
- [12] M. Doostmohammadian and N. Meskin, "Sensor fault detection and isolation via networked estimation: Full-rank dynamical systems," *IEEE Trans. Control of Net. Systems*, vol. 8, no. 2, pp. 987–996, 2021.
- [13] L. Wang, J. Liu, A. S. Morse, and B. D. O. Anderson, "A distributed observer for a discrete-time linear system," in *58th IEEE Conference on Decision and Control*, 2019, pp. 367–372.
- [14] D. Marelli, M. Zamani, M. Fu, and B. Ninness, "Distributed kalman filter in a network of linear systems," *Systems & Control Letters*, vol. 116, pp. 71–77, 2018.
- [15] M. Doostmohammadian and U. Khan, "On the genericity properties in distributed estimation: Topology design and sensor placement," *IEEE J. of Sel. Topics in Signal Processing*, vol. 7, no. 2, pp. 195–204, 2013.
- [16] J. Giraldo, D. Urbina, A. Cardenas, J. Valente, M. Faisal, J. Ruths, N. O. Tippenhauer, H. Sandberg, and R. Candell, "A survey of physics-based attack detection in cyber-physical systems," *ACM Computing Surveys (CSUR)*, vol. 51, no. 4, pp. 1–36, 2018.
- [17] A. I. Rikos, T. Charalambous, K. H. Johansson, and C. N. Hadjicostis, "Privacy-preserving event-triggered quantized average consensus," in *IEEE Conference on Decision and Control*, 2020, pp. 6246–6253.
- [18] A. Mitra, F. Ghawash, S. Sundaram, and W. Abbas, "On the impacts of redundancy, diversity, and trust in resilient distributed state estimation," *IEEE Transactions on Control of Network Systems*, 2021.
- [19] J. Kim, C. Lee, H. Shim, Y. Eun, and J. H. Seo, "Detection of sensor attack and resilient state estimation for uniformly observable nonlinear systems having redundant sensors," *IEEE Transactions on Automatic Control*, vol. 64, no. 3, pp. 1162–1169, 2018.
- [20] M. Doostmohammadian, A. Aghasi, T. Charalambous, and U. Khan, "Distributed support vector machines over dynamic balanced directed networks," *IEEE Control Systems Let.*, vol. 6, pp. 758 – 763, 2021.
- [21] M. Doostmohammadian and U. A. Khan, "Vulnerability of CPS inference to DoS attacks," in *48th Annual Asilomar Conference on Signals, Systems, and Computers*, 2014, pp. 2015–2018.
- [22] S. Gracy, J. Milošević, and H. Sandberg, "Security index based on perfectly undetectable attacks: Graph-theoretic conditions," *Automatica*, vol. 134, pp. 109925, 2021.
- [23] D. Guichard, "An introduction to combinatorics and graph theory," *Whitman College-Creative Commons*, 2017.
- [24] S. Pequito, S. Kar, and A. P. Aguiar, "A structured systems approach for optimal actuator-sensor placement in linear time-invariant systems," in *American Control Conference*, 2013, pp. 6123–6128.
- [25] M. Doostmohammadian and U. A. Khan, "Measurement partitioning and observational equivalence in state estimation," in *IEEE Conference on Acoustics, Speech and Signal Processing*, 2016, pp. 4855–4859.
- [26] U. A. Khan and A. Jadbabaie, "Coordinated networked estimation strategies using structured systems theory," in *49th IEEE Conference on Decision and Control*, Orlando, FL, Dec. 2011, pp. 2112–2117.
- [27] U. A. Khan and A. Jadbabaie, "Collaborative scalar-gain estimators for potentially unstable social dynamics with limited communication," *Automatica*, vol. 50, no. 7, pp. 1909–1914, 2014.
- [28] J. Bausch, "On the efficient calculation of a linear combination of chi-square random variables with an application in counting string vacua," *Journal of Physics A*, vol. 46, pp. 505202, 2013.
- [29] M. Doostmohammadian, T. Charalambous, M. Shafie-khah, H. R. Rabiee, and U. A. Khan, "Analysis of contractions in system graphs: Application to state estimation," in *IEEE International Conference on Autonomous Systems*, 2021, pp. 359–363.
- [30] K. Murota, *Matrices & matroids for systems analysis*, Springer, 2000.
- [31] Y. Wu and Y. Li, "Construction algorithms for k-connected m-dominating sets in wireless sensor networks," in *9th ACM symposium on Mobile ad hoc networking and computing*, 2008, pp. 83–90.
- [32] A. Jabal Ameli, *Approximation algorithms for survivable network design*, Ph.D. thesis, Università della Svizzera italiana, 2021.
- [33] M. Doostmohammadian, M. Pirani, U. A. Khan, and T. Charalambous, "Consensus-based distributed estimation in the presence of heterogeneous, time-invariant delays," *IEEE Control Systems Letters*, vol. 6, pp. 1598 – 1603, 2021.
- [34] C. N. Hadjicostis and T. Charalambous, "Average consensus in the presence of delays in directed graph topologies," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 763–768, 2013.