2018 IEEE Conference on Decision and Control (CDC)
Miami Beach, FL, USA, Dec. 17-19, 2018

# Attack Detection in Stochastic Interconnected Systems: Centralized vs Decentralized Detectors

Rajasekhar Anguluri, Vaibhav Katewa, and Fabio Pasqualetti

*Abstract*— **In this paper we consider a security problem for stochastic interconnected systems, where the objective is to detect attacks on the system dynamics from sparse measurements. We consider two classes of detectors, namely centralized and decentralized detectors, which differ primarily in their knowledge of the system model. In particular, a decentralized detector has a model of the dynamics of the isolated subsystems, but is unaware of the interconnections between them. Instead, a centralized detector has a model of the entire dynamical system. We characterize the performance of the two detectors, and show that, depending on the systems parameters and attack strategy, each of the detectors can outperform the other. In other words, it may be possible for a decentralized detector to outperform a centralized detector, which, as we discuss, is due to the considered hypothesis testing problem. Finally, we illustrate our results through a set of numerical examples.**

## I. Introduction

Modern dynamical systems are increasingly becoming more distributed, diverse, complex, and integrated with cyber components. Typically, these systems are composed of multiple subsystems, which are interconnected via physical couplings among the states of the subsystems. An example of such systems is the smart city, which comprises of power grid, transportation system, and a water distribution network, among other [1], [2], [3]. Although these subsystems are interconnected, they are often operated independently. As a result, they may have limited information about the overall system dynamics. Further, the subsystem operators may not be willing to share information with the other subsystems due to security and privacy concerns, and it may be difficult to directly measure the interconnection signals between the subsystems. For these reasons, interconnected cyber-physical systems may be more vulnerable targets to cyber/physical attacks, which greatly degrade the performance [4].

During the last few years there have been many studies on analyzing the effect of different types of attacks on dynamical systems, and possible remedial strategies [5]. A key component of these strategies is the assumption that a detector responsible for making decisions about attacks has complete knowledge of the overall system dynamics. Yet, due to the limited information available to subsystems or their unwillingness to share information, it is not always feasible to test for attacks on the global system. One possible strategy to tackle this situation is to allocate the detection tasks to each subsystem independently. In this paper we develop such a local detection strategy, which enables each subsystem to detect local attacks in the absence of information regarding other subsystems. We consider a scenario in which the local detectors cooperate to detect attacks on the interconnected system, thereby resulting in a decentralized attack detection strategy. We compare this decentralized strategy with a centralized strategy, in which a central detector detect attacks on the whole system based on the complete knowledge of the overall system model.

**Related Work:** Since the past decade, researchers extensively studied both the detection and the estimation of the attacks in the stochastic systems, under the realm of cyber-physical system (CPS) security. A few notable works in this direction includes [6], [7], [8], [9]. For more interesting results in the security of CPS, we refer the reader to [5].

We note that the most of the existing works on the attack detection assumes a centralized detector. Although, the decentralized detection strategies are well studied in the realm of communications and signal processing communities [10], [11], in the context of dynamical systems framework, decentralized detectors have recently been proposed [12], [13], [14]. In all these works, authors mainly studied the distributed attack detection for the deterministic dynamical systems using geometric control techniques. Hence, many of the existing tools cannot be readily extended to the stochastic systems and thus, in this paper, we take a step forward to develop a decentralized detector for attack detection in the stochastic systems along with a centralized detector.

**Contributions:** The main contributions of this work are as follows. First, we present two classes of detectors, namely centralized and decentralized detectors, to decide on the presence of deterministic additive attacks against interconnected systems driven by noise. Our detection schemes are based on the decision theoretic framework that falls under the category of *simple vs composite* statistical hypotheses testing. We characterize the detection probability of the detectors as a function of the system knowledge available to a detector, i.e., system dynamics, noise statistics, and attack parameters. Although, intuitively one may believe that a centralized detector must perform better than a decentralized detector, contrast to that belief, we show that there exist some scenarios, depending on the system dynamics and attack parameters, such that a decentralized detector can outperform a centralized detector.

**Paper Organization:** The rest of the paper is organized as follows. Section II contains the problem setup and some preliminary notions. Section III presents our measurement processing methods and sets up our hypothesis testing frame-

This material is based upon work supported in part by ARO award 71603NSYIP, and in part by NSF awards ECCS1405330 and BCS1631112.
Rajasekhar Anguluri, Vaibhav Katewa, Fabio Pasqualetti are with the Department of Mechanical Engineering, University of California at Riverside, {ranguluri,vkatewa,fabiopas}@engr.ucr.edu.

work. The performance comparisons of the centralized and decentralized detectors are detailed in Section IV. In Section V we provide numerical examples to support our theoretical findings. Finally, in Section VI we conclude the paper.

## II. Problem setup and preliminary notions

We consider the following discrete-time, linear, stochastic interconnected system composed of $N$ subsystems:

$$\begin{aligned} x_i(k+1) &= A_{ii}x_i(k) + \sum_{j\neq i}^{N} A_{ij}x_j(k) + w_i(k), \\ y_i(k) &= C_i x_i(k) + v_i(k), \end{aligned} \tag{1}$$

where $x_i \in \mathbb{R}^{n_i}$ is the state of the $i$-th subsystem, and $w_i \sim \mathcal{N}(0, \Xi_i)$, $v_i \sim \mathcal{N}(0, \Upsilon_i)$ are the process and measurement noises. We assume that the noise vectors are mutually independent, that is, $w_i(k) \perp\!\!\!\perp v(k)$ $(\forall k \in \mathbb{N})$, and that $w_i \perp\!\!\!\perp w_j$, $v_i \perp\!\!\!\perp v_j$ whenever $i \neq j$. Let $n = \sum_{i=1}^{N} n_i$ and, to simplify the notation, let $B_i u_i = \sum_{j\neq i}^{N} A_{ij}x_j$, where

$$\begin{aligned} B_i &= \begin{bmatrix} A_{i1} & \cdots & A_{i,i-1} & A_{i,i+1} & \cdots & A_{iN} \end{bmatrix} \text{ and }, \\ u_i^{\mathsf{T}} &= \begin{bmatrix} x_1^{\mathsf{T}} & \cdots & x_{i-1}^{\mathsf{T}} & x_{i+1}^{\mathsf{T}} & \cdots & x_N^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{n-n_i}. \end{aligned}$$

We allow for the presence of attackers that compromise the dynamics of the subsystems, and we model such attacks as exogenous unknown inputs. In particular, the system dynamics of the $i$-th subsystem when subject to the attack $B_i^a \in \mathbb{R}^{n_i \times m_i}$, $u_i^a \in \mathbb{R}^{m_i}$ become

$$\begin{aligned} x_i(k+1) &= A_i x_i(k) + B_i u_i(k) + B_i^a u_i^a(k) + w_i(k), \\ y_i(k) &= C_i x_i(k) + v_i(k). \end{aligned} \tag{2}$$

In vector form, the dynamics of the interconnected system (2) read as

$$x(k+1) = Ax(k) + B^a u^a(k) + w(k) \tag{3}$$

where $x = \begin{bmatrix} x_1^{\mathsf{T}} & \dots & x_N^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}}$, $w = \begin{bmatrix} w_1^{\mathsf{T}} & \dots & w_N^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^n$, $u^a = \begin{bmatrix} (u_1^a)^{\mathsf{T}} & \dots & (u_N^a)^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^m$, $m = \sum_{i=1}^{N} m_i$, and

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{N1} & \cdots & A_{NN} \end{bmatrix}, B^a = \begin{bmatrix} B_1^a & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & B_N^a \end{bmatrix}.$$

We assume that a decentralized detector is composed of $N$ local detectors indexed by $i \in \{1 \dots N\}$, where each $i-$th local detector knows the local matrices $A_{ij}$ for all $j \in \{1 \dots N\}$, the statistics of $w_i$ and $v_i$, and the measurements $y_i(k)$ collected over discrete times in the interval $[1, T]$. Based on the local measurements, each $i-$th local detector decides against the presence of attacks in the $i-$th subsystem. The decision about the presence of attacks on the overall system is made by the decentralized detector based on the decisions reported by the local detectors.

Instead, a centralized detector knows the matrix $A$ and makes decision about attacks based on the measurements $\{y_{c,1}(k), \cdots, y_{c,N}(k)\}$ over the interval $[1, T]$, where

$$y_{c,i}(k) = \underbrace{\begin{bmatrix} 0 \cdots C_i \cdots 0 \end{bmatrix}}_{\triangleq C_{c,i}} x(k) + v_i(k) \tag{4}$$

From the equation (4) we notice that the local measurements obtained by the centralized detector from the sensors of the $i-$th subsystem sensor are dependent on the state $x(k)$ rather than just $x_i(k)$. Hence, we assume that the centralized detector has knowledge about the statistics of the noise vectors $w_i$ and $v_i$ for all $i \in \{1 \cdots N\}$.

The following procedures summarizes our detection mechanisms for detecting attacks on the interconnected system:

**1. Centralized detection scheme:**

(i) *Collect measurements:* For all $i \in \{1 \cdots N\}$, collect the measurements using (4) over the interval $[1, T]$:

$$\begin{aligned} Y_{c,i} &\triangleq \begin{bmatrix} y_{c,i}^{\mathsf{T}}(1) & y_{c,i}^{\mathsf{T}}(2) & \cdots & y_{c,i}^{\mathsf{T}}(T) \end{bmatrix}^{\mathsf{T}} \\ Y_c &\triangleq \begin{bmatrix} Y_{c,1}^{\mathsf{T}} & Y_{c,2}^{\mathsf{T}} & \cdots & Y_{c,N}^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} \end{aligned}$$

(ii) *Process measurements:* Process the measurements $Y_c \to \widetilde{Y}_c$ to perform detection (see below).

(iii) *Attack decision:* A suitable statistical test is conducted on $\widetilde{Y}_c$ to decide against attacks.

**2. Decentralized detection scheme:**

(i) *Collect measurements:* For every $i \in \{1 \cdots N\}$, collect measurements (2) over the interval $[1, T]$

$$Y_i \triangleq \begin{bmatrix} y_i^{\mathsf{T}}(1) & y_i^{\mathsf{T}}(2) & \cdots & y_i^{\mathsf{T}}(T) \end{bmatrix}^{\mathsf{T}}.$$

(ii) *Process measurements:* Process the measurements $Y_i \to \widetilde{Y}_i$ to perform local detection using a suitable statistical test (see below).

(iii) *Attack decision:* A decision is made based on pooling the decisions from the local detectors.

*Remark 1:* (**Similarity of measurements $Y_i$ and $Y_{c,i}$**) We note that the local and centralized detectors possess the same set of measurements, i.e., $Y_i = Y_{c,i}$ for all $i \in \{1, \cdots, N\}$ (this is evident from (2) and (4)). But the way the detectors process their measurements is different due to their different information about the interconnected system dynamics. □

## III. Detection framework

### A. Processing of measurements

Define the following observability and forced response matrices associated with the $i_{th}$ subsystem

$$\mathcal{O}_i = \begin{bmatrix} C_i A_{ii} \\ \vdots \\ C_i A_{ii}^T \end{bmatrix}, \mathcal{F}_i^{(u)} = \begin{bmatrix} C_i B_i & \dots & 0 \\ \vdots & \ddots & \vdots \\ C_i A_{ii}^{T-1} B_i & \dots & C_i B_i \end{bmatrix},$$

$$\mathcal{F}_i^{(a)} = \begin{bmatrix} C_i B_i^a & \dots & 0 \\ \vdots & \ddots & \vdots \\ C_i A_{ii}^{T-1} B_i^a & \dots & C_i B_i^a \end{bmatrix}, \mathcal{F}_i^{(w)} = \begin{bmatrix} C_i & \dots & 0 \\ \vdots & \ddots & \vdots \\ C_i A_{ii}^{T-1} & \dots & C_i \end{bmatrix}.$$

Analogously, define $\mathcal{O}_{c,i}$, $\mathcal{F}_{c,i}^{(a)}$, and $\mathcal{F}_{c,i}^{(w)}$ for the interconnected system (3) with respect to the local measurements $y_{c,i}$ (4). Then, $Y_{c,i}$ and $Y_i$ can be explicitly written as

$$\begin{aligned} Y_i &= \mathcal{O}_i x_i(0) + \mathcal{F}_i^{(u)} U_i + \mathcal{F}_i^{(a)} U_i^a + \mathcal{F}_i^{(w)} W_i + V_i, \\ Y_{c,i} &= \mathcal{O}_{c,i} x(0) + \mathcal{F}_{c,i}^{(a)} U^a + \mathcal{F}_{c,i}^{(w)} W + V_i, \end{aligned} \tag{5}$$

where

$$
\begin{aligned}
U_i &= \begin{bmatrix} u_i(0)^\mathsf{T} & \cdots & u_i(T-1)^\mathsf{T} \end{bmatrix}^\mathsf{T}, \\
U_i^a &= \begin{bmatrix} u_i^a(0)^\mathsf{T} & \cdots & u_i^a(T-1)^\mathsf{T} \end{bmatrix}^\mathsf{T}, \\
U^a &= \begin{bmatrix} u^a(0)^\mathsf{T} & \cdots & u^a(T-1)^\mathsf{T} \end{bmatrix}^\mathsf{T}, \\
W_i &= \begin{bmatrix} w_i(0)^\mathsf{T} & \cdots & w_i(T-1)^\mathsf{T} \end{bmatrix}^\mathsf{T}, \\
W &= \begin{bmatrix} w(0)^\mathsf{T} & \cdots & w(T-1)^\mathsf{T} \end{bmatrix}^\mathsf{T}, \text{ and} \\
V_i &= \begin{bmatrix} v_i(0)^\mathsf{T} & \cdots & v_i(T-1)^\mathsf{T} \end{bmatrix}^\mathsf{T}.
\end{aligned}
$$

The terms $x_i(0)$, $U_i$, and $U_i^a$ in (5) are unknown to $i-$th local detector, while the terms $x(0)$ and $U^a$ in (5) are unknown to a centralized detector. As the goal of the detectors is to detect if $U^a \neq 0$, we let the detectors pre-process the measurements to cancel the effect of the unknown parameters that do not carry information about the attack vector. Let $N_i^\mathsf{T}$ and $N_{c,i}^\mathsf{T}$ be a basis for the left null space of $\begin{bmatrix} \mathcal{O}_i & \mathcal{F}_i^{(u)} \end{bmatrix}$ and $\mathcal{O}_c$, respectively, and let

$$
\begin{aligned}
\widetilde{Y}_i &\triangleq N_i^\mathsf{T} Y_i = N_i^\mathsf{T} \left[ \mathcal{F}_i^{(a)} U_i^a + \mathcal{F}_i^{(w)} W_i + V_i \right], \text{ and} \\
\widetilde{Y}_{c,i} &\triangleq N_{c,i}^\mathsf{T} Y_{c,i} = N_{c,i}^\mathsf{T} \left[ \mathcal{F}_{c,i}^{(a)} U^a + \mathcal{F}_{c,i}^{(w)} W + V_i \right].
\end{aligned} \tag{6}
$$

Notice that, in the absence of attacks, the measurements (6) depend only on the system noise. Instead, in the presence of attacks, such measurements also depend on the attack vector, which may leave a statistical signature for the detectors[1].

Before presenting our testing procedures we express $\widetilde{Y}_{c,i}$ and $\widetilde{Y}_i$ in terms of $Y_c = \begin{bmatrix} Y_{c,1}^\mathsf{T} & Y_{c,2}^\mathsf{T} & \cdots & Y_{c,N}^\mathsf{T} \end{bmatrix}^\mathsf{T}$. We begin with the following proposition.

*Proposition 3.1:* Let

$$
\begin{aligned}
\mathcal{N}_i^L &= \left\{ z : z^\mathsf{T} \begin{bmatrix} \mathcal{O}_i & \mathcal{F}_i^{(u)} \end{bmatrix} = 0^\mathsf{T} \right\}, \text{ and} \\
\mathcal{N}_{c,i}^L &= \left\{ z : z^\mathsf{T} \mathcal{O}_{c,i} = 0^\mathsf{T} \right\}.
\end{aligned}
$$

Then, $\mathcal{N}_i^L \subseteq \mathcal{N}_{c,i}^L$.

*Proof:* Without loss of generality we can consider the case when $i = 1$. Consider the linear system defined in (2) without the attack and noise inputs i.e., $x(k+1) = Ax(k)$, and let $x(k) = \begin{bmatrix} x_1(k)^\mathsf{T} & \widetilde{x}_1(k)^\mathsf{T} \end{bmatrix}^\mathsf{T}$. Now, let

$$
A = \begin{bmatrix} A_{11} & B_1 \\ \widetilde{B}_1 & \widetilde{A}_{11}. \end{bmatrix}
$$

Then $x(k+1) = Ax(k)$ can be decomposed as,

$$
\begin{aligned}
x_1(k+1) &= A_{11} x_1(k) + B_1 \widetilde{x}_1(k), \\
\widetilde{x}_1(k+1) &= \widetilde{A}_{11} \widetilde{x}_1(k) + \widetilde{B}_1 x_1(k).
\end{aligned} \tag{7}
$$

Then by straightforward computation we have,

$$
\mathcal{O}_{c,1} x(0) = \mathcal{O}_1 x_1(0) + \mathcal{F}_1^{(u)} \begin{bmatrix} \widetilde{x}_1(0)^\mathsf{T} & \ldots & \widetilde{x}_1(T-1)^\mathsf{T} \end{bmatrix}^\mathsf{T}.
$$

Consider an arbitrary vector $z$ such that $z^\mathsf{T} \begin{bmatrix} \mathcal{O}_1 & \mathcal{F}_1^{(u)} \end{bmatrix} = 0^\mathsf{T}$ holds true, then from the above decomposition it can be seen that $z$ also satisfies the property that $z^\mathsf{T} \mathcal{O}_{c,1} = 0^\mathsf{T}$ ■

[1]If $\mathrm{Im} B_i^a) \subseteq \mathrm{Im}(B_i)$, then $N_i^\mathsf{T} \mathcal{F}_i^{(a)} = 0$, for all $i \in \{1, \ldots, N\}$. Thus for these type of attacks, the aforementioned measurement processing technique cannot retain the attack signature to perform detection.

Proposition 3.1 is due to the fact that the $i-$th local detector has more uncertainty about the system (3) knowledge than a centralized detector. Since $\mathcal{N}_i^L \subseteq \mathcal{N}_{c,i}^L$, we have $N_i = P_i N_{c,i}$, for all $i \in \{1, \ldots, N\}$, for some full row rank matrix $P_i$. By invoking the fact that $Y_i = Y_{c,i}$ (see Remark 1) it now follows that $\widetilde{Y}_i = (P_i N_{c,i})^\mathsf{T} Y_{c,i}$ and $\widetilde{Y}_{c,i} = N_{c,i}^T Y_{c,i}$. Finally, let $\widetilde{Y}_c = \begin{bmatrix} \widetilde{Y}_{c,1}^\mathsf{T}, \cdots, \widetilde{Y}_{c,N}^\mathsf{T} \end{bmatrix}^\mathsf{T}$ and notice that

$$
\widetilde{Y}_i = (M_i \widetilde{N}_i)^\mathsf{T} Y_c \text{ and } \widetilde{Y}_c = N_c^T Y_c, \tag{8}
$$

where $\widetilde{N}_i = P_i N_{c,i}$, $M_i^\mathsf{T} = \begin{bmatrix} 0 & \cdots & I_i & \cdots & 0 \end{bmatrix}$, $I_i$ is the identity matrix, and $N_c^\mathsf{T} = \mathrm{blkdiag} \begin{bmatrix} N_{c,1}^\mathsf{T} & \cdots & N_{c,N}^\mathsf{T} \end{bmatrix}$.

The above representation helps us to express the statistical properties of $\widetilde{Y}_i$ and $\widetilde{Y}_c$ in the terms of $Y_c$. Hence, for the purpose of inference, analogous to (5), we express $Y_c$ as

$$
Y_c = \mathcal{O}_c x(0) + \mathcal{F}_c^{(a)} U^a + \mathcal{F}_c^{(w)} W + V, \tag{9}
$$

where,

$$
\begin{aligned}
\mathcal{O}_c &= \begin{bmatrix} \mathcal{O}_{c,1}^\mathsf{T} & \cdots & \mathcal{O}_{c,N}^\mathsf{T} \end{bmatrix}^\mathsf{T}, \\
\mathcal{F}_c^{(a)} &= \begin{bmatrix} (\mathcal{F}^{(a)})_{c,1}^\mathsf{T} & \cdots & (\mathcal{F}^{(a)})_{c,N}^\mathsf{T} \end{bmatrix}^\mathsf{T}, \\
\mathcal{F}_c^{(w)} &= \begin{bmatrix} (\mathcal{F}^{(w)})_{c,1}^\mathsf{T} & \cdots & (\mathcal{F}^{(w)})_{c,N}^\mathsf{T} \end{bmatrix}^\mathsf{T}, \text{ and} \\
V &= \begin{bmatrix} V_1^\mathsf{T} & \cdots & V_N^\mathsf{T} \end{bmatrix}^\mathsf{T}.
\end{aligned}
$$

### *B. Statistical properties of the processed measurements*

Let $\beta$ and $\Sigma$ be the mean vector and the covariance matrix of $Y_c$, respectively. Then, from (9) we have

$$
\begin{aligned}
\beta &= \mathcal{O}_c x(0) + \mathcal{F}_c^{(a)} U^a \text{ and} \\
\Sigma &= (\mathcal{F}_c^{(w)})(I \otimes \Xi_c)(\mathcal{F}_c^{(w)})^\mathsf{T} + (I \otimes \Upsilon_c),
\end{aligned}
$$

where, $\Xi_c \triangleq \mathrm{blkdiag} \begin{pmatrix} \Xi_1 & \cdots & \Xi_N \end{pmatrix}$ and $\Upsilon_c \triangleq \mathrm{blkdiag} \begin{pmatrix} \Upsilon_1 & \cdots & \Upsilon_N \end{pmatrix}$. Moreover, $Y_c$ is normally distributed and so are $\widetilde{Y}_i$ and $\widetilde{Y}_c$. Thus we have

$$
\begin{aligned}
\beta_c &\triangleq \mathbb{E}[\widetilde{Y}_c] = N_c^\mathsf{T} \beta, \\
\beta_i &\triangleq \mathbb{E}[\widetilde{Y}_i] = (M_i \widetilde{N}_i)^\mathsf{T} \beta, \\
\Sigma_c &\triangleq \mathrm{Cov}[\widetilde{Y}_c] = N_c^\mathsf{T} \Sigma N_c, \text{ and} \\
\Sigma_i &\triangleq \mathrm{Cov}[\widetilde{Y}_i] = (M_i \widetilde{N}_i)^\mathsf{T} \Sigma (M_i \widetilde{N}_i).
\end{aligned} \tag{10}
$$

Observe that the vectors $\beta_i$ and $\beta_c$ depend on the attack vector $U^a$, while the covariance matrices $\Sigma_i$ and $\Sigma_c$ are independent of the attack. This motivates us to use a detection algorithm based on the mean vectors of the measurements.

### *C. Statistical hypothesis testing*

We assume that a centralized detector and all the local detectors employ the generalized likelihood ratio tests [15] for deciding against attacks. Let $H_0$ be the null hypothesis, where the system is not affected by attacks, and let $H_1$ be the alternative hypothesis. To decide which hypothesis is true, each detector performs the following test:

$$
\Lambda_i \triangleq \widetilde{Y}_i^\mathsf{T} \Sigma_i^{-1} \widetilde{Y}_i \underset{H_0}{\overset{H_1}{\gtrless}} \tau_i \text{ and } \Lambda \triangleq \widetilde{Y}_c^\mathsf{T} \Sigma_c^{-1} \widetilde{Y}_c \underset{H_0}{\overset{H_1}{\gtrless}} \tau_c, \tag{11}
$$

where $\tau_i$ and $\tau_c$ are suitable thresholds.

Let $P_i^F$ and $P_c^F$ be the false alarm probabilities of a $i-$th local detector and a centralized detector, respectively, then $P_c^F \triangleq \text{Prob}\left[\Lambda_c \geq \tau_c | H_0\right]$ and $\text{Prob}\left[\Lambda_i \geq \tau_i | H_0\right]$. Similarly, we define the probabilities of detection, that is the probability of deciding for $H_1$ when $H_1$ is true, as $P_c^D \triangleq \text{Prob}\left[\Lambda_c \geq \tau_c | H_1\right]$ and $P_i^D \triangleq \text{Prob}\left[\Lambda_i \geq \tau_i | H_1\right]$.

The following result [16, Theorem 3.3.3] helps to compute the detection and false alarm probabilities

(i) under $H_0$: for all $i \in \{1 \dots N\}$, $\Lambda_i \sim \chi^2(p_i)$, where $p_i = \text{rank}(\Sigma_i)$, and,
(ii) under $H_1$: for all $i \in \{1 \dots N\}$, $\Lambda_i \sim \chi^2(p_i, \lambda_i)$, where $p_i = \text{rank}(\Sigma_i)$ and $\lambda_i = \beta_i^\mathsf{T} \Sigma_i^{-1} \beta_i / 2$.

The parameters $p_i$ and $\lambda_i$ are referred to as degrees of freedom and non-centrality parameter, respectively. Analogously, for the centralized detector's test statistic we have $p_c = \text{rank}(\Sigma_c)$ and $\lambda_c = \beta_c^\mathsf{T} \Sigma_c^{-1} \beta_c / 2$.

*Remark 2:* ***(System theoretic interpretation of detection probability parameters)*** The following discussions are equivalently valid in the case of the centralized detector as well.

*(i) Degrees of freedom of $p_i$* : Intuitively, the degrees of freedom $p_i$ measure the amount of information possessed by a detector. The probability of detection is an increasing function of the detectors information as captured by $p_i$. Formally, for fixed $\tau_i$ and $\lambda_i$, we have $\lim_{p_i \to \infty} P_i^D(\tau_i, p_i, \lambda_i) = 1$ [17]. Further, it should be observed that $p_i$ depends on the rank of $N_i^\mathsf{T}$ in (6) through the matrix $\Sigma_i$, and that the rank of $N_i^\mathsf{T}$ is inversely proportional to the rank of the matrix $B_i$ in (2). Thus, the more the detector knows about the system dynamics, the smaller the rank of $B_i$, the larger the rank of $N_i^\mathsf{T}$, the value of $p_i$ and, ultimately, the better the detection performance of the decentralized detector.

*(ii) Non-centrality parameter $\lambda_i$* The non-centrality parameter $\lambda_i$ measures the intensity of the signature of the attack signal on the measurements. In fact, from (10) we have $2\lambda_i = \beta_i^\mathsf{T} \Sigma_i^{-1} \beta_i = (U^a)^\mathsf{T} \widetilde{\Sigma}_i^{-1} (U^a)$, where $\widetilde{\Sigma}_i^{-1} = \left[(M_i \widetilde{N}_i)^\mathsf{T} \mathcal{F}_i^{(2)}\right]^\mathsf{T} \Sigma_i^{-1} \left[(M_i \widetilde{N}_i)^\mathsf{T} \mathcal{F}_i^{(2)}\right]$. Although the above expression depends on $U^a$, by expanding $\widetilde{\Sigma}_i^{-1}$we can see that the result depends on $U_i^a$. It also follows that $P_i^D$ increases monotonically with $\lambda_i$, and it tends to $P_i^F$ as $\lambda_i \to 0$. □

## IV. Centralized and decentralized detection of attacks

In this section we aim to compare the performance of decentralized and centralized detectors. We assume that the decentralized detector decide on attacks if any of the local detector detects an attack (see Section II). We now define the false alarm and detection probabilities as

$$P_d^F \triangleq \text{Prob}\left[\text{ at least one local detector decide } H_1 \middle| H_0\right],$$
$$P_d^D \triangleq \text{Prob}\left[\text{ at least one local detector decide } H_1 \middle| H_1\right].$$

We state a lemma that relates above the probability measures with such quantities of the local detectors.

***Lemma* 4.1:** ***(Decentralized detector false alarm and detection probabilities)*** Let $P_d^F$ and $P_i^F$ be the false alarm probabilities of the decentralized and $i_{th}$ local detector, respectively. Let $P_d^D$ and $P_i^D$ be the detection probabilities of the decentralized and $i_{th}$ local detector. Then,

$$P_d^F = 1 - \prod_{i=1}^{N} \left(1 - P_i^F\right) \text{ and } P_d^D = 1 - \prod_{i=1}^{N} \left(1 - P_i^D\right).$$

*Proof:* Trivially follows from the statistical independence of $\widetilde{Y}_i$, for all $i \in \{1, \dots, N\}$. ■

The above lemma serves two purposes: (i) it allows us to compute the decentralized detector false alarm probability based on the local detector's false alarm probabilities, and (ii) it allows us to compare the detection probabilities of centralized and decentralized detectors via the local detector's detection probabilities. In what follows, to have a fair comparison between the detection probabilities of the decentralized and centralized detectors, we let their false alarm probabilities be equal, i.e., $P_c^F = P_d^F$. We now state a lemma that provides a relationship between a local and centralized detector's detection probability parameters.

***Lemma* 4.2:** ***(Non-centrality and degrees of freedom: centralized vs a local detector)*** For all $i \in \{1 \cdots N\}$ the following inequalities hold:

(i) $p_i < p_c$
(ii) $\lambda_i \leq \lambda_c$

*Proof:* See the Appendix. ■

The Lemma 4.2 states that a centralized detector has more information about the attack vector than the local detector, and this fact is measured by the degrees of freedom and the non-centrality parameter. We make the following assumption for computing the detection probabilities of the decentralized and the centralized detectors.

**A1:** The detectors collect sufficiently many measurements, that is, $T$ is sufficiently large.

In **A1**, because $T$ is large, the detection probabilities can be approximated using the standard normal distribution with appropriate mean and variance [17]. Now we state our first result, which provides conditions under which a decentralized detector outperforms a centralized detector.

***Theorem* 4.3:** ***(Sufficient conditions for $P_d^D \geq P_c^D$)*** Let $\tau_i$ and $\tau_c$ be the decision thresholds of a $i-$th local detector and a centralized detector, respectively, and let

$$\gamma_i = \sqrt{\frac{p_i + 2\lambda_i}{p_c + 2\lambda_c}}.$$

Under the assumption of **A1**, if $\tau_i - (\lambda_i + p_i) \leq \gamma_i \left[\tau_c - (\lambda_c + p_c)\right]$ for some $i \in \{1, \dots, N\}$, then $P_d^D \geq P_c^D$.

*Proof:* Let the above relationship hold true for any of the $i_{th}$ local detector. Under the assumption of **A1**, we now approximate detection probabilities using the CDF of standard normal, i.e., $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$. Define $\mu_i \triangleq \lambda_i + p_i$ and $\sigma_i^2 \triangleq \sqrt{2(p_i + 2\lambda_i)}$. Then we have

$$P_i^D \triangleq \text{Prob}\left[\frac{\Lambda_i - \mu_i}{\sqrt{\sigma_i^2}} \geq \frac{\tau_i - \mu_i}{\sqrt{\sigma_i^2}}\right] \approx 1 - \Phi\left(\frac{\tau_i - \mu_i}{\sqrt{\sigma_i^2}}\right).$$

Similarly, for the centralized detector, define $\mu_c \triangleq \lambda_c + p_c$ and $\sigma_c^2 \triangleq \sqrt{2(p_c + 2\lambda_c)}$. Then,

$$P_c^D \triangleq \text{Prob}\left[\frac{\Lambda_c - \mu_c}{\sqrt{\sigma_c^2}} \geq \frac{\tau_c - \mu_c}{\sqrt{\sigma_c^2}}\right] \approx 1 - \Phi\left(\frac{\tau_c - \mu_c}{\sqrt{\sigma_c^2}}\right).$$

From the hypothesis of the theorem and the monotonicity property of standard normal CDF we now conclude that

$$\frac{\tau_i - \mu_i}{\sqrt{\sigma_i^2}} \leq \frac{\tau_c - \mu_c}{\sqrt{\sigma_c^2}} \implies \Phi\left(\frac{\tau_i - \mu_i}{\sqrt{\sigma_i^2}}\right) \leq \Phi\left(\frac{\tau_c - \mu_c}{\sqrt{\sigma_c^2}}\right),$$

which further implies that $P_i^D > P_c^D$. By the definition of $P_d^D$ (see Lemma 4.1), for any $i \in \{1, \ldots, N\}$ we have

$$1 - P_d^D = \prod_{i=1}^{N} \left(1 - P_i^D\right) \leq (1 - P_i^D). \tag{12}$$

The inequality follows because for any $i \in \{1, \ldots, N\}$, $0 \leq (1 - P_j^D) \leq 1$. Since, we already noticed that $P_i^D \geq P_c^D$, by invoking the inequality (12) the result follows. ■

The condition in Theorem 4.3 guarantees better performance of a decentralized detector over a centralized detector in a subset of the region spanned by detection probability parameters. Intuitively, by the way we defined the test (11), if the test statistic and the decision threshold are far apart, i.e., $\lambda_c >> \tau_c$ ($\lambda_i >> \tau_i$), a centralized (local) detector easily decides attacks and, else otherwise. As $\Lambda_i$ and $\Lambda_c$ are random quantities, Theorem 4.5 places conditions on the expected value of $\Lambda_i$, where $\mathbb{E}[\Lambda_i] = \lambda_i + p_i$, such that a $i-$th local detector can easily detect attacks, and hence contributing to better performance of the decentralized detector.

***Lemma* 4.4: *(Upper bound on $P_d^D$)*** Let $\tau_i$, $p_i$, and $\lambda_i$ be the detection probability parameters of a $i-$th local detector. Also let, $S_d \sim \chi^2(p_{\text{sum}}, \lambda_{\text{sum}})$, where $p_{\text{sum}} = \sum_{i=1}^{N} p_i$, $\lambda_{\text{sum}} = \sum_{i=1}^{N} \lambda_i$, and $\tau_{\min} = \min_{1 \leq i \leq N} \tau_i$. Then,

$$P_d^D \leq \text{Prob}\left[S_d > \tau_{\min}\right],$$

*Proof:* See the Appendix. ■

The above lemma helps us in providing sufficient conditions, under which a centralized detector over a decentralized detector, similar to that of the conditions in Theorem 4.3.

***Theorem* 4.5: *(Sufficient conditions for $P_c^D \geq P_d^D$)*** Let $\tau_{\min} = \min_{1 \leq i \leq N} \tau_i$ and $\tau_c$ be the decision threshold of the centralized detector, respectively. Define $p_{\text{sum}} = \sum_{i=1}^{N} p_i$, $\lambda_{\text{sum}} = \sum_{i=1}^{N} \lambda_i$, and

$$\gamma_c = \sqrt{\frac{p_c + 2\lambda_c}{p_{\text{sum}} + 2\lambda_{\text{sum}}}},$$

Under the assumption of **A1**, if $\tau_c - (\lambda_c + p_c) \leq \gamma_c \left[\tau_{\min} - (\lambda_i + p_i)\right]$ for all $i \in \{1 \cdots N\}$, then $P_c^D \geq P_d^D$.

*Proof:* By invoking the upper bound on $P_d^D$ obtained in Lemma 4.4, the proof of the statement follows the similar lines of the argument as in Theorem 4.3. Hence, omitted. ■

Similar to the condition in Theorem 4.3, the sufficient condition in Theorem 4.5 restricts the deviation of the centralized detector's decision threshold from the expected value of $\Lambda_c$, where $\mathbb{E}[\Lambda_c] = \lambda_c + p_c$. It places conditions on the magnitude of mean deviation of the centralized detector, i.e., $\mathbb{E}[\Lambda_c] - \tau_c$ with respect to the mean deviations of all the local detectors, i.e., $\mathbb{E}[\Lambda_i] - \tau_i$. Since, the performance of a decentralized detector is influenced by all of the local detectors, for a centralized detector to perform better, it is intuitive to expect restrictions on the region spanned by the detection probability parameters of all the local detectors.

*Remark 3:* ***(Sub optimality of GLR test)*** It is interesting to note that a decentralized detector can outperform the centralized detector. For instance, consider the case $\lambda_i = \lambda_c$. Then, from the properties of non-central chi square distribution it, follows that even with fewer measurements a $i-$th local detector can outperform the centralized detector, and thus resulting in a superior performance of a decentralized detector. We note that this kind of behavior is because of the suboptimal nature of the GLR test [15] and the equal false alarm probability constraint we imposed for the detection. □

## V. Illustrative example

We consider an interconnected system composed of 3 subsystems, with equal local dynamics, i.e., $A_{11} = A_{22} = A_{33}$, where $A_{ii} \in \mathbb{R}^{4\times4}$ and $i \in \{1, 2, 3\}$. Instead, the interconnection matrices are different, i.e., $B_1 \neq B_2 \neq B_3$, where $B_i \in \mathbb{R}^{4\times8}$ and $i \in \{1, 2, 3\}$. In particular,

$$A_{ii} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \text{Col}(B_i) \in \text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}\right\},$$

where $\text{Col}(\cdot)$ denotes the columns of the matrix. The measurement horizon is $T = 10$, and the false alarm probability is $P_i^F = 0.05$ for all $i \in \{1, 2, 3\}$. Let $u_1^a$, $u_2^a$, and $u_3^a$ be the attacks on the subsystems 1, 2, and 3 respectively.

**Case 1 (unequal attacks)** Let $u = [0, 1, 1, 0]^\mathsf{T}$. For all times in the interval $[1, T]$, let $u_1^a = u$ and $u_2^a = u_3^a = 0.1u$. Notice that the impact of attack signal on the subsystem 1 is higher than the others. In Fig.1, we illustrate this fact through the mean value (proportional to attack intensity) test statistics. Further, as shown in Fig. 1, this attack vector leads to mean value of the test statistic that satisfies the conditions in Theorem 4.3 and hence, it follows that $P_d^D \geq P_c^D$.

**Case 2 (equal attacks)** Let $u_1^a = u_2^a = u_3^a = [0, 1, 1, 0]^\mathsf{T}$, for all times in the interval $[1, T]$. In Fig. 2, we showed only the performance of local detector-1 vs as centralized detector, as the local detector-1 has higher value of non-centrality parameter than the other local detector. As illustrated in Fig. 3, we note that this choice of attack vector lead to lesser mean deviation of centralized detector and thus Theorem 4.5 guarantees the inequality $P_c^D \geq P_d^D$ .

## VI. Conclusions

This paper studies the attack detection problem for the interconnected stochastic systems. We developed centralized and decentralized detection strategies for detecting attacks and, characterize the detection performances based on their

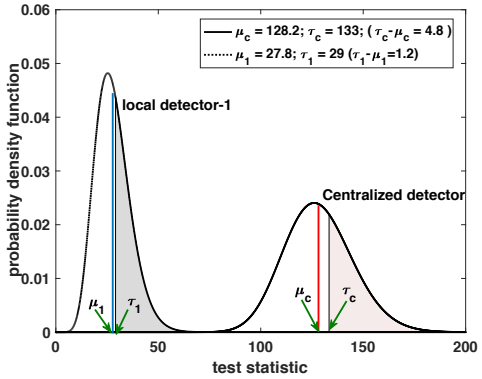


Fig. 1. The plot shows the densities of the test statistics of local detector-1 and a centralized detector, under the scenario of unequal attacks. The shaded area depicts the detection probabilities of the detectors. Notice that the mean deviation of local detector-1 ($\mu_1 - \tau_1 = 1.2$, see green ticks on the density of local detector-1) is less than the mean deviation of the centralized detector ($\mu_c - \tau_c = 4.8$, see green ticks on the density of centralized detector).

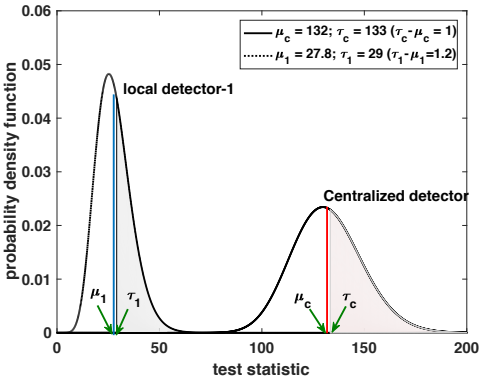


Fig. 2. The plot shows the densities of the test statistics of local detector-1 and a centralized detector in the case of equal attacks. The shaded area depicts the detection probabilities of the detectors. Notice that the mean deviation of the centralized detector ($\mu_c - \tau_c = 1$, see green ticks on the density of centralized detector) is less than the mean deviation of local detector-1 ($\mu_1 - \tau_1 = 1.2$, see green ticks on the density of local detector-1).

knowledge of the interconnected system. We derived conditions that guarantees better performance of a detector over another and, our results shows that the system dynamics, attack parameters, and the choice of statistical test influence the performance of the detectors. Finally, we demonstrated the trade-offs between the detection performance of the detectors with the aid of numerical examples.

## REFERENCES


[1] S. Massoud Amin and B. F. Wollenberg, "Toward a smart grid: power delivery for the 21st century," *IEEE Power and Energy Magazine*, vol. 3, no. 5, pp. 34–41, 2005.

[2] A. Su, H. Eichi, W. Zeng, and M. Y. Chow, "A survey on the electrification of transportation in a smart grid environment," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 1, pp. 1–10, 2012.

[3] S. M. Rinaldi, "Modeling and simulating critical infrastructures and their interdependencies," in *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*, 2004.

[4] A. A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *The 28th International Conf. on Distributed Comp. Systems Workshops*, June 2008, pp. 495–500.

[5] G. Wu, J. Sun, and C. J, "A survey on the security of cyber-physical systems," *Control Theory and Technology*, pp. 2–10, 2016.

[6] K. Ding, X. Ren, D. E. Quevedo, S. Dey, and L. Shi, "Dos attacks on remote state estimation with asymmetric information," *IEEE Transactions on Control of Network Systems*, pp. 1–1, 2018.

[7] Y. Li, L. Shi, and T. Chen, "Detection against linear deception attacks on multi-sensor remote state estimation," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 846–856, 2018.

[8] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1454–1467, 2014.

[9] C. Bai, F. Pasqualetti, and V. Gupta, "Security in stochastic control systems: Fundamental limitations and performance bounds," in *2015 American Control Conference (ACC)*, July 2015, pp. 195–200.

[10] J. F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 407–416, Feb 2003.

[11] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals and Systems*, pp. 167–182, 1988.

[12] F. Pasqualetti, F. Drfler, and F. Bullo, "A divide-and-conquer approach to distributed attack identification," in *2015 54th IEEE Conference on Decision and Control (CDC)*, 2015, pp. 5801–5807.

[13] F. Pasqualetti, F. Dörfler, and F. Bullo, "Cyber-physical security via geometric control: Distributed monitoring and malicious attacks," in *IEEE Conf. on Decision and Control*, USA, Dec. 2012, pp. 3418–3425.

[14] N. H and I. H, "Distributed detection of cyber attacks and faults for power systems," *IFAC Proceedings*, vol. 47, pp. 11 932 – 11 937, 2014.

[15] L. Wasserman, *All of Statistics*. Springer-Verlag New York, 2004.

[16] T. Anderson, *An Introduction to Multivariate Statistical Analysis*. Wiley, New York, 1958.

[17] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous univariate distributions, Volume 2*. Wiley & Sons, 1995.


## APPENDIX

*Proof of Lemma 4.2*:

(i) The first inequality is rather trivial, as the degrees of freedom, $p_i$ and $p_c$, are associated with dimensions of processed measurements, $\widetilde{Y}_i$ and $\widetilde{Y}_c$, and as $\widetilde{Y}_c$ includes processed measurements from all subsystems it's dimension should be at the least equal to the dimension of $\widetilde{Y}_i$.

(ii) Without loss of generality we consider the case when $i = 1$. From the section III-C and (10) we have,

$$\lambda_1 = \beta^\mathsf{T}(M_1\widetilde{N}_1)\left[(M_1\widetilde{N}_1)^\mathsf{T}\Sigma_c(M_1\widetilde{N}_1)\right]^{-1}(M_1\widetilde{N}_1)^\mathsf{T}\beta,$$
$$\lambda_c = \beta^\mathsf{T}N_c\Sigma_c^{-1}N_c^\mathsf{T}\beta.$$

Let $\Sigma_c = \begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{12}^\mathsf{T} & \Sigma_{22}\end{bmatrix}$ and substitute $M_1^T = \begin{bmatrix}I_1 & 0\end{bmatrix}$ in $\lambda_i$. Then by simple algebraic manipulations it follows that

$$\lambda_1 = \beta^\mathsf{T}N_c\begin{bmatrix}I & -\Sigma_{11}^{-1}\Sigma_{12}\\ 0 & I\end{bmatrix}\begin{bmatrix}\Sigma_{11}^{-1} & 0\\ 0^\mathsf{T} & 0\end{bmatrix}\begin{bmatrix}I & 0\\ -\Sigma_{12}^\mathsf{T}\Sigma_{11}^{-1} & I\end{bmatrix}N_c^\mathsf{T}\beta.$$

Instead, by Schur's complement, we can express $\Sigma_c^{-1}$ as

$$\Sigma_c^{-1} = \begin{bmatrix}I & -\Sigma_{11}^{-1}\Sigma_{12}\\ 0 & I\end{bmatrix}\begin{bmatrix}\Sigma_{11}^{-1} & 0\\ 0 & (\Sigma_c/\Sigma_{11})^{-1}\end{bmatrix}\begin{bmatrix}I & 0\\ -\Sigma_{12}^\mathsf{T}\Sigma_{11}^{-1} & I\end{bmatrix},$$

where $\Sigma_c/\Sigma_{11} = \Sigma_{22} - \Sigma_{12}^\mathsf{T}\Sigma_{11}^{-1}\Sigma_{12} > 0$. By substituting $\Sigma_c^{-1}$ in $\lambda_c$, the statement of the lemma follows. □

*Proof of Lemma 4.4:*

For all $i \in \{1, \ldots, N\}$, let $\mathcal{V}_i \triangleq \{\Lambda_i \geq \tau_i\}$ and $\mathcal{V} \triangleq \left\{S_d \triangleq \sum_{i=1}^N \Lambda_i \geq \tau_{\min}\right\}$ be the events associated with the test statistics. Then, we can see that $\bigcup_{i=1}^N \mathcal{V}_i \subseteq \mathcal{V}$, which implies that $\text{Prob}\left[\bigcup_{i=1}^N \mathcal{V}_i\right] \leq \text{Prob}\left[\mathcal{V}\right]$. Further, under the hypothesis $H_1$, we note that the term left to the inequality is $P_d^D$. Moreover, from the properties of the chi squared distribution [17], it follows that $S_d \sim \chi^2(p_{\text{sum}}, \lambda_{\text{sum}})$. □