UNIVERSITY OF CALIFORNIA
RIVERSIDE

Security of Interconnected Stochastic Dynamical Systems

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Mechanical Engineering

by

Rajasekhar Anguluri

December 2019

Dissertation Committee:

Dr. Fabio Pasqualetti, Chairperson
Dr. Wei Ren
Dr. Jun Sheng

The Dissertation of Rajasekhar Anguluri is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

It is an immense pleasure to thank all the people (except possibly on a set of measure zero[1]) who made this thesis possible and this phase of my life productive, responsible, and scintillating.

Foremost, I would like to express my sincere thanks to my advisor, Dr. Fabio Pasqualetti, for accepting me as a student, and providing time varying feedback, encouragement which is almost persistent, and a static high DC gain support. He is definitely the one who has suffered the most throughout my academic, but also emotional, development as a PhD student. In any case, he never gave up on me, even when things did not seem as gleaming as they do today! Thank you very much Fabio, not only for showing me how to do research, but teaching me how to be disciplined and organized, and also why it is not good to eat PRINGLES but to eat at GOODWINS!

To all my teachers, I owe a great debt of thanks. I was also fortunate that both my mother and father are in this profession. They have influenced every aspect of my view towards research and also the need to inspire younger generations. Special thanks to Dr. Ravi Kumar Jatoth, who mentored me during my undergraduate studies in India. If not for him, I would have never made an attempt to purse a PhD.

I am also grateful to my committee members, Dr. Wei Ren and Dr. Jun Sheng. I would also like to thank my collaborators, who helped me complete my research, encouraged me towards challenging problems, and also guided me from being stuck at various road blocks. Thank your Dr. Qi Zhu for allowing me to be a part of the cross-layer codesign project. Thank you Dr. Vijay Gupta for your inputs on the periodic coordinated attack

[1] but wait! what is a zero measure set in discrete space?

problem for cyber-physical security (my first CDC paper). Thank you Dr. Sandip Roy for suggesting me to work on the detection problem for network systems. I am especially grateful to Vaibhav Katewa. Vaibhav has not only contributed several important results contained in this thesis, but also provided valuable feedback for every problem I worked on and every paper I wrote in the last three years. With probability one, I can say that I would not have been able to finish my dissertation thesis without your help.

Many thanks to my friends and colleagues from the University of California, Riverside for all the wonderful moments spent together. In particular[2], thanks to my labmates throughout the past 5 years, Mika, Gianluca, Tommy, Akila, and Abed, for making the work place exciting, and to all those who visited the lab and left good memories to cherish, Chiara, Giacomo, Riccardo, Christian Cuba, and Adriano! Also, special thanks to Ravi Raj, Rakesh, Pavan, Subodh, Keerthana, Sri Lekha, Tushar, Devashree, and Rameswar for being wonderful friends and guiding me in many ventures.

Many thanks to Venki and Sid for being so nice and helpful to me all the time, and I am proud of your daughters Nivedita and Arpita, and thanks Nivedita for allowing me to mentor your high school math!

My heartfelt thanks to my parents for giving me all their support and love, and providing me the best possible education. Also, a gigantic hug to my wonderful twin sisters. Last, but not the least, I thank Priya (cute scientist), from the bottom of my heart!

Acknowledgment of previously published or accepted materials: The text of this dissertation, in part or in full, is a reprint of the material as appeared in four previously published papers and one paper that is currently under review. The co-author Dr. Fabio

[2] Of course, the list is not exhaustive.

Pasqualetti, listed in all the publications, directed and supervised the research which forms the basis for this dissertation[3]. The papers are as follows:

1. R. Anguluri, V. Katewa, and F. Pasqualetti, Centralized Vs Decentralized Detection of Attacks in Stochastic Interconnected Systems, IEEE TAC 2018 (accepted).

2. R. Anguluri, V. Katewa, and F. Pasqualetti, Network Theoretic Analysis of Maximum-a-Posteriori Detectors for Optimal Sensor Placement, Automatica, 2019 (submitted).

3. R. Anguluri, V. Katewa, and F. Pasqualetti, A Probabilistic Approach to Design Switching Attacks against Interconnected Systems, ACC 2019, Philadelphia, USA.

4. R. Anguluri, V. Katewa, and F. Pasqualetti, Attack Detection in Stochastic Interconnected Systems: Centralized vs Decentralized Detectors, CDC 2018, Miami, USA.

5. R. Anguluri, R. Dhal, S. Roy, and F. Pasqualetti, Network Invariants for Optimal Input Detection, ACC 2016, Boston, USA.

[3] I fear that errors lurk among the details collected here, and I want to correct them as soon as possible. Please send me an email (rangu003@ucr.edu) if you find any

To my friend Prudhvi Raj Kandepi, a victim of senseless act of gun violence.

ABSTRACT OF THE DISSERTATION

Security of Interconnected Stochastic Dynamical Systems

by

Rajasekhar Anguluri

Doctor of Philosophy, Graduate Program in Mechanical Engineering
University of California, Riverside, December 2019
Dr. Fabio Pasqualetti, Chairperson

Modern dynamical systems are large and inevitably comprise different subsystems that are often integrated with cyber (computation and communication) components. The applications of these systems are far reaching, ranging from power and water networks, to telecommunication and transportation systems etc. Recently, researchers and hackers have shown that these systems are vulnerable to attacks targeting their physical infrastructure or the signals exchanged between the physical and cyber layers. Given the interconnected nature of dynamical systems, and the fact that each subsystem usually has only partial knowledge or measurements of other interconnected units, the security question arises as to whether sophisticated attackers can hide their action to the individual subsystems while inducing system-wide critical perturbations.

This thesis addresses problems concerning security of interconnected systems that are subject to random (stochastic) disturbances. Our contribution is twofold. First, we investigate whether, and to what extent, coordination among different subsystems and knowledge of the global system dynamics is necessary to detect attacks in interconnected systems. We consider centralized and decentralized detectors, which differ primarily in their knowledge of the system model, and characterize the performance of the two detectors and show that, depending on the system and attack parameters, each of the detectors can outperform the other. Hence, it may be possible for the decentralized detector to outperform its centralized counterpart, despite having less information about the system dynamics. We provide an explanation for this counter-intuitive result and illustrate our results through simulations. Second, we study an attack design problem for interconnected systems where the attacker compromises a subsystem at each time, based on a pre-computed probabilistic rule. The goal of the attacker is to degrade the system performance, which is measured based on a quadratic function of the system state, while remaining undetected from a centralized detector. We show that selectively compromising different subsystems over time increases the severity of the attacks with respect to compromising a fixed subsystem at each time.

We study another related security problem for network systems, where changes in the statistical properties of an input driving certain network nodes has to be detected by remotely located sensors. To detect the changes, we associate a maximum-a-posteriori detector for a given set of sensors, and study its detection performance as function of the network topology, and the graphical distance between the input and sensor locations. We derive conditions under which the detection performance obtained when sensors are located on a network cut is not worse (resp. not better) than the performance obtained by measuring all nodes of the subnetwork induced by the cut and not containing the input nodes. Our results provide insights into the sensor placement from a detection-theoretic point of view.

Contents

List of Figures

Chapter 1

Introduction

Modern dynamical systems are increasingly becoming more distributed, diverse, complex and integrated with cyber (computation and communication) components. Typically, these systems are composed of multiple subsystems, which are interconnected via physical or virtual couplings among the states of subsystems. An example such system is the smart grid, which arises from the interconnection of smaller power systems at different geographical locations, and whose performance depends on other critical infrastructures including the transportation network and the water system. In contrast with legacy dynamical systems (which include traditional control systems), typically isolated from the outer world, the cyber and physical components of modern dynamical systems are interconnected via local data networks, and connected to the outer world via the Internet. This poses significant risks to personal privacy, economic security, and infrastructure. Further, the fact that each subsystem usually has only partial knowledge or measurements of other interconnected units, the security question arises as to whether sophisticated attackers can hide their action to the individual subsystems while inducing system-wide critical perturbations. Due to the vital role of these systems in everyday life, ensuring security of these systems, as of today, is one of the main focuses of the National Security [59].

Concerns regarding security of dynamical systems are not new, as several publications on systems fault diagnosis isolation, and recovery testify; see for example [4, 34]. Interconnected dynamical systems integrated with cyber and communication components, however, suffer from specific vulnerabilities which do not affect classical dynamical systems, and for which appropriate attack detection and isolation techniques need to be developed. For instance, relying on wireless communication networks to transmit measurements and control signals heavily increases the possibility of intentional and malignant attacks against various subsystems. Instead, traditional IT security methods, such as authentication and cryptography methods, are not always adequate for satisfactory protection of dynamical systems. The famous Stuxnet attack [26] is an excellent example to demonstrate the fact that IT security methods do not exploit the compatibility of the measurements with the underlying physical process and control mechanism, which are the primary objective of a protection scheme [10]. In fact, the supreme level of sophistication of Stuxnet attack prevented some well known anti-virus software to detect it initially [42]. Moreover, the existing IT security methods are also not effective against insider attacks carried out by authorized entities, for instance the Maroochy Water Breach case [73], and they also fail against attacks targeting directly the physical dynamics [19].

With security emerging as a major concern for dynamical systems, different modeling frameworks and protection schemes have been proposed for a variety of systems and attacks; see for example survey papers [69, 48, 30, 21]. Contrary to these existing works, in this thesis, we study the performance limitations of several attack detection schemes for both interconnected and networked dynamical systems. The main contributions of each chapter are summarized below:

In chapter 2, we consider a security problem for interconnected systems governed by linear, discrete, time-invariant, stochastic dynamics, where the objective is to detect exogenous attacks by processing the measurements at different locations. 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 interconnection signals that are exchanged among subsystems. 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 system and attack parameters, each of the detectors can outperform the other. In particular, it may be possible for the decentralized detector to outperform its centralized counterpart, despite having less information about the system dynamics, and this surprising property is due to the nature of the considered attack detection problem. To complement our results on the detection of attacks, we propose and solve an optimization problem to design attacks that maximally degrade the system performance while maintaining a pre-specified degree of detectability. Finally, we validate our findings via numerical studies on an electric power system.

In chapter 3, we study an attack design problem for interconnected systems where the attacker compromises a subsystem at each time, which is selected randomly based on a pre-computed probabilistic policy. The objective of the attacker is to degrade the system performance, which is measured based on a quadratic function of the system state, while remaining undetected from a centralized detector. First, we derive an explicit expression for the detection probability, analyze its properties, and compute an upper bound. Then, we use our upper bound to formulate and numerically solve an optimization problem for the computation of optimal attack strategies. Finally, we validate our results and show that our probabilistic attack strategy outperforms a constant attack strategy that compromises a fixed subsystem at each time.

In chapter 4, we consider the optimal sensor placement problem in a network system for detecting abrupt changes in the statistical properties of a stationary stochastic input driving certain network nodes, under the constraint that the sensor locations should be at least a specified distance apart from the input nodes. This specific detection problem is motivated by emerging monitoring needs for cyber-physical networks, where intrusions or abnormalities in cyber and human components may cause subtle changes in stochastic driving or input signals, and ultimately incur significant risk to the network operation. As it may be impractical or impossible to directly monitor these input signals, we exploit network structure to identify changes in the driving signals. We consider two scenarios: one in which the changes occurs in the mean of the input, and the other where the changes are allowed to happen in the covariance of the input. In both the scenarios, to detect the changes, we associate a maximum-a-posteriori (MAP) detector for a given set of sensors, and study its detection performance as function of the network topology, and the graphical distance between the input nodes and the sensors location. When the input and measurement noise follows a Gaussian distribution, we show that, as the number of measurements goes to infinity, the MAP detectors' performance can be studied using the input-output transfer function matrix gain of the network system. We derive conditions under which the detection performance obtained when sensors are located on a network cut is not worse (respectively not better) than the performance obtained by measuring all nodes of the subnetwork induced by the cut and not containing the input nodes. Our results provide structural insights into the sensor placement from a detection-theoretic viewpoint. Finally, we illustrate our findings through numerical examples.

We conclude the thesis with chapter 5, in which we also discuss some aspects for future research in the area of secure dynamical systems. We will also indicate possibilities of harnessing the power of Random Matrix Theory to develop suitable tools to analyze security in data-driven dynamical systems.

Chapter 2

Centralized Versus Decentralized Detection of Attacks in Stochastic Interconnected Systems

2.1 Introduction

In this chapter we investigate whether, and to what extent, coordination among different subsystems and knowledge of the global system dynamics is necessary to detect attacks in interconnected systems. In fact, while existing approaches for the detection of faults and attacks typically rely on a centralized detector [61, 49, 13], the use of local detectors would not only be computationally convenient, but it would also prevent the subsystems from disclosing private information about their plants. As a counterintuitive result, we will show that local and decentralized detectors can, in some cases, outperform a centralized detector, thus supporting the development of distributed and localized theories and tools for the security of cyber-physical systems.

Main contributions: This work features three main contributions. First, we propose centralized and decentralized schemes to detect unknown and unmeasurable sensor attacks in stochastic interconnected systems. Our detection schemes are based on the statistical decision theoretic framework that falls under the category of simple versus composite hypotheses testing. We characterize the probability of false alarm and the probability of detection for both detectors, as a function of the system and attack parameters. Second, we compare the performance of the centralized and decentralized detectors, and show that each detector can outperform the other for certain system and attack configurations. We discuss that this counterintuitive phenomenon is inherent with the simple versus composite nature of the considered attack detection problem, and provide numerical examples of this behavior. Third, we formulate and solve an optimization problem to design attacks against interconnected systems that maximally affect the system performance as measured by the mean square deviation of the state while remaining undetected by the centralized and decentralized detectors with a pre-selected probability. Finally, we validate our theoretical findings on the IEEE RTS-96 power system model.

Related Work: Centralized attack detectors have been the subject of extensive research in the last years [92, 27, 74, 46, 94, 55, 53, 14, 41], where the detector has complete knowledge of the system dynamics and all measurements. Furthermore, these studies use techniques from various disciplines including game theory, information theory, fault detection and signal processing, and have a wide variety of applications [49]. Instead, decentralized attack detectors, where each local detector decides on attacks based on partial information and measurements about the system, and local detectors cooperate to improve their detection capabilities, have received only limited and recent attention [22, 58, 40, 95, 33]. Decentralized detection schemes have also been studied for fault detection and isolation (FDI). In such schemes, multiple local detectors make inferences about either the global or local process, and transmit their local decisions to a central entity, which uses appropriate fusion rules to make the global decision[78, 84, 11, 3, 66]. Methods to improve the detection performance by exchanging information among the local detectors have also been proposed [90, 71, 29]. These decentralized algorithms are typically complex [61], their effectiveness in detecting unknown and unmeasurable attacks is difficult to characterize, and their performance is believed to be inferior when compared to their centralized counterparts. To the best of our knowledge, a rigorous comparison of centralized and decentralized attack detection schemes is still lacking, which prevents us from assessing whether, and to what extent, decentralized and distributed schemes should be employed for attack detection.

2.2 Problem setup and preliminary notions

We consider an interconnected system with N subsystems, where each subsystem obeys the discrete-time linear dynamics

$$\begin{aligned} x_i(k+1) &= A_{ii}x_i(k) + B_iu_i(k) + w_i(k), \\ y_i(k) &= C_ix_i(k) + v_i(k), \end{aligned} \tag{2.1}$$

with $i \in \{1, \dots, N\}$. In the above equation, the vectors $x_i \in \mathbb{R}^{n_i}$ and $y_i \in \mathbb{R}^{r_i}$ are the state and measurement of the $i-$th subsystem, respectively. The process noise $w_i(k) \sim \mathcal{N}(0, \Sigma_{w_i})$

and the measurement noise $v_i(k) \sim \mathcal{N}(0, \Sigma_{v_i})$ are independent stochastic processes, and w_i is assumed to be independent of v_i, for all $k \geq 0$. Further, the noise vectors across different subsystems are assumed to be independent at all times. The $i-$th subsystem is coupled with the other subsystems through the term $B_i u_i$, which takes the form

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

The input $B_i u_i = \sum_{j \neq i}^{N} A_{ij} x_j$ represents the cumulative effect of subsystems j on subsystem i. Hence, we refer to B_i as to the interconnection matrix, and to u_i as to the interconnection signal, respectively. We allow for the presence of attacks compromising the dynamics of the subsystems, and model such attacks as exogenous unknown inputs. In particular, the dynamics of the $i-$th subsystem under the attack u_i^a with matrix B_i^a read as

$$x_i(k+1) = A_{ii} x_i(k) + B_i u_i(k) + B_i^a u_i^a(k) + w_i(k), \tag{2.2}$$

where $u_i^a \in \mathbb{R}^{m_i}$. In vector form, the dynamics of overall system under attack read as

$$\begin{aligned} x(k+1) &= Ax(k) + B^a u^a(k) + w(k), \\ y(k) &= Cx(k) + v(k), \end{aligned} \tag{2.3}$$

where $\phi = \begin{bmatrix} \phi_1^\mathsf{T} & \dots & \phi_N^\mathsf{T} \end{bmatrix}$, with ϕ standing for $x \in \mathbb{R}^n$, $w \in \mathbb{R}^n$, $u^a \in \mathbb{R}^m$, $y \in \mathbb{R}^r$, $v \in \mathbb{R}^r$, $n = \sum_{i=1}^N n_i$, $m = \sum_{i=1}^N m_i$, and $r = \sum_i^N r_i$. Moreover, as the components of the vectors w and v are independent and Gaussian, $w \sim \mathcal{N}(0, \Sigma_w)$ and $v \sim \mathcal{N}(0, \Sigma_v)$, respectively, where

$\Sigma_w = \text{blkdiag}\left(\Sigma_{w_1}, \dots, \Sigma_{w_N}\right)$ and $\Sigma_v = \text{blkdiag}\left(\Sigma_{v_1}, \dots, \Sigma_{v_N}\right)$. Further,

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

and $C = \text{blkdiag}\left(C_1, \dots, C_N\right)$.

We assume that each subsystem is equipped with a *local detector*, which uses the local measurements and knowledge of the local dynamics to detect the presence of local attacks. In particular, the $i-$th local detector has access to the measurements y_i in (2.1), knows the matrices A_{ii}, B_i, and C_i, and the statistical properties of the noise vectors w_i and v_i. Yet, the $i-$th local detector does not know or measure the interconnection input u_i, and the attack parameters B_i^a and u_i^a. Based on this information, the $i-$th local detector aims to detect whether $B_i^a u_i^a \neq 0$. The decisions of the local detectors are then processed by a *decentralized detector*, which aims to detect the presence of attacks against the whole interconnected system based on the local decisions. Finally, we assume the presence of a *centralized detector*, which has access to the measurements y in (2.3), and knows the matrix A and the statistical properties of the overall noise vectors w and v. Similarly to the local detectors, the centralized detector does not know or measure the attack parameters B^a and u^a, and aims to detect whether $B^a u^a \neq 0$. We postpone description of our detectors to Section 2.3. To conclude this section, note that the decentralized and centralized detectors have access to the same measurements. Yet, these detectors differ in their knowledge of the system dynamics, which determines their performance as explained in Section 2.4.

Remark 1 ***(Control input and initial state)*** *The system setup in* (2.2) *and* (2.3) *typ-*

ically includes a control input. However, assuming that each subsystem knows its control input, it can be omitted without affecting generality. Further, as the detectors do not have information about the initial state, we assume without loss of generality, that the initial state is deterministic and unknown to the detectors. □

2.3 Local, decentralized, and centralized detectors

In this section we formally describe our local, decentralized, and centralized detectors, and characterize their performance as a function of the available measurements and knowledge of the system dynamics. To this aim, let $T > 0$ be an arbitrary time horizon and define the vectors

$$Y_i = \begin{bmatrix} y_i^{\mathsf{T}}(1) & y_i^{\mathsf{T}}(2) & \cdots & y_i^{\mathsf{T}}(T) \end{bmatrix}^{\mathsf{T}}, \tag{2.4}$$

which contains the measurements available to the $i-$th detector, and

$$Y_c = \begin{bmatrix} y^{\mathsf{T}}(1) & y^{\mathsf{T}}(2) & \cdots & y^{\mathsf{T}}(T) \end{bmatrix}^{\mathsf{T}}, \tag{2.5}$$

which contains the measurements available to the centralized detector. Both the local and centralized detectors perform the following three operations in order:

1. Collect measurements as in (2.4) and (2.5), respectively;

2. Process measurements to filter unknown variables; and

3. Perform statistical hypotheses testing to detect attacks (locally or globally) using the processed measurements.

The decisions of the local detectors are then used by the decentralized detector, which triggers an alarm if any of the local detectors does so. We next characterize how the detectors process their measurements and perform attack detection via statistical hypothesis testing.

2.3.1 Processing of measurements

The measurements (2.4) and (2.5) depend on parameters that are unknown to the detectors, namely, the system initial state and the interconnection signal (although the process and measurement noises are also unknown, the detectors know their statistical properties). Thus, to test for the presence of attacks, the detectors first process the measurement vectors to eliminate their dependency on the unknown parameters. To do so, using equations (2.1) and (2.2), define the observability matrix and the attack, interconnection, and noise forced response matrices of the $i-$th subsystem as

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

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

Analogously, for the system model (2.3) define the matrices $\mathcal{O}_c$, $\mathcal{F}_a^{(w)}$, and $\mathcal{F}_c^{(w)}$, which are constructed as above by replacing A_i, B_i^a, and C_i with A, B^a, and C, respectively. The measurements (2.4) and (2.5) can be written as follows:

$$Y_i = \mathcal{O}_i x_i(0) + \mathcal{F}_i^{(u)} U_i + \mathcal{F}_i^{(a)} U_i^a + \mathcal{F}_i^{(w)} W_i + V_i, \tag{2.6}$$

$$Y_c = \mathcal{O}_c x(0) + \mathcal{F}_c^{(a)} U^a + \mathcal{F}_c^{(w)} W + V, \tag{2.7}$$

where $U_i = \begin{bmatrix} u_i^\mathsf{T}(0) & u_i^\mathsf{T}(1) & \cdots & u_i^\mathsf{T}(T-1) \end{bmatrix}^\mathsf{T}$. The vectors U_i^a, U^a, W_i and W are the time aggregated signals of u_i^a, u^a, w_i, and w, respectively, and are defined similarly to U_i. Instead, $V_i = \begin{bmatrix} v_i^\mathsf{T}(1) & v_i^\mathsf{T}(2) & \cdots & v_i^\mathsf{T}(T) \end{bmatrix}^\mathsf{T}$, and V is defined similarly to V_i. To eliminate the dependency from the unknown variables, let N_i and N_c be bases of the left null spaces of the matrices $\begin{bmatrix} \mathcal{O}_i & \mathcal{F}_i^{(u)} \end{bmatrix}$ and $\mathcal{O}_c$, respectively, and define the processed measurements as

$$\begin{aligned} \widetilde{Y}_i &= N_i Y_i = N_i \left[\mathcal{F}_i^{(a)} U_i^a + \mathcal{F}_i^{(w)} W_i + V_i \right], \\ \widetilde{Y}_c &= N_c Y_c = N_c \left[\mathcal{F}_c^{(a)} U^a + \mathcal{F}_c^{(w)} W + V \right], \end{aligned} \tag{2.8}$$

where the expressions for $\widetilde{Y}_i$ and $\widetilde{Y}_c$ follows from (2.6) and (2.7). Notice that, in the absence of attacks ($U^a = 0$), the measurements $\widetilde{Y}_i$ and $\widetilde{Y}_c$ depend only on the system noise. Instead, in the presence of attacks, such measurements depend on the attack vector, which may leave a signature for the detectors.[1] We now characterize the statistical properties of $\widetilde{Y}_i$ and $\widetilde{Y}_c$.

Lemma 2 ***(Statistical properties of the processed measurements)*** *The processed measurements $\widetilde{Y}_i$ and $\widetilde{Y}_c$ satisfy*

$$\begin{aligned} &\widetilde{Y}_i \sim \mathcal{N}(\beta_i, \Sigma_i), \textit{ for all } i \in \{1, \dots, N\}, \textit{ and} \\ &\widetilde{Y}_c \sim \mathcal{N}(\beta_c, \Sigma_c), \end{aligned} \tag{2.9}$$

[1] If $\operatorname{Im}(B_i^a) \subseteq \operatorname{Im}(B_i)$, then $N_i \mathcal{F}_i^{(a)} = 0$ and the processed measurements do not depend on the attack. Thus, our local detection technique can only be successful against attacks that do not satisfy this condition.

where

$$\begin{aligned}
\beta_i &= N_i \mathcal{F}_i^{(a)} U_i^a, \\
\beta_c &= N_c \mathcal{F}_c^{(a)} U^a, \\
\Sigma_i &= N_i \left[\left(\mathcal{F}_i^{(w)} \right) (I_T \otimes \Sigma_{w_i}) \left(\mathcal{F}_i^{(w)} \right)^{\mathsf{T}} + (I_T \otimes \Sigma_{v_i}) \right] N_i^{\mathsf{T}}, \\
\Sigma_c &= N_c \left[\left(\mathcal{F}_c^{(w)} \right) (I_T \otimes \Sigma_{w}) \left(\mathcal{F}_c^{(w)} \right)^{\mathsf{T}} + (I_T \otimes \Sigma_{v}) \right] N_c^{\mathsf{T}}.
\end{aligned} \tag{2.10}$$

A proof of Lemma 2 is postponed to the Appendix. From Lemma 2, the mean vectors β_i and β_c depend on the attack vector, while the covariance matrices Σ_i and Σ_c are independent of the attack. This observation motivates us to develop a detection mechanism based on the mean of the processed measurements, rather the covariance matrices.

2.3.2 Statistical hypothesis testing framework

In this section we detail our attack detection mechanism, which we assume to be the same for all local and centralized detectors, and we characterize its false alarm and detection probabilities. We start by analyzing the test procedure of the $i-$th local detector. Let H_0 be the null hypothesis, where $\beta_i = 0$ and the system is not under attack, and let H_1 be the alternative hypothesis, where $\beta_i \neq 0$ and the system is under attack. To decide which hypothesis is true, or equivalently whether the mean value of the processed measurements is zero, we resort to the generalized log-likelihood ratio test (GLRT):

$$\Lambda_i \triangleq \widetilde{Y}_i^{\mathsf{T}} \Sigma_i^{-1} \widetilde{Y}_i \underset{H_0}{\overset{H_1}{\gtrless}} \tau_i, \tag{2.11}$$

where the threshold $\tau_i \geq 0$ is selected based on the desired false alarm probability of the test (2.11) [65]. For a statistical hypothesis testing problem, the false alarm probability equals the probability of deciding for H_1 when H_0 is true, while the detection probability equals the probability of deciding for H_1 when H_1 is true. While the former is used for tuning the threshold, the latter is used for measuring the performance of the test. Formally, the false alarm and detection probabilities of (2.11) are the probabilities that are conditioned on the hypothesis H_0 and H_1, respectively, and are symbolically denoted as

$$P_i^F = \Pr\left[\Lambda_i \geq \tau_i | H_0\right] \text{ and } P_i^D = \Pr\left[\Lambda_i \geq \tau_i | H_1\right].$$

Similarly, the centralized detector test is defined as

$$\Lambda_c \triangleq \widetilde{Y}_c^{\mathsf{T}} \Sigma_c^{-1} \widetilde{Y}_c \underset{H_0}{\overset{H_1}{\gtrless}} \tau_c, \tag{2.12}$$

where $\tau_c \geq 0$ is a preselected threshold, and its false alarm and detection probabilities are denoted as P_c^F and P_c^D. We next characterize the false alarm and detection probabilities of the detectors with respect to the system and attack parameters.

Lemma 3 ***(False alarm and detection probabilities of local and centralized detectors)*** *The false alarm and the detection probabilities of the tests* (2.11) *and* (2.12) *are, respectively,*

$$\begin{aligned} &P_i^F = Q(\tau_i; p_i, 0), \; P_i^D = Q(\tau_i; p_i, \lambda_i), \; \textit{and} \\ &P_c^F = Q(\tau_c; p_c, 0), \; P_c^D = Q(\tau_c; p_c, \lambda_c), \end{aligned} \tag{2.13}$$

where

$$\begin{aligned} &p_i = \operatorname{Rank}(\Sigma_i), \; p_c = \operatorname{Rank}(\Sigma_c), \\ &\lambda_i = (U_i^a)^{\mathsf{T}} M_i (U_i^a), \; \lambda_c = (U^a)^{\mathsf{T}} M_c (U^a), \end{aligned} \tag{2.14}$$

and

$$\begin{aligned} M_i &= \left(N_i \mathcal{F}_i^{(a)}\right)^{\mathsf{T}} \Sigma_i^{-1} \left(N_i \mathcal{F}_i^{(a)}\right), \\ M_c &= \left(N_c \mathcal{F}_c^{(a)}\right)^{\mathsf{T}} \Sigma_c^{-1} \left(N_c \mathcal{F}_c^{(a)}\right). \end{aligned} \tag{2.15}$$

Lemma 3, whose proof is postponed to the Appendix, allows us to compute the false alarm and detection probabilities of the detectors using the decision thresholds, the system parameters, and the attack vector. Moreover, for fixed P_i^F and P_c^F, the detection thresholds are computed as $\tau_c = Q^{-1}(P_c^F; p_c, 0)$ and $\tau_i = Q^{-1}(P_i^F; p_i, 0)$, where $Q^{-1}(\cdot)$ is the inverse of the complementary Cumulative Distribution Functions (CDF) that is associated with a central chi-squared distribution. The parameters p_i, p_c and λ_i, λ_c in Lemma 3 are referred to as *degrees of freedom* and *non-centrality* parameters of the detectors.

Remark 4 ***(System theoretic interpretation of detection probability parameters)*** *The degrees of freedom and the non-centrality parameters quantify the knowledge of the detectors about the system dynamics and the energy of the attack signal contained in the processed measurements. In particular:*

(Degrees of freedom p_i) The detection probability and the false alarm probability are both increasing functions of the degrees of freedom p_i, because the Q function in (2.13) *is an increasing function of p_i. Thus, increasing p_i by increasing the number of sensors or the horizon T, does not necessarily lead to an improvement of the detector performance.*

(Non-centrality parameter λ_i) The non-centrality parameter λ_i measures the energy of the attack signal contained in the processed measurements. In the literature of communication and signal processing, the non-centrality parameter is often referred to as signal to noise

ratio (SNR) [65]. For fixed τ_i and p_i, the detection probability increases monotonically with λ_i, and approaches the false alarm probability as λ_i tends to zero.

(Decision threshold τ_i) For fixed λ_i and p_i, the probability of detection and the false alarm probability are monotonically decreasing functions of the detection threshold τ_i. This is due to the fact that the complementary CDFs, which define the false alarm and detection probabilities, are decreasing functions of τ_i. As we show later, because of the contrasting behaviors of the false alarm and detection probabilities with respect to all individual parameters, the decentralized detector can outperform the centralized detector. □

We now state a result that provides a relation between the degrees of freedom (p_i and p_c) and the non-centrality parameters (λ_i and λ_c) of the local and the centralized detectors. This result plays a central role in comparing the performance of these centralized and decentralized detectors.

Lemma 5 ***(Degrees of freedom and non-centrality parameters)*** *Let p_i, p_c and λ_i, λ_c be the degrees of freedom and non-centrality parameters of the $i-$th local and centralized detectors, respectively. Then, $p_i \leq p_c$ and $\lambda_i \leq \lambda_c$ for all $i \in \{1, \dots, N\}$.*

A proof of Lemma 5 is postponed to the Appendix. In loose words, given the interpretation of the degrees of freedom and noncentrality parameters in Remark 4, Lemma (5) states that a centralized detector has more knowledge about the system dynamics ($p_i \leq p_c$) and its measurements contain a stronger attack signature ($\lambda_i \leq \lambda_c$) than any of the $i-$th local detector. Despite these properties, we will show that the decentralized detector can outperform the centralized one.

2.4 Comparison of centralized and decentralized detection of attacks

In this section we characterize the detection probabilities of the decentralized and centralized detectors, and we derive sufficient conditions for each detector to outperform the other. Recall that the decentralized detector triggers an alarm if any of the local detectors detects an alarm. In other words,

$$\begin{aligned} P_d^F &= \Pr\left[\Lambda_i \geq \tau_i, \text{ for some } i \in \{1, \dots, N\} \mid H_0\right], \\ P_d^D &= \Pr\left[\Lambda_i \geq \tau_i, \text{ for some } i \in \{1, \dots, N\} \mid H_1\right], \end{aligned} \tag{2.16}$$

where P_d^F and P_d^D denote the false alarm and detection probabilities of the decentralized detector, respectively.

Lemma 6 ***(Performance of the decentralized detector)*** *The false alarm and detection probabilities in* (2.16) *satisfy*

$$P_d^F = 1 - \prod_{i=1}^{N} \left(1 - P_i^F\right), \text{ and } P_d^D = 1 - \prod_{i=1}^{N} \left(1 - P_i^D\right). \tag{2.17}$$

A proof of Lemma 6 is postponed to the Appendix. As shown in Fig. 2.1, for the case when $P_i^F = P_j^F$, for all $i, j \in \{1, \dots, N\}$, P_d^F increases with increase in P_i^F and N. To allow for a fair comparison between the decentralized and centralized detectors, we assume that $P_c^F = P_d^F$. Consequently, for a probability P_c^F, the probabilities P_i^F satisfy

$$P_c^F = 1 - \prod_{i=1}^{N} \left(1 - P_i^F\right).$$

We now derive a sufficient condition for the centralized detector to outperform the decentralized detector.

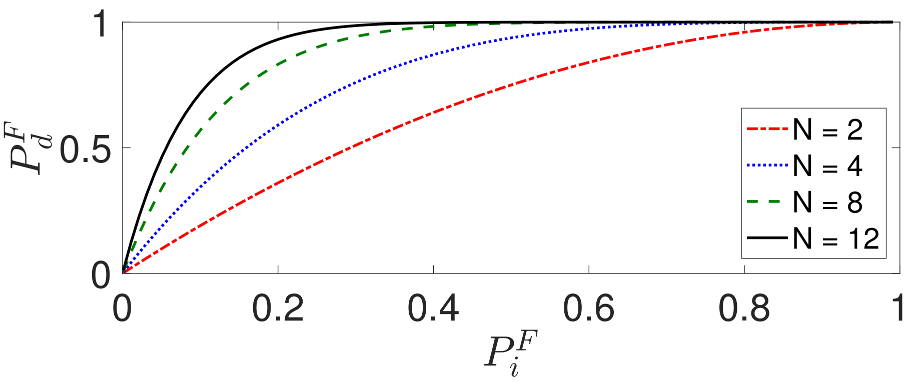


Figure 2.1: This figure shows the false alarm probability of the decentralized detector, P_d^F, as a function of the identical false alarm probabilities of the local detectors, P_i^F, for different numbers of local detectors.

Theorem 7 ***(Sufficient condition for*** $P_c^D \geq P_d^D$***)*** *Let* $P_c^F = P_d^F$*, and assume that the following condition is satisfied:*

$$\tau_c \leq p_c + \lambda_c - \sqrt{4N(p_c + 2\lambda_c)\ln\left(\frac{1}{1 - P_{max}^D}\right)}, \tag{2.18}$$

where $P_{max}^D = \max\{P_1^D, \dots, P_N^D\}$*. Then,* $P_c^D \geq P_d^D$*.*

A proof of Theorem 7 is postponed to the Appendix. We next derive a sufficient condition for the decentralized detector to outperform the centralized detector.

Theorem 8 ***(Sufficient condition for*** $P_d^D \geq P_c^D$***)*** *Let* $P_c^F = P_d^F$*, and assume that the following condition is satisfied:*

$$\begin{aligned}\tau_c \geq p_c + \lambda_c + \sqrt{4\,(p_c + 2\lambda_c)\ln\left(\frac{1}{1-(1-P_{min}^D)^N}\right)} + \\ + 2\ln\left(\frac{1}{1-(1-P_{min}^D)^N}\right),\end{aligned} \tag{2.19}$$

where $P_{min}^D = \min\{P_1^D, \dots, P_N^D\}$*. Then* $P_d^D \geq P_c^D$*.*

A proof of Theorem 7 is postponed to the Appendix. Theorems 7 and 8 provide sufficient conditions on the detectors and attack parameters that result in one detector outperforming the other. In particular, from (2.18) and (2.19) we note that, depending on decision threshold τ_c, a centralized detector may or may not outperform a decentralized detector. This is intuitive as the Q function, which quantifies the detection probability, is a decreasing function of the detection threshold (see Remark 4). To clarify the effect of attack and detection parameters on the performance trade-offs of the detectors, we now express (2.18) and (2.19) using the mean and standard deviation of Λ_c in (2.12). Let

$$\mu_c \triangleq \mathbb{E}\left[\Lambda_c\right] = \lambda_c + p_c, \quad \text{and} \quad \sigma_c \triangleq \mathrm{SD}[\Lambda_c] = \sqrt{2(p_c + 2\lambda_c)}.$$

where the expectation and standard deviation (SD) of Λ_c follows from the fact that under H_1, $\Lambda_c \sim \chi^2(p_c, \lambda_c)$ (see proof of Lemma 3). Hence, (2.18) and (2.19) can be rewritten, respectively, as

$$\tau_c \leq \mu_c - \sigma_c \underbrace{\sqrt{2N \ln\left(\frac{1}{1 - P^D_{\max}}\right)}}_{\triangleq \kappa_c}, \text{ and} \tag{2.20a}$$

$$\tau_c \geq \mu_c + \sigma_c \underbrace{\sqrt{2 \ln\left(\frac{1}{1 - (1 - P^D_{\min})^N}\right)}}_{\triangleq \kappa_d} + \kappa_d^2. \tag{2.20b}$$

From (2.20a) and (2.20b) we note that a centralized detector outperforms the decentralized one if τ_c is κ_c standard deviations smaller than the mean μ_c. Instead, a decentralized detector outperforms the centralized detector if τ_c is at least κ_d standard deviations larger than the mean μ_c. See Fig. 2.2 for a graphical illustration of this interpretation. Theorems 7 and 8 are illustrated in Fig. 2.3 as a function of the non-centrality parameters. It can be observed that (i) each of the detectors can outperform the other depending on the values of the noncentrality parameter values, (ii) the provided bounds qualitatively capture the actual performance of the centralized and decentralized detectors as the non-centrality parameters

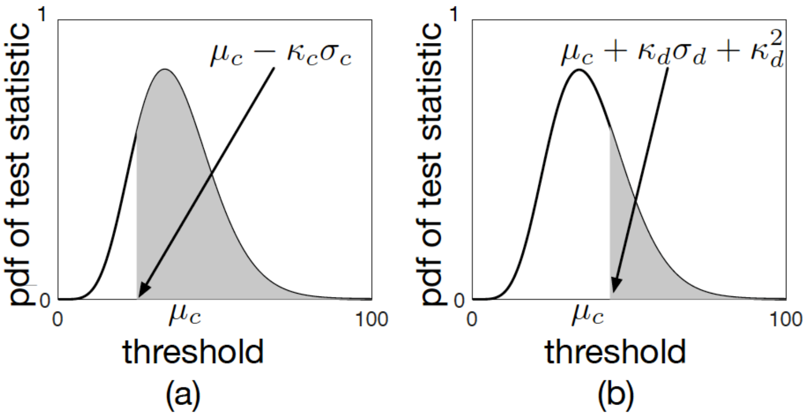


Figure 2.2: This figure shows the probability density function of Λ_c under H_1, as a function of threshold τ_c. For $\tau_c = \mu_c - \kappa_c\sigma_c$ and $\tau_c = \mu_c + \kappa_d\sigma_d + \sigma_d^2$, the shaded area in panels (a) and (b) indicates the detection probability of the centralized detector. As seen in panels (a) and (b), an increase in κ_c results in larger area (larger detection probability) while a increase in κ_d results in smaller area (smaller detection probability).

increase, and (iii) the provided bounds are rather tight over a large range of non-centrality parameters. In Fig. 2.4 we show that the difference of the detection probabilities of the centralized and decentralized detectors can be large, especially when the non-centrality parameters satisfy $\lambda_c \approx \lambda_i$, as evident in Fig. 2.4 (a).

2.5 Design of optimal attacks

In this section we consider the problem of designing attacks that deteriorate the performance of the interconnected system (2.1) while remaining undetected from the centralized and decentralized detectors. We measure the degradation induced by an attack with the expected value of the deviation of the state trajectory from the origin. We assume that the attack is a deterministic signal, and thus independent of the noise affecting the system dynamics and measurements. In particular, for a fixed value of the probability P_c^F

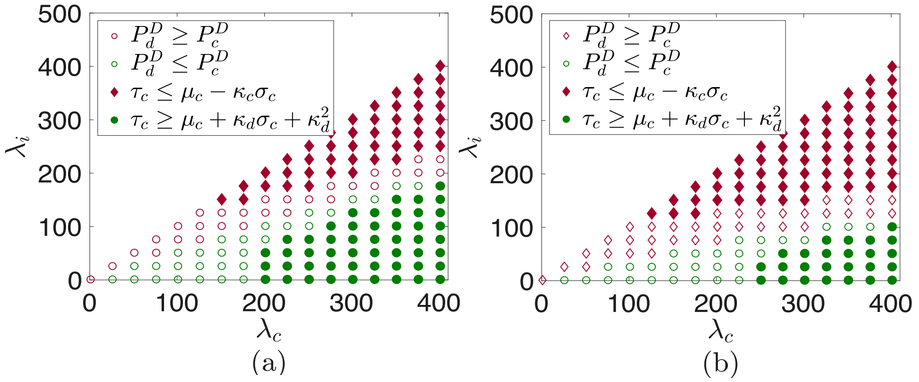


Figure 2.3: This figure shows when the decentralized, which comprises identical local detectors, and centralized detectors outperform their counterpart, as a function of the noncentrality parameters. The regions identified by solid markers correspond to the conditions in Theorems 7 and 8. Instead, regions identified by empty markers are identified numerically. Since $\lambda_i \leq \lambda_c$, the white region is not admissible. For $P_c^F = P_d^F = 0.01$, (a) corresponds to $N = 2$ and (b) corresponds to the case of $N = 4$. When $N = 4$, the decentralized detector outperforms the centralized one for a larger set of noncentrality parameters.

and a threshold $P_c^F \leq \delta_c \leq 1$, we consider the optimization problem

$$\begin{aligned}
\text{(P.1)} \quad \max_{U^a} \quad & \mathbb{E}\left[\sum_{k=1}^{T} x(k)^\mathsf{T} x(k)\right], \\
\text{subject to} \quad & P_c^D \leq \delta_c, \\
& x(k+1) = Ax(k) + B^a u^a(k) + w(k),
\end{aligned}$$

where U^a is the deterministic attack input over time horizon T (see (2.7)). Notice that, because the attack is deterministic, the objective function in (P.1) can be simplified by bringing the expectation inside the summation, and replacing the state equation constraint with the mean state response. Further, because the system parameters and P_c^F are fixed, τ_c and p_c are also fixed, which ensures that P_d^D only depends on noncentrality parameter. This observation along with the fact that $Q(\cdot)$ is increasing function in noncentrality parameter (see Remark 4) allows us to express the detection constraint in terms of λ_c. Specifically,

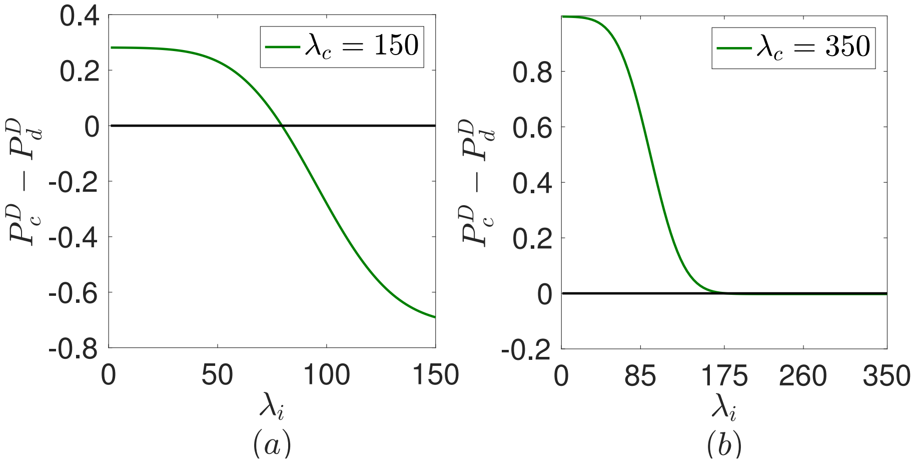


Figure 2.4: This figure shows the difference of the centralized and decentralized detection probabilities as a function of λ_i for different values of λ_c. For small values of λ_c, the detection probability of the decentralized detector can be very large than its centralized counterpart.

the optimization problem (P.1) can be rewritten as

$$\begin{aligned} \text{(P.2)} \quad \max_{U^a} \quad & \sum_{k=1}^{T} \overline{x}(k)^{\mathsf{T}} \overline{x}(k) \\ \text{subject to} \quad & (U^a)^{\mathsf{T}} M_c (U^a) \leq \widetilde{\delta}_c, \\ & \overline{x}(k+1) = A\overline{x}(k) + B^a u^a(k), \end{aligned}$$

where we have used that $\text{Cov}\left[x(k)\right]$ is independent of the attack $u^a(k)$, and

$$\mathbb{E}[x(k)^{\mathsf{T}} x(k)] = \overline{x}(k)^{\mathsf{T}} \overline{x}(k) + \text{Trace}\left(\text{Cov}\left[x(k)\right]\right),$$

with $\overline{x}(k) = \mathbb{E}[x(k)]$. Further, we have $\widetilde{\delta}_c = Q^{-1}_{p_c,\tau_c}(\delta_c)$, where $Q^{-1}_{p_c,\tau_c}(\alpha) : [0,1] \rightarrow [0,\infty]$ denotes the inverse of $Q(\tau_c; p_c, \lambda_c)$ for fixed p_c and τ_c, and $\lambda_c = (U^a)^{\mathsf{T}} M_c (U^a)$, with M_c as in (2.15). It should be noticed that the attack constraint in (P.2) essentially limits the energy of the attack signal. We next characterize the solution to the problem (P.2).

Theorem 9 ***(Optimal attack vectors)*** *Let U_c^* be any solution of (P.2). Then, there exist a $\gamma_c > 0$ such that the pair (U_c^*, γ_c) solves the following optimality equations:*

$$\left[\mathcal{B}_a^{\mathsf{T}}\mathcal{B}_a - \gamma_c M_c\right] U_c^* + \mathcal{B}_a^{\mathsf{T}}\mathcal{A}x(0) = 0, \tag{2.21a}$$

$$\left(U_c^*\right)^{\mathsf{T}} M_c(U_c^*) = \widetilde{\delta}_c, \tag{2.21b}$$

where

$$\mathcal{A} = \begin{bmatrix} A \\ \vdots \\ A^T \end{bmatrix} \text{ and } \mathcal{B}_a = \begin{bmatrix} B^a & \cdots & 0 \\ \vdots & \ddots & \vdots \\ A^{T-1}B^a & \dots & B^a \end{bmatrix}. \tag{2.22}$$

A proof of Theorem 9 is postponed to the Appendix. Theorem 9 not only guarantees the existence of optimal attacks, but it also provides us with necessary conditions to verify if an attack is (locally) optimal. When the system initial state is zero, we can also quantify the performance degradation induced by an optimal attack. Let $\rho_{\max}(A, B)$ and $\nu_{\max}(A, B)$ denote a largest generalized eigenvalue of a matrix pair (A, B) and one of its associated generalized eigenvectors [18].

Lemma 10 ***(System degradation with zero initial state)*** *Let $x(0) = 0$. Then, the optimal solution to (P.2) is*

$$U_c^* = \left(\sqrt{\frac{\widetilde{\delta}_c}{(\nu^*)^{\mathsf{T}} M_c(\nu^*)}}\right)\nu^*, \tag{2.23}$$

and its associated optimal cost is

$$J_c^* = \widetilde{\delta}_c\, \rho_{max}\left(\mathcal{B}_a^{\mathsf{T}}\mathcal{B}_a, M_c\right), \tag{2.24}$$

where $\nu^* = \nu_{max}\left(\mathcal{B}_a^{\mathsf{T}}\mathcal{B}_a, M_c\right)$.

A proof of Lemma 10 is postponed to the Appendix. From (2.24), note that the system degradation caused by an optimal attack depends on the detector's tolerance, as measured by $\widetilde{\delta}_c$, and the system dynamics, as measured by $\rho_{\max}(\cdot)$. See Remark 12 for the influence of noise uncertainty on the system degradation due to optimal attacks.

Remark 11 ***(Optimal attack vector against decentralized detector)*** *To characterize the performance degradation of the system analytically, we consider a relaxed form of detection constraint. Specifically, we design optimal attacks subjected to $\overline{P}_d^D \leq \delta_d$ instead of $P_d^D \leq \delta_d$, where $\overline{P}_d^D$ is an upper bound on P_d^D (see Lemma 14). The design of optimal attacks that are undetectable from the decentralized detector can be formulated as:*

$$\begin{aligned} \textit{(P.3)} \quad \max_{U^a} \quad & \sum_{k=1}^{T} \overline{x}(k)^{\mathsf{T}} \overline{x}(k) \\ \textit{subject to} \quad & \sum_{i=1}^{N} (U_i^a)^{\mathsf{T}} M_i (U_i^a) \leq \widetilde{\delta}_d, \\ & \overline{x}(k+1) = A\overline{x}(k) + B^a u^a(k), \end{aligned}$$

where the summation in the detectability constraint follows from Lemma 14 and the fact that $\overline{P}_d^D \leq \delta_d$ becomes equivalent to $\sum_{i=1}^N \lambda_i \leq \widetilde{\delta}_d$, where $\widetilde{\delta}_d = Q^{-1}_{p_{sum},\tau_{min}}(\delta_d)$, $p_{sum} = \sum_{i=1}^N p_i$, and $\tau_{min} = \min\limits_{1 \leq i \leq N} \tau_i$. Let Π_i be a permutation matrix such that $U_i^a = \Pi_i U^a$, and let $\Pi = \left[\Pi_1^{\mathsf{T}}, \ldots, \Pi_N^{\mathsf{T}}\right]^{\mathsf{T}}$ and $M_d = \Pi^{\mathsf{T}} \mathrm{blkdiag}(M_1, \ldots, M_N)\Pi$. For any solution U_d^* of (P.2), there exist $\gamma_d > 0$ such that the pair (U_d^*, γ_d) solves the following optimality equations:

$$\begin{aligned} \left[\mathcal{B}_a^{\mathsf{T}} \mathcal{B}_a - \gamma_d M_d\right] U_d^* + \mathcal{B}_a^{\mathsf{T}} \mathcal{A} x(0) = 0, \ \textit{and} \\ (U_d^*)^{\mathsf{T}} M_d (U_d^*) = \widetilde{\delta}_d. \end{aligned}$$

Further, if $x(0) = 0$, then the largest degradation is $J_d^* = \widetilde{\delta}_d \, \rho_{max}\left(\mathcal{B}_a^{\mathsf{T}} \mathcal{B}_a, M_d\right)$. □

Remark 12 ***(Maximum degradation of the system performance with respect to system noise)*** *To see the role of noise level, in the processed measurements, on the system degradation, we consider the following covariance matrices:* $\Sigma_{w_i} = \sigma^2 I_{n_i}$ *and* $\Sigma_{v_i} = \sigma^2 I_{r_i}$*, for* $i \in \{1, \dots, N\}$*. Then, from* (2.24) *we have*

$$J_c^* = \sigma^2 \, \widetilde{\delta}_c \, \left[\rho_{max} \left(\mathcal{B}_a^{\mathsf{T}} \mathcal{B}_a, \widetilde{M}_c \right) \right], \tag{2.25}$$

where $\widetilde{M}_c = \left(N_c \mathcal{F}_c^{(a)} \right)^{\mathsf{T}} \left[\mathcal{F}_c^{(w)} \left(\mathcal{F}_c^{(w)} \right)^{\mathsf{T}} + I \right]^{-1} \left(N_c \mathcal{F}_c^{(a)} \right)$*. From* (2.25) *we note that the system degradation increases with the increase in the noise level, i.e.,* σ^2*.* □

2.6 Numerical comparison of centralized and decentralized detectors

In this section, we demonstrate our theoretical findings on the IEEE RTS-96 power network model [32], which we partition into three subregions as shown in Fig. 2.5. We followed the approach in [23] to obtain a linear time-invariant model of the power network, and then discretized it using a sampling time of 0.01 seconds. For a false alarm probability $P_c^F = P_d^F = 0.05$, we consider the family of attacks $U^a = \sqrt{\theta / (\mathbf{1}^{\mathsf{T}} M_c \mathbf{1})} \mathbf{1}$, where $\mathbf{1}$ is the vector of all ones and $\theta > 0$. It can be shown that the noncentrality parameters satisfy $\lambda_c = \theta$ and $\lambda_i = \theta(\mathbf{1}^{\mathsf{T}} M_i \mathbf{1})/(\mathbf{1}^{\mathsf{T}} M_c \mathbf{1})$, and moreover, the choice of vector $\mathbf{1}$ is arbitrary and it does not affecting the following results.

(Illustration of Theorem 7) For the measurement horizon of $T = 100$ seconds, the values of p_c and τ_c are 5130 and 5480.6, respectively. Fig. 2.6 show that the detection probabilities of the centralized and decentralized detectors increase monotonically with the attack parameter

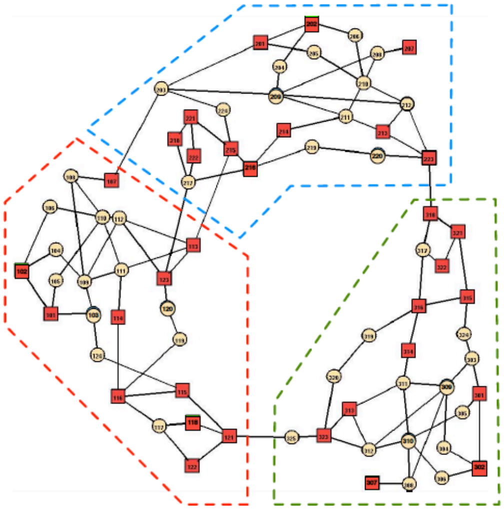
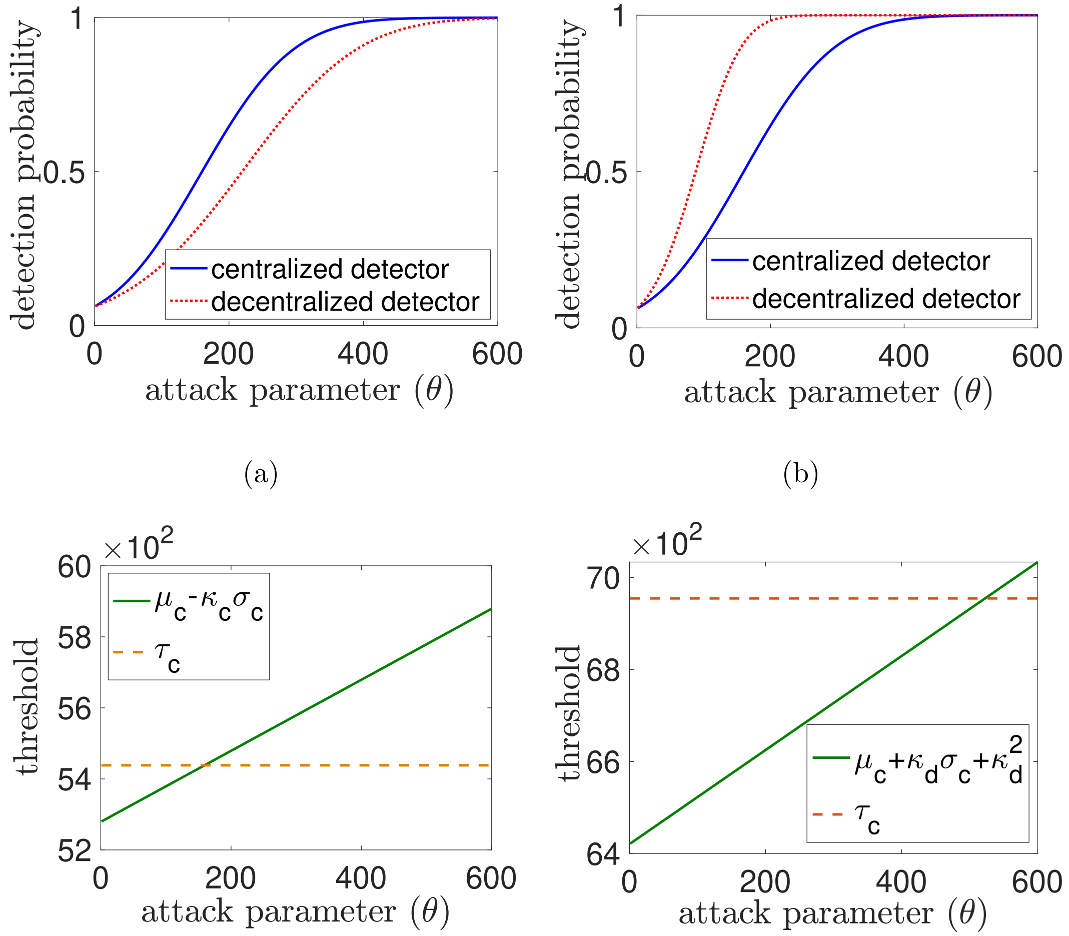


Figure 2.5: The figure shows a single-line diagram of IEEE RTS-96 power network, which is composed of three weakly-coupled areas (subsystems). The square nodes denote the generators, while the circular nodes denotes the load buses of the network [23].

θ. As predicted by the condition (2.20a) and shown in Fig. 2.6, the centralized detector is guaranteed to outperform the decentralized detector when $\theta > 173$. This figure shows that our result is conservative, because $P_c^D \geq P_d^D$ for all values of θ as shown in Fig. 2.6.

(Illustration of Theorem 8) Contrary to the previous example, by letting $T = 125$ seconds, we obtain $p_c = 6755$ and $\tau_c = 6947.3$. For these parameters, the decentralized detector is guaranteed to outperform the centralized one when $\theta \leq 511$. This behavior is predicted by our sufficient condition (2.20b), and is illustrated in Fig. 2.6. The estimation provided by our condition (2.20b) is conservative, as illustrated in Fig. 2.6.

(Illustration of Lemma 10) In Fig. 2.7 we compare the performance degradation induced by the optimal attacks designed according to the optimization problems (P.2) and (P.3) with zero initial conditions. In particular, we plot the optimal costs J_c^* and J_d^* against the tolerance levels $\widetilde{\delta}_c$ and $\widetilde{\delta}_d$, respectively. As expected, the performance degradation is



(a)

(b)

(c)

(d)

Figure 2.6: Scenarios in which the centralized detector outperforms the decentralized detector (a), and vice versa (b), on the IEEE RTS-96 power network, for a range of attack parameter (θ) values. In panels (c) and (d) we plot the right (solid line) and left hand expressions (dashed line) of the inequalities (2.20a) and (2.20b), respectively, as a function of θ. For attacks such that the time horizon $T = 100$ sec and $\theta > 200$, the sufficient condition (2.20a) holds true, it guarantees that $P_c^D \geq P_d^D$. Instead, when $T = 125$ sec and $\theta < 500$, the sufficient condition (2.20b) holds true, it guarantees that $P_d^D \geq P_c^D$.

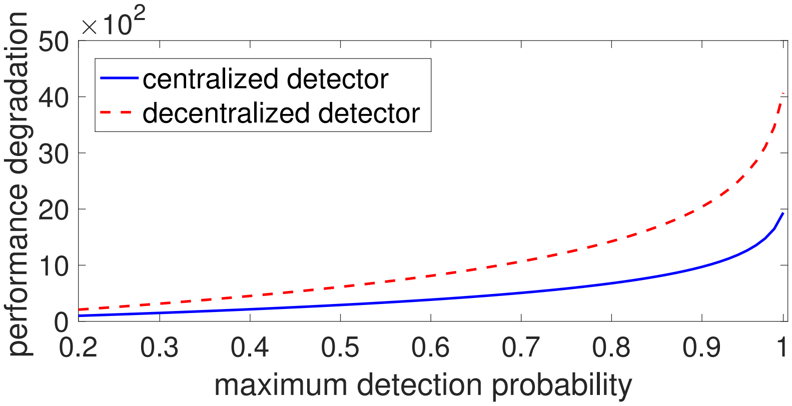


Figure 2.7: This figure shows the performance degradation induced by undetectable optimal attacks on the IEEE RTS-96 power network. The performance degradation is computed using the optimal cost J_c^* and J_d^* derived in Lemma 10 and Remark 11, respectively. Instead, the maximum detection probability is given by the tuning parameters δ_c and δ_d in the detection probability constraints of the optimization problems (P.2) and (P.3), respectively.

proportional to the tolerance levels and, for the considered setup, it is larger in the case of the decentralized detector.

2.7 Summary

In this work we compare the performance of centralized and decentralized schemes for the detection of attacks in stochastic interconnected systems. In addition to quantifying the performance of each detection scheme, we prove the counterintuitive result that the decentralized scheme can, at times, outperform its centralized counterpart, and that this behavior results due to the simple versus composite nature of the attack detection problem. We illustrate our findings through academic examples and a case study based on the IEEE RTS-96 power system.

2.8 Appendix

Proof of Lemma 2: Since the attack vectors U_i^a and U^a are deterministic, and W_i, V_i, V, and W are zero mean random vectors, from the linearity of the expectation operator it follows from (2.8) that

$$\beta_i \triangleq \mathbb{E}[\widetilde{Y}_i] = N_i \mathcal{F}_i^{(a)} U_i^a, \text{ and } \beta_c \triangleq \mathbb{E}[\widetilde{Y}_c] = N_c \mathcal{F}_c^{(a)} U_c^a.$$

Further, from the properties of Cov[$\cdot$], we have the following:

$$\begin{aligned}\Sigma_i \triangleq \text{Cov}\left[\widetilde{Y}_i\right] &= N_i \text{Cov}\left[Y_i\right] N_i^{\mathsf{T}} \\ &\overset{(a)}{=} N_i \left[\text{Cov}\left[\mathcal{F}_i^{(w)} W_i\right] + \text{Cov}[V_i]\right] N_i^{\mathsf{T}} \\ &\overset{(b)}{=} N_i \left[\left(\mathcal{F}_i^{(w)}\right) \text{Cov}\left[W_i\right] \left(\mathcal{F}_i^{(w)}\right)^{\mathsf{T}} + \text{Cov}[V_i]\right] N_i^{\mathsf{T}} \\ &= N_i \left[\left(\mathcal{F}_i^{(w)}\right) \left(I_T \otimes \Sigma_{w_i}\right) \left(\mathcal{F}_i^{(w)}\right)^{\mathsf{T}} + \left(I_T \otimes \Sigma_{v_i}\right)\right] N_i^{\mathsf{T}},\end{aligned}$$

where (a) follows because the measurement and process noises are independent of each other. Instead, (b) is due to the fact that the noise vectors are independent and identically distributed. Similar analysis also results in the expression of Σ_c, and hence the details are omitted. Finally, by invoking the fact that linear transformations preserve Gaussianity, the distribution of $\widetilde{Y}_i$ and $\widetilde{Y}_c$ is Gaussian as well. □

Proof of Lemma 3:

From the statistics and distributional form of $\widetilde{Y}_i$ and $\widetilde{Y}_c$ (see (2.9)), and threshold tests defined in (2.11) and (2.12), it follows from [1, Theorem 3.3.3] that, under

1. H_0: $\Lambda_i \sim \chi^2(p_i)$ and $\Lambda_c \sim \chi^2(p_c)$, where p_i and p_c are defined in (2.14).

2. H_1: $\Lambda_i \sim \chi^2(p_i, \lambda_i)$ and $\Lambda_c \sim \chi^2(p_c, \lambda_c)$, where $\lambda_i = \beta_i^{\mathsf{T}} \Sigma_i^{-1} \beta_i$ and $\lambda_c = \beta_c^{\mathsf{T}} \Sigma_c^{-1} \beta_c$.

By substituting $\beta_i = N_i \mathcal{F}_i^{(a)} U_i^a$ and $\beta_c = N_c \mathcal{F}_c^{(a)} U_c^a$ (see Lemma 2) and rearranging the terms, we get the expressions of λ_i and λ_c in (2.14). Finally, from the aforementioned distributional forms of Λ_i and Λ_c, it now follows that the false alarm and the detection probabilities of the tests (2.11) and (2.12) are the right tail probabilities (represented by $Q(\cdot)$ function) of the central and noncentral chi-squared distributions, respectively. Hence, the expressions in (2.13) follow. □

Proof of Lemma 5: Without loss of generality let $i = 1$. Thus, it suffices to show that a) $p_1 \leq p_c$ and b) $\lambda_1 \leq \lambda_c$.

Case (a): For brevity, define

$$\begin{aligned} \widetilde{\Sigma}_i &= \left[\left(\mathcal{F}_i^{(w)}\right)(I_T \otimes \Sigma_{w_i})\left(\mathcal{F}_i^{(w)}\right)^{\mathsf{T}} + (I_T \otimes \Sigma_{v_i})\right] \text{ and} \\ \widetilde{\Sigma}_c &= \left[\left(\mathcal{F}_c^{(w)}\right)(I_T \otimes \Sigma_{w})\left(\mathcal{F}_c^{(w)}\right)^{\mathsf{T}} + (I_T \otimes \Sigma_{v})\right], \end{aligned} \tag{2.26}$$

and note that $\widetilde{\Sigma}_i > 0$ and $\widetilde{\Sigma}_c > 0$. From Lemma 2, Lemma 3, and (2.26), we have

$$\begin{aligned} p_c = \text{Rank}(\Sigma_c) &= \text{Rank}\left(\left(N_c \widetilde{\Sigma}_c^{1/2}\right)\left(N_c \widetilde{\Sigma}_c^{1/2}\right)^{\mathsf{T}}\right) \\ &= \text{Rank}\left(N_c \widetilde{\Sigma}^{1/2}\right) = \text{Rank}\left(N_c\right). \end{aligned}$$

Similarly, $p_1 = \text{Rank}\left(N_1\right)$. Since, N_1^{T} and N_c^{T} are a basis vectors of the null spaces $\mathcal{N}_1^L$ and $\mathcal{N}_c^L$ (see (2.37)) respectively, it follows from Proposition 13 that $p_1 \leq p_c$.

Case (b): As the proof for this result is rather, we break it down in to multiple steps:

- *Step 1*: Express λ_1 and λ_c using the statistics of a permuted version of Y_c.
- *Step 2*: Obtain lower bound on λ_c, which depends on the statistics of the measurements pertaining to Subsystem 1.

- *Step 3*: Show that λ_1 is less than bound in Step 2.

Step 1 (alternative form of λ_1 and λ_c): Notice that λ_1 and λ_c in (2.14) can be expressed as $\lambda_1 = \beta_1^\mathsf{T}\Sigma_1^{-1}\beta_1$ and $\lambda_c = \beta_c^\mathsf{T}\Sigma_c^{-1}\beta_c$, respectively, where β_1, β_c, Σ_1, and Σ_c are defined in Lemma 2. For convenience, we express λ_1 and λ_c in an alternative way. Let $i \in \{1, \dots, N\}$ and consider the $i-$th sensor measurements of (2.3)

$$y_{c,i}(k) = \underbrace{\begin{bmatrix} 0 & \dots & C_i & \dots & 0 \end{bmatrix}}_{\triangleq C_{c,i}} x(k) + v_i(k). \tag{2.27}$$

Also, define $Y_{c,i} = \begin{bmatrix} y_{c,i}^\mathsf{T}(1) & \dots & y_{c,i}^\mathsf{T}(T) \end{bmatrix}^\mathsf{T}$ and $\widehat{Y}_c = \begin{bmatrix} Y_{c,1}^\mathsf{T} & \dots & Y_{c,N}^\mathsf{T} \end{bmatrix}$. Now, from (2.27) and state equation in (2.3), $Y_{c,i}$ can be expanded as

$$Y_{c,i} = \mathcal{O}_{c,i}x(0) + \mathcal{F}_{c,i}^{(a)}U^a + \mathcal{F}_{c,i}^{(w)}W + V_i,$$

where the matrices $\mathcal{O}_{c,i}$, $\mathcal{F}_{c,i}^{(a)}$, and $\mathcal{F}_{c,i}^{(w)}$ are similar to the matrices defined in Section II-A. By substituting the above decomposition of $Y_{c,i}$ in $\widehat{Y}$ we have

$$\widehat{Y}_c = \underbrace{\begin{bmatrix} \mathcal{O}_{c,1} \\ \vdots \\ \mathcal{O}_{c,N} \end{bmatrix}}_{\widehat{\mathcal{O}}_c} x(0) + \underbrace{\begin{bmatrix} \mathcal{F}_{c,1}^{(a)} \\ \vdots \\ \mathcal{F}_{c,N}^{(a)} \end{bmatrix}}_{\widehat{\mathcal{F}}_c^a} U^a + \underbrace{\begin{bmatrix} \mathcal{F}_{c,1}^{(w)} \\ \vdots \\ \mathcal{F}_{c,N}^{(w)} \end{bmatrix}}_{\widehat{\mathcal{F}}_c^w} W + V.$$

Moreover, from the distributional assumptions on W and V, it readily follows that (similarly to the proof of Lemma 2),

$$\widehat{Y}_c \sim \mathcal{N}\left(\widehat{\mathcal{O}}_c x(0) + \widehat{\mathcal{F}}_c^a U^a, \Sigma\right), \tag{2.28}$$

where $\Sigma = \left(\widehat{\mathcal{F}}_c^w\right)(I_T \otimes \Sigma_w)\left(\widehat{\mathcal{F}}_c^w\right)^\mathsf{T} + (I_T \otimes \Sigma_v)$, and Σ_w and Σ_v are defined in Lemma 2.

Now, consider the measurement equation $y_i(k)$ in (2.1) and note that $C_{c,i}x(k) = C_i x_i(k)$. Thus, $y_i(k) = y_{c,i}(k)$, for all $i \in \{1, \dots, N\}$ and $k \in \mathbb{N}$. From this observation it follows that $Y_i = Y_{c,i} = \Pi_i \widehat{Y}_c$, where Π_i is a selection matrix. Let $\widetilde{N}_i = N_i \Pi_i$ and note that $\widetilde{N}_i \widehat{\mathcal{O}} = N_i \mathcal{O}_{c,i}$. Further from Proposition 13 we have $N_i \mathcal{O}_{c,i} = 0$. With these facts in place, from Lemma 2 we now have

$$\beta_i = \widetilde{N}_i \widehat{\mathcal{F}}_c^a U^a \text{ and } \Sigma_i = \widetilde{N}_i \Sigma \widetilde{N}_i^{\mathsf{T}}. \tag{2.29}$$

Similarly, since $\widehat{Y}_c$ is just a rearrangement of Y_c (see (2.5)), there exists a permutation matrix Q such that $Y_c = Q\widehat{Y}_c$, and, ultimately $\widetilde{Y}_c = N_c Y_c = N_c Q \widehat{Y}_c$. Thus,

$$\beta_c = N_c Q \widehat{\mathcal{F}}_c^a U^a \text{ and } \Sigma_c = N_c Q \Sigma (N_c Q)^{\mathsf{T}}. \tag{2.30}$$

Let $z = \widehat{\mathcal{F}}_c^a U^a$. From (2.29) and (2.30) we have

$$\begin{aligned} \lambda_1 &= z^{\mathsf{T}} \widetilde{N}_1^{\mathsf{T}} \left[\widetilde{N}_1 \Sigma \widetilde{N}_1^{\mathsf{T}} \right]^{-1} \widetilde{N}_1 z, \\ \lambda_c &= z^{\mathsf{T}} \left(N_c Q\right)^{\mathsf{T}} \left[\left(N_c Q\right) \Sigma \left(N_c Q\right)^{\mathsf{T}} \right]^{-1} \left(N_c Q\right) z. \end{aligned} \tag{2.31}$$

Step 2 (lower bound on λ_c): Since, $Y_c = N_c Y_c = N_c Q \widehat{Y}_c$, it follows that $N_c Q$ is the basis of the null space $\widehat{\mathcal{O}}_c$. Further, the row vectors of $\mathcal{O}_{c,i}$ and $\mathcal{O}_{c,j}$ are linearly independent, whenever $i \neq j$. Using these facts we can define $N_{c,i} = \begin{bmatrix} N_{c,i}^i & \cdots & N_{c,i}^N \end{bmatrix}$ such that $N_c Q = \begin{bmatrix} N_{c,1}^{\mathsf{T}} & \cdots & N_{c,N}^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}}$, where $N_{c,i}^i \mathcal{O}_{c,i} = 0$. Let $P_1 = \begin{bmatrix} (N_{c,2})^{\mathsf{T}} & \cdots & (N_{c,N})^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}}$ and note that

$$\left(N_c Q\right) \Sigma \left(N_c Q\right)^{\mathsf{T}} = \begin{bmatrix} N_{c,1} \\ P_1 \end{bmatrix} \Sigma \begin{bmatrix} N_{c,1}^{\mathsf{T}} & P_1^{\mathsf{T}} \end{bmatrix} = \begin{bmatrix} N_{c,1} \Sigma N_{c,1}^{\mathsf{T}} & N_{c,1} \Sigma P_1^{\mathsf{T}} \\ N_{c,1}^{\mathsf{T}} \Sigma P_1 & P_1^{\mathsf{T}} \Sigma P_1 \end{bmatrix}.$$

Let $S_1 = N_{c,1} \Sigma N_{c,1}^{\mathsf{T}}$. Since $\Sigma > 0$, it follows that both the matrices S_1 and $P_1^{\mathsf{T}} \Sigma P_1$ are invertible. Hence, from Schur's complement, there exists a matrix $X \geq 0$ such that

$$\left[(N_c Q)\,\Sigma\,(N_c Q)^{\mathsf{T}}\right]^{-1} = \begin{bmatrix} S_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} + X. \tag{2.32}$$

Similarly, consider the following partition of Σ:

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

where $\Sigma_{11} > 0$ and $\Sigma_{22} > 0$, and let $S_2 = (N_{c,1}^1)\Sigma_{11}(N_{c,1}^1)^{\mathsf{T}}$. Invoking Schur's complement, we have the following:

$$S_1^{-1} = \begin{bmatrix} S_2^{-1} & 0 \\ 0 & 0 \end{bmatrix} + Y, \tag{2.33}$$

where $Y \geq 0$. Substituting(2.32) and (2.33) in (2.31), it follows that

$$\begin{aligned}
\lambda_c &= z^{\mathsf{T}}(N_c Q)^{\mathsf{T}} \begin{bmatrix} S_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} (N_c Q)z + \underbrace{z^{\mathsf{T}}(N_c Q)^{\mathsf{T}} X (N_c Q) z}_{\geq 0} \\
&\geq \left[(N_{c,1}z)^{\mathsf{T}}\,(P_1 z)^{\mathsf{T}}\right] \begin{bmatrix} S_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} N_{c,1}z \\ P_1 z \end{bmatrix} \\
&= z^{\mathsf{T}}\left(N_{c,1}^{\mathsf{T}} S_1^{-1} N_{c,1}\right) z \\
&= z^{\mathsf{T}}(N_{c,1})^{\mathsf{T}} \begin{bmatrix} S_2^{-1} & 0 \\ 0 & 0 \end{bmatrix} (N_{c,1})z + \underbrace{z^{\mathsf{T}}(N_{c,1})^{\mathsf{T}} Y (N_{c,1}) z}_{\geq 0} \\
&\geq z^{\mathsf{T}}(N_{c,1})^{\mathsf{T}} \begin{bmatrix} S_2^{-1} & 0 \\ 0 & 0 \end{bmatrix} (N_{c,1})z = z^{\mathsf{T}} \begin{bmatrix} (N_{c,1}^1)^{\mathsf{T}} S_2^{-1} N_{c,1}^1 & 0 \\ 0 & 0 \end{bmatrix} z.
\end{aligned} \tag{2.34}$$

Instead, λ_1 in (2.31) can be shown as

$$\lambda_1 = z^{\mathsf{T}} \begin{bmatrix} N_1^{\mathsf{T}} \left[N_1 \Sigma_{11} N_1^{\mathsf{T}}\right]^{-1} N_1 & 0 \\ 0 & 0 \end{bmatrix} z, \tag{2.35}$$

where we used the fact that $\widetilde{N}_1 = N_1 \Pi_1$.

Step 3 ($\lambda_c \geq \lambda_1$): For $\lambda_c \geq \lambda_1$ to hold true, it suffices to show the following:

$$(N_{c,1}^1)^{\mathsf{T}} S_2^{-1} N_{c,1}^1 \geq N_1^{\mathsf{T}} \left[N_1 \Sigma_{11} N_1^{\mathsf{T}}\right]^{-1} N_1.$$

By invoking Proposition 13, we note that there exists a full row rank matrix F_1, such that $N_1 = F_1 N_{c,1}^1$. Since F_1^{T} is a full column rank matrix, we can define an invertible matrix $\widetilde{F}_1^{\mathsf{T}} \triangleq \left[F_1^{\mathsf{T}}\ M_1^{\mathsf{T}}\right]$, where M_1 forms a basis for null space of F_1, such that the following holds

$$S_2^{-1} = \widetilde{F}_1^{\mathsf{T}} \left[\widetilde{F}_1 S_2 \widetilde{F}_1^{\mathsf{T}}\right]^{-1} \widetilde{F}_1 = \widetilde{F}_1^{\mathsf{T}} \begin{bmatrix} F_1 S_2 F_1^{\mathsf{T}} & F_1 S_2 M_1^{\mathsf{T}} \\ M_1 S_2 F_1^{\mathsf{T}} & M_1 S_2 M_1^{\mathsf{T}} \end{bmatrix}^{-1} \widetilde{F}_1.$$

By invoking Schur's complement, it follows that

$$\begin{bmatrix} F_1 S_2 F_1^{\mathsf{T}} & F_1 S_2 M_1^{\mathsf{T}} \\ M_1 S_2 F_1^{\mathsf{T}} & M_1 S_2 M_1^{\mathsf{T}} \end{bmatrix}^{-1} = \begin{bmatrix} \left(F_1 S_2 F_1^{\mathsf{T}}\right)^{-1} & 0 \\ 0 & 0 \end{bmatrix} + Y,$$

where $Z \geq 0$. Hence,

$$(N_{c,1}^1)^{\mathsf{T}} S_2^{-1} N_{c,1}^1 = (\widetilde{F}_1 N_{c,1}^1)^{\mathsf{T}} \begin{bmatrix} \left(F_1 S_2 F_1^{\mathsf{T}}\right)^{-1} & 0 \\ 0 & 0 \end{bmatrix} (\widetilde{F}_1 N_{c,1}^1) + (\widetilde{F}_1 N_{c,1}^1)^{\mathsf{T}} Z (\widetilde{F}_1 N_{c,1}^1).$$

By substituting $\widetilde{F}_1^{\mathsf{T}} = [F_1^{\mathsf{T}}\ M_1^{\mathsf{T}}]$ in the above expression, and rearranging the terms we have

$$(N_{c,1}^1)^{\mathsf{T}} S_2^{-1} N_{c,1}^1 = (F_1 N_{c,1}^1)^{\mathsf{T}} \left(F_1 S_2 F_1^{\mathsf{T}}\right)^{-1} (F_1 N_{c,1}^1) + (\widetilde{F}_1 N_{c,1}^1)^{\mathsf{T}} Z (\widetilde{F}_1 N_{c,1}^1).$$

The required inequality follows by substituting $S_2 = (N_{c,1}^1)\Sigma_{11}(N_{c,1}^1)^\mathsf{T}$ and $N_1 = F_1 N_{c,1}$, and recalling the fact that the sum of two positive semi definite matrices is greater than or equal to either of the matrices. $\square$

Proof of Lemma 6: Let $\mathcal{E}_i$ be an event that the $i-$th local detector decides H_1 when the true hypothesis is H_0. Then, $P_i^F = \Pr\left[\mathcal{E}_i\right]$. Let $\mathcal{E}_i^\complement$ be the complement of $\mathcal{E}_i$. Then, from (2.16) it follows that

$$\begin{aligned} P_d^F &= \Pr\left(\bigcup_{i=i}^{N}\mathcal{E}_i\right) = 1 - \Pr\left(\bigcap_{i=i}^{N}\mathcal{E}_i^\complement\right) \overset{(a)}{=} 1 - \prod_{i=1}^{N}\Pr\left(\mathcal{E}_i^\complement\right) \\ &= 1 - \prod_{i=1}^{N}\left(1 - \Pr\left(\mathcal{E}_i\right)\right) = 1 - \prod_{i=1}^{N}\left(1 - P_i^F\right), \end{aligned}$$

where for the (a) we used the fact that the events $\mathcal{E}_i$ are mutually independent for all $i \in \{1, \dots N\}$. To see this fact, notice that the event $\mathcal{E}_i$ is defined on $\widetilde{Y}_i$ (see (2.8)). Further, $\widetilde{Y}_i$ depends only on the deterministic attack signal U_i^a and the noise vectors V_i and W_i, but not on the interconnection signal U_i (see (2.6)). Now, by invoking the fact that noises variables across different subsystems are independent, it also follows that the events $\mathcal{E}_i$ are also mutually independent. Similar procedure will lead to the analogous expression for P_d^D and hence, the details are omitted. $\square$

Proof of Theorem 7: Let $\mu_c = p_c + \lambda_c$ and $\sigma_c = \sqrt{2(p_c + 2\lambda_c)}$, and assume that (2.18) holds true. Then, from the monotonicity property of the CDF associated with the test statistic Λ_c, which follows $\chi^2(p_c, \lambda_c)$, we have the following inequality

$$\Pr\left[\Lambda_c \leq \tau_c\right] \leq \Pr\left[\Lambda_c \leq \mu_c - \sigma_c\sqrt{2N\ln\left(\frac{1}{1 - P_{\max}^D}\right)}\right].$$

From the inequality (2.41b), it now follows that

$$\Pr\left[\Lambda_c \leq \tau_c\right] \leq \exp\left(-N \ln\left(\frac{1}{1-P^D_{\max}}\right)\right) = \exp\left(\ln\left(1-P^D_{\max}\right)^N\right) \leq \prod_{i=1}^{N}\left(1-P^D_i\right),$$

where for the last inequality we used the fact that $P^D_i \leq P^D_{\max}$ for all $i \in \{1, \dots, N\}$. By using the above inequality and Lemma 3, under hypothesis H_1, we have

$$P^D_c = 1 - \Pr\left[\Lambda_c \leq \tau_c | H_1\right] \geq 1 - \prod_{i=1}^{N}\left(1-P^D_i\right) = P^D_d. \quad \square$$

Proof of Theorem 8: Let $\mu_c = p_c + \lambda_c$ and $\sigma_c = \sqrt{2(p_c + 2\lambda_c)}$, and assume that (2.19) holds true. Then, from the monotonicity property of the CDF associated with the test statistic Λ_c, which follows $\chi^2(p_c, \lambda_c)$, we have the following inequality

$$\Pr\left[\Lambda_c \leq \tau_c\right] \geq \Pr\left[\Lambda_c \leq \mu_c + \sigma_c\sqrt{2\ln\left(\frac{1}{1-(1-P^D_{\min})^N}\right)} + 2\ln\left(\frac{1}{1-(1-P^D_{\min})^N}\right)\right].$$

From the inequality (2.41a), it now follows that

$$\begin{aligned}\Pr\left[\Lambda_c \leq \tau_c\right] &\geq 1 - \exp\left(-\ln\left(\frac{1}{1-(1-P^D_{\min})^N}\right)\right)\\ &= 1 - \exp\left(\ln\left(1-\left(1-P^D_{\min}\right)^N\right)\right) \geq \prod_{i=1}^{N}\left(1-P^D_i\right) = 1 - \underbrace{\left[1-\prod_{i=1}^{N}\left(1-P^D_i\right)\right]}_{P^D_d}.\end{aligned}$$

The result follows by substituting $P^D_c = 1 - \Pr\left[\Lambda_c \leq \tau_c | H_1\right]$ in the above inequality. $\square$

Proof of Theorem 9 By recursively expanding the equality constraint of the optimization problem (P.2) we have $\begin{bmatrix}\overline{x}(1) & \dots & \overline{x}(T)\end{bmatrix} = \mathcal{A}x(0) + \mathcal{B}_a U^a$. From this identity, (P.2) can also be expressed as

$$\begin{aligned}&\max_{U^a} && \underbrace{\left[\mathcal{A}x(0) + \mathcal{B}_a U^a\right]^{\mathsf{T}}\left[\mathcal{A}x(0) + \mathcal{B}_a U^a\right]}_{f(U^a)}\\ &\text{subject to} && (U^a)^{\mathsf{T}} M_c (U^a) \leq \widetilde{\delta}_c.\end{aligned}$$

From the first-order necessary conditions [16] we now have

$$\nabla \left(f(U_c^*) - \gamma (U_c^*)^{\mathsf{T}} M_c (U_c^*) \right) = 0, \tag{2.36a}$$

$$\gamma \left(\widetilde{\delta}_c - (U_c^*)^{\mathsf{T}} M_c (U_c^*) \right) = 0, \tag{2.36b}$$

$$\gamma \geq 0, \tag{2.36c}$$

$$(U_c^*)^{\mathsf{T}} M_c (U_c^*) \leq \widetilde{\delta}_c, \tag{2.36d}$$

where the gradient ∇ is with respect to U^a.

Case (i): Suppose $(U_c^*)^{\mathsf{T}} M_c (U_c^*) < \widetilde{\delta}_c$. Then $\gamma = 0$ should hold true to ensure the complementarity slackness condition (2.36b). Using these observations in the KKT conditions we now have $\nabla f(U_c^*) = 0$. Further, since, $f(U^a)$ is a convex function of U^a, by evaluating the second derivative of $f(U^a)$ at U_c^*, it can be easily seen that the obtained U_c^* results in minimum value of (P.2) rather than the maximum.Thus, for any U_c^* of (P.2), the condition $(U_c^*)^{\mathsf{T}} M_c (U_c^*) < \widetilde{\delta}_c$ cannot hold true.

Case (ii): Suppose $(U_c^*)^{\mathsf{T}} M_c (U_c^*) = \widetilde{\delta}_c$. Then the KKT conditions can be simplified as:

$$\nabla \left(f(U_c^*) - \gamma (U_c^*)^{\mathsf{T}} M_c (U_c^*) \right) = 0,$$

$$(U_c^*)^{\mathsf{T}} M_c (U_c^*) = \widetilde{\delta}_c.$$

The result follows by evaluating the derivative on the left hand side of the first equality. □

Proof of Lemma 10: By substituting $x(0) = 0$ in (2.21a), we note that any optimal attack U_c^* is of the form $k\nu$, where ν is the generalized eigenvector of the pair $(\mathcal{B}_a^{\mathsf{T}} \mathcal{B}, M_c)$ [18], and the scalar $k = \sqrt{\widetilde{\delta}_c / \nu^{\mathsf{T}} M_c \nu}$ is obtained from (2.21b). Let J_c be the optimal cost associated with an attack of the form $U_c^* = k\nu$. Then,

$$J_c = (k\nu)^\mathsf{T} \mathcal{B}_a^\mathsf{T} \mathcal{B}_a (k\nu) = \gamma (k\nu)^\mathsf{T} M_c (k\nu) = \gamma \widetilde{\delta}_c,$$

where the first equality follows from the fact that the objective function $\sum_{k=1}^{T} \overline{x}^\mathsf{T}(k)\overline{x}(k)$ in (P.2) can be expressed as $(U_c^*)^\mathsf{T} \mathcal{B}_a^\mathsf{T} \mathcal{B}_a (U_c^*)$, and the second equality follows from (2.21a). Since ν is a generalized eigenvector of the pair $(\mathcal{B}_a^\mathsf{T} \mathcal{B}, M_c)$, it follows that γ is the eigenvalue corresponding to ν and hence, J_c is maximized when γ is maximum, which is obtained for $v = v^\star$. The result follows since, $\gamma = \rho_{\max}$, for $v = v^\star$. □

Proposition 13 *Let $\mathcal{O}_i$ $\mathcal{F}_i^{(u)}$ be the observability and impulse response matrices defined in (2.6). Define the matrices $\mathcal{O}_{c,i} = \begin{bmatrix} (C_{c,i}A)^\mathsf{T} & \cdots & \left(C_{c,i}A^T\right)^\mathsf{T} \end{bmatrix}^\mathsf{T}$ and $C_{c,i} = \begin{bmatrix} 0 & \cdots & C_i & \cdots & 0 \end{bmatrix}$, and the following left null spaces:*

$$\mathcal{N}_i^L = \left\{ z : z^\mathsf{T} \begin{bmatrix} \mathcal{O}_i & \mathcal{F}_i^{(u)} \end{bmatrix} = 0^\mathsf{T} \right\}, \mathcal{N}_{c,i}^L = \left\{ z : z^\mathsf{T} \mathcal{O}_{c,i} = 0^\mathsf{T} \right\}, \text{ and } \mathcal{N}_c^L = \bigcup_{i=1}^{N} \mathcal{N}_{c,i}^L. \quad (2.37)$$

Then, $\mathcal{N}_i^L \subseteq \mathcal{N}_{c,i}^L \subseteq \mathcal{N}_c^L$, for all $i \in \{1, \ldots, N\}$.

Proof. Without loss of generality, let $i = 1$. By definition, the set inclusion $\mathcal{N}_{c,1}^L \subseteq \mathcal{N}_c^L$ is trivial. For the other inclusion, consider the system defined in (2.3) without the attack and noise, i.e., $x(k+1) = Ax(k)$. Let $x(k) = \begin{bmatrix} x_1^\mathsf{T}(k) & u_1^\mathsf{T}(k) \end{bmatrix}^\mathsf{T}$, where $x_1(k)$ and $u_1(k)$ are the state and the interconnection signal of Subsystem 1. Also, let

$$A = \begin{bmatrix} A_{11} & B_1 \\ \widetilde{B}_1 & \widetilde{A}_{11} \end{bmatrix}. \quad (2.38)$$

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

$$\begin{aligned} x_1(k+1) &= A_{11}x_1(k) + B_1u_1(k), \\ u_1(k+1) &= \widetilde{A}_{11}u_1(k) + \widetilde{B}_1x_1(k). \end{aligned} \quad (2.39)$$

By letting $\widetilde{C}_1 = \begin{bmatrix} C_1 A_{11} & C_1 B_1 \end{bmatrix}$ and recursively expanding $x_1(k)$ using (2.39), we have

$$\begin{aligned} C_{c,1} A^k x(0) &= \begin{bmatrix} C_1 & 0 \end{bmatrix} A A^{k-1} x(0) = \widetilde{C}_1 A^{k-1} x(0) \\ &= \widetilde{C}_1 \begin{bmatrix} x_1(k-1) \\ \\ u_1(k-1) \end{bmatrix} = \widetilde{C}_1 \begin{bmatrix} A_{11}^{k-1} x_1(0) + \sum_{j=0}^{k-2} A_{11}^{k-2-j} B_1 u_1(j) \\ \\ u_1(k-1) \end{bmatrix} \\ &= C_1 A_{11}^k x_1(0) + \sum_{j=0}^{k-1} C_1 A_{11}^{k-1-j} B_1 u_1(j), \end{aligned} \tag{2.40}$$

where the second, third, and fourth equalities follows from (2.38), system $x(k+1) = Ax(k)$, and (2.39), respectively. By recalling that $\mathcal{O}_{c,1} x(0) = \begin{bmatrix} (C_{c,1} A)^\mathsf{T} & \cdots \left(C_{c,1} A^T\right)^\mathsf{T} \end{bmatrix}^\mathsf{T} x(0)$, it follows from (2.40) that

$$\mathcal{O}_{c,1} x(0) = \mathcal{O}_1 x_1(0) + \mathcal{F}_1^{(u)} \begin{bmatrix} u_1^\mathsf{T}(0) & \cdots & u_1^\mathsf{T}(T-1) \end{bmatrix}^\mathsf{T}.$$

Let z be any vector such that $z^\mathsf{T} \begin{bmatrix} \mathcal{O}_1 & \mathcal{F}_1^{(u)} \end{bmatrix} = 0^\mathsf{T}$. Then, z also satisfies $z^\mathsf{T} \mathcal{O}_{c,1} = 0^\mathsf{T}$. Thus, $\mathcal{N}_1^L \subseteq \mathcal{N}_{c,1}^L$. ■

Lemma 14 ***(Upper bound on P_d^D)*** *Let p_i and λ_i be defined as in* (2.14)*, and τ_i be defined as in* (2.11)*. Let $p_{sum} = \sum_{i=1}^N p_i$, $\lambda_{sum} = \sum_{i=1}^N \lambda_i$, and $\tau_{min} = \min_{1 \le i \le N} \tau_i$. Then, $P_d^D \le \underbrace{\Pr\left[S_d > \tau_{min}\right]}_{\triangleq \overline{P}_d^D}$, where $S_d \sim \chi^2(p_{sum}, \lambda_{sum})$.*

Proof. Consider the following events:

$$\begin{aligned} \mathcal{V}_i &= \left\{ \widetilde{Y}_i^\mathsf{T} \Sigma_i^{-1} \widetilde{Y}_i \ge \tau_i \right\} \text{ for all } i \in \{1, \ldots, N\}, \text{ and} \\ \mathcal{V} &= \left\{ \sum_{i=1}^N \widetilde{Y}_i^\mathsf{T} \Sigma_i^{-1} \widetilde{Y}_i \ge \tau_{\min} \right\}, \end{aligned}$$

where the event $\mathcal{V}_i$ is associated with the $i-$th local detector's threshold test. From the definition of the above events, it is easy to note that $\bigcup_{i=1}^N \mathcal{V}_i \subseteq \mathcal{V}$. By the monotonicity of the probability measures, it follows that

$$P_d^D \triangleq \Pr\left[\bigcup_{i=1}^{N} \mathcal{V}_i \mid H_1\right] \leq \Pr\left[\mathcal{V} \mid H_1\right].$$

From the reproducibility property of the noncentral chi-squared distribution [36], it follows that $\sum_{i=1}^{N} \widetilde{Y}_i^{\mathsf{T}} \Sigma_i^{-1} \widetilde{Y}_i$ equals S_d in distribution and hence, $\Pr[\mathcal{V}|H_1] = \Pr[S_d > \tau_{\min}]$. ■

Lemma 15 ***(Exponential bounds on the tails of*** $\chi^2(p, \lambda)$***)*** *Let* $Y \sim \chi^2(p, \lambda)$, $\mu = p + \lambda$, $\sigma = \sqrt{2(p + 2\lambda)}$. *For all* $x > 0$,

$$\Pr\left[Y \geq \mu + \sigma\sqrt{2x} + 2x\right] \leq \exp(-x) \tag{2.41a}$$

$$\Pr\left[Y \leq \mu - \sigma\sqrt{2x}\right] \leq \exp(-x) \tag{2.41b}$$

Proof. See [6]. ■

Chapter 3

A Probabilistic Approach to Design Switching Attacks against Interconnected Systems

3.1 Introduction

In this chapter, we study a security problem for interconnected systems, where the objective of the attacker is to compromise the performance of the system by tampering with the individual subsystems, while maintaining undetectability. In particular, we develop a probabilistic rule to randomly select an attacked subsystem over time, and optimize over the switching probabilities to maximize degradation and maintain undetectability from a centralized detector. Overall, our results show that the ability to selectively compromise different parts of a system over time greatly increases the severity of the attacks, thereby motivating the development of advanced detection schemes for interconnected system [2].

Main Contributions: This work features three main contributions. First, we develop an attack model which randomly, through some pre-assigned probabilistic rule, compromise a subsystem. Second, we characterize the detection probability of a centralized detector, with respect to these attacks, and, derive upper bounds on the detection probabilities, both in the finite and asymptotic cases. Third, we formulate and numerically solve an optimization problem for computing optimal probabilistic rules with constraints on the detection probability. Finally, we demonstrate the superiority of using our optimal probabilistic strategy against attacking fixed subsystem strategy on a real time example.

Related Work: In the last few years, with security emerging as a major concern for real time dynamical systems, different attack models and possible remedial frameworks have been studied by researchers to a great extent [49, 77, 76, 28, 61, 56]. Although, these works provide deep insights into the attackers capabilities in compromising systems, several of these works mainly restrict their attention to the attacks that target fixed subparts or the overall system. Thereby undermining the vulnerabilities posed by the interconnected systems, at various subsystem and interconnection levels. However, only recently, researchers started to study attack models in the context of interconnected systems, of which, the switching attack locations model have received considerable attention. A few notable works in this direction are as follows: Exploiting the sparsity structure in deterministic systems, authors in [45] proposed dynamic decoders to estimate the initial state accurately. Instead, for the stochastic systems, several authors proposed robust state estimation techniques exploiting the tools from framework hidden mode switching systems and hidden Markov models [91, 39, 72]. Using the variable structure systems theory, authors in [47] demonstrated switching attacks that can disrupt operation of power grid within a short period of time. Instead, authors in [25] considered game-theoretic approach based power system stabilizers to counter attack switching attacks in smart grids. Finally, authors in [88, 2] studied the detrimental effects of switching and coordinated type of integrity attacks in general cyber-physical systems.

3.2 Problem setup and preliminary notions

3.2.1 Nominal system model

We consider an interconnected system composed of N interacting subsystems whose dynamics are as follows:

$$\begin{aligned} x_i(k+1) &= A_{ii}x_i(k) + \sum_{j\neq i}^{N} A_{ij}x_j(k) + w_i(k), \\ y_i(k) &= C_i x_i(k) + v_i(k), \end{aligned} \tag{3.1}$$

where $x_i \in \mathbb{R}^{n_i}$ and $y_i \in \mathbb{R}^{m_i}$ is the state and measurements of the i-th subsystem, and $w_i \sim W_i$, $v_i \sim \mathcal{N}(0, V_i)$ are the process and measurement noise affecting the i-th subsystem dynamics. Let $n = \sum_{i=1}^{N} n_i$ and $m = \sum_{i=1}^{N} m_i$. In the vector form, the dynamics read as

$$\begin{aligned} x(k+1) &= Ax(k) + w(k), \\ y(k) &= Cx(k) + v(k), \end{aligned} \tag{3.2}$$

where the state x, measurements y, and the noise vectors w and v of the interconnected system are given by $x = \begin{bmatrix} x_1^\mathsf{T} & \dots & x_N^\mathsf{T} \end{bmatrix}^\mathsf{T}$, $y = \begin{bmatrix} y_1^\mathsf{T} & \dots & y_N^\mathsf{T} \end{bmatrix}^\mathsf{T}$, $w = \begin{bmatrix} w_1^\mathsf{T} & \dots & w_N^\mathsf{T} \end{bmatrix}^\mathsf{T} \in$

$\mathbb{R}^n$, and $v = \begin{bmatrix} v_1^\mathsf{T} & \dots & v_N^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^m$. Further, the dynamical matrices A and C are given by

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{N1} & \cdots & A_{NN} \end{bmatrix} \text{ and } C = \begin{bmatrix} C_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & C_N \end{bmatrix},$$

respectively. The initial state $x(0) \sim \mathcal{N}(0, \Sigma_0)$, the noises $w \sim \mathcal{N}(0, W)$ and $v \sim \mathcal{N}(0, V)$ are uncorrelated, for all $k \in \mathbb{N}$, where the noise covariance matrices are $W \triangleq \text{blkdiag}\,(W_1, \cdots, W_N)$ and $V \triangleq \text{blkdiag}\,(V_1, \cdots, V_N)$.

We assume that (3.2) is operating in steady state and, we allow for the presence of attackers that compromise the dynamics of the subsystems, and we model such attacks as exogenous unknown inputs (see Section 3.2.2). We task an interconnected system with a detector whose role is to trigger an alarm, based on the innovations signals generated by a Kalman filter. Under the assumption that (C, A) is observable and (A, W) is controllable, a steady Kalman filter employs the following recursion:

$$\begin{aligned} \widehat{x}(k) &= A\widehat{x}(k-1) + Kz_k, \\ z(k) &= y(k) - CA\widehat{x}(k-1), \end{aligned} \tag{3.3}$$

where $\widehat{x}(k) \triangleq \mathbb{E}[x(k)|y(0), \dots, y(k)]$ is the MMSE estimate of the state $x(k)$, and the matrices $K \triangleq PC^T[CPC^T + V]^{-1}$ and $P \triangleq A(I - KC)PA^T + W$ are the steady state Kalman gain and the error covariance matrix, respectively. Further, the innovations $z(k) \sim \mathcal{N}(0, \Sigma)$ forms an i.i.d sequence with covariance $\Sigma \triangleq CPC^T + V$.

3.2.2 Objectives of attacker and attacked system model

We assume that the main objective of an attacker is to inject malicious inputs into the system (3.2) with the following objectives:

1. at any given time, an attacker selects a subsystem, with a probabilistic rule, to inject a malicious inputs and,
2. the rule used in (i) should maximize the state deviation of (3.2) from the origin, and, should result in minimum detection probability.

Our motivation to consider this kind of objective stems from the following reasoning: often in practice, attackers need to compromise systems with limited amount of resources and, ingenuous attackers might look for some clever mechanisms to tamper systems such that the degradation of system performance is worse even with the limited resources.

Let $\{a_k\}_{k=0}^{\infty}$ be a scalar valued i.i.d stochastic process, where a_k takes value in the finite set $\{1, \ldots, N\}$, at every time $k \in \mathbb{N}$, with probability $\mathbb{P}[a_k = i] \triangleq p_i$, for all $i \in \{1, \ldots, N\}$, such that $\sum_{i=1}^{N} p_i = 1$. Let $p \triangleq [p_1, \ldots, p_N]^{\mathsf{T}}$, and note that p denotes the probabilities of selecting subsystems. Thus, by specifying p, the attack process $\{a_k\}_{k=0}^{\infty}$, realizes a subsystem index, for any given time k. Let $\delta_i(a_k)$ be a indicator random variable of a_k, i.e., $\delta_i(a_k) = 1$ if $a_k = 1$, else $\delta_i(a_k) = 0$ otherwise. Let $\delta(a_k) \triangleq [\delta_1(a_k), \ldots, \delta_N(a_k)]^{\mathsf{T}}$. Then, the attacked system dynamics can be modeled as

$$\begin{aligned} x^e(k+1) &= Ax^e(k) + w(k) + \Pi(k)\delta(a_k), \\ y^e(k) &= Cx^e(k) + v(k) + \Psi(k)\delta(a_k), \end{aligned} \tag{3.4}$$

where, $x^e(k)$ and $y^e(k)$ denote the state and the measurement of the system under attack. The attack matrices are $\Pi \triangleq \text{blkdiag}\,(\Pi_1 u_1, \ldots, \Pi_N u_N)$ and $\Psi \triangleq \text{blkdiag}\,(\Psi_1 \tilde{u}_1, \ldots, \Psi_N \tilde{u}_N)$, respectively, where $\Pi_i u_i(k)$ and $\Psi_i \tilde{u}(k)$ are the malicious inputs that an attacker wants to inject into the $i-$th subsystem at time k. Further, we assume that the process $\{a_k\}_{k=0}^{\infty}$ is independent of $w(k)$ and $v(k)$. Thus it follows that the random variables $\delta(a_k)$, $w(k)$, and $v(k)$ are mutually independent, for all $k \in \mathbb{N}$. Let $P^D(k)$ be the detection probability of a detector. Then, the attacker's objective can be casted as a following optimization problem:

$$\begin{aligned}
\textbf{(P.1)} \quad & \arg\max_{p} && \mathbb{E}\left[\sum_{k=0}^{T-1} x^e(k+1)^{\mathsf{T}} x^e(k+1)\right], \\
& \text{subject to} && p \geq 0, && (3.5)\\
& && \mathbf{1}^{\mathsf{T}} p = 1, \text{ and} && (3.6)\\
& && P^D(k) \leq \zeta \quad \forall k \in \{0, \dots, T-1\}, && (3.7)
\end{aligned}$$

where the expectation $\mathbb{E}[\cdot]$ is taken over the noise variables and the process $\{a_k\}_{k=0}^{T-1}$. We make the following assumption for an attacker to design an optimal probabilistic rule.

Assumption 16 *The attacker has complete information about the matrices of the system* (3.2) *and of the Kalman filter* (3.3).

3.2.3 Relation between nominal and attacked system

In this section we characterize the bias accumulated in the interconnected system dynamics (3.2) and the Kalman filter dynamics (3.3) due to the attacks. Let $\gamma(k)$ and $\beta(k)$ denote the bias in the state and the measurements of (3.2), respectively. Then, $x^e(k) = x(k) + \gamma(k)$ and $y^e(k) = y(k) + \beta(k)$, where

$$\begin{aligned}
\gamma(k+1) &= A\gamma(k) + \Pi(k)\delta(a_k), \\
\beta(k) &= C\gamma(k) + \Psi(k)\delta(a_k).
\end{aligned} \tag{3.8}$$

To compute the bias in the state estimate and innovations, we consider the following filter under attacks

$$\begin{aligned} \widehat{x}^e(k) &= A\widehat{x}^e(k-1) + Kz^e(k), \\ z^e(k) &= y^e(k) - CA\widehat{x}^e(k-1), \end{aligned} \tag{3.9}$$

where $\widehat{x}_k^e$ and z_k^e are analogous to the state estimate and the innovations defined in (3.3). Let $\widehat{x}^e(k) = \widehat{x}(k) + \alpha(k+1)$ and $z^e(k) = z(k) + \epsilon(k)$. Then the biases $\alpha(k)$ and $\epsilon(k)$ can be obtained from the following linear system:

$$\begin{aligned} \alpha(k+1) &= (I - KC)A\alpha(k) + K\beta(k), \\ \epsilon(k) &= C\left[\gamma(k) - A\alpha(k)\right] + \Psi(k)\delta(a_k), \end{aligned} \tag{3.10}$$

Notice that in the absence of attacks, the bias $\epsilon(k) = 0$, and $z^e(k) = z(k)$ for all $k \in \mathbb{N}$. Instead, in the presence of attacks, the bias $\epsilon(k) \neq 0$ and $z^e(k) \neq z(k)$ (at least for one k). Based on this observation, in the following section we develop a procedure that will be employed by a detector to decide against the attacks.

3.3 Detection framework

Let H_0 and H_1 be the null and the alternative hypothesis corresponding to the presence and absence of attacks, respectively. The attack detection problem can be casted using the following hypothesis testing framework:

$$H_0 \text{ (attack absent)} : \mathbb{E}\left[z^e(k)\right] = 0,$$

$$H_1 \text{ (attack present)} : \mathbb{E}\left[z^e(k)\right] \neq 0.$$

We assume that the detector uses a chi-squared test statistic [54, 15], which is a quadratic transformation of the innovations $z^e(k)$, to compare with a threshold and decide against the attacks. Formally, we have the following test procedure for all $k \in \mathbb{N}$:

$$\Lambda(k) \triangleq z^e(k)^\mathsf{T} \Sigma^{-1} z^e(k) \underset{H_0}{\overset{H_1}{\gtrless}} \tau, \tag{3.11}$$

where τ is a suitable threshold. The false alarm probability (P^F) and the detection probability (P^D) of the test (3.11) are defined in the following way:

$$P^F(k) \triangleq \mathbb{P}\left[\Lambda(k) \geq \tau | H_0\right] \text{ and } P^D(k) \triangleq \mathbb{P}\left[\Lambda(k) \geq \tau | H_1\right].$$

We assume that $P^F(k)$ is identical for all $k \in \mathbb{N}$ and, thus we omit its dependence on time and refer to it as P^F. Now, by recalling that under the null hypothesis H_0 the bias $\epsilon(k) = 0$, we have $z^e(k) \sim \mathcal{N}(0, \Sigma)$. It follows that $\Lambda(k) \sim \chi^2(m)$, where m is the degrees of freedom. Further, we assume that P^F is predetermined and the threshold τ is computed by the inverse CDF of $\chi^2(m)$.

3.3.1 Characterization of the detection probability

Notice that the detector cannot compute the detection probability $P^D(k)$, as it does not know the bias $\epsilon(k)$ accumulated in the innovations $z^e(k)$. Although, this is a limitation of the detector, it is not very beneficial to an attacker either, since, if the magnitude of attack input is large, larger is the bias $\epsilon(k)$, greater the test statistic $\Lambda(k)$, and hence easier for the attack to satisfy (3.11). Instead, in order to inject attacks that can evade the detector, i.e., bypass the threshold test (3.11), the attacker needs to know $P^D(k)$. In this section we derive an expression for $P^D(k)$, which helps attacker decide what type of attacks to be casted on the interconnected system. We now state a proposition that expresses the bias $\epsilon(k)$ in terms of attack inputs until time k, for all $k \in \mathbb{N}$.

Proposition 17 *Let $\mathcal{A} = A\,(I - KC)$ and $\mathcal{B}(k) = \Pi(k) - AK\Psi(k)$. Then*

$$\epsilon(k) = \underbrace{\begin{bmatrix} C\mathcal{A}^{k-1}\mathcal{B}(0) & \dots & C\mathcal{B}(k-1) & \Psi(k) \end{bmatrix}}_{\triangleq \mathcal{E}_k} \underbrace{\begin{bmatrix} \delta(a_0) \\ \vdots \\ \delta(a_k) \end{bmatrix}}_{\triangleq \delta(a_{0:k})}. \tag{3.12}$$

Proof. *See the Appendix.* ■

We now introduce the following notation that will be useful for characterizing detection probability $P^D(k)$. Consider the truncation $\{a_j\}_{j=0}^{k}$ of the actual attack process $\{a_j\}_{j=0}^{\infty}$. Let $\mathcal{S}_k$ be the set of all possible realizations of $\{a_j\}_{j=0}^{k}$, and π_k be an element of $\mathcal{S}_k$, where the components of π_k can be enumerated as $[\pi_k^0, \pi_k^1 \dots, \pi_k^k]$. With slight abuse of notation we define $\delta(\pi_k) \triangleq [\delta_1(\pi_k^0)^\mathsf{T}, \dots, \delta_N(\pi_k^k)^\mathsf{T}]^\mathsf{T}$. To avoid confusion, we emphasize that $\delta(a_{0:k})$ is a random vector but $\delta(\pi_k)$ is a deterministic (realized) vector.

Lemma 18 ***(Detection probability)*** *The detection probability of the test statistic* (3.11) *is given by*

$$P^D(k) = \sum_{\pi_k \in \mathcal{S}_k} Q(\tau; m, \lambda(\pi_k)) p_{\pi_k^0} p_{\pi_k^1} \cdots p_{\pi_k^k}, \tag{3.13}$$

where $Q(\tau; r, \lambda(\pi_k))$ is the complementary CDF of $\chi^2(\tau, \lambda(\pi_k))$, $\lambda(\pi_k) \triangleq \widetilde{\delta}(\pi_k)^\mathsf{T}\mathcal{E}_k^\mathsf{T}\Sigma^{-1}\mathcal{E}_k\widetilde{\delta}(\pi_k)$.

Proof. See the Appendix. ■

For the attacks that randomly select a subsystem, Lemma 18, states that the $P^D(k)$ is a weighted sum of detection probability, i.e., $Q(\tau; m, \lambda(\pi_k))$, associated with all possible ways of selecting the locations. Moreover, for a given k, these weights are nothing but the binomial coefficients of the expansion $(p_1 + \ldots + p_N)^k$. Finally, notice that the expression (3.13) depends on the matrices of the interconnected system and the KF through the impulse response $\mathcal{E}_k$ in (3.12) of $\lambda(\pi_k) = \widetilde{\delta}(\pi_k)^{\mathsf{T}}\mathcal{E}_k^{\mathsf{T}}\Sigma^{-1}\mathcal{E}_k\widetilde{\delta}(\pi_k)$. Hence, the following assumption ensures that an attacker has the capability to compute $P^D(k)$ for $k \in \mathbb{N}$.

3.3.2 Upper bound on the detection probability

Although the formula of $P^D(k)$ we obtained in Lemma 18 is exact, the number of summands in (3.13) increases exponentially with time k. Hence, for practical purposes, computing the detection probability using (3.13) is not efficient. In this section we provide an upper bound on $P^D(k)$ using Markov's inequality. We now define the following matrices that will be helpful in expressing our bound compactly:

$$\mathcal{E}_k^{\mathsf{T}}\Sigma^{-1}\mathcal{E}_k \triangleq \begin{bmatrix} L_k(0,0) & L_k(0,1) & \cdots & L_k(0,k) \\ L_k(1,0) & L_k(1,1) & \cdots & L_k(1,k) \\ \vdots & \vdots & \ddots & \vdots \\ L_k(k,0) & L_k(k,1) & \dots & L_k(k,k) \end{bmatrix}, \tag{3.14}$$

where $L_k(i,j)$, $0 \leq i, j \leq k$ is obtained by performing block wise multiplication of matrices in $\mathcal{E}_k^{\mathsf{T}}$ with those in $\Sigma^{-1}\mathcal{E}_k$. Moreover, this construction results in $L_k(i,j) = L_k(j,i)^{\mathsf{T}}$, for all i, j. Further, define $\overline{L}_k$ and $\widehat{L}_k$ as

$$\overline{L}_k \triangleq \sum_{i=j} L_k(i,j) \text{ and } \widehat{L}_k \triangleq \sum_{i\neq j} L_k(i,j), \tag{3.15}$$

respectively. Further, $\overline{L}_k$ is a positive semi definite matrix, while $\widehat{L}_k$ is a symmetric matrix.

Lemma 19 ***(Upper bound of the detection probability)*** *Let $p \triangleq [p_1, p_2, \dots, p_N]^{\mathsf{T}}$ be the vector of probabilities with p_i denoting the probability of attacking an $i-th$ subsystem,*

$\forall i \in \{1, \ldots, N\}$. *Then, for all $k \in \mathbb{N}$, it holds that*

$$P^D(k) \leq \underbrace{\frac{m + \operatorname{diag}(\overline{L}_k)^\mathsf{T} p + p^\mathsf{T} \widehat{L}_k p}{\tau}}_{\overline{P}^D(k)}. \tag{3.16}$$

Proof. See the Appendix. ■

Notice that, unlike the expression in (3.13), the upper bound $\overline{P}^D(k)$ is a quadratic expression in the probability vector p. Further, $\overline{P}^D(k)$ dose not depend on the Q function, which is an infinite series. Rather it depends directly on the impulse response (3.12) through the matrices $\overline{L}_k$ and $\widehat{L}_k$. Finally, we note that the bound (3.16) becomes loose if the threshold τ is not sufficiently large, or equivalently, for the higher values of P^F.

3.3.3 Asymptotic upper bound

In this section we aim to provide an asymptotic expression for the bound $\overline{P}^D(k)$ when $k \to \infty$. For this purpose, we assume that the attack matrices are constant for all the times, i.e., $\Pi(k) \triangleq \Pi$ and $\Psi(k) \triangleq \Psi$ for all $k \in \mathbb{N}$.

Lemma 20 (Asymptotic upper bound of the detection probability) *Let $\overline{P}^D(k)$ be as in* (3.16)*. Then,*

$$\overline{P}^D_\infty \triangleq \lim_{k \to \infty} P^D(k) = \frac{m + \operatorname{diag}(\overline{L}_\infty)^\mathsf{T} p + p^\mathsf{T} \widehat{L}_\infty p}{\tau}.$$

where $\overline{L}_\infty = \mathcal{B}^\mathsf{T} \mathcal{O} \mathcal{B} + \Psi^\mathsf{T} \Sigma^{-1} \Psi$, $\widehat{L}_\infty = \mathcal{B}^\mathsf{T} \left[\mathcal{O} - \mathcal{M}\right] \mathcal{B} - \Psi \Sigma^{-1} \Psi$, $\mathcal{B} = \Pi - AK\Psi$, $\mathcal{O} \triangleq \sum_{j=0}^{\infty} (\mathcal{A}^j)^\mathsf{T} C^\mathsf{T} \Sigma^{-1} C \mathcal{A}^j$, *and* $\mathcal{M} \triangleq (I - \mathcal{A})^{-\mathsf{T}} C^\mathsf{T} C (I - \mathcal{A})^{-1}$.

Proof. See the Appendix. ■

As $P^D(k) \leq \overline{P}^D(k)$, for all $k \in \mathbb{N}$, from the above Lemma we can observe that for large k, the upper bound of detection probability is constant. Intuitively, what it means is that, if the attacker is not detected during the transience period of the filter (3.3), since the beginning of attacks, then it is unlikely for an attacker to be detected once the filter (3.3) reaches the steady state. Finally, if the attack matrices Π and Ψ are chosen such that the constraint (3.19), i.e., $P^D(k) \leq \zeta$ for $k \in \{0, \ldots, T_0 - 1\}$, where T is sufficiently large, the Lemma 20 guarantees that $P^D(k) \leq \zeta$, for all times $k \in \mathbb{N}$. Thus, this type of asymptotic analysis helps an attacker to carefully chose the attack matrices *a priori* that yields minimum detection probability.

3.4 Design of an optimal probabilistic strategy

In this section we solve the optimization problem (P.1) described in Section 3.2 with the help of numerical optimization techniques. First, we rewrite the cost function of (P.1) in such way that it depends explicitly on the variable p. Under Assumption **??**, consider the following impulse response matrices associated with the system (3.8):

$$\mathcal{H}_{T-1} \triangleq \begin{bmatrix} \Pi & 0 & \cdots & 0 \\ A\Pi & \Pi & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{T-1}\Pi & A^{T-2}\Pi & \ldots & \Pi \end{bmatrix} \quad \text{and}$$

$$\mathcal{H}_{T-1}^{\mathsf{T}}\mathcal{H}_{T-1} \triangleq \begin{bmatrix} G(0,0) & \cdots & G(0,T-1) \\ \vdots & \ddots & \vdots \\ G(T-1,0) & \ldots & G(T-1,T-1) \end{bmatrix}, \tag{3.17}$$

and let

$$\overline{G}_{T-1} \triangleq \sum_{i=j} G(i,j) \text{ and } \widehat{G}_{T-1} \triangleq \sum_{i\neq j} G(i,j). \tag{3.18}$$

By construction, $\overline{G}_{T-1}$ is a positive definite matrix, while $\widehat{G}_{T-1}$ is a symmetric matrix. The following proposition express the cost function of (P.1) in terms of $\overline{G}_{T-1}$ and $\widehat{G}_{T-1}$.

Proposition 21 *The cost function of (P.1) is equivalent to* $\operatorname{diag}(\overline{G}_{T-1})^{\mathsf{T}}p + p^{\mathsf{T}}\widehat{G}_{T-1}p$*.*

Proof. See the Appendix. ■

As $P^D(k)$ is inefficient for computational purposes we relax the constraint (3.7) of (P.1) by replacing it with constraint on the upper bound $\overline{P}^D(k)$. By incorporating the aforementioned changes in (P.1) we now have the following quadratically constrained quadratic programming type problem, whose solution yields a sub-optimal probabilistic attack strategy, with respect to the original problem (P.1).

$$\begin{aligned} \textbf{(P.2)} \quad & \arg\max_{p} && \operatorname{diag}(\overline{G}_{T-1})^{\mathsf{T}}p + p^{\mathsf{T}}\widehat{G}_{T-1}p, \\ & \text{subject to} && \mathbf{1}^{\mathsf{T}}p = 1, p \geq 0 \\ &&& p \geq 0, \\ &&& \operatorname{diag}(\overline{L}_k)^{\mathsf{T}}p + p^{\mathsf{T}}\widehat{L}_k p \leq \tau\,\zeta - m \\ &&& \qquad \forall k \in \{0,\ldots,T-1\} \end{aligned} \tag{3.19}$$

Notice that (P.2) is a non-convex optimization problem, since the matrices $\widehat{L}_k$, for all $k \in 0, \ldots, T-1$, and $\widehat{G}_{T-1}$ are only symmetric matrices. Thus, the standard convex optimization techniques/analysis cannot be applicable. Hence, to obtain a feasible solution to the maximization problem (P.2) we use standard numerical solvers. We also note that this optimal solution might not be a global maximum.

3.4.1 Numerical Example

We consider a chemical reactor consisting of two continuous stirred-tank reactors [35]. The discretized system matrices, with sampling time $T_s = 1$, are given by

$$A_{11} = \begin{bmatrix} 0.2603 & -0.1862 \\ 0.1862 & 0.2603 \end{bmatrix}, A_{12} = \begin{bmatrix} -0.0188 & -0.0230 \\ 0.0232 & -0.00188 \end{bmatrix},$$

$$A_{21} = \begin{bmatrix} -0.0215 & -0.0266 \\ 0.0263 & -0.0215 \end{bmatrix}, A_{12} = \begin{bmatrix} -0.3120 & 0.2713 \\ -0.2713 & -0.3120 \end{bmatrix}.$$

We consider the state and measurement attack matrices as

$$\Pi = \Psi = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}^{\mathsf{T}}.$$

Our results are illustrated in Fig. 3.1 and Fig. 3.2. For the probabilistic rule $p = [0.5, 0.5]^{\mathsf{T}}$ and $P^F = 0.01$, in Fig. 3.1 we report the actual detection probability $P^D(k)$ (3.13) and the upper bound $\overline{P^D}(k)$, for a given system and Kalman filter. As discussed in Section 3.3, we can see that the bound (3.16) converges to a constant when T increases. In Fig. 3.2 we report the values of the cost function (P.2) for the optimal probabilistic rule $p = p*$ and the fixed location rule, i.e., the degenerate probability vectors $p = [1, 0]^{\mathsf{T}}$ and

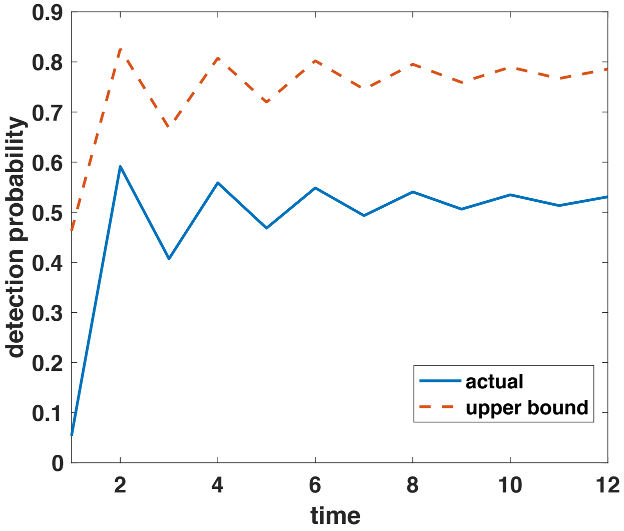


Figure 3.1: This figure shows the detection probability as a function of time. The black solid corresponds to the detection probability evaluated using (3.13). The dashed orange line is obtained using the upper bound (3.16). For the $P^F = 0.01$, degrees of freedom $m = 4$, and the attack matrices Π and Ψ, we notice that although there is an initial transience, due to the dynamics of Kalman filter, as discussed in Section 3.3, the actual value and the bound converges to a constant.

$p = [0, 1]^\mathsf{T}$, respectively. From Fig. 3.2 and as expected, the optimal rule results in higher degradation of the system performance. This work shows that the use of probabilistic rule for switching location attacks benefits attacker, as opposed to attacking fixed locations. As seen in the formulation of (P.2), one can see that, the optimal probabilistic rule depends on the subsystem dynamics, interconnection signals, and the choice of attack matrices. We leave these characterizations for our future research.

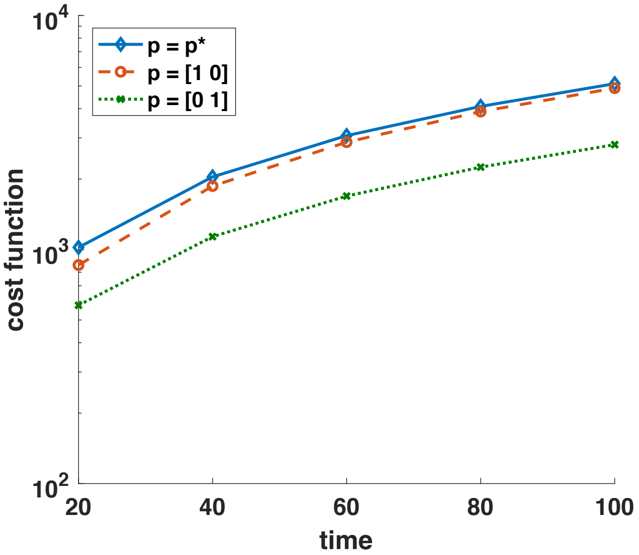


Figure 3.2: This figure shows the performance degradation of interconnected systems, evaluated by the cost function value of (P.2), for various subsystem selecting rules. The blue solid line correspond to the cost function associated with optimal probabilistic rule, that was obtained by solving (P.2) using numerical solver. The dashed orange (resp. dotted green) is obtained by using fixed attack locations. As expected the cost function values for all the rules increase with time horizon. In particular, the optimal probabilistic rule resulted in worst performance degradation than the rest.

3.5 Summary

This paper studies a security problem for interconnected systems, where the attacker objective is to randomly compromise subsystems, such that the performance degradation of interconnected system is maximum. We developed a probabilistic rule for attacking subsystems and characterized the bias accumulated in the system due to these attacks. We also characterized the detection probability of a centralized detector, and formulated an optimization problem to find an optimum probabilistic rule that maximizes system degradation, while maintaining minimum detection probability.

3.6 Appendix

Proof of Proposition 17: Let $\theta(k) \triangleq \gamma(k) - A\alpha(k)$, then from (3.8) and (3.10) it follows that

$$\begin{aligned}\theta(k+1) &= \gamma(k+1) - A\alpha(k+1)\\ &= A\gamma(k) + \Pi(k)\delta(a_k) - A(I-KC)A\alpha(k) - AK\beta(k)\\ &= A\left[\gamma(k) - A\alpha(k)\right] - AKC\left[\gamma(k) - A\alpha(k)\right] + \Pi(k)u(k) - AK\Psi(k)u(k)\\ &= \underbrace{A\left(I-KC\right)}_{\triangleq \mathcal{A}}\theta(k) + \underbrace{\left[\Pi(k) - AK\Psi(k)\right]}_{\triangleq \mathcal{B}(k)}\delta(a_k)\end{aligned}$$

Now, from (3.10), $\epsilon(k)$ can be easily computed by using the following linear system

$$\theta(k+1) = \mathcal{A}\theta(k) + \mathcal{B}(k)\delta(a_k)$$

$$\epsilon(k) = C\theta(k) + \Psi(k)\delta(a_k)$$

The result follows by recursively expanding $\theta(k)$ and observing that $\theta(0) = 0$, since $\gamma(0) = 0$ and $\alpha(0) = 0$. □

Proof of Lemma 18: For any $k \in \mathbb{N}$, let $I_{\{\Lambda(k)\geq\tau\}}$ be the indicator function of the event $\{\Lambda(k) \geq \tau\}$. Then, by using the iterated expectations formula we have

$$P^D(k) = \mathbb{P}\left[\Lambda(k) > \tau | H_1\right] = \mathbb{E}\left[I_{\{\Lambda(k)\geq\tau\}} | H_1\right] = \mathbb{E}\left[\mathbb{E}\left[I_{\{\Lambda(k)\geq\tau\}} \mid H_1, \delta(a_{0:k})\right] \mid H_1\right], \quad (3.20)$$

where the outer expectation is with respect to the truncated process $a_{0:k} \triangleq \{a_j\}_{j=0}^{k}$. Let $\widetilde{\delta}(\pi_k)$ be a realization of $\delta(a_{0:k})$, where $\pi_k = [\pi_k^0, \ldots, \pi_k^k]^\mathsf{T}$. Then, under the hypothesis H_1, we note that $\epsilon(k) = \mathcal{E}_k\widetilde{\delta}(\pi_k)$ is a deterministic quantity and further it follows that

$$z^e(k)|_{H_1,\delta(a_{0:k})=\widetilde{\delta}(\pi_k)} \sim \mathcal{N}(\epsilon(k), \Sigma).$$

Now, from the definition of noncentral chi-squared distribution, we also have

$$\Lambda(k) \,|_{H_1,\delta(a_{0:k})=\widetilde{\delta}(\pi_k)} \sim \chi^2(m, \lambda(\pi_k)),$$

Thus, from the above characterizations and the fact that conditional expectation of an indicator function is equal to the conditional probability, we conclude that

$$\mathbb{E}\left[I_{\{\Lambda(k)\geq\tau\}}\mid H_1, \delta(a_{0:k}) = \widetilde{\delta}(\pi_k)\right] = Q(\tau; m, \lambda(\pi_k)),$$

Substituting above expression in (3.20), and taking the expectation over all realizations of $a_{0:k}$ yields us the following:

$$P^D(k) = \sum_{\pi_k\in\mathcal{S}_k} Q(\tau; m, \lambda(\pi_k))\,\mathbb{P}(a_{0:k} = \pi_k) \tag{3.21}$$

Since, the attack process $\{a_k\}_{k=0}^{\infty}$ is i.i.d it follows that $\mathbb{P}(a_{0:k} = \pi_k) = \prod_{j=0}^{k}\mathbb{P}\left(a_j = \pi_k^j\right) = \prod_{j=0}^{k} p_{\pi_k^j}$. By substituting above expression in (3.21) the statement of Lemma follows. □

Proof of Lemma 19: From the definition of $P^D(k)$ we have

$$P^D(k) = \mathbb{P}[\underbrace{(z^e(k))^{\mathsf{T}}\Sigma^{-1}z^e(k)}_{\Lambda(k)} \geq \tau | H_1].$$

As $\Lambda(k) \geq 0$, from Markov's inequality it follows that

$$\mathbb{P}[(z^e(k))^{\mathsf{T}}\Sigma^{-1}z^e(k) \geq \tau | H_1] \leq \underbrace{\frac{\mathbb{E}\left[(z^e(k))^{\mathsf{T}}\Sigma^{-1}z^e(k) | H_1\right]}{\tau}}_{\overline{P}^D(k)}.$$

Notice that under hypothesis H_1, $z^e(k) = z(k) + \epsilon(k)$ and, from Proposition 17, $\epsilon(k) = \mathcal{E}_k\delta(a_{0:k}) \neq 0$. As the process $\{a_k\}_{k=0}^{\infty}$ is independent of noise random variables affecting the system dynamics, it also follows that $z(k)$ and $\delta(a_{0:k})$ are independent, and we have

$$\begin{aligned}\mathbb{E}\left[(z^e(k))^{\mathsf{T}}\Sigma^{-1}z^e(k)\right] &= \mathbb{E}[z(k)^{\mathsf{T}}\Sigma^{-1}z(k) + 2z(k)^{\mathsf{T}}\Sigma^{-1}\epsilon(k) + \epsilon(k)^{\mathsf{T}}\Sigma^{-1}\epsilon(k)]\\ &\overset{(a)}{=} \mathbb{E}[z(k)^{\mathsf{T}}\Sigma^{-1}z(k) + \epsilon(k)^{\mathsf{T}}\Sigma^{-1}\epsilon(k)], \qquad (3.22)\end{aligned}$$

where (a) follows because $z(k)$ is independent of $\epsilon(k)$ and $\mathbb{E}[z(k)]=0$. Now, consider,

$$\begin{aligned}\mathbb{E}[z(k)^\mathsf{T}\Sigma^{-1}z(k)] &= \operatorname{Tr}\left(\mathbb{E}[z(k)^\mathsf{T}\Sigma^{-1}z(k)]\right)\\ &= \operatorname{Tr}\left(\Sigma^{-1}\mathbb{E}[z(k)z(k)^\mathsf{T}]\right) = \operatorname{Tr}\left(\Sigma^{-1}\Sigma\right) = m \end{aligned} \tag{3.23}$$

where the last equality follows because the innovations $z(k)$ and the measurements $y(k)$ has the same dimension. Before simplifying the second term of (3.22) we note that

$$\mathbb{E}\left[\delta_{a_{0:k}}\delta_{a_{0:k}}^\mathsf{T}\right] = \begin{bmatrix}\mathbb{E}[\delta(a_0)\delta(a_0)^\mathsf{T}] & \dots & \mathbb{E}[\delta(a_0)\delta(a_k)^\mathsf{T}]\\ \vdots & \ddots & \vdots\\ \mathbb{E}[\delta(a_k)\delta(a_0)^\mathsf{T}] & \dots & \mathbb{E}[\delta(a_k)\delta(a_k)^\mathsf{T}]\end{bmatrix} = \begin{bmatrix}\operatorname{diag}(p) & pp^\mathsf{T} & \dots & pp^\mathsf{T}\\ pp^\mathsf{T} & \operatorname{diag}(p) & \dots & pp^\mathsf{T}\\ \vdots & \vdots & \ddots & \vdots\\ pp^\mathsf{T} & pp^\mathsf{T} & \dots & \operatorname{diag}(p)\end{bmatrix}, \tag{3.24}$$

where the second equality follows because the process $\{a_k\}_{k=0}^\infty$ is i.i.d and $\mathbb{E}[\delta(a_k)]=p$, for all $k\in\mathbb{N}$. Further, $p=[p_1,\dots,p_N]^\mathsf{T}$ and, with slight abuse of notation, we denote $\operatorname{diag}(p)$ as the diagonal matrix with diagonal elements as components of p. Consider the following:

$$\begin{aligned}\mathbb{E}[\epsilon(k)^\mathsf{T}\Sigma^{-1}\epsilon(k)] = \operatorname{Tr}\left(\Sigma^{-1}\mathbb{E}[\epsilon(k)\epsilon(k)^\mathsf{T}]\right) &= \operatorname{Tr}\left(\Sigma^{-1}\mathbb{E}[\mathcal{E}_k\delta(a_{0:k})\delta(a_{0:k})^\mathsf{T}\mathcal{E}_k^\mathsf{T}]\right)\\ &= \operatorname{Tr}\left(\mathcal{E}_k^\mathsf{T}\Sigma^{-1}\mathcal{E}_k\mathbb{E}[\delta_{a_{0:k}}\delta_{a_{0:k}}^\mathsf{T}]\right)\end{aligned}$$

By invoking (3.14) and (3.24) in the above expression, and, followed by block multiplication of entries in $\mathcal{E}_k^\mathsf{T}\Sigma^{-1}\mathcal{E}_k$ and $\mathbb{E}[\delta_{a_{0:k}}\delta_{a_{0:k}}^\mathsf{T}]$ we can see that

$$\mathbb{E}[\epsilon(k)^\mathsf{T}\Sigma^{-1}\epsilon(k)] = \operatorname{Tr}\left(\overline{L}_k\operatorname{diag}(p)\right) + \operatorname{Tr}\left(\widehat{L}_kpp^\mathsf{T}\right) = p^\mathsf{T}\operatorname{diag}(\overline{L}_k) + p^\mathsf{T}\widehat{L}_kp \tag{3.25}$$

By substituting (3.23) and (3.25) in $P^D(k)$, the statement of Lemma follows. □

Proof of Lemma 20: From (3.16) we note the following:

$$\lim_{k\to\infty}\overline{P}^D(k) = \frac{m + \lim\limits_{k\to\infty}\operatorname{diag}(\overline{L}_k)^\mathsf{T}p + \lim\limits_{k\to\infty}p^\mathsf{T}\widehat{L}_kp}{\tau} \tag{3.26}$$

Under Assumption **??** and from (3.15) it follows that

$$\begin{aligned}\overline{L}_k &= \sum_{j=0}^{k-1} \mathcal{B}(j)^{\mathsf{T}}(\mathcal{A}^j)^{\mathsf{T}} C^{\mathsf{T}} \Sigma^{-1} C \mathcal{A}^j \mathcal{B}(j) + \Psi(k)^{\mathsf{T}} \Sigma^{-1} \Psi(k),\\ &= \sum_{j=0}^{k-1} \mathcal{B}^{\mathsf{T}}(\mathcal{A}^j)^{\mathsf{T}} C^{\mathsf{T}} \Sigma^{-1} C \mathcal{A}^j \mathcal{B} + \Psi^{\mathsf{T}} \Sigma^{-1} \Psi,\end{aligned}$$

Since $\mathcal{A} = A(I - KC)$ is stable, we note that $\lim\limits_{k\to\infty} \sum_{j=0}^{k-1} (\mathcal{A}^j)^{\mathsf{T}} C^{\mathsf{T}} \Sigma^{-1} C \mathcal{A}^j$ exists. Thus

$$\overline{L}_\infty \triangleq \lim_{k\to\infty} \overline{L}_k = \mathcal{B}^{\mathsf{T}} \mathcal{O} \mathcal{B} + \Psi^{\mathsf{T}} \Sigma^{-1} \Psi \tag{3.27}$$

Now, let $E_k \triangleq \sum_{j=0}^{k-1} C\mathcal{A}^j\mathcal{B}$ and note that $\lim\limits_{k\to\infty} E_k = \lim\limits_{k\to\infty} \sum_{j=0}^{k-1} C\mathcal{A}^j\mathcal{B} = C(I - \mathcal{A})^{-1}\mathcal{B}$, where the last equality because $\mathcal{A}$ is a stable matrix. Moreover, a straightforward computation results in $\widehat{L}(k) = E_k^{\mathsf{T}} E_k - \overline{L}_k$. By taking limits on both sides we have

$$\begin{aligned}\overline{L}_\infty &\triangleq \lim_{k\to\infty} \widehat{L}_k = \lim_{k\to\infty} \left(E_k^{\mathsf{T}} E_k - \overline{L}_k \right)\\ &= \mathcal{B}^{\mathsf{T}} \underbrace{(I - \mathcal{A})^{-\mathsf{T}} C^{\mathsf{T}} C (I - \mathcal{A})^{-1}}_{\triangleq \mathcal{M}} \mathcal{B} - \overline{L}_\infty = \mathcal{B}^{\mathsf{T}} \left[\mathcal{O} - \mathcal{M} \right] \mathcal{B} - \Psi \Sigma^{-1} \Psi \end{aligned} \tag{3.28}$$

Now, by substituting (3.27), and (3.28) in (3.26) it follows that

$$\overline{P}^D_\infty \triangleq \lim_{k\to\infty} \overline{P}^D(k) = \frac{m + \operatorname{diag}(\overline{L}_\infty)^{\mathsf{T}} p + p^{\mathsf{T}} \widehat{L}_\infty p}{\tau}. \quad \square$$

Proof of Proposition 21: Recall that $x^e(k) = x(k) + \gamma(k)$, $\mathbb{E}[x(0)] = 0$ and $x(k)$ is independent of $\gamma(k)$, for all $k \in \mathbb{N}$. Hence,

$$\mathbb{E}\left[\sum_{k=0}^{T-1} x^e(k+1) x^e(k+1) \right] = \sum_{k=0}^{T-1} \mathbb{E}\left[x(k+1) x(k+1) \right] + \mathbb{E}\left[\gamma(k+1) \gamma(k+1) \right]$$

As the first term does not depend on the optimization variable p, for purpose of optimization, we can treat it as a constant. Instead, from (3.8) and (3.17) it follows that

$$\sum_{k=0}^{T-1} \gamma(k+1)^{\mathsf{T}} \gamma(k+1) = \delta(a_{0:T-1})^{\mathsf{T}} \mathcal{H}_{T-1}^{\mathsf{T}} \mathcal{H}_{T-1} \delta(a_{0:T-1})$$

Now, by taking the expectation on both sides of the above equation and following the same procedure as we did in Proof of Lemma 19 (for analyzing $\mathbb{E}[\epsilon(k)^{\mathsf{T}}\Sigma^{-1}\epsilon(k)]$), the statement of Proposition follows. □

Chapter 4

Network Theoretic Analysis of Maximum a Posteriori Detectors for Optimal Sensor Placement

In this chapter we consider a sensor placement problem using the input-output properties of the network dynamics [67, 89, 68, 51]. In contrast to the existing works, which primarily focused on either estimating or reconstructing the network dynamics, in this work, we formulate the placement problem from a detection theoretic view point. In particular, we are concerned with detecting statistical abnormalities in a local stochastic input to a network, using a detector that only relies on the remote time-course measurements of the network dynamics. The basic idea is that the network topological structure enforce relationships between signals at different locations in the network, which means that changes to the dynamics at one location can potentially be detected using measurement signatures at other locations in the network. Our analysis leads to the interesting results that, depending on the network weights and structure, and the intensity of sensor noise, the detection performance may improve as the graphical distance between the input nodes and the sensors location increases. In fact, our results (i) inform the optimal positioning of sensors for the detection of failure of system components or malicious tampering modeled by unknown stochastic inputs, (ii) allow the detection of unexpected modification of the system structure, because such changes would modify the original detection profile, and (iii) provide network design guidelines to facilitate or prevent measurability of certain network signals.

Main Contributions: We briefly summarize our main contributions as follows. First, we consider a binary hypothesis testing problem for a discrete time Gaussian process driving the linear network dynamics through certain network nodes. We primarily consider the scenario where hypothesis on either the mean or the covariance of the input process must be detected using the measurements (possibly corrupted with white Gaussian noise) collected from nodes (referred as to output nodes) that are at least at a specified distance apart from the input nodes. We characterize the *maximum a posteriori* (MAP) detector, and quantify its performance as a function of input-output transfer function matrix gain of the network system. These results are significant in their own rights. In fact, in cyber-physical security literature, there are only limited works related to detecting changes in the covariance of the input, modeled as attacks, and our results contributes to this area as well.

Second, we study the MAP detector's performance as a function of the sensors locations. In particular, in the absence of noise, regardless of the network structure and edge weights, we show that the performance of the detector associated with a set of sensors forming a cut of the network (nodes on the cut shall be referred as to cutset nodes) is as good as the performance obtained by measuring all nodes of the subnetwork identified by the cut and not containing the nodes affected by the input nodes (referred as partitioned set nodes). Conversely, in the presence of noise, depending upon the internal transfer function matrix gain between the cutset nodes and the partitioned nodes, we show that the detection performance of sensors on the cutset nodes may be better or worse than those of sensors on the partitioned nodes. Finally, we demonstrate our theoretical findings on Toeplitz line networks and some illustrative numerical examples.

Related Work: As the optimal sensor placement problem belongs to the class of combinatorial optimization problem, several research communities focused on developing heuristic algorithms and convex relaxation methods [86, 37, 38, 57, 20]. The computational complexity of selecting minimum input and output sets for achieving controllability and observability has been addressed by the authors in [93, 60, 63], and the authors in [60, 81] showed that these are NP-hard problems. A significant amount of work has also been devoted to the observability of linear systems from a graph theoretic view point [17, 12, 50]. In recent years, there is also an incipient research effort to optimize some useful Grammian based metrics in order to ensure the controllability (and dually, observability) [75, 62, 96, 82]. Finally, we also notice the current advancements in partial control of network's dynamics, which are refereed under the names as target controllability or reachability [85, 82].

4.1 Preliminaries and problem setup

Consider a network represented by the digraph $\mathcal{G} := (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \{1, \dots, n\}$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ are the node and edge sets. Let $g_{ij} \in \mathbb{R}$ be the weight assigned to the edge $(i, j) \in \mathcal{E}$, and define the *weighted adjacency matrix* of $\mathcal{G}$ as $G := [g_{ij}]$, where $g_{ij} = 0$ whenever $(i, j) \notin \mathcal{E}$. Let $\mathcal{K} := \{k_1, \dots, k_r\} \subseteq \mathcal{V}$ be the set of input nodes, which receive r inputs. Let $w(i, j)$ denote a path on $\mathcal{G}$ from node i to j, and let $|w(i, j)|$ be the number of edges of $w(i, j)$. Define the distance between input node set $\mathcal{K}$ and a set of nodes $\mathcal{S} \subseteq \mathcal{V}$ as $\text{dist}(\mathcal{K}, \mathcal{S}) := \min\{|w(i, j)| : i \in \mathcal{K}, j \in \mathcal{S}\}$. We associate to each node i a state $x_i \in \mathbb{R}$, and let the network evolve with discrete linear dynamics

$$\mathbf{x}[k+1] = G\mathbf{x}[k] + \Pi\mathbf{w}[k], \tag{4.1}$$

where $\mathbf{x} = [x_1 \cdots x_n]^T \in \mathbb{R}^n$ contains the states of the nodes at time $k \in \mathbb{N}$, $\mathbf{x}[0] \sim \mathcal{N}(\mathbf{0}, \Sigma_0)$ is the initial state, and $\mathbf{w}[k] \in \mathbb{R}^r$ is the input vector. The input matrix $\Pi = [\mathbf{e}_{k_1}, \dots, \mathbf{e}_{k_r}]$ indicates the location of the input nodes. Let the input $\mathbf{w}[k]$ be governed by one of the following two competing statistical hypotheses:

$$\begin{aligned} H_1: \quad & \mathbf{w}[k] \overset{\text{i.i.d}}{\sim} \mathcal{N}(\boldsymbol{\mu}_1, \Sigma_1), \quad k = 0, 1, \dots, N, \\ H_2: \quad & \mathbf{w}[k] \overset{\text{i.i.d}}{\sim} \mathcal{N}(\boldsymbol{\mu}_2, \Sigma_2), \quad k = 0, 1, \dots, N, \end{aligned} \tag{4.2}$$

where the moments $\boldsymbol{\mu}_i \in \mathbb{R}^r$ and $\Sigma_i \in \mathbb{R}^{r \times r} (\succ 0)$, $i \in \{1, 2\}$, are completely known. In other words, the competing hypotheses are simple. However, the true hypothesis is assumed to be unknown. We are concerned with detecting the true hypothesis on the input signal, using measurements from the sensors that are not collocated with the input nodes.

We assume that the nodes $\mathcal{J} := \{j_1, \dots, j_m\} \subseteq \mathcal{V}$ are accessible for sensor placement (one sensor for each node), if $\text{dist}(\mathcal{K}, \mathcal{J}) \geq d$, where $d \in \mathbb{N}$. We refer to $\mathcal{J}$ as the sensor set. The output of these sensors is given by

$$\mathbf{y}_{\mathcal{J}}[k] = C\mathbf{x}[k] + \mathbf{v}[k], \tag{4.3}$$

where $C = [\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_m}]^{\mathsf{T}}$ and $\mathbf{v}[k] \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 I)$. Let the process $\{\mathbf{x}[0], \mathbf{w}[0], \mathbf{v}[0], \mathbf{w}[1], \mathbf{v}[1], \ldots\}$ be uncorrelated. To detect the true hypothesis, we task sensors with a detector, which maps the following time aggregated measurements

$$\mathbf{Y}_{\mathcal{J}}^{\mathsf{T}} = \begin{bmatrix} \mathbf{y}_{\mathcal{J}}^{\mathsf{T}}[1] & \mathbf{y}_{\mathcal{J}}^{\mathsf{T}}[2] & \cdots \mathbf{y}_{\mathcal{J}}^{\mathsf{T}}[N] \end{bmatrix}, \tag{4.4}$$

to a detected hypothesis $\widehat{H}$. We will consider the *maximum a posteriori probability* (MAP) detector, which is given by the following decision rule:

$$\Pr(\{H_2 \,\text{is true}\} | \mathbf{Y}_{\mathcal{J}}) \underset{\widehat{H}=H_1}{\overset{\widehat{H}=H_2}{\gtrless}} \Pr(\{H_1 \,\text{is true}\} | \mathbf{Y}_{\mathcal{J}}). \tag{4.5}$$

For a predetermined set of input nodes $\mathcal{K}$, the focus of our analysis is to characterize the performance of the detector (4.5), in terms of the network's adjacency matrix G. The performance of the detector (4.5) is measured by its error probability, which is given by

$$\mathbb{P}_e(\mathcal{J}) = \sum_{i \in \{1,2\}} \Pr(\widehat{H} \neq H_i | \{H_i \,\text{is true}\}) \pi_i. \tag{4.6}$$

where $\pi_i = \Pr(\{H_i \text{ is true}\})$ is the prior probability.

For any sensor set $\mathcal{J}$ that statisifes $\operatorname{dist}(\mathcal{K}, \mathcal{J}) \geq d$, one expects that the MAP detector's performance (4.6) is maximum when $\operatorname{dist}(\mathcal{K}, \mathcal{J}) = d$. However, for certain network configurations, studies have shown that the gain of transfer function, which is closely related to the *signal-to-noise* ratio (SNR) of a detector, is maximum when the input and output nodes are farther apart [79]. Hence, it remains unclear whether the closeness of the sensors to the input nodes improves the performance of the detector.

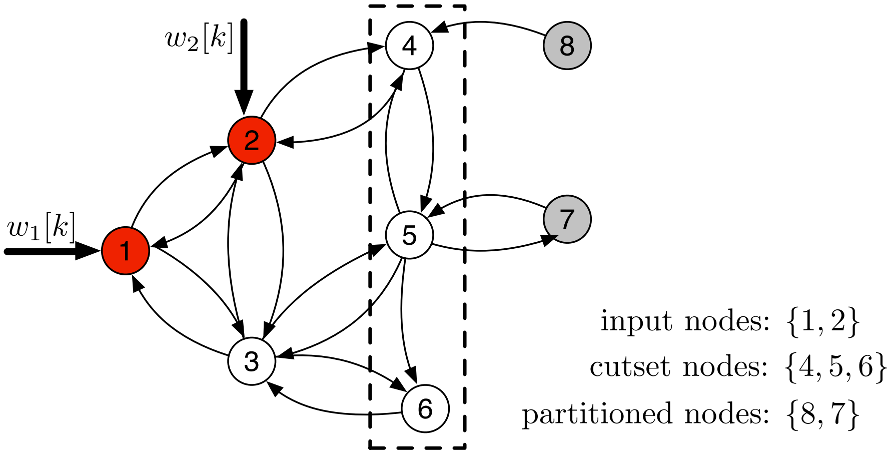


Figure 4.1: Illustration of network partitions induced by a node cutset.

In this paper, we show that the graphical proximity indeed modulate the MAP detector's performance, for certain classes of the detection problems (4.2). In particular, we characterize networks for which the detection performance obtained when sensors are located on a *node cutset* is better (or worse) than the performance obtained when sensors are placed on nodes of the subnetwork induced by the *node cutset* that does not contain the input nodes (precise statements are provided in Section 4.3). See Fig 4.1 for an illustration of *node cutset* and the subnetwork (partitioned nodes) induced by it. Throughout the paper, we will distinguish the performance of sensors with measurement noise $\sigma_v^2 = 0$ and without noise $\sigma_v^2 > 0$. The essential reason, as we will show later, is that only in the presence of noise, we have a scenario where the performance of a MAP detector associated with the cutset nodes may be worse than that of the partitioned nodes. We end this section with an example that shows the impact of noise on the detection performance.

Example 22 *Let $H_i : w[k] \sim \mathcal{N}(\mu_i, \sigma^2)$ and $y[k] = \alpha w[k] + v[k]$, where α is the tuning* *parameter and* $v(k) \sim \mathcal{N}(0, \sigma_v^2)$*. For* $\pi_i = 0.5$*, the MAP detector's error probability, based on N samples of* y*, is* $0.5\, Q_{\mathcal{N}}(0.5\,\eta)$*. Here* $\eta^2 = N(\mu_1 - \mu_2)^2/(\sigma^2 + \alpha^{-2}\sigma_v^2)$ *denotes the SNR, and* $Q_{\mathcal{N}}(0.5\,\eta)$ *is decreasing in* η*. For* $\sigma_v^2 = 0$*, the error probability does not depend on* α*, and is lesser than the case where* $\sigma^2 > 0$*. However, when* $\sigma_v^2 > 0$*,* $Q_{\mathcal{N}}(\cdot)$ *depends on* α*. In particular, when* α *is small,* $Q_{\mathcal{N}}(\cdot)$ *is high, and vice versa. Thus, in the presence of noise, the error probability can be reduced by properly tuning* α*.* □

Remark 23 ***(Non-testable hypotheses and zero dynamics)*** *The hypothesis testing framework considered in this paper may not applicable to the situations such as the system* (4.1) *associated with the measurement model* (4.3) *has zero dynamics, i.e.,* $\mathbf{Y}_{\mathcal{J}}[k] = \mathbf{0}$ *for non-zero input or, some components of the input gets canceled in the operation of* $C\mathbf{x}[k]$*. However, in the latter case, our framework can modified to analyze the input signals that lies in the low dimensional subspace.* □

4.2 Detection performance of the MAP detector

In this section, we derive the algebraic expressions of the MAP decision rules and their error probabilities for two special cases of the hypotheses in (4.2). The first case is the *mean shift model*, in which the covariance matrices in (4.2) are equal, but the mean vectors are different. The second case is the *covariance shift model* in which the mean vectors in (4.2) are equal, but the covariance matrices are different. For this latter case, we will rely on the LD-MAP detector (see below) for deciding the hypothesis. The reason for working with these models is twofold: (i) the error probability expressions are analytically tractable, and (ii) these models are widely used for detection and classification problems that arise in practice [80, 70]. The probability expressions we derive in this section will be used for the network analysis of the MAP detector's performance (details in Section 4.3). Finally, it should be noted that, although we deal with two detection frameworks, all our results will be stated in common theorem environment, i.e., we don't state our results for mean and covariance shift models separately (for instance, see Lemma 27).

Definition 24 ***(Linear discriminant function-based MAP detector: LD-MAP)*** *A LD-MAP detector is as in* (4.5) *with* $\mathbf{Y}_{\mathcal{J}}$ (4.4) *replaced by the discriminant* $y = \mathbf{b}^{\mathsf{T}}\mathbf{Y}_{\mathcal{J}}$*, where the vector* [1] $\mathbf{b} \in \mathbb{R}^{mN}$ *is the maximizer of the information divergence criterion:*

$$I = \pi_1 \mathbb{E}\left[\ln \frac{f_{H_1}(y)}{f_{H_2}(y)}\middle| H_1\right] + \pi_2 \mathbb{E}\left[\ln \frac{f_{H_2}(y)}{f_{H_1}(y)}\middle| H_2\right], \tag{4.7}$$

where $f_{H_i}(y)$ *is the density of* y *given* H_i *and* $I > 0$.

Remark 25 ***(Optimal discriminant vector)*** *For any arbitrary vector* $\mathbf{b}$*, the* I*-divergence measure* (4.7) *indicates how well a LD-MAP detector is performing in deciding between* H_1 *and* H_2*. Thus by maximizing* (4.7)*, we are finding an optimal detector among the class of LD-MAP detectors parameterized by* $\mathbf{b}$ *[70].* □

We now state a lemma that provides us with the algebraic expressions of the MAP detectors associated with the mean and covariance shift models.

Proposition 26 ***(Mean and covariance of*** $\mathbf{Y}_{\mathcal{J}}$***)*** *Let* $\mathbf{Y}_{\mathcal{J}}$ *and* H_i *be defined as in* (4.4) *and* (4.2)*, resp. Then,*

$$\begin{aligned} \overline{\boldsymbol{\mu}}_i &\triangleq \mathbb{E}[\mathbf{Y}_{\mathcal{J}}|H_i] = \mathcal{F}\left(\mathbf{1}_N \otimes \boldsymbol{\mu}_i\right) \;\; \textit{and} \\ \overline{\Sigma}_i &\triangleq \mathit{Cov}[\mathbf{Y}_{\mathcal{J}}|H_i] = \mathcal{O}\Sigma_0\mathcal{O}^{\mathsf{T}} + \mathcal{F}\left(I_N \otimes \Sigma_i\right)\mathcal{F}^{\mathsf{T}} + \sigma_v^2 I, \end{aligned} \tag{4.8}$$

[1] In the literature of pattern recognition and communications, $\mathbf{b}$ is commonly referred as to the Fisher's discriminant and optimal SINR beam former, respectively [24, 8].

where, the observability and impulse response matrices are

$$\mathcal{O}=\begin{bmatrix} CG \\ CG^2 \\ \vdots \\ CG^N \end{bmatrix} \text{ and } \mathcal{F}=\begin{bmatrix} C\Pi & 0 & \dots & 0 \\ CG\Pi & C\Pi & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CG^{N-1}\Pi & CG^{N-2}\Pi & \dots & C\Pi \end{bmatrix}.$$

Lemma 27 ***(MAP detectors)*** *Let π_1 and π_2 be non-zero priors, and define $\gamma = \ln(\pi_1/\pi_2)$. Let $\mathbf{Y}_{\mathcal{J}}$ be as in (4.4), and let $(\boldsymbol{\mu}_i, \Sigma_i)$ and $(\overline{\boldsymbol{\mu}}_i, \overline{\Sigma}_i)$ be as in (4.2) and (4.8), resp.*

1. *The MAP detector associated with the mean shift model ($\Sigma_1 = \Sigma_2$ but $\boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$) is given by:*

$$\left(2\,\overline{\boldsymbol{\mu}}_{\Delta}^{\mathsf{T}}\overline{\Sigma}_c^{-1}\right)\mathbf{Y}_{\mathcal{J}} \underset{\widehat{H}=H_1}{\overset{\widehat{H}=H_2}{\gtrless}} 2\gamma + \overline{\boldsymbol{\mu}}_{\Delta}^{\mathsf{T}}\overline{\Sigma}_c^{-1}\left(\overline{\boldsymbol{\mu}}_1 + \overline{\boldsymbol{\mu}}_2\right), \tag{4.9}$$

where $\overline{\boldsymbol{\mu}}_{\Delta} = \overline{\boldsymbol{\mu}}_2 - \overline{\boldsymbol{\mu}}_1$ and $\overline{\Sigma}_c \triangleq \overline{\Sigma}_1 = \overline{\Sigma}_1$.

2. *The LD-MAP detector associated with the covariance shift model ($\Sigma_1 \neq \Sigma_2$ but $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$) is given by:*

$$\ln\left(\frac{d_1}{d_2}\right) - 2\gamma \underset{\widehat{H}=H_1}{\overset{\widehat{H}=H_2}{\gtrless}} (y - \mathbf{b}^{\mathsf{T}}\overline{\boldsymbol{\mu}}_c)^2\left[\frac{1}{d_2} - \frac{1}{d_1}\right], \tag{4.10}$$

where $y = \mathbf{b}^{\mathsf{T}}\mathbf{Y}_{\mathcal{J}}$, $d_i = \mathbf{b}^{\mathsf{T}}\overline{\Sigma}_i\mathbf{b}$, and $\overline{\boldsymbol{\mu}}_c \triangleq \overline{\boldsymbol{\mu}}_1 = \overline{\boldsymbol{\mu}}_2$.

The detectors (4.9) and (4.10) are functions of the sufficient statistics $2\,\overline{\boldsymbol{\mu}}_{\Delta}^{\mathsf{T}}\overline{\Sigma}_c^{-1}\mathbf{Y}_{\mathcal{J}}$ and $y - \mathbf{b}^{\mathsf{T}}\overline{\boldsymbol{\mu}}_c$, respectively. This means that, given these statistics, other information in $\mathbf{Y}_{\mathcal{J}}$ is not needed for deciding between H_1 and H_2. In order to characterize the error probabilities of the detectors in Lemma 27, we make the following assumption:

Assumption 28 *The LTI system* (4.1) *is stable. Further,*

1. *for the mean shift model,* $\lim_{N\to\infty} N\|\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1\|_2 = c$*, where* $0 < c < \infty$*, and* $G^k = 0$ *for some* $k \in \mathbb{N}$*, and*

2. *for the covariance shift model,* $\Sigma_1 \succ 0$ *and* $\Sigma_2 = 0$*.*

Lemma 29 ***(Error probability: infinite horizon)*** *Let* $\pi_1 = \pi_2 = 0.5$ *and* $\mathbf{x}[0] = 0$*. Let* $T(z) = C(zI - G)^{-1}\Pi$*, where* $z \notin$ *spec*(G)*. The error probability of the MAP detector* (4.9) *and the LD-MAP detector* (4.10) *as* $N \to \infty$ *are*

$$\mathbb{P}_{e_m}(\mathcal{J}) = 0.5\, Q_{\mathcal{N}}(0.5\,\eta) \text{ and} \tag{4.11}$$

$$\mathbb{P}_{e_v}(\mathcal{J}) = 0.5\left[1 - Q_{\chi^2}(1,\tau)\right] + 0.5\, Q_{\chi^2}(1, \tau R), \tag{4.12}$$

respectively, where $\tau = \ln R/(R-1)$*. The SNRs are*

$$\eta^2 = N\widetilde{\boldsymbol{\mu}}_\Delta^{\mathsf{T}}\left([L^{\mathsf{T}}L + \sigma_v^2 I]^{-1}L^{\mathsf{T}}L\right)\widetilde{\boldsymbol{\mu}}_\Delta \text{ and} \tag{4.13}$$

$$R = 1 + \sigma_v^{-2}\|T(z)\Sigma_1^{\frac{1}{2}}\|_\infty^2, \tag{4.14}$$

where $L = T(1)\Sigma_c^{\frac{1}{2}}$ *and* $\widetilde{\boldsymbol{\mu}}_\Delta = \Sigma_c^{-\frac{1}{2}}[\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1]$*, and* $\Sigma_c^{\frac{1}{2}}$ *and* $\Sigma_1^{\frac{1}{2}}$ *are the positive square roots of* Σ_c *and* Σ_1*, respectively.*

The assumptions $\pi_i = 0.5$ and $\mathbf{x}[0] = 0$ are for the ease of presentation, and the probability expressions can be easily adjusted to include other priors and initial conditions. The assumption $N\|\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1\|_2 \to c$ ensures that $\mathbb{P}_{e_m}(\mathcal{J}) < 0.5$. Instead, the assumption $G^k = 0$ is to eliminate the remainder terms in the computation of η. We emphasize that the only restriction on k is that it should be finite, but can be arbitrarily large. We now state a corollary to the above lemma in which we do not assume $\Sigma_2 = 0$ in the covariance shift model (see Remark 32).

Corollary 30 ***(SNRs: identical input statistics)*** *Let H_i in* (4.2) *be* $\mathbf{w}[k] \sim \mathcal{N}(\mu_i \mathbf{1}, \sigma_i^2 D)$, *where μ_i and σ_i^2 are scalars, and $D > 0$. For the covariance shift model let $\sigma_1^2 > \sigma_2^2$. Then,*

$$\eta_s^2 = \left(N\mu_\Delta^2\right) \mathbf{1}^\mathsf{T}[\sigma_c^2 L^\mathsf{T} L + \sigma_v^2 I]^{-1} L^\mathsf{T} L \mathbf{1}, \tag{4.15}$$

$$R_s = \frac{\sigma_1^2 \|T(z) D^{\frac{1}{2}}\|_\infty + \sigma_v^2}{\sigma_2^2 \|T(z) D^{\frac{1}{2}}\|_\infty + \sigma_v^2}, \tag{4.16}$$

where $\mu_c = \mu_i$, $\sigma_c = \sigma_i$, $L = T(1)D^{\frac{1}{2}}$, and $\mu_\Delta = \mu_2 - \mu_1$.

The error probabilities for the identical statistics case can be obtained by substituting η_s and R_s to η and R in (4.11) and (4.12), respectively. The effect of sensor noise is also evident from the SNR expressions in the above corollary. In particular, by setting $\sigma_v^2 = 0$ in (4.15) and (4.16), the probabilities do not depend on the network matrix G.

Notice that the expressions of $\mathbb{P}_{e_m}(\mathcal{J})$ and $\mathbb{P}_{e_m}(\mathcal{J})$ in above lemma are valid even when N is finite. However, in this case, η and R are complicated functions of the adjacency matrix G. Instead, the elegance of SNRs in Lemma 29 and Corollary 30 is that they depend on the adjacency matrix G through the well understood transfer matrix $T(z)$. Thus, when $N \to \infty$, one can easily understand the impact of network structure on the detection performance by analyzing $T(z)$. By interpreting the quadratic function in η (or η_s) and $\|\cdot\|_\infty$ in R (or R_s) as a measure of gain, one expects that higher gains results in minimum error probabilities. This intuition is made precise in the following proposition:

Proposition 31 $\mathbb{P}_{e_m}(\mathcal{J})$ *and* $\mathbb{P}_{e_v}(\mathcal{J})$ *are decreasing in η (or η_s) and R (or R_s), resp.*

The above proposition also helps us to compare the performance of the MAP detectors associated with different sensor sets. This fact will be exploited greatly in the next section on network analysis of the MAP detector's performance.

Remark 32 ***(LD-MAP detector's error probability for other covariance matrix structures)*** *We now comment on extending $\mathbb{P}_{e_v}(\mathcal{J})$ (4.12) for including other covariance matrices. The case $\Sigma_1 = 0$ and $\Sigma_2 > 0$ can be handled using the proof of Lemma 29. For the scenario where neither of Σ_1 or Σ_2 is zero, if we have $N < \infty$ and $\lambda_{max}(\overline{\Sigma}_1\overline{\Sigma}_2^{-1}) > \lambda_{min}(\overline{\Sigma}_1\overline{\Sigma}_2^{-1})$, then $\mathbb{P}_{e_v}(\mathcal{J})$ remains the same as in (4.12), with $R = \lambda_{max}(\overline{\Sigma}_1\overline{\Sigma}_2^{-1})$. For other cases we refer the reader to [70]. However, the main difficulty in analyzing any of these error probabilities lies in the fact that resulting expressions of SNRs (R) are not amenable to analysis. If one assumes $\overline{\Sigma}_1$ and $\overline{\Sigma}_2$ to be simultaneously diagonalizable, as is the case with Corollary 30, an expression of R similar to (4.16) may be obtained.* □

4.3 Network analysis of the MAP detector

In this section, we characterize networks for which the MAP detector's performance associated with the sensors that are close to the input nodes is better (or worse) than those of sensors that are farther apart. We distinguish two separate cases when the sensors are without noise ($\sigma_v^2 > 0$) and with noise ($\sigma_v^2 = 0$). To make the notion of closeness precise, we introduce the notion of a node cutset.

Definition 33 ***(Node cutset)*** *For the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with input nodes $\mathcal{K}$, the nodes $\mathcal{C}_d \subseteq \mathcal{V}$, with $d > 1$, form a node cutset if there exist a non empty source set $\mathcal{S} \subseteq \mathcal{V}$ and a partitioned set $\mathcal{P} \subseteq \mathcal{V}$ such that $\mathcal{V} = \mathcal{S} \sqcup \mathcal{C}_d \sqcup \mathcal{P}$, where $\sqcup$ denotes the disjoint union, and*

(i) $\mathcal{K} \subseteq \mathcal{S}$ and dist$(\mathcal{K}, \mathcal{C}_d) \geq d$, and

(ii) every path from $\mathcal{S}$ to $\mathcal{P}$ contains a node in $\mathcal{C}_d$.

The requirement (i) ensures that the node cutset is at least d edges away from the input nodes. To illustrate Definition 33, consider the network in Fig 4.1. For the input nodes $\mathcal{K} = \{1, 2\}$, the nodes $\mathcal{C}_1 = \{4, 5, 6\}$ forms a node cutset. However, the nodes $\{5, 6, 7\}$ ceases to form a node cutset, since they failed to satisfy requirement (ii) in the above definition.

Assumption 34 *Given a graph $\mathcal{G}$, the input node set $\mathcal{K}$ and the distance d is predetermined.*

We are now ready to state our results on network theoretic characterization of the MAP detectors. It should be noted that, if a result holds true for the general detection problem (4.2), we do not state the analogous result for the *mean* and *covariance shift models.* We begin with the case of noiseless measurements ($\sigma_v^2 = 0$).

Theorem 35 ***(Performance of sensors on the node cutset vs the partitioned set: noiseless measurements)*** *Consider the general detection problem* (4.2). *Let $\mathcal{C}_d$ and $\mathcal{P}$ be as in Definition 33, and assume that the measurements from both these node sets are noiseless ($\sigma_v^2 = 0$). Associated with these measurements, let $\mathbb{P}_e(\mathcal{C}_d)$ and $\mathbb{P}_e(\mathcal{P})$ be the respective error probabilities that are computed using* (4.6). *Then, $\mathbb{P}_e(\mathcal{C}_d) \leq \mathbb{P}_e(\mathcal{P})$.*

This above theorem is a consequence of the following result in the binary detection problem, known as *theorem of irrelevance* [87] and the *invariance of MAP rule* [52].

Lemma 36 ***(Error probability of the MAP detector: dependent measurements)*** *Let M_1 and M_2 be any two arbitrary simple hypotheses with non-zero priors. Let δ_1 be the error probability of a MAP detector relying on the measurement $\mathbf{Y} \in \mathbb{R}^{p_1}$, and δ_2 be such a quantity associated with the measurement $\mathbf{Z} = g(\mathbf{Y}) + \mathbf{v}$, where $g(.) : \mathbb{R}^{p_1} \to \mathbb{R}^{p_2}$ and $\mathbf{v}$ is stochastically independent of the hypotheses. Then, $\delta_1 \leq \delta_2$.*

From Lemma 36, it also follows that Theorem 35 holds true even (i) for the case of non-Gaussian input and measurements (provided that the joint density exists), and (ii) if the set $\mathcal{P}$ is replaced with $\mathcal{P} \cup \widetilde{\mathcal{C}}_d$, where $\widetilde{\mathcal{C}}_d \subseteq \mathcal{C}_d$. Theorem 35 implies that, in the absence of noise, nodes near the input location achieve better detection performance compared to those far away from the inputs, *irrespective of the edge weights in the adjacency matrix G and the measurement horizon N*. Here, the notion of closeness is to be understood in the sense of *node cutsets*, since, $d \leq \text{dist}(\mathcal{K}, \mathcal{C}_d) < \text{dist}(\mathcal{K}, \mathcal{P})$. Thus, if node cutsets exist in a graph and the measurements are noiseless, one should always place sensors on the cutsets. Thus, if a budget is associated with the sensor placement, it makes sense to find a cutset $\mathcal{C}_d$ of minimum cardinality.

Proposition 37 ***(Error probability of the oracle detector)*** *Consider the general detection problem (4.2), and let δ_1 be the error probability of a MAP detector which can directly access the inputs $\mathbf{w}[k]$, $k = 0, \ldots, N$. For any sensor set $\mathcal{J}$, let δ_2 and δ_3 be the error probabilities associated with the noiseless ($\sigma_v^2 = 0$) and noisy ($\sigma_v^2 > 0$) measurements $\mathbf{Y}_{\mathcal{J}}$ (4.4), respectively. Then, $\delta_1 \leq \delta_2 \leq \delta_3$.*

Proposition 37 states that sensor noise degrades the performance of the MAP detector (this fact is also illustrated in Example 22). It also implies that measuring the inputs directly is always better than measuring the noisy/noiseless states (dynamics) of the nodes. Of course, given this fact, it is always beneficial to place the sensors at the input nodes, rather than dealing with the *node cutsets* and the *partitioned sets.*

We now consider the case of noisy measurements ($\sigma_v^2 > 0$). Notice that our results will be specific to the MAP and LD-MAP detectors associated with the *mean* and *covariance shift models*, respectively. Possible extensions to the general detection problem (4.2) are briefly mentioned in the remarks. We now introduce some additional notation. For a cutset $\mathcal{C}_d$, let $\mathbf{x}_c[k]$, $\mathbf{x}_s[k]$, and $\mathbf{x}_p[k]$ denote the states of the node sets $\mathcal{C}_d$, $\mathcal{S}$, and $\mathcal{P}$, respectively. Let M be a permutation matrix such that $\mathbf{x}[k] = M[\mathbf{x}_s[k]^\mathsf{T}, \mathbf{x}_c[k]^\mathsf{T}, \mathbf{x}_p[k]^\mathsf{T}]^\mathsf{T}$, where $\mathbf{x}[k]$ is the state vector of (4.1). Then, from (4.1) it also follows that

$$\begin{bmatrix} \mathbf{x}_s[k+1] \\ \mathbf{x}_c[k+1] \\ \mathbf{x}_p[k+1] \end{bmatrix} = \underbrace{\begin{bmatrix} G_{ss} & G_{sc} & 0 \\ G_{cs} & G_{cc} & G_{cp} \\ 0 & G_{pc} & G_{pp} \end{bmatrix}}_{M^{-1}GM} \begin{bmatrix} \mathbf{x}_s[k] \\ \mathbf{x}_c[k] \\ \mathbf{x}_p[k] \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{w}_s[k] \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}}_{M^{-1}\Pi\mathbf{w}[k]}. \tag{4.17}$$

From the above relation, note that the states of $\mathcal{C}_d$ serve as an input for the states of partitioned nodes set $\mathcal{P}$, i.e.,

$$\mathbf{x}_p[k+1] = G_{pp}\mathbf{x}_p[k] + G_{pc}\mathbf{x}_c[k]. \tag{4.18}$$

Based on the transfer function matrix of subsystem (4.18), we now state a result that is analogous to Theorem 35, for the case $\sigma_v^2 > 0$.

Theorem 38 ***(Performance of sensors on the node cutset vs the partitioned set: noisy measurements)*** *Let G_{pp} and G_{pc} be as in (4.17), and assume that spec$(G_{pp}) \cap \{z \in \mathbb{C} : |z| = 1\} = \phi$. Let $\overline{\rho}(z)$ and $\underline{\rho}(z)$ be the maximum and minimum singular values of $T_s(z) = (zI - G_{pp})^{-1}G_{pc}$, respectively. Let $\mathbb{P}_{e_m}(\mathcal{C}_d)$ in (4.11) and $\mathbb{P}_{e_v}(\mathcal{C}_d)$ in (4.12) be the error probabilities obtained using the noisy measurements ($\sigma_v^2 > 0$) from the cutset $\mathcal{C}_d$. Instead, let $\mathbb{P}_{e_m}(\mathcal{P})$ and $\mathbb{P}_{e_v}(\mathcal{P})$ be the error probabilities associated with the partitioned set $\mathcal{P}$. Then we have:*

1a) If $\overline{\rho}(1) \leq 1$, then $\mathbb{P}_{e_m}(\mathcal{C}_d) \leq \mathbb{P}_{e_m}(\mathcal{P})$.

1b) If $\underline{\rho}(1) > 1$*, then* $\mathbb{P}_{e_m}(\mathcal{C}_d) > \mathbb{P}_{e_m}(\mathcal{P})$*.*

2a) If $\sup_{|z|=1} \overline{\rho}(z) \leq 1$ *then* $\mathbb{P}_{e_v}(\mathcal{C}_d) \leq \mathbb{P}_{e_v}(\mathcal{P})$*.*

2b) If $\inf_{|z|=1} \underline{\rho}(z) > 1$*, then* $\mathbb{P}_{e_v}(\mathcal{C}_d) > \mathbb{P}_{e_v}(\mathcal{P})$*.*

Hence, in the presence of noise, depending upon the entries in the matrix $[G_{pp}\, G_{pc}]$, measuring the cutset $\mathcal{C}_d$ might not be always optimal for the purposes of the detection. Instead, in the noiseless case, Theorem 35 states that measuring the cutset is always optimal, irrespective of the entries in G. We now explain the reason behind this contrasting behaviors.

Notice that, the quantities sup and inf of $\overline{\rho}(z)$ and $\underline{\rho}(z)$ in Theorem 38, respectively, are the maximum and minimum input to output gains of the transfer function matrix $T_s(z)$, associated with the system (4.18). Theorem 38 says that, if the gain between the states $\mathbf{x}_c[k]$ and the states $\mathbf{x}_p[k]$ is high (low), the detection performance with sensors in $\mathcal{P}$ should be better (worse) that that of $\mathcal{C}_d$. In fact, recall from Lemma 29 that the detectors associated with the noisy measurements of $\mathcal{C}_d$ and $\mathcal{P}$, respectively, depends on the SNRs of $\mathbf{x}_c[k]$ and $\mathbf{x}_c[k]$ (plus the sensor noise), respectively. Since $\mathbf{x}_p[z] = T_s(z)\mathbf{x}_c[z]$, it is clear that the SNRs are influenced by the gains of $T_s(z)$. In particular, a higher gain increases the SNR of the detector associated with $\mathcal{P}$, which results in a better performance compared to the detector associated with that of $\mathcal{C}_d$.

The above reasoning also holds in the case of noiseless measurements, however, the transfer function gain do not influence MAP detector's performance. In fact, this gain gets canceled in the error probability computations (this can be clearly seen in Example 22 by interpreting α as the gain). Theorem 38 provides conditions for placing sensors on or away from the cutset nodes. For general adjacency matrix, one needs to rely on the software (based on LMI based inequalities) to validate those conditions. However, for non-negative adjacency matrices, the conditions for placing (or not) sensors on the cutset nodes can be stated based on algebraic conditions on the entries of the adjacency matrix. In fact, we have the following result:

Lemma 39 ***(Non-negative adjacency matrix)*** *Let the matrix G in* (4.1) *be non-negative, and $\widetilde{G} = [G_{pp}\, G_{pc}] \in \mathbb{R}^{m_1 \times n_1}$, where G_{pp} and G_{pc} are defined in* (4.18)*.*

1. *If $\|\widetilde{G}\|_\infty \leq 1/\sqrt{m_1}$, then we have $\mathbb{P}_{e_m}(\mathcal{C}_d) \leq \mathbb{P}_{e_m}(\mathcal{P})$ and $\mathbb{P}_{e_v}(\mathcal{C}_d) \leq \mathbb{P}_{e_v}(\mathcal{P})$.*
2. *If $n_1 = 1$, and all row sums of $\widetilde{G}$ are greater than one, then $\mathbb{P}_{e_m}(\mathcal{C}_d) \geq \mathbb{P}_{e_m}(\widetilde{\mathcal{P}})$ and $\mathbb{P}_{e_v}(\mathcal{C}_d) \geq \mathbb{P}_{e_v}(\widetilde{\mathcal{P}})$, where $\widetilde{\mathcal{P}} \subseteq \mathcal{P}$.*

The inequality $\mathbb{P}_{e_m}(\mathcal{C}_d) \leq \mathbb{P}_{e_m}(\mathcal{P})$ can be obtained even without the non-negativity assumption on G. However, this might not be true for the case of $\mathbb{P}_{e_v}(\cdot)$. Thus, by ensuring that the maximum row sum of $\widetilde{G}$ is bounded by $1/\sqrt{m_1}$ (here m_1 refers to the cardinality of the partitioned set $\mathcal{P}$), one can guarantee that the detection performance of sensors on the cutset is always superior than that of the sensors on the partitioned nodes. The assumption $n_1 = 1$ in part 2) of above lemma implies that card$(\mathcal{C}_d) = 1$. For arbitrary n_1, the condition row sums of $\widetilde{G}$ greater than one may not be sufficient, and more assumptions on G are required to handle this case. For instance, when G is a diagonally dominant matrix, required sufficient conditions can be obtained using the lower bounds in [83]. Finally, we notice that the bounds presented in Lemma 39 depends on the cardinality of the node sets, and hence, our results on networks with non-negative edge weights may be conservative when these cardinalities are large.

Remark 40 ***(Extension of network theoretic results to the other detectors: noisy measurements)*** *In the cases where the analytical error probability calculation is difficult, eg., the general Gaussian or non-Gaussian detection problem and the covariance shift model with arbitrary covariance matrix structures, one relies on the Chernoff type bounds (for eg., see [80]) to quantify the detection performance. In both the cases, i.e., evaluating the performance directly or via bounds, Theorem 38 holds true for any detector whose performance (resp. bounds) is monotonically increasing in* $\|T(z) = C(zI - G)^{-1}\Pi\|_M$*, for some suitable* $M \succ 0$*. For instance, the Chernoff bounds on the error probability of the general Gaussian detection problem* (4.2) *depend on the moment generating function (MGF) of the test statistic of the MAP detector, which depends on the filtered mean and covariance matrices* (4.8)*, and our analysis becomes applicable. In the non-Gaussian case, the MGF might depend on other moments as well, and extending our results to this case will be challenging.* □

4.3.1 Single input single output (SISO) line networks

In this section, we validate our cutset based results, that we present in previous section, for the case of line networks by explicitly expressing the error probabilities as a function of the entires of G, and then compare the performance of sensors on $\mathcal{C}_d$ versus sensors on $\mathcal{P}$. We restrict our attention to the SISO systems.

We assume that a stochastic input enters the network through a fixed node $q \in \{1, \dots, n\}$, and we allow any node $l \in \{1, \dots, n\}$ with $\text{dist}(l, q) \geq d$ for sensor placement. For this setup, we assume that probabilities $\mathbb{P}_{e_m}(l)$ and $\mathbb{P}_{e_v}(l)$ are obtained by substituting the SNRs η_s (4.15) and R_s (4.16) in the expressions of (4.11) and (4.12), respectively. Notice that, in contrast to the previous analysis, in which we assume $\Sigma_2 = 0$ (see Assumption 28),

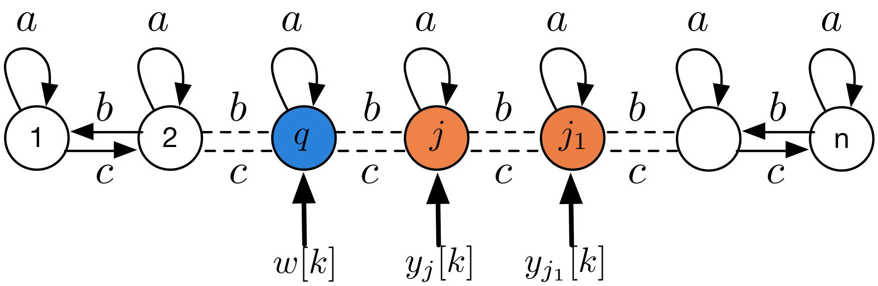


Figure 4.2: Toeplitz line network with n nodes. The q-th node is injected with the input, and the j-th node represents the cutset node.

in this section we do not assume $\sigma_2^2 = 0$ in R_s. For the ease of presentation, we assume the cutset to be a singleton set, i.e., $\mathcal{C}_d = \{j\}$. The following proposition is an extension of Lemma 39 for our SISO system setup with the revised error probabilities.

Proposition 41 *Let $\widetilde{G}$ be as in Lemma 39, and $\sigma_v^2 > 0$. Let $\{j\}$ and $\mathcal{P}$ be the cutset and partitioned sets, resp. If $\|\widetilde{G}\|_\infty \leq 1$, then for any $j_1 \in \mathcal{P}$, we have $\mathbb{P}_{e_m}(j) \leq \mathbb{P}_{e_m}(j_1)$ and $\mathbb{P}_{e_v}(j) \leq \mathbb{P}_{e_v}(j_1)$. The opposite inequality holds true if all row sums of $\widetilde{G}$ are greater than one.*

The proof of above proposition is similar to the proof of Lemma 39 and hence, the details are omitted. By not resorting to any proof techniques, i.e, the functional dependence arguments, that we used in previous section, we now validate assertions in above proposition by expressing the error probability in terms of the entries in the matrix G. To this aim, we consider a line network (see Fig. 4.2), whose adjacency matrix is given by the following matrix:

$$G = \begin{bmatrix} a & b & 0 & \cdots & 0 & 0 \\ c & a & b & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & b \\ 0 & 0 & 0 & \cdots & c & a \end{bmatrix}_{n \times n}, \tag{4.19}$$

where, $a, b, c \in \mathbb{R}_{\geq 0}$. We let the cutset node j be located on the right of the input node q, i.e., $1 \leq q < j < n$ (see Fig 4.2). The case when j is to the left side of the input node q follows similar analysis. Thus, we have the partitioned set $\mathcal{P} = \{j+1, \ldots, n\}$. We now show that, for any $l \in \mathcal{P}$, the error probabilities $\mathbb{P}_{e_m}(l)$ and $\mathbb{P}_{e_v}(l)$ are greater or smaller than those of the cutset node j. The following proposition helps us achieve the required goal:

Proposition 42 *Let G be as in* (4.19) *and $\overline{\lambda}(G) < 1$. Let $|(I-G)^{-1}_{l,q}|$ be the absolute value of (l,q)-th entry of $(I-G)^{-1}$. Let $\widetilde{G}$ be as in Lemma 39. Then, we have:*

i) If $\|\widetilde{G}\|_\infty < 1$, then $|(I-G)^{-1}_{q,q}| \geq |(I-G)^{-1}_{q+1,q}| \cdots \geq |(I-G)^{-1}_{n,q}|$.

ii) If all row sums of $\widetilde{G}$ are greater than one, then $|(I-G)^{-1}_{lq}| \geq |(I-G)^{-1}_{qq}|$ for all $q < l \leq n$. If $b = 0$, we have $|(I-G)^{-1}_{q+1,q}| \geq |(I-G)^{-1}_{q+2,q}| \cdots \geq |(I-G)^{-1}_{n,q}|$.

For a fixed input q, above proposition characterizes the qualitative behavior of the input-to-output transfer function gains associated with different output nodes. This fact can be easily seen by expressing $|(I-G)^{-1}_{lq}|$ as $|\mathbf{e}_l^{\mathsf{T}}(I-G)^{-1}\mathbf{e}_q|$. For the case of Toeplitz line networks, the assertion in Proposition 41 is now an easy consequence of Proposition 31

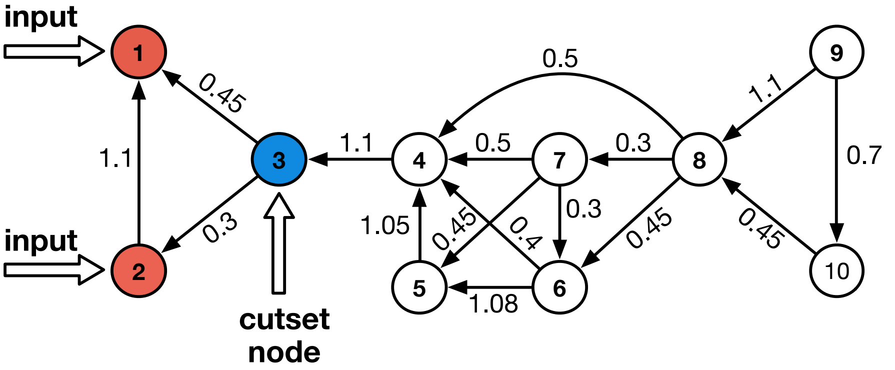


Figure 4.3: The graph of a network consisting of 10 nodes. The nodes that are to the right of the cutset node $\{3\}$ form the partitioned set. Instead, nodes 1 and 2 form the source set.

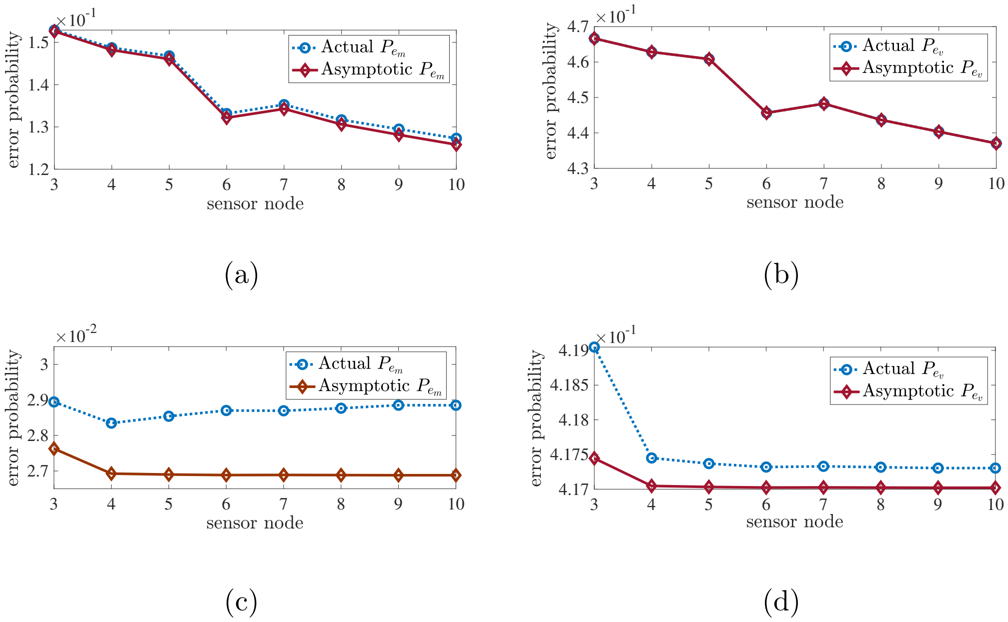
and 42. In particular, if $b = 0$ and $a+c > 1$, Proposition 42 also implies that, the node that is farthest from the input has better detection performance than any other node, including the cutset node. Similarly, assertion in Theorem 35 can be verified by letting $\sigma_v^2 = 0$.

The procedure illustrated above, evaluating the error probabilities via the entries of $(I - G)^{-1}$, becomes tedious and might not be even possible for arbitrary network structures. In such situations, one can use the proof techniques presented in Section 4.3 for understanding the detection performance of sensors on networks.

4.4 Simulation results

In this section, we present numerical simulations to validate the effectiveness of our cutset based characterization of MAP detection performance on networks.

(A. Detection performance of sensors on the partitioned nodes is better than that of the sensors on the cutset nodes): For this scenario, consider the network in Fig 4.3. The network has 10 nodes, with 1 and 2 being the input nodes, $\mathcal{C}_d = \{3\}$ is the cutset node,



(a)

(b)

(c)

(d)

Figure 4.4: Actual and asymptotic error probabilities (Lemma 29) of the MAP and LD-MAP detectors associated with various nodes of the network shown in Fig. 4.3. The panels (a) and (b) corresponds to the adjacency matrix that results in the shorter memory of the network dynamics (3.2). Instead, panels (c) and (d) are associated with an adjacency matrix that results in the longer memory of the network dynamics. The error probability associated with each node in the partitioned set $\mathcal{P} = \{4, \ldots, 10\}$ is less than that of the cutset node $\mathcal{C}_d = \{3\}$. This result is consistent with Lemma 39, because all row sums of submatrix $\widetilde{G}$ are greater than one.

and $\mathcal{P} = \{4, \ldots, 10\}$ is the partitioned node set. The adjacency matrix of this network is nilpotent, and as a result, system (4.1) evolving on this network will have a short memory (in fact $G^{10} = 0$). By short (resp. long) memory, we mean that the current state of the network depends on few (resp. several) past states. For the mean shift model, the input $\mathbf{w}_i[k] \sim \mathcal{N}(\mu_i \mathbf{1}, \sigma_i^2 I_{2\times 2})$, where $\mu_1 = 2$, $\mu_2 = 1$, and $\sigma_2^2 = \sigma_2^2 = 1.5$. Instead, for the covariance shift model, the input[2] $\mathbf{w}_i[k] \sim \mathcal{N}(\mathbf{0}, \sigma_i^2 I_{2\times 2})$, where $\sigma_1^2 = 2.0$ and $\sigma_2^2 = 1.0$. In both the models, $N = 200$ and the sensor noise variance $\sigma_v^2 = 1.2$.

[2] the choice of zero mean is arbitrary, since, the LD-MAP detector's error probability do not depend on the mean; see Lemma 29.

Fig. 4.4(a) and Fig. 4.4(b) illustrates the actual and asymptotic error probabilities of the mean and covariance shift models, respectively. The error probabilities are computed using the formulas in Lemma 29. In particular, for the asymptotic case, we use the SNRs in Corollary 30. In both figures, the error probability associated with the cutset node is greater than that of any node in the partitioned set. This must be the case since $G \geq 0$, and the row sums of the submatrix $\widetilde{G}$ are greater than one (see Lemma 39).

The error between the asymptotic and actual error probabilities in Fig. 4.4(a) and Fig. 4.4(b) is almost negligible, even when N is not large. This is because the adjacency matrix G is a nilpotent matrix, and as a result, the difference between the actual and asymptotic SNRs is minimum. However, this might not be the case when G has long memory, i.e., $G^k \approx 0$ only for a very large k. For $N = 800$, Fig. 4.4(c) and Fig. 4.4(d) illustrate this scenario for the network that is obtained by modifying some edges of the network in Fig. 4.3, such that $G^k \approx 0$ for very large k.

(B. Detection performance of sensors on the cutset nodes is better than that of the sensors on the partitioned nodes): Consider the network shown in Fig 4.5. The network has 50 nodes among which $\mathcal{K} = \{1, 2, 3, 5, 21, 26, 36, 43\}$ are the input nodes. The cutset $\mathcal{C}_d = \{22, 30, 38\}$ separates $\mathcal{K}$ from the partitioned set $\mathcal{P} = \{34, 35, 40, 42, 44, 48, 49\}$. For the mean shift model, the input $\mathbf{w}_i[k] \sim \mathcal{N}(\mu_i \mathbf{1}, \sigma_i^2 I_8)$, where $\mu_1 = 2$, $\mu_2 = 1$, and $\sigma_2^2 = \sigma_2^2 = 1.5$, and $\sigma_v^2 = 1.2$. Instead, for the covariance shift model, the input $\mathbf{w}_i[k] \sim \mathcal{N}(\mathbf{0}, \sigma_i^2 I_8)$, where $\sigma_1^2 = 25.0$, $\sigma_2^2 = 0.1$, and $\sigma_v^2 = 0.5$. In both the models, $N = 200$.

Consider all possible subsets of $\mathcal{C}_d \sqcup \mathcal{P}$ whose cardinalities are same as that of the cutset $\mathcal{C}_d$. It is easy to see that there are 120 such sets. For each of these sets, we associate

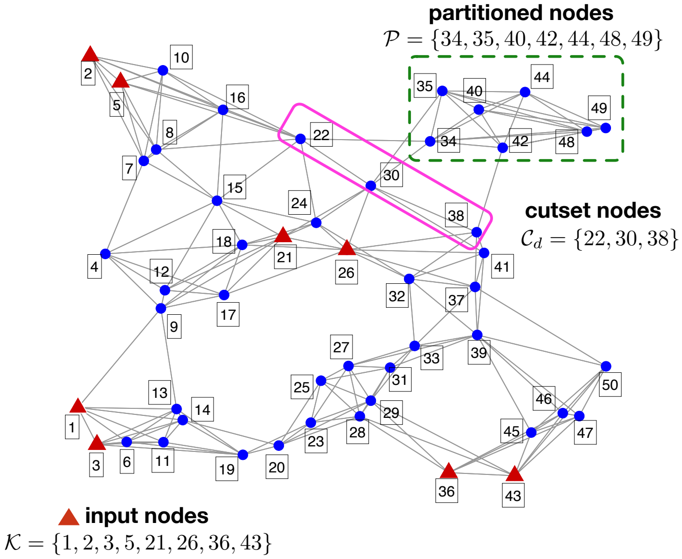
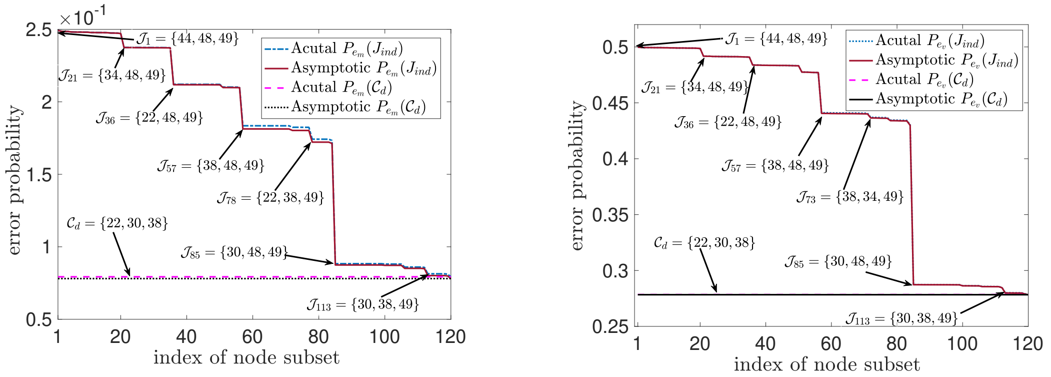


Figure 4.5: Graph associated with a randomly generated network consisting of 50 nodes [64]. A total of 8 nodes are subjected to stochastic inputs. Instead, sensors are placed on the cutset nodes and the partitioned nodes that are not collocated with the input nodes.

a label $\mathcal{J}_{\text{ind}}$, where ind $\in \{1, \ldots, 120\}$. The labels are given based on a decreasing order of the error probabilities associated with the subsets. In Fig. 4.6(a) and Fig. 4.6(b), we show the actual and asymptotic error probabilities of the mean and covariance shift models, respectively. In both figures, the error probability associated with the $\mathcal{C}_d$ is lesser than that of any $\mathcal{J}_{\text{ind}}$. This must be the case because $G \geq 0$, and the row sums of the submatrix $\|\widetilde{G}\|_\infty < 1/\sqrt{7} = 0.3780$ (see Lemma 39).

4.5 Summary

In this paper we formulate a sensor placement problem for linear dynamical systems defined over networks with unknown stochastic inputs, using the statistical hypothesis testing framework. In particular, the main technical contribution of this paper is to derive



(a) mean shift model.

(b) covariance shift model.

Figure 4.6: Actual and asymptotic error probabilities (Lemma 29) of the MAP and LD-MAP detectors associated with the node cutset $\mathcal{C}_d$ and all possible 3 node subsets of $\mathcal{C}_d \sqcup \mathcal{P}$ of the network shown in Fig. 4.5.The error probabilities of the detectors associated with cutset nodes is lower than that of the detectors associated with any subset of the nodes in the partitioned set. This result is consistent with Lemma 39, because the submatrix $\widetilde{G}$ row sums of the adjacency matrix G are less than $1/\sqrt{m_1}$ ($m_1 = 7$).

sufficient conditions under which a MAP detector associated with the sensors closer to the input nodes results in better (or worse) performance than compared to a MAP detector associated with the sensors that are far away from the input nodes. For networks with non-negative adjacency matrix, we show that the sensors should be placed either near the origin of the signal or, as far as possible from it, depending on the network parameters.

4.2 Appendix

Proof of proposition 26: From the network dynamics (4.1) and sensor measurements (4.3), $\mathbf{Y}_{\mathcal{J}}$ (4.4) can be expanded as

$$\mathbf{Y}_{\mathcal{J}} = \mathcal{O}\mathbf{x}[0] + \mathcal{F}\mathbf{w}_{0:N-1} + \mathbf{v}_{1:N}, \tag{4.20}$$

where $\mathbf{w}_{0:N-1} = [\mathbf{w}[0]^\mathsf{T}, \ldots, \mathbf{w}[N-1]^\mathsf{T}]^\mathsf{T}$ and $\mathbf{v}_{1:N} = [\mathbf{v}[1]^\mathsf{T}, \ldots, \mathbf{v}[N]^\mathsf{T}]^\mathsf{T}$, respectively. The matrices $\mathcal{O}$ and $\mathcal{F}$ are defined in the statement of the proposition. The expressions of $\overline{\boldsymbol{\mu}}_i$ and $\overline{\Sigma}_i$ in (4.8) follows by taking the expectation and covariance of $\mathbf{Y}_\mathcal{J}$. □

Proof of Lemma 27: Let ζ and z are the realizations of $\mathbf{Y}_\mathcal{J}$ and y, respectively. Since the input and measurement noises follows a Gaussian distribution, the probability density functions of $\mathbf{Y}_\mathcal{J}$ (4.4) and $y = \mathbf{b}^\mathsf{T}\mathbf{Y}_\mathcal{J}$ are

$$f(\zeta|H_i) \propto \frac{1}{\sqrt{|\overline{\Sigma}_i|}} \exp\left[-\frac{1}{2}(\zeta - \overline{\boldsymbol{\mu}}_i)^T \overline{\Sigma}_i^{-1}(\zeta - \overline{\boldsymbol{\mu}}_i)\right] \text{ and}$$

$$g(z|H_i) \propto \frac{1}{\sqrt{\mathbf{b}^\mathsf{T}\overline{\Sigma}_i\mathbf{b}}} \exp\left[-\frac{(z - \mathbf{b}^\mathsf{T}\overline{\boldsymbol{\mu}}_i)^2}{2\,\mathbf{b}^\mathsf{T}\overline{\Sigma}_i\mathbf{b}}\right], \tag{4.21}$$

respectively, where $|\cdot|$ denotes the determinant. Define the log likelihood ratios $\Psi(\zeta) = \ln(f(\zeta|H_2)/f(\zeta|H_1))$ and $\widehat{\Psi}(z) = \ln(f(z|H_2)/f(z|H_1))$. Then, from the mixed Bayes formula [52], the MAP decision rules based on ζ and z, respectively, are given by

$$\Psi(\zeta) \underset{\widehat{H}=H_1}{\overset{\widehat{H}=H_2}{\gtrless}} \gamma \text{ and } \widehat{\Psi}(z) \underset{\widehat{H}=H_1}{\overset{\widehat{H}=H_2}{\gtrless}} \gamma. \tag{4.22}$$

part 1) Since $\Sigma_1 = \Sigma_2$ and $\boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, from (4.8), it follows that $\overline{\Sigma}_1 = \overline{\Sigma}_2$ and $\overline{\boldsymbol{\mu}}_1 \neq \overline{\boldsymbol{\mu}}_2$. Invoking this observation in $f(\zeta|H_i)$, yields the following expression for $\psi(\zeta)$:

$$\Psi(\zeta) = -0.5\,\overline{\boldsymbol{\mu}}_\Delta^\mathsf{T}\overline{\Sigma}_2^{-1}\overline{\boldsymbol{\mu}}_\Delta + (y - \overline{\boldsymbol{\mu}}_1)^\mathsf{T}\,\overline{\Sigma}_2^{-1}\overline{\boldsymbol{\mu}}_\Delta. \tag{4.23}$$

Substitute (4.23) in the first decision rule of (4.22) and simplify the expression to obtain the MAP decision rule (4.9) for ζ. Finally, replacing ζ with $\mathbf{Y}_\mathcal{J}$ yields the required expression.

part 2) In this case we have $\overline{\boldsymbol{\mu}}_1 = \overline{\boldsymbol{\mu}}_2$ and $\overline{\Sigma}_1 \neq \overline{\Sigma}_2$. A similar procedure, as in part 1), based on $g(z|H_i)$ (4.22) and the second decision rule in (4.21), yields the LD-MAP detector's expression (4.9). Details are left to the reader. □

Proof of Lemma 29: We divide the proof into two parts. In part 1) we derive the expressions (4.11) and (4.13) Instead, in part 2) we derive the expressions (4.12) and (4.14).

part 1) Under the assumption that $N < \infty$, let $\widehat{\mathbb{P}}_{e_m}(\mathcal{J})$ be the error probability of (4.9).. Then, from (4.9), we have

$$\begin{aligned}\Pr\left(\widehat{H} = H_2 | H_1\right) &= \Pr\left(s > \overline{\boldsymbol{\mu}}_{\Delta}^{\mathsf{T}} \overline{\Sigma}_c^{-1} \left(\overline{\boldsymbol{\mu}}_1 + \overline{\boldsymbol{\mu}}_2\right) | H_1\right) \text{ and}\\ \Pr\left(\widehat{H} = H_1 | H_2\right) &= \Pr\left(s < \overline{\boldsymbol{\mu}}_{\Delta}^{\mathsf{T}} \overline{\Sigma}_c^{-1} \left(\overline{\boldsymbol{\mu}}_1 + \overline{\boldsymbol{\mu}}_2\right) | H_2\right),\end{aligned}$$

where $s = 2\,\overline{\boldsymbol{\mu}}_{\Delta}^{\mathsf{T}} \overline{\Sigma}_c^{-1} \mathbf{Y}_{\mathcal{J}}$ follows $\mathcal{N}(\overline{\boldsymbol{\mu}}_{\Delta}^{\mathsf{T}} \overline{\Sigma}_c^{-1} \overline{\boldsymbol{\mu}}_1, 4\overline{\boldsymbol{\mu}}_{\Delta}^{\mathsf{T}} \overline{\Sigma}_c^{-1} \overline{\boldsymbol{\mu}}_{\Delta})$ under H_i, because s is a linear transform of $\mathbf{Y}_{\mathcal{J}}|H_i$, which follows a Gaussian distribution. Define $\widehat{\eta}^2 = \overline{\boldsymbol{\mu}}_{\Delta}^{\mathsf{T}} \overline{\Sigma}_c^{-1} \overline{\boldsymbol{\mu}}_{\Delta}$, and notice that $\Pr(\widehat{H} = H_2|H_1) = Q_{\mathcal{N}}(0.5\,\widehat{\eta})$ and $\Pr(\widehat{H} = H_1|H_2) = 1 - Q_{\mathcal{N}}(0.5\,\widehat{\eta})$. Finally, from (4.6), we have $\widehat{\mathbb{P}}_{e_m}(\mathcal{J}) = 0.5\,Q_{\mathcal{N}}(\widehat{\eta})$. Define $\mathbb{P}_{e_m}(\mathcal{J}) = \lim_{N\to\infty} \widehat{\mathbb{P}}_{e_m}(\mathcal{J})$, and note that

$$\mathbb{P}_{e_m}(\mathcal{J}) = \lim_{N\to\infty} 0.5\,Q_{\mathcal{N}}(0.5\,\widehat{\eta}) = 0.5\,Q_{\mathcal{N}}\left(0.5 \lim_{N\to\infty} \widehat{\eta}\right)$$

where the final equality follows because $\widehat{\eta}$ is increasing in N (see Proposition A.43). We now show that $\lim_{N\to\infty} \widehat{\eta} = \eta$. From (4.8), it follows that

$$\widehat{\eta}^2 = (\mathcal{F}\mathbf{m})^{\mathsf{T}}\, \overline{\Sigma}_c^{-1} (\mathcal{F}\mathbf{m}), \tag{4.24}$$

where $\mathbf{m} = \mathbf{1}_N \otimes \boldsymbol{\mu}_{\Delta}$ and $\boldsymbol{\mu}_{\Delta} = \boldsymbol{\mu}_2 - \boldsymbol{\mu}_1$. Let $l = 1, 2, \ldots,$, and define $K(l) = CG^l\Pi$ and $S(i) = \sum_{l=0}^{i-1} K(l)$. With these definitions and the assumption $\overline{\lambda}(G) < 1$, we have $\lim_{i\to\infty} S(i) = C(I-G)^{-1}\Pi \triangleq \overline{K}$, and

$$\mathcal{F}\mathbf{m} = \underbrace{\begin{bmatrix}(S(1)-\overline{K})^{\mathsf{T}} & (S(2)-\overline{K})^{\mathsf{T}} & \cdots & (S(N)-\overline{K})^{\mathsf{T}}\end{bmatrix}^{\mathsf{T}}}_{S_N} \boldsymbol{\mu}_{\Delta} + \left[\mathbf{1}_N \otimes \overline{K}\right] \boldsymbol{\mu}_{\Delta}. \tag{4.25}$$

Let $t(S_N)=\boldsymbol{\mu}_\Delta^\mathsf{T}\left[S_N^\mathsf{T}\overline{\Sigma}_c^{-1}S_N+2S_N^\mathsf{T}\Sigma_c^{-1}\left[\mathbf{1}_N\otimes\overline{K}\right]\right]\boldsymbol{\mu}_\Delta$. From (4.25) and (4.24), we have

$$\widehat{\eta}^2=\boldsymbol{\mu}_\Delta^\mathsf{T}\underbrace{\left[\mathbf{1}_N\otimes\overline{K}\right]^\mathsf{T}\overline{\Sigma}_c^{-1}\left[\mathbf{1}_N\otimes\overline{K}\right]}_{F}\boldsymbol{\mu}_\Delta+t(S_N). \tag{4.26}$$

Since $\mathbf{x}[0]=0$, from (4.8), it follows that $\overline{\Sigma}_c=\left[\mathcal{F}\left(I_N\otimes\Sigma_c\right)\mathcal{F}^\mathsf{T}+\sigma_v^2I\right]$. Further,

$$\begin{aligned}\left[\mathbf{1}_N\otimes\overline{K}\right]^\mathsf{T}\overline{\Sigma}_c&=\left[\overline{K}^\mathsf{T}\overline{K}\Sigma_c+\sigma_v^2I\right]\left[\mathbf{1}_N\otimes\overline{K}\right]^\mathsf{T}+\\&\qquad\underbrace{\overline{K}^\mathsf{T}\left[\widetilde{S}_N^\mathsf{T}\left(I\otimes\Sigma_c\right)\mathcal{F}^\mathsf{T}+\overline{K}\Sigma_cS_N^\mathsf{T}\right]}_{\widetilde{M}},\end{aligned} \tag{4.27}$$

where $\widetilde{S}_N$ is obtained by permuting, bottom to top, the block matrices of S_N (4.25). Right multiplying either sides of (4.27) with $\overline{\Sigma}_c^{-1}\left[\mathbf{1}_N\otimes\overline{K}\right]$ gives us:

$$N\overline{K}^\mathsf{T}\overline{K}=\left[\overline{K}^\mathsf{T}\overline{K}\Sigma_c+\sigma_v^2I\right]F+P, \tag{4.28}$$

where $P=\widetilde{M}\overline{\Sigma}_c^{-1}[\mathbf{1}_N\otimes\overline{K}]$ and F is defined in (4.26). Since $\overline{K}^\mathsf{T}\overline{K}\Sigma_c\succeq 0$, from (4.28), it follows that $F=[\overline{K}^\mathsf{T}\overline{K}\Sigma_c+\sigma_v^2I]^{-1}[N\overline{K}^\mathsf{T}\overline{K}-P]$. Substituting F in (4.26) yields

$$\widehat{\eta}^2=N\boldsymbol{\mu}_\Delta^\mathsf{T}([\overline{K}^\mathsf{T}\overline{K}\Sigma_c+\sigma_v^2I]^{-1}\overline{K}^\mathsf{T}\overline{K})\boldsymbol{\mu}_\Delta+\epsilon(N),$$

where $\epsilon(N)=-\boldsymbol{\mu}_\Delta^\mathsf{T}[\overline{K}^\mathsf{T}\overline{K}\Sigma_c+\sigma_v^2I]^{-1}P\boldsymbol{\mu}_\Delta+t(S_N)$. Finally, substituting $K=L\Sigma_c^{-\frac{1}{2}}$ in the above expression, and manipulating the terms will give us

$$\widehat{\eta}^2=N\widetilde{\boldsymbol{\mu}}_\Delta^\mathsf{T}\left([L^\mathsf{T}L+\sigma_v^2I]^{-1}L^\mathsf{T}L\right)\widetilde{\boldsymbol{\mu}}_\Delta+\epsilon(N). \tag{4.29}$$

We claim that $\lim_{N\to\infty}\epsilon(N)=0$. To see this, rewrite $\epsilon(N)$ as $\boldsymbol{\mu}_\Delta^\mathsf{T}(Q_1(N)+Q_2(N)+Q_3(N))\boldsymbol{\mu}_\Delta$, where

$$\begin{aligned}Q_1(N)&=S_N^\mathsf{T}\left[\overline{\Sigma}_c^{-1}S_N+2\Sigma_c^{-1}\left[\mathbf{1}_N\otimes\overline{K}\right]\right]\\Q_2(N)&=[\overline{K}^\mathsf{T}\overline{K}\Sigma_c+\sigma_v^2I]^{-1}\overline{K}^\mathsf{T}\widetilde{S}_N^\mathsf{T}\left(I\otimes\Sigma_c\right)\mathcal{F}^\mathsf{T}\\Q_3(N)&=[\overline{K}^\mathsf{T}\overline{K}\Sigma_c+\sigma_v^2I]^{-1}\overline{K}^\mathsf{T}\overline{K}\Sigma_cS_N^\mathsf{T}.\end{aligned}$$

From part 1) of Assumption 28, there exist a $k \in \mathbb{N}$ such that for all $m \in \{k, k+1, \dots, N\}$, $S(m) - \overline{K} = 0$. Thus, all but finite rows of S_N (4.25) are zeros, i.e., we can express S_N^T as $[F_1^\mathsf{T}\ 0^\mathsf{T}]$ and $\widetilde{S}_N^\mathsf{T}$ as $[0^\mathsf{T}\ F_2^\mathsf{T}]$, where the dimension of F_1 and F_2 depends only k. Thus, for all $N > k$, $Q_i(N)$ is a constant matrix, say Q_i, and we may conclude that

$$\|\boldsymbol{\mu}_\Delta\|_2^2 \sum_{i=1}^{3} \lambda_{\min}(Q_i + Q_i^\mathsf{T}) \leq 2 \sum_{i=1}^{3} \boldsymbol{\mu}_\Delta^\mathsf{T} Q_i \boldsymbol{\mu}_\Delta \leq \|\boldsymbol{\mu}_\Delta\|_2^2 \sum_{i=1}^{3} \lambda_{\max}(Q_i + Q_i^\mathsf{T}),$$

where $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ are the maximum and minimum eigenvalues. Since $\lim_{N\to\infty} N\|\boldsymbol{\mu}_\Delta\|_2 = c$ (Assumption 28), it follows that $\lim_{N\to\infty} \|\boldsymbol{\mu}_\Delta\|_2 = 0$. Hence, $\lim_{N\to\infty} \epsilon(N) = 0$ and $\lim_{N\to\infty} \widehat{\eta} = \eta$ (4.13).

part 2) Under the assumption that $N < \infty$, let $\widehat{\mathbb{P}}_{e_v}(\mathcal{J})$ be the error probability of (4.10). Then, from (4.9), we have

$$\begin{aligned}
\Pr\left(\widehat{H} = H_2 | H_1\right) &= \Pr\left(\ln(\widehat{R}) > \left[\frac{Z^2}{\mathbf{b}^\mathsf{T}\overline{\Sigma}_2\mathbf{b}} - \frac{Z^2}{\mathbf{b}^\mathsf{T}\overline{\Sigma}_1\mathbf{b}}\right] | H_1\right), \\
\Pr\left(\widehat{H} = H_1 | H_2\right) &= \Pr\left(\ln(\widehat{R}) < \left[\frac{Z^2}{\mathbf{b}^\mathsf{T}\overline{\Sigma}_2\mathbf{b}} - \frac{Z^2}{\mathbf{b}^\mathsf{T}\overline{\Sigma}_1\mathbf{b}}\right] | H_2\right),
\end{aligned}$$

where $Z = \mathbf{b}^\mathsf{T}[\mathbf{Y}_J - \overline{\boldsymbol{\mu}}_c]$ and $\widehat{R} = (\mathbf{b}^\mathsf{T}\overline{\Sigma}_1\mathbf{b}/(\mathbf{b}^\mathsf{T}\overline{\Sigma}_2\mathbf{b}) > 1$ (since $\Sigma_2 = 0$; Assumption 28). Let $U \sim \mathcal{N}(0,1)$. Then, $Z|H_i \overset{d}{=} (\sqrt{\mathbf{b}^\mathsf{T}\overline{\Sigma}_i\mathbf{b}})U$, where $\overset{d}{=}$ means equality in the distribution. From this fact, we now have $\Pr(\widehat{H} = H_2|H_1) = \Pr\left(\widehat{\tau} > U^2\right)$ and $\Pr(\widehat{H} = H_1|H_2) = \Pr(U^2 > \widehat{\tau}\widehat{R})$, where $\widehat{\tau} = \ln(\widehat{R})/(\widehat{R} - 1)$. Since $U^2 \sim \chi^2(1)$, we finally have

$$\widehat{\mathbb{P}}_{e_v}(\mathcal{J}) = 0.5\left[1 - Q_{\chi^2}\left(1, \widehat{\tau}\right)\right] + 0.5\, Q_{\chi^2}(1, \widehat{\tau}\widehat{R}).$$

To simplify $\widehat{R}$, note the following: since $\mathbf{b}$ is the maximizer of I-divergence (4.7), from [70], we can also express $\widehat{R}$ as

$$\widehat{R} = \frac{\mathbf{b}^\mathsf{T}\overline{\Sigma}_1\mathbf{b}}{\mathbf{b}^\mathsf{T}\overline{\Sigma}_2\mathbf{b}} = \max_{\mathbf{d}\in\mathbb{R}^{mN}} \frac{\mathbf{d}^\mathsf{T}\overline{\Sigma}_1\mathbf{d}}{\mathbf{d}^\mathsf{T}\overline{\Sigma}_2\mathbf{d}}.$$

Let $\mathbf{c} = \overline{\Sigma}_2^{1/2}\mathbf{d}$, and note the following:

$$\begin{aligned}\widehat{R} &= \max_{\mathbf{c}\in\mathbb{R}^{mN}} \left(\frac{\mathbf{c}}{\|\mathbf{c}\|_2}\right)^{\mathsf{T}} \overline{\Sigma}_2^{-1/2}\overline{\Sigma}_1\overline{\Sigma}_2^{-1/2}\left(\frac{\mathbf{c}}{\|\mathbf{c}\|_2}\right)\\ &= \lambda_{\max}\left(\overline{\Sigma}_2^{-1/2}\overline{\Sigma}_1\overline{\Sigma}_2^{-1/2}\right) = \lambda_{\max}\left(\overline{\Sigma}_1\overline{\Sigma}_2^{-1}\right).\end{aligned}$$

Since $\widehat{R}$ is an increasing sequence, with respect to N (see Proposition A.43), the limits $R = \lim_{N\to\infty}\widehat{R}$, $\tau = \lim_{N\to\infty}\widehat{\tau}$ and $\lim_{N\to\infty}\widehat{\tau}\widehat{R} = \tau R$ are well defined. Now, consider

$$\begin{aligned}\mathbb{P}_{e_v}(\mathcal{J}) &= \lim_{N\to\infty}\widehat{\mathbb{P}}_{e_v}(\mathcal{J})\\ &= \lim_{N\to\infty} 0.5\left[1 - Q_{\chi^2}\left(1,\widehat{\tau}\right)\right] + 0.5\,Q_{\chi^2}(1,\widehat{\tau}\widehat{R})\\ &= 0.5\left[1 - Q_{\chi^2}\left(1,\tau\right)\right] + 0.5\,Q_{\chi^2}(1,\tau R),\end{aligned}$$

where the last equality follows because $\widehat{\tau}$ and $\widehat{\tau}\widehat{R}$ are decreasing and increasing in N (Proposition A.43), respectively. We now show that R is given by (4.14). Since $\Sigma_2 = 0$ and $\mathbf{x}[0] = 0$, we have $\overline{\Sigma}_2 = \sigma_v^2 I$ and $\overline{\Sigma}_1 = FF^{\mathsf{T}} + \sigma_v^2 I$, where $F = \mathcal{F}(I_N \otimes \Sigma_1^{\frac{1}{2}})$ and $\Sigma_1^{\frac{1}{2}}$ satisfies $\Sigma_1 = \Sigma_1^{\frac{1}{2}}\Sigma_1^{\frac{1}{2}}$. From these observations, we may conclude that

$$\begin{aligned}R = \lim_{N\to\infty}\widehat{R} &= \lim_{N\to\infty}\frac{\lambda_{\max}(FF^{\mathsf{T}} + \sigma_v^2 I)}{\sigma_v^2}\\ &= 1 + \sigma_v^{-2}\lim_{N\to\infty}\lambda_{\max}(FF^{\mathsf{T}}).\end{aligned}\tag{4.30}$$

It now suffices to evaluate $\lim_{N\to\infty}\lambda_{\max}(FF^{\mathsf{T}}$. Since $\overline{\lambda}(G) < 1$, we may define the following matrix valued function [31]:

$$A(\omega) = \sum_{l=0}^{\infty} K(l)\Sigma_1^{1/2}e^{jk\omega} \quad \omega \in [0, 2\pi],$$

where $K(l) = CG^l\Pi$ and $j = \sqrt{-1}$. Since the coefficients $K(l)\Sigma_1^{1/2}$ are absolutely summable, for any $l \in \mathbb{N}$, these coefficients can also be recovered as [31]:

$$K(l)\Sigma_1^{1/2} = \frac{1}{2\pi}\int_0^{2\pi} A(\omega)e^{-jl\omega}d\omega.$$

Let $\overline{z}$ be the conjugate of $z \in \mathbb{C}$. Then, from [7, Chapter 6.4], we have

$$\begin{aligned}\lim_{N\to\infty} \lambda_{\max}^{1/2}(FF^{\mathsf{T}}) &= \operatorname*{ess\ sup}_{\omega\in[0,2\pi]} \|A(\omega)\|_2 \\ &= \operatorname*{ess\ sup}_{\omega\in[0,2\pi]} \left\| C\left(\sum_{l=0}^{\infty} G^l e^{jl\omega}\right)\Pi\Sigma_1^{1/2}\right\|_2 \\ &= \operatorname*{ess\ sup}_{\omega\in[0,2\pi]} \left\| C\left(I - Ge^{jw}\right)^{-1}\Pi\Sigma_1^{1/2}\right\|_2 \\ &= \operatorname*{ess\ sup}_{\{z\in\mathbb{C}:|z|=1\}} \left\| C\left(\overline{z}I - G\right)^{-1}\Pi\Sigma_1^{1/2}\right\|_2 \\ &\overset{(a)}{=} \operatorname*{ess\ sup}_{\{z\in\mathbb{C}:|z|=1\}} \left\| C\left(zI - G\right)^{-1}\Pi\Sigma_1^{1/2}\right\|_2 \\ &= \operatorname*{ess\ sup}_{\{z\in\mathbb{C}:|z|=1\}} \left\| T(z)\Sigma_1^{\frac{1}{2}}\right\|_2 = ||T(z)\Sigma_1^{\frac{1}{2}}||_\infty.\end{aligned}$$

where (a) follows because, for any $A \in \mathbb{C}^{N\times N}$ with A^* denoting its complex conjugate transpose, $\|A\|_2 = \|A^{\mathsf{T}}\|_2 = \|A^*\|_2$. Substituting $\lim_{N\to\infty}\lambda_{\max}^{1/2}(FF^{\mathsf{T}})$ in (4.30) gives us $R = 1 + \sigma_v^{-2}||T^*(z)||_\infty^2$. □

Proof of Theorem 35: Let $\mathbf{y}_{\mathcal{P}}[k]$ and $\mathbf{y}_{\mathcal{S}}[k]$ denote the measurements of associated with the sensor sets $\mathcal{P}$ and $\mathcal{C}$, respectively. Since $\sigma_v^2 = 0$, from (4.18), we have

$$\mathbf{y}_{\mathcal{P}}[k+1] = G_{pp}\mathbf{y}_{\mathcal{P}}[k] + B\mathbf{y}_{\mathcal{C}}[k], \tag{4.31}$$

where $B = G_{pc}$. From (4.31), it follows that

$$\underbrace{\begin{bmatrix}\mathbf{y}_{\mathcal{P}}[1]\\ \mathbf{y}_{\mathcal{P}}[2]\\ \vdots\\ \mathbf{y}_{\mathcal{P}}[N]\end{bmatrix}}_{\mathbf{Y}_{\mathcal{P}}} = \underbrace{\begin{bmatrix}G_{pp} & B\\ G_{pp}^2 & G_{pp}B\\ \vdots & \vdots\\ G_{pp}^N & G_{pp}^{N-1}B\end{bmatrix}}_{M} \underbrace{\begin{bmatrix}\mathbf{y}_{\mathcal{P}}[0]\\ \mathbf{y}_{\mathcal{C}}[0]\end{bmatrix}}_{\widehat{\mathbf{Y}}[0]} + \underbrace{\begin{bmatrix}0 & 0 & \cdots & 0 & 0\\ B & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ G_{pp}^{N-2}B & G_{pp}^{N-3}B & \cdots & B & 0\end{bmatrix}}_{\widehat{M}} \underbrace{\begin{bmatrix}\mathbf{y}_{\mathcal{C}}[1]\\ \mathbf{y}_{\mathcal{C}}[2]\\ \vdots\\ \mathbf{y}_{\mathcal{C}}[N]\end{bmatrix}}_{\mathbf{Y}_{\mathcal{C}}}.$$

Since $\widehat{\mathbf{Y}}[0]$ is independent of H_i, the assertion of the theorem follows from Lemma 36. $\square$.

Proof of Lemma 36 We shall prove the result assuming that $\mathbf{Y}$ and $\mathbf{Z} = g(\mathbf{Y}) + \mathbf{v}$ admits density functions. With the expense of notation, the given proof can be adapted to handle random variables that do not have densities. Let $\mathbf{L} = [\mathbf{Y}^\mathsf{T}, \mathbf{Z}^\mathsf{T}]^\mathsf{T}$. Consider the following log likelihood ratio (LR) based on $\mathbf{L}$:

$$\begin{aligned}\frac{f(l|M_2)}{f(l|M_1)} &= \frac{f(y, g(y) + v|M_2)}{f(y, g(y) + v|M_1)} \\ &= \frac{f(y, g(y) + v|y, M_2)f(y|M_2)}{f(y, g(y) + v|y, M_1)f(y|M_1)} \\ &\overset{(a)}{=} \frac{f(y, g(y) + v|y)f(y|M_2)}{f(y, g(y) + v|y)f(y|M_1)} = \frac{f(y|M_2)}{f(y|M_1)},\end{aligned}$$

where (a) follows because $\mathbf{v}$ is independent of M_i. Since LRs of $\mathbf{L}$ and $\mathbf{Y}$ are equal, the error probabilities associated with their MAP rules should be the same. Instead, error probability of the MAP rule based on $\mathbf{L}$ is always superior to that of $\mathbf{Y}$ or $\mathbf{Z}$ alone. Thus $\delta_1 \leq \delta_2$. $\square$

Proof of Theorem 38: Consider the following deterministic analogue of (4.1): $\mathbf{x}[k+1] = G\mathbf{x}[k] + \Pi\mathbf{u}$, where $\mathbf{u}$ is arbitrary. Recall that $\mathbf{x}_p[k+1] = G_{pp}\mathbf{x}_p[k] + G_{pc}\mathbf{x}_c[k]$ (4.18). Since $\mathbf{x}[0] = 0$, for $z \notin \text{spec}(G) \cup \text{spec}(G_{pp})$, we have

$$\mathbf{x}[z] = (zI - G)^{-1}\Pi\mathbf{u} \text{ and} \tag{4.32a}$$

$$\mathbf{x}_p[z] = (zI - G_{pp})^{-1}G_{pc}\mathbf{x}_c[z] = T_s(z)\mathbf{x}_c[z]. \tag{4.32b}$$

From (4.32b), the following inequalities are obvious

$$\underline{\rho}(z)\|\mathbf{x}_c[z]\|_2 \leq \|\mathbf{x}_p[z]\|_2 \leq \overline{\rho}(z)\|\mathbf{x}_c[z]\|_2. \tag{4.33}$$

Let C_1 and C_2 be the sensor matrices associated with $\mathcal{C}$ and $\mathcal{P}$, respectively. Then,

$$\mathbf{x}_c[z] = C_1\mathbf{x}[z] \text{ and } \mathbf{x}_p[z] = C_1\mathbf{x}[z]. \tag{4.34}$$

part 1) We now consider the cases 1a) and 1b). Let $L_i = C_i(I-G)^{-1}\Pi\Sigma_c^{\frac{1}{2}}$, where $\Sigma_c = \Sigma_c^{\frac{1}{2}}\Sigma_c^{\frac{1}{2}}$ is defined in Lemma 27. Let $z=1$. Then, from (4.34) note that

$$\|\mathbf{x}_c[1]\|_2^2 = \|C_1\mathbf{x}[1]\|_2^2 = \mathbf{u}^\mathsf{T} L_1^\mathsf{T} L_1 \mathbf{u} \text{ and}$$

$$\|\mathbf{x}_p[1]\|_2^2 = \|C_2\mathbf{x}[1]\|_2^2 = \mathbf{u}^\mathsf{T} L_2^\mathsf{T} L_2 \mathbf{u}.$$

From (4.33) and above identities, it follows that

$$\begin{aligned} &\overline{\rho}(1) < 1 \implies L_1^\mathsf{T} L_1 + \sigma_v^2 I \succ L_2^\mathsf{T} L_2 + \sigma_v^2 I \text{ and} \\ &\underline{\rho}(1) > 1 \implies L_2^\mathsf{T} L_2 + \sigma_v^2 I \succ L_1^\mathsf{T} L_1 + \sigma_v^2 I. \end{aligned} \tag{4.35}$$

Let $\mathbf{u} = \widetilde{\boldsymbol{\mu}}_\Delta$, where $\widetilde{\boldsymbol{\mu}}_\Delta$ is defined in the statement of Lemma 29. Let η_1 and η_2 be the SNRs of $\mathbb{P}_{e_m}(\mathcal{C})$ and $\mathbb{P}_{e_m}(\mathcal{P})$, respectively. Then from (4.13), we have

$$\eta_i^2 = N\widetilde{\boldsymbol{\mu}}_\Delta^\mathsf{T}\left([L_i^\mathsf{T} L_i + \sigma_v^2 I]^{-1} L_i^\mathsf{T} L_i\right)\widetilde{\boldsymbol{\mu}}_\Delta.$$

Using the identity $[L_i^\mathsf{T} L_i + \sigma_v^2 I]^{-1} L_i^\mathsf{T} L_i = I - \sigma_v^2 [L_i^\mathsf{T} L_i + \sigma_v^2 I]^{-1}$, we can also express η_i^2 as

$$\eta_i^2 = \widetilde{\boldsymbol{\mu}}_\Delta^\mathsf{T}\widetilde{\boldsymbol{\mu}}_\Delta - \sigma_v^2\,\widetilde{\boldsymbol{\mu}}_\Delta^\mathsf{T}\left[L_i^\mathsf{T} L_i + \sigma^2 I\right]^{-1}\widetilde{\boldsymbol{\mu}}_\Delta^\mathsf{T}. \tag{4.36}$$

Finally, from (4.36) and (4.35), and Proposition 31, we have

$$\overline{\rho}(1) < 1 \implies \eta_1^2 \geq \eta_2^2 \implies \mathbb{P}_{e_m}(\mathcal{C}_d) \leq \mathbb{P}_{e_m}(\mathcal{P}) \text{ and}$$

$$\underline{\rho}(1) > 1 \implies \eta_1^2 \leq \eta_2^2 \implies \mathbb{P}_{e_m}(\mathcal{C}_d) \geq \mathbb{P}_{e_m}(\mathcal{P}).$$

part 2) We now consider the cases 2a) and 2b). Let $T_i(z) = C_i(zI-G)^{-1}$. Let $\mathbf{u} = \Sigma_1^{1/2}\mathbf{d}$, where $\Sigma_1^{1/2}$ is defined in the statement of Lemma 29. From (4.34) and (4.32a), we have $\mathbf{x}_c[k] = T_1(z)\Sigma_1^{1/2}\mathbf{d}$ and $\mathbf{x}_c[k] = T_2(z)\Sigma_1^{1/2}\mathbf{d}$. By invoking these two facts in (4.33), we may now conclude that

$$\sup_{|z|=1} \overline{\rho}(z) < 1 \implies \|T_2(z)\Sigma_1^{\frac{1}{2}}\mathbf{d}\|_2 \leq \|T_1(z)\Sigma_1^{\frac{1}{2}}\mathbf{d}\|_2 \text{ and}$$

$$\inf_{|z|=1} \underline{\rho}(z) > 1 \implies \|T_2(z)\Sigma_1^{\frac{1}{2}}\mathbf{d}\|_2 \geq \|T_1(z)\Sigma_1^{\frac{1}{2}}\mathbf{d}\|_2,$$

for all z that satisfies $|z| = 1$. Let R_1 and R_2 be the SNRs of $\mathbb{P}_{e_v}(\mathcal{C})$ and $\mathbb{P}_{e_v}(\mathcal{P})$, respectively. Then, from (4.14)

$$R_i - 1 = \frac{\|T_i(z)\Sigma_1^{\frac{1}{2}}\|_\infty^2}{\sigma_v^2} = \left(\underset{\{z\in\mathbb{C}:|z|=1\}}{\text{ess sup}} \|T_2(z)\Sigma_1^{1/2}\mathbf{d}\|_2\right)^2.$$

From Proposition 31, it follows that

$$\sup_{|z|=1} \overline{\rho}(z) < 1 \implies R_1 \geq R_2 \implies \mathbb{P}_{e_v}(\mathcal{C}_d) \leq \mathbb{P}_{e_v}(\mathcal{P}) \text{ and}$$

$$\inf_{|z|=1} \underline{\rho}(z) > 1 \implies R_1 \leq R_2 \implies \mathbb{P}_{e_v}(\mathcal{C}_d) \geq \mathbb{P}_{e_v}(\mathcal{P}). \quad \square$$

Proof of Corollary 39: We shall prove part 1) of the corollary, and part 2) can be derived using similar analysis (the details are omitted). The idea of the proof is to show that $||\widetilde{G}||_\infty \leq 1/\sqrt{m} \implies \overline{\rho}(1) < 1$, and thereby invoking Theorem 38 yields the result.

step 1) For $G \geq 0$, it follows that $\sup_{|z|=1} \overline{\rho}(z) = \overline{\rho}(1)$, where $\overline{\rho}(z)$ is $\|(zI - G_{pp})^{-1}G_{pc}\|_2$. To see this, note the following: For any $\mathbf{d} \in \mathbb{C}^{n_1}$, let $|\mathbf{d}| = (|d_1|, \dots, |d_{n_1}|)^\mathsf{T}$. Then, for any $l \in \mathbb{N}$ and z that satisfies $|z| = 1$, we have

$$|(\overline{z}G_{pp})^l G_{pc}\mathbf{d}| = |(G_{pp})^l G_{pc}\mathbf{d}| \leq (G_{pp})^l G_{pc}|\mathbf{d}|,$$

where the inequality, to be understood coordinate wise, follows because $[G_{pp}\, G_{pc}] \geq 0$. From the above inequality, and the fact $|\mathbf{y} + \mathbf{z}| \leq |\mathbf{y}| + |\mathbf{z}|$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{C}^p$, we have

$$\left|\sum_{l=0}^{\infty} (\overline{z}G_{pp})^l G_{pc}\mathbf{d}\right| \leq \sum_{l=0}^{\infty} |(\overline{z}G_{pp})^l G_{pc}\mathbf{d}| \leq \sum_{l=0}^{\infty} (G_{pp})^l G_{pc}|\mathbf{d}|.$$

Since G_{pp} is a submatrix of $G \geq$, which is a non-negative matrix, we have $|\lambda_{\max}(\overline{z}G_{pp})| = |\lambda_{\max}(G_{pp})| \leq |\lambda_{\max}(G)| < 1$ [5], the above inequality can also be expressed as

$$\left|(I - \overline{z}G_{pp})^{-1}G_{pc}\mathbf{d}\right| \leq (I - G_{pp})^{-1}G_{pc}\,|\mathbf{d}|\,.$$

Taking 2-norm on both sides of the inequality yields us:

$$\left\|\,\left|(I - \overline{z}G_{pp})^{-1}G_{pc}\mathbf{d}\right|\,\right\|_2 \leq \left\|\,(I - G_{pp})^{-1}G_{pc}\,|\mathbf{d}|\,\right\|_2.$$

Since the above inequality holds for any vector $\mathbf{d} \in \mathbf{R}^{n_1}$, using the identity $\|\,|\mathbf{x}|\,\|_2 = \|\,\mathbf{x}\,\|_2$ for any $\mathbf{x} \in \mathbb{C}^p$, the following inequality is now obvious:

$$\sup_{|z|=1}\ \sup_{\|\mathbf{d}\|_2=1} \|(I - \overline{z}G_{pp})^{-1}G_{pc}\mathbf{d}\|_2 \leq \left\|\,(I - G_{pp})^{-1}G_{pc}\,\right\|_2,$$

which can be expressed as $\sup_{|z|=1}\overline{\rho}(z) \leq \overline{\rho}(1)$. The equality is attained at $z = 1$.

step 2) Since $\sup_{|z|=1}\overline{\rho}(z) = \overline{\rho}(1)$, from Theorem 38 it readily follows that, both $\mathbb{P}_{e_m}(\mathcal{C}_d) \leq \mathbb{P}_{e_m}(\mathcal{P})$ and $\mathbb{P}_{e_v}(\mathcal{C}_d) \leq \mathbb{P}_{e_v}(\mathcal{P})$ holds true whenever $\overline{\rho}(1) < 1$. We now show that $\|\widetilde{G}\|_\infty = \|[G_{pp}\,G_{pc}]\|_\infty < 1/\sqrt{m_1}$ guarantees $\rho(1) < 1$. Let $\mathbf{1}$ denote the all ones vector, and note the following identity:

$$\begin{bmatrix} G_{pp} & G_{pc} \\ 0 & I \end{bmatrix}^k \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix} = \begin{bmatrix} G_{pp}^k & \sum_{l=0}^{k-1} G_{pp}^l G_{pc} \\ 0 & I \end{bmatrix}^k \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix}. \tag{4.37}$$

Since $\|[G_{pp}\ G_{pc}]\|_\infty < 1/\sqrt{m_1}$, for any $k \in \mathbb{N}$, we also have

$$\begin{bmatrix} G_{pp} & G_{pc} \\ 0 & I \end{bmatrix}^k \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix} \leq \begin{bmatrix} \frac{1}{\sqrt{m_1}}\mathbf{1} \\ \mathbf{1} \end{bmatrix}.$$

From the above inequality and (4.37), it follows that

$$\lim_{k\to\infty} \begin{bmatrix} G_{pp}^k & \sum_{l=0}^{k-1} G_{pp}^l G_{pc} \\ 0 & I \end{bmatrix}^k \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix} \leq \begin{bmatrix} \frac{1}{\sqrt{m_1}}\mathbf{1} \\ \mathbf{1} \end{bmatrix}.$$

Since $|\lambda_{\max}(G_{pp})| < 1$, as $k \to \infty$, we have $G_{pp}^k \to 0$ and $\sum_{l=0}^{k-1} G_{pp}^l G_{pc} \to (I - G_{pp})^{-1} G_{pc}$. Thus $(I - G_{pp})^{-1} G_{pc}\mathbf{1} = \|(I - G_{pp})^{-1} G_{pc}\|_\infty < 1/\sqrt{m_1}$, and hence, $\overline{\rho}(1) = ||(I - G_{pp})^{-1} G_{pc}||_2 < \sqrt{m_1}||(I - G_{pp})^{-1} G_{pc}||_\infty < 1$. □

Proof of Proposition 31: Since $Q_{\mathcal{N}}(x)$ is decreasing function of x, $\mathbb{P}_{e_m}(\mathcal{J})$ (4.11) is decreasing in SNR η, given by either (4.13) or (4.15). For $\mathbb{P}_{e_v}(\mathcal{J})$ (4.12) note the following: first, observe that $R > 1$ in both (4.14) and (4.16). Thus

$$\begin{aligned} \frac{d\tau}{dR} &= \frac{\left(\frac{R-1}{R}\right) - \ln R}{(R-1)^2} < 0, \text{ and} \\ \frac{d(\tau R)}{dR} &= \frac{(R-1) - \ln R}{(R-1)^2} > 0. \end{aligned} \tag{4.38}$$

Hence, we conclude that τ is deceasing in R. Instead, τR is increasing in R. From this observation and the fact that the $Q_{\chi^2}(1, z) = \Pr[Z \geq z]$, where $Z \sim \chi^2(1)$, is decreasing in z, it follows that $\mathbb{P}_{e_v}(\mathcal{J})$ is decreasing in R. □

Proof of Proposition 42: From (4.19), and the fact that $1 \leq q < j < \ldots < n$, where $\mathcal{C}_d = \{j\}$ and $\mathcal{P} = \{j+1, \ldots, n\}$, the row sums of $\widetilde{G}$ takes values in the set $\{a + c, a + b + c\}$. Let $|\overline{G}_{l,q}| = |(I - G)^{-1}_{l,q}|$. Using the principle of backward induction, we shall show that, when $\|\widetilde{G}\|_\infty = a + b + c < 1$, $\{|\overline{G}_{lq}|\}_{l=q}^n$ is monotonically decreasing. The proof of part (ii) is left to the reader as an exercise. Let $\widetilde{a} = 1 - a$, $\widetilde{b} = -b$, $\widetilde{c} = -c$. If $\widetilde{a} \neq 0$, then $\overline{G}_{l,q}$ of $(I - G)^{-1}$ are given by the following expressions [43]:

$$\overline{G}_{l,q} = \frac{1}{\theta_n} \begin{cases} (-1)^{l+q} \widetilde{b}^{q-l} \theta_{l-1} \phi_{q+1} & q \geq l \\ (-1)^{l+q} \widetilde{c}^{l-q} \theta_{q-1} \phi_{l+1} & q < l \end{cases} \tag{4.39}$$

where $l, q \in \{1, \ldots, n\}$, and θ_k and ϕ_k are governed by

$$\begin{aligned} \theta_k &= \widetilde{a}\theta_{k-1} - \widetilde{b}\widetilde{c}\theta_{k-2} \quad \text{for } k = 2, \ldots, n \\ \phi_k &= \widetilde{a}\phi_{k+1} - \widetilde{b}\widetilde{c}\phi_{k+2} \quad \text{for } k = n-1, \ldots, 1 \end{aligned} \tag{4.40}$$

where $\theta_0 = 1$, $\theta_1 = \widetilde{a}$, $\phi_n = \widetilde{a}$, $\phi_{n+1} = 1$ and $\theta_n = \det(I-G)$. Let $\mathcal{L} = \{q+1, \ldots, n\}$. Then, for any $l \in \mathcal{L} \cup \{q\}$,

$$|\overline{G}_{lq}| \triangleq |(I-G)^{-1}_{lq}| = \left|\frac{\theta_{q-1}\widetilde{c}^{-q}}{\theta_n}\right| \left|\widetilde{c}^{l}\phi_{l+1}\right|,$$

Let $l \in \mathcal{L}$, and define $\zeta(l) = |\overline{G}_{l,q}|/|\overline{G}_{l-1,q}|$. Since $\phi_{n+1} = 1$ and $\phi_n = \widetilde{a}$, for $l = n$ (base step), it follows that

$$\zeta(n) = \frac{|\overline{G}_{nq}|}{|\overline{G}_{n-1,q}|} = \frac{|\widetilde{c}^{n}\phi_{n+1}|}{|\widetilde{c}^{n-1}\phi_n|} = \frac{|\widetilde{c}|}{|\widetilde{a}|} = \frac{c}{1-a} \overset{(i)}{<} 1,$$

where (i) follows because $a, b, c > 0$, and $a + b + c < 1$. Let $q < l < n$ and $\zeta(l+1) < 1$ (inductive step). Then,

$$\zeta(l) = \frac{|\overline{G}_{l,q}|}{|\overline{G}_{l-1,q}|} = \frac{|\widetilde{c}|\,|\phi_{l+1}|}{|\phi_l|} \overset{(4.40)}{=} \frac{c}{\left|\widetilde{a} - \widetilde{b}\widetilde{c}\left(\frac{\phi_{l+2}}{\phi_{l+1}}\right)\right|} < 1,$$

To see the last inequality, consider the following:

$$\begin{aligned} b + c < 1 - a &\overset{(ii)}{\Longrightarrow} b\left(\frac{|\widetilde{c}||\phi_{l+2}|}{|\phi_{l+1}|}\right) + c < 1 - a \\ &\Longrightarrow b\left(\frac{c\phi_{l+2}}{\phi_{l+1}}\right) + c < 1 - a \\ &\Longrightarrow \frac{c}{\left|(1-a) - bc\left(\frac{\phi_{k+2}}{\phi_{k+1}}\right)\right|} < 1 \\ &\overset{(iii)}{\Longrightarrow} \frac{|\widetilde{c}|}{\left|\widetilde{a} - \widetilde{b}\widetilde{c}\left(\frac{\phi_{l+2}}{\phi_{l+1}}\right)\right|} < 1, \end{aligned}$$

where (ii) follows because the hypothesis $\zeta(l+1) < 1$ implies that $|\widetilde{c}|(|\phi_{l+2}|/|\phi_{l+1}|) < 1$, and (iii) from the fact that $\widetilde{a} = 1 - a$, $\widetilde{b} = -b$, and $\widetilde{c} = -c$. From the principle of finite induction, for all $l \in \mathcal{L}$, we have $\zeta(l) < 1$. Hence, $\{|\overline{G}_{lq}|\}_{l=q}^{n}$ is a decreasing sequence. □

Proposition A.43 *Let $\widehat{\eta}^2 = \overline{\boldsymbol{\mu}}_\Delta^\intercal \overline{\Sigma}_c^{-1} \overline{\boldsymbol{\mu}}_\Delta$, $\widehat{R} = \lambda_{max}(\overline{\Sigma}_1 \overline{\Sigma}_2^{-1})$ and $\widehat{\tau} = \ln(\widehat{R})/(\widehat{R}-1)$, where $(\overline{\boldsymbol{\mu}}_\Delta$, $\overline{\Sigma}_c$, $\overline{\Sigma}_1$, $\overline{\Sigma}_2)$ are defined in the statement of Lemma 27. Then, $\widehat{\eta}$, $\widehat{R}$, and τ are increasing in N. However, $\widehat{\tau}\widehat{R}$ is decreasing in N.*

Proof. Let $N < \infty$. Then, from Proposition 26, we have $\overline{\boldsymbol{\mu}}_{\Delta} = \mathbb{E}\left[\mathbf{Y}_{\mathcal{J}}|H_2\right] - \mathbb{E}\left[\mathbf{Y}_{\mathcal{J}}|H_1\right]$, $\overline{\Sigma}_c = \mathrm{Cov}[\mathbf{Y}_{\mathcal{J}}|H_1] = \mathrm{Cov}[\mathbf{Y}_{\mathcal{J}}|H_2]$. For clarity, we drop the existing subscripts and replace them with the total number of measurements. Let $N_2 = N_1 + k$, $k \in \mathbb{N}$, and consider $\mathbf{Y}_{N_2}^{\mathsf{T}} = \left[\mathbf{Y}_{N_1}^{\mathsf{T}},\, \mathbf{Z}_k^{\mathsf{T}}\right]$, where $\mathbf{Z}_k$ are the measurements collected after N_1. Then,

$$\overline{\boldsymbol{\mu}}_{N_2} = \left[\begin{array}{c} \overline{\boldsymbol{\mu}}_{N_1} \\ \hline \mathbf{m}_k \end{array}\right] \text{ and } \overline{\Sigma}_{N_2} = \left[\begin{array}{c|c} \overline{\Sigma}_{N_1} & D \\ \hline D^{\mathsf{T}} & M \end{array}\right],$$

where $\mathbf{m}_k = \mathbb{E}[\mathbf{Z}_k|H_2] - \mathbb{E}[\mathbf{Z}_k|H_1]$, $M = \mathrm{Cov}[\mathbf{Z}_k|H_1] > 0$, and $D = \mathrm{Cov}[\mathbf{Y}_k, \mathbf{Z}_k|H_1]$. Further, using the Schur complement, $\overline{\Sigma}_{N_2}^{-1}$ can be expressed as

$$\overline{\Sigma}_{N_2}^{-1} = \left[\begin{array}{c|c} \overline{\Sigma}_{N_1} & D \\ \hline D^{\mathsf{T}} & M \end{array}\right]^{-1} = \left[\begin{array}{c|c} \overline{\Sigma}_{N_1}^{-1} & 0 \\ \hline 0^{\mathsf{T}} & 0 \end{array}\right] + \underbrace{F}_{>0}.$$

From the above identity, it follows that

$$\begin{aligned} \widehat{\eta}_{N_2} = \left(\overline{\boldsymbol{\mu}}_{N_2}^{\mathsf{T}} \overline{\Sigma}_{N_2}^{-1} \overline{\boldsymbol{\mu}}_{N_2}\right)^{\frac{1}{2}} &= \left(\overline{\boldsymbol{\mu}}_{N_1}^{\mathsf{T}} \overline{\Sigma}_{N_1}^{-1} \overline{\boldsymbol{\mu}}_{N_1} + \overline{\boldsymbol{\mu}}_{N_2}^{\mathsf{T}} F \overline{\boldsymbol{\mu}}_{N_2}\right)^{\frac{1}{2}} \\ &\geq \left(\overline{\boldsymbol{\mu}}_{N_1}^{\mathsf{T}} \overline{\Sigma}_{N_1}^{-1} \overline{\boldsymbol{\mu}}_{N_1}\right) = \widehat{\eta}_{N_1}. \end{aligned}$$

Hence, we may conclude that $\widehat{\eta}$ is increasing in N. Instead, from the eigenvalue interlacing property for the symmetric matrix pencils [44], it follows that $\widehat{R} = \lambda_{\max}(\overline{\Sigma}_1 \overline{\Sigma}_2^{-1})$ is increasing in N. Finally, from (4.38), it follows that $\widehat{\tau}$ and $\widehat{\tau}\widehat{R}$ are decreasing and increasing in N, respectively. ■

Chapter 5

Conclusions and Future Work

This dissertation is concerned with security of dynamical systems with emphasis on interconnected and networked type dynamical systems. For interconnected systems, in chapter 2, we highlighted the trade-offs between the performance of centralized and decentralized detectors, and showed that a decentralized detector can outperform its centralized counterpart, despite having less information about the system dynamics, and this property is due to the nature of the considered attack detection problem. We illustrated our findings on a real time power system model. Instead, in chapter 3, we developed a probabilistic rule to randomly select a subsystem to attack, over time, and optimize over the switching probabilities to maximize degradation and maintain undetectability from a centralized detector. Overall, our results show that the ability to selectively compromise different parts of a system over time greatly increases the severity of the attacks, thereby motivating the development of advanced detection schemes for interconnected type dynamical systems.

For networked dynamical systems, in chapter 4, we studied a problem of detecting changes in the statistical properties of an input driving certain network nodes has to be detected by sparse and remotely located sensors. We explicitly derive the Maximum A Posteriori (MAP) detector, and characterize its performance as a function of the network parameters, and the location of the sensor nodes. We show that, in the absence of measurement noise, the detection performance obtained when sensors are located on a network cut is not worse than the performance obtained by measuring all nodes of the subnetwork induced by the cut and not containing the input node. Conversely, in the presence of measurement noise, we show that the detection performance may increase or decrease with the graphical distance between the input node and the sensors. Our analysis provided structural insights into the optimal sensor placement problem. Opportunities for future work, with regards to the current results in this thesis, includes the following:

- In chapter 1, we worked with a particular detection scheme and characterized performance degradation due to undetectable detects. Several questions remain of interest for future investigation, including the characterization of optimal detection schemes, an analytical comparison of the degradation induced by undetectable attacks as a function of the detection scheme, and the analysis of iterative detection strategies.
- In chapter 2, we were only able to demonstrate the superiority of probabilistic attacks over a deterministic attack through numerical examples (the main difficulty being the optimal probabilistic rule turned out to be a solution of non-convex optimization problem). As seen in the formulation of (P.2), one can see that, the optimal probabilistic rule depends on the subsystem dynamics, interconnection signals, and the choice of attack matrices. Understanding the role of these quantities on the optimal policy and providing exact theoretical characterizations is also of interest.

- In chapter 3, the analytical expressions of error probabilities are limited only to special cases of general Gaussian detection problem. Further, we considered only time invariant network models with deterministic edge weights. It would be interesting to extend our network analysis to the time varying networks with random edge weights, and also to the case where the input signal is non-Gaussian.

We end this section by commenting on the security aspects of dynamical systems whose mathematical models are not necessarily given by the classical physics principles, but are obtained using data-driven algorithms. For instance, Reinforcement or Statistical learning based methods. As a result, these models will have completely different mathematical as compared to thee existing parametric models. Hence, it is far from clear if the characterizations provided in the thesis, or existing characterizations in the cyber-physical security literature, extend naturally to these so called data-driven dynamical systems. As a preliminary step towards this line of research, we are currently developing a theoretical framework to understand the robustness of data driven closed loop systems using tools from non-asymptotic random matrix theory. In particular, we are trying to come up with sharp bounds on the *probability of stability* of a random closed loop LTI system, whose feedback matrix is random. The hope is that using these bounds we can understand the role of system theoretic properties on the behavior of the spectral radius of the closed loop system. The stability of random matrices is still an open problem, and we firmly believe the following spirit shared by authors in [9]: *the one who adventures in the field of random matrix theory, with emphasis on control applications, will encounter unexpected and exciting connections among different fields of science and beautiful branches of mathematics.*

Bibliography

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

[2] R. Anguluri, V. Gupta, and F. Pasqualetti. Periodic coordinated attacks against cyber-physical systems: Detectability and performance bounds. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 5079–5084, Dec 2016.

[3] S. Appadwedula, V. V. Veeravalli, and D. L. Jones. Energy-efficient detection in sensor networks. *IEEE Journal on Selected Areas in Communications*, 23(4):693–702, April 2005.

[4] M. Basseville and I. V. Nikiforov. *Detection of Abrupt Changes: Theory and Application.* Prentice Hall, 1993.

[5] A. Berman and R.J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences.* Society for Industrial and Applied Mathematics, 1994.

[6] Lucien Birgé. An alternative point of view on Lepski's method. *Institute of Mathematical Statistics*, 36:113–133, 2001.

[7] A. Böttcher and B. Silbermann. Block toeplitz matrices. In *Introduction to Large Truncated Toeplitz Matrices*, pages 185–219. Springer New York, 1999.

[8] Y. Bresler, V. U. Reddy, and T. Kailath. Optimum beamforming for coherent signal and interferences. *IEEE Trans. on Acou., Speech, and Signal Process.*, 36(6):833–843, June 1988.

[9] G. Calafiore and F. Dabbene. On the stability of random matrices. In *Unsolved Problems in Mathematical Systems and Control Theory*, pages 71–75. Princeton University Press, 2004.

[10] A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, and S. Sastry. Challenges for securing cyber physical systems. In *Workshop on Future Directions in Cyber-physical Systems Security.* DHS, July 2009.

[11] J. Chamberland and V. V. Veeravalli. Decentralized detection in sensor networks. *IEEE Transactions on Signal Processing*, 51(2):407–416, Feb 2003.

[12] A. Chamseddine, H. Noura, M. Ouladsine, and T. Raharijaona. Observability of complex systems: Minimal cost sensor network design. *IFAC Proceedings Volumes*, 41(2):13287–13292, 2008. 17th IFAC World Congress.

[13] Jie Chen and R.J. Patton. *Robust Model-Based Fault Diagnosis for Dynamic Systems.* Springer-Verlag New York, 1999.

[14] Y. Chen, S. Kar, and J. M. F. Moura. Dynamic attack detection in cyber-physical systems with side initial state information. *IEEE Transactions on Automatic Control*, 62(9):4618–4624, Sept 2017.

[15] Y. Chen, S. Kar, and J. M. F. Moura. Cyber-physical attacks with control objectives. *IEEE Transactions on Automatic Control*, 63(5):1418–1425, May 2018.

[16] Edwin K.P. Chong and Stanislaw H. Zak. *An Introduction to Optimization.* Wiley, New York, 2013.

[17] C. Commault, J. -M. Dion, and Do Hieu Trinh. Observability recovering by additional sensor implementation in linear structured systems. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 7193–7197, Dec 2005.

[18] Kressner Daniel. *Numerical Methods for General and Structure Eigenvalue Problems.* Springer-Verlag Berlin Heidelberg, 2005.

[19] C. L. DeMarco, J. V. Sariashkar, and F. Alvarado. The potential for malicious control in a competitive power systems environment. In *Proceeding of the 1996 IEEE International Conference on Control Applications IEEE International Conference on Control Applications held together with IEEE International Symposium on Intelligent Contro*, pages 462–467. Sep. 1996.

[20] N. K. Dhingra, M. R. Jovanović, and Z. Luo. An ADMM algorithm for optimal sensor and actuator selection. In *53rd IEEE Conf. on Decision and Control*, pages 4039–4044, Dec 2014.

[21] D. Ding, Q.-L. Han, Y. Xiang, X. Ge, and X.-M. Zhang. A survey on security control and attack detection for industrial cyber-physical systems. *Neurocomputing*, 275:1674 – 1683, 2018.

[22] F. Dörfler, F. Pasqualetti, and F. Bullo. Distributed detection of cyber-physical attacks in power networks: A waveform relaxation approach. In *Allerton Conf. on Communications, Control and Computing*, pages 1486–1491, Allerton, IL, USA, September 2011.

[23] F. Dörfler, F. Pasqualetti, and F. Bullo. Continuous-time distributed observers with discrete communication. *IEEE Journal of Selected Topics in Signal Processing*, 7(2):296–304, 2013.

[24] R. Duda, P. Hart, and D. Stork. *Pattern Classification.* New York: Wiley, 2000.

[25] A. K. Farraj, E. M. Hammad, A. A. Daoud, and D. Kundur. A game-theoretic control approach to mitigate cyber switching attacks in smart grid systems. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 958–963, Nov 2014.

[26] J. P. Farwell and R. Rohozinski. Stuxnet and the future of cyber war. *Survival*, 53(1):23–40, 2011.

[27] H. Fawzi, P. Tabuada, and S. Diggavi. Secure estimation and control for cyber-physical systems under adversarial attacks. *IEEE Transactions on Automatic Control*, 59(6):1454–1467, 2014.

[28] Hamza Fawzi, Paulo Tabuada, and Suhas Diggavi. Secure estimation and control for cyber-physical systems under adversarial attacks. *arXiv preprint arXiv:1205.5073*, 2012.

[29] R. M. G. Ferrari, T. Parisini, and M. M. Polycarpou. Distributed fault detection and isolation of large-scale discrete-time nonlinear systems: An adaptive approximation approach. *IEEE Transactions on Automatic Control*, 57(2), Feb 2012.

[30] J. Sun G. Wu and J. Chen. A survey on the security of cyber-physical systems. *Control Theory and Technology*, 14(1):2–10, Feb 2016.

[31] R. M. Gray. Toeplitz and circulant matrices: A review. *Foundations and Trends® in Communications and Information Theory*, 2(3):155–239, 2006.

[32] C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, S. Kuruganty, W. Li, R. Mukerji, D. Patton, N. Rau, D. Reppen, A. Schneider, M. Shahidehpour, and C. Singh. The IEEE reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee. *IEEE Transactions on Power Systems*, 14(3):1010–1020, Aug 1999.

[33] E. M. Hammad, A. K. Farraj, and D. Kundur. A resilient feedback linearization control scheme for smart grids under cyber-physical disturbances. In *2015 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5, Feb 2015.

[34] J.Chen and R. J. Patton. *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Springer Publishing Company, Incorporated, 2012.

[35] L. Jian and W. Chun-Yu. Finite-time robust fault detection filter design for interconnected systems concerning with packet dropouts and changing structures. *International Journal of Control*, pages 1–12, 2018.

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

[37] S. Joshi and S. Boyd. Sensor selection via convex optimization. *IEEE Trans., on Signal Process.*, 57(2):451–462, Feb 2009.

[38] V. Kekatos, G. B. Giannakis, and B. Wollenberg. Optimal placement of phasor measurement units via convex relaxation. *IEEE Transactions on Power Systems*, 27(3):1521–1530, Aug 2012.

[39] H. Kim, P. Guo, M. Zhu, and P. Liu. Attack-resilient estimation of switched nonlinear cyber-physical systems. In *2017 American Control Conference (ACC)*, pages 4328–4333, May 2017.

[40] M. N. Kurt, Y. Yılmaz, and X. Wang. Distributed quickest detection of cyber-attacks in smart grid. *IEEE Transactions on Information Forensics and Security*, 13(8):2015–2030, Aug 2018.

[41] C. Kwon and I. Hwang. Reachability analysis for safety assurance of cyber-physical systems against cyber attacks. *IEEE Transactions on Automatic Control*, 63(7):2272–2279, July 2018.

[42] A. Leedom. Stuxnet: Risk and uncertainty in the first salvo of global cyber warfare. *The SAIS Europe Journal*.

[43] J. W. Lewis. Inversion of tridiagonal matrices. *Numerische Mathematik*, 38(3):333–345, Oct 1982.

[44] R. C. Li. Rayleigh quotient based optimization methods for eigenvalue problems. In *Matrix Functions and Matrix Equations*, pages 76–108. World Scientific, 2015.

[45] C. Liu, J. Wu, C. Long, and Y. Wang. Dynamic state recovery for cyber-physical systems under switching location attacks. *IEEE Transactions on Control of Network Systems*, 4(1):14–22, March 2017.

[46] L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han. Detecting false data injection attacks on power grid by sparse optimization. *IEEE Transactions on Smart Grid*, 5(2):612–621, March 2014.

[47] S. Liu, S. Mashayekh, D. Kundur, T. Zourntos, and K. Butler-Purry. A framework for modeling cyber-physical switching attacks in smart grid. *IEEE Transactions on Emerging Topics in Computing*, 1(2):273–285, Dec 2013.

[48] Y. Z. Lun, A. D'Innocenzo, F. Smarra, I. Malavolta, and M. D. D. Benedetto. State of the art of cyber-physical systems security: An automatic control perspective. *Journal of Systems and Software*, 149:174 – 216, 2019.

[49] Y.L Lun, A. D'Innocenzo, I. Malavolta, M. Domenica, and D. Benedetto. Cyber-physical systems security: a systematic mapping study. *arxiv*, 2016. Available at https://arxiv.org/pdf/1605.09641.pdf.

[50] S. Martínez-Martínez, N. Messai, F. Hamelin, N. Manamanni, and T. Boukhobza. Graphic approach for the determination of the existence of sequences guaranteeing observability of switched linear systems. *Automatica*, 50(2):584–590, 2014.

[51] A. A. Maruf and S. Roy. Observability-blocking controllers for network synchronization processes. In *2019 American Control Conference (ACC)*, pages 2066–2071, July 2019.

[52] J. L. Melsa and D. L. Cohn. *Decision and Estimation Theory.* McGraw-Hill, 1978.

[53] F. Miao, Q. Zhu, M. Pajic, and G. J. Pappas. Coding schemes for securing cyber-physical systems against stealthy data injection attacks. *IEEE Transactions on Control of Network Systems*, 4(1):106–117, March 2017.

[54] Y. Mo and B. Sinopoli. Secure estimation in the presence of integrity attacks. *IEEE Transactions on Automatic Control*, 60(4):1145–1151, April 2015.

[55] Y. Mo and B. Sinopoli. On the performance degradation of cyber-physical systems under stealthy integrity attacks. *IEEE Transactions on Automatic Control*, 61(9):2618–2624, Sept 2016.

[56] Y. Mo, S. Weerakkody, and B. Sinopoli. Physical authentication of control systems. *IEEE Control Systems Magazine*, 35(1):93–109, 2015.

[57] U. Münz, M. Pfister, and P. Wolfrum. Sensor and actuator placement for linear systems based on H_2 and H_∞ optimization. *IEEE Trans. on Automatic Cont.*, 59(11):2984–2989, Nov 2014.

[58] H. Nishino and H. Ishii. Distributed detection of cyber attacks and faults for power systems. *IFAC Proceedings Volumes*, 47(3):1932 – 11937, 2014.

[59] B. Obama. Presidential policy directive 21: Critical infrastructure security and resilience. *Homeland Security Digital Library*, Feb 2013.

[60] A. Olshevsky. Minimal controllability problems. *IEEE Transactions on Control of Network Systems*, 1(3):249–258, Sep. 2014.

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

[62] F. Pasqualetti, S. Zampieri, and F. Bullo. Controllability metrics, limitations and algorithms for complex networks. *IEEE Transactions on Control of Network Systems*, 1(1):40–52, March 2014.

[63] S. Pequito, G. Ramos, S. Kar, A. P. Aguiar, and J. Ramos. The robust minimal controllability problem. *Automatica*, 82:261 – 268, 2017.

[64] N. Perraudin, J. Paratte, D. Shuman, L. Martin, V. Kalofolias, P. Vandergheynst, and D. K. Hammond. Gspbox: A toolbox for signal processing on graphs. *ArXiv e-prints*, Aug. 2014.

[65] H V. Poor. *An Introduction to Signal Detection and Estimation.* Springer-Verlag New York, 1994.

[66] V. Reppa, M. M. Polycarpou, and C. G. Panayiotou. Distributed sensor fault diagnosis for a network of interconnected cyberphysical systems. *IEEE Transactions on Control of Network Systems*, 2(1):11–23, March 2015.

[67] S. Roy, J. Abed Torres, and M. Xue. Sensor and actuator placement for zero-shaping in dynamical networks. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 1745–1750, Dec 2016.

[68] S. Roy, M. Xue, and S. Sundaram. Graph-theoretic analysis of estimators for stochastically-driven diffusive network processes. In *2018 Annual American Control Conference (ACC)*, pages 1796–1801, June 2018.

[69] H.S. Sánchez, D. Rotondo, T. Escobet, V. Puig, and J. Quevedo. Bibliographical review on cyber attacks from a control oriented perspective. *Annual Reviews in Control*, Sep. 2019.

[70] L. L. Scharf. *Statistical Signal Processing*. Reading, MA: Addition-Wesley, 1991.

[71] I. Shames, A. M. H. Teixeira, H. Sandberg, and K. H. Johansson. Distributed fault detection for interconnected second-order systems. *Automatica*, 47(12):2757 – 2764, 2011.

[72] D. Shi, R. J. Elliott, and T. Chen. On finite-state stochastic modeling and secure estimation of cyber-physical systems. *IEEE Transactions on Automatic Control*, 62(1):65–80, Jan 2017.

[73] J. Slay and M. Miller. Lessons learned from the Maroochy water breach. *Critical Infrastructure Protection*, 253:73–82, 2007.

[74] S. Sridhar and M. Govindarasu. Model-based attack detection and mitigation for automatic generation control. *IEEE Transactions on Smart Grid*, 5(2):580–591, March 2014.

[75] T. H. Summers and J. Lygeros. Optimal sensor and actuator placement in complex dynamical networks. *IFAC Proceedings Volumes*, 47(3):3784–3789, 2014. 19th IFAC World Congress.

[76] S. Sundaram and C. Hadjicostis. Distributed function calculation via linear iterative strategies in the presence of malicious agents. *IEEE Transactions on Automatic Control*, 56(7):1495–1508, 2011.

[77] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson. A secure control framework for resource-limited adversaries. *Automatica*, 51:135–148, Jan 2015.

[78] R. R. Tenney and N. R. Sandell. Detection with distributed sensors. *IEEE Transactions on Aerospace and Electronic Systems*, 17(4):501–510, July 1981.

[79] J. Abed Torres and S. Roy. Graph-theoretic analysis of network input-output processes: Zero structure and its implications on remote feedback control. *Automatica*, 61:73–79, 2015.

[80] H. L. Van Trees. *Detection, Estimation, and Modulation Theory. Part I:*. Wiley Press, 2004.

[81] V. Tzoumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie. Minimal actuator placement with bounds on control effort. *IEEE Transactions on Control of Network Systems*, 3(1):67–78, March 2016.

[82] H. J. Van Waarde, M. K. Camlibel, and H. L. Trentelman. A distance-based approach to strong target control of dynamical networks. *IEEE Transactions on Automatic Control*, 62(12):6266–6277, Dec 2017.

[83] J. M. Varah. A lower bound for the smallest singular value of a matrix. *Linear Algebra and its Applications*, 11(1):3–5, 1975.

[84] P.K. Varshney. *Distributed Detection and Data Fusion.* Springer-Verlag New York, 1997.

[85] A. Vosughi, C. Johnson, M. Xue, S. Roy, and S. Warnick. Target control and source estimation metrics for dynamical networks. *Automatica*, 100:412–416, 2019.

[86] M. Van De Wal and B. De Jager. A review of methods for input/output selection. *Automatica*, 37(4):487–510, 2001.

[87] J. M. Wozencraft and I. M. Jacobs. *Principles of Communication Engineering.* New York: Wiley, 1965.

[88] G. Wu and J. Sun. Optimal switching integrity attacks in cyber-physical systems. In *2017 32nd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pages 709–714, May 2017.

[89] M. Xue and S. Roy. Input-output properties of linearly-coupled dynamical systems: Interplay between local dynamics and network interactions. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 487–492, Dec 2017.

[90] Xing-Gang Yan and Christopher Edwards. Robust decentralized actuator fault detection and estimation for large-scale systems using a sliding mode observer. *Int. Journal of Control*, 81(4):591–606, 2008.

[91] S. Z. Yong, M. Zhu, and E. Frazzoli. Resilient state estimation against switching attacks on stochastic cyber-physical systems. In *2015 54th IEEE Conference on Decision and Control (CDC)*, Dec 2015.

[92] Y. Yuan, Q. Zhu, F. Sun, Q. Wang, and T. Basar. Resilient control of cyber-physical systems against denial-of-service attacks. In *2013 6th International Symposium on Resilient Control Systems (ISRCS)*, pages 54–59, Aug 2013.

[93] H. Zhang, R. Ayoub, and S. Sundaram. Sensor selection for kalman filtering of linear dynamical systems: Complexity, limitations and greedy algorithms. *Automatica*, 78:202–210, 2017.

[94] H. Zhang, P. Cheng, L. Shi, and J. Chen. Optimal denial-of-service attack scheduling with energy constraint. *IEEE Transactions on Automatic Control*, 60(11):3023–3028, Nov 2015.

[95] J. Zhao and L. Mili. Power system robust decentralized dynamic state estimation based on multiple hypothesis testing. *IEEE Transactions on Power Systems*, 33(4):4553–4562, July 2018.

[96] S. Zhao and F. Pasqualetti. Discrete-time dynamical networks with diagonal controllability gramian. *IFAC-PapersOnLine"*, 50(1):8297–8302, 2017. 20th IFAC World Congress.