2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Structural Analysis and Design for Security against Actuator Stealthy Attacks in Uncertain Systems

Kangkang Zhang, Andreas Kasis, Marios M. Polycarpou, Thomas Parisini

Abstract— **This paper considers the existence of stealthy integrity attacks for uncertain cyber-physical systems from a geometric point of view. We derive geometric structural conditions for the existence of stealthy integrity attacks and deduce the minimal actuator communication channels that, when protected, no stealthy integrity attacks exists. To examine different knowledge disclosure conditions for the attacker, we consider: (a) the attacker has full knowledge of the system linear terms but only the structure of the uncertain term, and (b) the attacker only knows the structures of the linear terms and the uncertain non-linear term. For scenario (a), the obtained existence condition of stealthy integrity attacks is that the uncertainty is decoupled with the maximal output-zeroing controlled-invariant subspace. In scenario (b), a graph is used to describe the uncertain system and we show that the existence of stealthy attacks is only possible if the uncertainty is decoupled with the fixed maximal output-zeroing controlled-invariant subspace. For each disclosure scenario, we deduce the minimum actuator communication channels to protect for guaranteeing the absence of stealthy integrity attacks. Our results are validated with a numerical example.**

I. INTRODUCTION

Motivation and literature review: The cyber-physical system (CPS) framework is applicable to a broad range of systems, such as smart power grids, intelligent transportation networks, and water distribution networks. However, the rising integration of control, computation and communication techniques makes CPS more vulnerable to malicious cyber attacks [1]. Hence, significant research focus is paid on the risk management of CPS [2]. This includes a broad variety of security topics such as attack scenario description [3], [4], attack analysis [5], [6] and attack mitigation [7]–[10]. An important topic of security analysis is that of attack prevention, which mainly aims to eliminate the vulnerability to stealthy integrity attacks in CPS.

Two approaches that have been utilized for attack prevention are information security techniques and system structure design. Information security approaches such as encryption of the communication channels, firewalls and adding stochastic noise, mainly related to secure information flow (privacy preservation) are well developed in [11]–[13]. On the other hand, suitably designing the system structure is an efficient way for securing the integrity of information flow, thus enabling the prevention of stealthy attack events. Recent studies have explored actuator protection design [14], [15], vertex communication structure design [16], and sensor placement design [6] for networked linear systems.

The authors in [14], [15] determine the actuator security indices for systems with known parameters and uncertain structural systems respectively. The structural left invertibility property is used to characterize perfectly undetectable integrity attacks in [16]. In [6], the optimal sensor placement problem is formulated as a zero-sum game.It should be noted that most of the current literature addresses perfectly undetectable attacks, which is not the most general class of stealthy attacks. In addition, the security structure design for uncertain systems has not been explored. Motivated by the aforementioned issues, this paper studies the problem of security structural design against general stealthy attacks.

Contribution: This paper studies the problem of the existence of stealthy integrity attacks for uncertain cyber-physical systems by considering their structural properties. To examine different knowledge disclosure conditions for the attacker, we consider: (a) the attacker has full knowledge of the system linear terms but only the structure of the uncertain term, and (b) the attacker only knows the structures of the linear terms and the uncertain non-linear term. In particular, for scenario (a), we show that stealthy attacks exist if and only if the uncertainty is decoupled with the maximal output-zeroing controlled-invariant subspace. Moreover, the uncertain system is considered as a structural system in scenario (b). We use a graph to describe this system and specify a fixed output-zeroing controlled-invariant subspace that is decoupled with the uncertainty using suitable graphic properties, and show that no stealthy integrity attacks exist if and only if this subspace is empty. In each scenario, we determine the minimum number of actuator communication channels that when protected, then the non-existence of stealthy integrity attacks is guaranteed.

Preliminaries: The notation e_k represents a vector with the kth element being 1 and all other elements being 0. Then, the vectors $e_1, \cdots, e_n$ constitute a unit basis for $\mathbb{R}^n$. The sets of real matrices of dimension $m \times n$ with real and binary entries are denoted by $\mathbb{R}^{n\times n}$ and $\mathbb{B}^{n\times n}$ respectively, where $\mathbb{B} = \{0, 1\}$. The kernel and image of $A \in \mathbb{R}^{n\times n}$ are denoted

This work has been supported by: the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101027980 (CSP-CPS-A-ICA), No. 891101 (SmarTher Grid) and No. 739551 (KIOS CoE-TEAMING), the Italian Ministry for Research in the framework of the 2017 Program for Research Projects of National Interest (PRIN) (grant No.2017YKXYXJ).

K. Zhang and T. Parisini are with the Dept. of Electrical and Electronic Engineering, Imperial College London, London, SW7 2AZ, UK. T. Parisini is also with the Dept. of Engineering and Architecture, University of Trieste, Trieste, 34127, Italy, kzhang5@ic.ac.uk;t.parisini@gmail.com

A. Kasis and M. Polycarpou are with the KIOS Research and Innovation Center of Excellence, and the Dept. of Electrical and Computer Engineering, University of Cyprus, Nicosia, 1678, Cyprus kasis.andreas@ucy.ac.cy;mpolycar@ucy.ac.cy

by $\ker A$ and $\mathrm{Im}A$ respectively. For $A \in \mathbb{R}^{n\times m}$ and a set S, $|A|_0$ is the number of the nonzero elements of A. For a set S, $|S|$ represents the cardinality of S. In addition, we use $\oplus$ to represent the direct sum of two subspaces. For a vector signal $x(t)$, $x(t) \equiv 0$ for $t \in [t_1, t_2]$ means that $x(t) = 0$ identically for all $t \in [t_1, t_2]$; $x(t) \not\equiv 0$ for $t \in [t_1, t_2]$ means $x(t) \neq 0$ for at least one time instant $t \in [t_1, t_2]$.

Some definitions for controlled invariant subspace are given below. To this end, the state and output under initial condition x_0 and input $\bar{u}$ are denoted by $x(x_0, \bar{u}, t)$ and $y(x_0, \bar{u}, t)$ respectively.

Definition 1. [17, Dfn. 4.1] A subspace $\mathcal{V}$ is controlled invariant for system (A, B, C) if for any $x_0 \in \mathcal{V}$, there exists an input $\bar{u}$ such that $x(x_0, \bar{u}, t) \in \mathcal{V}$ for all $t \geq 0$. □

Definition 2 ((A, B) Invariant Subspace Contained in $\ker C$)**.** For the system (A, B, C), the following statements are equal:

- A subspace $\mathcal{V}$ is a controlled invariant subspace contained in $\ker C$;
- $A\mathcal{V} \subset \mathcal{V} + \mathrm{Im}B$ and $C\mathcal{V} = 0$;
- There exists a matrix L with proper dimensions such that $(A + BL)\mathcal{V} \subset \mathcal{V}$ and $C\mathcal{V} = 0$. □

To describe structural matrices, we introduce the concept of pattern matrices with elements being in the set $\{0, *\}$. The set of all $m \times n$ pattern matrices is denoted by $\{0, *\}^{m\times n}$. For a given $m \times n$ pattern matrix $\mathscr{M}$, we define the pattern class of $\mathscr{M}$ as follows:

$$\mathscr{P}(\mathscr{M}) = \{M \in \mathbb{R}^{m\times n} | M_{ij} = 0 \ if \ \mathscr{M}_{ij} = 0, \\ M_{ij} \neq 0 \ if \ \mathscr{M}_{ij} = *\}.$$

II. Problem Formulation

In this section we provide a mathematical description of the CPS considered in this paper, define stealthy integrity attacks and state the problem examined in this paper.

A. System Description

A general structure of CPS subject to integrity type cyber attacks is depicted in Fig. 1. The attacker compromises the

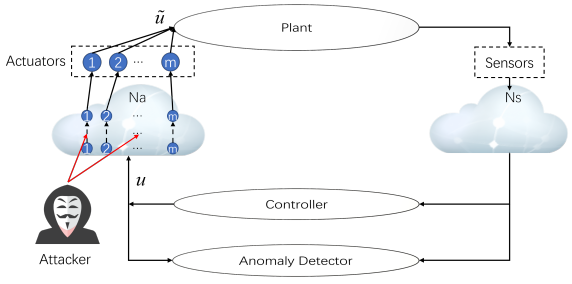

Fig. 1. General architecture of CPSs under potential integrity cyber attacks.

actuator communication network $\mathcal{N}_a$ by injecting additive false data. Suppose that the CPS has m actuator communication channels. Let $K \subseteq \{1, \cdots, m\}$ represent the set of disruption resources, i.e., the set of actuator communication channels that may be affected by the adversary. Throughout this paper, we use T_0 to denote the attack event occurrence time. Then, in the presence of an actuator attack associated with the disruption resource K at T_0, the transmitted and received control signals by $\mathcal{N}_a$, denoted by $\tilde{u}$ and u respectively, satisfy the following relation:

$$\tilde{u}(t) = u(t) + \Gamma a(t), \ \forall \ t \geq T_0, \tag{1}$$

where $a = [a_1, \cdots, a_k]^T \in \mathbb{R}^{|K|}$ is the actuator attack. For each $i \in K$, $a_i(t) \equiv 0$ for $t \geq 0$ if there is no attack occurring on the i-th $\mathcal{N}_a$ channel. The distribution matrix $\Gamma \in \mathbb{B}^{m\times|K|}$ is the binary diagonal matrix related to the disruption resources K.

The closed-loop CPS, denoted by $\mathcal{W}$ in the attack case, is described by

$$\mathcal{W}: \begin{cases} \dot{x} = Ax + Ff(t, x) + B\tilde{u} + \omega_1, \ x(0) = x_0, & \text{(2a)} \\ \tilde{u} = u(y, y_{\text{ref}}) + \Gamma a(t), & \text{(2b)} \\ y = Cx + \omega_2, & \text{(2c)} \end{cases}$$

where $x \in \mathbb{R}^n$ is the state, $\tilde{u}$, $u \in \mathbb{R}^m$ are the control data received by the actuator and computed by the controller respectively, and $y \in \mathbb{R}^p$ denote the sensor measurements received by the controller. The signal $y_{\text{ref}} \in \mathbb{R}^p$ is the reference signal, and $\omega_1 \in \mathbb{R}^n$ and $\omega_2 \in \mathbb{R}^m$ represent the lumped disturbances and noise. The function $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}^q$ represents the lumped unmodelled uncertainties. The function $u : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^m$ is the nonlinear output feedback control law. Moreover, u is continuous in y and y_{ref}, and satisfies $u(0, 0) = 0$. Finally, the matrices A, B, F and C have proper dimensions.

In the absence of attacks, $\mathcal{W}$ is denoted by $\mathcal{W}_n$ given by:

$$\mathcal{W}_n: \begin{cases} \dot{x}_n = Ax_n + Ff(t, x_n) + B\tilde{u}_n + \omega_1, & \text{(3a)} \\ \tilde{u}_n = u(y_n, y_{\text{ref}}), & \text{(3b)} \\ y_n = Cx_n + \omega_2, & \text{(3c)} \end{cases}$$

where x_n, y_n, and $\tilde{u}_n$, are the state, output, and the received control input in the nominal case, corresponding to x_n, y_n, and $\tilde{u}_n$, respectively.

B. Stealthy Integrity Attacks

In this section, we provide a rigorous definition of stealthy integrity attacks for systems described by (2). To this end, we let $y(x_0, a, t)$ and $y_n(x_0 + \Delta x_0, 0, t)$ denote the output of system $\mathcal{W}$ with the initial condition x_0 and input a and the output of system $\mathcal{W}_n$ with the initial condition $x_0 + \Delta x_0$ respectively. Below, we provide a description of the typical model-based anomaly detector depicted in Fig. 1. The residual r of the anomaly detector is a function of y. In addition, since the initial state x_0 is frequently not exactly known, the detector threshold, denoted by $\bar{r}$, is typically designed as a function of $x_0 + \Delta x_0$ for some bounded value of Δx_0. Typically, no alarm is triggered by the detector in the presence of the output $y_n(x_0 + \Delta x_0, 0, t)$, since $|r(t)| < \bar{r}(t)$ for any t and any Δx_0 that abides by certain bound conditions. Stealthy integrity attacks with respect to the typical model-based anomaly detector is defined as follows.

Definition 3 (Stealthy Integrity Attacks)**.** Consider the system $\mathcal{W}$. Then, an attack $a(t) \not\equiv 0$ for $t \geq T_0$ is a stealthy

integrity attack with respect to typical model-based anomaly detectors if there exists Δx_0 such that

$$y(x_0, a, t) = y_n(x_0 + \Delta x_0, 0, t), \ \forall\ t \geq T_0. \tag{4}$$

C. Problem Statement

Below we state the problems considered in this paper, which are explored in the following sections.

Problem 1. *By analyzing the structure of the nonlinear closed-loop system $\mathcal{W}$, described by (2), we intent to*

(a) *obtain geometric conditions that enable the nonexistence of stealthy integrity attacks;*

(b) *determine the minimum actuator communication channels that, when secured, it is guaranteed that no stealthy integrity attacks exist,*

when the attacker has knowledge of either of the following:

(i) *A, B, and C and the partial structure of $Ff(t,x)$;*

(ii) *the partial structures of A, B and C and $Ff(t,x)$.*

Cases (i) and (ii) are considered in Sections III and IV respectively.

III. Security analysis under uncertain nonlinear dynamics

This section explores the problem described in Section II-C when the attacker side has complete knowledge of the matrices A, B and C and the partial structure of $Ff(t,x)$, of the closed-loop system $\mathcal{W}$, described by (2).

A. Geometric Existence Condition

This subsection provides geometric conditions for the existence of stealthy integrity attacks when the attacker knows A, B, and C, but has uncertain knowledge of $Ff(t,x)$. More specifically, the attacker has knowledge of all elements of the matrices A, B, and C, and the zero entries of the matrix F, associated with (2), while the remaining elements of F are unknown. Recalling the knowledge resources known by the attacker, a pattern matrix $\mathscr{F}$ can be defined such that, from the attacker's point of view, F belongs to a pattern class $\mathscr{P}(\mathscr{F})$. For the matrix F with uncertain parameters $*$, its image $\mathrm{Im}F$ varies with respect to these uncertain parameters. In the analysis that follows, we denote the smallest fixed subspace containing generically $\mathrm{Im}F$ by $\mathcal{F}_F$. The space $\mathcal{F}_F$ can be described through the unit basis vector e_k.

In addition, in order to focus on the system security structure analysis and not be misled by the knowledge of system states, we assume that the attacker access to all system states at all times. Moreover, we assume the worst case that the attacker is able to disrupt only a subset of the inputs, which is denoted by K.

An incremental system of $\mathcal{W}_n$ in (3) with initial condition $x_0 + \Delta x_0$ is utilized to determine the existence of attacks. To define such a system, we denote $\Delta x = x - x_n$ and $\Delta y = y - y_n$. Then, from (2a), (2b), (3a) and (3b), the changes in x and y due to an attack can be expressed as

$$\Delta\mathcal{W}: \begin{cases} \Delta\dot{x} &= A\Delta x + F(f(t,x) - f(t,x_n)) + B\Gamma a, \\ \Delta y &= C\Delta x, \end{cases} \tag{5}$$

where $\Delta x(T_0) = \Delta x_0$. Note that $u(y, y_{\mathrm{ref}}) - u(y_n, y_{\mathrm{ref}}) = 0$ is used to derive the above incremental system.

Remark 1. It can be observed from (5) that the nonlinearity disappears if $f(t,x)$ is independent on x since $f(t,x) - f(t,x_n) = 0$. Hence, in the case that $f(t,x)$ is independent on x, the attack existence problem is trivial since many associated works have been done (see e.g., [4], [18]). The considered nonlinear function $f(t,x)$ that depends on x brings significant challenges. It is also worth pointing out that the existence of the attack does not rely on the form of the nonlinear function $f(t,x)$, while depends on the distribution matrix F. Hence, this paper does not require any limitation for the uncertain nonlinear function $f(t,x)$. ∇

The attack is conducted based on the incremental system (5). Recalling that all disclosure resources are supposed to be available to the attacker, the attack a is designed as $a = L\Delta x$ such that the output Δy is decoupled with the nonlinearity $Ff(t,x)$, i.e., in the context of the transfer function, $C(sI - A - B\Gamma L)^{-1}F = 0$ for any $F \in \mathscr{P}(\mathscr{F})$. The existence of such matrix L implies the existence of a stealthy integrity attack satisfying Definition 3. Hence, the existence of stealthy attacks is equivalent to the existence of a matrix L that satisfies the above condition.

Recalling Definition 2, we characterize the $(A, B\Gamma)$ invariant subspace $\mathcal{V}$ contained in $\ker C$ as follows:

$$A\mathcal{V} \subset \mathcal{V} + \mathrm{Im}B\Gamma,\ C\mathcal{V} = 0. \tag{6}$$

Let $\mathcal{V}^*$ be the maximal $(A, B\Gamma)$ invariant subspace contained in $\ker C$. Then, the following theorem provides conditions for the existence of stealthy integrity attacks for system $\mathcal{W}$. In addition, we remind that $\mathcal{F}_F$ denotes the smallest fixed subspace containing $\mathrm{Im}F$.

Theorem 1 (Geometric Existence Condition)**.** *There exists a stealthy integrity attack for the system $\mathcal{W}$ in* (2) *if and only if $\mathcal{F}_F \subset \mathcal{V}^*$.* ■

B. Minimal Actuator Protection for Security

The aim of this subsection is to determine the minimum key actuator communication channels to secure such that no stealthy integrity attack exists even though the attacker can disrupt all the remaining actuators. By protecting these key actuator communication channels, the attacker has limited actuator disruption resources to obtain the required space $\mathcal{V}^*$ and thus, the defender can manipulate the inclusion relationship between $\mathcal{F}_F$ and $\mathcal{V}^*$. Hence, following from the result in Theorem 1, the existence of the stealthy attack can be eliminated. The objective can be characterized by the following optimization problem:

$$\max_{\Gamma} |\Gamma|_0 \tag{7a}$$

$$s.\ t.\ \exists\ \mathcal{V} \neq \emptyset, \tag{7b}$$

$$\mathcal{F}_F \subset \mathcal{V},\ A\mathcal{V} \subset \mathcal{V} + \mathrm{Im}B\Gamma,\ C\mathcal{V} = 0. \tag{7c}$$

The above optimization problem is studied based on a system decomposition presented below. In particular, we consider the case where $\mathcal{F}_F \subset \ker C$, which is necessary for the

existence of stealthy attacks, as follows from Theorem 1. Suppose that the matrix C is full row rank and B is full column rank. Let $\mathcal{V} \subset \mathbb{R}^n$ be a state subspace satisfying

$$\mathcal{F}_F \subset \mathcal{V} \subset \ker C. \tag{8}$$

Then, we have $q \leq d \triangleq \dim(\mathcal{V}) \leq n-p$, and we can also define the quotient space $\mathbb{R}^n/\mathcal{V}$ and a linear map $\Pi : \mathbb{R}^n \to \mathbb{R}^n/\mathcal{V}$. (the notation Π represents the corresponding matrix of the linear map Π). Moreover, based on $\mathbb{R}^n/\mathcal{V}$, we can define two full column rank matrices $B_1 \in \mathbb{R}^{d\times m_1}$ and $B_2 \in \mathbb{R}^{(n-d)\times m_2}$ with $m_1+m_2=m$, satisfying

$$\mathrm{Im}\begin{bmatrix} B_1 \\ 0_{(n-d)\times m_1}\end{bmatrix} \subset \mathcal{V},\ \mathrm{Im}\begin{bmatrix} 0_{d\times m_2} \\ B_2\end{bmatrix} \subset \mathbb{R}^n/\mathcal{V},$$

$$\mathrm{Im}B = \begin{bmatrix} B_1 \\ 0_{(n-d)\times m_1}\end{bmatrix} \oplus \begin{bmatrix} 0_{d\times m_2} \\ B_2\end{bmatrix}.$$

Corresponding to the dimensions of B_1 and B_2, we have $\Gamma_1 \in \mathbb{R}^{m_1\times m_1}$ and $\Gamma_2 \in \mathbb{R}^{m_2\times m_2}$ such that $\Gamma = \mathrm{diag}(\Gamma_1, \Gamma_2)$.

Let $\{e_1, \cdots, e_d, \cdots, e_n\}$ be a basis of $\mathbb{R}^n$ with $\{e_1, \cdots, e_d\}$ being a basis of $\mathcal{V}$. Then, in this basis, the system $(A, [B\Gamma, F], C)$ can be written in the following form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ \Pi & A_{22}\end{bmatrix},\ B = \begin{bmatrix} B_1\Gamma_1 & 0 \\ 0 & B_2\Gamma_2\end{bmatrix}, \tag{9a}$$

$$F = \begin{bmatrix} F_1 \\ 0\end{bmatrix},\ C = [0\ \ C_2], \tag{9b}$$

where $\Pi \in \mathbb{R}^{(n-d)\times d}$, $A_{11} \in \mathbb{R}^{d\times d}$, $A_{22} \in \mathbb{R}^{(n-d)\times(n-d)}$, $F_1 \in \mathbb{R}^{d\times q}$ and $C_2 \in \mathbb{R}^{p\times(n-d)}$. It should be noted that C_2 is full row rank since C is full row rank. Hence, we are ready to present a result as follows:

Lemma 1. *The conditions (7b) and (7c) hold for some given fixed space $\mathcal{V}$ satisfying (8) if and only if the actuator communication channels accessed by the attacker satisfy* $\mathrm{Im}B_2\Gamma_2 \supseteq \mathrm{Im}\Pi$. *Moreover, the minimum number of such channels is given by* $|\Gamma|_0 = \mathrm{rank}\Pi$ *where* $\Gamma = \mathrm{diag}(\Gamma_1, \Gamma_2)$ *with* $\Gamma_1 = 0$. ■

Remark 2. It should be noted that the number of the $\mathcal{V}$ satisfying (8) is finite for a system with finite dimensions. Hence, based on the result of Lemma 1, the optimization problem (7) can be solved by using the brute-force algorithm [14]. In particular, we can derive the minimum number of the actuator communication channels to protect for each $\mathcal{V}$ satisfying (8). The infimum of the minimum numbers is the objective value $\min_\Gamma |I_m - \Gamma|_0$, and can be obtained by comparing them. However, this brute-force algorithm is not applicable to high-dimension systems due to the computation complexity of this algorithm. Hence, we derive the analytically optimal solution to (7) presented in the following theorem. ▽

The following theorem characterizes the solution to (9), by determining the minimum number of actuators that, when protected, the non-existence of stealthy attacks is guaranteed.

Theorem 2. *Consider the optimization problem* (7). *The minimum number of actuator communication channels to secure is given by* $|I_m - \Gamma^*|_0 + 1 = m - \mathrm{rank}\Pi^* + 1$ *where* $\Pi^* : \mathbb{R}^n \to \mathbb{R}^n/\mathcal{V}^*$ *with* $V^* = \ker C$ *and* $\Gamma^* = \mathrm{diag}(0, \Gamma_2^*)$ *with* Γ_2^* *satisfying* $\mathrm{Im}B_2\Gamma_2^* \supseteq \Pi^*$. ■

IV. Security analysis under uncertain linear and nonlinear dynamics

This section considers the case where the attacker is only aware of the partial structure of matrices A and B and the uncertainty $Ff(t,x)$ of system $\mathcal{W}$ in (2). This significantly complicates the analysis compared to the case presented in Section III, due to the lack of exact knowledge of the matrices A and B. In particular, the attacker knows only the structure of the matrices A, B and F. Let $\mathcal{A} \in \{0,*\}^{n\times n}$, $\mathcal{B} \in \{0,*\}^{n\times m}$ and $\mathscr{F} \in \{0,*\}^{n\times q}$. Then, from the attacker's point of view, A, B and F belong to the pattern classes given below:

$$A \in \mathscr{P}(\mathcal{A}),\ B \in \mathscr{P}(\mathcal{B}),\ F \in \mathscr{P}(\mathscr{F}). \tag{10}$$

In addition, the attacker knows the kernel of matrix C, i.e. is aware of which states are not measured.

A. Geometric Existence Condition

Let $\mathcal{V}$ be the (A,B) invariant subspace contained in $\ker C$ (see Definition 2), and let $\mathcal{V}^*$ be the maximal (A,B) invariant subspace contained in $\ker C$. Then, we define $\mathcal{V}_F$ as a fixed subspace contained in $\mathcal{V}$ and $\mathcal{V}_F^*$ as the maximal fixed subspace contained in $\mathcal{V}^*$ for all $A \in \mathscr{P}(\mathcal{A})$ and $B \in \mathscr{P}(\mathcal{B})$. The following result provides conditions for the existence of the stealthy integrity attack.

Theorem 3 (Geometric Existence Condition)**.** *Consider the structural system $\mathcal{W}$ described by* (2). *There exists a stealthy integrity attack for $\mathcal{W}$ with respect to $A \in \mathscr{P}(\mathcal{A})$, $B \in \mathscr{P}(\mathcal{B})$ and $F \in \mathscr{P}(\mathscr{F})$ if and only if $\mathcal{F}_F \subset \mathcal{V}_F^*$.* ■

B. Minimal Protection of Actuator Communication Channels

The aim of this subsection is to determine, for the case considered in this section, the minimal key actuator communication channels that, when protected, ensure that no stealthy attack exists. The latter is achieved by utilizing the inclusion relation $\mathcal{F}_F \subset \mathcal{V}_F$, which is a requirement for the existence of stealthy integrity attacks (see Theorem 3). In analogy to previous section, this aim is achieved by solving

$$\max_{\Gamma} |\Gamma|_0 \tag{11a}$$

$$s.\ t.\ \exists\ \mathcal{V}_F \neq \emptyset, \tag{11b}$$

$$\mathcal{F}_F \subset \mathcal{V}_F,\ A\mathcal{V}_F \subseteq \mathcal{V}_F + \mathrm{Im}B\Gamma,\ C\mathcal{V}_F = 0. \tag{11c}$$

Since $\mathcal{F}_F \subset \ker C$, we choose a fixed $\mathcal{V}_F \subset \mathbb{R}^n$ satisfying

$$\mathcal{F}_F \subset \mathcal{V}_F \subset \ker C. \tag{12}$$

Then, we have $q \leq d \triangleq \dim(\mathcal{V}_F) \leq n-p$. Similar to the previous section, by replacing $\mathcal{V}$ with $\mathcal{V}_F$, we define the matrices Π, B_1 and B_2, and the associated matrices Γ_1 and Γ_2. The subspace $\mathcal{V}_F$ satisfies $\mathcal{F}_F \subset \mathcal{V}$ in (11b) and $C\mathcal{V}_F = 0$ in (11c). Moreover, by using a state permutation based on $\mathscr{P}(\mathscr{F})$ and $\ker C$, the system can be transformed into the form presented in (9). Let $[w_1^T, w_2^T]^T$ be the state of the

system described in (9). Then, the subsystem that describes the evolution of state w_2 is given by

$$\hat{\Sigma}: \begin{cases} \dot{w}_2 = A_{22}w_2 + B_2\Gamma_2 u_2 + \Pi w_1, & \text{(13a)} \\ z \ \ = w_2. & \text{(13b)} \end{cases}$$

In the context of system $\hat{\Sigma}$, $A\mathcal{V}_F \subseteq \mathcal{V}_F + \mathrm{Im}B\Gamma$ in (11c) is equivalent to a disturbance decoupling problem of system $\hat{\Sigma}$, i.e., z is decoupled with w_1 by state feedback u_2 through the input distribution matrix $B_2\Gamma_2$. Hence, below we present a graphic condition based on system $\hat{\Sigma}$ that enables to deduce the minimum number of actuators to be protected such that no stealthy integrity attacks exist.

System $\hat{\Sigma}$ in (13) can be represented by a direct graph $\mathbb{G}(\hat{\Sigma}) = (\mathbb{V}, \mathbb{E})$ where the vertex set is $\mathbb{V} = \mathbb{X} \cup \mathbb{U} \cup \mathbb{D} \cup \mathbb{Z}$ with $\mathbb{X} = \mathbb{Z} = \{w_{21}, \cdots, w_{2(n-d)}\}$, $\mathbb{U} = \{u_{21}, \cdots, u_{2m_2}\}$ and $\mathbb{D} = \{w_{11}, \cdots, w_{1d}\}$. The edge set $\mathbb{E} \subseteq \mathbb{V} \times \mathbb{V}$ represents the connections/communications among w_1, w_2, u_2 and z. A path P in $\mathbb{G}(\hat{\Sigma})$ from a vertex i_0 to a vertex i_l is a sequence of edges $(i_0, i_1), (i_1, i_2), \cdots, (i_{l-1}, i_l)$ such that $i_j \in \mathbb{V}$ for $j = 0, 1, \cdots, l$ and $(i_{j-1}, i_j) \in \mathbb{E}$ for $j = 1, 2, \cdots, l$. If $i_0 \in \mathbb{V}_s$ and $i_l \in \mathbb{V}_e$, where $\mathbb{V}_s$ and $\mathbb{V}_e$ are two subsets of $\mathbb{V}$, P is called a $\mathbb{V}_s - \mathbb{V}_e$ path. Moreover, if i_0 is the only vertex belonging to $\mathbb{V}_s$ and $i_l \neq i_0$ is the only vertex belonging to $\mathbb{V}_e$, P is called a direct $\mathbb{V}_s - \mathbb{V}_e$ path. A set of paths with no common vertex is said to be vertex disjoint. A $\mathbb{V}_s - \mathbb{V}_e$ linking of size k is a set of k vertex disjoint $\mathbb{V}_s - \mathbb{V}_e$ paths. A linking is maximal when k is maximal.

Let Σ_1 and Σ_2 represent the system $\hat{\Sigma}$ without input Πw_1, and the system $\hat{\Sigma}$ without input $B_2\Gamma_2 u_2$, respectively, that is

$$\Sigma_1: \begin{cases} \dot{w}_2 = A_{22}w_2 + B_2\Gamma_2 u_2, & \text{(14a)} \\ z \ \ = w_2, & \text{(14b)} \end{cases}$$

$$\Sigma_2: \begin{cases} \dot{w}_2 = A_{22}w_2 + \Pi w_1, & \text{(15a)} \\ z \ \ = w_2. & \text{(15b)} \end{cases}$$

Correspondingly, $\mathbb{G}(\Sigma_1) = (\mathbb{V}_1, \mathbb{E}_1)$ and $\mathbb{G}(\Sigma_2) = (\mathbb{V}_2, \mathbb{E}_2)$ are the graphs of Σ_1 and Σ_2 respectively, where $\mathbb{V}_1 = \mathbb{X} \cup \mathbb{U} \cup \mathbb{Z}$, $\mathbb{E}_1 \subseteq \mathbb{V}_1 \times \mathbb{V}_1$, $\mathbb{V}_2 = \mathbb{X} \cup \mathbb{D} \cup \mathbb{Z}$ and $\mathbb{E}_2 \subseteq \mathbb{V}_2 \times \mathbb{V}_2$. To facilitate the subsequent analysis, we provide the following definitions for systems $\mathbb{G}(\Sigma_i), i \in \{1, 2\}$.

Definition 4. [19] A separator S of the graph $\mathbb{G}(\Sigma_i), i \in \{1, 2\}$ is a set of vertices of $\mathbb{G}(\Sigma_i)$ such that any $\mathbb{U} - \mathbb{Z}$ (or $\mathbb{D} - \mathbb{Z}$) path has at least one vertex in S. □

Definition 5. A separator S is an output separator if $|S'| > |S|$ for any separator S' such that any direct $S - \mathbb{Z}$ path contains a vertex in S'. □

Among all the output separators, the output separator which has the minimal dimension is unique [20], and such separator is referred to as the minimal output separator S_0. Then, we have the following result.

Lemma 2. *The conditions (11b) and (11c) hold for some given fixed space $\mathcal{V}$ if and only if the inputs in $\mathbb{U}$ connect to the vertices of a separator S' between $\mathbb{D}$ and S_0 satisfying $S' \cap \mathbb{D} = \emptyset$. Moreover, the minimal number of actuator communication channels that have to be accessed by the attacker is equal to $|S_0|$.* ■

Similar to Theorem 2, we derive the analytically optimal solution to the optimizaiton problem (11) below. Let S_0^* be the minimum output separator of $\mathbb{D} - \mathbb{Z}$ when $\mathcal{V}_F = \ker C$. Then, we have the following result.

Theorem 4. *Consider the optimization problem in (11). The minimum number of actuator communication channels to secure is given by $|I_m - \Gamma^*|_0 + 1 = m - |S_0^*| + 1$ where $\Gamma^* = \mathrm{diag}(\Gamma_1^*, \Gamma_2^*)$ with $\Gamma_1^* = 0$ and the actuators characterized by Γ_2^* connected to the minimal output separator S_0^*.* ■

V. Illustrative Numerical Example

In this section, we validate our analytical results with a suitable numerical example. In particular, we consider a system described by (2) with the following matrices:

$$A = \begin{bmatrix} 0 & \lambda_1 & \lambda_2 & 0 \\ \lambda_1 & 0 & 0 & \lambda_2 \\ -\lambda_3 & 0 & -\lambda_3 & -\lambda_1 \\ 0 & -\lambda_3 & -\lambda_1 & -\lambda_3 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 & 0 \\ \lambda_4 & 0 & 0 \\ 0 & -\lambda_5 & 0 \\ 0 & 0 & -\lambda_5 \end{bmatrix}, \quad \text{(16a)}$$

$$F = \begin{bmatrix} \lambda_6 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 & \lambda_7 & 0 \\ 0 & 0 & 0 & \lambda_7 \end{bmatrix}, \quad \text{(16b)}$$

where $\lambda_1, \cdots, \lambda_7$ are constant parameters. It can be observed from Theorems 2 and 4 that the form of the nonlinear uncertainty $f(t, x)$ does not affect the main results. Hence, a description of $f(t, x)$ is not required. Due to space limitation, we only validate the results in Theorem 4.

In this section, we also use the brute-force algorithm to calculate the minimum number of the actuator channels to protect and then compare the obtained result with that derived based on Theorem 4. The values of the parameters $\lambda_1, \cdots, \lambda_7$ are supposed to be unknown by the attacker. However, the attacker knows the locations of the zero entries in A, B, C and F. Based on the attacker's knowledge, this system has a structure given by the graph shown in Fig. 2. To this end, the procedure to use the brute-force algorithm

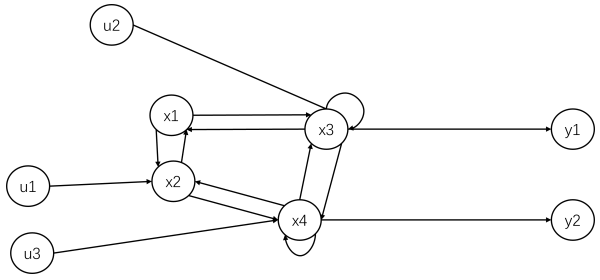

Fig. 2. Graphic description of the considered numerical system.

is given below. In order to satisfy (12), $\mathcal{V}$ can be selected as $\mathcal{V}_1 = \mathrm{Im}[1, 0, 0, 0]^T$ or $\mathcal{V}_2 = \mathrm{Im}[I_2, 0_{2\times 2}]^T$. In the case $\mathcal{V} = \mathcal{V}_1$, we have $w_1 = x_1$ and $w_2 = [x_2, x_3, x_4]^T$. The graph $\mathbb{G}(\hat{\Sigma})$ in this case is depicted in Fig. 3. We list the output separators of $\mathbb{G}(\Sigma_2)$ as follows:

- $S_1 = \{z_1, z_2, z_3\}$ is an output separator with $|S_1| = 3$;
- $S_2 = \{x_2, x_3, x_4\}$ is an output separator with $|S_2| = 3$;
- $S_3 = \{x_2, x_3\}$ is an output separator with $|S_3| = 2$.

Hence, the minimal output separator is $S_0 = S_3 = \{x_2, x_3\}$. Based on Lemma 2, the minimal number of actuators that the attacker has to access is $|S_0| = 2$, and these actuators have to connect to x_2 and x_3. Hence, the minimal number of the actuator communication channels to protect is $|I_3 - \Gamma|_0 + 1 = 2$ where $\Gamma = \mathrm{diag}(I_2, 0)$.

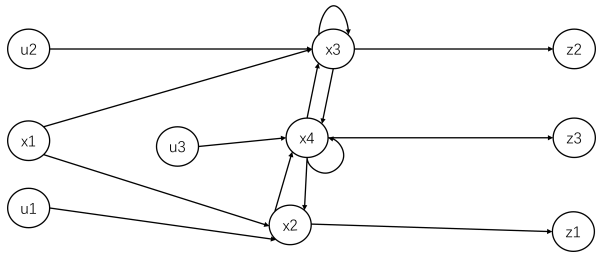

Fig. 3. Graph of $\mathbb{G}(\hat{\Sigma})$ in the case $\mathcal{V} = \mathcal{V}_1$.

In the case $\mathcal{V} = \mathcal{V}_2$, we have $w_1 = [x_1, x_2]^T$ and $w_2 = [x_3, x_4]^T$, and the associated graph $\mathbb{G}(\hat{\Sigma})$ is depicted in Fig. 4. Then, we list the output separators of $\mathbb{G}(\hat{\Sigma}_2)$ as follows:

- $S_1 = \{z_1, z_2\}$ is an output separator with $|S_1| = 2$;
- $S_2 = \{x_3, x_4\}$ is an output separator with $|S_2| = 2$.

Hence, the minimal output separator is $S_0 = S_1 = \{z_1, z_2\}$. Based on Lemma 2, the minimal number of the actuator communication channels that the attacker has to access is $|S_0| = 2$, and these actuators have to connect to z_1 and z_2 (x_3 and x_4). Hence, the minimal number of the actuator communication channels to protect is $|I_3 - \Gamma|_0 + 1 = 2$ where $\Gamma = \mathrm{diag}(0, I_2)$. It can be observed from the calculation in both cases $\mathcal{V} = \mathcal{V}_1$ and $\mathcal{V}_2$ that the minimum number of the actuator communication channels to protect is 2.

We turn to show the result derived based on Theorem 4. Given the minimal output separator $S_0^* = \{x_3, x_4\}$, based on Theorem 4, the minimal number of the actuator communication channels to protect is $3 - |S_0| + 1 = 2$, and these channels connect to x_3 and x_4. This result coincides with that obtained using the brute-force algorithm and validates Theorem 4.

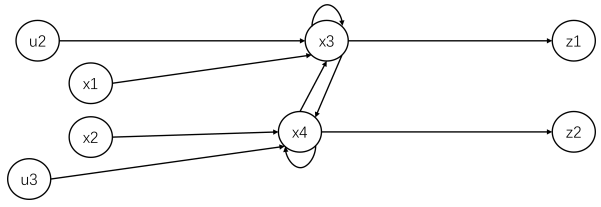

Fig. 4. Graph of $\mathbb{G}(\hat{\Sigma})$ in the case $\mathcal{V} = \mathcal{V}_2$.

VI. CONCLUSION

This paper explored the existence of stealthy integrity attacks for cyber-physical systems with uncertain nonlinear dynamics, when only incomplete information is available to the attacker side, by using tools from structural analysis and systems theory. We provide analytic geometric conditions for the existence of stealthy integrity attacks, by considering suitable controlled-invariant subspaces where state trajectories do not affect the system output. Moreover, we formulate and solve suitable optimization problems that obtain the minimum number of actuation channels that, when protected, guarantee that non-existence of stealthy integrity attacks.

REFERENCES

[1] A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *28th International Conference on Distributed Computing Systems Workshops*. IEEE, 2008, pp. 495–500.

[2] M. Chong, H. Sandberg, and A. Teixeira, "A tutorial introduction to security and privacy for cyber-physical systems," in *The 18th European Control Conference (ECC)*. IEEE, 2019, pp. 968–978.

[3] K. Zhang, C. Keliris, T. Parisini, and M. M. Polycarpou, "Stealthy integrity attacks for a class of nonlinear cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 67, no. 12, pp. 6723–6730, 2021.

[4] A. Teixeira, I. Shames, H. Sandberg, and K. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

[5] A. Teixeira, C. Sou, H. Sandberg, and K. Johansson, "Secure control systems: A quantitative risk management approach," *IEEE Control Systems Magazine*, vol. 35, no. 1, pp. 24–45, 2015.

[6] T. Nguyen, S. Anand, and A. Teixeira, "A zero-sum game framework for optimal sensor placement in uncertain networked control systems under cyber-attacks," in *IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 6126–6133.

[7] K. Zhang, C. Keliris, T. Parisini, and M. Polycarpou, "Identification of sensor replay attacks and physical faults for cyber-physical systems," *IEEE Control Systems Letters*, vol. 6, pp. 1178–1183, 2021.

[8] K. Zhang, C. Keliris, M. Polycarpou, and T. Parisini, "Detecting stealthy integrity attacks in a class of nonlinear cyber-physical systems: A backward-in-time approach," *Automatica*, vol. 141, p. 110262, 2022.

[9] K. Zhang, C. Keliris, T. Parisini, B. Jiang, and M. Polycarpou, "Passive attack detection for a class of stealthy intermittent integrity attacks," *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 4, pp. 1–18, 2023.

[10] K. Zhang, A. Kasis, M. M. Polycarpou, and T. Parisini, "A sensor watermarking design for threat discrimination," *in IFAC 11th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS*, vol. 55, no. 6, pp. 433–438, 2022.

[11] F. Farokhi, I. Shames, and N. Batterham, "Secure and private control using semi-homomorphic encryption," *Control Engineering Practice*, vol. 67, pp. 13–20, 2017.

[12] J. Wang, J.-F. Zhang, and X. He, "Differentially private distributed algorithms for stochastic aggregative games," *Automatica*, vol. 142, p. 110440, 2022.

[13] J.-F. Zhang, J. Tan, and J. WANG, "Privacy security in control systems," *SCIENCE CHINA Information Sciences*, vol. 64, no. 7, p. 176201, 2021.

[14] J. Milošević, A. Teixeira, K. Johansson, and H. Sandberg, "Actuator security indices based on perfect undetectability: Computation, robustness, and sensor placement," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3816–3831, 2020.

[15] S. Gracy, J. Milošević, and H. Sandberg, "Security index based on perfectly undetectable attacks: Graph-theoretic conditions," *Automatica*, vol. 134, p. 109925, 2021.

[16] S. Weerakkody, X. Liu, S. Son, and B. Sinopoli, "A graph-theoretic characterization of perfect attackability for secure design of distributed control systems," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 60–70, 2016.

[17] H. Trentelman and M. Stoorvogel, Antonand Hautus, *Control theory for linear systems*. Springer Science & Business Media, 2012.

[18] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[19] J. Vander-Woude, "The generic number of invariant zeros of a structured linear system," *SIAM Journal on Control and Optimization*, vol. 38, no. 1, pp. 1–21, 1999.

[20] C. Commault, J.-M. Dion, and Y. Agha, "Structural analysis for the sensor location problem in fault detection and isolation," *Automatica*, vol. 44, no. 8, pp. 2074–2080, 2008.