A Moving Target Defense Control Framework for Cyber-Physical Systems

Aris Kanellopoulos, *Student Member, IEEE*, and Kyriakos G. Vamvoudakis, *Senior Member, IEEE*

***Abstract*—This paper considers the problem of efficiently and securely controlling cyber-physical systems that are operating in uncertain, and adversarial environments. To mitigate sensor, actuator attacks, and performance loss due to such attacks, we formulate a secure control algorithm that consists of a proactive and a reactive defense mechanism. The proactive mechanism, which is based on the principles of moving target defense, utilizes a stochastic switching structure to dynamically and continuously alter the parameters of the system, while hindering the attacker's ability to conduct successful reconnaissance to the system. The unpredictability of the current actuator and sensor is optimized using an information entropy measure, which is induced by probabilistic switching. The reactive mechanism on the other side, detects potentially attacked components, namely sensors and actuators, by leveraging online data to compute an integral Bellman error. A rigorous mathematical framework is presented to guarantee the stability of the equilibrium point of the closed-loop system, and provide a quantified bound on the performance loss when utilizing both reactive and proactive mechanisms. Simulation results show the efficacy of the proposed approaches on a benchmark aircraft model.**

***Index Terms*—Cyber-physical systems (CPSs), entropy, moving target and proactive defense, reactive defense, security.**

I. Introduction

CYBER-PHYSICAL systems (CPSs) are complex platforms comprised of a physical layer, containing sensing and actuating devices, as well as communication and computational layers [1]. Such systems can be found in a number of areas ranging from military to civilian applications, namely healthcare and medicine [2], smart grids [3], [4], and transportation [5]. Due to the complex, and often large-scale nature of CPS, there is a plethora of attack angles that can be exploited by potential malicious agents/components. Numerous attacks on CPS have been reported, e.g., the Stuxnet virus, a malicious computer worm targeting programmable logic controllers [6] or the attack on the Maroochy water services in Australia [7]. Also, more complex attacks have been reported, e.g., the simultaneous communication jamming and GPS spoofing of a military U.S. drone [8]. In order to ensure the integration of CPS in our society, there is a need for robust defense mechanisms to counteract such malicious attacks.

Moving target defense (MTD) [9] is a defense paradigm, which aims to minimize the inherent advantage the attacker has, over the defender. While the security measures employed by the system's defender have to monitor all the vulnerable components at all times and mitigate against all kinds of attack approaches, the attacker herself may need to bypass those defenses only once. Moreover, most CPS operate statically with respect to their structure, goals, and constraints. Such vulnerabilities, offer to a persistent attacker the necessary time to exploit the system and develop appropriate strategies. MTD protocols aim to tackle this asymmetry by developing mechanisms that continually and unpredictably change the parameters of the system. Such unpredictability has three goals: to increase the cost of attacking; to limit the exposure of vulnerable components; and to deceit the opponent.

A. Related Work

The work in [10] questions the adequacy of the security approaches that operate only in the computational layer, such as encryption algorithms. Therefore, extensive research has been conducted on the behavior and security of complex CPS from a control-theoretic standpoint [11]–[13]. Furthermore, by leveraging models that are common in control theory, such as dynamical systems, we are able to better exploit the interconnection between the input and the output of a given system, which is often leveraged in CPS attacks. This has been addressed in [14] and is a valuable tool in defending against attacks such as in drone spoofing.

Among the different design approaches, optimal control, and game theory [15] have emerged as important frameworks due to their abilities to satisfy user-defined performances in the presence of cooperating and noncooperating agents. Mathematically, optimal feedback policies are computed by solving the socalled Hamilton–Jacobi–Bellman (HJB) equation. In [16], security problems are formulated as zero-sum games between attacking and defending agents. In [17], a graphical game is solved on a complex multiagent network under persistent adversaries.

Manuscript received October 25, 2018; revised March 19, 2019; accepted April 28, 2019. Date of publication May 9, 2019; date of current version February 27, 2020. This work was supported in part by ONR Minerva under Grant N00014-18-1-2160, in part by the NSF CAREER under Grant CPS-1851588, in part by the Army Research Office (ARO) under Grant W911NF-19-1-0270, and in part by the Department of Energy under Grant DE-EE0008453. Recommended by Associate Editor Prof. H. Lin. *(Corresponding author: Aris Kanellopoulos.)*

The authors are with the Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: ariskan@gatech.edu; kyriakos@gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2019.2915746

These approaches seek to mitigate the attacker's influence rather than dissuade her from attacking. Theodorakopoulos and Baras [18] defined trust metrics to analyze the interconnections between the agents of a network.

The research on MTD has mostly focused on its application to computer networks [19]–[21]. In [22], the authors apply the principles of MTD to constantly rotating Internet Protocol version 6 addresses. In [23], a proactive defense strategy was formulated to deceit an attacker targeting nodes in a wireless network. In the context of CPS, those approaches can be employed in the computational and communication layers of the system. In this paper, we leverage control-theoretic tools to achieve a system-wide proactive and reactive defense that, to the best of our knowledge, has not been addressed before. A more formalized approach to MTD was introduced in [24], leading to an MTD entropy hypothesis framework that is generally applicable. In [25], a multilayer zero-sum game was formulated between an attacker trying to maximize the damage to the system, and a defender randomizing over different configurations of the system, without considering the continuous-time dynamics due to the physics of the system. An MTD approach was used to enlarge the dimension of the state space in [26] for the purposes of attack detection, rather than proactive defense based on an unpredictability measure.

Vamvoudakis *et al.* [27] focused on the estimation problem of a binary variable in a network of sensors under Byzantine attacks, with no consideration of the system's dynamics. Fawzi *et al.* [28] showed that if the attacker is able to compromise less than half of the sensors, it is always possible to recover the state information. In order to relax this assumption, several switching-based schemes have been developed that, rather than guaranteeing robustness of estimation or operation under attack, they opt to identify the attacked components and take them offline. Following this line of research, an attack detection filter and a passivity-based switching mechanism were introduced in [29] and [30], where explicit knowledge of the dynamics was required for the detection mechanism, and the switching structure was utilized in a reactive fashion, without the advantages offered by a proactive MTD mechanism.

1) Contribution: The contributions of this paper are fourfold. First, we model the attacker's effect as a time-varying and unknown, but integrable degradation parameter. Then, multiple controllers and observers are designed for every admissible combination of actuators and sensors. Third, we use a probabilistic switching rule based on the entropy hypothesis to design a structure that offers proactive defense properties to the system. Moreover, we propose a performance evaluator based on the integral Bellman error of the closed-loop system to detect compromised actuators and sensors, and remove them from the switching queue. Finally, we show that the system under unpredictable switching, of either the actuating or the sensing components, has an asymptotically stable equilibrium point with a quantified dwell time and the performance and we present simulation results that highlight the operation of our approach as well as the tradeoff between optimality and security.

2) Structure: The remainder of this paper is structured as follows. Section II formulates the problem of defending a CPS from actuator and sensor attacks while also increasing the attacking surface to enhance uncertainty and unpredictability. In Section III, we focus on proactive and reactive defense against actuator attacks. Section IV extends the framework of Section III, to incorporate a proactive and reactive defense framework against sensor attacks. Simulation results are shown in Section V. Finally, Section VI concludes and discusses future work.

3) Notation: The notation used here is standard. $\bar{\lambda}(A)$ is the maximum eigenvalue of the matrix A and $\underline{\lambda}(A)$ is its minimum eigenvalue. $\|\cdot\|$ denotes the Euclidean norm of a vector and the Frobenius norm of a matrix. The superscript $\star$ is used to denote the optimal trajectories of a variable. $(\cdot)^{\mathrm{T}}$ denotes the transpose of a matrix. ∇_x and $\frac{\partial}{\partial x}$ are used interchangeably and denote the partial derivative with respect to a vector x. The cardinality of a set, i.e., the number of elements contained in the set, is denoted by $\mathrm{card}(\cdot)$. 2^A denotes the power set of a set A, i.e., the set containing all the subsets of A, including the empty set and A itself. Finally, $\mathrm{supp}(x)$ denotes the support of a vector, i.e., the number of its nonzero elements.

II. Problem Formulation

Consider the following linear time-invariant continuous-time system

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu_a(t),\ t \geqslant 0\\ y(t) &= C_a(t)x(t)\end{aligned} \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u_a(t) \in \mathbb{R}^m$ is the potentially attacked input of the system, $y(t) \in \mathbb{R}^p$ is the output, $A \in \mathbb{R}^{n\times n}$ is the plant matrix, $B \in \mathbb{R}^{n\times m}$ is the input matrix, and $C_a(t) \in \mathbb{R}^{p\times n}$ is the potentially attacked output matrix.

We can rewrite (1) as

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^{m} b_i u_i(t)$$

$$y_j(t) = c_j(t)x(t)\,,\ j \in \{1,\ldots,p\}$$

where b_i is a column vector corresponding to the ith actuator, u_i is the value of the input signal associated with this actuator, and y_j is the output given by a specific sensor c_j corresponding to the jth row of the output matrix.

The potentially compromised control input of (1) will be of the following form

$$u_a(t) = \rho(t)u(t),\ t \geqslant 0 \tag{2}$$

where $\rho(t) = \mathrm{diag}(\rho_{ii}(t))$, $\forall i \in \{1,\ldots,m\}$ is a time-varying actuator attack parameter controlled by an adversary and $u(t) \in \mathbb{R}^m$ is the nonattacked control input.

The output matrix of the system can be undermined by a signal $\rho^s(t)$ as

$$C_a(t) = \rho^s(t)C,\ t \geqslant 0 \tag{3}$$

where $\rho^s(t)$ is a diagonal matrix controlled by the attacker and $C \in \mathbb{R}^{p\times n}$ is the nonattacked output matrix.

Remark 1: Note that the focus of this paper is on the components of the CPS that can be modeled utilizing control theoretic techniques. Although there are attack angles that can affect the

software, which implements the proposed intrusion detection algorithms, those lie beyond the scope of our research. It is assumed that the computing elements are equipped with appropriate security measures, such as encryption mechanisms [31]. On the other hand, our approach considers attacks that leverage the dual nature of CPS, hence, we develop methods that take into account the cyber components that interact with the physics of the systems, i.e., sensors and actuators. □

Assumption 1: In order to offer a greater degree of freedom for deception purposes and to mitigate the effect of potential attacks, we will consider systems with redundant actuating and sensing components. □

Assumption 2: We will assume that the system's actuators are not compromised over a time interval $\tau \in [t_1, t_2]$ if and only if $\rho_{ii}(\tau) = 1$, $\forall i \in \{1, \ldots, m\}$, $\forall \tau$. Similarly, we consider the sensors as secure, if and only if $\rho^s_{jj}(\tau) = 1$, $\forall j \in \{1, \ldots, p\}$ $\forall \tau$. The signals (2) and (3) are assumed to be locally integrable over any closed time interval $[t_1, t_2]$, $0 \leqslant t_1 < t_2$. □

Remark 2: The assumption on the local integrability of the adversarial signals allows us to take into account a variety of realistic attack scenarios, such as impulses and other discontinuous signals, or constant bias injection, which is locally, but not globally, integrable. The underlying restriction on the signal excludes attacks that have infinite value on a specific time interval, which is a practical assumption on the adversarial capabilities. □

Assumption 3: We will assume that the attacker is not able to compromise all of the actuators and sensors at once. Therefore, $\text{supp}(\rho) < m$ and $\text{supp}(\rho^s) < p$. □

Remark 3: It should be noted that our formulation will make no assumptions on the structure, boundedness, and other Lipschitz continuity properties of the attacker's signal. Furthermore, attacks of the form (2) and (3), due to their time-varying nature, can describe a wide range of attacks, including additive and multiplicative attacks. □

We are, thus, interested in designing a proactive and a reactive defense mechanism that will operate well in the absence of attackers, and will detect and mitigate attacks while guaranteeing closed-loop stability of the equilibrium point.

III. Defense Against Actuator Attacks

We will initially focus our attention to the case of actuator attacks. We note that throughout this section, full state feedback is assumed. Let $\mathcal{B}$ denote the set containing the actuators of (1) by the vectors b_i, $i \in \{1, \ldots, m\}$. The power set of $\mathcal{B}$, denoted as $2^{\mathcal{B}}$, contains all possible combinations of the actuators acting on (1). Each of these combinations is expressed by the input matrix B_j, $j \in \{1, \ldots, 2^m\}$ whose columns are the appropriate vectors b_i.

The set of the candidate actuating modes $\mathcal{B}_c$ is defined as the set of the actuator combinations that renders system (1) fully controllable, i.e.,

$$\mathcal{B}_c = \left\{ B_j \in 2^{\mathcal{B}} : \text{rank}\left(\begin{bmatrix} B_j & AB_j & \cdots & A^{n-1}B_j \end{bmatrix}\right) = n \right\}. \quad (4)$$

System (1) assuming full state-feedback, with the actuating mode B_i can be rewritten as

$$\dot{x} = Ax + B_i u_i, \; i \in \{1, \ldots, \text{card}(\mathcal{B}_c)\}, \; t \geqslant 0. \quad (5)$$

Remark 4: Note that, we do not require different actuating modes to share common actuators. Moreover, while a single actuating mechanism might be able to control a system, two different—less potent—mechanisms might need to work cooperatively to control the same system. All these modes will belong to the set described in (4). □

A. Optimal Controllers Design

For each actuating operating mode B_i, $i \in \{1, \ldots, \text{card}(\mathcal{B}_c)\}$, we denote the candidate control law as $u_i(t)$.

We are interested in deriving optimal controllers for each of these modes by utilizing well-known optimal control approaches [32]. Toward that, we are interested in solving the following optimization:

$$V_i^\star(x(t_0)) = \min_{u_i} \int_{t_0}^{\infty} r_i(x, u_i) \mathrm{d}\tau \equiv \min_{u_i} \int_{t_0}^{\infty} (x^{\mathrm{T}} Q_i x + u_i^{\mathrm{T}} R_i u_i) \mathrm{d}\tau, \quad \forall x(t_0) \quad (6)$$

given (5), where $Q_i \succeq 0$, $R_i \succ 0$, $\forall i \in \{1, \ldots, \text{card}(\mathcal{B}_c)\}$.

Assumption 4: We assume that each pair $(A, \sqrt{Q_i})$ is detectable. □

The Hamiltonian associated with (5) and (6) is

$$H_i(x, u_i, \nabla V_i) = \nabla V_i^{\mathrm{T}}(Ax + B_i u_i) + x^{\mathrm{T}} Q_i x + u_i^{\mathrm{T}} R_i u_i, \quad \forall x, u_i$$

with V_i denoting the value function, not necessarily the optimal.

Applying the stationarity conditions $\frac{\partial H_i(x, u_i, \nabla V_i)}{\partial u_i} = 0$ yields

$$u_i = -R_i^{-1} B_i^{\mathrm{T}} \nabla V_i. \quad (7)$$

The optimal value functions $V_i^\star(\cdot)$ must satisfy the following HJB equation

$$x^{\mathrm{T}} Q_i x + \nabla V_i^{\star \mathrm{T}} Ax - \frac{1}{2} \nabla V_i^{\star \mathrm{T}} B_i R_i^{-1} B_i^{\mathrm{T}} \nabla V_i^\star = 0. \quad (8)$$

Since all the systems described by (5) are linear and the cost given by (6) is quadratic, all the value functions will be quadratic in the state x, i.e., $V_i^\star(x) = x^{\mathrm{T}} P_i x$, $P_i \succ 0$. Substituting this expression into (8) and the resulting optimal value function into (7) yields the feedback controller with optimal gain K_i

$$u_i^\star(x) = -K_i x := -R_i^{-1} B_i^{\mathrm{T}} P_i x, \quad \forall x$$

where P_i are the solutions to the following Riccati equations

$$A^{\mathrm{T}} P_i + P_i A - P_i B_i^{\mathrm{T}} R_i^{-1} B_i^{\mathrm{T}} P_i + Q_i = 0. \quad (9)$$

We introduce $\mathcal{K}$ the set containing all K_i, $i \in \{1, \ldots, m\}$, with the understanding that $\text{card}(\mathcal{K}) = \text{card}(\mathcal{B}_c)$. For the ease of exposition, with some abuse of notation, we will consider K_i to mean the optimal controller with this gain as well as its corresponding index.

Fact 1: Due to (4) and Assumption 4, for each B_i, the solution exists and is unique. □

Fact 2: Each K_i, with input given by (7) guarantees that (1) has an asymptotically stable equilibrium point. □

B. Switching-Based MTD Framework

We will now develop a framework to facilitate the deception of potential attackers based on the principles of MTD.

1) Maximization of Unpredictability: To formally define the switching law, we need to introduce a *probability simplex* $\mathbf{p}$, which denotes the probability that each controller K_i is active.

To incorporate ideas from the framework of MTD, we propose a switching rule that optimizes over the minimum cost that each controller is able to attain, as well as an unpredictability term quantified by the information entropy produced by the switching probability simplex $\mathbf{p}$. This way, we will achieve the desired tradeoff between overall optimality and unpredictability. The use of the information entropy is a standard practice in MTD design [33].

Theorem 1: Suppose that (1) is controlled by $N = \text{card}(\mathcal{K})$ candidate controllers with an associated cost given by (6). Then, the probability p_i that each controller K_i is active is given by

$$p_i = e^{\left(-\frac{V_i^\star}{\epsilon} - 1 - \epsilon \log\left(e^{-1}\sum_{i=1}^{N} e^{\frac{V_i^\star}{\epsilon}}\right)\right)} \tag{10}$$

with $\epsilon \in \mathbb{R}^+$ denoting the weight on unpredictability during the optimization process.

Proof: We formulate the following optimization problem:

$$\min_{\mathbf{p}} \left(\mathbf{V}^{\star\mathrm{T}}\mathbf{p} - \epsilon\mathcal{H}(\mathbf{p})\right)$$
$$\text{subject to: } \|\mathbf{p}\|_1 = 1 \text{ and } \mathbf{p} \succeq 0$$

where $\mathbf{V}^\star := [V_1^\star \ \cdots \ V_N^\star]^\mathrm{T} = [x(t_0)^\mathrm{T} P_1 x(t_0) \cdots x(t_0)^\mathrm{T} P_N x(t_0)]^\mathrm{T}$ denotes a column vector containing the value function of each candidate controller and $\mathcal{H}(\mathbf{p}) = -\mathbf{p}^\mathrm{T}\log(\mathbf{p})$ is the information entropy produced by the simplex.

Remark 5: The choice of this particular objective function allows us to combine the two required specifications. The linear term $\mathbf{V}^{\star\mathrm{T}}\mathbf{p}$ penalizes the deviations from the overall optimal controller, while the entropy term $\mathcal{H}(\mathbf{p})$ penalizes the use of a single controller throughout the operation of the system. The result is a compromise that is specified by the optimization weight ϵ. □

Furthermore, for the decision vector $\mathbf{p}$ to constitute a probability simplex, we constrain it to the nonnegative orthant (i.e., $p_i \geqslant 0$, $\forall i \in \{1, \ldots, N\}$) and we require its l_1 norm to satisfy, $\|p\|_1 = \sum_{i=1}^{N} \|p_i\| = 1$.

The entropy of a probability is a concave function [34], and therefore, the cost index, being a sum of a linear function of the probability and the negative entropy, is convex. Thus, we can define the Lagrangian of the optimization problem as

$$\begin{aligned} L &= \mathbf{V}^{\star\mathrm{T}}\mathbf{p} - \epsilon\mathcal{H}(\mathbf{p}) + \lambda(\mathbf{1}^\mathrm{T}\mathbf{p} - 1) + \boldsymbol{\beta}^\mathrm{T}\mathbf{p} \\ &= \mathbf{V}^{\star\mathrm{T}}\mathbf{p} + \epsilon\mathbf{p}^\mathrm{T}\log(\mathbf{p}) + \lambda(\mathbf{1}^\mathrm{T}\mathbf{p} - 1) + \boldsymbol{\beta}^\mathrm{T}\mathbf{p} \end{aligned}$$

where $\mathbf{1}$ denotes a vector consisting of ones and λ, $\boldsymbol{\beta}$ are the Karush–Kuhn–Tucker (KKT) multipliers.

The KKT conditions for the problem are

$$\nabla_\mathbf{p} L = \mathbf{V}^\star + \epsilon\mathbf{1} + \epsilon\log(\mathbf{p}) + \lambda\mathbf{1} + \boldsymbol{\beta}$$

and the complementarity conditions for the optimal solution $\mathbf{p}^\star$ are

$$\boldsymbol{\beta}^\mathrm{T}\mathbf{p}^\star = 0.$$

If there exists an i for which $p_i = 0$, then the term $\log(p_i)$ will be undefined. Consequently, for the optimization problem to be feasible, one of the following two conditions need to hold:

1) $\epsilon\log(p_i) = 0, \forall i \Rightarrow \epsilon = 0 \Rightarrow \mathbf{p}^\star = [\mathbf{0}_{i-1}, \ldots, 1, \ldots, \mathbf{0}_{N-i}]^\mathrm{T}$ where the K_i controller is the one with an overall less cost; and
2) $\boldsymbol{\beta} = 0$.

Consider now the nontrivial case, i.e., $\boldsymbol{\beta} = 0$, which yields

$$\nabla_\mathbf{p} L = \mathbf{V}^\star + \epsilon\log(\mathbf{p}) + \epsilon\mathbf{1} + \lambda\mathbf{1} = 0.$$

The N equations for each controller are independent, leading to the following system of equations:

$$V_i^\star + \epsilon\log(p_i) + \epsilon + \lambda = 0, \quad \forall i \in \{1, \ldots, N\}.$$

Solving now for the optimal probabilities p_i yields

$$p_i = e^{\left(-\frac{V_i^\star}{\epsilon} - \frac{\lambda}{\epsilon} - 1\right)}, \quad \forall i \in \{1, \ldots, N\}. \tag{11}$$

Taking into account that

$$\|\mathbf{p}\|_1 = 1 \Rightarrow \sum_{i=1}^{N} p_i = 1 \Rightarrow \sum_{i=1}^{N} e^{\left(-\frac{V_i^\star}{\epsilon} - \frac{\lambda}{\epsilon} - 1\right)} = 1$$

and solving for λ yields

$$\lambda = \epsilon\log\left(e^{-1}\sum_{i=1}^{N} e^{\left(-\frac{V_i^\star}{\epsilon}\right)}\right). \tag{12}$$

Substituting (12) in (11) provides the required result. ■

2) Switching-Based MTD Scheme: In order to analyze the behavior of the system under the proposed MTD framework, we shall formulate a switched system consisting of the different operating modes.

First, we introduce the switching signal $\sigma(t) = i$, $i \in \{1, \ldots, \text{card}(\mathcal{K})\}$, which denotes the active controller as a function of time. This way, the system is

$$\dot{x}(t) = \tilde{A}_{\sigma(t)}x(t) \tag{13}$$

where $\tilde{A}_{\sigma(t)} := A - B_{\sigma(t)}R_{\sigma(t)}^{-1}B_{\sigma(t)}^\mathrm{T}P_{\sigma(t)}$ denotes the closed-loop subsystem with the controller $K_{\sigma(t)}$ active.

Remark 6: Since the actual switching sequence is different under the designer's choice for unpredictability, we will constrain the switching signal to have a predefined average dwell time. This way, the stability of the overall system will be independent of the result of the optimization. Intuitively, as was initial shown in [35], a system with stable subsystems is stable if the switching is slow enough on an average sense. □

Definition 1: A switching signal has an average dwell time τ_D if over any time-interval $[t, T]$, $T \geqslant t$, the number of switches $S(T, t)$ is bounded above as

$$S(T, t) \leqslant S_0 + \frac{T - t}{\tau_D}$$

where S_0 is an arbitrary chatter bound and τ_D is the dwell time. □

Theorem 2: Consider system (1) in the absence of attacks. The switched system defined by the piecewise continuous switching signal $\sigma(t) = i,\ i \in \{1, \ldots, \text{card}(\mathcal{K})\}$, with active controller K_i given by (7) and continuous flow given by (5) has an asymptotically stable equilibrium point for every switching signal $\sigma(t)$ if the average dwell time is bounded by

$$\tau_D > \frac{\log\left(\max_{q,p\in\{1,\ldots,\text{card}(\mathcal{K})\}} \frac{\bar{\lambda}(P_p)}{\underline{\lambda}(P_q)}\right)}{\min_{p\in\{1,\ldots,\text{card}(\mathcal{K})\}} \frac{\underline{\lambda}(Q_p + P_p B_p R_p^{-1} B_p^{\mathrm{T}} P_p)}{\underline{\lambda}(P_p)}} \tag{14}$$

with an arbitrary chatter bound $S_0 > 0$.

Proof: For each $i \in \{1, \ldots, \text{card}(\mathcal{K})\}$ following [36], we choose the Lyapunov function for each subsystem

$$\mathcal{V}_i(x) = x^{\mathrm{T}} P_i x, \quad \forall x$$

where P_i is the solution to the Riccati equation (9). The Lyapunov functions are positive definite and radially unbounded $\forall x$.

According to the Rayleigh–Ritz inequality for symmetric matrices, one has

$$\underline{\lambda}(P_i)\|x\|^2 \leqslant x^{\mathrm{T}} P_i x = \mathcal{V}_i(x) \leqslant \bar{\lambda}(P_i)\|x\|^2. \tag{15}$$

The time derivative of $\mathcal{V}_i(x)$ along the solutions of the trajectory of the corresponding subsystem is

$$\begin{aligned}\dot{\mathcal{V}}_i(x) &= \dot{x}^{\mathrm{T}} P_i x + x^{\mathrm{T}} P_i \dot{x} \\ &= x^{\mathrm{T}}(A - B_i R_i^{-1} B_i^{\mathrm{T}} P_i)^{\mathrm{T}} P_i x \\ &\quad + xP(A - B_i R_i^{-1} B_i^{\mathrm{T}} P_i)x \\ &= x^{\mathrm{T}}(A^{\mathrm{T}} P_i - P_i B_i R_i^{-1} B_i^{\mathrm{T}} P_i \\ &\quad + P_i A - P_i B_i R_i^{-1} B_i^{\mathrm{T}} P_i)x.\end{aligned}$$

Taking into account (9) and denoting $\bar{H}_i := Q_i + P_i B_i R_i^{-1} B_i^{\mathrm{T}} P_i \succ 0,\ \forall i \in \{1, \ldots, \text{card}(\mathcal{K})\}$ yield

$$\dot{\mathcal{V}}_i(x) = -x^{\mathrm{T}} \bar{H}_i x.$$

Consequently, it holds that

$$\dot{\mathcal{V}}_i(x) \leqslant -\underline{\lambda}(\bar{H}_i)\|x\|^2. \tag{16}$$

Combining now (16) with (15) and noting that $\underline{\lambda}(P_i)\|x\|^2 \leqslant \mathcal{V}_i(x) \Rightarrow \|x\|^2 \leqslant \frac{1}{\underline{\lambda}(P_i)}\mathcal{V}_i(x)$ yield

$$\dot{\mathcal{V}}_i(x) \leqslant -\frac{\underline{\lambda}(\bar{H}_i)}{\underline{\lambda}(P_i)}\mathcal{V}_i(x). \tag{17}$$

For the inequality to hold for arbitrary modes, we have

$$\dot{\mathcal{V}}_i(x) \leqslant -\min_{i\in\{1,\ldots,\text{card}(\mathcal{K})\}} \frac{\underline{\lambda}(\bar{H}_i)}{\underline{\lambda}(P_i)}\mathcal{V}_i(x).$$

Following similar arguments, we show that it holds $\forall p, q \in \{1, \ldots, \text{card}(\mathcal{K})\}$

$$\mathcal{V}_p(x) \leqslant \frac{\bar{\lambda}(P_p)}{\underline{\lambda}(P_q)}\mathcal{V}_q(x).$$

For the inequality to hold for arbitrary pairs of modes, we further write

$$\mathcal{V}_p(x) \leqslant \max_{p,q\in\{1,\ldots,\text{card}(\mathcal{K})\}} \frac{\bar{\lambda}(P_p)}{\underline{\lambda}(P_q)}\mathcal{V}_q(x).$$

For ease of exposition, we denote $\nu := \min_{i\in\{1,\ldots,\text{card}(\mathcal{K})\}} \frac{\underline{\lambda}(\bar{H}_i)}{\underline{\lambda}(P_i)}$ and $\mu := \max_{p,q\in\{1,\ldots,\text{card}(\mathcal{K})\}} \frac{\bar{\lambda}(P_p)}{\underline{\lambda}(P_q)}$. Without loss of generality, we will consider that the switched system is evolving on the time interval $[0, t_f]$. Denote as $S(t_f, 0)$, the number of switches over this interval, which takes place at times $t_i,\ i \in [0, S(t_f, 0)]$ with $t_i < t_{i+1}$. The active mode will be the same over any interval $[t_i, t_{i+1}]$, i.e., the switching signal $\sigma(t) = i$ is piecewise constant.

Define the function

$$W(t) = e^{-\nu t}\mathcal{V}_{\sigma(t)}(x(t)). \tag{18}$$

Along the solutions of the switched system (13) over an interval $t \in [t_i, t_{i+1}]$, the time derivative of (18) is

$$\dot{W} = -\nu W + e^{-\nu t}\dot{V}_{\sigma(t)}(x(t))$$

which is nonpositive due to (17). Consequently, the function $W(t)$ is a nonincreasing function $\forall t \in [t_i, t_{i+1}]$.

At the jump instances t_i one has

$$\begin{aligned}W(t_{i+1}) &= e^{-\nu t_{i+1}}\mathcal{V}_{\sigma(t_{i+1})}(x(t_{i+1})) \\ &\leqslant \mu e^{-\nu t_{i+1}}\mathcal{V}_{\sigma(t_{i+1})}(x(t_{i+1})) \\ W(t_{i+1}) &\leqslant \mu e^{-\nu t_i}\mathcal{V}_{\sigma(t_i)}(x(t_i)) = \mu W(t_i)\end{aligned} \tag{19}$$

where we used the nonincreasing property of $W(t)$.

Over the whole interval $[0, t_f]$, by iterating (19) over the $S(0, t_f) - 1$ discontinuities yields

$$\begin{aligned}&W(t_f-) \leqslant \mu^{S[0,t_f]} W(0) \Rightarrow \\ &e^{\nu t_f}\mathcal{V}_{\sigma(t_f-)}(x(t_f)) \leqslant \mu^{S[0,t_f]} e^{\nu 0}\mathcal{V}_{\sigma(0)}(x(0)) \Rightarrow \\ &e^{\nu t_f}\mathcal{V}_{\sigma(t_f-)}(x(t_f)) \leqslant \mu^{S[0,t_f]}\mathcal{V}_{\sigma(0)}(x(0)).\end{aligned} \tag{20}$$

We can rewrite now (20) as

$$\mathcal{V}_{\sigma(t_f-)}(x(t_f)) \leqslant e^{S_0 \log\mu} e^{\left(\frac{\log\mu}{\tau_D} - \nu\right)t_f}\mathcal{V}_{\sigma(0)}(x(0)).$$

It is clear that choosing τ_D in a way that satisfies the bound (14), the exponential terms are such that $\mathcal{V}_{\sigma(t_f-)}(x(t_f)) \to 0$ as $t_f \to \infty$. Due to (15), we can conclude that $x(t_f) \to 0$, which is the required result. ■

C. Integral Bellman-Based Intrusion Detection Mechanism

In this section, an intrusion detection mechanism is designed to identify the potentially corrupted sets of controllers that belong to the set $\mathcal{K}$. The attack detection signal will rely on the optimality property as well as on data measured along the—possibly corrupted—trajectories of the system. Based on a sampling mechanism, we denote the measurements of the state at the sampling instances as $x_c(t)$ and define the functions $\hat{V}_i(\cdot) := x_c^{\mathrm{T}} P_i x_c,\ i \in \{1, \ldots, \text{card}(\mathcal{K})\}$. Intuitively, we obtain

a sampled version of the optimal value function along the system's real, and potentially compromised, trajectories.

Lemma 1: The error between the optimal trajectory and the real (potentially attacked) trajectory under the integrable attack signal $\rho(t)$ over a closed time interval $[t_0, t_1]$ is bounded as

$$\|e_x(t)\| \leqslant \alpha_i(t,\rho)\|x(t_0)\|$$

where

$$\alpha_i(t,\rho) = \int_{t_0}^{t} \delta_i(\tau)\|I-\rho(t)\| e^{\int_{t_0}^{t} \delta_i(\sigma)\|I-\rho(\sigma)\|d\sigma} \mathrm{d}\tau$$

and

$$\delta_i(\tau) = \|e^{(A-B_iK_i)(\tau-t_0)}\|\|B_i\|\|K_i\|$$

with $K_i = R_i^{-1}B_i^{\mathrm{T}}P_i$.

Proof: For ease of exposition, we denote the time interval under consideration as $[t_0, t_1]$. Let $x^\star(t)$ be the trajectory of the system driven by (7) under the absence of attacks, and by $x_c(t)$ the actual—possibly compromised—trajectory. Also we assume, without loss of generality, that $\forall t \leqslant t_0$ there is no attack on the system, therefore, the optimal trajectory of the system coincides with the actual trajectory $x^\star(t_0) = x_c(t_0) = x(t_0)$.

The optimal and actual trajectories evolve according to

$$\dot{x}^\star(t) = (A - B_iK_i)x^\star(t),\ x^\star(t_0) = x(t_0)$$
$$\dot{x}_c(t) = (A - B_i\rho(t)K_i)x_c(t),\ x_c(t_0) = x(t_0).$$

Since we take into account attack signals that may be integrable but discontinuous, the trajectories $x_c(t)$ are defined in the sense of Carathéodory.

First, we consider the actual trajectory of the system $x_c(t)$, $t \in [t_0, t_1]$ according to

$$\dot{x}_c = (A - B_i\rho K_i)x_c\ ,\ x_c(t_0) = x(t_0)$$
$$\dot{x}_c = (A - B_iK_i)x_c + (B_i(I-\rho)K_i)x_c. \tag{21}$$

The solution to (21), with $(B_i(I-\rho)K_i)x_c$ taken as a forcing term is

$$x_c(t) = e^{(A-B_iK_i)(t-t_0)}x_c(t_0) + \int_{t_0}^{t} e^{(A-B_iK_i)(t-\tau)}B_i(I-\rho(\tau))K_ix_c(\tau)\mathrm{d}\tau.$$

Taking norms yields

$$\|x_c(t)\| \leqslant \|e^{(A-B_iK_i)(t-t_0)}\|\|x(t_0)\| + \int_{t_0}^{t} \|e^{(A-B_iK_i)(t-\tau)}\|\|B_i\|\|(I-\rho(\tau))\|\|K_i\|)\|x_c(\tau)\|\mathrm{d}\tau.$$

Since each controller K_i renders the system stable, we know that the transition matrix of the closed-loop system will be upper bounded. Therefore, we can introduce $\gamma_i = \max_t \|e^{(A-B_iK_i)(t-t_0)}\|$. By denoting as $\delta_i(\tau) = \|e^{(A-B_iK_i)(\tau-t_0)}\|\|B_i\|\|K_i\|$, will have

$$\|x_c(t)\| \leqslant \gamma_i\|x_0\| + \int_{t_0}^{t} \delta_i(\tau)\|I-\rho(\tau)\|\|x_c(\tau)\|\mathrm{d}\tau.$$

It has been shown in [37] that under assumptions of integrability, Gronwall-type inequalities hold for discontinuous functions inside the integral, such as $\delta(\tau)\|I-\rho(\tau)\|$. Applying these results yields a bound on the norm of the actual trajectory

$$\|x_c(t)\| \leqslant \gamma_i\|x(t_0)\|e^{\int_{t_0}^{t} \delta_i(\tau)\|I-\rho(\tau)\|\mathrm{d}\tau}. \tag{22}$$

We can now define the error between the actual and the optimal trajectory as

$$e_x(t) = x_c(t) - x^\star(t) \tag{23}$$

with dynamics given by

$$\dot{e}_x = \dot{x}_c - \dot{x}^\star = (A - B_i\rho K_i)x_c - (A - B_iK_i)x^\star \Rightarrow$$
$$\dot{e}_x = (A - B_iK_i)e_x + B_i(\rho - I)K_ix_c.$$

Due to the assumption that $x^\star(t_0) = x_c(t_0)$, we have $e(t_0) = 0$ and

$$e_x(t) = \int_{t_0}^{t} e^{(A-B_iK_i)(t-\tau)}B_i(\rho(\tau)-I)K_ix_c(\tau)\mathrm{d}\tau$$

which, after taking norms, yields

$$\|e_x(t)\| \leqslant \int_{t_0}^{t} \|e^{(A-B_iK_i)(t-\tau)}\|\|B_i\|\|I - \rho(\tau)\|\|K_i\|\|x_c(\tau)\|\mathrm{d}\tau.$$

Utilizing now the bound (22), we can further write

$$\begin{aligned}\|e_x(t)\| \leqslant & \int_{t_0}^{t} \|B_i\|\|I-\rho(\tau)\|\|K_i\|\|e^{(A-B_iK_i)(t-\tau)}\| \\ & \times \|x(t_0)\|e^{\int_{t_0}^{\tau} \delta_i(\sigma)\|I-\rho(\sigma)\|d\sigma}\mathrm{d}\tau \\ = & \int_{t_0}^{t} \delta_i(\tau)\|I-\rho(\tau)\|\|x(t_0)\|e^{\int_{t_0}^{\tau} \delta_i(\sigma)\|I-\rho(\sigma)\|d\sigma}\mathrm{d}\tau.\end{aligned}$$

By using

$$\alpha_i(t,\rho) = \int_{t_0}^{t} \delta_i(t)\|I-\rho(t)\|e^{\int_{t_0}^{t} \delta_i(\sigma)\|I-\rho(\sigma)\|d\sigma}\mathrm{d}\tau$$

for which it holds that $\alpha_i(t,\rho) \neq 0$, $\forall \rho \neq I$, we can write the bound on the trajectory error as

$$\|e_x(t)\| \leqslant \alpha_i(\tau,\rho)\|x(t_0)\|.$$

■

Remark 7: It can be seen that $\alpha_i(\tau,\rho) = 0$, if and only if $\rho(t) = I$, $\forall t \in [t_0, t_1]$. □

Theorem 3: Consider that the system is operating with $K_i \in \mathcal{K}$, designed based on (7) and (8). Define the detection signal over a predefined time window $T > 0$

$$e(t) = \hat{V}_i(x_c(t-T)) - \hat{V}_i(x_c(t)) - \int_{t-T}^{t} (x_c^{\mathrm{T}}Q_ix_c + u_i^{\star\mathrm{T}}R_iu_i^\star)\mathrm{d}\tau. \tag{24}$$

Then, the system is under attack if and only if $e(t) \neq 0$. The optimality loss due to the attacks, quantified by $\|e(t)\|$, is bounded for any injected signal $\rho(t)$ that is integrable.

Proof: As was proven in [38], (24) is the integral form of the Bellman equation. For the sampled value of the state at

$t_1 = t - T$, we have that

$$\begin{aligned}\hat{V}_i(t-T) &= x^{\mathrm{T}}(t-T)P_i x(t-T)\\ &= \min_{u_i}\Big\{\int_{t-T}^{t}(x^{\mathrm{T}}Q_i x + u_i^{\mathrm{T}}R_i u_i)\mathrm{d}\tau\\ &\quad + \hat{V}_i(t)\Big\}.\end{aligned}$$

Since $P_i \succ 0$ we have

$$\begin{aligned}\hat{V}_i(t-T) &= x^{\mathrm{T}}(t-T)P_i x(t-T)\\ &= \min_{u_i}\Big\{\int_{t-T}^{t}(x^{\mathrm{T}}Q_i x + u_i^{\mathrm{T}}R_i u_i)\mathrm{d}\tau\Big\}\\ &\quad + x^{\mathrm{T}}(t)P_i x(t).\end{aligned}$$

For the accumulated cost utilizing the optimal input, and the cost utilizing an arbitrary input u_a, it holds that

$$\begin{aligned}&\int_{t-T}^{t}(x^{\mathrm{T}}Q_i x + u_i^{\star}R_i^{\mathrm{T}}u_i^{\star})\mathrm{d}\tau\\ &\quad = \min_{u_i}\Big\{\int_{t-T}^{t}(x^{\mathrm{T}}Q_i x + u_i^{\mathrm{T}}R_i u_i)\mathrm{d}\tau\Big\}\\ &\quad \leqslant \int_{t-T}^{t}(x^{\mathrm{T}}Q_i x + u_a^{\mathrm{T}}R_i u_a)\mathrm{d}\tau \Rightarrow\\ &\int_{t-T}^{t}(x^{\mathrm{T}}Q_i x + u_i^{\star}R_i^{\mathrm{T}}u_i^{\star})\mathrm{d}\tau\\ &\quad = \int_{t-T}^{t}(x^{\mathrm{T}}Q_i x + u_a^{\mathrm{T}}R_i u_a)\mathrm{d}\tau - I(\rho).\end{aligned}$$

Due to Assumption 4, the solution is unique. By extension, the optimal cost over any time interval is also unique. Consequently, the system is attack-free when $I(\rho) = 0$.

For the boundedness part of the proof, we adopt the notation of Lemma 1. Along the actual trajectory of the system within a time interval $[t_0, t_1]$, the intrusion detection signal is

$$\begin{aligned}e(t) &= x_c^{\mathrm{T}}(t_0)P_i x_c(t_0) - x_c^{\mathrm{T}}(t_1)P_i x_c(t_1)\\ &\quad - \int_{t_0}^{t_1}(x_c^{\mathrm{T}}(\tau)Q_i x_c(\tau) + u_i^{\star\mathrm{T}}(x_c)R_i u_i^{\star}(x_c))\mathrm{d}\tau.\end{aligned}$$

Since, the control signal utilized by the controller will be optimal, we can write, $\forall t \geqslant 0$

$$\begin{aligned}e(t) &= x_c^{\mathrm{T}}(t_0)P_i x_c(t_0) - x_c^{\mathrm{T}}(t_1)P_i x_c(t_1)\\ &\quad - \int_{t_0}^{t_1}\big(x_c^{\mathrm{T}}(\tau)Q_i x_c(\tau)\\ &\quad + \big(-R_i^{-1}B_i^{\mathrm{T}}P_i x_c(\tau)\big)^{\mathrm{T}}R_i\big(-R_i^{-1}B_i^{\mathrm{T}}P_i x_c(\tau)\big)\big)\mathrm{d}\tau \Rightarrow\\ e(t) &= x_c(t_0)^{\mathrm{T}}P_i x_c(t_0) - x_c(t_1)^{\mathrm{T}}P_i x_c(t_1)\\ &\quad - \int_{t_0}^{t_1}x_c^{\mathrm{T}}(\tau)\tilde{Q}_i x_c(\tau)\mathrm{d}\tau\end{aligned}$$

where $\tilde{Q}_i = Q_i + P_i B_i R_i^{-1}B_i^{\mathrm{T}}P_i \succ 0$. The positive definiteness is derived by the asymptotic stability property of the optimal closed-loop system.

We substitute the actual trajectory $x_c(t)$, utilizing the trajectory error (23), to write

$$\begin{aligned}e(t) &= x_c^{\mathrm{T}}(t_0)P_i x_c(t_0)\\ &\quad - \big(x^{\star}(t_1) + e_x(t_1)\big)^{\mathrm{T}}P_i\big(x^{\star}(t_1) + e_x(t_1)\big)\\ &\quad - \int_{t_0}^{t_1}\big(x^{\star}(\tau) + e_x(\tau)\big)^{\mathrm{T}}\tilde{Q}_i\big(x^{\star}(\tau) + e_x(\tau)\big)\mathrm{d}\tau\end{aligned}$$

which can be rewritten as

$$\begin{aligned}e(t) &= x_c^{\mathrm{T}}(t_0)P_i x_c(t_0) - x^{\star\mathrm{T}}(t_1)P_i x^{\star}(t_1)\\ &\quad - \int_{t_0}^{t_1}x^{\star}(\tau)^{\mathrm{T}}\tilde{Q}_i x^{\star}(\tau)\\ &\quad - \Big\{e_x^{\mathrm{T}}(t_1)P_i x^{\star}(t_1) + x^{\star\mathrm{T}}(t_1)P_i e_x(t_1)\\ &\quad + e_x^{\mathrm{T}}(t_1)P_i e_x(t_1)\\ &\quad + \int_{t_0}^{t}(e_x^{\mathrm{T}}(\tau)\tilde{Q}_i x^{\star}(\tau) + x^{\star\mathrm{T}}(\tau)\tilde{Q}_i e_x(\tau)\\ &\quad + e_x^{\mathrm{T}}(\tau)\tilde{Q}_i e_x(\tau))\mathrm{d}\tau\Big\}.\end{aligned}$$

It can be seen that the first three terms of the above-mentioned expression satisfy the integral form of the HJB equation. As a result, the residual terms that quantify the optimality loss due to the attack are

$$\begin{aligned}e(t) = &-\big\{e_x^{\mathrm{T}}(t_1)P_i x^{\star}(t_1) + x^{\star\mathrm{T}}(t_1)P_i e_x(t_1)\\ &+ e_x^{\mathrm{T}}(t_1)P_i e_x(t_1)\\ &+ \int_{t_0}^{t}(e_x^{\mathrm{T}}(\tau)\tilde{Q}_i x^{\star}(\tau)\\ &+ x^{\star\mathrm{T}}(\tau)\tilde{Q}_i e_x(\tau) + e_x^{\mathrm{T}}(\tau)\tilde{Q}_i e_x(\tau))\mathrm{d}\tau\big\}.\end{aligned}$$

Taking norms, and utilizing the fact that $x^{\star}(t) = e^{(A-B_i K_i)(t-t_0)}x(t_0)$ as well as Lemma 1, we can bound the norm of the intrusion detection signal as

$$\|e(t)\| \leqslant\ b_i(t,\rho)\|x(t_0)\|^2$$

where

$$\begin{aligned}b_i(t,\rho) &= 2\alpha_i(t,\rho)\|P_i\|\gamma_i + \alpha_i^2(t,\rho)\|P_i\|\\ &\quad + \int_{t_0}^{t}\big(2\alpha_i(\tau,\rho)\|\tilde{Q}_i\|\gamma_i + \alpha_i^2(\tau,\rho)\|\tilde{Q}_i\|\big)\mathrm{d}\tau\end{aligned}$$

with the property that $b_i(t,\rho) = 0$, $\forall t \in [t_0, t_1]$ if and only if $\rho(t) = I$. ■

Remark 8: We note that in the notation of Lemma 1 and Theorem 3, both the state trajectories $x^{\star}(t)$, $x_c(t)$ and the error signals $e_x(t)$, $e(t)$ were not dependent on the active mode i even if their dynamics were. This is due to the fact that since the different modes operate sequentially, the defender computes only a single error signal at each time instant. □

Algorithm 1: Proactive/Reactive Defense Mechanism for Actuator Attacks.

procedure
  Given an initial state $x(t_0)$, and a time window T.
  Find all permutations of actuators (columns of B) and derive the subset of controllable pairs (A, B_i), denoted by $\mathcal{K}$.
  for $i = 1, \ldots,$ card$(\mathcal{K})$
    Compute the optimal feedback gain and Riccati matrices K_i, P_i according to (7) and (9).
    Compute the optimal cost of each controller for the given $x(t_0)$.
  end for
  Solve for the optimal probabilities $p_i^\star$ using (10).
  At $t = t_0$, choose the optimal controller for which $\sigma(t_0) = \arg\min_i \left(x(t_0)^{\mathrm{T}} P_i x(t_0)\right)$.
  while $\sigma(t) = i$ and $t < \tau_D$
    Compute the integral Bellman error detection signal using (24).
    Propagate the system using (5).
  end while
  Choose the random mode $\sigma(t + \tau_D) = j$ and go to 9.
  if $\|e_i(t_c)\| > 0$
    Take the i-th controller offline.
    Switch to the controller with the best performance, $\sigma(t_c) = \arg\min_{i \in \mathcal{K} \setminus i} \left(x(t_0)^{\mathrm{T}} P_i x(t_0)\right)$ and go to 9.
  end if
end procedure

D. Proactive and Reactive Defense for Actuator Attacks

Under safe operation, the system switches between the available modes with MTD in order to have guaranteed stability and maximal unpredictability, according to (11). If we can find $i : e_i(t_k) \neq 0$, then we can conclude that the ith mode is considered under attack, and isolated. Specifically, the system switches to the controller with the best performance and the compromised ith mode is taken out of the queue for the MTD switching. The pseudocode for the proactive and reactive defense system is provided in Algorithm 1.

Fact 3: It has been shown that $p_i > 0$, $\forall i \in \{1, \ldots,$ card$(\mathcal{K})\}$. Consequently, there exists a $t_f^\star$ such that $\exists \tau \in [t_0, t_f^\star]$ with $\sigma(\tau) = i, \forall i \in \{1, \ldots,$ card$(\mathcal{K})\}$ and an arbitrary $t_0 > 0$. This means that, since the probability that all controllers will eventually be active, there is some time interval long enough, such that we have already switched through every available controller. □

Theorem 4: Suppose that the (1), uses the framework of Algorithm 1. Then, the closed-loop system has an asymptotically stable equilibrium point given that the attacker has not compromised all the available controllers, i.e., $\mathcal{K} \setminus \mathcal{K}_c \neq \emptyset$, where $\mathcal{K}_c$ is the subset of those controllers that have been compromised by an attacker.

Proof: We consider a trajectory of the system within the time interval $t \in [t_0, t_f]$, $t_f > t_f^\star$. Denote by $\mathcal{K}_u$ the set of safe controllers and $\mathcal{K}_c$ the set of compromised ones. Recall that according to Algorithm 1 and Theorem 4, the controller will stay at a compromised mode during $[t, t + T]$. Since the MTD algorithm is constrained by an average dwell time, for any part of the trajectory $t \in [t_k, t_{k+1}]$ where a compromised controller has not been utilized, it holds according to Theorem 2, that

$$\|x(t_{k+1})\| < \|x(t_k)\|. \tag{25}$$

We now need to take into account those instances where, after detecting a compromised controller K_i, the system immediately switches to another controller, which is also compromised. For $\eta \in \mathbb{N}$ subsequent switches to compromised controllers, due to Lemma 1, those parts of the trajectory for $t \in [t_k, t_{k+1}] = [t_k, t_k + \eta T]$ are bounded by a positive definite function $\beta_i(\rho, \tau)$ as

$$\begin{aligned}\|x(t_k + \eta T)\| &\leqslant \beta_{\sigma(t_k+\eta T)}(\rho, T)\|x(t_k + (\eta - 1)T)\| \\ &\leqslant \beta_{\sigma(t_k+\eta T)}(\rho, T)\beta_{\sigma(t_k+(\eta-1)T)}(\rho, T) \\ &\quad \times \|x(t_k + (\eta - 2)T)\| \\ &\leqslant \prod_{i=1}^{\eta} \beta_i(\rho, T)\|x(t_k)\|. \end{aligned} \tag{26}$$

Furthermore (26) can be upper bounded as

$$\|x(t_k + \eta T)\| \leqslant \left(\max_{i \in \mathcal{K}_c}(\beta_i(\rho, T))\right)^\eta \|x(t_k)\|.$$

Due to the fact that $\beta_i(\rho(t), T) > 0$, $\forall i$, we can conclude that the parts of the trajectory where the compromised and safe modes are interchanged, are upper bounded by the same trajectory driven only by compromised modes, by combining the inequalities (25) and (26).

By Assumption 3, the attacker under finite resources is able to compromise a number N of the available controllers, i.e., card$(\mathcal{K}_c) \leqslant N$. Using Algorithm 1, and Fact 3, there is a time $t_p < t_f$ such that all the compromised controllers have been detected by the integral Bellman detector and have been isolated from the switching queue of the MTD.

Consequently, recalling that every closed-loop matrix $\tilde{A}_i = A - B_i K_i$ is Hurwitz, there exist positive numbers $\bar{K}_i$, a_i such that, $\|e^{\tilde{A}_i t}\| \leqslant \bar{K}_i e^{-a_i t}$, one has

$$\begin{aligned}\|x(t_f)\| &\leqslant \bar{K}_i e^{-a_i t}\|x(t_p)\| \\ &\leqslant \bar{K}_i e^{-a_i t}\left(\max_{i \in \mathcal{K}_c}(\beta(\rho, T))\right)^\eta \|x(t_0)\|.\end{aligned}$$

It can be seen that the remaining trajectory converges to the origin exponentially fast with rate that depends on the slowest safe controller. As a result, if the set of the safe controllers is not empty, the trajectory is guaranteed to asymptotically go to zero as $t_f \to \infty$. ■

1) Intrusion Detection Under Actuation Noise: It is possible to extend the results of the previous section to take into account noise in the actuation mechanism, i.e., in (1)

$$u_a(t) = \rho(t)u^\star(t) + w(t)$$

where $w(t)$ is a bounded but otherwise unknown disturbance with $\|w(t)\| \leqslant \bar{w}$.

Theorem 5: System (1), equipped with the MTD control scheme described in Section III and the detection mechanism as

defined in Theorem 3, under the effect of a disturbance $w(t)$ is compromised if

$$\|e(t)\| \geqslant e_{i,\text{thres}}(t)$$

where $e_{i,\text{thres}}$ are the dynamic thresholds for each mode of the form

$$e_{i,\text{thres}}(t) = 2\|\bar{w}\| \int_{t-T}^{t} \|R_i u_i^\star(\tau)\| \mathrm{d}\tau + \bar{\lambda}(R_i)\|\bar{w}\|^2.$$

Proof: First, we will consider the system in the absence of attacks and formulate the intrusion detection signal based on the data collected along the trajectories of the system. In other words, we can write

$$\begin{aligned} e(t) &= \hat{V}_i(t-T) - \hat{V}_i(t) - \int_{t-T}^{t} (x^{\mathrm{T}} Q_i x + u_a^{\mathrm{T}} R_i u_a)\mathrm{d}\tau \\ &= \hat{V}_i(t-T) - \hat{V}_i(t) - \int_{t-T}^{t} \big(x^{\mathrm{T}} Q_i x \\ &\quad + (u_i^\star + w)^{\mathrm{T}} R_i (u_i^\star + w)\big)\mathrm{d}\tau \\ &= \hat{V}_i(t-T) - \hat{V}_i(t) - \int_{t-T}^{t} (x^{\mathrm{T}} Q_i x + u_i^{\star\mathrm{T}} R_i u_i^\star)\mathrm{d}\tau \\ &\quad - \int_{t-T}^{t} (w^{\mathrm{T}} R_i u^\star + u_i^{\star\mathrm{T}} R_i w \\ &\quad + w^{\mathrm{T}} R_i w)\mathrm{d}\tau. \end{aligned}$$

Leveraging the integral Bellman equality and taking norms yield

$$\begin{aligned} \|e(t)\| &\leqslant 2\int_{t-T}^{t} \|w^{\mathrm{T}} R_i u_i^\star\|\mathrm{d}\tau + \int_{t-T}^{t} \|w^{\mathrm{T}} R_i w\|\mathrm{d}\tau \Rightarrow \\ \|e(t)\| &\leqslant 2\|\bar{w}\| \int_{t-T}^{t} \|R_i u_i^\star(\tau)\|\mathrm{d}\tau + T\bar{\lambda}(R_i)\|\bar{w}\|^2 \end{aligned}$$

which is the adaptive threshold for the active controller i. ■

Remark 9: It should be noted that the adaptive threshold can be computed online utilizing only knowledge of the optimal input signal that the controller sends to the system (and not the potentially corrupted one). □

Remark 10: In noisy environments, the system may be under attack, and the integral Bellman error may not cross the adaptive threshold. Thus, the attack shall remain undetected. However, attacks that have so little effect on the system become indistinguishable from random noise and do not degrade the performance of the system in a significant way. □

IV. Defense Against Sensor Attacks

In this section, we show how the methods developed can be applied to securely estimate the state of a system with compromised measurements by employing sensor redundancy.

A. Candidate Sensors Sets

Similarly to the proposed framework for the actuators, we introduce the set of all sensors, denoted by $\mathcal{C}$, and the elements of its power set $\mathcal{C}_i \in 2^{\mathcal{C}}$, and $C_i \in \mathcal{C}_i$ is a combination of the different rows of C.

The set of candidate sensing modes $\mathcal{S}_o$ is defined as the set of the sensor combinations that renders system (1) fully observable

$$\mathcal{S}_o = \left\{C_j \in 2^{\mathcal{C}} : \text{rank} \begin{pmatrix} \begin{bmatrix} C_j \\ C_j A \\ \vdots \\ C_j A^{n-1} \end{bmatrix} \end{pmatrix} = n\right\}.$$

The system utilizing the sensor combination C_i is

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y_i &= C_i x. \end{aligned}$$

Remark 11: We note the distinction between the set of sensors $\mathcal{C}$ and the set of sensing modes $\mathcal{S}_o$. The set of sensors contains the different physical components that measure parts of the system's behavior. On the other hand, the set of sensing modes contains those cooperating sensors together with an observer scheme that reconstruct an estimate of the system state. □

B. Optimal Observer Design and MTD for Sensor Attacks

The observer of (1) will be now designed as a dynamic system sharing the same structural properties

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + Bu + B\bar{u}_i \\ \hat{y}_i &= C_i \hat{x} \end{aligned} \tag{27}$$

where $\hat{x}$, $\hat{y}_i$ are the estimates of the state and the output, respectively, $\bar{u}_i$ denotes a "fictional" input, i.e., a correction term, which forces the observer to track the actual system.

Remark 12: The state estimate $\hat{x}$ is independent of the active sensing mode. On the other hand, the output $\hat{y}_i$ and "fictional" input $\bar{u}_i$ are not. □

Following the work of [39], [40], to design the optimal $\bar{u}_i$, we define the optimization problem based on the following cost function $\forall t \geqslant 0$

$$U_i^\star(\hat{x}) = \min_{\bar{u}_i} \int_t^\infty \left[(\hat{y}_i - y_i)^{\mathrm{T}} Q_i (\hat{y}_i - y_i) + \bar{u}_i^{\mathrm{T}} R_i \bar{u}_i\right] \mathrm{d}\tau.$$

Defining the Hamiltonian of the system as

$$\begin{aligned} H_i(\hat{x}, \bar{u}_i^\star, U_i^\star) &= (\hat{y}_i - y_i)^{\mathrm{T}} Q_i (\hat{y}_i - y_i) + \bar{u}_i^{\star\mathrm{T}} R_i \bar{u}_i^\star \\ &\quad + \nabla U_i^{\star\mathrm{T}} (A\hat{x} + Bu + B\bar{u}_i^\star) = 0. \end{aligned} \tag{28}$$

We can now find the optimal control from the stationarity conditions $\frac{\partial H_i(\hat{x}, \bar{u}_i^\star, U_i^\star)}{\partial \bar{u}_i^\star} = 0$. This leads to

$$\bar{u}_i^\star = -R_i^{-1} B^{\mathrm{T}} \nabla U_i^\star(\hat{x}).$$

Due to the quadratic structure of the cost functional and the linear structure of the dynamic system, we assume that the value function is quadratic in $\hat{x}(t)$, i.e., $U_i^\star(\hat{x}) = \hat{x}^{\mathrm{T}} G_i \hat{x}$, $G_i \succ 0$, which means that the optimal "input" is

$$\bar{u}_i^\star = -R_i^{-1} B^{\mathrm{T}} G_i \hat{x}. \tag{29}$$

In this section, we show how the same techniques introduced and analyzed in the previous sections can be applied to detect and mitigate sensor attacks.

C. MTD for Sensor Attacks

Theorem 6: The state estimation scheme utilizing optimal observers as described by (27), for every sensing mode in $\mathcal{S}_o$ has an asymptotically stable equilibrium point under a switching-based MTD mechanism given that the switching signal has the average dwell time

$$\tau_D > \frac{\log\left(\max_{q,p\in\{1,\dots,\text{card}(\mathcal{S}_o)\}} \frac{\bar{\lambda}(G_p)}{\underline{\lambda}(G_q)}\right)}{\min_{p\in\{1,\dots,\text{card}(\mathcal{S}_o)\}} \frac{\underline{\lambda}(C_i^{\mathrm{T}}Q_pC_i+G_pB_pR_p^{-1}B_p^{\mathrm{T}}G_p)}{\underline{\lambda}(G_p)}}.$$

Proof: The proof follows closely Theorem 2 for the switched observer comprised of different sensing modes and by taking into account the optimal control problem formulated in this section by (29) and (28). ■

Remark 13: The optimization problem solved in Section III-C is identical for the case of sensor switching. As a result, the probability that a certain sensing mode S_i is active, obeys (10). □

D. Integral Bellman Based Intrusion Detection for Sensor Attacks

We will now introduce a detection signal based on the online, possibly compromised, estimations of the state, which we will denote $\hat{x}_c(t)$. For that reason, we formulate the function $\hat{U}_i(t) = \hat{x}_c^{\mathrm{T}}G_i\hat{x}_c,\ G_i \succ 0$.

Theorem 7: Consider system (1) operating with the sensor mode $S_i \in \mathcal{S}$, designed based on (28) and (29). Define the detection signal over a predefined time window $T > 0$ as

$$e^s(t) = \hat{U}_i(\hat{x}_c(t-T)) - \hat{U}_i(\hat{x}_c(t)) - \int_{t-T}^{t} \left((y_i-\hat{y}_i)^{\mathrm{T}}Q_i(y_i-\hat{y}_i) + \bar{u}_i^{\star\mathrm{T}}R_i\bar{u}_i^{\star}\right)\mathrm{d}\tau. \quad (30)$$

Then, the system is under attack if and only if $e^s(t) \neq 0$. Moreover, the optimality loss due to attacks, is bounded for any injected signal $\rho^s(t)$.

Proof: The first part of the proof follows from Theorem 2, for the optimal control problem formulated in this section, and is based on the uniqueness of optimal solutions for a given initial condition. However, to compute the bound on the optimality/observation loss, we define the measurement error $\tilde{y}_i = y_i - \hat{y}_i$. Then, the detection signal is

$$e^s(t) = \hat{U}_i(\hat{x}_c(t-T)) - \hat{U}_i(\hat{x}_c(t)) - \int_{t-T}^{t} (\tilde{y}_i^{\mathrm{T}}Q_i\tilde{y}_i + \bar{u}_i^{\mathrm{T}}R_i\bar{u}_i)\mathrm{d}\tau.$$

Note that, in the presence of attacks

$$y_i - \hat{y}_i = \rho^s C_i\hat{x} - C_i\hat{x} = \tilde{y}_i + \delta_i^s(\rho^s)$$

where $\delta_i^s(\rho^s) = (I-\rho^s)C_i\hat{x}$. Therefore, the detection signal under attack is

$$e^s(t) = \hat{U}_i(\hat{x}_c(t-T)) - \hat{U}_i(\hat{x}_c(t)) - \int_{t-T}^{t} \left((\tilde{y}_i+\delta_i^s(\rho^s))^{\mathrm{T}}Q_i(\tilde{y}_i+\delta_i^s(\rho^s)) + \bar{u}_i^{\mathrm{T}}R_i\bar{u}_i\right)\mathrm{d}\tau.$$

Algorithm 2: Proactive/Reactive Defense Mechanism for Sensor Attacks.

procedure
  Given initial state $x(t_0)$, system dynamics (1) and time window T.
  Find all permutations of sensors (rows of C) and derive the subset of observable pairs (A, C_i), denoted $\mathcal{S}_o$.
  for $i = 1, \dots, \text{card}(\mathcal{S}_o)$
    Compute the optimal 'fictional' input and value function according to (28) and (29).
    Compute the optimal cost of each observation mode for the given $x(t_0)$.
  end for
  Solve for the optimal probabilities $p_i^\star$ using (10).
  At $t = t_0$, choose the optimal observer.
  while $\sigma(t) = i$ and $t < \tau_D$
    Compute the integral Bellman error detection signal using (30).
    Propagate the system using the observer dynamics.
  end while
  Choose a random mode $\sigma(t+\tau_D) = j$ and go to 9.
  if $\|e^s(t_c)\| > 0$
    Take the i-th observer offline.
    Switch to the safe observer with the best performance and go to 9.
  end if
end procedure

Expanding the quadratic terms, the residual detection signal becomes

$$e^s(t) = -\int_{t-T}^{t} \left(\delta_i^{s\mathrm{T}}(\rho^s)Q_iC_i\hat{x} + (C_i\hat{x})^{\mathrm{T}}Q_i\delta_i^s(\rho^s) + \delta_i^{s\mathrm{T}}(\rho^s)Q_i\delta_i^s(\rho^s)\right)\mathrm{d}\tau.$$

Taking into account the Cauchy–Schwartz inequality, we can bound the norm of the error as

$$\|e^s(t)\| \leqslant 2\int_{t-T}^{t} \|C_i\hat{x}\|\|\delta_i^s(\rho^s)\| + T\bar{\lambda}(Q_i)\|\delta_i^s(\rho^s)\|.$$

■

Remark 14: The bound on the optimality/observation loss can be quantified more easily, because the injected attack does not directly affect the dynamics in system (1), rather it behaves like a noise in the cost term defined by the output error $(y_i - \hat{y}_i)^{\mathrm{T}}Q_i(y_i-\hat{y}_i)$. □

E. Proactive and Reactive Defense for Sensor Attacks

We will now combine the proactive defense mechanism with the intrusion detection system described above. The pseudocode for the operation is presented in Algorithm 2.

Remark 15: We can combine the algorithmic frameworks presented for both actuators and sensor attacks. However, the result would be conservative, since the two problems are coupled. Consequently, we cannot differentiate between integral Bellman errors caused by an actuator or a sensor attack. □

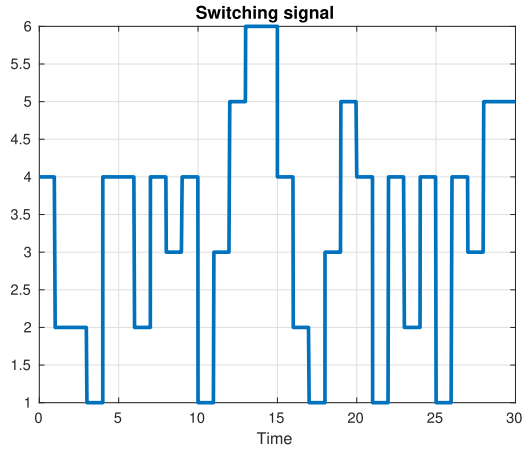

Fig. 1. Evolution of the MTD switching signal that guarantees actuator proactive security. It can be seen that controller with index 4 is preferred since it is the most optimal.

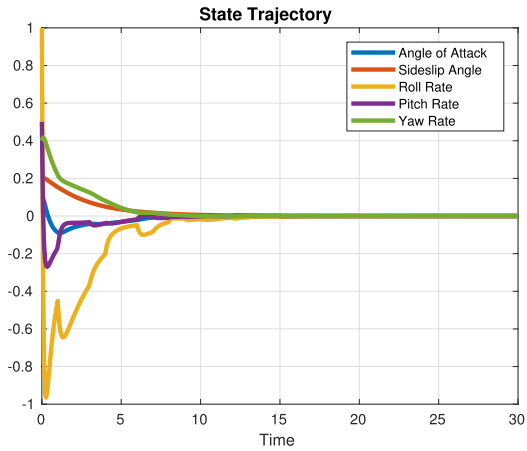

Fig. 2. Evolution of the MTD state that guarantees actuator proactive security. With the appropriate dwell-time, the system remains stable.

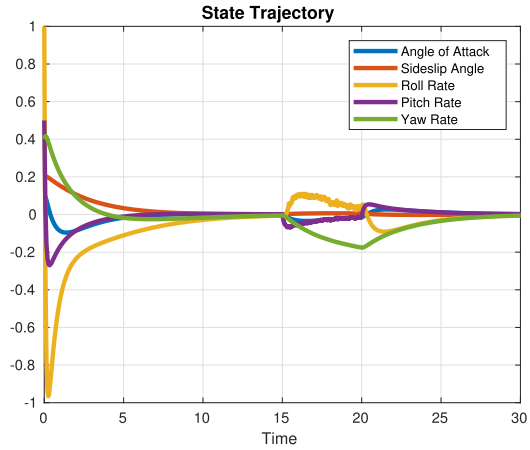

Fig. 3. Evolution of the states in the presence of actuator attacks. The attack takes place for $t \in [15, 20]$.

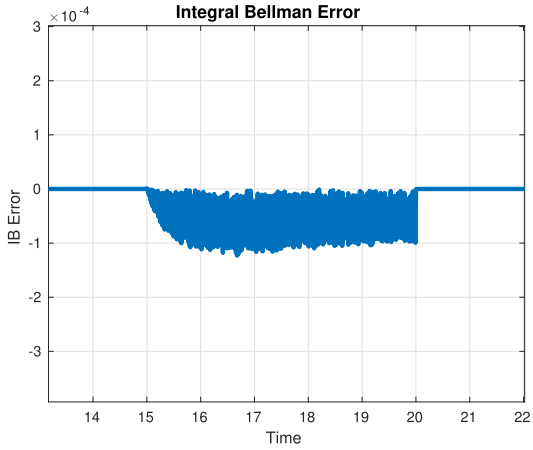

Fig. 4. Evolution of the integral Bellman error. For the time interval where the attacker inputs an adversarial signal, $t \in [15, 20]$, the integral Bellman error is nonzero, which is enough to achieve intrusion detection in the absence of stochastic noise.

V. Simulation

In order to show the effectiveness of our approaches we will use a linearized five-dimensional model of the ADMIRE benchmark aircraft [41]. The model has seven redundant actuators and two redundant sensors. Initially, we present results for the problem of controlling the plant in an adversarial environment. Fig. 1 shows the switching signal for the MTD framework applied to actuator attacks. It can be seen that the actuator with index 4 is the preferred one. This is due to its overall optimality compared to the rest of the actuator modes. Fig. 2 shows convergence of the states under actuator MTD. Under the appropriate dwell time, the switching system remains asymptotically stable. In Fig. 3, we can see the evolution of the states under an attack signal for $t \in [15, 20]$ where only the intrusion detection system was utilized. In Fig. 4, we see the evolution of the integral Bellman error. Although its magnitude is small, due to the absence of stochastic noise, the integral Bellman error is still able to detect the attack.

In Fig. 5, we combine the reactive and the proactive security system. The adversary manages to completely shut down one of the actuators belonging to the most optimal controller in $t = 6$ s. It is clear that the system is stabilized. In Fig. 6, we show the evolution of the random switching signal favoring the controller with the best performance. After an attack is detected the compromised component is taken out of the switching queue. However, even without the compromised mode, the MTD structure still operates. This way, we maintain some level of unpredictability, while guaranteeing attack-free operation of the system. We note that, the more compromised modes we have, the less unpredictable the system will be. However, once a mode has been taken out of the switching queue, offline methods may be utilized to repair it and reintroduce it to the MTD. In Fig. 7, we consider a system under actuator attacks in the presence of system noise. Specifically, the noise had known upper bound $\|w\| = 0.5$, while the attack was a random signal with maximum value of 0.3. We note from Fig. 8, that the intrusion

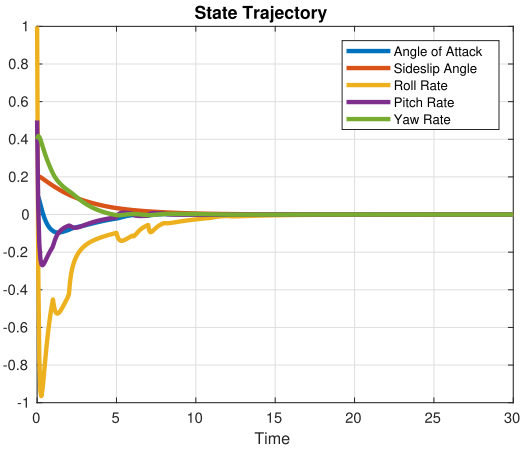

Fig. 5. Evolution of the state with both proactive and reactive defense. Even in the presence of attacks, the system converges to the origin, since the attacked components have been taken offline.

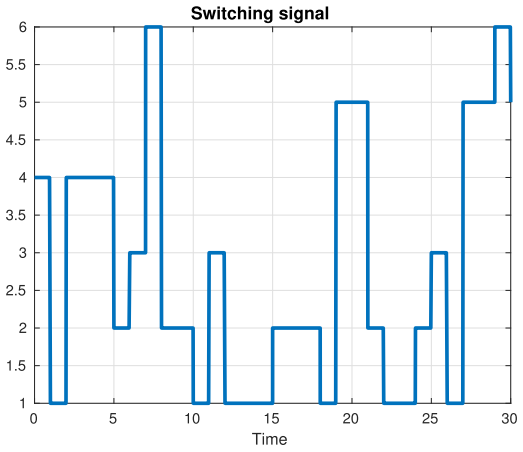

Fig. 6. Evolution of the switching signal with both proactive and reactive defense. It can be seen that when the adversarial signal is detected in the fourth controller, the random switching persists, but never chooses the compromised configuration.

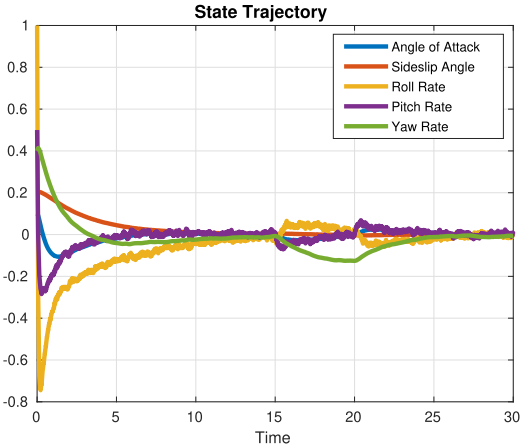

Fig. 7. State evolution of a system with noise and actuator attacks. We note that the attack takes place at $t = [15, 20]$.

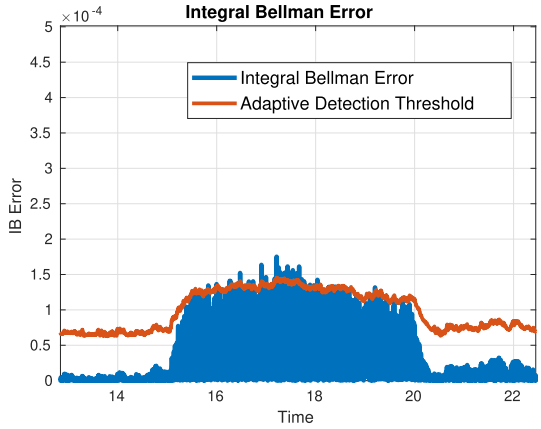

Fig. 8. Evolution of the integral Bellman error for the system under attack, as well as evolution of the adaptive threshold that takes into account the system noise. Despite the magnitude of the attack being smaller than the noise, the proposed algorithm is able to detect the intrusion.

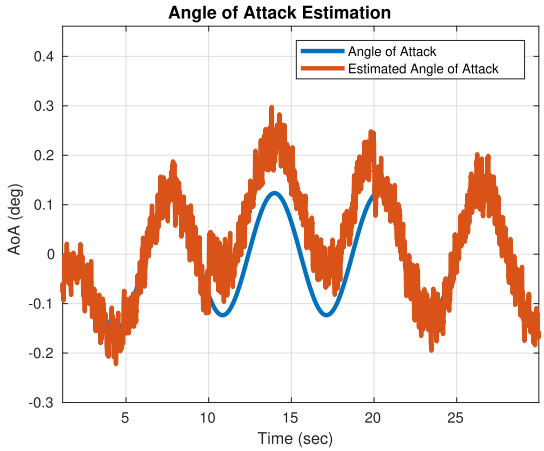

Fig. 9. Optimal state estimation under noisy measurements and injected sensor attack. The attacker corrupts the output of the sensor for $t \in [10, 20]$ by adding a constant bias.

detection signal is able to detect the injected signal despite the noise.

Also, we consider the optimal state estimation problem for the ADMIRE aircraft utilizing the optimal observer framework. The state evolution of the observer is shown in Fig. 9. We notice that the attacker injects a relatively small bias to the estimated signal. For the sensor attacks, we take into account noise (with known statistics) on the measurements and show the evolution of the integral Bellman error and of the adaptive threshold in Fig. 10. Even though the discrepancy between the estimated angle of attack and the actual one is small relative to the measurement noise, the integral Bellman error manages to detect the attack. Specifically, during $t \in [14, 16]$, we detect the difference between the estimated error (due to sensor noise) and the one induced by the attack.

Although the advantages of the model-free integral Bellman error based intrusion detection mechanism can be seen when the

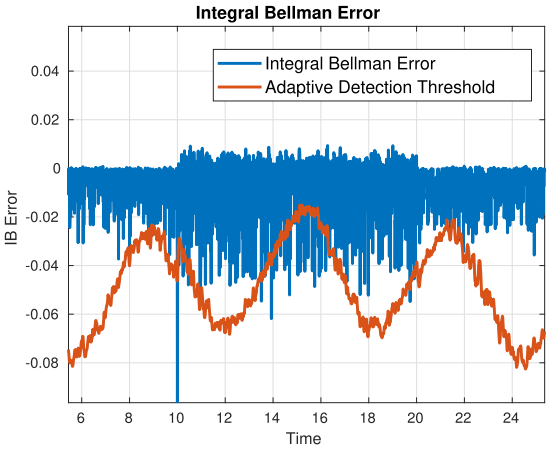

Fig. 10. Evolution of the integral Bellman error and adaptive threshold for successful state reconstruction in the presence of sensor attacks. Even taking into account the noise of the sensors, the attack is detected.

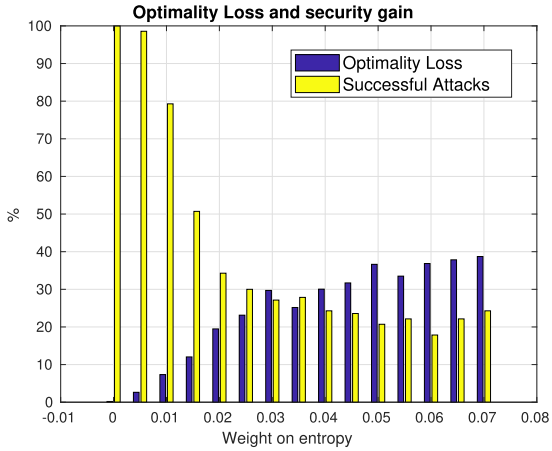

Fig. 12. Optimality loss and rate of successful attacks as a function of the entropy. Increase of the weight on unpredictability leads to performance degradation but manages to secure the system from attacks.

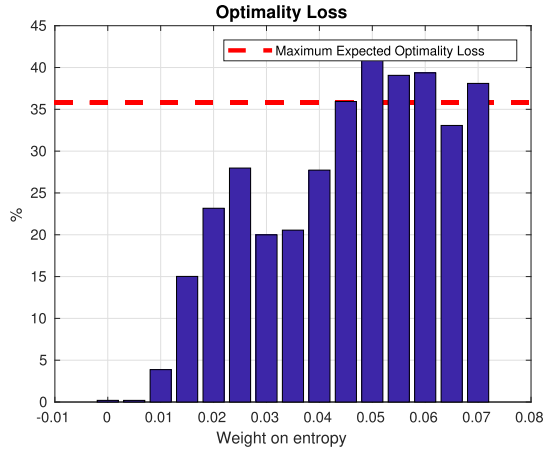

Fig. 11. Optimality loss (difference between actual cost during the system run, and the value function of the most optimal controller) induced by the unpredictable controllers for different entropy levels. By increasing the weight on the entropy, we reach a maximum optimality loss in the case of the uniform distribution.

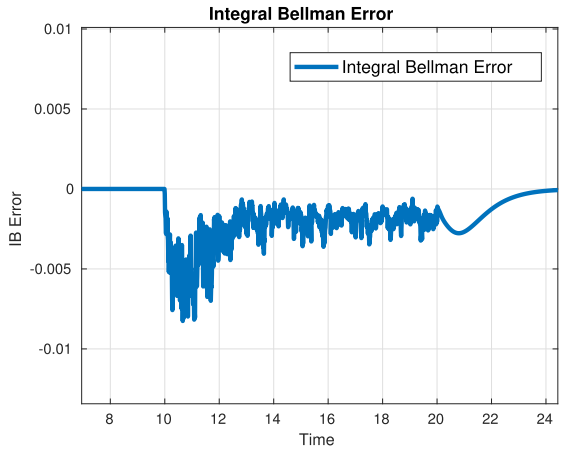

Fig. 13. Evolution of the integral Bellman error of the observer under actuator attacks. This signal is induced by the interconnection between the optimal control problems solved for control and estimation.

system is under attack, due to its proactive nature the success or failure of an MTD system is not easily obvious. Furthermore, the optimality loss induced by the use of non-overall-optimal controllers must be examined. The optimality loss was assessed as the difference between the actual cost during the system run, and the value function of the most optimal controller. To obtain some first validation results for our MTD algorithm, random attack vectors were considered for multiple runs of the system. In Fig. 11, we present the average cost of the system as the unpredictability increases. We can see that the cost converges to a maximum value for uniform distribution over all the available controllers. In Fig. 12, the compromise between security against attacks and optimality is highlighted. Specifically, as we increased the weight on entropy, i.e., the parameter ϵ, then the system switches more aggressively. This leads to the decrease of successful attacks, since the attack is less probable to affect a mode that is in use. However, as ϵ increases, the system utilizes the overall optimal controller less, until it reaches the uniform distribution over the different modes.

Fig. 13 shows the integral Bellman error of the observer under actuator attacks. We note that even though the design processes of the secure optimal controller and the secure optimal observer are separate, attacks in one subsystem (in this case, in the controller) may induce integral Bellman error to the other one (in this scenario, in the observer). However, we are still able to identify a specific controller/observer pair that has been compromised and switch to a different one as described.

VI. Conclusion and Future Work

This paper proposes a proactive and reactive defense switching mechanism against actuator and sensor attacks in CPS. The proactive defense is based on the framework of MTD and maximizes the unpredictability of the system. A novel intrusion detection mechanism, based on the performance evaluation, is used to isolate the attacked actuators and sensors. The system

utilizing both reactive and proactive defenses is proven to have an asymptotically stable equilibrium point provided the attacker did not compromise all the controllers and sensors simultaneously. Simulations on a linearized aircraft model are provided to show the efficiency of our approach.

Future research efforts will focus on incorporating learning mechanisms with attackers of different rationality, i.e., bounded reasoning. Furthermore, we will combine the proposed intrusion detection system with reinforcement learning techniques to develop a CPS framework, which is able to learn optimal behaviors and defend against attackers without knowledge of the model. Moreover, we will investigate attacks that operate between different layers of the CPS, such as the communication and the computation layers. Finally, experiments will be conducted to investigate the practicality of our approach in real-life environments and under realistic attacks.

References

[1] R. R. Rajkumar, I. Lee, L. Sha, and J. Stankovic, "Cyber-physical systems: The next computing revolution," in *Proc. 47th Des. Autom. Conf.*, 2010, pp. 731–736.

[2] I. Lee and O. Sokolsky, "Medical cyber physical systems," in *Proc. 47th ACM/IEEE Des. Autom. Conf.*, 2010, pp. 743–748.

[3] Y. Mo *et al.*"Cyber–physical security of a smart grid infrastructure," *Proc. IEEE*, vol. 100, no. 1, pp. 195–209, Jan. 2012.

[4] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, 2011, Art. no. 13.

[5] J. Kim, H. Kim, K. Lakshmanan, and R. R. Rajkumar, "Parallel scheduling for cyber-physical systems: Analysis and case study on a self-driving car," in *Proc. ACM/IEEE 4th Int. Conf. Cyber Phys. Syst.*, 2013, pp. 31–40.

[6] J. P. Farwell and R. Rohozinski, "Stuxnet and the future of cyber war," *Survival*, vol. 53, no. 1, pp. 23–40, 2011.

[7] J. Slay and M. Miller, "Lessons learned from the maroochy water breach," in *Proc. Int. Conf. Crit. Infrastructure Protection*, 2007, pp. 73–82.

[8] D. P. Shepard, J. A. Bhatti, T. E. Humphreys, and A. A. Fansler, "Evaluation of smart grid and civilian UAV vulnerability to GPS spoofing attacks," in *Proc. Radionavigat. Lab. Conf. Proc.*, 2012.

[9] S. Jajodia, A. K. Ghosh, V. Swarup, C. Wang, and X. S. Wang, *Moving Target Defense: Creating Asymmetric Uncertainty for Cyber Threats*, vol. 54. Berlin, Germany: Springer, 2011.

[10] A. A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *Proc. 28th IEEE Int. Conf. Distrib. Comput. Syst. Workshops,*2008, pp. 495–500.

[11] M. Pajic, J. Weimer, N. Bezzo, O. Sokolsky, G. J. Pappas, and I. Lee, "Design and implementation of attack-resilient cyberphysical systems: With a focus on attack-resilient state estimators," *IEEE Control Syst.Mag.*, vol. 37, no. 2, pp. 66–81, Apr. 2017.

[12] F. Pasqualetti, F. Dorfler, and F. Bullo, "Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems," *IEEE Control Syst.*, vol. 35, no. 1, pp. 110–127, Feb. 2015.

[13] B. Satchidanandan and P. R. Kumar, "Dynamic watermarking: Active defense of networked cyber–physical systems," *Proc. IEEE*, vol. 105, no. 2, pp. 219–240, Feb. 2017.

[14] D. I. Urbina *et al.*, *Survey and New Directions for Physics-Based Attack Detection in Control Systems*. U.S. Dept. Commerce, Nat. Inst. Standards Technol., Gaithersburg, MD, USA, 2016.

[15] K. G. Vamvoudakis, H. Modares, B. Kiumarsi, and F. L. Lewis, "Game theory-based control system algorithms with real-time reinforcement learning: How to solve multiplayer games online," *IEEE Control Syst.*, vol. 37, no. 1, pp. 33–52, Feb. 2017.

[16] T. Alpcan and T. Başar, *Network Security: A Decision and Game-Theoretic Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[17] K. G. Vamvoudakis and J. P. Hespanha, "Cooperative Q-learning for rejection of persistent adversarial inputs in networked linear quadratic systems," *IEEE Trans. Autom. Control*, vol. 63, no. 4, pp. 1018–1031, Apr. 2018.

[18] G. Theodorakopoulos and J. S. Baras, "On trust models and trust evaluation metrics for ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 2, pp. 318–328, Feb. 2006.

[19] S. Jajodia, A. K. Ghosh, V. Subrahmanian, V. Swarup, C. Wang, and X. S. Wang, *Moving Target Defense II: Application of Game Theory and Adversarial Modeling*, vol. 100. Berlin, Germany: Springer, 2012.

[20] V. Casola, A. De Benedictis, and M. Albanese, "A multi-layer moving target defense approach for protecting resource-constrained distributed devices," in *Integration of Reusable Systems*. Berlin, Germany: Springer, 2014, pp. 299–324.

[21] J. H. Jafarian, E. Al-Shaer, and Q. Duan, "Openflow random host mutation: Transparent moving target defense using software defined networking," in *Proc. 1st ACM Workshop Hot Topics Softw. Defined Netw.*, 2012, pp. 127–132.

[22] M. Dunlop, S. Groat, W. Urbanski, R. Marchany, and J. Tront, "Mt6d: A moving target IPv6 defense," in *Proc. IEEE Military Commun. Conf.*, 2011, pp. 1321–1326.

[23] Z. Lu, C. Wang, and M. Wei, "A proactive and deceptive perspective for role detection and concealment in wireless networks," in *Cyber Deception*. Berlin, Germany: Springer, 2016, pp. 97–114.

[24] R. Zhuang, S. A. DeLoach, and X. Ou, "Towards a theory of moving target defense," in *Proc. 1st ACM Workshop Moving Target Defense*, 2014, pp. 31–40.

[25] Q. Zhu and T. Başar, "Game-theoretic approach to feedback-driven multi-stage moving target defense," in *Proc. Int. Conf. Decis. Game Theory Secur.*, 2013, pp. 246–263.

[26] S. Weerakkody and B. Sinopoli, "Detecting integrity attacks on control systems using a moving target approach," in *Proc. IEEE 54th Annu. Conf. Decis. Control*, 2015, pp. 5820–5826.

[27] K. G. Vamvoudakis, J. P. Hespanha, B. Sinopoli, and Y. Mo, "Detection in adversarial environments," *IEEE Trans. Autom. Control*, vol. 59, no. 12, pp. 3209–3223, Dec. 2014.

[28] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Trans. Autom. Control*, vol. 59, no. 6, pp. 1454–1467, Jun. 2014.

[29] Y. Yan, P. Antsaklis, and V. Gupta, "A resilient design for cyber physical systems under attack," in *Proc. IEEE Amer. Control Conf.*, 2017, pp. 4418–4423.

[30] L. An and G.-H. Yang, "Secure state estimation against sparse sensor attacks with adaptive switching mechanism," *IEEE Trans. Autom. Control*, vol. 63, no. 8, pp. 2596–2603, Aug. 2017.

[31] S. E. McLaughlin, D. Podkuiko, A. Delozier, S. Miadzvezhanka, and P. D. McDaniel, "Embedded firmware diversity for smart electric meters," in *Proc. HotSec*, 2010.

[32] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*. Hoboken, NJ, USA: Wiley 2012.

[33] H. Okhravi, T. Hobson, D. Bigelow, and W. Streilein, "Finding focus in the blur of moving-target techniques," *IEEE Secur. Privacy*, vol. 12, no. 2, pp. 16–26, Mar./Apr. 2014.

[34] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2012.

[35] J. P. Hespanha and A. S. Morse, "Stability of switched systems with average dwell-time," in *Proc. 38th IEEE Conf. Decis. Control*, 1999, vol. 3, pp. 2655–2660.

[36] D. Liberzon, *Switching in Systems and Control*. Berlin, Germany: Springer, 2012.

[37] W. F. Ames and B. Pachpatte, *Inequalities for Differential and Integral Equations*, vol. 197. Amsterdam, The Netherlands: Elsevier, 1997.

[38] D. Vrabie, K. G. Vamvoudakis, and F. L. Lewis, *Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles*, vol. 2. London, U.K.: IET, 2013.

[39] J. Na, G. Herrmann, and K. G. Vamvoudakis, "Adaptive optimal observer design via approximate dynamic programming," in *Proc. IEEE Amer. Control Conf.*, 2017, pp. 3288–3293.

[40] V. Durbha and S. Balakrishnan, "New nonlinear observer design with application to electrostatic micro-actuators," in *Proc. ASME Int. Mech. Eng. Congr. Expo.*, 2005, pp. 101–107.

[41] X. Yu and J. Jiang, "Hybrid fault-tolerant flight control system design against partial actuator failures," *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 4, pp. 871–886, Jul. 2012.

Aris Kanellopoulos (S'17) was born in Athens, Greece. He received the Diploma in mechanical engineering from the National Technical University of Athens, Athens, Greece, in 2016. He is currently working toward the Ph.D. degree at The Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, USA.

In 2018, he was a Research Assistant with the Kevin T. Crofton Department of Aerospace and Ocean Engineering, Virginia Tech. He was a Research Assistant with The Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology. His current research interests include cyber-physical security, game theory, optimal, and learning-based control.

Kyriakos G. Vamvoudakis (SM'15) was born in Athens, Greece. He received the Diploma (a five year degree, equivalent to a Master of Science) in electronic and computer engineering from Technical University of Crete, Chania, Greece, in 2006 with highest honors and the M.S. and Ph.D. degrees in electrical engineering The University of Texas, Austin, TX, USA, in 2008 and 2011, respectively.

During the period from 2012 to 2016, he was a Project Research Scientist with the Center for Control, Dynamical Systems and Computation, University of California, Santa Barbara. He was an Assistant Professor with the Kevin T. Crofton Department of Aerospace and Ocean Engineering, Virginia Tech, in 2018. He is currently an Assistant Professor with the Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, USA. His research interests include optimal control, reinforcement learning, and game theory. Recently, his research has focused on cyber-physical security, and safe autonomy.

Prof. Vamvoudakis is the recipient of a 2019 ARO YIP Award, a 2018 National Science Foundation CAREER Award, the 2016 International Neural Network Society Young Investigator (INNS) Award, the Best Paper Award for Autonomous/Unmanned Vehicles at the 27th Army Science Conference in 2010, and the Best Researcher Award from the Automation and Robotics Research Institute in 2011. He currently is an Associate Editor for *Automatica*, an Associate Editor for the IEEE Computational Intelligent Magazine, an Associate Editor for the IEEE CONTROL SYSTEMS LETTERS, an Associate Editor for *Journal of Optimization Theory and Applications*, an Associate Editor for *Control Theory and Technology*, a Registered Electrical/Computer Engineer (PE), and a member of the IEEE Control Systems Society Conference Editorial Board.