On the Role of Information Sharing in the Security of Interconnected Systems

Rajasekhar Anguluri, Vaibhav Katewa, and Fabio Pasqualetti
University of California, Riverside, USA
E-mail: {ranguluri, vkatewa, fabiopas}@engr.ucr.edu

***Abstract*—In this paper we consider a security problem for interconnected dynamical systems, where each subsystem aims to detect local attacks. Each subsystem has knowledge of only its local dynamics and, therefore, the subsystems share information among themselves to aid local attack detection. We develop a technique that processes the local and shared measurements and detects attacks with provable guarantees. Interestingly, we show that for some instances of the attack, the detection performance deteriorates if the subsystems share their measurements. We provide an explanation for this counter-intuitive behavior and illustrate our results through a numerical example.**

I. Introduction

Dynamical systems are increasingly becoming more distributed, diverse, complex and integrated with cyber components. Usually, these systems are composed of multiple subsystems, which are interconnected among each other via physical, cyber and other types of couplings [1]. An example of such system is the smart city, which consists of subsystems such as the power grid, the transportation network, the water distribution network, and others. Although these subsystems are interconnected, it is usually difficult to directly measure the couplings and dependencies between them [1]. As a result, they are often operated independently without the knowledge of the other subsystems' models and dynamics.

Dynamical systems are usually vulnerable to cyber/physical attacks that can degrade their performance or may even render them inoperable [2]. There have been many recent studies on analyzing the effect of different types of attacks on dynamical systems and possible remedial strategies (see [3] and the references therein). A key component of these strategies is detection of attacks using the measurements generated by the system. Due to the autonomous nature of the subsystems, each subsystem is primarily concerned with detection of local attacks which affect its operation directly. However, local attack detection capability of each subsystem is limited due to the absence of knowledge of the dynamics and couplings of external subsystems. One way to mutually improve the detection performance is to share information and measurements among the subsystems. In this paper we develop a local attack detection strategy using the local measurements and the shared measurements from other subsystems. In some cases, these measurements may contain some confidential information about the subsystem and subsystem operators may not be willing to share them due to privacy concerns. We compare the detection performance of this case with the case when the subsystems share their measurements.

Related Work: Attack detection in dynamical systems has been studied both in centralized and decentralized settings. In [4], [5], centralized and decentralized monitor design was presented for deterministic attack detection and identification. Stochastic attack detection was studied in [6] and a comparison between centralized and decentralized attack detection schemes was presented in [7]. There have also been recent studies related to privacy in dynamical systems in the context of consensus, filtering and distributed optimization (see [8] and the references therein). These works develop additive noise-based privacy mechanisms, and characterize the trade-offs between the privacy level and the control performance. In [9], a privacy vs. cooperation trade-off for multi-agent systems was presented. The authors in [10] showed that the privacy mechanism can be used by an attacker to execute stealthy attacks in a centralized setting. In contrast to these studies, we present attack detection in an interconnected system in a decentralized manner.

Contributions: The contribution of this paper is two-fold. First, we propose a local attack detection scheme in an interconnected dynamical system which uses the local measurements and the measurements received from neighboring subsystems. Second, we present a comparison of the detection performance for the case in which the subsystems share the measurements against case where the subsystems do not share the measurements. Interestingly, our analysis shows that in some cases sharing no measurements can lead to better detection performance. We illustrate our theoretical results through numerical simulations.

Mathematical notation: $\mathrm{Tr}(\cdot)$, $\mathrm{Im}(\cdot)$, $\mathrm{Null}(\cdot)$ and $\mathrm{Rank}(\cdot)$ denote the trace, image, null space and rank of a matrix, respectively. $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{+}$ denote the transpose and pseudo-inverse of a matrix. A positive (semi)definite matrix A is denoted by $A > 0$ ($A \geq 0$). $\mathrm{diag}(A_1, A_2, \cdots, A_n)$ denotes a block diagonal matrix whose block diagonal elements are $A_1, A_2, \cdots, A_n$. The identity matrix is denoted by I (or I_n to denote its dimension explicitly). $\otimes$ denotes the Kronecker product. A zero mean Gaussian random variable y is denoted by $y \sim \mathcal{N}(0, \Sigma_y)$, where Σ_y denotes the covariance of y. The (central) chi-square distribution with q degrees of freedom is denoted by χ^2_q and the noncentral chi-square distribution with noncentrality parameter λ is denoted by $\chi^2_q(\lambda)$. For $x \geq 0$, let

This material is based upon work supported in part by ARO award 71603NSYIP, and in part by NSF awards ECCS1405330.

$\mathcal{Q}_q(x)$ and $\mathcal{Q}_q(x;\lambda)$ denote the right tail probabilities of a chi-square and noncentral chi-square distributions, respectively.

II. Problem Formulation

Let $\mathcal{S} \triangleq \{1,2,\cdots,N\}$ and $\mathcal{S}_{-i} \triangleq \{1,\cdots,i-1,i+1,\cdots,N\}$. We consider an interconnected discrete-time LTI dynamical system composed of N individual sub-subsystems. The dynamics of the subsystems are given by:

$$x_i(k+1) = A_i x_i(k) + A_{-i}x_{-i}(k) + w_i(k), \quad (1)$$
$$y_i(k) = C_i x_i(k) + v_i(k) \qquad i \in \mathcal{S}, \quad (2)$$

where $x_i \in \mathbb{R}^{n_i}$ and $y_i \in \mathbb{R}^{p_i}$ are the state and output/measurement of subsystem i, respectively. Let $n \triangleq \sum_{i=1}^{N} n_i$. Subsystem i is coupled with other subsystems through the interconnection term $A_{-i}x_{-i}(k)$, where $x_{-i} \triangleq [x_1^\mathsf{T},\cdots,x_{i-1}^\mathsf{T},x_{i+1}^\mathsf{T},\cdots,x_N^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^{n-n_i}$ denotes the aggregated states of all other subsystems. We refer to x_{-i} as the interconnection signal. Further, $w_i \in \mathbb{R}^{n_i}$ and $v_i \in \mathbb{R}^{p_i}$ are the process and measurement noise, respectively. We assume that $w_i(k) \sim \mathcal{N}(0,\Sigma_{w_i})$ and $v_i(k) \sim \mathcal{N}(0,\Sigma_{v_i})$ for all $k \geq 0$, with $\Sigma_{w_i} > 0$ and $\Sigma_{v_i} > 0$. The process and measurement noise are assumed to be white and independent for different subsystems. Finally, we assume that the initial state $x_i(0) \sim \mathcal{N}(0,\Sigma_{x_i(0)})$ is independent of $w_i(k)$ and $v_i(k)$ for all $k \geq 0$.

We make the following assumption regarding the interconnected system:

Assumption 1: Subsystem i has perfect knowledge of its dynamics, i.e., it knows (A_i, A_{-i}, C_i), the statistical properties of w_i, v_i and $x_i(0)$. However, it does not have the knowledge of the dynamics, states and the statistical properties of the noises of the other subsystems. □

Remark 1: (**Control input**) The dynamics in (1) typically includes a control input. However, since each subsystem has the knowledge of its control input, its effect can be easily included in the attack detection procedure. Therefore, for the ease of representation, we omit the control input. □

We consider the scenario where each subsystem can be under attack. We model the attacks as external linear additive input to the subsystems. Specifically, the dynamics of the subsystems under attack are given by

$$x_i(k+1) = A_i x_i(k) + A_{-i}x_{-i}(k) + B_i a_i(k) + w_i(k), \quad (3)$$
$$y_i(k) = C_i x_i(k) + v_i(k) \qquad i \in \mathcal{S}, \quad (4)$$

where $a_i \in \mathbb{R}^{r_i}$ is the attack input for subsystem i. We assume a_i to be a deterministic but unknown signal for all $i \in \mathcal{S}$.

Each subsystem i is equipped with an attack monitor whose goal is to detect the local attack a_i using the local measurements y_i. The detection procedure requires the knowledge of the statistical properties of y_i which depend on the interconnection signal x_{-i}. Since the subsystems do not have the knowledge of the interconnection signals (c.f. *Assumption 1*), they share their outputs among each other to aid the local detection of attacks (see Fig. 1). Let the parameters corresponding to the limited measurements of subsystem i be denoted by $\mathcal{P}_i \triangleq \{C_i, \Sigma_{v_i}\}$. We make the following assumption regarding measurement sharing:

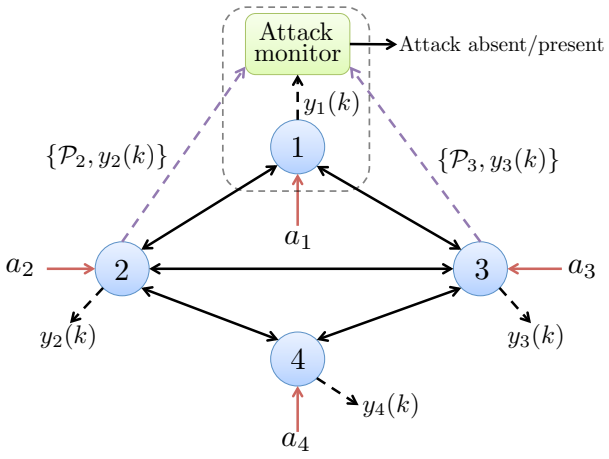

Fig. 1. An interconnected system consisting of $N = 4$ subsystems. The solid lines represent state coupling among the subsystems. For attack detection by Subsystem 1, its neighboring agents 2 and 3 communicate their output information to 1 (denoted by dashed lines). The attack monitor associated with Subsystem 1 uses the received information and the local measurements to detect attacks.

Assumption 2: Each subsystem i shares its measurements y_i in (4) and the parameters $\mathcal{P}_i$ with the other subsystems[1]. □

Under *Assumptions 1* and *2*, the goal of each subsystem i is to detect the local attack a_i using its local measurements y_i and the measurements $\{y_j\}_{j\in\mathcal{S}_{-i}}$ received from the other subsystems (see Fig. 1).

III. Local Attack Detection

In this section we present the local attack detection procedure of the subsystems and characterize their detection performance. For the ease or presentation, we describe the analysis for Subsystem 1 and remark that the procedure is same for other subsystems.

A. Measurement collection

We employ a batch detection scheme in which each subsystem collects the measurements for $k = 1,2,\cdots,T$, with $T > 0$ and performs detection based on the collected measurements. In this subsection, we model the collected local and shared measurements for Subsystem 1.

Local measurements: Let the time-aggregated local measurements, interconnection signals, attacks, process noises and measurement noises corresponding to Subsystem 1 be

[1]To be precise, this information sharing is required only between *neighboring* subsystems, i.e., between subsystems that are directly coupled with each other in (1).

respectively denoted by

$$\begin{aligned} y_L &\triangleq [y_1^\mathsf{T}(1), y_1^\mathsf{T}(2), \cdots, y_1^\mathsf{T}(T)]^\mathsf{T}, \\ x &\triangleq [x_{-1}^\mathsf{T}(0), x_{-1}^\mathsf{T}(1), \cdots, x_{-1}^\mathsf{T}(T-1)]^\mathsf{T}, \\ a &\triangleq [a_1^\mathsf{T}(0), a_1^\mathsf{T}(1), \cdots, a_1^\mathsf{T}(T-1)]^\mathsf{T}, \\ w &\triangleq [w_1^\mathsf{T}(0), w_1^\mathsf{T}(1), \cdots, w_1^\mathsf{T}(T-1)]^\mathsf{T}, \\ v &\triangleq [v_1^\mathsf{T}(1), v_1^\mathsf{T}(2), \cdots, v_1^\mathsf{T}(T)]^\mathsf{T}. \end{aligned}$$

By using (3) recursively and (4), the local measurements can be written as

$$y_L = Ox_1(0) + F_x x + F_a a + F_w w + v, \qquad \text{where,} \quad (5)$$

$$F_x \triangleq \begin{bmatrix} C_1A_{-1} & 0 & \cdots & 0 \\ C_1A_1A_{-1} & C_1A_{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C_1A_1^{T-1}A_{-1} & C_1A_1^{T-2}A_{-1} & \cdots & C_1A_{-1} \end{bmatrix},$$

$$F_a \triangleq \begin{bmatrix} C_1B_1^a & 0 & \cdots & 0 \\ C_1A_1B_1^a & C_1B_1^a & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C_1A_1^{T-1}B_1^a & C_1A_1^{T-2}B_1^a & \cdots & C_1B_1^a \end{bmatrix},$$

$$F_w \triangleq \begin{bmatrix} C_1 & 0 & \cdots & 0 \\ C_1A_1 & C_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C_1A_1^{T-1} & C_1A_1^{T-2} & \cdots & C_1 \end{bmatrix}, \; O \triangleq \begin{bmatrix} C_1A_1 \\ C_1A_1^2 \\ \vdots \\ C_1A_1^T \end{bmatrix}.$$

Note that $w \sim \mathcal{N}(0, \Sigma_w)$ and $v \sim \mathcal{N}(0, \Sigma_v)$ with

$$\Sigma_w = I_T \otimes \Sigma_{w_1} > 0 \quad \text{and} \quad \Sigma_v = I_T \otimes \Sigma_{v_1} > 0.$$

Let $v_L \triangleq Ox_1(0) + F_w w + v$ denote the effective local noise in the measurement equation (5). Using the fact that $\{x_1(0), w, v\}$ are independent, the overall local measurements of the subsystem are given by

$$\begin{aligned} &y_L = F_x x + F_a a + v_L, \quad \text{where} \qquad (6) \\ &v_L \sim \mathcal{N}(0, \Sigma_{v_L}),\ \Sigma_{v_L} = O\Sigma_{x_1(0)}O^\mathsf{T} + F_w\Sigma_w F_w^\mathsf{T} + \Sigma_v > 0. \end{aligned}$$

Shared measurements: Let $y_{-1}(k) \triangleq [y_2^\mathsf{T}(k), y_3^\mathsf{T}(k), \cdots y_N^\mathsf{T}(k)]^\mathsf{T}$ denote the measurements received by Subsystem 1 from all the other subsystems at time k. Further, let $v_{-1}(k) \triangleq [v_2^\mathsf{T}(k), v_3^\mathsf{T}(k), \cdots, v_N^\mathsf{T}(k)]^\mathsf{T}$. Then, from (4) we have

$$\begin{aligned} &y_{-1}(k) = Cx_{-1}(k) + v_{-1}(k), \quad \text{where} \qquad (7) \\ &\quad C \triangleq \text{diag}(C_2, C_3, \cdots, C_N), \quad \text{and} \\ &v_{-1}(k) \sim \mathcal{N}(0, \Sigma_{v_{-1}}),\ \Sigma_{v_{-1}} = \text{diag}(\Sigma_{v_2}, \cdots, \Sigma_{v_N}) > 0. \end{aligned}$$

Further, let the time-aggregated measurements received by the subsystem be denoted by $y_R \triangleq [y_{-1}^\mathsf{T}(0), y_{-1}^\mathsf{T}(1), \cdots, y_{-1}^\mathsf{T}(T-1)]^\mathsf{T}$, and let $v_R = [v_{-1}^\mathsf{T}(0), v_{-1}^\mathsf{T}(1), \cdots, v_{-1}^\mathsf{T}(T-1)]^\mathsf{T}$. Then, from (7), the overall measurements received by Subsystem 1 are given by

$$\begin{aligned} &y_R = Hx + v_R, \quad \text{where} \qquad (8) \\ &H \triangleq I_T \otimes C, \quad \text{and} \\ &v_R \sim \mathcal{N}(0, \Sigma_{v_R}) \quad \text{with} \quad \Sigma_{v_R} = I_T \otimes \Sigma_{v_{-1}} > 0. \end{aligned}$$

The goal of Subsystem 1 is to detect the local attack using the local and received measurements given by (6) and (8), respectively. Note that the subsystem knows (F_x, Σ_{v_L}) (c.f. *Assumption 1*) and (H, Σ_{v_R}) (c.f. *Assumption 2*). Further, the random variables v_L and v_R are independent, because they depend exclusively on the local and external subsystems' noises, respectively.

B. Measurement processing

Since Subsystem 1 does not have access to the interconnection signal x, it uses the received measurements to obtain an estimate of x. Note that x is stochastic since it depends on the process and measurement noise present in the external subsystems. However, since Subsystem 1 is oblivious to its statistics, it computes the estimate assuming x to be a deterministic signal.

The maximum likelihood (ML) estimate of x using the received measurements in (8) is given by (using Lemma A.1)

$$\begin{aligned} &\hat{x} = \tilde{H}^+ H^\mathsf{T}\Sigma_{v_R}^{-1} y_R + (I - \tilde{H}^+\tilde{H})d, \quad \text{where} \qquad (9) \\ &\tilde{H} \triangleq H^\mathsf{T}\Sigma_{v_R}^{-1} H, \end{aligned}$$

and d is any real vector of appropriate dimension. Note that if $\tilde{H}$ is not full column rank, then the estimate can lie anywhere in the null space of $\tilde{H}$ (shifted by $\tilde{H}^+ H^\mathsf{T}\Sigma_{v_R}^{-1} y$). Thus, the component of x that lies in the null space of $\tilde{H}$ cannot be estimated. We decompose x as:

$$\begin{aligned} x &= (I - \tilde{H}^+\tilde{H})x + \tilde{H}^+\tilde{H}x \\ &= (I - \tilde{H}^+\tilde{H})x + \tilde{H}^+ H^\mathsf{T}\Sigma_{v_R}^{-1} Hx \\ &\overset{(8)}{=} (I - \tilde{H}^+\tilde{H})x + \tilde{H}^+ H^\mathsf{T}\Sigma_{v_R}^{-1}(y_R - v_R). \qquad (10) \end{aligned}$$

Substituting x from (10) in (6), we get

$$\begin{aligned} y_L = {} & F_x(I - \tilde{H}^+\tilde{H})x + F_x\tilde{H}^+ H^\mathsf{T}\Sigma_{v_R}^{-1}(y_R - v_R) \\ & + F_a a + v_L. \qquad (11) \end{aligned}$$

Next, we process the local measurements in two steps. First, we subtract the known term $F_x\tilde{H}^+ H^\mathsf{T}\Sigma_{v_R}^{-1} y_R$. Second, we eliminate the component $(I - \tilde{H}^+\tilde{H})x$ (which cannot be estimated) by premultiplying (11) with a matrix M^T, where

$$\begin{aligned} &M = \text{Basis of Null}\left([F_x(I - \tilde{H}^+\tilde{H})]^\mathsf{T}\right), \\ &\Rightarrow M^\mathsf{T}F_x(I - \tilde{H}^+\tilde{H}) = 0. \qquad (12) \end{aligned}$$

Note that since the columns of M are basis vectors, M is full column rank. The processed measurements are given by

$$\begin{aligned} z &= M^\mathsf{T}(y_L - F_x\tilde{H}^+ H^\mathsf{T}\Sigma_{v_R}^{-1} y_R) \\ &\overset{(11),(12)}{=} M^\mathsf{T}F_a a + \underbrace{M^\mathsf{T}(v_L - F_x\tilde{H}^+ H^\mathsf{T}\Sigma_{v_R}^{-1} v_R)}_{\triangleq\, v_P}, \qquad (13) \end{aligned}$$

where $v_P \sim \mathcal{N}(0, \Sigma_{v_P})$. Since v_L and v_R are independent, we have

$$\Sigma_{v_P} = M^\mathsf{T}\left[\Sigma_{v_L} + F_x\tilde{H}^+ H^\mathsf{T}\Sigma_{v_R}^{-1}\Sigma_{v_R}\Sigma_{v_R}^{-\mathsf{T}} H(\tilde{H}^+)^\mathsf{T}F_x^\mathsf{T}\right]M \overset{\tilde{H}^\mathsf{T}=\tilde{H}}{=} M^\mathsf{T}\Sigma_{v_L}M + M^\mathsf{T}F_x\tilde{H}^+F_x^\mathsf{T}M \overset{(a)}{>} 0. \quad (14)$$

where (a) follows from the facts that M is full column rank and $\Sigma_{v_L} > 0$. The processed measurements z in (13) depend only on the local attack a and the Gaussian noise v_P, whose statistics is known to Subsystem 1, i.e. $z \sim \mathcal{N}(M^\mathsf{T}F_a a, \Sigma_{v_P})$. Thus, Subsystem 1 uses z to perform attack detection.

C. Statistical hypothesis testing

The goal of Subsystem 1 is to determine whether it is under attack or not (attack detection) using the processed measurements z in (13). We cast the attack detection problem as a binary hypothesis testing problem. Since Subsystem 1 does not know the attack a, we consider the following *composite* (simple vs. composite) testing problem

$$H_0: \quad a = 0 \quad \text{(Attack absent)} \qquad \text{vs}$$
$$H_1: \quad a \neq 0 \quad \text{(Attack present)}$$

We use the generalized likelihood ratio test (GLRT) criterion [11] for the above testing problem, which is given by

$$\frac{f(z|H_0)}{\sup_a f(z|H_1)} \underset{H_1}{\overset{H_0}{\gtrless}} \tau' \quad \text{where,} \quad (15)$$

$$f(z|H_0) = \frac{1}{\sqrt{2\pi|\Sigma_{v_P}|}} e^{-\frac{1}{2}z^\mathsf{T}\Sigma_{v_P}^{-1}z} \quad \text{and,}$$

$$f(z|H_1) = \frac{1}{\sqrt{2\pi|\Sigma_{v_P}|}} e^{-\frac{1}{2}(z-M^\mathsf{T}F_a a)^\mathsf{T}\Sigma_{v_P}^{-1}(z-M^\mathsf{T}F_a a)},$$

are the pdf of the multivariate Gaussian distribution of z under hypothesis H_0 and H_1, respectively, and τ' is a suitable threshold. Using the result in Lemma A.1 to compute the denominator in (15) and taking the logarithm, the test (15) can be equivalently written as

$$t(z) \triangleq z^\mathsf{T}\Sigma_{v_P}^{-1}M^\mathsf{T}F_a\tilde{M}^+F_a^\mathsf{T}M\Sigma_{v_P}^{-1}z \underset{H_0}{\overset{H_1}{\gtrless}} \tau, \quad (16)$$
$$\text{where} \quad \tilde{M} = F_a^\mathsf{T}M\Sigma_{v_P}^{-1}M^\mathsf{T}F_a.$$

Next, we derive the distribution of the test statistics $t(z)$ under both hypothesis.

Lemma 3.1: **(Distribution of test statistics)** The distribution of test statistics $t(z)$ in (16) is given by

$$t(z) \sim \chi_q^2 \quad \text{under } H_0, \quad (17)$$
$$t(z) \sim \chi_q^2(\lambda \triangleq a^\mathsf{T}F_a^\mathsf{T}M\Sigma_{v_P}^{-1}M^\mathsf{T}F_a a) \quad \text{under } H_1, \quad (18)$$

where $q = \text{Rank}(M^\mathsf{T}F_a)$.

Proof: See [12]. ∎

Remark 2: **(Interpretation of detection parameters** (q, λ)**)** The parameter q denotes the number of independent observations of the attack vector a in (13). The parameter λ can be interpreted as the signal to noise ratio (SNR) of the processed measurements in (13), where the signal of interest is the attack. □

Next, we characterize the performance of the test (16). Let the probability of false alarm and probability of detection for the test be respectively denoted by

$$P_F = \text{Prob}(t(z) > \tau | H_0) \quad \text{and,}$$
$$P_D = \text{Prob}(t(z) > \tau | H_1).$$

Inspired by the Neyman-Pearson test framework, we fix the desired size (P_F) of the test and determine the threshold τ which provides the desired size. Then, we use the threshold to perform the test and compute the detection probability. From Lemma 3.1, we have

$$\tau(q) = \mathcal{Q}_q^{-1}(P_F), \quad (19)$$
$$P_D(q, \lambda) = \mathcal{Q}_q(\tau(q); \lambda). \quad (20)$$

The detection probability is an indicator of the detection performance of Subsystem 1 and it depends on the detection parameters (q, λ). The next result characterizes this dependence.

Lemma 3.2: **(Dependence of detection performance on parameters** (q, λ)**)** For a fixed false alarm probability P_F, the detection probability $P_D(q, \lambda)$ is decreasing in q and increasing in λ.

Proof: It is a standard result that for a fixed q (and $\tau(q)$), the CDF $(= 1 - \mathcal{Q}_q(\tau(q); \lambda) = 1 - P_D(q, \lambda))$ of a noncentral chi-square random variable is decreasing in λ [12]. Thus, $P_D(q, \lambda)$ is increasing in λ.

Next, we have [12]

$$P_D(q, \lambda) = e^{-\lambda/2}\sum_{j=0}^{\infty}\frac{(\lambda/2)^j}{j!}\mathcal{Q}_{q+2j}(\tau(q)).$$

From [13, Corollary 3.1], it follows that $\mathcal{Q}_{q+2j}(\tau(q)) = \mathcal{Q}_{q+2j}(\mathcal{Q}_q^{-1}(P_F))$ is decreasing in q for all $j > 0$. Thus, $P_D(q, \lambda)$ is decreasing in q. ∎

Lemma 3.2 implies that for a fixed q, a higher SNR (λ) leads to a better detection performance, which is intuitive. However, for a fixed λ, an increase in the number of independent observations (q) results in degradation of the detection performance. This is due to the fact that the GLRT in (15) is not an uniformly most powerful (UMP) test for all values of the attack a.[2] This suboptimality is an inherent property of the GLRT in (15). It arises due to the composite nature of the test and the fact that the value of the attack vector a is not known to the attack monitor.

Next, we compare the above measurement sharing case with the case when the subsystems do not share information among each other, denoted as case 0. In the latter case, $\tilde{H} = 0$ and M is basis of Null(F_x). Let (q_0, λ_0) denote the detection parameters for case 0. Clearly, when the subsystems do not share information, both the SNR and number of observations at the detector decreases, i.e., $q_0 \leq q$ and $\lambda_0 \leq \lambda$. By Lemma 3.2, this implies that $P_D(q_0, \lambda_0)$ can be greater or smaller than

[2] A UMP test does not exist in this case [14].

$P_D(q, \lambda)$ depending on the detection parameters. Intuitively, if the decrease in P_D due to the decrease in the SNR[3] ($\lambda \to \lambda_0$) is larger than the increase in P_D due to the decrease in the number of measurements ($q \to q_0$), then the the detection performance decreases, and vice versa.

This is an interesting and counter-intuitive property and it implies that in certain cases sharing information can lead to worse detection performance. This phenomenon occurs because the GLRT for the considered hypothesis testing problem is a suboptimal test, as explained before. Next, we illustrate this behavior using a numerical example.

IV. Simulation Example

Consider an interconnected system with $N = 3$ subsystems with the following parameters:

$$A_1 = \frac{1}{3}\begin{bmatrix} -1 & -16 & 2 & -4 \\ 0 & -6 & 1 & -1 \\ 0 & 2 & 1 & 1 \\ 1 & 28 & -3 & 6 \end{bmatrix}, A_{12} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix},$$

$$A_{13} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}, B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 2 \end{bmatrix}, C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$A_{-1} = \begin{bmatrix} A_{12} & A_{13} \end{bmatrix}$, $\Sigma_{x_1(0)} = 0.2I_4$, $\Sigma_{w_1} = 0.1I_4$, $C_2 = I_3$ $C_3 = I_2$, $\Sigma_{v_1} = \Sigma_{v_2} = I_3$, $\Sigma_{v_3} = I_2$, $T = 6$.

We focus on the attack detection for Subsystem 1. The detection performance is completely characterized by P_F and the detection parameters (q, λ). Recalling (18), λ can vary between $[0, \infty)$ depending on the value of a. Thus, we present the results in this section in terms of λ.

First, we consider the case in which Subsystems 2 and 3 share their measurements with Subsystem 1 and denote it by case 1. In this case, we have $q = 11$. Fig. 2 shows the detection performance as a function of false alarm probability (typically known as the ROC curve) for different values of λ. We observe that for any given P_F, the detection performance becomes better as λ increases (c.f. Lemma 3.2).

Next, we compare the detection performance of case 0 (no measurement sharing) and case 1. The detection parameter $q_0 = 1$ for case 0 is less than $q = 11$ for case 1. Further, as stated previously, $\lambda \geq \lambda_0$. We choose $P_F = 0.05$ for both the cases for a fair comparison. Fig. 3 presents a comparison of the detection performance of cases 0 and 1. The blue circle region is characterized by pairs (λ_0, λ) for which $P_D(q, \lambda) \geq P_D(q_0, \lambda_0)$, and vice versa with red square region. We observe that case 1 performs better than case 0 if $\Delta_\lambda = \frac{\lambda - \lambda_0}{\lambda_0}$ is large, and vice versa. This shows that if the attack vector a is such that Δ_λ is small, then not sharing measurements can improve the detection performance. This counter-intuitive result is due to the suboptimality of the GLRT used to perform detection, as explained before.

[3]Note that the SNR depends upon the attack vector a (via (18)), which we do not know a-priori. Thus, depending on the actual attack value, the SNR can take any positive value.

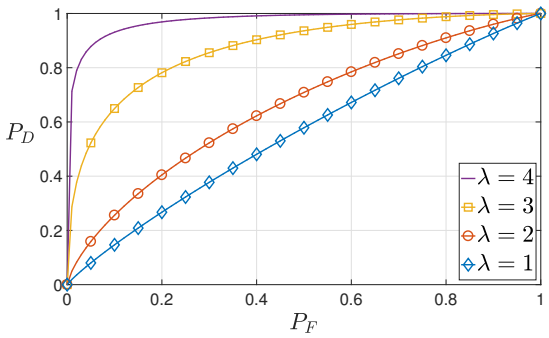

Fig. 2. Detection performance as a function of P_F for different values of λ.

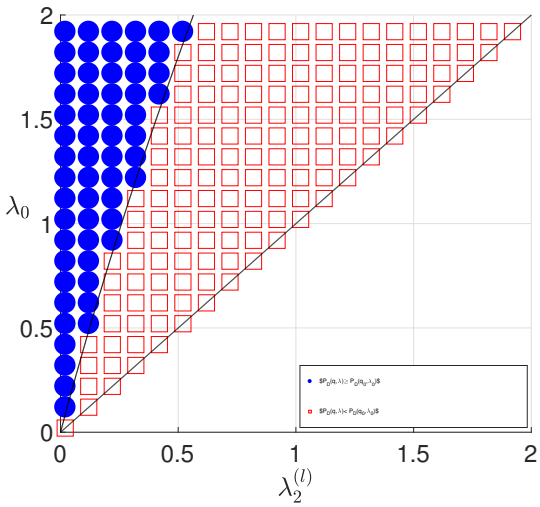

Fig. 3. Comparison between detection performance of case 0 and case 1. In the blue circle region, case 1 performs better than case 0, and vice versa in red square region. Since $\lambda \geq \lambda_0$, the white region is inadmissible.

V. Conclusion

We study an attack detection problem in interconnected dynamical systems wherein each subsystem is tasked with detection of local attacks without any knowledge of the dynamics of other subsystems and their interconnection signals. We present a measurement processing method which enables each subsystem to eliminate the unknown interconnection signal and perform attack detection using a composite hypothesis testing framework. The subsystems aid each other in attack detection by sharing measurements among each other. Interestingly, we show that in some cases, not sharing the measurements can improve the attack detection performance. This counter-intuitive behaviour is due to the sub-optimality of the composite nature of the considered hypothesis test.

REFERENCES

[1] S. M. Rinaldi, Peerenboom J. P., and T. K. Kelly. Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems Magazine*, 21(6):1125, 2001.

[2] A. A. Cardenas, S. Amin, and S. Sastry. Secure control: Towards survivable cyber-physical systems. In *International Conference on Distributed Computing Systems Workshops*, page 495500, Beijing, China, 2008.

[3] J. Giraldo, E. Sarkar, A. A. Cardenas, M. Maniatakos, and M. Kantarcioglu. Security and privacy in cyber-physical systems: A survey of surveys. *IEEE Design & Test*, 34(4):7–17, 2017.

[4] F. Pasqualetti, F. Dörfler, and F. Bullo. Attack detection and identification in cyber-physical systems. *IEEE Transactions on Automatic Control*, 58(11):2715–2729, 2013.

[5] F. Pasqualetti, F. Dörfler, and F. Bullo. A divide-and-conquer approach to distributed attack identification. In *IEEE Conf. on Decision and Control*, pages 5801–5807, Osaka, Japan, December 2015.

[6] Y. Mo, J. Hespanha, and B. Sinopoli. Resilient detection in the presence of integrity attacks. *IEEE Transactions on Signal Processing*, 62(1):3143, 2014.

[7] R. Anguluri, V. Katewa, and F. Pasqualetti. Attack detection in stochastic interconnected systems: Centralized vs decentralized detectors. In *IEEE Conf. on Decision and Control*, Miami, FL, 2018. Under review.

[8] J. Cortes, G. E. Dullerud, S. Han, J. Le Ny, S. Mitra, and G. J. Pappas. Differential privacy in control and network systems. In *IEEE Conf. on Decision and Control*, pages 4252–4272, Las Vegas, USA, 2016.

[9] V. Katewa, F. Pasqualetti, and V. Gupta. On privacy vs cooperation in multi-agent systems. *International Journal of Control*, pages 1–15, 2017.

[10] J. Giraldo, A. Cardenas, and M. Kantarcioglu. Security and privacy trade-offs in cps by leveraging inherent differential privacy. In *IEEE Conference on Control Technology and Applications*, pages 1313–1318, Hawaii, USA, 2017.

[11] L. Wasserman. *All of Statistics: A Concise Course in Statistical Inference*. Springer, 2004.

[12] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions, Volume 2*. Wiley-Interscience, 1995.

[13] E. Furman and R. Zitikis. A monotonicity property of the composition of regularized and inverted-regularized gamma functions with applications. *Journal of Mathematical Analysis and Applications*, 348(2):971–976, 2008.

[14] E. L. Lehmann and J. P. Romano. *Testing Statistical Hypotheses*. Springer-Verlag New York, 2005.

APPENDIX

Consider the following weighted least squares problem for a given y and $\Sigma > 0$:

$$\min_x \quad J(x) = (y - Hx)^\mathsf{T}\Sigma^{-1}(y - Hx). \tag{21}$$

Lemma A.1: The optimal solutions of the weighted least squares problem in (21) are given by

$$x^* = \tilde{H}^+ H^\mathsf{T}\Sigma^{-1}y + (I - \tilde{H}^+\tilde{H})d,$$

where $\tilde{H} = H^\mathsf{T}\Sigma^{-1}H$, and d is any real vector of appropriate dimension. Further, the optimal value of the cost is given by

$$J(x^*) = y^\mathsf{T}(\Sigma^{-1} - \Sigma^{-1}H\tilde{H}^+H^\mathsf{T}\Sigma^{-1})y.$$