2022 American Control Conference (ACC)
Atlanta, USA, June 8-10, 2022

Risk-averse controller design against data injection attacks on actuators for uncertain control systems

Sribalaji C. Anand[1] and André M. H. Teixeira[2]

Abstract— **In this paper, we consider the optimal controller design problem against data injection attacks on actuators for an uncertain control system. We consider attacks that aim at maximizing the attack impact while remaining stealthy in the finite horizon. To this end, we use the Conditional Value-at-Risk to characterize the risk associated with the impact of attacks. The worst-case attack impact is characterized using the recently proposed output-to-output ℓ_2-gain (OOG). We formulate the design problem and observe that it is non-convex and hard to solve. Using the framework of scenario-based optimization and a convex proxy for the OOG, we propose a convex optimization problem that approximately solves the design problem with probabilistic certificates. Finally, we illustrate the results through a numerical example.**

I. INTRODUCTION

Cyber-physical systems (CPSs) represent a large class of networked control systems where the physical world and the digital infrastructure are tightly coupled, such as smart cities, autonomous systems, transportation networks, and Internet Of Things. However, the trend towards increased usage of open-standard communication protocols among control systems has made these systems vulnerable to cyber-attacks such as Stuxnet [1], Industroyer [2], etc. Such cyber-attacks can negatively affect the operation of CPS [3].

Significant work is done in detecting and mitigating cyber-attacks (see [4], [5] and references therein). For instance, [6] designs an optimal controller in the presence of covert attacks. The limitation of [6] is, it approximates the risk metric Conditional Value-at-Risk (CVaR) empirically using samples, and it parameterizes the controller as a finite family of Finite Impulse Response (FIR) filters. Although these approximations simplify the problem, the validity of these approximations is not discussed except in the asymptotic case i.e., as the number of samples for empirical approximation and the number of FIR filters tends to infinity.

The article [7] proposes and solves two controller design problems. Firstly, it proposes a convex design problem such that the volume of the reachable set of states by the adversary is minimized. Secondly, it proposes a convex design problem that maximizes the Euclidean distance between the set of states reachable by the adversary and the set of critical states. A similar approach was also adopted in [8]. However, both of these works do not consider an uncertain system. The works [9] and [10] addresses the issue of jointly designing the controller and detector against false data injection (FDI) attacks. However, they also assume a deterministic system.

This paper addresses some of the existing limitations in the literature, by investigating the optimal controller design problem against FDI attacks on actuators for an uncertain control system. To this end, we adopt the following setup. We consider a discrete-time (DT) linear time-invariant (LTI) process with parametric uncertainty, a static output feedback controller, and an anomaly detector. An adversary with perfect system knowledge injects false data into the actuators. In reality, it is hard for the adversary to have perfect system knowledge, but this assumption helps to study the worst-case. The system operator (or the defender) knows only about the bounds of the uncertainty. Under this setup, we present the following contributions.

1) Firstly, we formulate the risk-averse design problem. Here, for a given realization of the uncertainty, we use the output-to-output ℓ_2-gain (OOG) [11] to characterize the worst-case impact. We then use the CVaR to characterize the risk associated with the attack impact. The advantages of using the OOG over the classical H_∞/H_- metrics were demonstrated in [9]. We also observe that the design problem corresponds to an untractable infinite non-convex optimization problem.
2) Secondly, extending the results of [12], we derive an upper bound for the OOG. Using this upper bound, we relax the infinite non-convex design problem into an infinite convex design problem.
3) Finally, by adopting the scenario-based approach [13], we modify the infinite convex optimization problem into its sampled counterpart. We also provide probabilistic guarantees on the infinite design problem based on the number of samples used to formulate the sampled optimization problem and the dimension of the controller. The advantage of using scenario-based approach over other approaches is discussed in [14].

To the best of the author's knowledge, the problem of risk-sensitive controller design for an uncertain control in the finite horizon against FDI attacks has not been addressed in the literature.

The remainder of this paper is organized as follows. Section II describes the problem background. The design problem is formulated in Section III. The problem is relaxed and convexified in Section IV. Section V approximates the problem empirically using the scenario-based approach. We

*This work is supported by the Swedish Research Council under the grant 2018-04396 and by the Swedish Foundation for Strategic Research.

[1] Sribalaji C. Anand is with the Department of Electrical Engineering, Uppsala University, PO Box 65, SE-75103, Uppsala, Sweden. sribalaji.anand@angstrom.uu.se

[2] André M. H. Teixeira is with the Department of Information Technology, Uppsala University, PO Box 337, SE -75105, Uppsala, Sweden. andre.teixeira@it.uu.se

Authorized licensed use limited to: Uppsala Universitetsbibliotek. Downloaded on February 19,2023 at 13:27:42 UTC from IEEE Xplore. Restrictions apply.

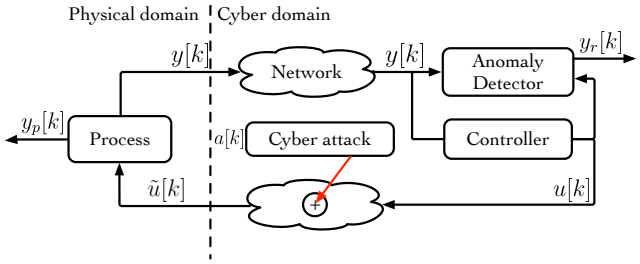

Fig. 1. Control system under data injection attack on actuators

illustrate the results using a numerical example in Section VI. Finally, we provide concluding remarks in Section VII.

II. Problem Background

In this section, we describe the control system structure and the goal of the adversary. Consider the general description of a finite horizon closed-loop DT LTI system with a process ($\mathcal{P}$) with parametric uncertainty, a static output feedback controller ($\mathcal{C}$) and an anomaly detector ($\mathcal{D}$) as shown in Fig. 1. The closed-loop system is represented by

$$\begin{aligned}
\mathcal{P}: &\begin{cases} x_p[k+1] &= A^{\Delta}x_p[k] + B\tilde{u}[k] \\ y[k] &= Cx_p[k] \\ y_p[k] &= C_J x_p[k] \end{cases} \\
\mathcal{C}: &\begin{cases} u[k] &= Ky[k] \end{cases} \\
\mathcal{D}: &\begin{cases} \hat{x}_p[k+1] &= A\hat{x}_p[k] + Bu[k] + Ly_r[k] \\ y_r[k] &= y[k] - C\hat{x}_p[k],\ k = 0, \ldots, N_h - 1. \end{cases}
\end{aligned} \tag{1}$$

Here $A^{\Delta} \triangleq A + \Delta A(\delta)$ with A representing the nominal system matrix and $\delta \in \Omega$ denoting the probabilistic parameter uncertainty with probability space $(\Omega, \mathcal{D}_a, \mathbf{P})$. We assume the uncertainty set $\Omega \subset \mathbb{R}^v$ to be closed, bounded, and to include the zero uncertainty yielding $\Delta A(0) = 0$. The state of the process is represented by $x_p[k]$, the output of the process is $y[k]$, $\tilde{u}[k]$ is the control signal received by the process, $u[k]$ is the control signal generated by the controller, $y_p[k]$ is the virtual performance output, and $y_r[k]$ is the residue generated by the detector. In this paper, we assume all signals have the same dimension n_x. That is, we consider a fully actuated square system. The system is said to have a good performance over the horizon N_h, when the energy of the performance output ($||y_p||^2_{\ell_2,[0,N_h]}$) is small. In the closed-loop system described above, we consider that an adversary is injecting false data into the actuators. An attack is said to be detected when the energy of the detection output ($||y_r||^2_{\ell_2,[0,N_h]}$) is higher than a predefined threshold (say ϵ_r). We assume that the detection threshold and the detector L is designed to be robust against all uncertainties.

Given this setup, we now discuss the resources the adversary has access to. For clarity, we establish the following:

Assumption 2.1: (A^{Δ}, B) is controllable $\forall \delta \in \Omega$. ◁

Assumption 2.2: Matrices B and C are invertible. ◁

A. Disruption and disclosure resources

The adversary can access the control channels and inject data. This is represented by $\tilde{u}[k] = u[k] + a[k]$, where $a[k] \in \mathbb{R}^{n_x}$ is the data injected by the adversary. The adversary cannot access the sensor channels. The adversary does not have access to any disclosure (eavesdropping) resources.

B. System knowledge

We assume that, at design time, the defender knows the bounds of the set Ω and that the system matrix A^{Δ} is known only up to the nominal system matrix A. Next, at operation time, we assume that the adversary has full system knowledge. That is, the adversary knows the system matrix A^{Δ} without any uncertainties. In reality, it is hard for the adversary to know the system matrices, but this assumption helps to study the worst case.

The system knowledge is used by the adversary to calculate the optimal data injection attacks. Defining $e[k] \triangleq x_p[k] - \hat{x}_p[k]$ and $x[k] \triangleq [x_p[k]^T \ \ e[k]^T]^T$, the closed-loop system under attack with the performance output and detection output as system outputs becomes

$$\mathcal{P}_{cl}: \begin{cases} x[k+1] &= A^{\Delta}_{cl}x[k] + B_{cl}a[k] \\ y_p[k] &= C_p x[k] \\ y_r[k] &= C_r x[k], \end{cases} \tag{2}$$

$$\text{where } A^{\Delta}_{cl} \triangleq \begin{bmatrix} A^{\Delta} + BKC & 0 \\ \Delta A & A - LC \end{bmatrix}, \quad B_{cl} \triangleq \begin{bmatrix} B \\ B \end{bmatrix},$$
$$C_p \triangleq \begin{bmatrix} C_J & 0 \end{bmatrix}, \quad C_r \triangleq \begin{bmatrix} 0 & C \end{bmatrix}.$$

C. Attack goals and constraints

Given the resources the adversary has access to, the adversary aims at disrupting the system's behavior whilst remaining stealthy. The system disruption is evaluated by the increase in energy of the performance output, and the attack signal is deemed to be stealthy when the energy of the detection output is less than ϵ_r. Next, we discuss the optimal attack policy of the adversary and the design problem of the defender when the system is deterministic.

D. Design for a deterministic system

From the previous discussions, it can be understood that the goal of the adversary is to maximize the energy of the performance output whilst remaining stealthy. When the system is deterministic ($\Omega = \{0\}$), the attack policy of the adversary can be formulated as

$$\begin{aligned} q(K,0) \triangleq \sup_{a \in \ell_{2e}} \quad & \|y_p(K,0)\|^2_{\ell_2} \\ \text{s.t.} \quad & \|y_r(K,0)\|^2_{\ell_2} \leq \epsilon_r,\ x(K,0)[0] = 0, \end{aligned} \tag{3}$$

where the subscript $[0, N_h]$ is dropped for clarity. In (3), $q(K,0)$ is the disruption caused by the attack signal on the nominal system, $y_p(K,0)$ and $y_r(K,0)$ are the performance output and the detection output under the given controller $K \in \mathbb{R}^{n_x \times n_x}$, and N_h is the horizon length. In (3), the constraint $x(K,0)[0] = 0$ is introduced since the system is at equilibrium before the attack commences.

Assumption 2.3: The system (2) is at equilibrium before the attack commences. ◁

The aim of the defender then is to design a controller K such that the disruption caused by the adversary ($q(K,0)$)

Authorized licensed use limited to: Uppsala Universitetsbibliotek. Downloaded on February 19,2023 at 13:27:42 UTC from IEEE Xplore. Restrictions apply.

is minimized. To this end, the design problem can be formulated as

$$K^* = \arg\inf_K q(K, 0) \tag{4}$$

The design problem (4) is optimal only when (2) is deterministic. By extending (4), we formulate the design problem when the system is uncertain in the next section.

III. Problem Formulation

Consider the data injection attack scenario where the parametric uncertainty $\delta \in \Omega$ of the system is known to the adversary but not to the defender. The defender knows only about the probabilistic description of the set Ω. In reality, it is hard for the adversary to know the system matrices, but this assumption helps to study the worst case. Under this setup, the adversary can cause high disruption by remaining stealthy as it will be able to inject attacks by solving

$$\begin{aligned} q(K,\delta) \triangleq \sup_{a \in \ell_{2e}} \quad & \|y_p(K,\delta)\|_{\ell_2}^2 \\ \text{s.t.} \quad & \|y_r(K,\delta)\|_{\ell_2}^2 \leq \epsilon_r,\ x(K,\delta)[0] = 0, \end{aligned} \tag{5}$$

where $y_p(K,\delta)$ and $y_r(K,\delta)$ are the performance and detection output corresponding to the controller K and uncertainty δ. Since the defender does not know the system completely, $q(K,\delta)$ becomes a random variable. Thus, from the defenders point of view, the best option is to choose a feedback policy K, such that the risk corresponding to the impact random variable $q(K,\delta)$ is minimized. This design problem can be formulated as *Problem 1*.

Problem 1: Find an optimal feedback controller K^* s.t:

$$K^* \triangleq \arg\inf_K \mathcal{R}_\Omega(q(K,\delta)),$$

where $\mathcal{R}_\Omega$ is a risk metric chosen by the defender. The subscript Ω denotes that the risk acts on the uncertainty whose probabilistic description is known to the defender. ◁

Problem 1 searches for a controller K such that the risk is minimized. Let us consider the setup where the defender evaluates the risk based on the risk metric CVaR. CVaR is used in the research community due to its numerous advantages [15] and is defined in *Definition 3.1*.

Definition 3.1 (CVaR [13]): Given a random variable X and $\alpha \in (0,1)$, the CVaR is defined as [1]

$$\text{CVaR}_\alpha(X) = \mathbb{E}\{X | X > \text{VaR}_\alpha(X)\},$$

$$\text{where} \qquad \text{VaR}_\alpha(X) = \inf\{x | \mathbb{P}[X \geq x] \leq \alpha\}.$$

$\text{CVaR}_\alpha(X) = \beta$ implies that $X \leq \beta$ at least $\alpha \times 100\%$ of the time on average. ◁

In our setting, the defender is interested in determining the controller such that the CVaR_α (given α) of the impact random variable ($q(K,\delta)$) is minimized. To this end, *Problem 1* can be reformulated as

$$K^* = \arg\inf_K \mathbb{E}_\Omega\{q(K,\delta) | q(K,\delta) > \text{VaR}_\alpha(q(K,\delta))\}. \tag{6}$$

[1]This Definition assumes the distribution of X has no point masses. For general definitions of CVaR see [16].

There are two difficulties in solving (6). Firstly, (5) is non-convex for any given δ. Secondly, since the operator $\mathbb{E}$ operates over the continuous space Ω, the optimization problem (6) is computationally intensive and in general NP-hard. To this end, in *Section IV*, we determine a convex approximation for (5). We then use this approximation, to recast (6) as a convex optimization problem. In *Section V*, we provide a method to approximate the expectation operator.

IV. Design problem formulation using a convex Impact proxy

In this section, we consider the function $q(K,\delta_j)$ for a given uncertainty $\delta_j \in \Omega$ and prove that it has an upper bound. We then show that the term of the upper bound that is dependent on the controller (say $\bar{q}(\cdot)$) is convex in K. The main objective of performing this step is that, once we determine the term $\bar{q}(\cdot)$, it can be used in (6) instead of $q(K,\delta_j)$ to formulate a relaxed convex design problem. To this end, we will refer to $\bar{q}(\cdot)$ as *Impact proxy* in the reminder of the paper.

To derive the upper bound, we begin by defining the vectors $\mathbf{a}_j \triangleq \left[a_j[0]^T, \ldots, a_j[N_h-1]^T\right]^T$, $\mathbf{x}_{p,j} \triangleq \left[x_{p,j}[1]^T, \ldots, x_{p,j}[N_h]^T\right]^T$, $\mathbf{e}_j \triangleq \left[e_j[1]^T, \ldots, e_j[N_h]^T\right]^T$, $\mathbf{y}_{p,j} \triangleq \left[y_{p,j}[1]^T, \ldots, y_{p,j}[N_h]^T\right]^T$, and $\mathbf{y}_{r,j} \triangleq \left[y_{r,j}[1]^T, \ldots, y_{r,j}[N_h]^T\right]^T$. Here $\mathbf{a}_j, \mathbf{x}_{p,j}, \mathbf{y}_{p,j}$ and $\mathbf{y}_{r,j}$ are the stacked attack vector, system state, performance output vector and the detection output vectors corresponding to the uncertainty δ_j respectively. Let us define the matrices $F_{xa}(K,\delta_j), F_{ea}(\delta_j), F_{ex}(\delta_j) \in \mathbb{R}^{n_x N_h \times n_x N_h}$, such that

$$\begin{aligned} &\mathbf{x}_{p,j} = F_{xa}(K,\delta_j)\mathbf{a}_j,\ \mathbf{e}_j = F_{ea}(\delta_j)\mathbf{a}_j + F_{ex}(\delta_j)\mathbf{x}_{p,j}, \\ &\mathbf{y}_{p,j} = F_p(K,\delta_j)\mathbf{a}_j,\ \ \mathbf{y}_{r,j} = F_r(K,\delta_j)\mathbf{a}_j, \\ &F_p(K,\delta_j) \triangleq (I_{N_h} \otimes C_J) F_{xa}(K,\delta_j), \\ &F_r(K,\delta_j) \triangleq (I_{N_h} \otimes C)(F_{ea}(\delta_j) + F_{ex}(\delta_j) F_{xa}(K,\delta_j)). \end{aligned}$$

Under the uncertainty δ_j, let us represent the system matrix of (1) by A_j. Then $F_{va}(K,\delta_j), \upsilon = \{x, e\}$ is given by

$$\begin{bmatrix} B & 0 & \ldots & 0 \\ A_{\upsilon,j}B & B & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{\upsilon,j}^{N_h-1}B & A_{\upsilon,j}^{N_h-2}B & \ldots & B \end{bmatrix},$$

where $A_{x,j} \triangleq A_j + BKC$ and $A_{e,j} \triangleq A_j - LC$. Similarly $F_{ex}(\delta_j)$ is given by

$$\begin{bmatrix} 0 & 0 & \ldots & 0 \\ \Delta A & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{e,j}^{N_h-2}\Delta A & A_{e,j}^{N_h-3}B\Delta A & \ldots & 0 \end{bmatrix}$$

Under these definitions, $q(K,\delta_j)$ can be obtained by the non-convex optimization problem

$$\begin{aligned} q(K,\delta_j) \triangleq \sup_{\mathbf{a}_j} \quad & \|F_p(K,\delta_j)\mathbf{a}_j\|_2^2 \\ \text{s.t.} \quad & \|F_r(K,\delta_j)\mathbf{a}_j\|_2^2 \leq \epsilon_r, x(K,\delta_j)[0] = 0. \end{aligned} \tag{7}$$

Authorized licensed use limited to: Uppsala Universitetsbibliotek. Downloaded on February 19,2023 at 13:27:42 UTC from IEEE Xplore. Restrictions apply.

Next, we derive the upper bound of $q(K,\delta_j)$ in *Lemma 4.1*.

Lemma 4.1: Let $\delta_j \in \Omega$ and $\kappa \triangleq F_p(\cdot)F_r^{-1}(\cdot)$ (Here the arguments of F_p and F_r are dropped for clarity). Let the matrix B be invertible. Then, it holds that

$$q(K,\delta_j) \leq \mu \bar{q}(K,\delta_j)^f,\ \bar{q}(K,\delta_j) \triangleq \eta||K||_F^2 + \sum_{i=2}^{n_x N_h} \sigma_i(\kappa^{-1}),$$

where $f \triangleq n_x N_h - 1, \mu$ is a term independent of K and η is a positive scalar weight on the regularization term.

Proof: See Appendix. ∎

In *Lemma 4.1*, we formulated an upper bound for $q(K,\delta_j)$. However, only the term $\bar{q}(K,\delta_j)$ of the bound is dependent on the variable K. Moreover, since $(\bar{q}(K,\delta_j))^{n_x N_h - 1}$ is a monotonically increasing function on $\bar{q}(K,\delta_j) > 0$, $\bar{q}(K,\delta_j)$ can be replaced as the term to be optimized. Next, we show that $\bar{q}(K,\delta_j)$ is strongly convex in the design variable K.

Theorem 4.2: For any given $\delta_j \in \Omega$, the function $\bar{q}(K,\delta_j)$ is strongly convex in the design variable K.

Proof: See Appendix. ∎

We have shown in this section that the term $\bar{q}(K,\delta_j)$ can be used as the convex proxy objective function for the attack impact $q(K,\delta_j)$. That is, we can recast (6) as

$$K^* = \arg\inf_K \mathbb{E}\{\bar{q}(K,\delta)|\bar{q}(K,\delta) > \text{VaR}_\alpha(\bar{q}(K,\delta))\}. \quad (8)$$

Although (8) is convex, it is computationally intensive due to the expectation operator. In *Section V*, we discuss a method to approximate the expectation operator.

Remark 1: We use Definition 3.1 to formulate (8). Thus (8) implicitly assumes that the distribution of $\bar{q}$ has no point masses. However, verifying this conditions is beyond the scope of this paper and is left for future work.

V. Empirical risk using scenario based approach

The optimization problem (8) is computationally intensive since it involves an expectation operator which acts on a continuum of uncertainties Ω. In this section, we provide a method to approximate the expectation operator using the scenario-based approach [13]. To begin with, let us establish the following:

Assumption 5.1: For any δ, $\bar{q}(\cdot,\delta)$ is a convex function in the design variable K. ◁

We have shown in *Theorem 4.2* that *Assumption 5.1* is satisfied. Next, we approximate the expectation operator in (8) empirically. To do this, let us begin by sampling the uncertainty set Ω with N samples. Let us consider that $(\delta_1,\ldots,\delta_N)$ is a collection of N independent realizations from Ω and let $\Omega_N \triangleq \{1,\ldots,N\}$. Then for any given K, and $i \in \Omega_N$, we denote by $\bar{q}_i(K)$, the value attained by $\bar{q}(K,\delta_i)$, and we denote by $\bar{q}_{(i)}(K)$, the $N-i+1^{th}$ order statistic. That is $\bar{q}_{(1)}(K) \geq \bar{q}_{(2)}(K) \geq \cdots \geq \bar{q}_{(N)}(K)$. Now we present the first result of the section.

Lemma 5.1: Let the dimension of the design variable K be $d = n_x^2$. Let $N \geq d$ and $m \triangleq \lceil N(1-\alpha) \rceil$. Then, under *Assumption 5.1*, the solution to (8) can be obtained empirically by solving the convex optimization problem

$$K^* = \arg\inf_K \frac{1}{m}\sum_{i=1}^{m} \bar{q}_{(i)}(K). \quad (9)$$

Proof: It was shown that (9) is the empirical version of (8) in [13] (See equations (3) and (5) in [13]). ∎

The design problem (9) is formulated using the empirical formulation of the risk metric CVaR. If we could order the functions $\bar{q}_{(i)}(K)$, then the problem of obtaining the optimal controller would be simple. However, $\bar{q}_{(i)}(K)$ is a function of the optimization variable K. Hence it is not possible to explicitly define the order without knowing K beforehand. Alternatively, the design problem (9) can be modified such that the controller is optimal to any possible ordering of the functions. To this end, (9) can be recast as

$$\begin{aligned} \arg\inf_{K,y} \quad & y \\ \text{s.t.} \quad & \frac{1}{m}\sum_{j=1}^{m} \bar{q}_{i_j}(K) \leq y, \\ & \forall \text{choice of } m \text{ indices} \{i_1,\ldots,i_m\} \subseteq \Omega_N. \end{aligned} \quad (10)$$

Since $\{i_1,\ldots,i_m\}$ is any subset of Ω_N with cardinality m, the optimization problem (10) has $\binom{N}{m}$ constraints. To summarize, (10) is the empirical reformulation of (8).

Using the solution obtained from (10), we provide probabilistic out-of-sample certificate on the random variable $\bar{q}(K,\delta), \delta \in \Omega$. To this end, let us represent the optimal controller of (10) as K_N^*, and let $N \geq m+d$. Let us define the out-of-sample *Probability of Shortfall (PS)* as follows:

Definition 5.1 (Probability of Shortfall): PS is defined as

$$PS(K_N^*) \triangleq \mathbf{P}\{\delta \in \Omega \mid \bar{q}(K_N^*,\delta) \geq \bar{q}_{(m+d)}(K_N^*)\}. \quad ◁$$

By ensuring that the $PS(K_N^*)$ is small (say ϵ), one can ensure that the impact proxy under the optimal controller for any new uncertainty drawn from the set Ω, $\bar{q}(K_N^*,\delta)$, exceeds a predefined valued $\bar{q}_{(m+d)}(K_N^*)$ (also called as the *shortfall threshold*) with a small probability ϵ. Now we are ready to present the main result in *Theorem 5.2*.

Theorem 5.2: It holds that $\mathbf{P}^N\{PS(K_N^*) \leq \epsilon\} \triangleq$

$$\int_0^\epsilon \frac{\Gamma(N+1)}{\Gamma(m+d)\Gamma(N+1-m-d)} p^{m+d-1}(1-p)^{N-m-d}dp,$$

where Γ is Euler's Gamma function and $\epsilon \in (0,\ 1)$.

Proof: See Appendix. ∎

Theorem 5.2 provides posteriori results on the confidence with which $PS(K_N^*)$ is below a small threshold ϵ. In other words, *Theorem 5.2* states that, the confidence of the $PS(K_N^*)$ can be evaluated by knowing the dimension of the decision variable (d), the number of samples (N), and m.

To recall, in this section we proposed an empirical version of (8) in (10). We also provided probabilistic guarantees on the out-of-sample PS. We conclude this section by providing **Algorithm 1** which depicts the outline for solving *Problem 1* approximately. In the next section, we depict the efficacy of the proposed algorithm using a numerical example.

VI. Numerical example

In this section, the efficacy of the design algorithm is depicted through a numerical example. Consider the system

Authorized licensed use limited to: Uppsala Universitetsbibliotek. Downloaded on February 19,2023 at 13:27:42 UTC from IEEE Xplore. Restrictions apply.

Initialization: $N_h, k, d, \Omega, \epsilon$;
Step 1: Choose N such that $N \geq m + d$.
Step 2: Extract N independent realizations from Ω.
Step 3: Build the matrix $\kappa^{-1}, \forall i \in \Omega_N$.
Step 4: Solve the convex optimization problem (10).
Step 5: Given ϵ, evaluate the confidence using *Theorem 5.2*.
Result: K_N^*, PS

Algorithm 1: Risk averse design algorithm

TABLE I
RISK AND SHORTFALL THRESHOLD

	$\text{CVaR}_\alpha(\bar{q}(K,\delta) + 0.1\|\|K\|\|_F^2)$	$\bar{q}_{(m+d)}(K_N^*)$
$m = 1$	20.7160	16.9069
$m = 2$	20.6436	16.8910

of the form (2) where $C_J = C = I_3$, A, B and L are

$$\begin{bmatrix} 2 & 0 & 1 \\ 1 & a & 0 \\ 0 & 1 & b \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0.3 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \text{and } \begin{bmatrix} 1.95 & 0 & 1 \\ 1 & 0.36 & 1 \\ 0 & 1 & -0.87 \end{bmatrix},$$

respectively. Here $a \in [0.5\ 1.5]$ and $b \in [-0.5\ 0.5]$ are uncertain parameters. The observer gain L is designed using pole placement method.

Let $N = 11, \epsilon_r = 1$, and $N_h = 5$. The risk and the corresponding shortfall threshold $\bar{q}_{(m+d)}(K_N^*)$ obtained from **Algorithm 1** for different values of m is shown in TABLE I. The corresponding confidence that the the out-of-sample PS is less than or equal to ϵ can be evaluated from the integral in *Theorem 5.2*. In reality, the confidence can be made higher by increasing N or decreasing m.

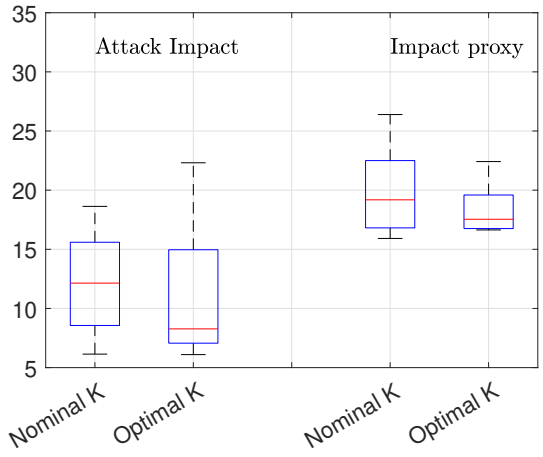

Fig. 2. Evaluation of controller performance: The plot depicts the distribution of the attack impact $q(\cdot)$ and the impact proxy $\bar{q}(\cdot)$ for 100 different uncertainties under the nominal and the optimal controller. The optimal controller is obtained from **Algorithm 1** by optimizing $\text{CVaR}_{0.8}(\bar{q}(\cdot))$ with parameters $N_h = 5, N = 11$, and $m = 2$. The nominal controller is obtained by optimizing $\bar{q}(K, 0)$ for the nominal system. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25^{th} and 75^{th} percentiles, respectively. The whiskers extend to the most extreme data points.

Next, we depict the efficiency of the proposed design framework using Fig. 2. It can be seen that: (a) compared to the nominal controller, the controller obtained from **Algorithm 1** lowers the impact proxy $\bar{q}(\cdot)$); (b) The value of attack impact $(q(\cdot))$ is shifted towards a lower median.

VII. CONCLUSION

The problem of risk-optimal controller design against FDI attacks on actuators of an uncertain control system was studied. We considered an attacker with perfect system knowledge that maximizes the system disruption whilst remaining stealthy. We quantified the system disruption using the OOG. We formulate a design problem where the defender aims at designing a controller such that its CVaR is minimized. We proposed a convex optimization problem that approximately solves the design problem with probabilistic certificates. Finally, we illustrated the results through a numerical example.

APPENDIX

PROOF OF *Lemma 4.1*

Before presenting the proof, we provide an intermediate result which helps to construct the proof.

Lemma A.1 (Inequality of arithmetic and geometric means): Given N real numbers $x_1, \ldots, x_N$, it holds that $\prod_{i=1}^N x_i \leq \left(\frac{\sum_{i=1}^N x_i}{N}\right)^N$

Proof: [Proof of *Lemma 4.1*] The convex dual problem of (7) can be formulated as (11). Furthermore, it was shown in [17, *Theorem 3.1*] that the duality gap is zero.

$$\inf\{\epsilon_r \gamma \mid F_p^T(\cdot)F_p(\cdot) - \gamma F_r^T(\cdot)F_r(\cdot) \preceq 0.\} \tag{11}$$

Pre-multiplying the constraint of (11) by $F_r^{-T}(\delta_j)$ and post-multiplying by $F_r^{-1}(\delta_j)$, (11) can be rewritten as

$$\epsilon_r \inf\{\gamma \mid \kappa^T \kappa \preceq \gamma I.\} \tag{12}$$

Since ϵ_r is a pre-defined constant, it can be moved outside the optimization problem (12). Using the definition of singular values, (12) can be re-written as $\epsilon_r \bar{\sigma}(\kappa)$. Thus we have shown that $q(K, \delta_j) = \epsilon_r \bar{\sigma}(\kappa)$. Next, we prove that $\bar{\sigma}(\kappa) \leq \mu \bar{q}(K, \delta_j)^{n_x N_h - 1}$. To this end, we can show that the matrix $F_r(\cdot)$ is a block lower triangular with the element CB in the leading diagonal. Then $F_r(\delta_j)^{-1}$ will be $\begin{bmatrix} B^{-1}C^{-1} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ * & \ldots & B^{-1}C^{-1} \end{bmatrix}$, where $*$ represents that its value is unimportant for now. Then, matrix κ will be of the form $\begin{bmatrix} C_J C^{-1} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ * & \ldots & C_J C^{-1} \end{bmatrix}$. Since κ is block lower triangular, its determinant is the product of determinant of diagonal blocks [12, Proof of *Theorem A.1*]. Thus $\det(\kappa) = |\det(C_J C^{-1})^{N_h}| \triangleq c$. Since the product of singular values of a matrix is equal to its determinant, we get

$$\bar{\sigma}(\kappa) = \frac{c}{\prod_{i=1}^{n_x N_h - 1} \sigma_i(\kappa)} \overset{1}{=} c \prod_{i=2}^{n_x N_h} \sigma_i(\kappa^{-1}). \tag{13}$$

Authorized licensed use limited to: Uppsala Universitetsbibliotek. Downloaded on February 19,2023 at 13:27:42 UTC from IEEE Xplore. Restrictions apply.

The equality 1 in (13) follows since the singular values of κ and κ^{-1} are reciprocals of each other. By using the result of *Lemma A.1*, the term $\epsilon_r\bar{\sigma}(\kappa)$ can be bounded as

$$\begin{aligned}\epsilon_r\bar{\sigma}(\kappa) &\leq \epsilon_r|\det(C_JC)^{N_h}|\Big(\frac{\sum_{i=2}^{n_xN_h}\sigma_i(\kappa^{-1})}{n_xN_h-1}\Big)^{n_xN_h-1}\\ &= \underbrace{\frac{\epsilon_r|\det(C_JC)^{N_h}|}{(n_xN_h-1)^{n_xN_h-1}}}_{\mu}\Big(\sum_{i=2}^{n_xN_h}\sigma_i(\kappa^{-1})\Big)^{n_xN_h-1}\\ &\leq \mu\underbrace{\Big(\sum_{i=2}^{n_xN_h}\sigma_i(\kappa^{-1})+\eta||K||_F^2\Big)}_{\bar{q}(K,\delta_j)}^{n_xN_h-1} \qquad (14)\end{aligned}$$

where the last inequality follows since μ and $\eta||K||_F^2$ are non-negative terms. In (14), only the term $\bar{q}(K,\delta_j)$ is dependent on the design variable K. This concludes the proof. ■

PROOF OF *Theorem 4.2*

Proof: Since $\kappa = F_p(\cdot)F_r(\cdot)^{-1}$, it follows that $\kappa^{-1} = F_r(K,\delta_j)F_{xa}(K,\delta_j)^{-1}(I_{N_h}\otimes C_J)^{-1}$

$$= (I_{N_h}\otimes C)(F_{ea}(\delta_j)F_{xa}(K,\delta_j)^{-1}+F_{ex}(\delta_j))(I_{N_h}\otimes C_J)^{-1}.$$

Here all the matrices are independent of the design variable K except $F_{xa}(K,\delta_j)^{-1}$. It can be verified by matrix multiplication that $F_{xa}(K,\delta_j)^{-1}$ is of the form

$$\begin{bmatrix} B^{-1} & 0 & 0 & \dots & 0\\ -B^{-1}A_{x,j} & B^{-1} & 0 & \dots & 0\\ 0 & -B^{-1}A_{x,j} & B^{-1} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \dots & B^{-1}\end{bmatrix}$$

The matrix $A_{x,j}$ is affine in K and thus the same holds for κ^{-1}. It can also be shown that the sum of singular values of a matrix, $X \to \sum\sigma(X)$, is convex [18]. Thus we have proven that the term $\sum\sigma(\kappa^{-1})$ is convex in K. Since the regularization term $||K||_F^2$ is strictly convex in K, we then conclude that $\bar{q}_{i_j}(K,\delta)+\eta||K||_F^2$ is a strictly convex function in K which concludes the proof. ■

PROOF OF *Theorem 5.2*

Before presenting the proof, we provide an intermediate result: *Theorem A.2*, which follows from applying [13, Theorem 3.1] to our problem setup.

Theorem A.2: Let us suppose that (a) a solution to (10) exists and it is unique almost surely, and (b) for a sample $(\delta_1,\dots,\delta_N)$ of independent realizations from Ω, the event that $\{\ \exists\ K|\bar{q}_i(K), \forall\ i \in \Omega_N$ has the same value$\}$ has zero probability. Then $\mathbf{P}^N\{PS(K_N^*) \leq \epsilon\}$ has a Beta$(m+d, N+1-m-d)$ distribution.

That is, *Theorem A.2* states that the result of *Theorem 5.2* follows if (a) and (b) hold, which we prove next.

Proof: **Proof that** (a) **holds:** We intend to show that the (10) has a unique solution. However, as discussed in the paper, (10) is simply a reformulation of (9). Thus, we can equivalently show that the solution to (9) is unique.

In *Theorem 4.2*, we have shown that the term $\bar{q}_{i_j}(K)$ is strongly convex function in K. The uniqueness of the optimal solution, K_N^*, then follows from the strict convexity of the objective function.

Proof that (b) **holds:** We prove by contradiction. Let us assume that $\exists K, l$ and independent samples $(\delta_1,\dots,\delta_N)$ such that $\mathbf{P}^N\{\bar{q}_i(K) = l, i = 1,2,\dots,N\} = \theta \geq 0$. Then by definition, $\exists\ \delta_j$ such that $\mathbf{P}\{\delta_j|\bar{q}_j(K) = l\} \neq 0$. However this contradicts the assumption that the distribution of $\bar{q}(\cdot)$ has no point masses. This concludes the proof. ■

REFERENCES

[1] R. Langner, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Security & Privacy*, vol. 9, no. 3, pp. 49–51, 2011.

[2] A. Cherepanov and R. Lipovsky, "Industroyer: Biggest threat to industrial control systems since stuxnet," *WeLiveSecurity, ESET*, vol. 12, 2017.

[3] D. U. Case, "Analysis of the cyber attack on the ukrainian power grid," *Electricity Inf. Sharing and Analy. Center (E-ISAC)*, vol. 388, 2016.

[4] S. M. Dibaji, M. Pirani, D. B. Flamholz, A. M. Annaswamy, K. H. Johansson, and A. Chakrabortty, "A systems and control perspective of cps security," *Annu. Reviews in Contr.*, vol. 47, pp. 394–411, 2019.

[5] P. Griffioen, S. Weerakkody, B. Sinopoli, O. Ozel, and Y. Mo, "A tutorial on detecting security attacks on cyber-physical systems," in *2019 18th Eur. Contr. Conf. (ECC)*, pp. 979–984, IEEE, 2019.

[6] M. I. Müller, J. Milošević, H. Sandberg, and C. R. Rojas, "A risk-theoretical approach to $\mathcal{H}_2$-optimal control under covert attacks," in *2018 IEEE Conf. on Decision and Contr. (CDC)*, pp. 4553–4558, IEEE, 2018.

[7] C. Murguia, I. Shames, J. Ruths, and D. Nešić, "Security metrics and synthesis of secure control systems," *Automatica*, vol. 115, p. 108757, 2020.

[8] N. Hashemi and J. Ruths, "Gain design via LMIs to minimize the impact of stealthy attacks," in *2020 Am. Contr. Conf. (ACC)*, pp. 1274–1279, IEEE, 2020.

[9] S. C. Anand and A. M. Teixeira, "Joint controller and detector design against data injection attacks on actuators," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 7439–7445, 2020.

[10] S. D. Bopardikar, A. Speranzon, and J. P. Hespanha, "An H_∞ approach to stealth-resilient control design," in *2016 Resilience Week (RWS)*, pp. 56–61, IEEE, 2016.

[11] A. Teixeira, H. Sandberg, and K. H. Johansson, "Strategic stealthy attacks: the output-to-output ℓ_2-gain," in *2015 54th IEEE Conf. on Decision and Contr. (CDC)*, pp. 2582–2587, IEEE, 2015.

[12] K. Dvijotham, E. Todorov, and M. Fazel, "Convex structured controller design in finite horizon," *IEEE Trans. on Contr. of Network Systems*, vol. 2, no. 1, pp. 1–10, 2014.

[13] F. A. Ramponi and M. C. Campi, "Expected shortfall: Heuristics and certificates," *Eur. Journal of Operational Research*, vol. 267, no. 3, pp. 1003–1013, 2018.

[14] M. C. Campi and S. Garatti, *Introduction to the scenario approach*. SIAM, 2018.

[15] R. T. Rockafellar, S. Uryasev, *et al.*, "Optimization of conditional value-at-risk," *Journal of risk*, vol. 2, pp. 21–42, 2000.

[16] R. T. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of banking & finance*, vol. 26, no. 7, pp. 1443–1471, 2002.

[17] S. C. Anand and A. M. H. Teixeira, "Stealthy cyber-attack design using dynamic programming," in *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 3474–3479, 2021.

[18] S. Subramani, *On sums of singular values*. University of New South Wales, 1993.

Authorized licensed use limited to: Uppsala Universitetsbibliotek. Downloaded on February 19,2023 at 13:27:42 UTC from IEEE Xplore. Restrictions apply.