Security Metrics of Networked Control Systems under Sensor Attacks (extended preprint)

Carlos Murguia\textsuperscript{1,a}, Iman Shames\textsuperscript{1,b}, Jusin Ruths\textsuperscript{2,c}, and Dragan Nešić\textsuperscript{1,d}

\textsuperscript{1}Department of Electrical and Electronic Engineering, University of Melbourne, Australia
\textsuperscript{2}Departments of Mechanical and Systems Engineering, University of Texas at Dallas, USA
\textsuperscript{a}carlos.murguia@unimelb.edu.au; \textsuperscript{b}iman.shames@unimelb.edu.au; \textsuperscript{c}jruths@utdallas.edu; \textsuperscript{d}dnesic@unimelb.edu.au

1 Abstract

As more attention is paid to security in the context of control systems and as attacks occur to real control systems throughout the world, it has become clear that some of the most nefarious attacks are those that evade detection. The term \textit{stealthy} has come to encompass a variety of techniques that attackers can employ to avoid being detected. In this manuscript, for a class of perturbed linear time-invariant systems, we propose two \textit{security metrics} to quantify the potential impact that \textit{stealthy attacks} could have on the system dynamics by tampering with sensor measurements. We provide analysis mathematical tools (in terms of linear matrix inequalities) to quantify these metrics for given system dynamics, control structure, system monitor, and set of sensors being attacked. Then, we provide synthesis tools (in terms of semidefinite programs) to redesign controllers and monitors such that the impact of stealthy attacks is minimized and the required attack-free system performance is guaranteed.

\textbf{Keywords:} Network Control Systems; Model-based fault/attack monitors, Security Metrics, Secure Control Systems, Attacker Capabilities.

2 Introduction

There has recently been significant interest and work in the broad area of security of Networked Control Systems (NCSs), see, e.g., \cite{1-13}. This topic investigates the properties of conventional control systems in the presence of adversarial disturbances. Control theory has shown great ability to robustly deal with disturbances and uncertainties. However, adversarial attacks raise all-new issues due to the aggressive and strategic nature of the disturbances that attackers might inject into the system.

This paper focuses on quantifying and minimizing attacker capabilities in NCSs. A majority of the work on attack detection leverages the established literature of fault detection \cite{3-15}. A fault detection approach uses an \textit{estimator} to forecast the evolution of the system dynamics. When the residual (the difference between a measurement and its estimate), or some function of the residual, is larger than a predetermined threshold, an alarm is raised. Arguably the most insidious attacks are those that occur without our knowledge. Fault detectors impose limits on the attacker, if the attacker aims to avoid being identified. Beyond retooling these existing methods for the new attack detection context, a fundamental question is: given a chosen fault detection approach, how does this method constrain the influence of an attacker? More specifically, what is an attacker able to accomplish when a system employs certain fault detection procedure?

Different methodologies exist for evaluating the impact of attacks. Most of the existing work uses some measure of state deviation. A number of groups have studied the system response when the attacks are constrained by the detector, i.e., they investigate the system trajectories that can be induced due to \textit{stealthy attacks} – attacks such that the detector threshold is never crossed \cite{11-20}. In this manuscript, for given system dynamics, control structure, and fault detector, we provide mathematical tools for \textit{quantifying and minimizing} the potential impact of sensor stealthy attacks on the system dynamics. We consider the set of states that stealthy attacks can induce in the system (the attacker’s reachable set) and use the "size" of this set as a \textit{security metric} for the NCS. Because it is not mathematically tractable to compute this set exactly, we provide synthesis tools – in terms of Linear Matrix Inequalities (LMIs) – for computing \textit{ellipsoidal outer approximations} of the attacker’s reachable set. The obtained approximations quantify the attacker’s potential impact when it is constrained to stay hidden from the detector. We use the size (in terms of volume) of these ellipsoidal approximations to approximate the proposed security metric. As a second security metric, we propose the minimum distance from the attacker’s reachable set to a possible set of \textit{critical states} – states that, if reached, compromise the integrity or safe operation of the system. We approximate this distance by the minimum distance between the ellipsoidal approximations and the critical states. This distance gives us intuition on how far the actual attacker’s reachable set is from the critical states. Once we have provided a complete set of analysis tools to approximate the aforementioned security metrics, we use these tools to derive...
synthesis tools (in terms of semidefinite programs) to redesign controllers and fault detectors such that the impact of stealthy attacks is minimized and the required attack-free system performance is guaranteed.

There are a few results in this direction already; chiefly the work in [21] (and the preliminary paper [8]), where the authors provide a recursive algorithm to compute ellipsoidal approximations of attacker’s reachable sets for LTI systems subjected to Gaussian noise. The authors in [21] give analysis-only results for a very particular structure of controllers and fault-detectors. They consider Kalman-filter based fault detectors and use the state of the filter to construct output feedback controllers. Although this results in compact designs of controllers and fault detectors, the flexibility of having dedicated controllers and detectors (mainly for synthesis of secure control systems) is limited. We remark that, in the stochastic setting considered in [21], the detector threshold is always crossed even when there are no attacks. This is due to the infinite support of the Gaussian noise they consider. Thus, they do not consider stealthy attacks in the sense described above. Instead, they consider attacks that increase the alarm rate of the detector by a small amount only. Then, they approximate the attacker’s reachable set corresponding to this small increase. In contrast to the work in [21], in this manuscript, we provide a set of mathematical tools in terms of semidefinite programs to approximate reachable sets induced by stealthy attacks for LTI systems driven by peak bounded deterministic perturbations. We provide both analysis and synthesis results for dedicated general dynamic output feedback controllers and observer-based fault detectors. We propose two security metrics to assess the vulnerability of systems to attacks; and optimize these metrics (enhancing thus the system resilience to attacks) by synthesizing optimal controllers and detectors. The synthesis part considers the attack-free performance of the closed-loop dynamics, i.e., we optimize the security metrics subject to certain prescribed attack-free system performance. In our preliminary work [17], we also approximate reachable sets of false-data-injection attacks but we consider the same stochastic framework as the one proposed in [21], i.e., Gaussian noise, joint Kalman-filter based fault detectors and controllers, and attacks increasing the alarm rate of the detector. Thus, the problems considered in this manuscript (and the obtained results) and the ones addressed in [17] are fundamentally different; and the set of results (and the tools used to obtained them) are different too. Moreover, in [17], we consider attacks to all the sensors. Although the latter case provides a worse-case scenario, we lose the capability of quantifying the sensitivity of the system dynamics to attacks on specific sensors. As in [21], the results in [17] mainly focus on analysis (although they hint how to address synthesis for joint Kalman-filter based detectors and controllers).

The remainder of the paper is organized as follows. In Section 3, we present some preliminaries results needed for the subsequent sections. We provide tools for computing outer time-varying bounds on the trajectories of a class of perturbed nonlinear discrete-time systems. Then, we use these tools to obtain outer ellipsoidal approximations of reachable sets of LTI systems driven by multiple peak bounded perturbations. The system dynamics, monitor, and controller descriptions are given in Section 4. Our proposed security metrics and analysis tools, together with some numerical results, are given in Section 5; and the corresponding synthesis results are given in Section 6. Finally, conclusions and recommendations are stated in Section 7.

3 Preliminaries

In this section, we present some preliminary results needed for the subsequent sections. First, in Lemma 1 we present a tool for computing time-varying outer bounds on the trajectories of a class of perturbed nonlinear discrete-time systems. Next, in Proposition 1, we use this lemma to compute outer ellipsoidal approximations of reachable sets of LTI systems driven by multiple peak bounded perturbations.

Consider the nonlinear system driven by $N$ perturbations:

$$\xi_{k+1} = f(k, \xi_k, \omega_k^1, \ldots, \omega_k^N),$$

with time-index $k \in \mathbb{N}$, state $\xi_k \in \mathbb{R}^{n\xi}$, perturbation $\omega_k^i \in \mathbb{R}^{p_i}$ satisfying $(\omega_k^i)^T W_k^i \omega_k^i \leq 1$ for some time-varying positive definite matrix $W_k^i \in \mathbb{R}^{p_i \times p_i}$, $i = 1, \ldots, N$, $N \in \mathbb{N}$, and function $f : \mathbb{N} \times \mathbb{R}^{n\xi} \times \mathbb{R}^{p_1} \times \cdots \times \mathbb{R}^{p_N} \rightarrow \mathbb{R}^{n\xi}$.

**Lemma 1** For a given $a \in (0, 1)$, if there exist functions $a_k^i : \mathbb{N} \rightarrow (0, 1)$, $i = 1, \ldots, N$, and $V : \mathbb{R}^{n\xi} \rightarrow \mathbb{R}_{\geq 0}$ satisfying $\sum_{i=1}^{N} a_k^i \geq a$ and the inequality:

$$V(f(k, \xi_k, \omega_k^1, \ldots, \omega_k^N)) - aV(\xi_k) - \sum_{i=1}^{N} (1 - a_k^i)(\omega_k^i)^T W_k^i \omega_k^i \leq 0, \forall k \in \mathbb{N};$$

then, $V(\xi_k) \leq \alpha_k$, where $\alpha_k := a^{k-1}V(\xi_1) + \frac{(N - a)(1 - a^{k-1})}{1 - a}$, and $\lim_{k \rightarrow \infty} V(\xi_k) \leq \frac{N - a}{1 - a}$. 


Proof: By assumption, for \( i = 1, \ldots, N \), \((\omega_k)^T W_k^i \omega_k \leq 1\), then, from (2), we have
\[
V(\xi_{k+1}) \leq a V(\xi_k) + \sum_{i=1}^{N} (1 - a_k) \left( \omega_k^T W_k^i \omega_k \right) \leq a V(\xi_k) + (N - a),
\]
(3)
because \( \sum_{i=1}^{N} a_k^i \geq a \). It follows that
\[
V(\xi_k) \leq a V(\xi_{k-1}) + (N - a),
\]
(4)
and
\[
V(\xi_{k-1}) \leq a V(\xi_{k-2}) + (N - a).
\]
(5)
Using (5) to upper bound (4) and continuing the recursion yields
\[
V(\xi_k) \leq a^{k-1} V(\xi_1) + \frac{(N - a)(1 - a^{k-1})}{1 - a}.
\]
Therefore, \( \lim_{k \to \infty} V(\xi_k) \leq (N - a)/(1 - a) \) because \( a \in (0, 1) \).

Next, we present a tool to identify outer ellipsoidal approximations of reachable sets of LTI systems driven by multiple peak bounded perturbations.

Consider the perturbed LTI system
\[
\xi_{k+1} = A \xi_k + \sum_{i=1}^{N} B^i \omega_k^i,
\]
(6)
with \( k \in \mathbb{N} \), state \( \xi_k \in \mathbb{R}^{n_c} \), initial condition \( \xi_1 \in \mathbb{R}^{n_c} \), perturbation \( \omega_k^i \in \mathbb{R}^{n_p} \) satisfying \((\omega_k)^T W_i \omega_k^i \leq 1\) for some positive definite matrix \( W_i \in \mathbb{R}^{n_p \times n_p} \), \( i = 1, \ldots, N \), \( N \in \mathbb{N} \), and matrices \( A \in \mathbb{R}^{n_c \times n_c} \) and \( B^i \in \mathbb{R}^{n_c \times n_p} \). Denote by \( \psi_k(k, \xi_1, \omega_1(\cdot), \ldots, \omega_N(\cdot)) := A^{k-1} \xi_1 + \sum_{i=0}^{k-2} A^i B^i \omega_{k-1-i} \) the solution of (6) at time instant \( k > 1 \) given the initial condition \( \xi_1 \) and the infinite disturbance sequence \( \omega(\cdot) := \{\omega_1, \omega_2, \ldots\} \).

Definition 1 The reachable set \( \mathcal{R}_k^\xi \) at time instant \( k > 1 \) from the initial condition \( \xi_1 \) is the set of states \( \psi_k(k, \xi_1, \omega_1(\cdot), \ldots, \omega_N(\cdot)) \) reachable in \( k \) steps by system (6) through all possible perturbations satisfying \((\omega_k)^T W_i \omega_k^i \leq 1\), i.e., \( \mathcal{R}_k^\xi := \{ \xi \in \mathbb{R}^{n_c} \mid \psi_k(k, \xi_1, \omega_1(\cdot), \ldots, \omega_N(\cdot)), \xi_1 \in \mathbb{R}^{n_c}, \text{ and } (\omega_k)^T W_i \omega_k^i \leq 1\} \).

Proposition 1 Consider the perturbed LTI system (6) and the reachable set \( \mathcal{R}_k^\xi \) introduced in Definition 1. For a given \( a \in (0, 1) \), if there exist constants \( a_1 = \tilde{a}_1, \ldots, a_N = \tilde{a}_N \) and matrix \( P = \tilde{P} \in \mathbb{R}^{n_c \times n_c} \) satisfying:
\[
\begin{bmatrix}
a P & A^T P & 0 \\
ap A & P & \tilde{P} B \\
0 & B^T \tilde{P} & W_a
\end{bmatrix} \geq 0;
\]
(7)
with \( B := (B^1, \ldots, B^N) \in \mathbb{R}^{n_c \times \sum_{i=1}^{N} p_i} \) and \( W_a := \text{diag}([1 - a_1])W_1, \ldots, (1 - a_N)W_N] \in \mathbb{R}^{\sum_{i=1}^{N} p_i \times \sum_{i=1}^{N} p_i} \); then, \( \mathcal{R}_k^\xi \subseteq \tilde{E}_k := \{ \xi \in \mathbb{R}^{n_c} \mid \tilde{P} \xi \leq \tilde{a}_k \} \), where \( \tilde{a}_k := a^{k-1} \xi_1^T \tilde{P} \xi_1 + ((N - a)/(1 - a^{k-1}))/1 - a \).

Proof: For a positive definite matrix \( P \in \mathbb{R}^{n_c \times n_c} \), let \( V_k = \tilde{P} \xi_k \) in Lemma 3. Substituting this \( V_k \), the dynamics \( \xi_{k+1} = A \xi_k + B \omega_k \), where \( \omega_k := ((\omega_k^1)^T, \ldots, (\omega_k^N)^T)^T \), and the inequality \( a_1 + \cdots + a_N \geq a \) in (2) yields
\[
\nu_k^T \begin{bmatrix}
P - A^T PA & -A^T PB & 0 \\
0 & W_a & 0 \\
PA & \tilde{PB} & P
\end{bmatrix} 
\nu_k \geq 0,
\]
with \( \nu_k := (\xi_1^T, \omega_k^T)^T \). This inequality is satisfied if and only if \( Q \) is positive semidefinite. This \( Q \) can be written as the Schur complement of a higher dimensional matrix \( Q' \); it follows that \( Q \geq 0 \iff Q' \geq 0 \) where
\[
Q' := \begin{bmatrix}
P & 0 & A^T P \\
0 & W_a & B^T \tilde{P} \\
PA & \tilde{PB} & P
\end{bmatrix}.
\]
Consider the congruence transformation \( Q' \to T^T Q'T \) with
\[
T := \begin{bmatrix}
I & 0 & 0 \\
0 & I & 0 \\
0 & 0 & I
\end{bmatrix}.
\]
Hence, \( Q \geq 0 \iff Q' \geq 0 \iff T^T Q'T \geq 0 \), see [22] for details. Inequality \( T^T Q'T \geq 0 \) equals the last inequality in (7). Then, by Lemma 3 we have \( \xi_1^T \tilde{P} \xi_1 \leq a^{k-1} \xi_1^T \tilde{P} \xi_1 + ((N - a)/(1 - a^{k-1}))/1 - a = \tilde{a}_k \) for any \( a_1 = \tilde{a}_1 \),
$i = 1, \ldots, m$, and $\mathcal{P} = \mathcal{P}$ satisfying $\mathcal{P}$. It follows that the trajectories $\xi_k$ generated by $\xi_{k+1} = A\xi_k + \sum_{i=1}^{N} B^i \omega_k^i$, the initial condition $\xi_1$, and the perturbation $\omega_k$, are always contained in the time-varying ellipsoid $\bar{E}_k^\xi$, i.e., $\bar{E}_k^\xi \subseteq \bar{E}_k^\xi$.

Remark 1 Note that the contribution of the initial condition $\xi_1$ and the constant $a \in (0, 1)$ to the sequence $\alpha_k^\xi$ vanishes exponentially. We have that $\lim_{k \to \infty} \alpha_k^\xi = (N - a)/(1 - a)$; therefore $\lim_{k \to \infty} \bar{E}_k^{\xi_k} = \{ \xi \in \mathbb{R}^n | \xi^T \mathcal{P} \xi \leq (N - a)/(1 - a) \} =: \bar{E}_\infty^{\xi_k}$.

That is, $\bar{E}_\infty^{\xi_k}$ provides an ultimate bound [23] for the time-varying ellipsoidal approximation $\bar{E}_k^{\xi_k}$.

Proposition 1 provides a tool for computing time-varying ellipsoidal outer approximations $\bar{E}_k^\xi$ of $\mathcal{R}_k^\xi$. Note that $\bar{E}_k^\xi$ could be an arbitrarily conservative approximation of $\mathcal{R}_k^\xi$ as long as $\mathcal{R}_k^\xi \subseteq \bar{E}_k^\xi$. Then, to make $\bar{E}_k^\xi$ less conservative, we aim at obtaining ellipsoids with minimal volume, i.e., the tightest possible ellipsoid bounding $\mathcal{R}_k^\xi$ among all the ellipsoids generated by Proposition 1. To find such an ellipsoid, we look to minimize $(\det[\mathcal{P}])^{-1/2}$ subject to (7) because $(\det[\mathcal{P}])^{-1/2}$ is proportional to the volume of $\xi^T \mathcal{P} \xi = \alpha_k$ for any $\alpha_k > 0$ [22]. We minimize $\log \det[\mathcal{P}^{-1}]$ instead as it shares the same minimizer with $(\det[\mathcal{P}])^{-1/2}$ and because for positive definite $\mathcal{P}$ this objective is convex [22]. This is stated in the following corollary of Proposition 1.

Corollary 1 Consider the perturbed LTI system (6) and the reachable set $\mathcal{R}_k^\xi$ introduced in Definition 1. For a given $a \in (0, 1)$, if there exist constants $a_1 = a_1, \ldots, a_N = a_N$ and matrix $\mathcal{P} = \mathcal{P}^*$ solution of the convex optimization:

$$
\begin{align*}
\min_{\mathcal{P}, a_1, \ldots, a_N} & -\log \det[\mathcal{P}], \\
\text{s.t.} & (7)
\end{align*}
$$

then, $\mathcal{R}_k^\xi \subseteq \bar{E}_k^\xi := \{ \xi \in \mathbb{R}^n | \xi^T \mathcal{P} \xi \leq \alpha_k^\xi \}$, where $\alpha_k^\xi := a_k^{-1} \xi_1^T \mathcal{P}^* \xi_1 + ((N - a)/(1 - a)) / (1 - a)$. Moreover, for any $a_1 = \bar{a}_1 \neq a_1^*$ and $\mathcal{P} = \mathcal{P} \neq \mathcal{P}^*$ satisfying the constraints in (7) and corresponding ellipsoidal approximation $\bar{E}_k$, the volume of $\bar{E}_k$ is strictly less than the volume of $\bar{E}_k$, i.e., $\bar{E}_k$ has the minimum volume among all the outer ellipsoidal approximations $\bar{E}_k$ generated by Proposition 1.

Proof: The solution space of the objective function is convex because the constraints are linear [24]. Moreover, the function $\log \det[\mathcal{P}^{-1}]$ is convex for any positive definite matrix $\mathcal{P}$ [22]. Hence, Corollary 1 follows from Proposition 1, convexity of the solution space, and convexity of the objective function.

Remark 2 Corollary 1 implies that the solution to the optimization problem (6) may exist for some, but not all, values of the parameter $a \in (0, 1)$. Out of these feasible values, we are interested in selecting the one that leads to the ellipsoid with minimum volume. We employ a straightforward grid search to find this value of $a \in (0, 1)$.

4 System & Monitor Description
4.1 System

Consider the LTI perturbed system

$$
\begin{align*}
x^p(t_{k+1}) &= A^p x^p(t_k) + B^p u(t_k) + Ev(t_k), \\
y(t_k) &= C^p x^p(t_k) + F y(t_k),
\end{align*}
$$

Figure 1: Cyber-physical system under sensor attacks.
with sampling time-instants $t_k, k \in \mathbb{N}$, state $x^p \in \mathbb{R}^n$, output $y \in \mathbb{R}^m$, control input $u \in \mathbb{R}^l$, matrices $A^p, B^p, C^p, E$, and $F$ of appropriate dimensions, and unknown system and sensor perturbations $v \in \mathbb{R}^q$ and $\eta \in \mathbb{R}^m$, respectively. The perturbations are assumed to be peak bounded, i.e., $v_k^T v_k \leq \bar{v}$ and $\eta_k^T \eta_k \leq \bar{\eta}$ for some known $\bar{v}, \bar{\eta} \in \mathbb{R}_{\geq 0}$ and all $k \in \mathbb{N}$. The pair $(A^p, B^p)$ is stabilizable and $(A^p, C^p)$ is detectable. At the time-instants $t_k, k \in \mathbb{N}$, the output of the process $y(t_k)$ is sampled and transmitted over a communication network. The received output $\hat{y}(t_k)$ is used to compute control actions $u(t_k)$ which are sent back to the actuators. The complete control-loop is assumed to be performed instantaneously, i.e., sampling, transmission, and arrival time-instants are equal. In this manuscript, we focus on false data injection attacks on sensor measurements. That is, in between transmission and reception of sensor data, an attacker may inject data to the signals coming from sensors to the controller, see Fig. 1. The opponent compromises up to $s$ sensors, $s \in \{1, \ldots, m\}$ of the system. Denote the attacker’s sensor selection matrix $\Gamma \in \mathbb{R}^{m \times s}$, $\Gamma \subseteq \{\gamma_1, \ldots, \gamma_m\}$ where $\gamma_i \in \mathbb{R}^{m \times 1}$ denotes the $i$-th vector of the canonical basis of $\mathbb{R}^m$. After each transmission and reception, the networked output $\hat{y}$ takes the form:

$$\hat{y}(t_k) := y(t_k) + \Gamma \delta(t_k),$$

where $\delta(t_k) \in \mathbb{R}^s$ denotes additive sensor attacks/faults. Denote $x_k := x(t_k)$, $u_k := u(t_k)$, $v_k := v(t_k)$, $\bar{y}_k := \hat{y}(t_k)$, $\eta_k := \eta(t_k)$, and $\delta_k := \delta(t_k)$. Using this new notation, the attacked system is written in the following compact form:

$$\begin{align*}
\dot{x}_{k+1} & = A^p x_k + B^p u_k + E v_k, \\
\bar{y}_k & = C^p x_k + F \eta_k + \Gamma \delta_k.
\end{align*}$$

4.2 Filter and Residual

In this manuscript, we aim at characterizing the effect that false data injection attacks can induce in the system without being detected by standard fault-detectors. The main idea behind fault detection is the use of an estimator to forecast the evolution of the system state. If the difference between what it is measured and the output estimation is larger than expected, there may be a fault in or an attack on the system. Here, to estimate the state of the process, we use the filter:

$$\dot{x}_{k+1} = A^p \hat{x}_k + B^p u_k + L(\bar{y}_k - C^p \hat{x}_k),$$

with estimated state $\hat{x} \in \mathbb{R}^n$, $\hat{x}_1 = (C^p)^+ y_1$, where $(C^p)^+$ denotes the Moore-Penrose inverse of $C^p$, and filter gain matrix $L \in \mathbb{R}^{n \times m}$. Define the estimation error $e_k := x_k^p - \hat{x}_k$. Given the system dynamics (11) and the filter (12), the estimation error is governed by the following difference equation

$$e_{k+1} = (A^p - LC^p) e_k - L \delta_k - LF \eta_k + Ev_k.$$  

4.3 Distance Measure, Anomaly Detection, and System Monitor

The input to any detection procedure is a distance measure $z_k \in \mathbb{R}$, $k \in \mathbb{N}$, i.e., a measure of how deviated the estimator is from the attack-free system dynamics [25]. Here, we use a quadratic form of the residual as distance measure. Consider the residual sequence $r_k$ and some positive definite matrix $\Pi \in \mathbb{R}^{m \times m}$. Define the distance measure $z_k := r_k^T \Pi r_k$ and consider the following monitor.

**System Monitor:**

$$z_k = r_k^T \Pi r_k > 1, \quad \bar{k} = k.$$  

**Design parameter:** positive semidefinite matrix $\Pi \in \mathbb{R}^{m \times m}$.

**Output:** alarm time(s) $\bar{k}$.

Thus, the monitor is designed so that alarms are triggered if $z_k$ exceeds one. The matrix $\Pi$ must be selected such that, after sufficiently large number of time-steps (enough to allow transients to settle down), $z_k \leq 1$ in the attack-free case. That is, after transients due to initial conditions have decreased to a desired level, the ellipsoid

$$r_k^T \Pi r_k = 1$$

must contain all the possible trajectories that the perturbations $v_k$ and $\eta_k$ can induce in the residual given Eq. (15) and the inequalities $v_k^T v_k \leq \bar{v}$ and $\eta_k^T \eta_k \leq \bar{\eta}$. Note that the tighter the ellipsoidal bound, the less opportunity the attacker has to manipulate the system without being detected. Here, we use Corollary 1 to design
an optimal matrix II (in terms of tightness of the ellipsoidal bound). In particular, using Corollary 1, we obtain an outer time-varying ellipsoidal approximation of the reachable set of the estimation error \( r_k^* \) driven by \( r_k \) and \( \eta_k \) and \( \delta_k \) in the attack-free case (\( \delta_k = 0 \)). Once we have this ellipsoid, using the S-procedure [22], we project it onto the residual hyperplane to get the ellipsoid \( r_k^* \Pi r_k = 1 \) of the monitor. For transparency, these results are presented in the appendix. We need, however, the following assumption for the subsequent sections.

**Assumption 1** In the attack-free case (\( \delta_k = 0 \)), there exists some \( k^* \in \mathbb{N} \) such that the matrix II of the monitor satisfies \( r_k^* \Pi r_k \leq 1 \) for all \( k \geq k^* \) and \( r_k \) solution of (15). In the appendix, we give tools for obtaining a matrix II satisfying Assumption 1 for a desired \( k^* \) as a function of the initial estimation error \( e_1 \) and a desired tightness level of the ellipsoidal bound.

### 4.4 Dynamic Output Feedback Controller

We consider general dynamic output feedback controllers of the form:

\[
\begin{align*}
4.4 \text{ Dynamic Output Feedback Controller} \\
\text{where} \\
\Pi \\
\text{We consider general dynamic output feedback controllers of the form:} \\
4.4 \text{ Dynamic Output Feedback Controller} \\
\begin{equation}
\begin{aligned}
x_{k+1}^c &= A^c x_k^c + B^c y_k, \\
u_k &= C^c x_k^c + D^c y_k,
\end{aligned}
\end{equation}
\]

with controller state \( x^c \in \mathbb{R}^n \), networked output \( y \), control input \( u \), and controller matrices \( (A^c, B^c, C^c, D^c) \) of appropriate dimensions. For simplicity, we only consider controllers with the same order as the plant. This is particularly important in the synthesis section of the manuscript (however, results for general order controllers can be derived following the same approach). The closed-loop system (11), (12), (17) can be written in terms of the estimation error \( e_k = x_k^c - \hat{x}_k \) as follows:

\[
\begin{align*}
x_{k+1}^c &= (A^p + B^p D^p C^p)x_k^c + B^p D^p C^p \eta_k + Ev_k + B^p D^p \Gamma \delta_k, \\
e_{k+1} &= (A^p - LC^p)e_k - LF \eta_k + Ev_k - L \delta_k.
\end{align*}
\]

Note that, due to linearity, the estimation error dynamics and the residual (15) do not depend on the system and filter states.

### 5 Analysis Tools: Stealthy Attacker’s Reachable Sets

In this section, we provide tools for quantifying (for given \( (L, A^c, B^c, C^c, D^c) \)) and minimizing (by redesigning \( (L, A^c, B^c, C^c, D^c) \)) the impact of the attack \( \delta_k \) on the state of the system when the monitor (16) is used for attack detection. We are interested in attacks that keep the monitor from raising alarms. This class of attacks is what we refer to as stealthy attacks. Here, we characterize ellipsoidal bounds on the set of states that stealthy attacks can induce in the system. In particular, we provide tools based on Linear Matrix Inequalities (LMIs) for computing ellipsoidal bounds on the reachable set of the attack sequence given the system dynamics, the control strategy, the system monitor, and the set of sensors being attacked.

**Assumption 2** We assume that the attack to system (11), (12), (17) starts at \( k = k^* \) (the monitor convergence time), i.e., the system has been operating without attacks for sufficiently long time so that the residual trajectories \( r_k \), for \( k \geq k^* \), are contained in the monitor ellipsoid \( \{ r \in \mathbb{R}^n | r^T \Pi r \leq 1 \} \) before an attack occurs.

The attacker can compromise up to \( s \) sensors, \( s \in \{1, \ldots, m\} \), of the system. Consider the monitor (16) and write \( z_k \) in terms of the estimation error \( e_k \) and \( \delta_k \):

\[
z_k = r_k^T \Pi r_k = \left\| \Pi^\frac{1}{2} (C^p e_k + F \eta_k + \Gamma \delta_k) \right\|^2,
\]

where \( \Pi^\frac{1}{2} \) is the symmetric square root matrix of II and \( \| \cdot \| \) denotes Euclidian norm. The set of feasible attack sequences that the attacker can launch while satisfying \( z_k \leq 1 \) (i.e., without raising alarms by the monitor) can be written as the following constrained control problem on \( \delta_k \):

\[
\delta_k \in \mathbb{R}^m \quad \text{Eq. (18) and} \quad \left\| \Pi^\frac{1}{2} (C^p e_k + F \eta_k + \Gamma \delta_k) \right\|^2 \leq 1, \ \forall \ k \geq k^*.
\]

Define the extended state \( \zeta_k := (x_k^T, x_k^T, e_k^T)^T \) and denote by \( \psi_k^\delta(k, \zeta_k, \eta(\cdot), v(\cdot), \delta(\cdot)) \) the solution of (18) at time instant \( k \geq k^* \). Given the extended state at the starting attack instant \( \zeta_k^* \) and the infinite disturbance and attack sequences \( \eta(\cdot) := \{ \eta_1, \eta_2, \ldots \} \), \( v(\cdot) := \{ v_1, v_2, \ldots \} \), and \( \delta(\cdot) := \{ \delta_1, \delta_2, \ldots \} \). Denote by \( \psi_k^\delta(k, \zeta_k^*, \eta(\cdot), v(\cdot), \delta(\cdot)) \) the partition of \( \psi_k^\delta(k, \zeta_k, \eta(\cdot), v(\cdot), \delta(\cdot)) \) corresponding to the plant trajectories, i.e., the solution \( x_k^* \) of (18). We are interested in the state trajectories that the attacker can induce in the system
Define the matrices computed efficiently using LMIs, we use the volume of the set of sensors being compromised (the attacker’s sensor selection matrix).

In (15), the residual is given by

\[ x_k^p = \psi^p_{k}(k, \zeta_k, \eta(\cdot), v(\cdot), \delta(\cdot)), \quad x_k^c, e_k^c \in \mathbb{R}^n, \]

\[ \delta_k, \zeta_k \text{ satisfy } \{ x_k^p \in \mathbb{R}^n | x_k^p \leq \bar{v}, \eta_k^p \leq \bar{\eta}, \text{ and } \bar{v}, \bar{\eta} \in \mathbb{R}_{>0}, \forall k \geq k^* \}. \]  

In this manuscript, we propose to use the volume of the set \( R_{\Gamma,k}^R \) as a security metric. However, in general, it is not tractable to compute \( R_{\Gamma,k}^R \) exactly. Instead, for some positive definite \( P_{\Gamma}^R \in \mathbb{R}^{n \times n} \) and nonnegative function \( \alpha_k^R \), we look for outer ellipsoidal approximations of the form \( E_{\Gamma,k}^R = \{ x_k \in \mathbb{R}^n | (x_k)^T P_{\Gamma}^R x_k \leq \alpha_k^R \} \) such that \( R_{\Gamma,k} \subseteq E_{\Gamma,k}^R \). Because, for LTI systems \( E_{\Gamma,k}^R \) is a good approximation of \( R_{\Gamma,k}^R [22] \), and because \( E_{\Gamma,k}^R \) can be computed efficiently using LMIs, we use the volume of \( E_{\Gamma,k}^R \) as an approximation of the proposed security metric.

This approximation allows us to quantify the potential "damage" that sensor attacks can induce to the system in terms of the set of sensors being compromised (the attacker’s sensor selection matrix \( \Gamma \)). In Figure 2, we depict a schematic representation of the proposed ideas.

5.1 Analysis Tools

In [15], the residual is given by \( r_k = C^p e_k + \Gamma \delta_k + F \eta_k \). Because \( \Gamma \) has full column rank by construction, we can write the attack sequence as \( \delta_k = \Gamma^+ (r_k - C^p e_k - F \eta_k) \), where \( \Gamma^+ \) denotes the Moore-Penrose inverse of \( \Gamma \), and the closed-loop dynamics (18) as

\[ x_{k+1}^c = A^c x_k^c + B^c C^p x_k^p + B^c D^c F \eta_k + E v_k + B^c D^c \Gamma^+ r_k, \]

\[ e_{k+1} = (A^p - L(I_m - \Gamma^+) C^p) e_k - L(I_m - \Gamma^+) F \eta_k + E v_k - L \Gamma^+ r_k. \]

Define the matrices

\[ A := \begin{bmatrix} A^p + B^p D^c C^p & B^p C^p & -B^p D^c \Gamma^+ C^p \\ B^c C^p & A^c & -B^c \Gamma^+ C^p \\ 0 & 0 & A^p - L(I_m - \Gamma^+) C^p \end{bmatrix}, \]

\[ B^1 := \begin{bmatrix} B^p D^c (I_m - \Gamma^+) F \\ B^c (I_m - \Gamma^+) F \\ -L(I_m - \Gamma^+) \end{bmatrix}, B^2 := \begin{bmatrix} E \\ 0 \end{bmatrix}, B^3 := \begin{bmatrix} B^p D^c \Gamma^+ \\ B^c \Gamma^+ \\ -L \Gamma^+ \end{bmatrix}, B := [B^1 \ B^2 \ B^3]. \]

Then, the closed-loop dynamics can be written in terms of the extended state \( \zeta_k = ((x_k^p)^T, (x_k^c)^T, e_k^c)^T \) and the residual sequence:

\[ \zeta_{k+1} = A \zeta_k + B^1 \eta_k + B^2 v_k + B^3 r_k, \]

\[ k \geq k^*. \]

Denote by \( \psi^p_{k}(k, \zeta_k, \eta(\cdot), v(\cdot), r(\cdot)) \) the solution of (26) at time instant \( k \geq k^* \) given the extended state at the starting attack instant \( \zeta_{k^*} \) and the infinite residual and disturbance sequences \( r(\cdot) := \{ r_1, r_2, \ldots \} \), \( \eta(\cdot) \), and \( v(\cdot) \).

Define the reachable set

\[ R_{\Gamma,k}^\zeta := \left\{ \zeta \in \mathbb{R}^{3n} | \zeta = \psi^p_{k}(k, \zeta_k, \eta(\cdot), v(\cdot), r(\cdot)), \zeta_k = ((x_k^p)^T, (x_k^c)^T, e_k^c)^T \in \mathbb{R}^{3n}, \right. \]

\[ \left. r_k^T \Pi r_k \leq 1, v_k^T v_k \leq \bar{v}, \eta_k^T \eta_k \leq \bar{\eta}, \text{ and } \bar{v}, \bar{\eta} \in \mathbb{R}_{>0}, \forall k \geq k^*. \right\}. \]

The set \( R_{\Gamma,k}^\zeta \) is the reachable set of an LTI system driven by peak-bounded perturbations. Therefore, we can use Corollary 1 to obtain approximations of the form \( E_{\Gamma,k}^\zeta = \{ \zeta \in \mathbb{R}^{3n} | \zeta^T P_{\Gamma}^\zeta \zeta \leq \alpha_k^\zeta \} \) such that \( R_{\Gamma,k}^\zeta \subseteq E_{\Gamma,k}^\zeta \).
Remark 3 We are ultimately interested in the stealthy reachable set of the plant states $\mathcal{R}_{1,k}^c$ introduced in \([21]\).
Note that $\mathcal{R}_{1,k}^c$ is the projection of $\mathcal{R}_{1,k}^C$ onto the $x^p$-hyperplane. Hence, if $\mathcal{E}_{1,k}^C \subseteq \mathcal{R}_{1,k}^C$, then $\mathcal{R}_{1,k}^c \subseteq \mathcal{E}_{1,k}^c$ where $\mathcal{E}_{1,k}^c$ denotes the projection of $\mathcal{E}_{1,k}^C$ onto the $x^p$-hyperplane. Therefore, to obtain the ellipsoid $\mathcal{E}_{1,k}^c$, we can first obtain the ellipsoid $\mathcal{E}_{1,k}^C$ containing $\mathcal{R}_{1,k}^C$ and then take $\mathcal{E}_{1,k}^c$. 

Theorem 1 Consider the closed-loop dynamics \([22]-[24]\) with system matrices $(A^p, B^p, C^p)$, observer gain $L$, controller matrices $(A^c, B^c, C^c, D^c)$, monitor matrix $\Gamma$, perturbations of $\bar{v}$, $\bar{\eta} \in \mathbb{R}_{>0}$, and attack selection matrix $B$. For a given $a \in (0,1)$, if there exist constants $a_1 = a_1^*, \ldots, a_N = a_N^*$ and matrix $P = P^*$ solution of \([8]\) with $A = A$, $N = 3$, $B^1 = B^3$, $B^2 = B^2$, $A = (A, B)$ as defined in \([25]\), $W_1 = (1/\eta)I_m$, $W_2 = (1/\bar{\eta})I_m$, $W_3 = \Pi$, $p_1 = m$, $p_2 = n$, and $p_3 = m$; then, for all $k \geq k^*$, $\mathcal{R}_{1,k}^c \subseteq \mathcal{E}_{1,k}^c := \{ \zeta \in \mathbb{R}^{3n|C^T P^C_1 \zeta \leq \alpha_k^c}\}$, with $P^C := P^*$ and $\alpha_k^c := a^{k-1}C^T_k P^C k^* \alpha_k^c + ((3-a)(1-a^{k-1}))/1-a$, and the ellipsoid $\mathcal{E}_{1,k}^c$ has minimum volume in the sense of Corollary 1.

Proof: Consider the reachable set $\mathcal{R}_{1,k}^c$ in \([27]\). By Corollary 1, under the conditions of Theorem 1, for $k \geq k^*$, $\mathcal{R}_{1,k}^c \subseteq \mathcal{E}_{1,k}^c = \{ \zeta \in \mathbb{R}^{3n|C^T P^C_1 \zeta \leq \alpha_k^c}\}$, where $\mathcal{E}_{1,k}^c$ is the solution of \([8]\). Having this ellipsoid, we look for the projection $\mathcal{E}_{1,k}^c$ onto the $x^p$-hyperplane. Therefore, to obtain the ellipsoid $\mathcal{E}_{1,k}^c$ such that $\mathcal{R}_{1,k}^c \subseteq \mathcal{E}_{1,k}^c$, we use Lemma 10 in the Appendix to obtain this projection.

Corollary 2 Let the conditions of Theorem 1 be satisfied and consider the corresponding matrix $P^C_1$ and function $\alpha_k^c$. Let $P^C_1$ be partitioned as

$$P^C_1 = \begin{cases} P^C_1 = \begin{bmatrix} P^C_1^1 & P^C_1^2 \end{bmatrix}, & P^C_1^1 \in \mathbb{R}^{n \times n}, \ P^C_1^2 \in \mathbb{R}^{n \times 2n}, \ P^C_3 \in \mathbb{R}^{2n \times 2n}. \end{cases}$$

Then, for $k \geq k^*$, $\mathcal{R}_{1,k}^c \subseteq \mathcal{E}_{1,k}^c := \{ x^p \in \mathbb{R}^{n|C^T P^C_1 x^p \leq \alpha_k^c}\}$ with $P^C_k := P^C_1 - P^C_2 (P^C_3)^{-1}(P^C_2)^T$ and $\alpha_k^c := \alpha_k^c$.

Proof: By Theorem 1, the trajectories of \([26]\) satisfy $\zeta^T P^C_1 \zeta \leq \alpha_k^c$ for $k \geq k^*$. By Lemma 10 in the Appendix, the projection of $\zeta^T P^C_1 \zeta \leq \alpha_k^c$ onto the $x^p$-hyperplane is given by $\mathcal{E}_{1,k}^c$ defined above. Thus, in light of Remark 3, the trajectories of the plant dynamics are contained in $\mathcal{E}_{1,k}^c$, i.e., $\mathcal{R}_{1,k}^c \subseteq \mathcal{E}_{1,k}^c$ for all $k \geq k^*$.

5.2 Distance to Critical States

As a second security metric, we propose to use the minimum distance between $\mathcal{R}_{1,k}^c$ and a possible set of critical states $C^c$—states that, if reached, compromise the integrity or safe operation of the system. Such a region might represent states in which, for example, the pressure of a holding vessel exceeds its pressure rating or the level of a liquid in a tank exceeds its capacity. However, because $\mathcal{R}_{1,k}^c$ is not known exactly, this distance cannot be directly computed. Instead, once the ellipsoidal bound $\mathcal{E}_{1,k}^c$ on $\mathcal{R}_{1,k}^c$ is obtained, we compute the minimum distance $d_{1,k}^c$ from $\mathcal{E}_{1,k}^c$ to $C^c$ and use this $d_{1,k}^c$ as an approximation of the distance between $\mathcal{R}_{1,k}^c$ and $C^c$ in terms of the set of sensors being compromised (the attacker's sensor selection matrix $\Gamma$). The distance $d_{1,k}^c$ gives us intuition of how far the actual reachable set $\mathcal{R}_{1,k}^c$ is from $C^c$. Actually, because $\mathcal{E}_{1,k}^c$ is an outer bound on $\mathcal{R}_{1,k}^c$, the distance $d_{1,k}^c$ provides a worst-case quantification of the actual distance to critical states.

The set of critical states in many practical applications can be captured through the union of half-spaces defined by their boundary hyperplanes:

$$C^c := \left\{ x^p \in \mathbb{R}^n \left| \bigcup_{i=1}^N c_i^T x^p \geq b_i \right. \right\},$$

where each pair $(c_i, b_i)$, $c_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, $i = 1, \ldots, N$ quantifies a hyperplane that defines a single half-space.

Corollary 3 Consider the set of critical states $C^c$ defined in \([28]\) and the matrix $P^C_1$ and the function $\alpha_k^c$ obtained in Theorem 1. The minimum distance $d_{1,k}^c$ between the outer ellipsoidal approximation of $\mathcal{R}_{1,k}^c$, $\mathcal{E}_{1,k}^c = \{ x^p \in \mathbb{R}^{n|C^T P^C_1 x^p \leq \alpha_k^c}\}$, and $C^c$ is given by the formula:

$$d_{1,k}^c := \min \left( \frac{|b_i| - \sqrt{c_i^T (P^C)^{-1} c_i/\alpha_k^c}}{c_i^T c_i} \right), \quad i = 1, \ldots, N.$$
Remark 4: If the minimum distance between an ellipsoid centered at the origin
\[ \eta_k \leq \bar{\eta} \]
and the hyperplane
\[ \{ x \in \mathbb{R}^n | x^T P x = 1 \} \],
P \in \mathbb{R}^{n \times n}, P > 0 \text{ and a hyperplane } \{ x \in \mathbb{R}^n | c^T x = b \}, c \in \mathbb{R}^n, b \in \mathbb{R} \text{ is given by the formula } ((|b| - \sqrt{c^T P^{-1} c})/c^T c, \bar{\eta}, \bar{\eta})^T.

It follows that the minimum distance between \( D^\infty \), conformed by the \( N \) hyperplanes in \( \{ 28 \} \), and \( \mathcal{E}_{T, k}^x \) is simply given by \( d_{T, k}^x \) in \( \{ 29 \} \).

Proof: The minimum distance between an ellipsoid centered at the origin \( \{ x \in \mathbb{R}^n | x^T P x = 1 \} \), \( P \in \mathbb{R}^{n \times n}, P > 0 \) and a hyperplane \( \{ x \in \mathbb{R}^n | c^T x = b \} \), \( c \in \mathbb{R}^n, b \in \mathbb{R} \) is given by the formula \( ((|b| - \sqrt{c^T P^{-1} c})/c^T c, \bar{\eta}, \bar{\eta})^T \).

Remark 4: If \( d_{T, k}^x > 0 \), the ellipsoid \( \mathcal{E}_{T, k}^x \) bounding \( \mathcal{R}_{T, k}^x \) and the set of critical states \( D^\infty \) do not intersect; if \( d_{T, k}^x = 0 \), they touch at a point only; and if \( d_{T, k}^x < 0 \) implies that they intersect. In Figure 3, we depict a schematic representation of these ideas.

5.3 Simulation Results

Consider the closed-loop system \( \{ 18 \} \) with matrices as in \( \{ 30 \} \), \( \bar{\eta} = \sqrt{\bar{\tau}} \), and \( \bar{\bar{v}} = 1 \). The controller matrices \( (A^c, B^c, C^c, D^c) \) are designed to guarantee that the \( L_2 \)-gain \( \{ 28 \} \) from the vector of perturbations \( (\eta_k^T, \bar{\eta}_k)^T \) to the performance output \( \Upsilon_k = 0.25 x_k^{p, 3} + \eta_k^3 \) is upper bounded by \( \gamma = 3 \). We use the results in the appendix to design the monitor matrix \( \Pi \) so that, for \( k > k^* = 10 \), \( r_k \Pi r_k \leq 1 \). Using Theorem 1, we obtain \( \mathcal{E}_{T, k}^x \) for all the possible combinations of the sensor attack selection matrix \( \Gamma \). Once we have \( \mathcal{E}_{T, k}^x \), using Corollary \( \{ 2 \} \), we project \( \mathcal{E}_{T, k}^x \) onto the \( x^p \)-hyperplane to obtain \( \mathcal{E}_{T, k}^p \). Note that we have \( k \)-dependent approximations \( \mathcal{E}_{T, k}^p \) of \( \mathcal{R}_{T, k}^p \); however, because \( a < 1 \), the function \( \alpha_k^* \) comprising \( \mathcal{E}_{T, k}^p \) converge exponentially to \( (3 - a)/(1 - a) \). It follows that, in a few time steps, \( \mathcal{E}_{T, k}^x \approx \mathcal{E}_{T, \infty}^x = \{ x \in \mathbb{R}^n | x^T P x \leq (3 - a)/(1 - a) \} \), and thus, \( \mathcal{E}_{T, k}^x \approx \mathcal{E}_{T, \infty}^x \). We present \( \mathcal{E}_{T, \infty}^x \) instead of the time-dependent \( \mathcal{E}_{T, k}^x \). In Figure 4, we show the projection of \( \mathcal{E}_{T, \infty}^x \) onto the \( (x^{p, 1}, x^{p, 2}) \)-hyperplane for different sets of sensor being attacked. Figure 5 depicts the projection of \( \mathcal{E}_{T, \infty}^x \) onto the \( (x^{p, 1}, x^{p, 2}) \)-hyperplane and the distance to the set of critical states \( \mathcal{C}^x = \{ x^p \in \mathbb{R}^3 | x^{p, 1} \leq -15 \} \). In Table 1, we give the numerical values of the volume of \( \mathcal{E}_{T, \infty}^x \) and the distance to the critical states depicted in Figure 5 for different sensors being attacked. Assume, for instance, that two out of the three sensors can be completely secured, i.e., attacks to these sensors are impossible. From Table 1, we note that attacks to sensor two lead to the largest volume of \( \mathcal{E}_{T, \infty}^x \) and the smallest distance to critical states \( d_{T, \infty}^x \). Therefore, if only two sensors can be secured, they should be sensors two and three. Following the same logic, if only one sensor can be secured, then sensor two must be selected because attacks to the remaining sensors, one and three, lead to the smallest \( \mathcal{E}_{T, \infty}^x \) and the largest \( d_{T, \infty}^x \). Thereby, our tools can be used to allocate security equipment to sensors when limited resources are available.
Figure 4: Projection of $E_x^{\infty}$ onto the $(x^{p,2}, x^{p,3})$-hyperplane for different sets of sensor being attacked.

Figure 5: Projection of $E_x^{\zeta}$ onto the $(x^{p,1}, x^{p,2})$-hyperplane for different sets of sensor being attacked and distance to critical states.

$$L = \begin{pmatrix} 0.52 & 0.21 & 0.03 \\ 0.08 & 0.52 & 0.54 \\ 0.02 & 0.02 & 0.35 \end{pmatrix}, \quad \Pi = \begin{pmatrix} 9.50 & -0.76 & -0.05 \\ -0.76 & 7.69 & -0.95 \\ -0.05 & -0.95 & 8.14 \end{pmatrix} \times 10^{-2}, \quad E = I_n, F = I_m. \tag{30}$$

6 Synthesis Tools: Stealthy Attacker’s Reachable Sets

Next, we derive tools for designing $\kappa := (L, \Pi, A^c, B^c, C^c, D^c)$ such that the impact of stealthy attacks on the system dynamics is minimized. In particular, we design $\kappa$ to minimize the volume of $E_x^{\zeta}$ (thus decreasing the size of $R_x^{\zeta, k}$) while guaranteeing some attack-free prescribed performance of the closed-loop system. Consider the extended attacker’s reachable set $R_x^{\zeta, k}$ defined in (27) with matrices $(A, B)$ as in (25). Note that, for every realization of $\kappa = (L, \Pi, A^c, B^c, C^c, D^c)$, using Theorem 1 and Corollary 2, we can obtain $E_x^{\zeta}$ containing $R_x^{\zeta, k}$. Here, we aim at finding the $\kappa^*$ leading to the smallest possible volume of $E_x^{\zeta}$ among all realizations of $(L, \Pi, A^c, B^c, C^c, D^c)$. If we let $\kappa$ be optimization variables rather than given parameters, by Proposition 1, to find $\kappa^*$, we have to find $$(L, A^c, B^c, C^c, D^c)$$ comprising the matrices $(A, B)$, the constants $(a_1, a_2, b)$, and the matrices.
Table 1: Volume of the approximation $E_{T,\infty}^x$ of $R_{T,\infty}^x$ and distance $d_{T,\infty}^x$ to the critical states $C^x$ for different sensors being attacked.

| Attacked Sensors | Volume of $E_{T,\infty}^x$ | Distance to Critical States $d_{T,\infty}^x$ |
|------------------|-----------------------------|------------------------------------------|
| \{1\}           | 150.72                     | 8.07                                     |
| \{2\}           | 453.51                     | 4.20                                     |
| \{3\}           | 219.43                     | 8.60                                     |
| \{1,2\}         | 952.95                     | -2.38                                    |
| \{1,3\}         | 279.50                     | 6.85                                     |
| \{2,3\}         | 2063.46                    | -6.67                                    |
| \{1,2,3\}       | 4300.32                    | -23.01                                   |

To address the synthesis problem, we impose some structure on the matrix $P$ such that, in the new variables $P\rightarrow T$, $PA$ blocks $P$, sensors being attacked.

Then, the transformations $P \rightarrow T^IPPT_1$ and $L \rightarrow T^ILT_2$ take the form:

$$T^IPPT_1 = \begin{bmatrix} Y & I \ 0 & 0 \end{bmatrix} = P(\nu),$$
Therefore, under $P$, we have that the block $T_1^T P A T_1$ is linear in $X, Y,$ and $S$. Next, using the definition of $(A, B)$ in (25), we expand the blocks $T_1^T P A T_1$ and $T_1^T P B$. Note that the matrix $A$ is upper triangular. Let $A$ be partitioned as

$$A = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix},$$

and define the change of controller, observer, and monitor variables:

$$\begin{bmatrix} K \\ M \\ O \end{bmatrix} := \begin{bmatrix} X A P Y \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} U \\ 0 \\ I \end{bmatrix} \left( \begin{array}{c} A^c \\ B^c \\ C^c \\ D^c \end{array} \right) \begin{bmatrix} V^T \\ 0 \end{bmatrix},$$

and

$$R := SL, \quad G := \Pi.$$

Then, $T_1^T P A T_1$ can be written as

$$T_1^T P A T_1 = \begin{bmatrix} Y^T X A_1 Y & Z A_2 \\ 0 & S A_3 \end{bmatrix}$$

$$= \begin{bmatrix} A^P Y + B^P M & A^P + B^P N C^P & -B^P N T^T C^P \\ K & X A^P + O C^P & -O T^T C^P \\ 0 & 0 & S A^P - R(I_m - T^T C^P) \end{bmatrix} =: A(\nu),$$

the block $T_1^T P B$ as

$$T_1^T P B = \begin{bmatrix} Z \\ 0 \\ 0 \\ S \end{bmatrix} B$$

$$= \begin{bmatrix} B^P N(I_m - T^T C^P) & E & B^P N T^T \\ O(I_m - T^T C^P) & X E & O T^T \\ -R(I_m - T^T C^P) & S E & -R T^T \end{bmatrix} =: B(\nu),$$

and the block $W_a$ as

$$W_a = \text{diag} \left[ \frac{1 - a_1}{\eta} I_m, \frac{1 - a_2}{\epsilon} I_n, (1 - b) G \right] =: W(\nu).$$

Therefore, under $T_1, T_2$, and the change of variables (40), the blocks transforms as

$$P \rightarrow P(\nu), \quad T_1^T P A T_1 \rightarrow A(\nu), \quad T_1^T P B \rightarrow B(\nu), \quad W_a \rightarrow W(\nu),$$

with $P(\nu), A(\nu), B(\nu), W(\nu)$ as defined in (37), (42), (44), and (43), respectively. That is, the original blocks, $P, A$ and $P B$, that depend non-linearly on the decision variables $(\kappa, P)$ are transformed into blocks that are affine functions of the new variables $\nu$. If $\nu$ is given and $U$ and $V$ are invertible, the change of variables in (40) and the matrix $T_1$ are invertible and thus $(\kappa, P)$ can be constructed from $\nu$ and they are unique. Moreover, invertible $V$ implies that $T_1$ and $T_2$ are nonsingular and thus the transformations $P \rightarrow T_1^T P T_1$ and $L \rightarrow T_2^T L T_2$ are congruent. The latter implies that

$$P > 0 \text{ and } L \geq 0 \iff P(\nu) > 0 \text{ and } L(\nu) := T_2^T L T_2 = \begin{bmatrix} A(\nu) & 0 \\ 0 & B(\nu) \end{bmatrix} W(\nu) \succeq 0.$$  

If the matrix $P(\nu)$ is positive definite, by the Schur complement, $Y > 0$ and $X - Y^{-1} > 0$, and because $Y X + V U^T = I$ by construction (see Eq. (34)), $V U^T = I - Y X < 0$, i.e., the matrix $V U^T$ is nonsingular. Therefore, if $P(\nu) > 0$, it is always possible to find nonsingular $U$ and $V$ satisfying $Y X + V U^T = I$. In the following lemma, we summarize the discussion presented above.

**Lemma 2** Consider the observer, monitor, and controller matrices $\kappa = (L, \Pi, A^c, B^c, C^c, D^c)$, and the matrices $L$ and $P$ as defined in (51) and (33), respectively. If there exists $\nu = (X, Y, S, R, G, K, O, M, N)$ satisfying $P(\nu) > 0$ and $L(\nu) \succeq 0$ with $P(\nu)$ and $L(\nu)$ as defined in (37) and (47), respectively; then, there exists $(\kappa, P)$ satisfying $P > 0$ and $L \geq 0$. Moreover, for every $\nu$ such that $P(\nu) > 0$ and $L(\nu) \succeq 0$, the change of variables in (40) and matrix $T_1$ are invertible and the $(\kappa, P)$ obtained by inverting (37) and (40) is unique.

**Proof:** Assume that $\nu$ is such that $P(\nu) > 0$ and $L(\nu) \succeq 0$. Because $P(\nu) > 0$, by the Schur complement, $Y > 0$
and $X - Y^{-1} > 0$. Since $YX + VU^T = I$, then $VU^T = I - YX < 0$, i.e., the matrix $VU^T$ is invertible. Hence, it is always possible to factorize $I - YX$ as $VU^T = I - YX$ with square and nonsingular $U$ and $V$. Invertible $U$ and $V$ implies that $T_1$ and $T_2$ are square and nonsingular and thus the transformations $P \to T_2^TP T_1 = P(\nu)$ and $\mathcal{L} \to T_2^T\mathcal{L} T_2 = \mathcal{L}(\nu)$ are congruent. It follows that $P(\nu) > 0$ and $\mathcal{L}(\nu) > 0$ imply $P > 0$ and $\mathcal{L} > 0$ because $P(\nu)$ and $\mathcal{L}(\nu)$ have the same signature as $P$ and $\mathcal{L}$, respectively. Because $P(\nu) > 0$, the matrices $U$, $V$, and $S$ are nonsingular. This implies that the change of variables in (40) and $T_1$ are invertible and lead to unique $(\kappa, P)$ by inverting (37) and (40).

So far, we have derived from the analysis inequalities, $P > 0$ and $\mathcal{L} \geq 0$ in (31), the synthesis inequalities $P(\nu) > 0$ and $\mathcal{L}(\nu) \geq 0$ defined in (37) and (47). If we find a realization of $\nu$ satisfying the synthesis inequalities, we factorize $I - YX$ into nonsingular matrices $V$ and $U$ satisfying $I - YX = VU^T$, use these $V$ and $U$ to solve the equations in (40) to obtain the controller, observer, and monitor matrices, and invert (37) to obtain the ellipsoid $P$. By Lemma 2 this $(\kappa, P)$ satisfies the analysis inequalities in (31).

We aim at minimizing the number of states that the attacker can induce in the system while remaining stealthy, i.e., we want to make the "size" of $R^I_{\kappa, k}$ defined in (21) as small as possible by selecting $\nu$. To achieve this, we seek for the $\nu$ that minimizes the volume of $E^I_{\kappa, k}$ (which would decrease the size of $R^I_{\kappa, k}$). In the analysis case, we look for the matrix $P$ satisfying $P > 0$ and $\mathcal{L} \geq 0$ leading to the minimum volume ellipsoid $E^\kappa_{\ k} = \{\xi \in \mathbb{R}^n | \mathcal{L}^TP \xi \leq \alpha_k^\kappa\}$ bounding $R^I_{\kappa, k}$ (defined in (27)) and then, using Corollary 2 we project this $E^\kappa_{\ k}$ onto the $x^p$-hyperplane to obtain $E^\kappa_{\ k}$. To follow the same approach for synthesis, we would need to minimize the volume of $\xi^TP(\nu)\xi = \alpha_k^\kappa$ subject to $P(\nu) > 0$ and $\mathcal{L}(\nu) \geq 0$. However, the matrix $P$ cannot be written in terms of $\nu$ and minimizing the volume of $\xi^TP(\nu)\xi = \alpha_k^\kappa$ is not an equivalent objective. Instead, because the projection $E^\kappa_{\ k}$ can be written in terms of $\nu$, we seek to minimize the volume of $E^\kappa_{\ k}$ directly.

**Lemma 3** Consider the time-varying ellipsoid $E^\kappa_{\ k} = \{\xi \in \mathbb{R}^n | \mathcal{L}^TP \xi = \alpha_k^\kappa\}$ with matrix $P \in \mathbb{R}^{3n \times 3n}$ as defined in (33), $\zeta = ((x^p)^T, (x^p)^T, \varepsilon)^T$, and $\alpha_k^\kappa \in \mathbb{R}_{>0}, k \in \mathbb{N}$. The projection of $E^\kappa_{\ k}$ onto the $x^p$-hyperplane is given by the ellipsoid $E^\kappa_{\ k} = \{x^p \in \mathbb{R}^n | (x^p)^T Y - 1 x^p = \alpha_k^\kappa\}$ with $Y$ as defined in (34).

**Proof:** For $P$ as defined in (33), by Lemma 10 in the appendix, the boundary of the projection of $E^\kappa_{\ k}$ onto the $x^p$-hyperplane, $E^\kappa_{\ k}$, is given by $(x^p)^T (X - UX^{-1}U^T)x^p = \alpha_k^\kappa$. Using standard block matrix inversion formulas (see, e.g., (30)) and the definition of $Y$ in (34), we have $Y = (X - UX^{-1}U^T)^{-1}$ and therefore $E^\kappa_{\ k}$ can be written in terms of $\nu$ as $E^\kappa_{\ k} = \{x^p \in \mathbb{R}^n | (x^p)^T Y - 1 x^p = \alpha_k^\kappa\}$.

**Lemma 3** implies that, in the new variables, we can minimize the volume of $(x^p)^T Y^{-1} x^p = \alpha_k^\kappa$ to reduce the size of $R^I_{\kappa, k}$. Therefore, in the synthesis case, we seek to minimize the volume of $(x^p)^T Y^{-1} x^p = \alpha_k^\kappa$ subject to $P(\nu) > 0$ and $\mathcal{L}(\nu) \geq 0$. The volume of $E^\kappa_{\ k}$ is proportional to $\sqrt{\det[Y]}$ for any $\alpha_k^\kappa > 0$ (26). Moreover, the function $\sqrt{\det[Y]}$ shares the same minimizer with $\log \det[Y]$ (22). However, the function $\log \det[Y]$ is concave for any positive definite matrix $Y$. To overcome this obstacle, we look for a convex upper bound on $\sqrt{\det[Y]}$ and minimize this bound instead. In order to derive this bound, we use the Arithmetic Mean-Geometric Mean (AM-GM) Inequality which states the following: For any sequence of positive real numbers, $c_1, c_2, \ldots, c_n$, the inequality $(\prod_{j=1}^n c_j)^{1/n} \leq \frac{1}{n} \sum_{j=1}^n c_j$ is satisfied (31).

**Lemma 4** For any positive definite matrix $Y \in \mathbb{R}^{n \times n}$, the following is satisfied:

$$\det[Y]^{1/n} \leq (1/n) \text{trace}[Y] \Rightarrow \sqrt{\det[Y]} \leq (1/n^{n/2}) \text{trace}[Y]^{n/2}. \tag{48}$$

Moreover, because $Y > 0$, $\arg \min\{\text{trace}[Y]^{n/2}\} = \arg \min\{\text{trace}[Y]\}$, i.e., $\text{trace}[Y]^{n/2}$ and $\text{trace}[Y]$ share the same minimizer. Therefore, by minimizing $\text{trace}[Y]$, we minimize an upper bound on $\sqrt{\det[Y]}$.

**Proof:** Let $\lambda_j[Y]$ denote the $j$-th eigenvalue of $Y$, $j = 1, \ldots, n$. Because $Y$ is positive definite, the eigenvalues of $Y$ are strictly positive. Then, because $\det[Y] = \prod_{j=1}^n \lambda_j[Y]$ and $\text{trace}[Y] = \sum_{j=1}^n \lambda_j[Y]$, we have $\left(\prod_{j=1}^n \lambda_j[Y]\right)^{1/n} \leq \frac{1}{n} \sum_{j=1}^n \lambda_j[Y]$ as a direct consequence of the (AM-GM) inequality (31), i.e., the left-hand side of (48) is satisfied for any positive definite $Y$. Given that both $\det[Y]$ and $\text{trace}[Y]$ are strictly positive, the right-hand side of (48) follows from the left-hand side inequality by raising it to the power $n/2$. The function $g(x) := x^{n/2}$ is strictly positive and convex for $x > 0$. Hence, the upper bound $(1/n^{n/2}) \text{trace}[Y]^{n/2}$ in (48) is monotonically increasing in $\text{trace}[Y]$. It follows that, for $Y > 0$, $\arg \min\{\text{trace}[Y]^{n/2}\} = \arg \min\{\text{trace}[Y]\}$ for any $n \in \mathbb{N}$, and the assertion follows.

Up to this point, we have the necessary tools for selecting $\nu$ to reduce the size of the stealthy reachable set $R^I_{\kappa, k}$. That is, we have the constraints, $P(\nu) > 0$ and $\mathcal{L}(\nu) \geq 0$, and the objective function, $\text{trace}[Y]$, needed...
to cast the optimization problem to minimize the volume of $\mathcal{E}^*_{T,k}$. There is, however, one last ingredient to be considered before posing the complete synthesis optimization problem; namely, the attack-free performance of the closed-loop dynamics.

6.2 Attack-Free Observer, Monitor, and Controller Performance

As we now move towards posing the complete syntheses optimization problem, we note that as $\|L\| \to 0$, $\|B^c\| \to 0$, and $\|D^c\| \to 0$, the reachable set $\mathcal{R}^*_{T,k}$ converges to the empty set because the attack-dependent terms in $[18]$ vanish. To make this concrete, without any other considered criteria, the matrices $(L, A^c, B^c)$ leading to the smallest $\mathcal{E}^*_{T,k}$ are trivially given by $(L, A^c, B^c) = 0$. While this is effective at eliminating the impact of the attacker, it implies that we discard the observer and the controller altogether and, therefore, forfeit any ability to control the system and build a reliable estimate of the state. If there are performance specifications that the observer, monitor, and controller must satisfy in the attack-free case (e.g., convergence speed, perturbation-output gain, and closed-loop dynamics spectrum), they have to be added as extra constraints into the minimization problem posted to minimize the volume of $\mathcal{E}^*_{T,k}$.

Several time and frequency domain performance specifications for LTI systems have been expressed as LMI constraints on the closed-loop state-space matrices and quadratic Lyapunov functions [29]. Here, our goal is to compute a single observer [12], monitor [16], and controller [17] that: 1) meets the required attack-free performance specifications, and 2) decreases the set of states reachable by stealthy attackers. For LTI systems and some of the most frequently used performance specifications (e.g., general quadratic performance [29]), there are analysis and synthesis results of the form: System $\Sigma$ satisfies the performance specification $\gamma_j$ if there exists a Lyapunov matrix $P_j$ that satisfies some LMIs in $P_j$. If our synthesis problem involves $N$ specifications, $\gamma_1, \ldots, \gamma_N$, by collecting the LMIs of each specification, we end up having a set of matrix inequalities whose variables are the observer, monitor, and controller matrices, and the Lyapunov matrices, $P_1, \ldots, P_N$, of the specifications (plus auxiliary variables depending on the performance criteria). To pose a tractable co-design considering the volume of $\mathcal{E}^*_{T,k}$ and the specification $\gamma_j$, we must rewrite the specification Lyapunov matrix $P_j$ and its corresponding LMIs in terms of the synthesis variables $\nu$. This can be achieved by imposing $P_j = T_j^* P_j T_j$, where $P$ is the Lyapunov-like matrix associated with $\mathcal{E}^*_{T,k}$ in (33) and $T_j$ denotes some linear transformation. By doing so, we can write the specification LMIs in terms of $P$ and use the change of variables in (40) and the transformations $T_1$ and $T_2$ in (35)-(36) to write these LMIs in terms of $\nu$.

Remark 5 In this manuscript, as attack-free performance specifications, we consider the spectrum of the estimation error dynamics for the observer and, for the controller, the $L_2$ gain from the vector of perturbations to some performance output. We remark that any other specification $\gamma_j$ can be considered in our framework as long as the corresponding Lyapunov matrix $P_j$ and the LMIs can be written in terms of the synthesis variables $\nu$. In Ref. [29], the authors provide a synthesis framework for general quadratic performance – which covers $H_2/H_\infty$ performance, passivity, asymptotic disturbance rejection, peak impulse response, peak-to-peak gain, nominal/robust regulation, and closed-loop pole location. The framework here and the one in [29] are compatible in the sense that any performance specification considered in [29] can be written as LMIs in terms of our synthesis variables $\nu$.

Attack-Free Monitor Feasibility. Note that the observer gain $L$ and the monitor matrix $M$ must be chosen such that Assumption 1 is satisfied. That is, the pair $(L, M)$ must be such that, in the attack-free case ($\delta_k = 0$), there exists some $k^* \in \mathbb{N}$ satisfying $r_k^T \Pi r_k \leq 1$ for all $k \geq k^*$ and $r_k$ solution of (15). Next, we provide constraints in the syntheses variables $\nu$ that have to be fulfilled to satisfy Assumption 1.

Lemma 5 Consider the system matrices $(A^p, C^p, E, F)$ and the perturbation bounds $\bar{v}, \bar{\eta} \in \mathbb{R}_{>0}$. Assume no attacks to the system, i.e., $\delta_k = 0$. For a given $a \in (0, 1)$, constant $\alpha_\infty := (2 - a)/(1 - a)$, and $\epsilon \in \mathbb{R}_{>0}$, if there exist constants $a_1, a_2 \in \mathbb{R}$ and matrices $S \in \mathbb{R}^{n \times n}$, $G \in \mathbb{R}^{n \times m}$, and $R \in \mathbb{R}^{n \times m}$ satisfying:

$$\begin{cases}
  a_1, a_2 \in (0, 1), \\
  a_1 + a_2 \geq a,
\end{cases}
$$

$$S > 0,
\begin{bmatrix}
  aS - (SA^p - RC^p)^T S & 0 & 0 \\
  0 & -(RF)^T & -RF & SE \\
  0 & ET S & \frac{1 - a}{\bar{\eta}} I_m & 0 \\
  0 & \frac{1 - a}{\bar{\eta}} I_m & \frac{1 - a}{\bar{\eta}} I_m & \frac{1 - a}{\bar{\eta}} I_m
\end{bmatrix} \geq 0,
$$

$$G > 0,
\begin{bmatrix}
  \alpha_\infty + \epsilon \eta & - (C^p)^T G & -(C^p)^T G \\
  -GC^p & \frac{1}{\alpha_\infty + \epsilon \eta} I_m - G
\end{bmatrix} \geq 0;
$$

then, for $L = S^{-1} R$ and $M = G$, the residual dynamics (15) satisfies $r_k^T \Pi r_k \leq 1$ for all $k \geq k^*(\alpha, \epsilon, c_1, S)$, where $k^*(\alpha, \epsilon, c_1, S) := \min \{ k \in \mathbb{N} | \alpha^{k-1}(c_1 S e_1 - \alpha_\infty) \leq \epsilon \}$ and $c_1$ denotes the initial estimation error in (15).

14
The proof of Lemma 5 is given in the appendix. The constant $\epsilon$ determines the tightness of the monitor, i.e., the smaller the $\epsilon$ the tighter the bound $r_k^2 \Pi r_k \leq 1$ for $k \geq k^*$. Note, however, that depending on the initial condition $c_k$, too small $\epsilon$ might result in very large $k^* = \min\{k \in \mathbb{N} | a^{k-1}(c_1^T S c_1 - a_\infty^c) \leq \epsilon\}$. See Remark 9 in the appendix for further details.

**Attack-Free Observer Performance.** For the observer, we simply consider the speed of convergence of the estimation error to steady state as a performance criteria. This is quantified by the eigenvalues of the matrix $(A^p - LC^p)$. We restrict the values that $L$ might take by enforcing that the eigenvalues of $(A^p - LC^p)$ are contained in a disk, $\text{Disk}[\beta, \tau]$, centered at $\beta + 0i$ with radius $\tau$. We give a necessary and sufficient condition in terms of the synthesis variables, $L$, and $S$, to achieve this performance.

**Lemma 6 [Observer Performance]** Consider the system matrices $(A^p, C^p)$. If there exist $S \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{n \times m}$ satisfying:

$$S > 0 \quad \text{and} \quad \begin{bmatrix} (\alpha S - S A^p + R C^p)^T & \beta S - S A^p + R C^p \\ \tau^2 S & 0 \end{bmatrix} \geq 0;$$

then, the eigenvalues of $(A^p - LC^p)$ with $L = S^{-1} R$ are contained in the closed disk $\text{Disk}[\beta, \tau]$ centered at $\beta + 0i$ with radius $\tau$.

**Attack-Free Closed-loop Controller Performance.** For the controller, we consider the $L_2$ gain of the closed-loop system from the vector of perturbations, $d_k := (u_k^T, v_k^T) \in \mathbb{R}^{m+n}$, to some performance output, say $s_k \in \mathbb{R}^p$, in the attack-free case (i.e., $\delta_k = 0$). Define the matrices

$$\begin{bmatrix} \hat{A} := \begin{bmatrix} A^p + B^p D^p C^p & B^p C^c \\ B^p C^c & A^c \end{bmatrix} & \hat{B} := \begin{bmatrix} B^p D^F & E \end{bmatrix} \end{bmatrix},$$

and the performance output $s_k := C^s x_k^T + D^s u_k + D_T y_k + D_T v_k$, for some matrices $C^s \in \mathbb{R}^{q \times n}$, $D^s \in \mathbb{R}^{q \times l}$, $D_1 \in \mathbb{R}^{q \times m}$, and $D_2 \in \mathbb{R}^{q \times n}$. Then, the closed-loop dynamics (11) can be written in terms of the extended state $\tilde{z}_k := ((x_k^p)^T, (x_k^c)^T)^T \in \mathbb{R}^{2n}$, the vector of perturbations $d_k$, and the performance output $s_k$:

$$\begin{align*}
\frac{\zeta_{k+1}}{\zeta_k} &= \hat{A} \tilde{z}_k + \hat{B} d_k, \\
\zeta_k &= \hat{C} \tilde{z}_k + \tilde{D} d_k,
\end{align*}$$

with $\hat{C} := (C^s + D^s D^F C^c, D^s C^c)$ and $\hat{D} := (D_1 + D^s D^F, D_2)$. The $L_2$ gain from $d_k$ to $s_k$ of system (52) is given by $\sup_{d_k \in \mathbb{L}_2, e_k \neq 0} \|s_k\|_2 / \|d_k\|_2$ for $\tilde{z}_1 = 0$, where, for any sequence $p_k \in \mathbb{R}^n; \|p_k\|_2 := \sum_{k=1}^{\infty} (\rho_k^2 \rho_k)^{\frac{1}{2}}$. The $L_2$ gain of system (52) equals the $H_\infty$ norm of the transfer matrix $H(s) := \tilde{D} + \hat{C}(s I - \hat{A})^{-1} \hat{B}$, see [33].

**Lemma 7 [Bounded-Real Lemma]** Consider the closed-loop system (52) with input $d_k$ and output $s_k$. If there exist $X \in \mathbb{R}^{2n \times 2n}$ and $\gamma \in \mathbb{R}_{>0}$ satisfying:

$$X > 0, \quad S := \begin{bmatrix} X & \bar{A} \bar{X} & 0 & \bar{C} T \\ \bar{X} \bar{A} & \bar{X} \bar{B} & 0 & \bar{D} T \\ 0 & \bar{B} \bar{X} & \gamma^2 I & \bar{D} \\ \bar{C} & 0 & \bar{D} & I \end{bmatrix} \geq 0;$$

then, the $L_2$ gain of system (52) is less than or equal to $\gamma$, i.e., $\sup_{d_k \in \mathbb{L}_2, d_k \neq 0} \|s_k\|_2 / \|d_k\|_2 \leq \gamma$ for $\tilde{z}_1 = 0$.

The proof of Lemma 7 is omitted here. It is a standard result and details about the proof can be found in, for instance, [22], [33], and references therein.

Using the analysis inequalities in (53), we derive the corresponding synthesis constraints in terms of the synthesis variables $\nu$. Consider the matrices $\bar{X}$ and $\bar{Y}$ introduced in (34), the change of variables in (40a), and the attack-free closed-loop system matrices $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ above defined. Define the matrices:

$$\begin{align*}
\bar{X}(\nu) := \bar{Y}^T \bar{X} \bar{Y} &= \begin{bmatrix} Y & I \\ I & X \end{bmatrix}, \\
\bar{A}(\nu) := \bar{Y}^T \bar{X} \bar{A} \bar{Y} &= \begin{bmatrix} A^p Y + B^p M & A^p + B^p N C^p \\ \bar{K} & X A^p + O C^p \end{bmatrix}, \\
\bar{B}(\nu) := \bar{Y}^T \bar{X} \bar{B} &= \begin{bmatrix} B^p N F & E \\ O F & X E \end{bmatrix}, \\
\bar{C}(\nu) := \bar{C} \bar{Y} &= \begin{bmatrix} C^s Y + D^s M & C^s + D^s N C^p \end{bmatrix}, \\
\bar{D}(\nu) := \bar{D} &= \begin{bmatrix} D_1 + D^s N F & D_2 \end{bmatrix}.
\end{align*}$$

15
Lemma 8 [H∞-Performance] Consider the system matrices \((A^p, B^p, C^p, E, F)\). If there exist \(O \in \mathbb{R}^{n \times l}\), \(X, Y, K \in \mathbb{R}^{n \times n}\), \(M \in \mathbb{R}^{m \times n}\), and \(N \in \mathbb{R}^{l \times m}\), and constant \(\gamma \in \mathbb{R}_{>0}\) satisfying:
\[
\begin{bmatrix}
\dot{X}(\nu) & A(\nu) & B(\nu) & 0 \\
A(\nu) & X(\nu) & 0 & 0 \\
0 & B(\nu)^T & \gamma^2 I & D(\nu)^T \\
\end{bmatrix} \geq 0; \tag{55}
\]
then, the change of variables in \((40a)\) and the matrix \(Y \) in \((54)\) are invertible and the matrices \((X, A^c, B^c, C^c, D^c)\) obtained by inverting \((40a)\) and \(\dot{X}(\nu) = Y^T X Y \) in \((54)\) satisfy \((53)\) and lead to \(\sup_{d_k \in \mathcal{E}_2, d_k \neq 0} \|s_k\|_2 / \|d_k\|_2 \leq \gamma \) for \(\zeta_1 = 0\).

The proof of Lemma 8 is given in the appendix.

6.3 Synthesis of Secure Control Systems

Finally, combining the results presented above, we solve the complete optimization problem to minimize the volume of the ellipsoidal approximation \(E_{T_1,k}^f\) of the security metric \(E_{T_1}^f\) as a function of the set of sensor being attacked (the sensor selection matrix \(\Gamma\)) while guaranteeing certain attack-free system performance.

Theorem 2 Consider the system matrices \((A^p, B^p, C^p, E, F)\), the perturbations bounds \(\bar{e}, \bar{f} \in \mathbb{R}_{>0}\), and the attack sensor selection matrix \(\Gamma\). For given \(a, b \in (0, 1), \alpha_\infty^a = (2 - a)/(1 - a), \tau, \beta \in (0, 1), \gamma \in \mathbb{R}_{>0}\), if there exist \(a_1, a_2 \in \mathbb{R}\) and matrices \(\nu = (X, Y, S, R, G, K, O, M, N)\), \(X, Y, S, K \in \mathbb{R}^{n \times n}\), \(R \in \mathbb{R}^{n \times n}\), \(G \in \mathbb{R}^{m \times n}\), \(O \in \mathbb{R}^{n \times l}\), and \(M \in \mathbb{R}^{m \times n}\), \(N \in \mathbb{R}^{l \times m}\), solution of the convex optimization:

\[
\begin{align}
\min_{\nu, a_1, a_2} & \quad \text{trace}[Y], \\
\text{s.t.} & \quad P(\nu) > 0, \quad L(\nu) > 0, \quad (\text{attacker’s reachable set}), \\
& \quad P(\nu), \quad L(\nu), \quad X(\nu), \quad S(\nu) \quad \text{as defined in} \quad (51), (47), (54), \quad \text{and} \quad (55), \quad \text{respectively; then, the transformation} \quad T_1 \quad \text{in} \quad (55) \quad \text{and the change of variables in} \quad (40) \quad \text{are invertible and the matrices} \quad \{P, L, \Pi, A^c, B^c, C^c, D^c\} \quad \text{obtained by inverting} \quad (40) \quad \text{and} \quad T_1^T P T_1 = P(\nu) \quad \text{in} \quad (57) \quad \text{lead to:} \\
& \quad \text{1) a feasible monitor in the sense of Lemma 5} \quad \text{2) for} \quad k \geq k^*(a, \epsilon, e_1, S) = \min\{k \in \mathbb{N} | \epsilon^k \leq e \} \quad \text{and initial estimation error} \quad e_1 \quad \text{in} \quad (15) \quad \text{R}_{T_1,k} \subseteq E_{T_1,k} = \{x^p \in \mathbb{R}^{3n} \mid (x^p)^T P_{T_1} x^p \leq \alpha_k^p\} \quad \text{with} \quad P_{T_1} := X - U X^{-1} U^T \quad \text{and} \quad \alpha_k^p := a^{k-1} \gamma, \quad \text{and leads to} \quad E_{T_1,k} \quad \text{the eigenvalues of} \quad (A^p - L C^p) \quad \text{being contained in} \quad \text{Disk}[\beta, \tau]; \quad \text{and} \quad 3) \quad \text{sup}_{d_k \in \mathcal{E}_2, d_k \neq 0} \|s_k\|_2 / \|d_k\|_2 \leq \gamma \quad \text{for} \quad \zeta_1 = 0. \\
\end{align}
\]

Moreover, by minimizing \text{trace}[Y], we are minimizing an upper bound on the volume of \(E_{T_1,k}^f\).

\textbf{Proof:} Assume that \((\nu, a_1, a_2)\) satisfy the constraints in \((56)\). By Lemma 2, because \(P(\nu) > 0 \) and \(L(\nu) > 0\), the transformation \(T_1\) in \((55)\) and the change of variables in \((40)\) are invertible, and the \((P, L, \Pi, A^c, B^c, C^c, D^c)\) obtained by inverting \((40)\) and \(T_1^T P T_1 = P(\nu)\) in \((57)\) satisfy the analysis inequalities \(P > 0 \) and \(L \geq 0\) defined in \((31)\) and \((33)\), respectively, and lead to, by assumption, \((49)\), the residual dynamics \((45)\) satisfies \(\epsilon^k L_r e_k \leq 1 \) for all \(k \geq k^*(a, \epsilon, e_1, S)\), and \(\Pi = G \) and \(L = S^{-1} R\). Therefore, by Lemma 2, Lemma 3, and Lemma 5 \(R_{T_1,k} \subseteq \mathcal{E}_{T_1,k} \) with \(P_{T_1} = X - U X^{-1} U^T \) and \(\alpha_k^p := a^{k-1} \gamma, \quad \text{and} \quad \text{the eigenvalues of} \quad (A^p - L C^p) \quad \text{with} \quad L = S^{-1} R \quad \text{are contained in} \quad \text{Disk}[\beta, \tau]; \quad \text{and} \quad \text{4) sup}_{d_k \in \mathcal{E}_2, d_k \neq 0} \|s_k\|_2 / \|d_k\|_2 \leq \gamma \quad \text{for} \quad \zeta_1 = 0. \]

\textbf{Observer, Monitor, Controller, and Ellipsoidal-Approximation Reconstruction.} Given a solution \((\nu, a_1, a_2)\) of the optimization problem in \((56)\):

1. For given \(X \) and \(Y \), compute via singular value decomposition a full rank factorization \(V U^T = I - X Y \) with square and nonsingular \(V \) and \(U\).

2. For given \(\nu\) and invertible \(V \) and \(U\), solve the system of equations \(T_1^T P T_1 = P(\nu)\) and \((40)\) to obtain the matrices \((P, L, \Pi, A^c, B^c, C^c, D^c)\).
3. For given \( S, Y, \mathcal{P}, e_1, \epsilon, \) and \( a \), obtain the monitor convergence time \( k^* \), and \( \mathcal{P}^E_k \) and \( \alpha_k^E \) comprising the ellipsoidal approximation \( \mathcal{E}^E_{k,k} \) of \( \mathcal{R}^E_{k,k} \) as: \( k^* = \min \{ k \in \mathbb{N} | a^{k-1} (\epsilon_1^T S e_1 - \alpha_k^E) \leq 0 \} \), \( \mathcal{P}^E_k = Y^{-1} \), and \( \alpha_k^E = a^{k-1} \lambda_k^E \mathcal{P} \xi_k + \frac{3-a}{1-a} (1-a^{k-1}) \).

By Theorem 2 the reconstructed matrices \((L, I, A^c, B^c, C^c, D^c)\) satisfy the attack-free observer, monitor, and controller performance, and minimize an upper bound on the volume of \( \mathcal{E}^E_{k,k} \).

Remark 6 To obtain tighter approximations \( \mathcal{E}^{\tau}_{k,k} \) of \( \mathcal{R}^E_{k,k} \), once the matrices \((L, I, A^c, B^c, C^c, D^c)\) are computed using Theorem 2 and the above reconstruction procedure, we can close the loop using these matrices and use the analysis result in Theorem 4 to obtain tighter ellipsoidal approximations. That is, Theorem 2 could be used for synthesis only, and then, once \((L, I, A^c, B^c, C^c, D^c)\) are computed, we could use the analysis result in Theorem 7 to obtain less conservative approximations of \( \mathcal{R}^E_{k,k} \).

Remark 7 Note that the constants \( a, b, \epsilon, \tau, \beta, \) and \( \gamma \) in Theorem 2 must be fixed before solving the synthesis optimization problem in (56). The constants \((\tau, \beta, \gamma)\) determine the attack-free observer and controller performance. The constant \( \epsilon \) determines the tightness of the monitor in the attack-free case. The smaller \( \epsilon \) is, the tighter the monitor (see Remark 3 in the Appendix for details). Finally, \( a, b \in (0, 1) \) are, in fact, variables of the optimization problem. However, to linearize some of the constraints, we fix their value before solving (56) and search over \( a, b \in (0, 1) \) to find the optimal \( \nu \). The latter increases the computations needed to find the optimal \( \nu \); however, because \( a, b \in (0, 1) \) (a bounded set), the required grid in \((a, b)\) is of reasonable size.

6.4 Distance to Critical States

As a second objective function for synthesis, we consider the distance between \( \mathcal{R}^E_{k,k} \) and a possible set of critical states \( \mathcal{C}^E \). Because \( \mathcal{R}^E_{k,k} \) is not known exactly, we consider the distance from the approximation \( \mathcal{E}^E_{k,k} \) to \( \mathcal{C}^E \) and use this distance as objective function. We capture the set of critical states through the union of half-spaces defined by their boundary hyperplanes as introduced in (28). In the analysis case, we compute the minimum distance \( d_{k,k}^E \), between \( \mathcal{E}^E_{k,k} \) and \( \mathcal{C}^E \) and use this distance to approximate the proposed security metric (the distance between \( \mathcal{R}^E_{k,k} \) and \( \mathcal{C}^E \)). For synthesis, however, the distance \( d_{k,k}^E \) is highly nonlinear and not convex/concave in the matrices \( \rho \). Instead, we consider the minimum distance between each hyperplane comprising \( \mathcal{C}^E \) and the asymptotic ellipsoidal approximation, \( \mathcal{E}^E_{k,\infty} := \lim_{k \to \infty} \mathcal{E}^E_{k,k} \), and use the weighted sum of these distances as the objective function to be maximized.

Proposition 2 Consider the ellipsoidal approximation \( \mathcal{E}^E_{k,k} \) as introduced in Lemma 3 with matrix \( Y \) and function \( \alpha_k^E \), and the set of critical states:

\[
\mathcal{C}^E = \left\{ x^P \in \mathbb{R}^n \left| \bigcup_{i=1}^N c_i^T x^P \geq b_i \right\} \right.,
\]

where each pair \((c_i, b_i)\), \( c_i \in \mathbb{R}^n \), \( b_i \in \mathbb{R} \), \( i = 1, \ldots, N \) quantifies a hyperplane that defines a single half-space.

The minimum distance \( d_{k,k}^E \) between \( \mathcal{E}^E_{k,k} \) and the hyperplane \( c_i^T x^P = b_i \) is given by \( d_{k,k}^{c,i} = \frac{|b_i| - \sqrt{c_i^2 Y c_i / \alpha_k^E}}{c_i^2 c_i} \).

Proof: The assertion follows by the same arguments as in the proof of Corollary 3.

For synthesis, we aim at maximizing \( \sum_{i=1}^N \rho_i d_{k,k}^{c,i} \), for some \( \rho_i \in \mathbb{R}_{\geq 0} \) satisfying \( \sum_{i=1}^N \rho_i = 1 \), by selecting \((\nu, a_1, a_2)\) subject to (56b). The constant \( \rho_i \) assigns a priority weight to the distance \( d_{k,k}^{c,i} \). Note, however, that because \( \alpha_k^E = a^{k-1} \lambda_k^E \mathcal{P} \xi_k + \frac{3-a}{1-a} (1-a^{k-1}) \) and \( \mathcal{P}^{-1} = \text{diag} \left[ \begin{bmatrix} Y & V \\ V_T & Y \end{bmatrix} \right], \) the term \( c_i^2 Y c_i / \alpha_k^E \) is highly nonlinear and not convex/concave in the matrix \( Y \). Nevertheless, because \( a \in (0, 1) \), we can maximize the weighted sum of the asymptotic minimum distances between \( \mathcal{E}^E_{k,k} \) and \( c_i^2 x^P = b_i \), \( i = 1, \ldots, N \), i.e., \( \tilde{d}_k := \lim_{k \to \infty} \sum_{i=1}^N \rho_i d_{k,k}^{c,i} = \sum_{i=1}^N \rho_i (|b_i| - (\frac{3-a}{1-a} c_i^2 Y c_i)^{-1/2}) / c_i^2 c_i \). Because \((1-a)/(3-a)\) is strictly positive and \( Y \) is positive definite, maximizing \( \tilde{d}_k \) is equivalent to minimizing the linear function: \( \sum_{i=1}^N \rho_i (c_i^2 Y c_i) \). Next, as a corollary of Theorem 2 we pose the optimization problem required to maximize \( \tilde{d}_k \) while guaranteeing the required attack-free performance.
Corollary 4 Consider the setting stated in Theorem 2, the set of critical states $C^e$ defined in (28), and $\tilde{d}_\Gamma$ above defined for some $\rho_i \in \mathbb{R}_{\geq 0}$, $\sum_{i=1}^{N} \rho_i = 1$, $i = 1, \ldots, N$. If there exists $(\nu, a_1, a_2)$ solution of the convex optimization:

$$\begin{align*}
\min_{\nu, a_1, a_2} \sum_{i=1}^{N} \rho_i c_i^T Y c_i,
\text{s.t.} & \quad (56b),
\end{align*}$$

then, the matrices $(P, L, \Pi, A^c, B^c, C^c, D^c)$ obtained by inverting (40) and $T_1^T P T_1 = P(\nu)$ in (57), maximize $\tilde{d}_\Gamma$ and satisfy the desired attack-free system performance in the sense of Theorem 2.

Proof: Let the constraints in (56b) be satisfied. Then, by the same arguments as stated in the proof of Theorem 2, the matrices $(P, L, \Pi, A^c, B^c, C^c, D^c)$ obtained by inverting (40) and $T_1^T P T_1 = P(\nu)$ in (57) lead to a closed-loop dynamics that satisfies the attack-free performance considered in Theorem 2. Also, by the arguments above presented, minimizing $\sum_{i=1}^{N} \rho_i c_i^T Y c_i$ and maximizing $\tilde{d}_\Gamma$ are equivalent objectives.

6.5 Simulation Results

Consider the system matrices $(A^p, B^p, C^p, E, F)$ in (30), and the perturbation bounds $\bar{\eta} = \sqrt{\pi} a$ and $\bar{\nu} = 1.0$. Let $\epsilon = 0.1$ and $(\beta, \tau) = (0, 0.99)$, i.e., the monitor constant $\epsilon$ is fixed to 0.1 and the eigenvalues of the observer closed-loop matrix $(A^p - LCP^p)$ are required to be contained in the disk centered at $0 + 0i$ of radius 0.80, Disk[0, 0.80]. Consider the performance output matrices $C_s = (0, 0, 0.25)$, $D^s = 0_{1 \times 2}$, $D_1 = (0, 0, 1)$, and $D_2 = 0_{1 \times 3}$, and the set of critical states $C^e = \{x^p \in \mathbb{R}^3 \mid x^p,1 \leq -15\}$. The controller must guarantee, in the attack-free case, that the $L_\infty$-gain from the vector of perturbations $d_k = (\eta_k^1, \eta_k^2)^T$ to $s_k = C_s x^p_k + D_s u_k + D_1 \eta_k + D_2 v_k = 0.25 x^p_k + \eta_k^3$ is less than or equal to $\gamma = 3.0$ (as the controller given in (30) for the analysis section). We use Theorem 3 and Corollary 4 to obtain optimal $\kappa = (L, \Pi, A^c, B^c, C^c, D^c)$ minimizing $E_{\infty}^e$ and maximizing $\tilde{d}_\Gamma$, respectively, for all possible combinations of sensors being attacked (all the possible sensor attack selection matrices $\Gamma$). Once we have these $\kappa$, we use the analysis results in Theorem 1 and Corollary 2 to obtain tighter approximations $E_{\infty}^e$ of $\mathcal{R}_{\Gamma, k}$ and use these $E_{\infty}^e$ to obtain tightened $\tilde{d}_\Gamma$. As in the analysis case, we have $k$-dependent approximations $E_{\infty}^e$: however, because $\alpha < 1$, the function $\alpha_k\tilde{c}$ comprising $E_{\infty}^e$ converge exponentially to $(3 - \alpha)/(1 - \alpha)$, hence, in a few time steps, $E_{\infty}^e \approx E_{\infty}^e \approx \{x \in \mathbb{R}^n \mid x^T \mathcal{R}_{\Gamma, k} x \leq (3 - \alpha)/(1 - \alpha)\}$, and thus, $E_{\infty}^e \approx E_{\infty}^e$. We present $E_{\infty}^e$ instead of the time-dependent $E_{\infty}^e$. In Table 2, we present the volume of the asymptotic approximation $E_{\infty}^e$ and the distance $\tilde{d}_\Gamma$ between $E_{\infty}^e$ and the critical states $C^e$ for all possible combinations of sensors being attacked. We show results for the original $\kappa$ in (30); and for the optimal $\kappa$ obtained using Theorem 2 and Corollary 4. Note that the improvement is remarkable using the optimal $\kappa$. To illustrate this improvement, in Figure 6, we show the projection of $E_{\infty}^e$ onto the $(x^{p,1}, x^{p,2})$-hyperplane for sensors $\{2\}, \{2,3\}$, and $\{1,2,3\}$ being attacked. We depict the projections for both the original $\kappa$ in (30) and the optimal one (minimizing trace[$Y$]). For sensor $\{2\}$, we have a 67% improvement in volume and 142% in distance; for $\{2,3\}$, 88% and 247%; and for $\{1,2,3\}$, 67% and 92%, respectively.

7 Conclusion

We have provided mathematical tools – in terms of LMIs – for quantifying the potential impact of sensor stealthy attacks on the system dynamics. In particular, we have given a result for computing ellipsoidal outer approximations on the set of states that stealthy attacks can induce in the system. We have proposed to use the volume of these approximations and the distance to possible dangerous states as security metrics for NCSSs. Then, we have provide synthesis tools (in terms of semidefinite programs) to redesign controllers and monitors such that the impact of stealthy attacks is minimized and the required attack-free system performance is guaranteed. We have presented extensive computer simulations to illustrate the performance of our results.
Figure 6: Projection of $\mathcal{E}_{\infty}^c$ onto the $(x^{p,1}, x^{p,2})$-hyperplane for different sets of sensor being attacked and distance to critical states. Continuous-lines correspond to the original $\kappa$ in (50) and dashed-lines to the optimal $\kappa$ obtained using Theorem 2.

| Attacked Sensors | Original $\kappa$ | Volume | Distance | Optimal $\kappa$ | Volume | Distance |
|------------------|------------------|--------|----------|------------------|--------|----------|
| \{1\}           | 150.72           | 8.07   | 116.94   | 150.16           | 9.12   |
| \{2\}           | 453.51           | 4.20   | 145.31   | 151.80           | 10.15  |
| \{3\}           | 219.43           | 8.60   | 130.62   | 194.50           | 11.92  |
| \{1,2\}         | 952.95           | -2.38  | 456.06   | 487.79           | 5.15   |
| \{1,3\}         | 279.95           | 6.85   | 137.44   | 186.75           | 9.29   |
| \{2,3\}         | 2063.46          | -6.67  | 235.72   | 222.52           | 9.83   |
| \{1,2,3\}       | 4300.32          | -23.01 | 1394.31  | 1371.94          | -1.69  |

Table 2: Volume of the approximation $\mathcal{E}_{\infty}^c$ of $\mathcal{R}_{\infty}^c$ and distance $\hat{d}_r$ to the critical states $C^c$ for different sensors being attacked. We show results for the original $\kappa$ in (50) and for the optimal $\kappa$ obtained using Theorem 2 and Corollary 4.

8 Appendix

8.1 Monitor Design

We use Corollary 1 to obtain outer time-varying ellipsoidal approximations of the reachable set of the estimation error driven by $v_k$ and $\eta_k$ in the attack-free case ($\delta_k = 0$). Once we have this ellipsoid, we project it onto the residual hyperplane to get the ellipsoid $r_k^T \Pi r_k = 1$ of the monitor. Denote by $\psi^c(k, e_1, \eta(k), v(\cdot))$ the solution of (13) at time instant $k$ given the initial estimation error $e_1$ and the infinite disturbance sequences $\eta(\cdot) := \{\eta_1, \eta_2, \ldots\}$ and $v(\cdot) := \{v_1, v_2, \ldots\}$. The reachable set we seek to quantify is given by

$$\mathcal{R}_k^c := \left\{ e \in \mathbb{R}^n \mid e = \psi^c(k, e_1, \eta(k), v(\cdot)); e_1 = x_1 - (C^p)^T y_1; \right\}$$

Lemma 9 Consider the estimation error dynamics (13) with matrices $(A^p, C^p, E, F, L)$, the perturbation bounds $\tilde{v}, \tilde{\eta} \in \mathbb{R}_{>0}$, and assume no attacks to the system, i.e., $\delta_k = 0$. For a given $a \in (0, 1)$, if there exist constants $a_1 = a_1^*, \ldots, a_N = a_N^*$ and matrix $P = P^*$ solution of (38) with $A = (A^p - LC^p), N = 2$, $B^1 = -LF, B^2 = E, W_1 = (1/\tilde{v})I_m, W_2 = (1/\tilde{\eta})I_n, p_1 = m$, and $p_2 = n$; then, $\mathcal{R}_k^c \subseteq \mathcal{E}_k^c := \{ e \in \mathbb{R}^n \mid e^T P^c e \leq \alpha_k^c \}$, with $P^c = P^*$ and $\alpha_k^c := a^{k-1} A^T P^c A^* + (\delta - a)/(1 - a(k-1))/(1 - a)$, and the ellipsoid $\mathcal{E}_k^c$ has minimum volume in the sense of Corollary 1.

Proof: The result follows from Proposition 1 and Corollary 1.

By Lemma 9, the trajectories of the estimation error dynamics are contained in the time-varying ellipsoid $e^T P^c e = \alpha_k^c$. Having this ellipsoid, we look for the matrix $P$ of the monitor leading to the minimum-volume ellipsoid $r_k^T \Pi r_k = 1$ satisfying, for $k \geq k^*$ and some $k^* \in \mathbb{N}$, $r_k^T \Pi r_k = (C^p e_k + \eta_k)^T \Pi (C^p e_k + \eta_k) \leq 1$ for $e_k \in \mathcal{E}_k^c$ and $\eta_k$ such that $\eta^T_k \eta_k \leq \tilde{\eta}$. 

19
Proposition 3 Consider the function $\alpha^c_k$ defined in Lemma 9 and define $\alpha^c_k := \lim_{k \to \infty} \alpha^c_k = (2-a)/(1-a)$. For every $\epsilon \in \mathbb{R}_{>0}$, there exists $k^*(a, \epsilon, e_1, P^c) \in \mathbb{N}$ such that $\alpha^c_k \leq \alpha^c_k + \epsilon$ for all $k \geq k^*$.

**Proof:** The function $\alpha^c_k$ can be written in terms of the constant $\alpha^c_k$ as $\alpha^c_k = a^{-1} e^T_k P^c e_1 + (1-a^{-1}) \alpha^c_k$. Moreover, $\alpha^c_k \leq \alpha^c_k + \epsilon \Leftrightarrow \alpha^c_k - \alpha^c_k \leq \epsilon$ and $\alpha^c_k - \alpha^c_k = a^{-1} (e^T_k P^c e_1 - \alpha^c_k)$. Because $a < 1$, inequality $a^{-1} (e^T_k P^c e_1 - \alpha^c_k) \leq \epsilon$, can always be satisfied for any $\epsilon \in \mathbb{R}_{>0}$ and sufficiently large $k$.

**Remark 8** By Proposition 3, for every $\epsilon \in \mathbb{R}_{>0}$, there exists $k^* \in \mathbb{N}$ such that $\alpha^c_k \leq \alpha^c_k + \epsilon$ for all $k \geq k^*$. The least $k^*$ satisfying $\alpha^c_k \leq \alpha^c_k + \epsilon$ is given by $k^*(a, \epsilon, e_1, P^c) = \min\{k \in \mathbb{N} | a^{-1} (e^T_k P^c e_1 - \alpha^c_k) \leq \epsilon\}$ (see the proof of Proposition 3 above). Note that $\alpha^c_k \leq \alpha^c_k + \epsilon$ for $k \geq k^*$ implies $\mathcal{E}^c_k \subseteq \mathcal{E}^c_k := \{e \in \mathbb{R}^n|e^T P^c e \leq \alpha^c_k + \epsilon\}$ for all $k \geq k^*$. Therefore, for a fixed $\epsilon$ (and corresponding $k^*$), the problem of finding II of the monitor amounts to finding II such that $(C P^c e_k + \eta_k)^T (C^T P^c e_k + \eta_k) \leq 1$ for all $e_k$ and $\eta_k$, satisfying $e^T_k P^c e_k \leq \alpha^c_k + \epsilon$ and $e^T_k P^c e_k \leq \alpha^c_k + \epsilon$ and $\eta_k^T \eta_k \leq \eta_k$. This can be posed as a convex optimization problem using the S-procedure.

Proposition 4 Let the conditions of Lemma 9 be satisfied and consider the corresponding matrix $P^c \in \mathbb{R}^{n \times n}$, the function $\alpha^c_k$, the constant $\alpha^c_k = \lim_{k \to \infty} \alpha_k = (2-a)/(1-a)$, and some $\epsilon \in \mathbb{R}_{>0}$. If there exist $\tau_1, \tau_2 \in \mathbb{R}$ and $\Pi \in \mathbb{R}^{m \times n}$ solution of the following convex optimization:

$$
\begin{align*}
\min_{\Pi, \tau_1, \tau_2} & -\log \det[\Pi], \\
\text{s.t.} & \Pi \geq 0, \ \tau_1 \geq 0, \ \tau_2 \geq 0, \ \text{and} \\
& \begin{bmatrix}
\tau_1 P^c - (C^T \Pi C^p) & - (C^T \Pi C^p) - \Pi C^p \\
- \Pi C^p & \tau_2 I_m - \Pi
\end{bmatrix} \geq 0,
\end{align*}
$$

then, for $\delta_k = 0$ and $k \geq k^*(a, \epsilon, e_1, P^c) = \min\{k \in \mathbb{N} | \alpha^c_k - \alpha^c_k \leq \epsilon\}$, the monitor inequality $r_k^T \Pi r_k \leq 1$ is satisfied for all $e_k$ and $\eta_k$ satisfying $e^T_k P^c e_k \leq \alpha^c_k + \epsilon$ and $\eta_k^T \eta_k \leq \eta_k$.

**Proof:** By Lemma 9, Proposition 3 and Remark 8, for any $\epsilon \in \mathbb{R}_{>0}$ and corresponding $k^*$ satisfying $\alpha^c_k \leq \alpha^c_k + \epsilon$, the trajectories of estimation error dynamics (13) satisfy $e^T_k P^c e_k \leq \alpha^c_k + \epsilon$ for all $k \geq k^*$. By the S-procedure [22], if there exist $\tau_1, \tau_2 \in \mathbb{R}_{>0}$ satisfying

$$(C P^c + \eta_k)^T (C^T P^c e_k + \eta_k) - 1 - \tau_1 (e^T_k P^c e_k - \alpha^c_k - \epsilon) - \tau_2 (\eta_k^T \eta_k - \eta_k) \leq 0,$$

then, $(C P^c + \eta_k)^T (C^T P^c e_k + \eta_k) \leq 1$ is satisfied for all $e_k$ and $\eta_k$ satisfying $e^T_k P^c e_k \leq \alpha^c_k + \epsilon$ and $\eta_k^T \eta_k \leq \eta_k$. Inequality (60) can be written as

$$
\begin{bmatrix}
\tau_1 P^c & - (C^T \Pi C^p) \\
- \Pi C^p & \tau_2 I_m - \Pi
\end{bmatrix} \begin{bmatrix}
\tau_1 P^c - (C^T \Pi C^p) & - (C^T \Pi C^p) - \Pi C^p \\
- \Pi C^p & \tau_2 I_m - \Pi
\end{bmatrix} \geq 0,
$$

with $v_k := (e^T_k, \eta_k^T, 1)^T$. The above inequality is satisfied if and only if $Q$ is positive semidefinite. Therefore, for $k \geq k^*$, $r_k^T \Pi r_k \leq 1$ for any II solution of (59). Again, to ensure that the ellipsoidal bound is as tight as possible, we minimize $\log \det[\Pi^{-1}]$ as this objective shares the same minimizer with $(\det[\Pi])^{-1/2}$ and because for a positive definite II it is convex [22].

**Remark 9** Using Proposition 4, we can design monitors for every $\epsilon \in \mathbb{R}_{>0}$. If we want tight monitors, we need small $\epsilon$ because $\epsilon \approx 0$ yields $\mathcal{E}^c_k \subseteq \mathcal{E}^c_k \approx \mathcal{E}^c_k$ for $k \geq k^*$. That is, the contribution of initial conditions to the outer bound $\mathcal{E}^c_k$ on $\mathcal{E}^c_k$ used in Proposition 4 (see Remark 3) to compute the monitor matrix II has decreased to a small value and mainly the effect of the perturbations $\eta_k$ and $v_k$ is taken into account when designing the monitor matrix II. However, depending on the initial conditions, too small $\epsilon$ might result on very large $k^*$. The values of $\epsilon$ and $k^*$ are related through the expression $k^* = \min\{k \in \mathbb{N} | \alpha^c_k - \alpha^c_k = a^{-1} (e^T_k P^c e_1 - \alpha^c_k) \leq \epsilon\}$ introduced in Proposition 4. Note that, for $e^T_k P^c e_1 \leq \alpha^c_k$, $k^* = 1$ for any $\epsilon \in \mathbb{R}_{>0}$, i.e., $\epsilon$ can be selected arbitrarily small. On the other hand, $e^T_k P^c e_1 > \alpha^c_k$ implies that $k^* \to \infty$ as $\epsilon \to 0$. That is, in this case, there is a trade-off between conservative monitors and convergence time when selecting $\epsilon$.

8.2 Proof of Lemma 5

Assume that the conditions of Lemma 5 are satisfied for some $a \in (0, 1)$, $\epsilon \in \mathbb{R}_{>0}$, $a_1, a_2 \in \mathbb{R}$, and matrices $(S, G, R)$. Because $L = S^{-1} R$ and $\Pi = G$, then $R = S L$, $G = \Pi$, and the matrix inequalities in (49) take the
form:

\[
S > 0, \quad \begin{bmatrix}
\alpha S \\ S(\alpha P - \lambda C P) \\ (\alpha P - \lambda C P)^T S \\ -S L F \\ S \\ -L F^T S \\ \frac{1}{\alpha} \tilde{I}_m \\ 0 \\
0 \\ \frac{1}{\alpha - \epsilon} \tilde{I}_n
\end{bmatrix} \geq 0,
\]

(61)

\[
\Pi > 0, \quad \begin{bmatrix}
\frac{1}{\alpha_0 + \epsilon} S - (C \Pi)^T \Pi (C \Pi) \\ -\Pi \Pi \\
\frac{1}{\alpha_0 + \epsilon} \tilde{I}_m - \Pi 
\end{bmatrix} \geq 0.
\]

(62)

The inequalities in (61) are of the form (7) in Proposition II with \( P = S, A = (\alpha P - \lambda C P), N = 2, B^1 = -L F, B^2 = E, W_1 = (1/\bar{\nu}) I_m, W_2 = (1/\bar{\nu}) I_n, p_1 = m, \) and \( p_2 = n. \) Hence, because \( a, a_1, a_2 \in (0, 1) \) and \( a + a_2 \geq a, \) by Proposition II \( \epsilon_1^k \epsilon_k \leq \alpha_k^2 \) for all \( k \in \mathbb{N}, \epsilon_k \) solution of (15) with \( \tilde{\theta} = 0, \) \( \alpha_k^2 = a_k^2 = a_k^2 \epsilon_1^k \epsilon_1^k + \alpha_k^2 (1 - \alpha_k^2), \) and \( \alpha_k^2 = (2 - a)/(1 - a). \) Note that, for every \( \epsilon > 0, \) we have \( \alpha_k^2 \leq \alpha_k^2 + \epsilon \Rightarrow \alpha_k^2 - \alpha_k^2 = a_k^2 \epsilon_1^k \epsilon_1^k - \alpha_k^2 \leq \epsilon, \) and thus, because \( a \in (0, 1), \) \( \alpha_k^2 \leq \alpha_k^2 + \epsilon \) for all \( k \geq k^* (a, \epsilon, e_1, S) = \min(k \in \mathbb{N}|a_k^2 \epsilon_1^k \epsilon_1^k - \alpha_k^2 \leq \epsilon). \) Inequality \( \alpha_k^2 \leq \alpha_k^2 + \epsilon \) for \( k \geq k^* \) implies \( \epsilon_1^k \epsilon_k \leq \alpha_k^2 + \epsilon \) for \( k \geq k^*, \) i.e., for any \( \epsilon > 0, \) the estimation error \( \epsilon_k \) satisfies \( \epsilon_1^k \epsilon_k \leq \alpha_k^2 + \epsilon \) for all \( k \geq k^*. \) Moreover, because \( \eta_{k^*}^k \eta_k \leq \tilde{\eta} \) for \( k \in \mathbb{N}, \) it is easy to verify that \( w_k^T Q_1 w_k \leq q \) for \( k \geq k^*, \) where \( w_k := (\epsilon_1^k \epsilon_k^T)^T, \) \( Q_1 := \text{diag}[S, I_m] > 0, \) and \( q := \alpha_k^2 + \epsilon + \tilde{\eta} \in \mathbb{R}_{>0}. \) Since \( r_k = C \epsilon_k + \eta_k, \) the monitor inequality, \( r_k^T \Pi r_k \leq 1, \) can be written in terms of \( w_k \) as \( w_k^T Q_2 w_k \leq 1, \) where

\[
Q_2 := \begin{bmatrix} (C \Pi)^T \Pi (C \Pi) \\ \Pi \Pi \end{bmatrix}.
\]

Note that \( w_k^T Q_1 w_k \leq q \Leftrightarrow w_k^T \left( \frac{1}{\bar{\nu}} Q_1 \right) w_k \leq 1, \) because \( q \in \mathbb{R}_{>0} \) and \( Q_1 > 0, \) and thus, if \( w_k^T Q_2 w_k \leq \bar{q} \left( \frac{1}{\bar{\nu}} Q_1 \right) w_k, \) then \( w_k^T Q_2 w_k \leq 1 \) for \( k \geq k^* \) (because \( w_k^T Q_1 w_k \leq q \) only for \( k \geq k^* \)). Inequality \( w_k^T Q_2 w_k \leq \bar{q} \left( \frac{1}{\bar{\nu}} Q_1 \right) w_k \) is satisfied for any \( w_k \in \mathbb{R}_{n+m}^{>0} \) if and only if \( \frac{1}{\bar{\nu}} Q_1 - Q_2 \geq 0. \) The latter inequality equals the right-hand side inequality in (62) and it is satisfied by assumption. Therefore, \( w_k^T Q_2 w_k = r_k^T \Pi r_k \leq 1 \) for \( k \geq k^*, \) \( \Pi = G, \) \( L = S^{-1} R, \) and \( (a, a_1, a_2, \epsilon, S, G, R) \) satisfying (49).

8.3 Proof of Lemma 8

Let \( \nu \) be such that \( \tilde{X} (\nu) > 0 \) and \( S (\nu) > 0. \) Because \( \tilde{X} (\nu) > 0, \) by the Schur complement, \( Y > 0 \) and \( X - Y^{-1} > 0. \) Since \( Y X + V U^T = I \) (see (34)), \( V U^T = I - Y X < 0, \) i.e., the matrix \( V U^T \) is invertible. Hence, it is always possible to factorize \( I - Y X \) as \( V U^T = I - Y X \) with square and nonsingular \( U \) and \( V. \) Invertible \( U \) and \( V \) implies that \( Y \) and the matrix \( T_3 := \text{diag}[Y, Y, I, I] \) are invertible. It follows that the transformations \( X \rightarrow Y X Y = \tilde{X} (\nu) \) and \( S \rightarrow T_3^T S T_3 = S (\nu) \) are congruent. Therefore, \( \tilde{X} (\nu) > 0 \) and \( S (\nu) > 0 \) imply \( X > 0 \) and \( S > 0 \) because \( \tilde{X} (\nu) \) and \( S (\nu) \) have the same signature as \( X \) and \( S, \) respectively. Because \( \tilde{X} (\nu) > 0, \) the matrices \( U \) and \( V \) are nonsingular. The latter implies that the change of variables in (40a) and \( \gamma \) are invertible and lead to unique \( (A', B', C', D') \) by inverting (40c) and \( \tilde{X} (\nu) = Y^T X Y \) in (64), and by Lemma 7 this \( (A', B', C', D') \) leads to \( \sup_{\epsilon_{1,2}, \epsilon_{3,4} \neq 0} ||\epsilon_{1,2}|| / ||\epsilon_{3,4}|| \leq \gamma \) for \( \zeta = 0. \)

8.4 Projection of High Dimensional Ellipsoids onto Coordinate Hyperplanes

**Lemma 10** Consider the ellipsoid:

\[
E := \left\{ x \in \mathbb{R}^n, y \in \mathbb{R}^m \left| \begin{bmatrix} x \\ y \\ \tilde{Q}_1 Q_2 \\ \tilde{Q}_2 Q_3 \\ \tilde{Q}_3 \end{bmatrix} \right| \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \\ Q_5 \\ Q_6 \end{bmatrix} = \alpha \right\},
\]

for some positive definite matrix \( Q \in \mathbb{R}^{(n+m) \times (n+m)} \) and constant \( \alpha \in \mathbb{R}_{>0}. \) The projection of \( E \) onto the \( x \)-hyperplane is given by the ellipsoid:

\[
E' := \left\{ x \in \mathbb{R}^n \left| x^T \left[ Q_1 - Q_2 Q_3^{-1} Q_2^T \right] x = \alpha \right. \right\}.
\]

**Proof:** The matrix \( Q \) is positive definite and thus \( Q_1 \in \mathbb{R}^{n \times n} \) and \( Q_3 \in \mathbb{R}^{m \times m} \) are nonsingular. It follows that \( Q \) can be factorized as:

\[
\begin{bmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -Q_2^T Q_1 I_m \end{bmatrix}^T \begin{bmatrix} Q_1 - Q_2 Q_3^{-1} Q_2^T & 0 \\ 0 & Q_3 \end{bmatrix} \begin{bmatrix} I_n & 0 \\ -Q_2^T Q_1 I_m \end{bmatrix}.
\]

Introduce the change of coordinates:

\[
\begin{bmatrix} x \\ y \end{bmatrix} := \begin{bmatrix} I_n \\ -Q_2^T Q_1 I_m \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.
\]

(63)
In these coordinates, the ellipsoid $E$ is given by
\[
E = \left\{ \begin{array}{l}
\bar{x} \in \mathbb{R}^n, \bar{y} \in \mathbb{R}^m \\
\left[ \begin{array}{c}
\bar{x} \\
\bar{y}
\end{array} \right]^T \left[ \begin{array}{ccc}
Q_1 - Q_2 Q_3^{-1} Q_2^T & 0 \\
0 & Q_1
\end{array} \right] \left[ \begin{array}{c}
\bar{x} \\
\bar{y}
\end{array} \right] = \alpha
\end{array} \right\}.
\]

The matrix $\tilde{Q}$ is block diagonal; therefore, in the new coordinates, the projection of $E$ onto $\bar{y} = 0$ (the $\bar{x}$-hyperplane) and the intersection of $E$ with $\bar{y} = 0$ (and thus the projection onto $\bar{y} = 0$) are equal. The intersection with $\bar{y} = 0$ is simply given by $E^o := \{ (x, \bar{y}) \in E | \bar{y} = 0 \} = \{ x \in \mathbb{R}^n | x^T [Q_1 - Q_2 Q_3^{-1} Q_2^T] x = \alpha \}$. This $E^o$ provides an expression for all the points of $E$ that lie on the $x$-hyperplane. However, from (63), note that $\bar{x} = x$; therefore, $E^o = E'$ and $E'$ provides the locus for all the points of $E$ that lie on the $x$-hyperplane.

References

[1] A. Teixeira, I. Shames, H. Sandberg, and H. Johansson, “A secure control framework for resource-limited adversaries,” *Automatica*, vol. 51, pp. 135 – 148, 2015.

[2] C. Murguia, I. Shames, F. Farokhi, and D. Nešić, “On privacy of quantized sensor measurements through additive noise,” in *proceedings of the 57th IEEE Conference on Decision and Control (CDC)*, 2018.

[3] A. Cardenas, S. Amin, Z. Lin, Y. Huang, C. Huang, and S. Sastry, “Attacks against process control systems: risk assessment, detection, and response,” in *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, 2011, pp. 355–366.

[4] C. M. Ahmed, C. Murguia, and J. Ruths, “Model-based attack detection scheme for smart water distribution networks,” in *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, ser. ASIA CCS ‘17, 2017, pp. 101–113.

[5] E. Rothstein Morris, C. Murguia, and M. Ochoa, “Design-time quantification of integrity in cyber-physical-systems,” in *Proceedings of the 2017 ACM SIGSAC Workshop on Programming Languages and Analysis for Security*, (accepted), 2017.

[6] F. Pasqualetti, F. Dorfler, and F. Bullo, “Attack detection and identification in cyber-physical systems,” *IEEE Transactions on Automatic Control*, vol. 58, pp. 2715–2729, 2013.

[7] C. Murguia and J. Ruths, “On reachable sets of hidden cps sensor attacks,” in *proceedings of the American Control Conference (ACC)*, 2018, 2018.

[8] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli, “False data injection attacks against state estimation in wireless sensor networks,” in *Decision and Control (CDC), 2010 49th IEEE Conference on*, 2010, pp. 5967–5972.

[9] C. Kwon, W. Liu, and I. Hwang, “Security analysis for cyber-physical systems against stealthy deception attacks,” in *American Control Conference (ACC)*, 2013, 2013, pp. 3344–3349.

[10] F. Miao, Q. Zhu, M. Pajic, and G. J. Pappas, “Coding sensor outputs for injection attacks detection,” in *Decision and Control (CDC)*, 2014 IEEE 53rd Annual Conference on, 2014, pp. 5776–5781.

[11] C. Murguia and J. Ruths, “Characterization of a cusum model-based sensor attack detector,” in *proceedings of the 55th IEEE Conference on Decision and Control (CDC)*, 2016.

[12] ——, “Cusum and chi-squared attack detection of compromised sensors,” in *proceedings of the IEEE Multi-Conference on Systems and Control (MSC)*, 2016.

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

[14] J. Chen and R. J. Patton, *Robust model-based fault diagnosis for dynamic systems*. Norwell, MA, USA: Kluwer Academic Publishers, 1999.

[15] E. Kyriakides and M. M. Polycarpou, Eds., *Intelligent monitoring, control, and security of critical infrastructure systems*, ser. Studies in Computational Intelligence. Springer, 2015, vol. 565.
[16] N. Hashemil, C. Murguia, and J. Ruths, “A comparison of stealthy sensor attacks on control systems,” in proceedings of the American Control Conference (ACC), 2018, 2018.

[17] C. Murguia, N. van de Wouw, and J. Ruths, “Reachable sets of hidden cps sensor attacks: Analysis and synthesis tools,” in proceedings of the IFAC World Congress, 2016.

[18] 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, 2009.

[19] S. H. Kafash, J. Giraldo, C. Murguia, A. A. Cardenas, and J. Ruths, “Constraining attacker capabilities through actuator saturation,” in proceedings of the American Control Conference (ACC), 2018, 2018.

[20] Z. Guo, D. Shi, K. H. Johansson, and L. Shi, “Optimal linear cyber-attack on remote state estimation,” IEEE Transactions on Control of Network Systems, vol. PP, no. 99, pp. 1–10, 2016.

[21] Y. Mo and B. Sinopoli, “On the performance degradation of cyber-physical systems under stealthy integrity attacks,” IEEE Transactions on Automatic Control, vol. 61, pp. 2618–2624, 2016.

[22] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, Linear matrix inequalities in system and control theory, ser. Studies in Applied Mathematics. Philadelphia, PA: SIAM, 1994, vol. 15.

[23] H. K. Khalil, Nonlinear Systems, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[24] S. Boyd and L. Vandenberghe, Convex optimization. New York, NY, USA: Cambridge University Press, 2004.

[25] F. Gustafsson, Adaptive filtering and change detection. West Sussex, Chichester, England: John Wiley and Sons, LTD, 2000.

[26] A. B. Kurzhanskiy and I. Valyi, Ellipsoidal calculus for estimation and control. Laxenburg, Austria : IIASA ; Boston : Birkhauser Boston, 1997.

[27] A. A. Kurzhanskiy and P. Varaiya, “Ellipsoidal toolbox (et),” in Proceedings of the 45th IEEE Conference on Decision and Control, 2006, pp. 1498–1503.

[28] A. van der Schaft, L2-Gain and Passivity Techniques in Nonlinear Control. Berlin: Springer, 1999.

[29] C. Scherer, P. Gahinet, and M. Chilali, “Multiobjective output-feedback control via lmi optimization,” IEEE Transactions on Automatic Control, vol. 42, pp. 896–911, 1997.

[30] R. A. Horn and C. R. Johnson, Matrix analysis, 2nd ed. New York, NY, USA: Cambridge University Press, 2012.

[31] J. M. Steele, The Cauchy-Schwarz master class: an introduction to the art of mathematical inequalities. New York, NY, USA: Cambridge University Press, 2004.

[32] G. Garcia and J. Bernussou, “Pole assignment for uncertain systems in a specified disk by state feedback,” IEEE Transactions on Automatic Control, vol. 40, pp. 184–190, 1995.

[33] C. Scherer and S. Weiland, Linear matrix inequalities in control. The Netherlands: Springer-Verlag, 2000.