An H-infinity Approach to Stealth-resilient Control Design

Shaunak D. Bopardikar Alberto Speranzon João P. Hespanha

***Abstract*— We revisit the problem of stealth attack – a coordinated attack through several possible entry points – on a closed loop linear time invariant dynamical system. We propose a notion of the impact of such an attack on the system and consider a novel metric related to the H-infinity norm of the inverse of the system dynamics (assuming it exists) as a measure of the security of the system. We show that the problem can be cast as a linear matrix inequality optimization with the system parameters (observer gain) as the variable. This formulation allows a user to re-design the system from the perspective of minimizing the impact of any stealth attack. Numerical results on simulated data illustrate the security and performance tradeoff using the proposed approach.**

***Index Terms*— Stealth attacks, resilient control, H-infinity design**

I. Introduction

Cyber-physical systems (CPSs) have become ubiquitous in modern times. A CPS is composed of a set of devices directly interacting with the physical layer of the system, such as sensors, actuators, controllers, monitors, etc., and can exchange data by cyber means, such as the intra or internet. CPSs are used in both safety critical systems, such as power networks, water purification systems, airplanes, and in less critical infrastructures, such as heating and lighting systems. As CPSs are inherently interconnected, mostly using IP-based protocols, attacks on such system are growing and they can be rather elaborate such as STUXNET [5] or the latest attack on Ukraine's power grid [1]. Since an attack on a CPS can have major physical consequences, it is highly desirable to develop *design tools* to assess the impact of security measures on the performance of such systems: clearly there are tradeoffs between control performance and security.

In the CPSs security research literature, authors have been considering various models of attacks. In [2] and [6], the authors have considered *false data injections* on static estimators. This attack is modeled as corruption of measurements that are used for state estimation. Conditions on the systems properties that prevent these attacks are derived in [8], where the authors show that an attack exists if and only if the system dynamics have an unstable mode and the associated eigenvector satisfies a technical assumption. More recent work, [12] and [3], provide methods to change the system model so that a class of *stealthy attacks* on the actuators or sensors can be detects. In [10] the authors provide a more general framework to analyze several types of attacks on power systems and networks. In particular, general conditions for attack detection and identifiability for descriptor linear time-invariant systems are defined.

The problem of trading off security and control performance has been researched in recent years. One of the first papers to explore such a tradeoff is [7], where an additive Gaussian noise with zero mean and known covariance is added to the control input. The addition of such noise deteriorates the control performance – measured with respect to an LQG cost – but enables the detectability of the attack. More recently, in [9], the authors have extended the results considering such noise signature (watermarking) to be the output of a hidden Markov model. In a recent paper [11] the authors propose a framework to design secure and computationally efficient cyber-physical systems. In particular, they explore a tradeoff between the sampling period in a control system and the probability of an attacker to be able to decode an encrypted sensor message as a function of the number of bits in the encryption key and the importance such sensor has in the observability of the system. In [14], the authors leverage such framework to develop a control/security design method – they also consider schedulability of processes in a CPU as part of the cost – and formulate an optimization problem from where, for various security levels, one can obtain the control performance vs security Pareto curve. In a similar spirit [13] considers such tradeoffs, where the control performance is related to the tracking error and the security level is associated to the number of bits used to encrypt the sensor and actuator signals.

In this paper, we revisit the problem of stealth attack – a coordinated attack through several possible entry points – on a closed loop linear time invariant dynamical system. The main difference with respect to our previous work [3] is that we now consider the case of noise in the system, and thereby focus on minimizing the consequences of a stealth attack on the system. We propose a

Shaunak D. Bopardikar (email: bopardsd@utrc.utc.com) is with the United Technologies Research Center, Berkeley, CA, USA. Alberto Speranzon (email: alberto.speranzon@gmail.com) was with United Technologies Research Center, East Hartford, CT, USA when this work was performed. He is presently with Honeywell Aerospace – Advanced Technologies. João P. Hespanha (email: hespanha@ece.ucsb.edu) is with the Electrical and Computer Engineering Department at University of California Santa Barbara, CA, USA. This work was supported in part by United Technologies Research Center, under the Cyber-Physical Systems Security Initiative.

notion of the impact of such an attack on the system and consider a novel metric related to the H-infinity norm of the inverse of the system dynamics (assuming it exists) as a measure of the security of the system. We show that the problem can be cast as a linear matrix inequality optimization problem, under certain assumptions, with the system parameters (observer gains) as decision variables. This formulation allows a user to re-design the control loop so that the impact of any stealth attack is minimized. Numerical results on simulated data illustrate the security (maximum impact of stealth attack) and performance (ℓ_2 gain of the original system) tradeoff using the proposed approach.

This paper is organized as follows. Section II describes how the system and the attack are modeled. Section III provides an analysis and introduces the H-infinity norm based metric for measuring the impact of stealth attacks. Section IV describes our proposed approach on how to re-design the system parameters trading off security and closed-loop performance. Section V summarizes the results on synthetic data. Finally, Section VI summarizes our conclusions and directions for future research.

II. Problem Formulation

A. System Modeling

We consider the discrete-time version of a linear time invariant dynamical system given by

$$x_{k+1} = Ax_k + Bu_k + B_n n_k\,, \tag{1}$$
$$y_k = Cx_k + Du_k + D_n n_k\,, \tag{2}$$

where at every discrete time instant $k \in \mathbb{Z}_{\geq 0}$. The state is $x_k \in \mathbb{R}^n$ and the measurement vector $y_k \in \mathbb{R}^m$. The noise vector $n_k \in \mathbb{R}^q$ is assumed to be bounded. The control input is $u_k \in \mathbb{R}^p$ and the systems matrices are $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{m\times n}$, $D \in \mathbb{R}^{m\times p}$, $B_n \in \mathbb{R}^{n\times q}$, $D_n \in \mathbb{R}^{m\times q}$. We will assume that the system is stabilizable and that it operates in closed-loop, i.e., the control u_k is given by a state feedback law,

$$u_k = K\hat{x}_k,$$

where the matrix $K \in \mathbb{R}^{p\times n}$ is designed in a manner that the matrix $A + BK$ is stable and where $\hat{x}_k$ is the estimate of the state x_k. We thus assume that there is an observer that estimates the state from the measurements. The observer dynamics are given by

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k - L(y_k - C\hat{x}_k - Du_k)\,, \tag{3}$$

where the matrix $L \in \mathbb{R}^{n\times m}$ is designed so that the matrix $A + LC$ is stable and with eigenvalues smaller (in absolute value) of those of $A + BK$.

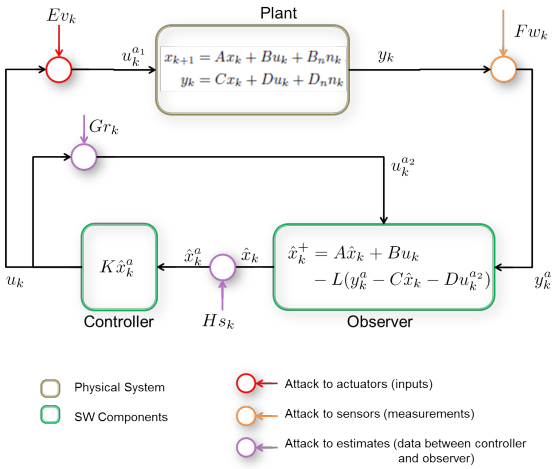

Fig. 1. A general closed loop controlled system with a plant a controller and observer. Attacks to the actuators and sensors are shown in red and orange, respectively, whereas an attack to the controller and observer signals is shown in purple.

B. Attack Modeling

Attacks may occur in the form of additive signals at different points of the closed loop system. Various cyber security mechanisms are assumed to be present in a CPS to prevent an attacker to tamper with the system. Such mechanisms can reduce the number of attack points and for each attack point reduce the degree of freedom of an attack, namely the components and/or signals that can be attacked.

For example, using authentication methods with different levels of security, certain sensors could be more difficult to attack. Thus given a CPS, it is possible to rank the attacks points from "likely" to "unlikely" and for a given attack point determine what could be the attack "progression", i.e., rank the number of devices/signals that could be corrupted as a function of time. Examples of attack points for a closed-loop control system are shown in Figure 1 and we refer the reader to [3] for a discussion on deployment of security countermeasures a function of the ranking of attack points. In this work, we will assume that the software modules that implement the control law and the observer are secure, i.e., it is very unlikely for an attacker to modify system parameters and/or control/estimation algorithms and therefore, $G = H = 0$ in Figure 1.

The closed loop system in Figure 1 is described by:

$$\begin{bmatrix} x_{k+1} \\ \hat{x}_{k+1} \end{bmatrix} = \begin{bmatrix} A & BK \\ -LC & A + BK + LC \end{bmatrix} \begin{bmatrix} x_k \\ \hat{x}_k \end{bmatrix} + \begin{bmatrix} BE \\ -LDE \end{bmatrix} v_k + \begin{bmatrix} 0 \\ -LF \end{bmatrix} w_k + \begin{bmatrix} B_n \\ LD_n \end{bmatrix} n_k, \tag{4}$$

and

$$y_k^a = \begin{bmatrix} C & DK \end{bmatrix} \begin{bmatrix} x_k \\ \hat{x}_k \end{bmatrix} + DEv_k + Fw_k + D_n n_k \,, \tag{5}$$

where the attack signals are $v \in \mathbb{R}^e$, $w \in \mathbb{R}^f$. The matrices E, F, are of compatible dimensions and their column dimensions represent the degree of freedom of the attacker. For example, if the matrix $E \in \mathbb{R}^{p\times e}$ with $p \geq e$, then the attacker has e degrees of freedom to affect the p control signals.

The matrices E and F serve as a lumped model of the cyber layer and capture the capability of an attacker to affect the actual physical signals u, y and $\hat{x}$. These matrices could be the identity of appropriate dimensions, which would model the capability of an attacker to directly corrupt each individual component of u, y and $\hat{x}$. Without loss of generality, we assume that both of these matrices are full column rank. Otherwise, they only provide redundant degrees of freedom to the attacker.

System (4) can be transformed into

$$\begin{aligned} \hat{x}_{k+1} - x_{k+1} = (A+LC)(\hat{x}_k - x_k) - (LDE + BE)v_k \\ - LFw_k + (LD_n - B_n)n_k, \end{aligned} \tag{6}$$

by subtracting the first component of (4) from the second.

From the point of view of monitoring attacks in the system, we can define the following output

$$\begin{aligned} y_k^m &:= \hat{y}_k - y_k^a \\ &= C\hat{x}_k + DK\hat{x}_k - Cx_k - DK\hat{x}_k - DEv_k \\ &\quad - Fw_k - D_n n_k \\ &= C(\hat{x}_k - x_k) - DEv_k - Fw_k - D_n n_k. \end{aligned} \tag{7}$$

C. Problem Statement

We first formalize the notion of stealth attacks considered in this paper. For the system (6)-(7) the output is a function of the error $\hat{x}_k - x_k$, the attack vectors v_k, w_k and the noise vectors n_k; namely we have that $y_k^m(\hat{x}_k - x_k, v_k, w_k, n_k)$. For short, define $\mathbf{a}_k = (v_k', w_k')'$.

In this paper, we consider noise vectors $n : \mathbb{Z}_{\geq 0} \to \mathbb{R}$, as signals with a maximum finite ℓ_2 norm, i.e.,

$$\mathcal{N}_\epsilon := \{n \,:\, \sum_{k=0}^{+\infty} \|n_k\|^2 \leq \epsilon, \forall k \in \{1, \ldots, +\infty\}\} \,,$$

where $\epsilon \geq 0$, is a given parameter.

We can then define an attack as stealth in the following manner, similar in spirit to what was proposed in [10].

Definition II.1 (Stealth Attack)**.** *Given a linear-time invariant system* (6)–(7) *and* $\epsilon, \delta \geq 0$*, an attack* $\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_T$*, is defined to be* stealth *if there exists an admissible noise signal* $n \in \mathcal{N}_\epsilon$*, such that*

$$\sum_{k=1}^{\infty} \|y_k^m(\hat{x}_k - x_k, \mathbf{a}_k, n_k)\|^2 \leq \delta \,.$$

If such a stealth attack exists, then its impact *is defined as the* ℓ_2 *norm of* $\mathbf{a}$*, i.e.,*

$$I := \sum_{k=1}^{\infty} \|\mathbf{a}_k\|^2 \,.$$

In other words, this means that there exists an action of the attacker on the internal signals of the system that is perceived as the effect of an admissible noise signal within the system. Specifically, we will address the problem of designing a technique to modify the system parameters (specifically the observer gain L) so that the modified system admits only a minimal impact of any stealth attack.

III. Analysis

This section provides an analysis of the system dynamics and related conditions, which will allow us to characterize the impact of a stealth attack. We begin the analysis by defining the following matrices:

$$\begin{aligned} \mathbf{A} &:= A + LC \,, \\ \mathbf{B} &:= \begin{bmatrix} -(B+LD)E & -LF \end{bmatrix} \\ &= -\begin{bmatrix} BE & 0 \end{bmatrix} - L\begin{bmatrix} DE & F \end{bmatrix} \,, \\ \mathbf{C} &:= C \,, \quad \mathbf{D} := -\begin{bmatrix} DE & F \end{bmatrix} \,, \\ \mathbf{B}_n &:= LD_n - B_n \,, \quad \mathbf{D}_n := -D_n \,. \end{aligned}$$

then the system (6)–(7) can be rewritten as

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{a}_k + \mathbf{B}_n n_k \,, \\ y_k^m &= \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{a}_k + \mathbf{D}_n n_k \,, \end{aligned} \tag{8}$$

with $\mathbf{x}_k := \hat{x}_k - x_k$ being the error vector and $\mathbf{a}_k = (v_k', w_k')'$ the attack vector.

In what follows, we will first consider the special case of zero noise, which arises when $\epsilon = \delta = 0$, and then present the general case with noise.

A. Special case of no noise

Let $\mathbf{V} := \begin{bmatrix} \mathbf{D}' & \mathbf{B}'\mathbf{C}' & \mathbf{B}'\mathbf{A}'\mathbf{C}' & \ldots & \mathbf{B}'\mathbf{A}'^{m-1}\mathbf{C}' \end{bmatrix}'$, where m is the dimension of $\mathbf{A}$. Then, the following result summarizes conditions for the extreme case, namely under the zero noise assumption.

Theorem III.1 (Zero Noise)**.** *If* $\epsilon = \delta = 0$*, then a stealth attack on the system* (6)–(7) *has impact*

$$I = \begin{cases} 0, & \text{if } \mathbf{D} \text{ has full column rank,} \\ +\infty, & \text{if the null space of } \mathbf{V} \text{ is non-empty.} \end{cases}$$

Proof. We assume without loss of generality that $\mathbf{x}_0 = 0$. Since $\epsilon = \delta = 0$, from the definition of stealth attack,

we have that $y_k^a = 0$ for any $k \geq 0$. This holds true if and only if for any finite $T \geq 0$,

$$\begin{aligned}
&\mathbf{D}\mathbf{a}_0 = 0\,,\\
&\mathbf{C}\mathbf{B}\mathbf{a}_0 + \mathbf{D}\mathbf{a}_1 = 0\,,\\
&\quad\vdots\\
&\mathbf{C}\mathbf{A}^{T-1}\mathbf{B}\mathbf{a}_0 + \cdots + \mathbf{C}\mathbf{B}\mathbf{a}_{T-1} + \mathbf{D}\mathbf{a}_T = 0\,.
\end{aligned}$$

For the first part, if $\mathbf{D}$ has full column rank, then $\mathbf{a}_0 = 0$. This implies that $\mathbf{a}_1 = 0$ from the second equation, and so on. Thus, $\mathbf{a}_k = 0$, for any $k \geq 0$, which implies that the impact is zero.

For the second part, let z denote a vector in the null space of $\mathbf{V}$. Consider the following attack: $\mathbf{a}_k = z, \forall k \geq 0$. Clearly this attack is stealth until time $t = m$. For $t \geq m+1$, we can invoke Cayley-Hamilton Theorem to write

$$\mathbf{A}^t = \sum_{i=0}^{m-1} \alpha_i \mathbf{A}^i,$$

$$\Rightarrow \mathbf{C}\mathbf{A}^t\mathbf{B}z = \mathbf{C}(\sum_{i=0}^{m-1} \alpha_i \mathbf{A}^i)\mathbf{B}z = \sum_{i=0}^{m-1} \alpha_i \mathbf{C}\mathbf{A}^i\mathbf{B}z = 0.$$

Therefore, the attack remains stealth for every $k \geq 0$. Since the vector z can have arbitrarily large magnitude, the impact of the attack is infinite. □

Remark III.1 (Connection with existing results)**.** *Theorem III.1 has been known in literature in several forms, cf. [12] and other references, related to output nulling attacks using geometric control theory. For completeness, we included a version of this result rephrased in the context of the impact of stealth attacks. Given that zero noise is an extreme case, it was natural to expect that the impact can be extreme as well, namely either zero or infinite, depending on whether an attacker can choose signals in a certain subspace or not.*

B. Noise Case: Redesign of Gain matrices

In this section, we analyze how the attacker can leverage the presence of noise, which may be present in the system dynamics and/or in the measurement processes, to carry out a stealth attack.

As a consequence of the condition in Theorem III.1, we assume here that $\mathbf{D}$ has full column rank. This implies that[1] $\mathbf{D}^\dagger\mathbf{D} = I$, so we can invert the system (8) and interpret the attack vector $\mathbf{a}$ as the output of the following inverse system:

$$\begin{aligned}
\mathbf{x}^+ &= (\mathbf{A} - \mathbf{B}\mathbf{D}^\dagger\mathbf{C})\mathbf{x} + \mathbf{B}\mathbf{D}^\dagger y^m\\
&\quad + (\mathbf{B}_n - \mathbf{B}\mathbf{D}^\dagger\mathbf{D}_n)n\,, \qquad (9)\\
\mathbf{a} &= -\mathbf{D}^\dagger\mathbf{C}\mathbf{x} + \mathbf{D}^\dagger y^m - \mathbf{D}^\dagger\mathbf{D}_n n\,.
\end{aligned}$$

[1]We denote with $D^\dagger$ the Moore-Penrose pseudoinverse of the matrix D.

Define,

$$\begin{aligned}
&\mathcal{A} := \mathbf{A} - \mathbf{B}\mathbf{D}^\dagger\mathbf{C}\,, \quad \mathcal{B} := \begin{bmatrix}\mathbf{B}\mathbf{D}^\dagger & (\mathbf{B}_n - \mathbf{B}\mathbf{D}^\dagger\mathbf{D}_n)\end{bmatrix}\,,\\
&\mathcal{C} := -\mathbf{D}^\dagger\mathbf{C}\,, \quad \mathcal{D} := \begin{bmatrix}\mathbf{D}^\dagger & -\mathbf{D}^\dagger\mathbf{D}_n\end{bmatrix}\,,
\end{aligned}$$

with the noise vector $\mathbf{n} := [y'\ n']'$.

Then, the following result yields an upper bound on the impact of stealth attack on (9).

Theorem III.2 (Bound on maximal impact)**.** *Suppose that the matrix $\mathcal{A}$ is Hurwitz and that the noise vector in the system (6)–(7) belongs to the admissible set $\mathcal{N}_\epsilon$. Then, the system (6)–(7) admits a stealth attack having impact at most*

$$\gamma(\epsilon + \delta),$$

where γ is the H_∞ norm of the system (9). If $\mathcal{A}$ is not Hurwitz, then the impact of the stealth attack is infinite.

Proof. Let us assume again, without loss of generality, that $\mathbf{x}_0 = 0$. From the definition of the stealth attack and from the admissible noise set, the noise vector $\mathbf{n}$ has finite ℓ_2 norm given by $\epsilon+\delta$. The claim now follows from the definition of the H_∞ norm for system (9). □

To minimize the impact of the admissible stealth attack in the system, we seek to minimize the H_∞ norm of the system (9) that, in general, is a function of the observer gain matrix L. If, for some value of L, the H_∞ norm of (9) is finite, then $\mathcal{A}(L)$ is guaranteed to be Hurwitz. However, it does not guarantee that the matrix $\mathbf{A}(L)$ is Hurwitz. This criterion needs to be specified separately as a constraint and is the topic of the next section.

IV. Re-design to Optimize for Security

The goal is to design the estimator gain L such that the H_∞ norm of system (9) is minimized. We will need to add the constraint that the closed-loop stability of the original system (8) is preserved. In other words, the maximum eigenvalue of $A + LC$ can only take values that are strictly less than one.

To minimize the H_∞ norm of the dynamical system (9) (equivalently, the ℓ_2 gain of the system), we need to solve the following optimization [4],

$$\begin{aligned}
\min_{\gamma,\bar{P},L}\quad & \gamma\\
\text{s.t.}\quad & \bar{P} \succ 0,\\
& \begin{bmatrix}\mathcal{A} & \mathcal{B}\\ \mathcal{C} & \mathcal{D}\end{bmatrix}' \begin{bmatrix}\bar{P} & 0\\ 0 & I\end{bmatrix} \begin{bmatrix}\mathcal{A} & \mathcal{B}\\ \mathcal{C} & \mathcal{D}\end{bmatrix} \prec \begin{bmatrix}\bar{P} & 0\\ 0 & \gamma^2 I\end{bmatrix}\,, \qquad (10)
\end{aligned}$$

From the definitions,

$$\begin{aligned}
\mathcal{A} &= \mathbf{A} - \mathbf{B}\mathbf{D}^\dagger\mathbf{C}\\
&= A - \mathbf{B}\mathbf{D}^\dagger\mathbf{C} + LC\\
&= A + \begin{bmatrix}BE & 0\end{bmatrix}\mathbf{D}^\dagger C + L\begin{bmatrix}DE & F\end{bmatrix}\mathbf{D}^\dagger C + LC\\
&= \bar{A} + L\bar{C}\,,
\end{aligned}$$

where

$$\bar{A} := A + \begin{bmatrix} BE & 0 \end{bmatrix} \mathbf{D}^\dagger C\,, \quad \bar{C} := -\mathbf{D}\mathbf{D}^\dagger C + C\,.$$

Similarly, we can write $\mathcal{B}$ as

$$\begin{aligned}\mathcal{B} &= \begin{bmatrix} \mathbf{B}\mathbf{D}^\dagger & (\mathbf{B}_n - \mathbf{B}\mathbf{D}^\dagger \mathbf{D}_n) \end{bmatrix} \\ &= \begin{bmatrix} 0 & I \end{bmatrix} \mathbf{B}_n + \mathbf{B} \begin{bmatrix} \mathbf{D}^\dagger & \mathbf{D}^\dagger \mathbf{D}_n \end{bmatrix} \\ &= \begin{bmatrix} 0 & I \end{bmatrix} (LD_n - B_n) + \begin{bmatrix} BE & 0 \end{bmatrix} \begin{bmatrix} \mathbf{D}^\dagger & \mathbf{D}^\dagger \mathbf{D}_n \end{bmatrix} \\ &\quad + L \begin{bmatrix} DE & F \end{bmatrix} \begin{bmatrix} \mathbf{D}^\dagger & \mathbf{D}^\dagger \mathbf{D}_n \end{bmatrix} \\ &= \bar{B} + L\bar{E},\end{aligned}$$

where we defined

$$\begin{aligned}\bar{B} &:= -\begin{bmatrix} 0 & I \end{bmatrix} B_n - \begin{bmatrix} BE & 0 \end{bmatrix} \begin{bmatrix} \mathbf{D}^\dagger & \mathbf{D}^\dagger \mathbf{D}_n \end{bmatrix} \\ \bar{E} &:= \begin{bmatrix} 0 & I \end{bmatrix} D_n - \mathbf{D} \begin{bmatrix} \mathbf{D}^\dagger & \mathbf{D}^\dagger \mathbf{D}_n \end{bmatrix}.\end{aligned}$$

Now, the maximum eigenvalue constraint on $A + LC$ is satisfied if and only if there exists a $\tilde{P} \succ 0$ for which

$$(A + LC)' \tilde{P} (A + LC) - \tilde{P} \prec 0\,.$$

This constraint can be combined with the constraint (10), to obtain:

$$\begin{bmatrix} \bar{A} + L\bar{C} & 0 & \bar{B} + L\bar{E} \\ 0 & A + LC & 0 \\ -\mathbf{D}^\dagger C & 0 & \mathcal{D} \end{bmatrix}' \cdot \begin{bmatrix} \bar{P} & 0 & 0 \\ 0 & \tilde{P} & 0 \\ 0 & 0 & I \end{bmatrix} \cdot \begin{bmatrix} \bar{A} + LC & 0 & \bar{B} + L\bar{E} \\ 0 & A + LC & 0 \\ -\mathbf{D}^\dagger C & 0 & \mathcal{D} \end{bmatrix} \prec \begin{bmatrix} \bar{P} & 0 & 0 \\ 0 & \tilde{P} & 0 \\ 0 & 0 & \gamma I \end{bmatrix}.$$

Employing a standard trick from [4], define $\bar{Q} := \bar{P}^{-1}, Y := \bar{Q}L, \tilde{Q} := \tilde{P}^{-1}, X := \tilde{Q}L$, we obtain

$$\begin{bmatrix} \bar{Q}\bar{A} + Y\bar{C} & 0 & \bar{Q}\bar{B} + Y\bar{E} \\ 0 & \tilde{Q}A + XC & 0 \\ -\mathbf{D}^\dagger C & 0 & \mathcal{D} \end{bmatrix}' \begin{bmatrix} \bar{Q}^{-1} & 0 & 0 \\ 0 & \tilde{Q}^{-1} & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \bar{Q}\bar{A} + Y\bar{C} & 0 & \bar{Q}\bar{B} + Y\bar{E} \\ 0 & \tilde{Q}A + XC & 0 \\ -\mathbf{D}^\dagger C & 0 & \mathcal{D} \end{bmatrix} \prec \begin{bmatrix} \bar{Q} & 0 & 0 \\ 0 & \tilde{Q} & 0 \\ 0 & 0 & \gamma^2 I \end{bmatrix}.$$

By taking the Schur complement, this can be written as

$$\begin{bmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12}^T \\ \mathcal{Q}_{12} & \mathcal{Q}_{22} \end{bmatrix} \succ 0, \tag{11}$$

where the matrices

$$\mathcal{Q}_{11} := \begin{bmatrix} \bar{Q} & 0 & 0 \\ 0 & \tilde{Q} & 0 \\ 0 & 0 & \gamma^2 I \end{bmatrix}, \mathcal{Q}_{22} := \begin{bmatrix} \bar{Q} & 0 & 0 \\ 0 & \tilde{Q} & 0 \\ 0 & 0 & I \end{bmatrix},$$

$$\mathcal{Q}_{12} := \begin{bmatrix} \bar{Q}\bar{A} + Y\bar{C} & 0 & \bar{Q}\bar{B} + Y\bar{E} \\ 0 & \tilde{Q}A + XC & 0 \\ -\mathbf{D}^\dagger C & 0 & \mathcal{D} \end{bmatrix}.$$

This problem is not convex since we need to ensure that $\bar{Q}^{-1}Y = \tilde{Q}^{-1}X = L$, which leads to a non-linear equality constraint. To convexify the problem and therefore make it tractable, we introduce additional constraints

$$\bar{Q} = \tilde{Q}\,, \qquad Y = X\,,$$

yielding an LMI in the variables $\tilde{Q}, \bar{Q}, X, Y$. Such convexication provides a sufficient condition for the satisfaction of (10) and the maximum eigenvalue constraint, leading to the following convex problem:

$$\begin{aligned} \min_{\gamma, \bar{Q}, \tilde{Q}, X, Y} \quad & \gamma \\ \text{s.t.} \quad & \begin{bmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12}^T \\ \mathcal{Q}_{12} & \mathcal{Q}_{22} \end{bmatrix} \succ 0, \\ & X = Y\,, \quad \bar{Q} = \tilde{Q}, \\ & \tilde{Q} \succ 0\,, \quad \bar{Q} \succ 0\,. \end{aligned} \tag{12}$$

The optimal solution to problem (12) (if it exists) is an upper bound on the optimal solution to the original problem (i.e., problem (12) without the additional constraints of $\bar{Q} = \tilde{Q}$ and $X = Y$). Therefore, infeasibility of (12) only implies that one is required to solve the original problem for that particular system.

A. Re-design under Performance Constraints

We now address the case when there is a performance constraint is expressed as a desired bound, η^2, on the ℓ_2 gain between the noise n and y^m for the system (8). We can proceed as before with the maximum eigenvalue constraint replaced by

$$\begin{bmatrix} A + LC & LD_n - B_n \\ C & \mathbf{D}_n \end{bmatrix}' \begin{bmatrix} \tilde{P} & 0 \\ 0 & I \end{bmatrix} \cdot \begin{bmatrix} A + LC & LD_n - B_n \\ C & \mathbf{D}_n \end{bmatrix} \prec \begin{bmatrix} \tilde{P} & 0 \\ 0 & \eta I \end{bmatrix}, \quad \bar{P} \succ 0\,,$$

which leads to a Semi-Definite Program similar to (12), with the difference that

$$\mathcal{Q}_{11} := \begin{bmatrix} \bar{Q} & 0 & 0 & 0 \\ 0 & \tilde{Q} & 0 & 0 \\ 0 & 0 & \gamma^2 I & 0 \\ 0 & 0 & 0 & \eta I \end{bmatrix}, \mathcal{Q}_{22} := \begin{bmatrix} \bar{Q} & 0 & 0 & 0 \\ 0 & \tilde{Q} & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix},$$

$$\mathcal{Q}_{12} := \begin{bmatrix} \bar{Q}\bar{A} + Y\bar{C} & 0 & \bar{Q}\bar{B} + Y\bar{E} & 0 \\ 0 & \tilde{Q}A + XC & 0 & XD_n - \tilde{Q}B_n \\ -\mathbf{D}^\dagger C & 0 & \mathcal{D} & 0 \\ 0 & C & 0 & \mathbf{D}_n \end{bmatrix}.$$

This is now a *multi-objective* problem for which we determine the (γ, η) Pareto curve by computing optimal values $\gamma^*(\eta)$ for different fixed values of η. We demonstrate this evaluation through a numerical example in the next section.

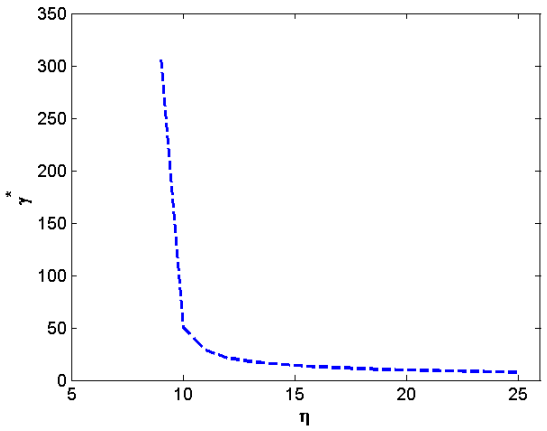

Fig. 2. Security versus Performance plot. Numerical results of the optimal value of problem (12) for different values of the parameter η evaluated on the system described in Section V.

V. Numerical Simulations

We now report the results of the proposed formulation from Section IV-A on synthetic problem. In particular, we chose the following values for the system matrices.

$$A = \begin{bmatrix} 1.1 & 0.2 \\ 0 & 0.9 \end{bmatrix}, B = \begin{bmatrix} 0.8147 & 0.1270 \\ 0.9058 & 0.9134 \end{bmatrix},$$
$$C = \begin{bmatrix} 0.6324 & 0.2785 \\ 0.0975 & 0.5469 \end{bmatrix}, D = \begin{bmatrix} 3.5784 & -1.3499 \\ 2.7694 & 3.0349 \end{bmatrix},$$
$$B_n = \begin{bmatrix} 0.9572 & 0.8003 \\ 0.4854 & 0.1419 \end{bmatrix}, D_n = \begin{bmatrix} -0.1241 & 1.4090 \\ 1.4897 & 1.4172 \end{bmatrix}.$$

and the following choice for the attack matrices

$$E = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Figure 2 summarizes how the optimal value γ^* of (12) varies with different choices for the parameter η.

We observe that below the value of $\eta < 9$, problem (12) is infeasible, and as η increases, the optimal value decreases as is expected. The decrease is very sharp for lower values of η, but the curve flattens out as η increases suggesting diminishing improvements for higher values of η. A curve such as this one may be used by a designer to study the tradeoff between security and performance of a given system.

VI. Conclusion and Future Directions

We considered the problem of stealth attack, i.e., a coordinated attack through several possible entry points on a closed loop linear time invariant dynamical system. This work extends previous results [3] to the case of noise in the system, and focuses on minimizing the consequences of a stealth attack on the system. We proposed a notion of the impact of such an attack on the system and considered a novel metric related to the H_∞ norm of the inverse of the system dynamics (assuming it exists) as a measure of the security of the system. We showed that the problem can be cast as a linear matrix inequality with the system parameters (observer gain) as the variable. This formulation allows a user to re-design of the system from the perspective of minimizing the impact of any stealth attack. Numerical results on simulated data illustrate the security (maximum impact of stealth attack) and performance (ℓ_2 gain of the original system) tradeoff using the proposed approach.

Future directions include the development of methods to minimize the gap between the convexified and the original optimization problem and optimal co-design of both the control and estimation gains for multiple objectives.

References

[1] Jose A Bernat Bacet. Inside the cunning, unprecedented hack of Ukraine's power grid. Available online at http://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/.

[2] R. B. Bobba, K. M. Rogers, Q. Wang, H. Khurana, K. Nahrstedt, and T. J. Overbye. Detecting false data injection attacks on DC state estimation. In *Preprints of the First Workshop on Secure Control Systems, CPSWEEK*, 2010.

[3] S. D. Bopardikar and A. Speranzon. On analysis and design of stealth-resilient control systems. In *6th International Symposium on Resilient Control Systems*, pages 48–53, 2013.

[4] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in Systems and Control Theory*. SIAM, 1994.

[5] J. Leyden. Stuxnet 'a game changer for malware defence', 2010. Available online at: http://www.theregister.co.uk/2010/10/09/stuxnet_enisa_response/.

[6] Y. Liu, M. K. Reiter, and P. Ning. False data injection attacks against state estimation in electric power grids. In *ACM conference on Computer and communications security*, 2009.

[7] Y. Mo and B. Sinopoli. Secure control against replay attacks. In *Forty-Seventh Annual Allerton Conference on Communication, Control, and Computing - Allerton House*, 2009.

[8] Y. Mo and B. Sinopoli. False data injection attacks in control systems. In *First Workshop on Secure Control Systems, CPS Week*, 2010.

[9] Y. Mo, S. Weerakkody, and B. Sinopoli. Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs. *IEEE Control Systems*, 35(1), 2015.

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

[11] F. Pasqualetti and Q. Zhu. Design and operation of secure cyber-physical systems. *IEEE Embedded Systems Letters*, 7(1), 2014.

[12] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson. Revealing stealthy attacks in control systems. In *50th Annual Allerton Conf. on Communication, Control and Comp.*, 2012.

[13] W. Zeng and M.-Y. Chow. Optimal tradeoff between performance and security in networked control systems based on coevolutionary algorithms. *IEEE Transaction on Industrial Electronics*, 59(7), 2012.

[14] B. Zheng, P. Deng, R. Anguluri, Q. Zhu, and F. Pasqualetti. Cross-layer codesign for secure cyber-physical systems. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2016. To Appear.