Joint controller and detector design against data injection attacks on actuators

Sribalaji C. Anand ∗ Andr´ e M. H. Teixeira ∗

∗ Department of Electrical Engineering, Uppsala University, PO Box 534, SE-75121, Uppsala, Sweden
(e-mail: {sribalaji.anand, andre.teixeira}@angstrom.uu.se)

Abstract: This paper addresses the issue of data injection attacks on actuators in control systems. Considering attacks that aim at maximizing impact while remaining undetected, the paper revisits the recently proposed output-to-output gain, which is compared to classical sensitivity metrics such as $H_\infty$ and $H_\infty$. In its original formulation, the output-to-output gain is unbounded for strictly proper systems. This limitation is further investigated and addressed by modifying the performance output of the system and ensuring that the system from attack signal to performance output is also strictly proper. With this system description, and by using the theory of dissipative systems, a Bi-linear Matrix Inequality (BMI) is formulated for system design. Using this BMI, a design algorithm is proposed based on the heuristic of alternating minimization. Through numerical simulations of the proposed algorithm, it is found that the output-to-output gain presents advantages over the other metrics: the effect of the attack is reduced in the performance output and increased in the detection output in a relatively large spectrum of frequencies.

Keywords: System security, Quadratic performance indices, Fault detection, $H_\infty$ control, Optimization.

1. INTRODUCTION

The trend towards increased usage of open-standard communication protocols among industrial control systems has made these systems vulnerable to online cyber-attacks such as Stuxnet (Langner, 2011). The issue of cyber-attacks has been addressed in detail for classical Information Technology (IT) systems (Bishop, 2002). In IT systems, cyber-security deals with properties such as confidentiality, integrity, and availability. Although these properties are essential for control systems, other key features such as stability and safe operation are not addressed. Hence the results from classical IT security cannot be directly extended to control systems.

Security of control systems has been studied in detail from different contexts such as (a) Modelling of various possible attacks, (b) Detection of attacks, (c) Quantifying the impact of attacks and (d) Prevention and treatment of attacks (Chong et al., 2019).

Possible attack scenarios such as eavesdropping attack, denial-of-service attack, replay attack, bias injection attack, zero dynamics attack are described in Cárdenas et al. (2011). A common thread in these scenarios is that adversaries are considered to be rational, with given objectives, resources, and constraints. Detection techniques of attacks was studied for data injection attacks (Teixeira et al., 2012), replay attacks (Mo et al., 2015) and routing attacks (Ferrari and Teixeira, 2017). The context of the attack undetectability was studied in Pasqualetti et al. (2015). The impact caused by the aforementioned attacks on control system was studied in Milošević et al. (2018b) and Urbina et al. (2016). Detectability and impact of attacks are key aspects in the security of control systems since they characterize the robustness/vulnerability of the control system against attacks. Given that the control system is under attack, attack treatment/mitigation through secure state estimation both in continuous-time (CT) and discrete-time (DT) has been studied in Fawzi et al. (2014).

Nonetheless, there are still significant gaps in the existing literature. First, most papers have focused on mitigating sensor attacks, while stealthy attacks on actuators have not been as much investigated (Ye and Luo, 2019). Second, most of the work combining detection and impact has focused on system analysis (Milošević et al., 2018a), and these approaches are not amenable to design controllers and detectors for increased security. Third, the joint design of controllers and detectors has received little attention, partly due to the decoupled nature of the sensitivity metrics used in the related literature (Tan and Patton, 2015), (Ding et al., 2002).

This paper addresses the above mentioned research gaps by investigating the joint design of controllers and detectors against stealthy attacks on actuators. The contribution of this article is as follows: Firstly, we look into the general DT control system representation and investigate the shortcoming faced by certain sensitivity metrics when applied to strictly proper systems (Teixeira, 2019). Secondly, we look into a different approach to address this limitation in the following way: when control systems are sampled from CT to DT, certain classes of systems end
up having algebraic loops which are not entirely causal, as supported by Blachuta (1999). Hence, we approach the sensitivity metrics with an altered system description (which respects causality) which in principle also addresses the aforementioned shortcomings of the sensitivity metric. Finally, we leverage the recent sensitivity metric, the output-to-output gain (OOG), and the proposed system description to cast the joint detector and controller design as an optimization problem with BMI constraints, which is tackled through an alternating minimization approach.

We conclude this section by providing the notations that are used in the paper. A formal problem background is provided in Section 2. Thereafter, a design problem based on sensitivity metrics is formulated in Section 3. The shortcoming faced by this design problem is discussed in Section 4. An altered system description is proposed to address this shortcoming. A design algorithm to the design problem based on this altered system description is presented. Section 5 provides numerical examples comparing the algorithm with the classical sensitivity metrics. Concluding remarks are provided in Section 6.

**Notation:**

Throughout this article, $\mathbb{R}, \mathbb{C}, \mathbb{Z}$ and $\mathbb{Z}^+$ represent the set of real numbers, complex numbers, integers and non-negative integers respectively. A positive (semi-)definite matrix $A$ is denoted by $A \succ 0, (A \succeq 0)$. The maximum and minimum singular values of a matrix $A$ is denoted by $\sigma(A)$ and $\sigma(A)$ respectively. The set of eigenvalues of a matrix $A$ is represented by $\lambda(A)$. Let $x : \mathbb{Z} \rightarrow \mathbb{R}^n$ be a discrete time signal with $x[k]$ as the value of the signal $x$ at the time step $k$. Let the time horizon be $[0, N] = \{k \in \mathbb{Z}^+ | 0 \leq k \leq N \}$. The $\ell_2$-norm of $x$ over the horizon $[0, N]$ is represented as $||x||_{\ell_2}^2 = \sum_{k=0}^{N} |x[k]|^2$. Let the set of square integrable signals be defined as $\ell_2 \triangleq \{ x : \mathbb{Z}^+ \rightarrow \mathbb{R}^n | ||x||_{\ell_2}^2 = \sum_{k=0}^{\infty} |x[k]|^2 < \infty \}$ and the extended signal space be defined as $\ell_2 \triangleq \{ x : \mathbb{Z}^+ \rightarrow \mathbb{R}^n | ||x||_{\ell_2, [0,N]}^2 = \sum_{k=0}^{N} |x[k]|^2 < \infty, \forall N \in \mathbb{Z}^+ \}$. $m \times n (m \times n)$ represents a matrix of size $m \times n$ where all the entries are zero (one).

## 2. PROBLEM BACKGROUND

In this section, we describe the control system structure and the goal of the stealthy adversary. Consider the general description of a closed-loop DT linear time-invariant system with a plant ($P$), output feedback controller ($C$) and anomaly detector ($D$). For the sake of simplicity, we assume a static output feedback controller (2). The closed-loop system is represented by

$$P : \begin{align*}
x_{p}[k + 1] & = A x_{p}[k] + B u[k] \\
y_{p}[k] & = C x_{p}[k]
\end{align*}$$

(1)

$$C : u[k] = L y[k]$$

(2)

$$D : \begin{align*}
\hat{x}_{p}[k + 1] & = A \hat{x}_{p}[k] + B u[k] + K y_{r}[k] \\
y_{r}[k] & = y[k] - C \hat{x}_{p}[k]
\end{align*}$$

(3)

where $x_{p}[k] \in \mathbb{R}^{n_z}$ is the state of the plant, $u[k] \in \mathbb{R}^{n_u}$ is the control signal applied to the actuator, $y[k] \in \mathbb{R}^{n_y}$ is the measurement output generated by the controller, $y_{r}[k] \in \mathbb{R}^{n_y}$ is the virtual performance output, $\hat{x}[k] \in \mathbb{R}^{n_z}$ is the state estimate produced by the observer based detector, $y_{r}[k] \in \mathbb{R}^{n_y}$ is the residue generated by the detector, $L$ and $K$ are the controller and detector gains respectively. In general, the system is considered to have a good performance when the energy of the performance output ($||y_{p}||_{\ell_2}^2$) is small and an anomaly is considered to be detected when the energy of the residue ($||y_{r}||_{\ell_2}^2$) is greater than a predefined threshold (say $\epsilon_r$). Without loss of generality, we assume $\epsilon_r = 1$ in the rest of this paper.

### 2.1 Data injection attack scenario

In the closed-loop system described above, we consider that an adversary is trying to inject false data into the actuator of the plant. Given this setup, we now discuss the resources the adversary has access to.

**Disruption resources:** The adversary can inject data on all the control channels. This is represented by:

$$\tilde{u}[k] \triangleq u[k] + a[k],$$

where $a[k] \in \mathbb{R}^{n_u}$ is the data injected by the adversary.

**Model knowledge:** The adversary has full system knowledge. This system knowledge is used by the adversary to calculate the optimal data injection attacks. Defining $e[k] \triangleq x_{p}[k] - \tilde{x}_{p}[k]$ and $\tilde{x}[k] \triangleq [x_{p}[k]^T e[k]^T]^T$, the closed-loop system under attack with the performance output and detection output as system outputs becomes:

$$\mathcal{P}_{cl} : \begin{align*}
x[k + 1] & = A x[k] + B_{cl} a[k] \\
y_{p}[k] & = C_{p} x[k] + D_{cl} a[k] \\
y_{r}[k] & = C_{r} x[k] + D_{r} a[k]
\end{align*}$$

(4)

where

$$A_{cl} \triangleq \begin{bmatrix} A + BLC & 0 \\ 0 & A - K C \end{bmatrix}, \quad B_{cl} \triangleq \begin{bmatrix} B \\ 0 \end{bmatrix}$$

$$C_{p} \triangleq [C_{p} + D_{cl} L C] 0, \quad D_{p} \triangleq D_{cl}$$

$$C_{r} \triangleq [0 C], \quad D_{r} \triangleq 0.$$

**Attack goals and constraints:** The adversary aims at deteriorating the system performance while remaining undetected. Hence, the adversary injects attack signals to maximize the energy of the performance output while keeping the energy of the detection output lower than $\epsilon_r$. This objective can be translated into an attack policy, which is formulated as the following optimization problem:

$$||\Sigma||_{\ell_2}^2 \triangleq \max_{a \in \ell_{2u}} ||y_{p}||_{\ell_2}^2$$

s.t. $||y_{r}||_{\ell_2}^2 \leq 1, \quad x[0] = 0$ (5)

where $||\Sigma||_{\ell_2}^2 \in \mathbb{R}$ represents the OOG.

### 2.2 Dissipative systems theory

The OOG resulting from the optimization problem (5) can be used for capturing the disruption induced by an attack signal. This optimization problem (5) is non-convex and can be reformulated to its convex dual counterpart as:

$$||\Sigma||_{\ell_2}^2 \triangleq \min_{\gamma \geq 0} \gamma$$

s.t. $||y_{p}||_{\ell_2}^2 \leq \gamma ||y_{r}||_{\ell_2}^2 \forall a \in \ell_{2u}.$ (6)
The strong duality between (5) and the above optimization problem was shown to be zero in Teixeira et al. (2015). One drawback of the above optimization problem is that, it operates in signal space which is infinite dimensional. It is advantageous, if the optimization problem is in the form of finite dimensional matrix inequalities. This transformation can be done by dissipative systems theory. For brevity, a detailed presentation of dissipative system theory is omitted (interested reader can refer to Teixeira (2019) and references therein). But one of the important propositions is presented below to keep the presentation self-contained.

**Definition 1.** A system $\Sigma \triangleq (A,B,C,D)$ is said to be dissipative with respect to the supply function $s(\cdot)$ if there exists a real valued function $V(\cdot)$ such that the following equation holds:

$$V(x[k_1]) - V(x[k_0]) \leq \sum_{k=k_0}^{k_1-1} s(x[k],u[k]), \quad \forall k_1 \geq k_0$$

**Remark 2.** Without any loss of generality, for this setting, the storage function can be taken of the form

$$V(x) = \frac{1}{2} x^T P x$$

where $P = P^T > 0$. The BMI approach for the optimization problem (10) exists in the literature (Hilhorst et al., 2014) and is given as:

$$A^2 = 0, \quad \exists P > 0, \quad \gamma > 0$$

$s.t. \quad R(P,L,K) + \sum_{i=1}^{N} (C_i^T P^T C_i P) - \gamma (C_i^T D_i^T C_i D_i) \leq 0$ (10)

**Proposition 3.** Consider a DT system $\Sigma$ which is controllable and observable. Let $y_i[k] = C_i x[k] + D_i u[k], \quad i = \{1,2\}, \quad s(x,u) = ||y_1[k]||^2 - ||y_2[k]||^2$. Then the following statements are equivalent:

1. The system $\Sigma$ is dissipative w.r.t. $s(x,u)$
2. For all trajectories of the system with $N > 0$ and $x[0] = 0$, it holds that $\sum_{k=0}^{N} s(x[k],u[k]) \geq 0$
3. There exists a $P > 0$ such that

$$A^2 P A - P A^T P B^2 B^T P A - B^2 P B \preceq 0 \quad \ldots \quad (7)$$

*Remark 4.* Define $G_1(z) = C_1(zI - A)^{-1} B + D_1$ and $G_2(z) = C_2(zI - A)^{-1} B + D_2$. Under these definitions, the necessary condition for dissipativity (cyclo-dissipativity) is given by the following inequality:

$$G_1(z)^2 G_1(z) - G_2(z)^2 G_2(z) \succeq 0, \quad \forall z \in \mathbb{C} \text{ with } z \notin \lambda(A), |z| = 1$$

### 2.3 Output-to-output gain

Let $s(\cdot) \triangleq ||y_r||^2 - ||y_p||^2$. Under this definition, using (7), the squared OOG $||\Sigma||^2_{z_{r,s},y_r,y_p}$ is the optimal value of (6), can be obtained by:

$$\min_{P \succeq 0, \gamma > 0} \gamma \quad \ldots \quad (8)$$

$s.t. \quad R(P,L,K) + \sum_{i=1}^{N} (C_i^T P^T C_i P) - \gamma (C_i^T D_i^T C_i D_i) \leq 0$

where $R(P) \triangleq \sum_{i=1}^{N} A_i^2 P A_i - P A_i^T P B_i^2 B_i^T P A_i - B_i^2 P B_i$. Here, the optimal value of the optimization problem denotes the squared-OOG, which represents the disruption induced by the attack vector (optimizer of (5)) on the system.

### 3. DESIGN BASED ON SENSITIVITY METRICS

Sensitivity metrics can be defined as metrics quantifying the influence of an anomaly (attack in this case) signal on the system. One of the sensitivity metrics, the OOG, which captures the worst-case effect of the anomaly in the detection and performance output, was introduced in the previous section. Currently, there are a few other approaches in the literature that deals with defining impact metrics for the system under attack in the following ways: the $H_\infty$ norm captures the worst-case (highest) effect of an anomaly in the performance output of the system, the $H_2$ index captures the worst-case (lowest) effect of an anomaly in the detection output of the system.

These sensitivity metrics can be used as an objective function to design the controller and/or detector of the system by minimizing the impact on performance and maximizing the detectability of the anomaly in the system outputs. This section is aimed at describing the sensitivity metric and the corresponding design problems.

#### 3.1 Output-to-Output gain (OOG) based design approach

The OOG can be used as an objective function to find a controller ($L$) and detector ($K$) such that the effect of the attack signal is increased in the detection output and reduced in the performance output simultaneously. The OOG optimal controller and detector is obtained by solving the optimization problem (9), where $||\Sigma||_{z_{r,s},y_r,y_p} = \gamma^*$ is the worst-case ratio of the effect of the attack signal on the performance output to the detection output.

$$\min_{P,L,K} \gamma^*$$

$s.t. \quad P \succeq 0, \gamma > 0$

$$R(P,L,K) + \sum_{i=1}^{N} (C_i^T P^T C_i P) - \gamma (C_i^T D_i^T C_i D_i) \leq 0$$

where $R(P,L,K) \triangleq \sum_{i=1}^{N} A_i^2 P A_i - P A_i^T P B_i^2 B_i^T P A_i - B_i^2 P B_i$ and the matrices $A_i, B_i, C_i$ are functions of $L$ and $K$.

*Remark 5.* The main advantage of using such a metric is that: from Remark 4, it is evident that the detector ($G_r(z)$) and the controller ($G_p(z)$) concentrate on the same frequency ($\omega$). This means that the attack detection and robustification occur at the same frequency, as opposed to the classical $H_\infty$ and $H_2$ methods. Hence the OOG design objective focuses on improving the detectability only when the impact of the attack signal on the performance output is sufficiently high at the same frequency. In other words, in frequency regions where the performance output is low, the OOG design problem (9) does not focus to improve the detectability. These observations are later illustrated through a numerical example in Section 5.

#### 3.2 Classical $H_\infty$ and $H_2$ metrics based design approach

The objective of the $H_\infty$ design approach is to find a feedback controller $L$, so as to reduce the effect of the disturbance/attack signal in the performance output. The classical $H_\infty$ design optimization problem is:

$$||\Sigma||_{H_\infty}^2 \triangleq \min_L \max_{a \in \mathbb{C}} ||y_p||^2_{z_r}$$

$s.t. \quad ||a||^2 = 1, \quad x[0] = 0 \quad \ldots \quad (10)$

where the optimal value of the $H_\infty$ norm is $||\Sigma||_{H_\infty}$. The BMI approach for the optimization problem (10) exists in the literature (Hillorst et al., 2014) and is given as:
The objective of the \( H_\infty \) design problem is to find an observer gain \( K \), so as to increase the effect of the disturbance/attack signal in the detection output. The classical \( H_\infty \) design optimization problem is:

\[
\min_{P, \beta, L} \beta \\
\text{s.t. } P \succeq 0, \; \beta > 0 \\
\begin{bmatrix}
P & P(A + BLC) P B T P \\
(A + BLC) T P & P & 0 & C_T^T \\
0 & \beta I & D_T^T \\
\end{bmatrix} \succeq 0
\]

where \( \beta^2 = \gamma \).

Theorem 6. The \( H_\infty \) design approach reduces the gain of the system \( \Sigma_p \) at a frequency \( z_1 \) where the attack impact on the performance output is the highest. The \( H_\infty \) design approach increases the gain of the system \( \Sigma_r \) at a frequency \( z_2 \) where the attack impact on the detection output is the lowest. The disadvantage of using these metrics for system design is that, the frequency at which the fault detection occurs might not necessarily be the same frequency at which robustification occurs (\( z_1 \neq z_2 \)).

4. JOINT DESIGN FOR STRICTLY PROPER SYSTEMS

Before introducing the design procedure for solving the non-convex BMIs in the previous section, there are certain limitations faced by these design metrics when applied to certain classes of systems. These limitations are discussed in this section.

Theorem 7. Consider a DT system \( \Sigma \triangleq (A, B, C, D) \) which is assumed to be controllable and observable. Let \( y_i[k] = C_i x[k] + D_i u[k] \), \( i = 1, 2, D_1 = 0, D_2 \neq 0 \) and \( s(\cdot) = \gamma \| y_1[k]\|^2 - \| y_2[k]\|^2, \gamma \geq 0 \). Under these definitions, the system \( \Sigma \) cannot be dissipative with respect to the supply function \( s(\cdot) \).

Proof. For the system dynamics defined in the statement of the theorem, using Proposition 3, the dissipation inequality (7) becomes:

\[
\begin{bmatrix}
A T P A - P & A T P B \\
B T P A & B T P B \\
\end{bmatrix} - \gamma \begin{bmatrix}
C T^T \\
C_1 \\
\end{bmatrix} \\
\begin{bmatrix}
A T P A & A T P B \\
B T P A & B T P B \\
\end{bmatrix} \preceq 0.
\]

For this inequality (12) to be satisfied, the following inequality should be satisfied as well (one of the elements of the above inequality): \( B T P B + D_T^2 D_2 \preceq 0 \). We know that, with the condition \( P \succeq 0 \), this inequality is impossible to solve (since \( D_T^2 D_2 \succeq 0 \)). Hence the LMI (12) is infeasible, which concludes the proof.

Remark 8. For \( H_\infty \) control, in accordance with the definitions of Theorem 7, we have \( y_1 = y_r \) for which the system matrix \( D_1 \) becomes 0 for any system of the form (4). This satisfies the assumption made in the theorem above (\( D_1 = 0 \)). Hence, Theorem 7 imposes a limitation on the \( H_\infty \) control/design approach.

Remark 9. For \( H_\infty \) control, in accordance with the definitions of Theorem 7, we have \( y_1 = y_r \) for which the system matrix \( D_1 \) becomes 0 for any system of the form (4). This does not satisfy the assumption made in the theorem above (\( D_1 = 0 \)). Hence, Theorem 7 does not impose any limitations on the \( H_\infty \) control/design approach.

4.1 Revised system description

This limitation described by Theorem 7 was also a result pointed out in a previous work (see Teixeira (2019)) where an approach to circumvent this issue was proposed based on cyclo-dissipativity which unfortunately leads to a complex design problem. In this section, we will concentrate on an alternative approach. Let us momentarily consider a CT plant described by:

\[
\begin{align*}
\dot{x}(t) &= \text{Az}(t) + \text{Bu}(t) \\
y(t) &= \text{Cx}(t) + \text{Du}(t). \\
\end{align*}
\]

The DT version of the plant is obtained by a zero sample hold (ZOH). The ZOH method can be represented by:

\[
y(t = kh) = C \sum_{(k-1)h}^{kh} \dot{x}(t) dt + \text{Du}(t \in [(k-1)h, kh])
\]

where \( k \) is the discrete time step and \( h \) represents the sample time.

This effect, mentioned in (13), is not captured in the system description (1)-(3). Hence an system description with the following variation which respects causality (Blachuta, 1999) is adopted.

\[
y[k] = C x[k] + D u[k - 1] \\
yr[k] = C x[k] + D r[k - 1]
\]

Defining \( e[k] \triangleq x[k] - x_r[k] \) and \( \dot{x}[k] \triangleq [x_p[k] T, e[k] T] \), the closed-loop system under attack with the performance output and detection output as system outputs becomes:

\[
\begin{align*}
P_d : & \begin{cases}
\dot{x}[k+1] = Ax[k] + Bu[k] \\
y_p[k] = C x[k] \\
yr[k] = C x_r[k].
\end{cases}
\end{align*}
\]

where

\[
A = \begin{bmatrix}
A + BLC & 0 & BLD \\
0 & -KC & -KD \\
LC & 0 & LD
\end{bmatrix}, \\
B = \begin{bmatrix}
B \\
0 \\
I
\end{bmatrix}
\]

\[
C_p = [C_0 \ C_1 \ D], \\
C_r = [0 \ C \ D \ D].
\]

Henceforth in this article, the altered system description (14) will be used unless stated otherwise.

For the altered system description (14), the results discussed in Section 3 still holds which is summarized by the following lemma.
Lemma 10. The OOG for any given system of the form (14) is the solution of the optimization problem:

\[ ||\Sigma||_{2x,\gamma}^2 = \min_{P = P^T \geq 0, \gamma} \gamma \]

s.t. \( R(P) \leq 0 \),

where

\[
R(P) = \begin{bmatrix} \bar{A}^T \bar{P} \bar{A} - P \bar{A}^T \bar{P} \bar{B} & \gamma \bar{C}_r^T \bar{C}_r - \bar{C}_p^T \bar{C}_p \end{bmatrix} \leq 0.
\]

4.2 Joint controller and detector design algorithm

In the previous section, an altered system description was proposed so that the OOG is bounded for any system of the form (4). With this altered system description, we aim to design an OOG optimal controller and detector in this section. This design algorithm is discussed below.

Theorem 11. Consider the closed-loop system described in (14). Let \( D = 0 \). Define \( P \equiv \begin{bmatrix} P_{x} & P_{zu} \end{bmatrix} \) where \( P_x \in \mathbb{R}^{(2n_x+n_u) \times (2n_x+n_u)} \), \( P_{zu} \in \mathbb{R}^{(2n_x+n_u) \times n_z} \), \( \bar{A} \equiv \begin{bmatrix} \bar{A} + BLC & 0 & 0 \\ 0 & A - K \bar{C} & 0 \\ \bar{C}_r & \bar{C}_p & -\bar{K} \end{bmatrix} \), \( \bar{C}_r \equiv \begin{bmatrix} C & 0 & D \end{bmatrix} \)

and \( \bar{C}_r \equiv \begin{bmatrix} 0 & C & 0 \end{bmatrix} \). The OOG optimal controller and detector for the system (14) are obtained by solving the following optimization problem:

\[
\min_{P = P^T, L, K, \gamma} \gamma \\
\text{s.t.} \quad P_x > 0 \\
P \bar{B} = 0
\]

\[
\begin{bmatrix} -P_x & P_x \bar{A} & 0 \\ \bar{A}^T P_x & -P_x - P_{zu} & 0 \\ 0 & -P_{zu}^T & -P_u \end{bmatrix} + \bar{Q}(\gamma) \leq 0,
\]

where \( \bar{Q}(\gamma) \equiv \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\gamma \bar{C}_r^T \bar{C}_r + \bar{C}_p^T \bar{C}_p & \bar{C}_r^T \bar{D}_J \end{bmatrix} \). The optimal solution variables \( L^* \) and \( K^* \) represent the optimal control and observer gains, where as the optimal value \( \gamma^* \) corresponds to the optimal OOG.

Proof. The dissipation inequality for any system of the form (14) can be written as the BMI \( \bar{R} \leq 0 \). For the above BMI to be satisfied, the necessary condition must be satisfied as well: \( \bar{B}^T \bar{P} \bar{B} \leq 0 \). Recalling the constraint \( P_x \geq 0 \), as required for dissipativity, the former necessary condition can be rewritten as \( PB = 0 \).

Under the constraint \( PB = 0 \), the BMI \( \bar{R} \leq 0 \) reduces to

\[
\bar{A}^T \bar{P} \bar{A} - P + \bar{C}_r^T \bar{C}_r - \gamma \bar{C}_p^T \bar{C}_p \leq 0
\]

Under the definitions \( D = 0, \bar{A} \equiv \begin{bmatrix} \bar{A} & 0 & 0 \\ 0 & 0 \end{bmatrix}, P = \begin{bmatrix} P_x & P_{zu} \end{bmatrix} \)

and by using the Schur lemma (Boyd and Vandenberghe, 2004), (16) becomes the last constraint of (15), which concludes the proof.

This optimization problem (15) is non-convex due to the BMI in \( K, L \) and \( P_z \). Hence, an alternating minimization approach is proposed in Algorithm 1 (Li et al., 2019).

**Algorithm 1**

**Joint design of controller and detector**

The advantages of designing the controller and detector with Algorithm 1 are: (a) as a direct consequence of a unified objective function, a joint optimal design of \( K \) and \( L \) is possible, (b) imposing the condition \( P_z > 0 \) guarantees that the system is stable, which is an added advantage (c) if the objective only concerns with the design of controller (or detector), this algorithm still applies but the steps (iii) and (iv) of Algorithm 1 is an optimization problem w.r.t. \( L \) (or \( K \) only).

5. Numerical Example

In this section, the effectiveness of the proposed Algorithm 1 is depicted through a numerical example.

**Example 12**. Consider the system described in (14) with

\[
A = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 0.5 & 0 \\ 0 & 1 & -0.5 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad C = 1_{1 \times 3}, \quad C^T = \begin{bmatrix} I_3 \\ 0_{3 \times 1} \end{bmatrix},
\]

and \( D_J = [0_{3 \times 1}]^T \). The sub-optimal gains obtained by Algorithm 1 for the altered system description (14) are \( L^* = -0.6208 \) and \( K^* = [1.2058 \ 0.5283 \ 0.3440] \). For comparison, the classical design problems are solved as follows:

- The \( H_\infty \) optimal controller for the altered system description (14) is traditionally computed by solving the BMI (Hilhorst et al., 2014). Since this problem is non-convex, Algorithm 1 is employed to solve this BMI. The only difference is that the optimization problem is now only a function of \( L, \gamma \) and \( P \).
- When designing the detector using the \( H_\infty \) method for a strictly proper system as in our case, the limitation mentioned in Remark 8 occurs. This implies that an optimal detector cannot be designed. Hence a sub-optimal detector design for a finite frequency range is adopted (see Theorem 1, Wang and Yang (2008)).

A frequency range in which a solution exist for this system \( (\omega \in [1, 50]) \) is chosen.

The design gains obtained from these methods are \( L = -0.5787 \) and \( K^T = [1.6997 \ 0.6185 \ 0.3083] \). The singular values of the altered system are shown in (a) Fig. 1 for performance output (\( \sigma(\Sigma_p) \)) and (b) Fig. 2 for detection output (\( \sigma(\Sigma_r) \)). Fig. 1 and Fig. 2 represent the singular values of the system on the unit circle of the complex plane i.e., the x-axis of the figures represent \( z = e^{j\omega T_a} \) with \( \omega \in [0, 1] \). The dark line at \( \pi \) rad/sec represents the Nyquist sampling frequency.

5.1 Discussion

As mentioned before, the objective of the \( H_\infty \) design is to minimize the effect of the attack on the performance.
is shown in Fig. 4. It can be seen that the detection and the performance energy is higher than the classical metrics. Hence, this method can help to improve attack detection, at the expense of an increased deterioration of the performance. However, in practice, the performance deterioration may be prevented by timely switching to a fault-tolerant controller (Gao, 2015) when the attack is detected.

The worst-case attack for the detection output is at $\omega = 1.42 \text{ rad/sec}$. Applying this worst-case input signal $a[k] = \sqrt{2} \sin(1.42T_xk)$, the performance of the system is shown in Fig. 5. Although the detection energy is lower than the classical metrics, it can be noted that the detection energy at time intervals $[0, 5]$s is very close to that of the classical metric which can help in detection. Moreover, although the detection is harder in this case, the performance degradation is also less significant when compared to the other attack signals examined in this section. By shifting the focus from attacks with small impact on performance, the detector has managed to improve its detection capabilities for other attacks with higher impact.

With the $H_\infty$ and $H_2$ optimal controller and detector obtained from Algorithm 1, we obtain $\gamma = 238.4265$ by solving (8). With the OOG optimal controller and detector parameters from Algorithm 1, we obtain $\gamma = 74.7920$ by solving (8). This is an indication of the bound imposed by the choice of the design metric chosen on the disruption induced by the attack signal.

6. CONCLUSION

This paper considers actuator attacks that aim at maximizing impact while remaining undetected, and proposes a joint detector and controller design approach based on dissipative systems theory and alternating minimization. Numerical examples are provided to compare the effectiveness of the proposed approach. Future work directions include the extension to other classes of attacks, as well as to other controller structures.


Teixeira, A.M. (2019). Optimal stealthy attacks on actuators for strictly proper systems. In 2019 IEEE 58th Conf. on Decision and Control (CDC), 4385–4390. IEEE.

