Risk Assessment of Stealthy Attacks on Uncertain Control Systems

Sribalaji C. Anand, Student Member, IEEE, André M. H. Teixeira, Member, IEEE, and Anders Ahlén, Senior Member, IEEE

Abstract—In this article, we address the problem of risk assessment of stealthy attacks on uncertain control systems. Considering the data injection attacks that aim at maximizing the impact while remaining undetected, we use the recently proposed output-to-output gain to characterize the risk associated with the impact of attacks under a limited system knowledge attacker. The risk is formulated using a well-established risk metric, namely the maximum expected loss. Under this setup, the risk assessment problem corresponds to an untractable infinite nonconvex optimization problem. To address this limitation, we adopt the framework of scenario-based optimization to approximate the infinite nonconvex optimization problem by a sampled nonconvex optimization problem. Then, based on the framework of dissipative system theory and S-procedure, the sampled nonconvex risk assessment problem is formulated as an equivalent convex semidefinite program. Additionally, we derive the necessary and sufficient conditions for the risk to be bounded. Finally, we illustrate the results through numerical simulation of a hydro-turbine power system.

Index Terms—Optimization, risk analysis, security, uncertainty.

I. INTRODUCTION

Research in the security of industrial control systems has received considerable attention [1] due to an increased number of cyber-attacks such as the one on the Ukrainian power grid [2], Kemuri water company [3] among others. One of the common recommendations for improving the security of control systems is to follow the risk management cycle: risk assessment, risk response, and risk monitoring [4]. This article focuses on risk assessment, the formal definition of which will be introduced later as a function of the attack impact.

Risk is often a combination of attack impact and/or likelihood. For instance, the risk is characterized in terms of average impact in [5] for different types of attacks. The consequences of data injection attacks were quantified using the conditional value-at-risk in [6]. The calculated risk can later be used to compute optimal defense-allocation strategies [7] and/or design robust controllers/detectors [8]. Risk assessment of combined data integrity and availability attacks against the power system state estimation is conducted in [9]. From this brief discussion, it can be realized that characterizing risk in terms of attack impact and likelihood is critical for the efficient allocation of protection resources. In the literature, the problem of risk assessment of stealthy attacks on uncertain control systems has not been addressed. To the best of the authors’ knowledge, the works that are closely related to this problem are [10], [11], and [12].

First, the authors in [10] designed a stealthy attack against an uncertain system using disclosure resources. Next, the authors in [11] focused on attack detection based on plant and model mismatch for the adversary. The results of both the above works could not facilitate the optimal allocation of protection resources.

Furthermore, the authors in [12] proposed an impact metric by computing a bound on the reachable set of states by an attacker for perturbed system dynamics. They also proposed a second metric by computing the distance between the reachable set of states for the adversary and the set of critical states. However, they considered a deterministic system.

The advantage of our study is multifold. First, we consider a generic modeling framework similar to [13]. Then, unlike many previous works, a system description with parametric uncertainty is considered. Next, similar to [7], we adopt an adversarial setup where the system knowledge of the adversary is limited. Finally, we consider a recently proposed impact metric: output-to-output gain (OOG) [14]. The main advantage of using this impact metric, as opposed to the classical $H_{\infty}$ and $H_2$ metrics, is that the OOG metric-based design problem focuses on improving detectability only when the impact of the attack is sufficiently high at the same frequency [15]. In other words, the OOG metric is more amenable to risk-optimal system design for increased security. Under the described setup, we present the following contributions.

1) Using OOG as an impact metric, and the maximum expected loss as a risk metric, we formulate the risk assessment problem. We observe that the risk assessment problem corresponds to an untractable infinite nonconvex robust optimization problem, which is NP-hard in general.

2) We propose a convex semidefinite program (SDP) that solves the risk assessment problem approximately by sampling the uncertainty set. Additionally, we provide probabilistic guarantees on the feasibility of the original robust optimization problem.

3) We derive the necessary and sufficient conditions for the risk metric to be bounded.

4) Our approach results in the risk of an open-loop attack, which is robust against uncertainties.

The rest of this article is organized as follows. The uncertain system and the adversary are described in Section II. The problem is formulated in Section III, to which an approximate solution is presented in Section IV. The efficacy of the proposed optimization problem is illustrated through a numerical example in Section V. Finally, Section VI concludes this article.

Notation: Throughout this article, $\mathbb{R}$, $\mathbb{R}^+$, $\mathbb{C}$, $\mathbb{Z}$, and $\mathbb{Z}^+$ represent the set of real numbers, positive real numbers, complex numbers, integers, and nonnegative integers, respectively. A positive semidefinite matrix $A$ is denoted by $A \succ 0 (A \succeq 0)$.

Let $\pi : \mathbb{Z} \to \mathbb{R}^+$ 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 $\|x\|_2 = \sqrt{\sum_{k=0}^{N} x[k]^2}$. The extended signal space is defined as $\mathbb{Z}^+ \to \mathbb{R}^+ \big| \|x\|_2, [0,N] < \infty \}$ and the extended signal space be defined

Manuscript received 31 March 2022; revised 15 August 2023; accepted 18 September 2023. Date of publication 22 September 2023; date of current version 26 April 2024. This work was supported in part by the Swedish Research Council under Grant 2018-04396 and in part by the Swedish Foundation for Strategic Research. Recommended by Associate Editor T. Yuvelen. (Corresponding author: Sribalaji C. Anand.) Sribalaji C. Anand and Anders Ahlén are with the Department of Electrical Engineering, Uppsala University, SE-75103 Uppsala, Sweden (e-mail: sribalaji.anand@angstrom.uu.se; anders.ahlen@angstrom.uu.se).

André M. H. Teixeira is with the Department of Information Technology, Uppsala University, SE-75105 Uppsala, Sweden (e-mail: andre.teixeira@it.uu.se).

Digital Object Identifier 10.1109/TAC.2023.3318194
In this section, we describe the control system structure and the goal of the adversary. Consider (P), a general description of a closed-loop DT LTI system with a process (P), feedback controller (C), and an anomaly detector (D) as shown in Fig. 1 and represented by

\[
\begin{align*}
\mathcal{P} & : \begin{cases} x_{p}[k+1] = A^\Delta x_{p}[k] + D^\Delta \delta \tilde{u}[k] \\ y_{p}[k] = C_j x_{p}[k] + D_j \tilde{u}[k] \end{cases} \\
\mathcal{C} & : \begin{cases} z[k+1] = A_z z[k] + B_z \tilde{y}[k] \\ \bar{u}[k] = C_z z[k] + D_z \tilde{y}[k] \end{cases} \\
\mathcal{D} & : \begin{cases} s[k+1] = A_s s[k] + B_s \bar{u}[k] + K_s \tilde{y}[k] \\ \bar{y}_c[k] = C_c s[k] + D_c \bar{u}[k] + E_c \tilde{y}[k] \end{cases}
\end{align*}
\]

where \(A^\Delta = A + \Delta A(\delta)\) with \(A\) representing the nominal system matrix and \(\delta \in \Omega\). Additionally, we assume \(\Omega\) to be closed, bounded, and to include the zero uncertainty yielding \(\Delta A(0) = 0\). The other matrices are similarly expressed. The state of the process is represented by \(x_{p}[k] \in \mathbb{R}^{n_z}\), \(z[k] \in \mathbb{R}^{n_z}\) is the state of the controller, \(s[k] \in \mathbb{R}^{n_s}\) is the state of the observer, \(u[k] \in \mathbb{R}^{n_u}\) is the control signal generated by the process, \(\tilde{u}[k] \in \mathbb{R}^{n_u}\) is the measurement output produced by the process, \(\tilde{y}[k] \in \mathbb{R}^{n_y}\) is the measurement signal received by the controller and the detector, \(y_{c}[k] \in \mathbb{R}^{n_y}\) is the virtual performance output, and \(\bar{y}_c[k] \in \mathbb{R}^{n_y}\) is the residue generated by the detector. In general, the system is considered to have a good performance when the energy of the performance output \(||y_{c}[k]||_2^2\) is small and an anomaly is considered to be detected when the detector output energy \(||\bar{y}_c[k]||_2^2\) is greater than a predefined threshold, say \(\epsilon_\alpha\). Without loss of generality (w.l.o.g.), we assume \(\epsilon_\alpha = 1\) in the sequel.

\[1\] In this article, the Cartesian product is considered over the same probability space \((\Delta)\). But this can be generalized to arbitrary probability spaces.

A. Data Injection Attack Scenario

In the system described in (1)–(3), we consider an adversary injecting false data into a combination of the sensors and actuators. Next, we discuss the resources the adversary has access to.

1) Disruption and Disclosure Resources: The adversary can access (eavesdrop) the control or sensor channels and can inject data. This is represented by

\[
\begin{align*}
\tilde{u}[k] & = u[k] + B_a a[k], \\
\tilde{y}[k] & = y[k] + B_a a[k],
\end{align*}
\]

where \(a[k] \in \mathbb{R}^{n_a}\) is the attack signal injected by the adversary. The matrix \(E_a(F_a)\) is a diagonal matrix with \(E_a(i,i) = 1\) if the actuator (sensor) channel \(i\) is under attack and 0 otherwise.

2) System Knowledge: In general, the system operator knows about the parameters of the controller and detector as she/he is the one who designs it. We assume that the adversary can obtain these parameters, but does not know the true parameters of the process. To this end, we consider a realistic adversary whose system knowledge is limited. We defined this adversary as a rational adversary.

Definition II.1 (Rational adversary): An adversary is defined to be rational if it knows the matrices of (1) with bounded uncertainty.

Defining \(x[k] = [x_{p}[k]\ z[k]\ s[k]\ ^T]_P\), the closed-loop system under attack with the performance output and detection output as system outputs becomes

\[
x[k+1] = A^\Delta x[k] + B^\Delta \delta \tilde{a}[k] \\
y_{p}[k] = C^\Delta x[k] + D^\Delta \tilde{a}[k] \\
y_{c}[k] = C^\Delta x[k] + D^\Delta \tilde{a}[k]
\]

where the nominal matrices are given by

\[
\begin{align*}
(A + BD_eC) & = A_c \\
(BC_e + BD_eD_a) & = 0 \\
(B(D_e + K_e)C_e + B_eD_e) & = 0 \\
C_e & = D_e + K_eD_e \\
\Delta A_{cl} & = \begin{cases} DA + BD_e \Delta C + \Delta B(D_eC + D_e \Delta C) & \text{if } A_e \neq 0 \\
0 & \text{otherwise}
\end{cases} \\
\Delta B_{cl} & = \begin{cases} \Delta BB_e + \Delta BD_e D_a & \text{if } A_e \neq 0 \\
0 & \text{otherwise}
\end{cases} \\
\Delta C_{cl} & = \begin{cases} (D_eD_e + E_e)C_e & \text{if } A_e \neq 0 \\
0 & \text{otherwise}
\end{cases}
\end{align*}
\]

and the uncertain matrices are given as

\[
\begin{align*}
\Delta A_{cl} & = \begin{cases} DA + BD_e \Delta C + \Delta B(D_eC + D_e \Delta C) & \text{if } A_e \neq 0 \\
0 & \text{otherwise}
\end{cases} \\
\Delta B_{cl} & = \begin{cases} \Delta BB_e + \Delta BD_e D_a & \text{if } A_e \neq 0 \\
0 & \text{otherwise}
\end{cases} \\
\Delta C_{cl} & = \begin{cases} (D_eD_e + E_e)C_e & \text{if } A_e \neq 0 \\
0 & \text{otherwise}
\end{cases}
\end{align*}
\]

Here, the matrices \(B_a\) and \(D_a\) are chosen by the operator. If all the diagonal elements of \(B_a\) or \(D_a\) are 0, the operator believes that an adversary might attack all (none of the) data channels. Next, we assume the following for clarity.

Assumption II.1: The closed-loop system (4) is stable \(\forall \delta \in \Omega\).

Assumption II.2: The input matrix has full column rank, i.e., \(\hat{s} \in \mathbb{R}^{n_u} \neq 0\) such that \(B^\Delta \hat{s} = 0\).<ref>

3) Attack Goals and Constraints: Given the resources the adversary has access to, the adversary aims at disrupting the system’s behavior while staying stealthy. The system disruption is evaluated by the increase in energy of the performance output. Moreover, similar to the fault detection approaches [16], the adversary is stealthy if the energy of the detection output is below a predefined threshold \(\epsilon_\alpha = 1\). Thus, we define a stealthy attack as follows.

Definition II.2 (Stealthy attack): An attack vector \(a\) is said to be stealthy if it satisfies \(||\tilde{a}[k]||_2^2 \leq 1\).

We discuss the attack policy for a deterministic system next.
B. Optimal Attack Policy for the Nominal System

From the previous discussions, it can be understood that the goal of the adversary is to maximize the performance cost while staying undetected. With this setup, we characterize the attack policy of the adversary next before which we introduce the following assumptions.

Assumption II.3: The closed-loop system (4) is at equilibrium \( x[0] = 0 \) before the attack commences.

Similar to [17] and [18], we consider that the adversary has finite energy and thus attacks the system for a long but finite time, say \( T \).

Although this time \( T \) is unknown \textit{a priori}, it can be ensured that

\[
\alpha[\infty] = \lim_{k \to \infty} a[k] = 0
\]

is computationally intensive or in general NP-hard \([\ref{17}]\). The maximum expected loss associated with the impact-random variable \( X^A(\cdot) \), defined in (6), is defined as

\[
\text{MEL}[X^A] = \sup_{a \in \mathcal{A}_2} E_{\Omega} \{ X^A(a, \delta) \}
\]

where \( X^A(\delta, a) \) is the loss on scenario \( \delta \) and \( E_{\Omega} \) represents the expectation operator over the set \( \Omega \).

Thus, by determining the attack vector that solves for maximal expected loss, one can ensure that the attack vector is stealthy with respect to all uncertainties whilst maximizing the performance loss. Using Definition III.2, the risk associated with the impact caused by a bounded-Rational Adversary can be characterized as

\[
\gamma_{RA} = \sup_{a \in \mathcal{A}_2} E_{\Omega} \{ X^A(a, \delta) \}. \tag{7}
\]

Since the operator \( E \) in (7) operate over the continuous space \( \Omega \), (7) is computationally intensive or in general NP-hard \([\ref{21}, \text{Sec. 3}]\).

Besides, the problem is also nonconvex. In the rest of this article, we discuss methods to solve the optimization problem approximately and efficiently.

IV. RISK ASSESSMENT FOR A BOUNDED-RATIONAL ADVERSARY

In this section, we focus on describing a scenario-based approach to the optimization problem (7).

A. Approximating the Uncertainty Set

To recall, we are interested in determining the maximum expected loss associated with the impact caused by a rational adversary. Unfortunately, this problem is computationally intensive or in general NP-hard. Thus, as a first step toward solving (7), we approximate the objective function in Lemma IV.1.

Lemma IV.1: Let \( \delta_i \) be sampled uncertainties from \( \Omega \). Let us define

\[
\hat{E}_{\Omega}\sum_{i=1}^{N_2} X^A(a_i, \delta_i)
\]

Then, it holds that \( \lim_{N_2 \to \infty} \hat{E}_{\Omega}\sum_{i=1}^{N_2} X^A(a_i, \delta_i) = E_{\Omega}(X^A(\delta, a)) \).

Proof: The proof follows from applying [\ref{22}, Th. 7.2] to approximate the expectation operator in (7).

Lemma IV.1 states that the continuous set \( \Omega \) can be approximated by a discrete set \( \Omega_{N_2} \) of cardinality \( N_2 \). The approximation becomes more accurate as \( N_2 \to \infty \). For most cases, this approximation introduces a curse of dimensionality to obtain a good estimate of the risk and obtain a well-feasible attack vector. To circumvent this practical issue, we next show that an attack vector obtained by solving (7) with a discrete uncertainty set as mentioned in Lemma IV.1 is partially feasible to the original optimization problem (7) with a continuous uncertainty set.

\[2\]In general, the knowledge of the probability distribution of \( \Omega \) is necessary to determine the expectation. However, as we show in Theorem IV.3, the results of this article are independent of the distribution of \( \Omega \).
It might not be immediately apparent that the notion of feasibility applies to (7) since there are no external constraints present. Thus, we begin by simplifying the optimization problem (7).

**Lemma IV.2:** The optimization problem (7) is equivalent to (8)

$$\sup_{a \in \mathcal{E}\omega} \mathbb{E}\left\{ \left\| y_\gamma \delta \right\|_2^2 \left( \left\| y_\gamma \delta \right\|_2^2 \leq 1 \right) \right\} \left(1 - \beta \right)$$

subject to

$$\mathbb{P}_\Omega \left( \left\| y_\gamma \delta \right\|_2^2 \leq 1 \right) \geq 1 - \beta. \quad \tag{8}$$

**Proof:** Consider the function $X^A(a, \delta)$ in (6). By expanding its indicator function, we can write $\mathbb{E}_\mathbb{X}(X^A(a, \delta))$ as

$$\mathbb{E}_\Omega \left\{ \left\| y_\gamma \delta \right\|_2^2 \left( \left\| y_\gamma \delta \right\|_2^2 \leq 1 \right) \right\} \mathbb{P}_\Omega \left( \left\| y_\gamma \delta \right\|_2^2 \leq 1 \right)$$

Using the above definition in (7), we obtain $\gamma_{RA}$ as

$$\sup_{a \in \mathcal{E}\omega} \mathbb{E}_\Omega \left\{ \left\| y_\gamma \delta \right\|_2^2 \left( \left\| y_\gamma \delta \right\|_2^2 \leq 1 \right) \right\} \mathbb{P}_\Omega \left( \left\| y_\gamma \delta \right\|_2^2 \leq 1 \right)$$

which can be rewritten as (8). This concludes the proof.

**Lemma IV.2** uncovers the constraints present in the optimization problem (7). We can now discuss the notion of feasibility in regard to the optimization problem (8). So, we continue by simplifying (8) as follows. In reality, $\beta$ represents the fraction of the uncertainty set with respect to which the adversary is not stealthy. Let us consider an adversary that wishes to be stealthy w.r.t. all bounded uncertainties. Thus, we could set $\beta = 0$. Motivated by the above argument, we rewrite (8) as

$$\sup_{a \in \mathcal{E}\omega} \mathbb{E}_\Omega \left\{ \left\| y_\gamma \delta \right\|_2^2 \left( \left\| y_\gamma \delta \right\|_2^2 \leq 1 \right) \right\}$$

subject to

$$\left\| y_\gamma \delta \right\|_2^2 \leq 1 \quad \forall \delta \in \Omega. \quad \tag{9}$$

Recalling the approximation result in Lemma IV.1, assume that $\Omega$ is replaced with a discrete uncertainty set $\Omega_N$, so that (9) is approximated by (10) whose value is denoted by $\gamma_{RA}(N_1)$

$$\sup_{a \in \mathcal{E}\omega} \left\{ \mathbb{E}_{\Omega N_1} \left\| y_\gamma \delta \right\|_2^2 \right\} \left( \left\| y_\gamma \delta \right\|_2^2 \leq 1 \right) \forall \delta \in \Omega N_1. \quad \tag{10}$$

Let the resulting optimal attack vector be denoted by $a_{\gamma(N_1)}$. Then, the following theorem provides a posteriori results on the feasibility of the attack vector $a_{\gamma(N_1)}$.

**Theorem IV.3:** Let the number of samples $N_2$ and the confidence level $\lambda \in (0, 1)$ be predefined constants. Define $e(\cdot)$ such that

$$e(N_2) = 1 - N_2 \sum_{k=0}^{N_2-1} \left( \begin{array}{c} N_2 \\ k \end{array} \right) (1 - e(k))^{N_2-k} = \lambda. \quad \tag{11}$$

Let $s_{\gamma(N_2)}^a$ represent the cardinality of the support subsample for $(\delta_1, \ldots, \delta_{N_2})$ (see [23, Definition 2]). Then, it holds that

$$\mathbb{P}_{\Omega N_2} \left\{ \mathbb{P}_\Omega \left[ \delta \in \Omega \left| a_{\gamma(N_2)} \notin \Theta \right. \right] > e(s_{\gamma(N_2)}) \right\} \leq \lambda$$

where $a_{\gamma(N_2)}$ is the argument of the optimization problem (10) and $\Theta$ is defined as

$$\Theta \triangleq \bigcap_{\delta \in \Omega} \Theta_\delta, \text{ where } \Theta_\delta \triangleq \left\{ a \in \mathcal{E}\omega \left| \left\| y_\gamma \delta \right\|_2^2 \leq 1 \right) \quad \forall \delta \in \Omega \right\}. \quad \tag{12}$$

**Proof:** See Appendix A.

In words, Theorem IV.3 states that the attack vector obtained by solving (10) is stealthy and state-vanishing for all the closed-loop system of the form (4) with uncertainties belonging to the set $\Omega$. Moreover, it also states that the $e(\cdot)$ and $\lambda$ are independent of the distribution of $\Omega$. This result is a direct consequence of [23, Th. 1].

In conclusion, it follows that (9) can be solved approximately with a discrete set $\Omega_{N_2}$ of arbitrary but bounded cardinality. Thus, the next section will focus on solving the optimization problem (10).

**B. Risk Assessment**

The optimization problem (10) has the following two main disadvantages: 1) it is a nonconvex optimization problem and 2) it constraints lie in the infinite-dimensional attack space. To resolve these disadvantages, we adopt the S-procedure and dissipative system theory and revisit the optimization problem (10) in the theorem below.

Let us consider the cases where the optimization problem is of the form (10) and we can rewrite as

$$\begin{align*}
\min_{\gamma, p, r, p, \mathbb{P}} & \quad \mathbb{I}^T \begin{bmatrix} \gamma_1 & \ldots & \gamma N_2 \end{bmatrix}^T \\
\text{s.t.} & \quad \begin{bmatrix} A^T P A - P & A^T P B \\ B^T P A & B^T P B \end{bmatrix} + \Psi(\gamma) \leq 0
\end{align*} \quad \tag{13}$$

where $\Psi(\gamma) \triangleq \frac{1}{2} \begin{bmatrix} C_p^T D_p^T & C_p^T D_p^T \end{bmatrix} \Gamma(\gamma) \begin{bmatrix} C_p & D_p \end{bmatrix}$

$$\begin{bmatrix} A & B \\ C_p & D_p \end{bmatrix} = \begin{bmatrix} A_{d,1} & \ldots & 0 \\ A_{d,2} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & A_{d,N_2} \\ 0 & \ldots & 0 \end{bmatrix}$$

and $\Gamma(\gamma) = I_{N_r} \otimes \text{diag}(\gamma_1, \ldots, \gamma N_2)$. The dimension of each matrix is given in Table 1.

**Proof:** See Appendix B.

The optimization problem (10) is the primal problem with its optimizer being the attack vector $a$. This optimization problem is nonconvex. With the help of the S-procedure and dissipative system theory, (10) is converted to its equivalent dual SDP form (13) with its optimizer $\gamma, P$, which is convex. This equivalency also helps us to conclude that the duality gap is zero. The necessary and sufficient conditions for the value of (13) to be finite is given Lemma IV.5.

**Lemma IV.5 (Boundedness):** Consider $N_2$ i.i.d. realizations of the closed-loop system (4), each with an uncertainty $\delta_i$. The optimal value of (13) with the abovementioned system realizations is bounded if and only if the system with $\Sigma_\gamma = (A, B, C_p, D_p)$ and $\Sigma_\delta = (A_d, B_{d,1}, C_{r,1}, D_{r,1})$ satisfy one of the following.

1) The system $\Sigma_a$ has no zeros on the unit-circle.
2) The zeros on the unit circle of the system $\Sigma_a$ (including multiplicity and input direction) are also zeros of $\Sigma_p$.

**Proof:** See Appendix C.

---

3With abuse of notation, we denote that every element of the vector $\gamma$ is nonnegative by $\gamma \geq 0$. 

---

**Table I**

| Matrix | Dimension | Matrix | Dimension |
|--------|-----------|--------|-----------|
| $A$    | $n_cN_2 \times n_cN_2$ | $B$    | $n_cN_2 \times n_a$ |
| $C_p$  | $n_pN_2 \times n_cN_2$ | $D_p$  | $n_pN_2 \times n_a$ |
| $C_r$  | $n_rN_2 \times n_cN_2$ | $D_r$  | $n_rN_2 \times n_a$ |
Let the outputs of $\Sigma_s$ and $\Sigma_r$ be represented by $\hat{y}_s$ and $\hat{y}_r$, respectively. Then, in words, Lemma IV.5 states that the maximum expected loss (10) is bounded if either, there does not exist an attack vector which makes the output $\hat{y}_i$ identically zero, or for all attack vectors which yield $\hat{y}_i$ identically zero, it also yields $\hat{y}_r$ identically zero.

It is important to study the conditions for unbounded risk because of the following: 1) if the conditions of the lemma do not hold, it means that there exists an attack vector that can remain stealthy but cause very huge system disruptions, and 2) it depicts that the conditions for unbounded risk are restrictive in comparison to the deterministic case [24, Sec. 4] as shown in Fig. 3. Here, the function $\epsilon(\cdot)$ is evaluated according to [23, eq. (7)] where $\lambda = 10^{-2}$ and $s_{N_2} = \text{supp}(\gamma_i)^2$. And $\epsilon(\cdot)$ should be interpreted as follows: The attack vector obtained by solving (13) with $N_2$ samples, with probability $1 - \lambda$, will be at most feasible for the fraction $(1 - \epsilon(\cdot))\Omega$ of the set $\Omega$.

Since $\epsilon(\cdot)$ represents the fraction of the uncertainty set to which the attack vector is infeasible, it is intuitive to expect this value to be close to zero. Numerically we have observed from Table III that $\text{supp}(\gamma)$ is always one. So, assuming again that $s_{N_2} = 1$, the number of samples required to guarantee that the attack vector, with a probability $1 - \lambda$, is feasible for $1 - \epsilon(\cdot)$ of the uncertainty set can be obtained by picking $N_2$ such that (16) holds [23, eq. (7)]

$$N_2^{-1} \sqrt{\frac{\lambda}{N_2}} = 1 - \epsilon(\cdot).$$

Thus, (16) gives an idea of how $N_2$ increases as $\epsilon(\cdot)$ decreases. And consequently gives an idea of the scalability of the proposed approach as the dimension of the matrix inequality (13) depends on $N_2$. We also depict the computation complexity by providing the time taken to solve (13) in the last column of Table III.

V. NUMERICAL EXAMPLE

Consider a power generating system [10, Sec. 4] as shown in Fig. 2 and represented by

$$\begin{align*}
\begin{bmatrix}
\eta_1 \\
\eta_2 \\
\eta_3
\end{bmatrix} &=
\begin{bmatrix}
-\frac{1}{T_{\text{in}}} & \frac{K_{im}}{T_{\text{in}}^2} & -\frac{2K_{im}}{T_{\text{in}}^3} \\
\frac{1}{T_{\text{in}}} & 0 & \frac{1}{T_{\text{in}}^2} \\
0 & 0 & 0
\end{bmatrix}
+ \begin{bmatrix}
0 \\
0 \\
1/T_{\text{g}}
\end{bmatrix} \hat{u} \\
\eta
\end{align*}$$

(14)

and

$$y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\
\eta_2 \\
\eta_3
\end{bmatrix}, \quad y_p = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\
\eta_2
\end{bmatrix}. \quad (15)$$

Here, $\eta \triangleq [df; dp + 2dx; dx]$, $df$ is the frequency deviation in Hz, $dp$ is the change in the generator output per unit (p.u.), and $dx$ is the change in the valve position p.u. The parameters of the plant are listed in Table II.

| $K_{im}$ | $T_{\text{in}}$ | $T_{\text{g}}$ | $R$ |
|---------|----------------|--------------|-----|
| 1       | 4.6            | 6            | 0.1 |
| $T_{\text{h}}$ | 0.08          |              |     |

Table II

TABLE III

RATIONAL ADVERSARY—A POSTERIORI GUARANTEES

| $N_2$ | $\supp(\gamma)$ | $\epsilon(s_{N_2}^+)$ | $\gamma_{RA}(N_2)$ | Time (seconds) |
|-------|-----------------|------------------------|-------------------|---------------|
| 8     | 1               | 0.7141                 | 638.04            | 17            |
| 15    | 1               | 0.5112                 | 628.60            | 116           |
| 21    | 1               | 0.4142                 | 617.26            | 578           |

In view of Lemma IV.1, we choose $N_2 = 21$ to approximate the set $\Omega$. With this approximation, by solving the convex SDP (13), we obtain $\gamma_{RA}(N_2) = 617.267$. To recall, $\gamma_{RA}$ represents the maximum expected performance loss of the system operator. In this implementation, the value of risk was obtained for $\beta = 0$ since we considered a maximally robust adversary.

Next, we discuss the validity of the approximation $\Omega_{N_2}$. For varying values of $N_2$, the number of nonzero $\gamma_{i}$s obtained while solving (13) is shown in the first two columns of Table III. In view of Theorem IV.3, if we solve the problem (13) with an arbitrary $N_2$, we can provide guarantees on the optimization problem (7) as shown in column 3 of Table III. Here, the function $\epsilon(\cdot)$ is evaluated according to [23, eq. (7)] where $\lambda = 10^{-2}$ and $s_{N_2} = \text{supp}(\gamma_i)^2$. And $\epsilon(\cdot)$ should be interpreted as follows: The attack vector obtained by solving (13) with $N_2$ samples, with probability $1 - \lambda$, will be at most feasible for the fraction $(1 - \epsilon(\cdot))\Omega$ of the set $\Omega$.

Since $\epsilon(\cdot)$ represents the fraction of the uncertainty set to which the attack vector is infeasible, it is intuitive to expect this value to be close to zero. Numerically we have observed from Table III that $\text{supp}(\gamma)$ is always one. So, assuming again that $s_{N_2} = 1$, the number of samples required to guarantee that the attack vector, with a probability $1 - \lambda$, is feasible for $1 - \epsilon(\cdot)$ of the uncertainty set can be obtained by picking $N_2$ such that (16) holds [23, eq. (7)]

$$N_2^{-1} \sqrt{\frac{\lambda}{N_2}} = 1 - \epsilon(\cdot).$$

Thus, (16) gives an idea of how $N_2$ increases as $\epsilon(\cdot)$ decreases. And consequently gives an idea of the scalability of the proposed approach as the dimension of the matrix inequality (13) depends on $N_2$. We also depict the computation complexity by providing the time taken to solve (13) in the last column of Table III.

VI. CONCLUSION

In this article, we study the problem of risk assessment on uncertain control systems under a bounded-rational adversary. Using the OOG as an impact metric, we formulated the risk assessment problem and observe that it corresponds to a nonconvex robust optimization problem. A scenario-based approach was used to relax the robust optimization problem to their sampled counterpart. Using dissipative system theory, the nonconvex sampled problem in signal space was converted to its convex dual problem in matrix inequalities. Detailed proof of the zero-duality gap was provided using the S-procedure. We additionally provide necessary and sufficient conditions for risks to be bounded, which highlights the important role of uncertainty and how it is incorporated in attack scenarios. The results are depicted with numerical examples. Future work includes the following: 1) investigating the risk assessment problem where the uncertainty set can be approximated as a polytopic set, and 2) studying the relation between our risk measure (7) and a coherent risk measure [24].

APPENDIX A

PROOF OF THEOREM IV.3

Before presenting the proof, an introduction to scenario-based constraint satisfaction is provided.

The confidence is denoted by $\lambda$ here whereas it is denoted by $\beta$ in [23].

1. The confidence is denoted by $\lambda$ here whereas it is denoted by $\beta$ in [23].
A. Scenario-Based Constraint Satisfaction [23]

Consider the constrained nonconvex optimization problem
\[
\inf_{\theta, \delta} f(\theta, \delta), \quad \text{where} \quad \delta \in \Delta \quad \text{is the uncertainty and} \quad \theta \quad \text{is the infinite-dimensional decision parameter which lies in the set}\]
\[
\Theta \triangleq \bigcap_{\Delta} \Theta_{\delta}, \quad \text{where} \quad \Theta_{\delta} \quad \text{is the constraint set, which includes all the admissible parameters for the isolated uncertainty} \quad \delta.
\]

Definition A.1.1 (Violation probability): Let us define the violation probability as \(\mathbb{V}(\theta) \triangleq \mathbb{P}_{\Delta}(\delta \in \Delta \setminus \Theta_{\delta})\).

Definition A.1.2 (\(\epsilon\)-level solution): Let \(\epsilon \in (0, 1)\), then, \(\theta \in \Theta\) is an \(\epsilon\) level solution if \(\mathbb{V}(\theta) \leq \epsilon\).

Definition A.1.3 (Confidence level : \(\lambda\)): Let \(\lambda \in (0, 1)\), then, the confidence level \(\lambda\) represents the probability that \(\theta\) is not an \(\epsilon\) level solution, i.e., \(\lambda \triangleq \mathbb{P}(\mathbb{V}(\theta) > \epsilon)\).

Proof of Theorem IV.3: The optimization problem (9) can be reformulated as
\[
- \inf_{a \in \mathcal{E}_2} \left\{ \mathbb{E}_{\Omega} \left[ -\|y_\delta(\delta)\|_2^2 \right] : \left\| x_{[-\infty]} \right\|_2 \leq 1, \quad x_{[\infty]} = 0 \right\} \quad \forall \delta \in \Omega \right\}.
\]

In view of [23, Th. 1], we define the objective function as \(f(\alpha, \delta) \triangleq \mathbb{E}_{\Omega} \left[ -\|y_\delta(\delta)\|_2^2 \right], \) where \(y_\delta(\delta)\) is also a function of the attack vector \(\alpha\). Let us define the set \(\Theta\) as in (12). Let us define a confidence level \(\lambda \in (0, 1)\), a constant \(N_2\) and \(\epsilon(\cdot)\) such that (11) holds. Then, applying [23, Th. 1], we obtain that
\[
\mathbb{P}_{N_2} \{ \Omega \in \Omega \mid a_{\star N_2} \notin \Theta \} > \epsilon(s_{N_2}) \leq \lambda.
\]
Thus, with a probability level \(1 - \lambda\), the solution \(a_{\star N_2}\) is \((s_{N_2})\) feasible to the optimization problem (17). In our problem setting, \(a_{\star N_2}\) is the optimal argument of the optimization problem
\[
- \inf_{a \in \mathcal{E}_2} \left\{ \frac{1}{N_2} \sum_{i=1}^{N_2} \|y_{\delta,i}\|_2^2 : \left\| y_{\delta,i} \right\|_2 \leq 1, \quad x_{[\infty]} = 0 \right\} \quad \forall \delta \in \mathcal{S}
\]
where \(\mathcal{S} \triangleq \{1, \ldots, N_2\}\), which can be rewritten as (10). This concludes the proof.

APPENDIX B
PROOF OF THEOREM IV.4

A core step of the proof relies on [25, Th. 4.3.1], which, in turn, leverages the notion of S-system. Therefore, before presenting the proof, an introduction to the S-system is provided. Readers interested in a more detailed proof are referred to the extended preprint of this article [26, Sec. A.5].

A. S-System [25, Definition 4.3.1]

Let \(\mathcal{L}\) be a real Hilbert space with a well-defined inner product denoted by \(\langle \cdot, \cdot \rangle\). Let \(\mathcal{G}_0(\omega), \ldots, \mathcal{G}_k(\omega)\) be quadratic functional mappings \(\mathcal{L} \to \mathbb{R}\). Let \(\omega\) be a discrete-time signal.

Definition A.2.1 (S-System, [25, Definition 4.3.1]): The quadratic functionals \(\mathcal{G}_0(\omega), \mathcal{G}_1(\omega), \ldots, \mathcal{G}_k(\omega)\) form an S-system if there exist a bounded linear operator \(\mathcal{T}_i : \mathcal{L} \to \mathcal{L}, i = 1, 2, \ldots,\) such as follows.
1) \(\mathcal{T}_i(\omega, \omega) \to 0 \) as \(i \to \infty\), \(\forall \omega, \omega \in \mathcal{L}\).
2) If \(\omega \in \mathcal{M}\), then \(\mathcal{T}_i \omega, \omega \in \mathcal{M}, \forall i = 1, 2, \ldots,\) where \(\mathcal{M}\) is a linear subspace of \(\mathcal{L}\).
3) \(\mathcal{G}_j(\mathcal{T}_i \omega) \to \mathcal{G}_j(\omega)\) as \(i \to \infty\), \(\forall \omega \in \mathcal{L}\) and \(j = 0, 1, \ldots, k\). We next present a theorem that helps in proving Theorem IV.4.

Theorem A.2.1: Let us define a stable discrete-time linear time-invariant system of the form
\[
\eta[k + 1] = \Phi \eta[k] + \Lambda \mu[k],
\]
\[
\sigma[k] = \Pi \eta[k] + \Upsilon \mu[k],
\]
\[
\eta[0] = \eta_0, \quad \eta[\infty] = 0.
\]
Let us define the set \(\mathcal{L}\) as
\[
\mathcal{L} \triangleq \left\{ \omega = \begin{bmatrix} \sigma \end{bmatrix} : \begin{bmatrix} \mu \end{bmatrix} \in \ell_2, \quad \eta[0] = \eta_0, \quad \eta[\infty] = 0 \right\}.
\]
Let us also define the functionals \(\mathcal{G}_0(\omega) \triangleq \sum_{k=0}^{\infty} \omega[k]^T \mathcal{M}_0 \omega[k] + \zeta_0, \ldots, \mathcal{G}_k(\omega) \triangleq \sum_{k=0}^{\infty} \omega[k]^T \mathcal{M}_k \omega[k] + \zeta_k, \) where \(\mathcal{M}_0, \ldots, \mathcal{M}_k\) are given matrices and \(\zeta_0, \ldots, \zeta_k\) are given constants. Then, the functionals \(\mathcal{G}_0(\cdot), \ldots, \mathcal{G}_k(\cdot)\) form an S-system.

Proof of Theorem A.2.1: In view of Definition A.2.1, let us define the operator \(\mathcal{T}_i\) as
\[
\mathcal{T}_i \omega[k] = \begin{cases} 0, & \text{if } 0 \leq k \leq i, \\ \omega[k - i], & \text{if } k > i. \end{cases}
\]
For any \(\omega_1, \omega_2 \in \mathcal{L}\), \(\mathcal{T}_i \omega_1, \omega_2\) satisfies the time-invariance property of (18), \(\mathcal{T}_i \omega, \omega \in \mathcal{M}\). This proves that \(\mathcal{M}\) is invariant under the operator \(\mathcal{T}_i\). Thus condition 2) of Definition A.2.1 holds.

Let us consider a set \(\mathcal{M} = \mathcal{L}_{|_{\eta[0]=0}}\). Then, if \(\omega \in \mathcal{M}, \) due to the time-invariant property of (18), \(\mathcal{T}_i \omega, \omega \in \mathcal{M}\). This proves that \(\mathcal{M}\) is invariant under the operator \(\mathcal{T}_i\). Thus condition 3) of Definition A.2.1 holds. Since we have shown that the functionals \(\mathcal{G}_0(\omega), \mathcal{G}_1(\omega), \ldots, \mathcal{G}_k(\omega)\) satisfy conditions 1), 2), and 3) of Definition A.2.1, they form an S-system. This concludes the proof.

Now, we are ready to present the proof of Theorem IV.4.

Proof: Step 1: [Problem reformulation] Using the hypograph formulation, (10) can be rewritten as
\[
\sup_{v, \omega \in \ell_2} \left\{ v \mathcal{N} \sum_{i=1}^{N_2} \|y_{\delta,i}\|_2^2 \geq v, \quad \|y_{\delta,i}\|_2 \leq 1, \quad x_{[\infty]} = 0, \quad \forall i \in \{1, \ldots, N_2\} \right\}.
\]

From now, in the next two steps of the proof, we focus on the optimization problem (20) without the state constraints
\[
\sup_{v, \omega \in \ell_2} \left\{ v \mathcal{N} \sum_{i=1}^{N_2} \|y_{\delta,i}\|_2^2 \geq v, \quad \|y_{\delta,i}\|_2 \leq 1, \quad \forall i \in \{1, \ldots, N_2\} \right\}.
\]

The reason to focus on (21) rather than (20) is that the S-procedure becomes convenient to apply when there are no equality constraints. Thus, we drop the state constraints now and introduce them back at the end of Step 3. Equivalently, (21) can be reformulated as
\[
\inf_{\eta} \left\{ v \mathcal{N} \sum_{i=1}^{N_2} \|y_{\delta,i}\|_2^2 - v > 0, \quad 1 - \|y_{\delta,i}\|_2^2 \geq 0, \quad \forall i \right\}.
\]
Step 2: Let the system with the isolated uncertainty $\delta_i$, with attack vector as input and performance and detection outputs as outputs be $\Sigma_{\delta_i} \triangleq (\bar{A}_{d,i}, \bar{B}_{d,i}, \bar{C}_{r,i}, \bar{D}_{r,i})$ and $\Sigma_{\ell,i} \triangleq (\bar{A}_{d,i}, \bar{B}_{d,i}, \bar{C}_{r,i}, \bar{D}_{r,i})$. Let us consider a linear time-invariant system of the form (18) with the attack vector as input $\mu(\cdot)$ and the vector \( [y_{p,1}, y_{r,1}, \ldots, y_{p,N_2}, y_{r,N_2}]^T \) as output \( \sigma(\cdot) \). This system will be stable due to Assumption II.1 and the system matrices of (18) would read

\[
\begin{bmatrix}
\Phi & \Lambda \\
\Pi & \Upsilon
\end{bmatrix} = \begin{bmatrix}
A_{d,1} & 0 & 0 & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \ddots & \ddots & \ddots \\
C_{p,1} & 0 & 0 & D_{p,1} \\
C_{r,1} & 0 & 0 & D_{r,1} \\
\vdots & \ddots & \ddots & \ddots \\
0 & \ddots & \ddots & \ddots \\
C_{p,N_2} & 0 & 0 & D_{p,N_2} \\
C_{r,N_2} & 0 & 0 & D_{r,N_2}
\end{bmatrix}
\]

For this system, let us define the set $\mathcal{L}$ as in (19) where $\omega = [a^T \gamma]^T \in \mathbb{R}^{n_+ + N_2(n_+ + n_2)}$. In view of (22), let us define

\[
G_0(\omega) \equiv \frac{1}{N_2} \sum_{i = 1}^{N_2} \|y_{p,i}\|^2_{\ell_2} - v = \sum_{k = 0}^{\infty} \omega[k]^T M_0 \omega[k] + \zeta_0
\]

(23)

where $M_0 \in \mathbb{R}^{(n_+ + N_2(n_+ + n_2) \times (n_+ + N_2(n_+ + n_2))}$, $\zeta_0 = -v$ and $M_0(i, j)$ is 1, if $i = j, n_+ + i$ is odd and 0 elsewhere. Similarly, let

\[
G_k(\omega) \equiv -\|y_{r,k}\|^2_{\ell_2} + 1 = \sum_{k = 0}^{\infty} \omega[k]^T M_k \omega[k] + \zeta_k
\]

(24)

\[\forall k = 1, 2, \ldots, N_2, \text{ where } \zeta_k = -1 \text{ and } M_k(i, j) = -1, \text{ if } i = j, i = n_+ + 2k - 1 \text{ and } 0 \text{ elsewhere. Here, } M_k \text{ has the same dimension as } M_0 \forall k = 1, 2, \ldots, N_2. \]

Therefore, we have shown that the constraints of (22) can be rewritten as functionals of the set $\mathcal{L}$. We can now see that the functionals $-G_0(\cdot), G_1(\cdot), \ldots, G_{N_2}(\cdot)$ along with Lemma II.1 satisfy the conditions under which Theorem A.2.1 holds. Thus, by applying Theorem A.2.1, it follows that the functionals $-G_0(\cdot), G_1(\cdot), \ldots, G_{N_2}(\cdot)$ form an S-system. Let this be argument 1.

In the case the adversary chooses not to attack the system, i.e., $a = 0 \in \ell_{2e}$, it follows that $\|y_{r,1}\|^2_{\ell_2}$ is strictly zero since there might be residual outputs due to the difference in initial condition between the system and the detector. The threshold ($\epsilon_r = 1$) is chosen in such a way that $\|y_{r,1}\|^2_{\ell_2} \ll \delta_1$ when $a = 0$. Thus, it holds that $\exists \omega_0 = [a, \omega] = [0, \omega]$ s.t. $-\|y_{r,1}\|^2_{\ell_2} + 1 = G_k(\omega_0) > 0 \forall k = 1, \ldots, N_2$. Here, $\omega_0$ represents a real number close to zero. Let this be argument 2. Finally, for any given $\omega \in \mathcal{L}$, let $\nu^*$ be the corresponding optimal solution from (21). Then, we know from (22) that the set of inequalities $G_0(\omega) > 0, G_i(\omega) \geq 0$ for $i \in \{1, \ldots, k\}$ (where $G_k(\omega)$ in (23) is constructed using the optimal $\nu^*$) is not solvable. Let this be argument 3. Using the above arguments (1)-(3), and [25, Th. 4.3.1], we can conclude that, given the functionals in (23) and (24), there exists constants $\gamma_1 \geq 0, \gamma_2 \geq 0, \ldots, \gamma_{N_2} \geq 0$ such that (25) holds

\[
G_0(\omega) + \sum_{i = 1}^{N_2} \gamma_i G_i(\omega) \leq 0, \forall \omega \in \mathcal{L}.
\]

(25)

To conclude, in this step we have shown that (25) holds if the constraint of (22) holds. Additionally, using [27, Th. 1] (see also [25, Remark 4.3.1]), we observe that (25) implies that the constraints of (22) hold. As a result, we have that (22) and (25) are equivalent.

Step 3: We have shown that the constraint of (22) holds if (25) is true. Then, we reformulate (22) as

\[
\inf_{\nu, \gamma_1, \ldots, \gamma_{N_2} \geq 0} \left\{ \nu \left| G_0(\omega) + \sum_{i = 1}^{N_2} \gamma_i G_i(\omega) \leq 0, \forall \omega \in \mathcal{L} \right. \right\}
\]

Substituting the definition of $G_0(\omega)$, we obtain

\[
\inf_{\gamma, \gamma_1, \ldots, \gamma_{N_2} \geq 0} \left\{ \nu \left| \sum_{i = 1}^{N_2} \left( \frac{1}{N_2} \|y_{p,i}\|^2_{\ell_2} + \gamma_i G_i(\omega) \right) \leq \nu, \forall \omega \right. \right\}
\]

(26)

where $\gamma = [\gamma_1, \ldots, \gamma_{N_2}]^T$. The inner optimization problem of (26) resembles an epigraph formulation which can be rewritten as

\[
\inf_{\gamma \geq 0} \left\{ \sup_{\omega} \left[ \sum_{i = 1}^{N_2} \left( \frac{1}{N_2} \|y_{p,i}\|^2_{\ell_2} + \gamma_i G_i(\omega) \right) \right] \right. \}
\]

(27)

Observe that $\kappa$ is a maximization problem with a quadratic term in its objective. Thus, it holds that

\[
\kappa = \left\{ 1^T \gamma, \text{ if } \sum_{i = 1}^{N_2} \left( \frac{1}{N_2} \|y_{p,i}\|^2_{\ell_2} - \gamma_i \|y_{r,i}\|^2_{\ell_2} \right) \leq 0, \right. \left. +\infty, \right. \text{ otherwise} \}
\]

Using the above result in (27), we obtain

\[
\inf_{\gamma \geq 0} \left\{ 1^T \gamma, \sum_{i = 1}^{N_2} \left( \frac{1}{N_2} \|y_{p,i}\|^2_{\ell_2} - \gamma_i \|y_{r,i}\|^2_{\ell_2} \right) \leq 0, \forall \omega \right\}
\]

(28)

Thus, in this step, we have shown using $S$-procedure that the optimization problems (20) and (28) are equivalent.

Step 4: Define $\bar{x}[\infty] = [x_1[\infty]^T, \ldots, x_{N_2}[\infty]^T]^T, \bar{x}[0] = [x_1[0]^T, \ldots, x_{N_2}[0]^T]^T, \bar{y}_p = [y_{p,1}^T, \ldots, y_{p,N_2}^T]^T$ and $\bar{y}_r = [y_{r,1}^T, \ldots, y_{r,N_2}^T]^T$. Using these definitions, the constraint of (28) can be rewritten as

\[
-\frac{1}{N_2} \|\bar{y}_p\|^2_{\ell_2} + \left\| \bar{x}[\infty] \right\|^2_{\ell_2} \geq 0, \forall \omega \in \ell_{2e}, \bar{x}[\infty] = 0
\]

(29)

where $\Gamma(\gamma)$ is defined in the theorem statement. Additionally due to Assumption II.2, we have $\bar{x}[0] = 0$. Next, let us define $y_1 = \sqrt{\Gamma} \bar{y}_r, y_2 = \frac{\bar{y}_r}{\sqrt{\Gamma}},$ and the supply rate $s(\cdot) \triangleq \|y_{p,1}\|^2_{\ell_2} - \|y_{r,1}\|^2_{\ell_2}$. Then, we have shown from (29) that $\bar{y}_r = [y_{r,1}^T, \ldots, y_{r,N_2}^T]^T$. Using these definitions, $\Sigma_p = (\bar{A}, \bar{B}, \bar{C}_p, \bar{D}_p)$ and $\Sigma_r = (\bar{A}, \bar{B}, \bar{C}_r, \bar{D}_r)$ represent the system with the attack as input and $\bar{y}_p$ and $\bar{y}_r$ as system outputs, respectively. Constructing these system matrices, as we did in Step 2 of this proof concludes the proof.
To recall, the optimization problem (13) was formulated using \( y_1 = \sqrt{\gamma} y_{\bar{g}} \) and \( y_2 = \frac{1}{\sqrt{N_2}} y_{\bar{g}_p} \). Here \( y_{\bar{g}} \) and \( y_{\bar{g}_p} \) represent the outputs of \( \Sigma_p \) and \( \Sigma_r \) respectively. Due to the equivalence between 3) and 4) of (28, Proposition 2, 3), the FDI (28, Proposition 2, 4) should hold \( \forall \gamma \in \mathbb{R} \). Since we know that \( y_1 = \sqrt{\gamma} y_{\bar{g}} \) and \( y_2 = \frac{1}{\sqrt{N_2}} y_{\bar{g}_p} \), we can deduce that \( G_1(z) \) \( \Gamma(\bar{z}) \) and \( G_2(z) \) \( \frac{1}{N_2} \) \( G_2(z) \) in (28, Proposition 2, 4), where \( \bar{G}_r(z) = \bar{C}_r(z I - A)^{-1} B + \bar{D}_r \) and \( \bar{G}_p(z) = \bar{C}_p(z I - A)^{-1} B + \bar{D}_p \). Thus, (13) can be rewritten as

\[
\inf \left\{ 1^T \gamma |G_r(z)|^2 \Gamma(\gamma) \bar{G}_r(z) - \bar{G}_p(z) \bar{p}(z) \geq 0, \forall \gamma \right\}
\]

(30)

Let us define the following sets such that \( C^{\text{sa}} = Z_p \cup \bar{Z} \cup \bar{Z}_r \cup Z_p \).

\[
\mathcal{Z}_p = \{ x \in C^{\text{sa}} | \bar{G}_r(z)x = 0, \bar{G}_p(z)x = 0 \},
\]

\[
\mathcal{Z}_s = \{ x \in C^{\text{sa}} | \bar{G}_r(z)x \neq 0, \bar{G}_p(z)x \neq 0 \},
\]

\[
\mathcal{Z}_p = \{ x \in C^{\text{sa}} | \bar{G}_r(z)x = 0, \bar{G}_p(z)x \neq 0 \},
\]

\[
\mathcal{Z}_p = \{ x \in C^{\text{sa}} | \bar{G}_r(z)x \neq 0, \bar{G}_p(z)x = 0 \}.
\]

**Sufficiency:** For any given \( z \) such that \( |z| = 1 \), if \( x \in \mathcal{Z}_p \) or \( x \in \mathcal{Z}_p \), choosing \( \Gamma(\gamma) = 0 \) satisfies the constraint of (30). Similarly, if \( x \in \mathcal{Z}_s \), let us pick \( \Gamma(\gamma) = cI_{n_r} \), where \( c \) is a constant. Then, the value of (30) is bounded if there exists a bounded \( c \) that makes \( \mathcal{Z}_s = \{ x \in C^{\text{sa}} | \bar{G}_r(z)x = 0, \bar{G}_p(z)x = 0 \} \) bounded. \( \mathcal{Z}_s \) is bounded since the denominator cannot become zero (since \( x \in \mathcal{Z} \) and \( \Gamma(\gamma) \) is full rank), and we have assumed that the \( \bar{G}_r(z) \) and \( \bar{G}_p(z) \) are stable (Assumption II.1). Next, we prove sufficiency when \( x \in \mathcal{Z}_s \).

When condition 1) of the lemma statement holds, by definition of a zero \( \forall |z| = 1, 2s \neq 0 \in C^{\text{sa}} \) such that \( \bar{G}_r(z)x = 0 \). Thus it follows that \( \mathcal{Z}_s = \mathcal{Z}_p \cup \bar{Z} \cup \bar{Z}_r \cup \bar{Z}_p \). When condition 2) of the lemma statement holds, by definition of a zero \( \forall |z| = 1, 2s \neq 0 \) such that \( \bar{G}_r(z)x = 0 \) and \( \bar{G}_p(z)x \neq 0 \). Thus it follows that \( \mathcal{Z}_s = 0 \).

**Necessity:** Assume that there exists a bounded \( \Gamma(\gamma) \) that solves (30). We also assume that there exists a complex number \( z_1 \) on the unit circle which is a zero of the system \( \Sigma_r \) (including multiplicity and input direction) but are not zeros of \( \Sigma_r \). By definition of a zero, it holds that \( \exists s \neq 0 \) such that \( \bar{G}_r(z_1)x = 0 \) and \( \bar{G}_p(z_1)x \neq 0 \). Under these assumptions, when \( z = z_1 \) and \( x = s \), the constraint of (30) can be rewritten as \( -s^H \bar{G}_r^T(z_1) \bar{G}_r(z_1) s \geq 0 \) which cannot hold since \( \bar{G}_r(z_1)x \neq 0 \). That is, the feasibility set of (30) is empty which contradicts our assumption. This concludes the proof.

### References

[1] S. M. Dibaji, M. Pirani, D. B. Flaholz, A. M. Annaswamy, K. H. Johansson, and A. Chakrabortty, “A systems and control perspective of CPS security,” Annu. Rev. Control, vol. 47, pp. 394–411, 2019.

[2] D. U. Case, “Analysis of the cyber attack on the ukrainian power grid,” Electric Inf. Sharing Anal. Center, vol. 388, pp. 1–29, 2016.

[3] K. E. Hemsley and Dr. E. F. Ronald, “History of industrial control system cyber incidents,” Dec. 2018. [Online]. Available: https://www.osti.gov/hribia/1505628

[4] R. Ross, “Guide for conducting risk assessment,” Nat. Inst. Stand. Technol., Gaithersburg, MD, USA, 2012, doi: 10.6028/NIST.SP.800-30r1.