2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Cyber-Attack Detection and Isolation of Nonlinear Cyber-Physical Systems: An Auxiliary Filter Approach

Hamed Kazemi and Khashayar Khorasani

Abstract— **This paper introduces a novel framework designed to bolster security and develop cyber-attack detection and isolation of nonlinear cyber-physical systems (CPS), focusing specifically on discrete-time CPS vulnerable to actuator and sensor cyber-attacks. The framework is particularly applicable in scenarios involving data exchanges among controllers in the command-and-control (C&C) center and the plant. It operates under the assumption that adversaries can inject false data into communication networks. The paper presents design and development of auxiliary filters that are tailored for cyber-attacks detection and isolation, along with a complete stability analysis and demonstration of their efficacy through associated design algorithms. Simulation studies are conducted by using a high-fidelity Unmanned Aerial Vehicle (UAV) model, which highlights the framework's effectiveness in detecting and isolating false data injection (FDI) and covert attacks.**

I. Introduction

Cyber-Physical Systems (CPS) represent critical modern infrastructures characterized by intricate integration of cyber and physical components, facilitating intelligent operation and control of critical subsystems [1]–[3]. This fusion of cyber and physical elements enhances CPS efficiency while simultaneously heightening susceptibility to cyber-attacks, which are actions by adversaries that exploit system vulnerabilities, leading to potential critical damages [1]. Due to their widespread availability and accessibility, CPS are at high risk and likelihood of being compromised through cyber-attacks, which can manifest at both cyber and physical layers [4]. Among the array of cyber-attacks, false data injection (FDI) stands out as serving the foundation for various deceptive cyber-attacks such as replay attacks, covert attacks, and even zero dynamics attacks, as detailed in studies such as [5]–[7].

In the FDI scenarios, an adversary gains access to and alters the physical system's state, sensor data, or control commands by introducing arbitrary errors or false information to them. Recent events highlight numerous instances of cyber-attacks targeting diverse infrastructure, underscoring the severity of the threat they pose and emphasizing the imperative to explore methods for prevention and protection of the critical infrastructure systems [8]–[10]. Prompt detection and isolation of cyber-attacks could mitigate the overall system damages to within acceptable limits.

In real world, numerous industrial processes exhibit nonlinear properties that are inherent to their characteristics and external environments. These complexities amplify the challenges of developing detection and isolation strategies when compared to linear systems. Despite the above, the literature contains only a limited number of works addressing cybersecurity of nonlinear CPS. For instance, [11] analyzed the vulnerability of nonlinear dynamical control systems to stealthy sensor attacks, while [12] developed a methodology for generating stealthy integrity attacks for a class of nonlinear CPS. However, both of these works primarily focus on the stealthiness of cyber-attacks and their generation in nonlinear systems, without addressing the detection and isolation of such attacks.

State estimation plays a pivotal role in cyber-attack detection, where precise system state estimation is crucial for developing various detectors on both the plant and the Command and Control (C&C) sides. Common examples of estimation methods include filtering approaches such as the Extended Kalman Filter (EKF) [13], fuzzy filtering [14], and the Unscented Kalman Filter (UKF) [15]. The Extended Kalman Filter (EKF) is widely used in nonlinear systems for performing state estimation. It is an extension of the traditional Kalman Filter (KF) that linearizes the system model to handle nonlinearities. Alternatively, the UKF is an unscented transform-based method that utilizes a series of formalized and predefined sample data to approximate the probability density function of the underlying processes for state estimation purposes. Fuzzy filtering [14] employs fuzzy logic to address uncertainties and nonlinearities in the system dynamics.

Furthermore, observer-based techniques have gained prominence in cyber-attack detection of nonlinear CPS. These techniques involve designing observers to estimate the system's internal states based on available measurements. They are particularly useful in scenarios where the system dynamics are not fully known or are subject to disturbances. Recent research has explored various observer-based designs tailored for nonlinear CPSs, such as adaptive observers and sliding mode observers. In these estimation methods, the nonlinear part is typically modeled as system disturbance or as a term that should be decoupled [16], [17].

The advantages and contributions of this paper include addressing the challenges posed by discrete-time nonlinear

K. Khorasani would like to acknowledge the financial support received from NATO under the Emerging Security Challenges Division program as well as the support received from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Department of National Defence (DND) under the Discovery Grant and DND Supplemental Programs. This work was also supported in part by funding from the Innovation for Defence Excellence and Security (IDEaS) program from the Department of National Defence (DND). Any opinions and conclusions in this work are strictly those of the authors and do not reflect the views, positions, or policies of - and are not endorsed by - IDEaS, DND, or the Government of Canada.

Hamed Kazemi (h_azemi@encs.concordia.ca) and Khashayar Khorasani (kash@ece.concordia.ca) are with the Department of Electrical and Computer Engineering, Concordia University, Montreal, Quebec, Canada.

979-8-3503-1632-2/24/$31.00 ©2024 IEEE

CPS that are vulnerable to adversarial FDI attacks on both measurement and actuation communication networks. These challenges are mitigated through the deployment of auxiliary filters on both the C&C and plant sides, aimed at detecting and isolating these threats, thereby eliminating the need for state estimation.

Specifically, we introduce auxiliary filters for actuator and sensor cyber-attack detection and isolation that are accompanied by formal stability analysis, to demonstrate their effectiveness. This approach is grounded on a formal design methodology outlined in a proposed algorithm. The efficacy of the proposed methodology is validated through its implementation in a CPS case study subject to FDI and covert attacks, which involves the nonlinear model of a high-fidelity Unmanned Aerial Vehicle (UAV). The problem of mitigation and resilient control recovery are beyond the scope of the current work and are topics of future research.

The remainder of this paper is structured as follows. In Section II, we elaborate on the problem formulation and provide preliminary information. Sections III and IV delve into the methodologies for actuator and sensor cyber-attack detection and isolation, respectively. Finally, we present simulation results of an case study in Section V.

II. Problem formulation and preliminaries

In this section, we present the formulation of the system under consideration, including its mathematical representation. We also outline the framework of the CPS and describe the strategy by which cyber-attacks impact the system. The mathematical model of a nonlinear CPS is expressed as follows:

$$\begin{aligned} x(k+1) &= f(x(k)) + g(x(k))u^*(k) \\ y(k) &= h(x(k)). \end{aligned} \tag{1}$$

Where, k denotes the discrete-time sample, $x(k) \in \mathbb{R}^n$ represents the system state, $u^*(k) \in \mathbb{R}^p$ represents the control input received at the plant side, and $y(k) \in \mathbb{R}^m$ denotes the measured output on the plant side. The functions $f(x(k))$, $g(x(k))$, and $h(x(k))$ denote nonlinearaties, each with the appropriate dimensions. The system states are assumed to be observable.

The control input to the plant and output of the plant at the C&C side, subject to cyber-attacks on the communication networks between the actuators and sensors, are expressed as follows:

$$\begin{aligned} u^a(k) &= u^*(k) = u(k) + B^a a^u(k) \\ y^a(k) &= h(x(k)) + D^a a^y(k), \end{aligned} \tag{2}$$

where $u(k)$ represents the control signal or reference signal generated on the C&C side, while $u^a(k) \in \mathbb{R}^p$ and $y^a(k) \in \mathbb{R}^m$ denote the control input received at the plant side and the output received at the C&C side, respectively. The signals $a^u(k) \in \mathbb{R}^{p_a}$ and $a^y(k) \in \mathbb{R}^{m_a}$ represent the FDI attack signals, and the matrices B^a and D^a describe the actuator and sensor cyber-attack signatures, respectively. It is important to note that the signals $a^u(k)$ and $a^y(k)$ can be arbitrarily manipulated by a malicious adversary.

The presence of the network layer allows for the observation of the CPS from both the plant and the C&C sides. The architecture of the CPS, along with the concept of observing it from both sides, is visually represented in Figure 1. In this figure, one can identify the auxiliary filters, as well as cyber-attack detectors on both the C&C and the plant sides. Furthermore, there are communication links that are depicted in the diagram, which are susceptible to cyber-attacks. Furthermore, a separate communication channel connects the plant side detector to the C&C, transmitting binary results from the detector that is used to inform C&C from cyber-attack that has been detected at the plant side, as depicted in Figure 1.

In the subsequent sections, we will provide specifics of these modules and their functionalities.

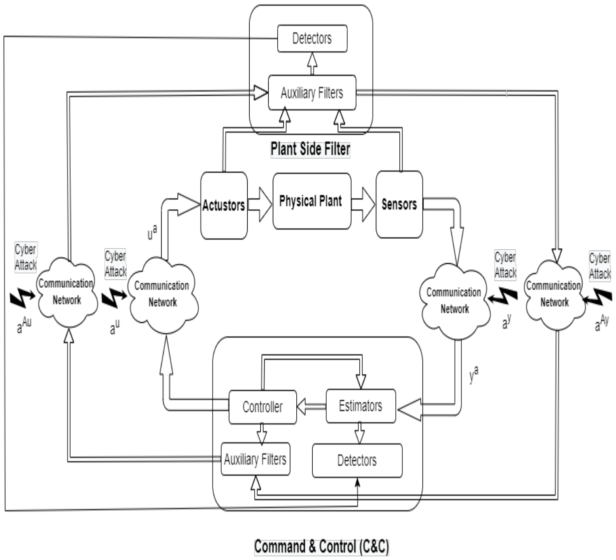

Fig. 1: The Cyber-Physical System (CPS) under actuator and sensor cyber-attacks with the C&C and the plant side filters and detectors.

Remark 1: For simplicity in analysis, we did not consider the presence of noise in the CPS dynamics. However, in the simulation results provided in Section V, we have indeed added noise and disturbances to the control and measurement channels for a more realistic simulation scenario.

III. Actuator Cyber-Attack Detection and Isolation

In this section, we outline the presence of two auxiliary filters, one on the C&C and the second on the plant sides. These filters serve the purpose of detecting actuator network cyber-attacks. To facilitate this, an extra communication link is introduced, allowing the C&C auxiliary filter output to be transmitted to the plant side for the purpose of comparison. However, it should be noted that this supplementary communication link is also susceptible to potential cyber-attacks. Importantly, the role of the auxiliary filters is not to estimate the system's states. Instead, their primary function is detection of actuator cyber-attacks. A similar structure will be proposed in the next section for detecting sensor network cyber-attacks.

Given the nonlinear dynamics of the plant (1), The formulation of the nonlinear actuator auxiliary filter on the C&C side is presented as follows:

$$\begin{aligned} z^{CC}_{Au}(k+1) &= A_{Au} z^{CC}_{Au}(k) + F_{Au}(z^{CC}_{Au}(k)) + T_{Au}u(k) \\ &+ K_{Au}y^{a}(k), \end{aligned} \tag{3}$$

where $z^{CC}_{Au}(k) \in \mathbb{R}^n$ represents the state of the actuator auxiliary filter at the C&C side. The nonlinear term $F_{Au}(.)$ and matrices A_{Au}, T_{Au}, and K_{Au} are of appropriate dimensions and will be designed and selected subsequently.

The proposed actuator auxiliary filters on the plant side are expressed in the following form:

$$\begin{aligned} z^{p}_{Au}(k+1) &= A_{Au} z^{p}_{Au}(k) + F_{Au}(z^{p}_{Au}(k)) + T_{Au}u^{a}(k) + K_{Au}y(k) \\ &+ L^{p}_{Au}\big(z^{p}_{Au}(k) - (z^{CC}_{Au}(k) + D^{a}_{Au}a^{Au})\big). \end{aligned} \tag{4}$$

In the above equation, $z^{p}_{Au}(k) \in \mathbb{R}^n$ denotes the state of the actuator auxiliary filter at the plant side, whereas $a^{Au} \in \mathbb{R}^{n_{a^{Au}}}$ denotes the cyber-attack on the communication link between the two actuator filters with the signature D^{a}_{Au}. The error signals between the estimated states on both sides can be defined as $e^{pc}_{Au}(k) = z^{p}_{Au} - z^{CC}_{Au}$. The gain of the filter, L^{p}_{Au}, will be designed subsequently.

The residual of actuator cyber-attacks is defined as follows,

$$res_{a^{u}(k)} = L^{p}_{Au}(z^{p}_{Au} - {}^{a}z^{CC}_{Au}). \tag{5}$$

where ${}^{a}z^{CC}_{Au}$ indicates the C&C-side auxiliary filter's output that is received at the plant side, which could also be subject to cyber-attacks.

Assumption 1: The malicious adversary does not have access to all the communication channels between the two side filters, i.e., $rank(D^{a}_{Au}) < n$.

Assumption 2: For $F_{Au}(x)$, there exists a positive constant γ for all x_1 and x_2 such that

$$||F_{Au}(x_1) - F_{Au}(x_2)|| \leq \gamma ||x_1 - x_2||. \tag{6}$$

Theorem 1: Under Assumptions 1 and 2 and if the following conditions hold, the residual $res_{a^{u}(k)}$ is affected by the actuator cyber-attacks $a^{u}(k)$, and is decoupled from other cyber-attacks and inputs, i.e. for $a^{u}(k) = 0$, e^{pc}_{Au} converges to zero asymptotically, provided that

1) $K_{Au}D^{a} = 0$.
2) $L^{p}_{Au}D^{a}_{Au} = 0$.
3) $T_{Au}B^{a} \neq 0$.
4) Given that the matrix L^{p}_{Au} satisfies the condition 2), we define the matrix A according to $A = A_{Au} + L^{p}_{Au}$, such that A is Hurwitz by properly selecting the matrix A_{Au} that is at our disposal.

Furthermore, if B^{a} is diagonal, each individual $res_{a^{u}_{i}(k)}$ is affected by the corresponding cyber-attack $a^{u}_{i}(k)$ and is decoupled from the others.

Proof: By subtracting (3) from (4), one can derive the state-space representation of the error dynamics between the two filter states as follows:

$$\begin{aligned} e^{pc}_{Au}(k+1) &= Ae^{pc}_{Au}(k) + F_{Au}(z^{p}_{Au}(k)) \\ &- F_{Au}(z^{cc}_{Au}(k)) + T_{Au}B^{a}a^{u}(k) \\ &- K_{Au}D^{a}a^{y}(k) - L^{p}_{Au}D^{a}_{Au}a^{Au}(k). \end{aligned} \tag{7}$$

Under conditions 1) and 2) above holding, (7) can be simplified to:

$$\begin{aligned} e^{pc}_{Au}(k+1) &= Ae^{pc}_{Au}(k) + F_{Au}(z^{p}_{Au}(k)) \\ &- F_{Au}(z^{cc}_{Au}(k)) + T_{Au}B^{a}a^{u}(k). \end{aligned} \tag{8}$$

To demonstrate the asymptotic convergence of $e^{pc}_{Au}(k)$, let us first assume that $a^{u}(k) = 0$. We choose a Lyapunov function candidate $V(k) = e^{pc}_{Au}k)^{T}Pe^{pc}_{Au}(k)$, where the matrix P is symmetric positive definite (PD) that satisfies

$$A^{T}PA + (\gamma^{2} - 1)P + \gamma PAA^{T}P + \gamma I < 0. \tag{9}$$

It follows from (8) and (6) that:

$$\begin{aligned} & V(k+1) - V(k) = \big(Ae^{pc}_{Au}(k) + F_{Au}(z^{p}_{Au}(k)) - F_{Au}(z^{cc}_{Au}(k))\big)^{T} \\ & P\big(Ae^{pc}_{Au}(k) + F_{Au}(z^{p}_{Au}(k)) - F_{Au}(z^{cc}_{Au}(k))\big) - e^{pc}_{Au}(k)^{T}Pe^{pc}_{Au}(k) \\ =\ & e^{pc}_{Au}(k)^{T}A^{T}PAe^{pc}_{Au}(k) + 2\big(Ae^{pc}_{Au}(k)\big)^{T}P\big(F_{Au}(z^{p}_{Au}(k)) \\ & -F_{Au}(z^{cc}_{Au}(k))\big) + \big(F_{Au}(z^{p}_{Au}(k)) - F_{Au}(z^{cc}_{Au}(k))\big)^{T}P \\ & \big(F_{Au}(z^{p}_{Au}(k)) - F_{Au}(z^{cc}_{Au}(k))\big) - e^{pc}_{Au}(k)^{T}Pe^{pc}_{Au}(k) \\ \leq\ & e^{pc}_{Au}(k)^{T}\big(A^{T}PA - P\big)e^{pc}_{Au}(k) + 2\gamma||e^{pc}_{Au}(k)^{T}A^{T}P||\ ||e^{pc}_{Au}(k)|| \\ & +\gamma^{2}e^{pc}_{Au}(k)^{T}Pe^{pc}_{Au}(k) \\ \leq\ & e^{pc}_{Au}(k)^{T}\big(A^{T}PA - P\big)e^{pc}_{Au}(k) + \gamma\big(||e^{pc}_{Au}(k)^{T}A^{T}P||^{2} \\ & +||e^{pc}_{Au}(k)||^{2}\big) + \gamma^{2}e^{pc}_{Au}(k)^{T}Pe^{pc}_{Au}(k) \\ =\ & e^{pc}_{Au}(k)^{T}\big(A^{T}PA + (\gamma^{2} - 1)P + \gamma PAA^{T}P + \gamma I\big)e^{pc}_{Au}(k). \end{aligned} \tag{10}$$

Note that regarding (9), $A^{T}PA + (\gamma^{2} - 1)P + \gamma PAA^{T}P + \gamma I < 0$ in (10) implies $e^{pc}_{Au}(k)$ tends to zero asymptotically for any initial condition $e^{pc}_{Au}(0)$, when $a^{u}(k) = 0$.

On the other hand, it follows from (8) that under the condition 3), $e^{pc}_{Au}(k)$ is affected by the actuator cyber-attack $a^{u}(k)$ for any $a^{u}(k) \neq 0$ and the impact of $e^{pc}_{Au}(k)$ will appear in the residual $res_{a^{u}(k)}$. Furthermore, if B^{a} is diagonal, i.e., each individual cyber-attack only affects a communication channel in the actuator communication network, according to Equation (8), only the corresponding element of the vector $e^{pc}_{Au}(k)$ is affected, while it is decoupled from other channels' cyber-attacks. This completes the proof of the theorem. ■.

In view of Theorem 1, we present below an algorithm for designing auxiliary filters. These filters are crucial for detecting actuator cyber-attacks, as discussed in this section. The algorithm provides a systematic approach to configuring these filters.

Algorithm 1 Actuator Auxiliary Filter Design Algorithm.

```
Inputs: Assumption 2 and conditions of Theorem 1
Output: $A$ and $A_{Au}$
1. Compute $K_{Au}$ according to condition 1). Setting $K_{Au} = 0$ is a valid selection.
2. Compute the matrix $L^{p}_{Au}$ to meet condition 2). It is essential to note that $L^{p}_{Au}$ cannot be zero due to the definition of $res_{a^u(k)}$.
3. Select a function $F_{Au}$ that satisfies the Assumption 2 with a desired positive value of $\gamma$ .
4. Choose $T_{Au}$ such that it does not nullify the potential cyber-attacks effect, i.e., $T_{Au}B^a \neq 0$.
5. Given the matrix $L^{p}_{Au}$ from the Step 2, then by properly selecting the matrix $A_{Au}$, one would select a new matrix $A$ as $A = A_{Au} + L^{p}_{Au}$, that is Hurwitz and satisfies $A^TPA + (\gamma^2 - 1)P + \gamma PAA^TP + \gamma I < 0$ for a PD symmetric matrix $P$ and the desired value of $\gamma$.
```

IV. Sensor Cyber-Attack Detection and Isolation

For the sensor cyber-attack detection and isolation methodology, and given the nonlinear dynamics of the plant (1) we first design the plant-side nonlinear auxiliary filter as follows:

$$\begin{aligned} z^{p}_{Ay}(k+1) &= A_{Ay}z^{p}_{Ay}(k) + F_{Ay}(z^{p}_{Ay}(k)) + T_{Ay}u^a(k) \\ &+ K_{Ay}y(k), \end{aligned} \tag{11}$$

where $z^{p}_{Ay}(k) \in \mathbb{R}^n$ represents the state of the sensor auxiliary filter at the plant side. The nonlinear term $F_{Ay}(.)$ and matrices A_{Ay}, T_{Ay}, and K_{Ay} are of appropriate dimensions and need to be designed subsequently.

On the C&C side, control inputs are obtained before the injection of actuator cyber-attacks. However, the measurement signal received at the C&C could be vulnerable to spositive densor cyber-attacks. Consequently, the proposed sensor auxiliary filters on the C&C side are formulated as follows:

$$\begin{aligned} z^{CC}_{Ay}(k+1) &= A_{Ay}z^{CC}_{Ay}(k) + F_{Ay}(z^{CC}_{Ay}(k)) + T_{Ay}y^a(k) \\ &+ K_{Ay}u(k) + L^{CC}_{Ay}\left(z^{CC}_{Ay}(k) - (z^{p}_{Ay}(k) + D^{a}_{Ay}y^{Ay})\right). \end{aligned} \tag{12}$$

where $z^{CC}_{Ay}(k) \in \mathbb{R}^n$ represents the state of the sensor auxiliary filter on the C&C side, and $a^{Ay} \in \mathbb{R}^{n_{a^{Ay}}}$ illustrates the cyber-attack occurring on the communication link between the two sensor filters, characterized by the signature D^{a}_{Ay}.

The error signals between the estimated states on both sides can be expressed as $e^{cp}_{Ay}(k) = z^{CC}_{Ay} - z^{p}_{Ay}$. Additionally, the residual of sensor cyber-attacks is defined as

$$res_{a^y(k)} = L^{CC}_{Ay}(z^{CC}_{Ay} - {}^{a}z^{p}_{Ay}) \tag{13}$$

Where ${}^{a}z^{p}_{Ay}$ indicates the output of the plant-side sensor auxiliary filter received at the C&C side, which could also be vulnerable to cyber-attacks.

***Theorem** 2:* Under Assumptions similar to 1 and 2 for D^{a}_{Ay} and F_{Ay} and provided that the following conditions hold, the residual $res_{a^y(k)}$ is affected by the sensor cyber-attack $a^y(k)$, and is decoupled from other cyber-attacks and inputs, i.e. for $a^y(k) = 0$, e^{cp}_{Ay} converges to zero asymptotically, provided that

1) $K_{Ay}B^a = 0$.
2) $L^{CC}_{Ay}D^{a}_{Ay} = 0$.
3) $T_{Ay}D^a \neq 0$.
4) Given that the matrix L^{CC}_{Ay} satisfies the condition 2), we define the matrix A according to $A = A_{Ay} + L^{CC}_{Ay}$, such that A is Hurwitz by properly selecting the matrix A_{Ay} that is at our disposal.

Furthermore, if D^a is diagonal, each individual $res_{a^y_i(k)}$ is affected by the corresponding cyber-attack $a^y_i(k)$ and is decoupled from the others.

Proof: The proof is similar to that of Theorem 1 and is omitted for brevity. .

Furthermore, the parameters of the filters (11) and (12) can be determined by using an algorithm similar to the Algorithm 1.

In terms of isolating different cyber-attacks, as mentioned in the above theorems, actuator and sensor residuals are detected separately and are decoupled from each other. Furthermore, if there are various individual cyber-attacks, each affecting the corresponding channel in the actuator/sensor, then only the corresponding element of $res_{a^u(k)}/res_{a^y(k)}$ is affected. Therefore, by utilizing residuals (5) and (13), one can detect and isolate different actuator and sensor cyber-attacks.

***Remark** 2:* In Theorems 1 and 2, no specific assumptions are made regarding the nature, characteristics, and type of cyber-attacks. This flexibility implies that the proposed method is well-equipped to detect and isolate various types of deception attacks that are rooted in the FDI, provided the corresponding assumptions are met. These may include readily detectable attacks such as replay attacks, as well as more subtle forms of cyber-attacks such as covert attacks and zero dynamics attacks that might lead to stealthy cyber-attack [18]. Typically, these cyber-attacks involve a component injected into the actuator communication network, where the effects caused by it are either unobservable or are compensated for at the outputs. Referring to (4), these components can be modeled as $a^u(k)$, which was shown to be detectable by the proposed actuator detector. This matter will be illustrated in the next section, focusing on the detection and isolation of different FDI and covert attacks in nonlinear CPS. Therefore, the proposed approach provides a versatile solution to enhance system security against various types of deceptive cyber-attacks.

***Remark** 3:* While existing literature has made significant strides in addressing the stealthiness and generation of cyber-attacks in nonlinear CPS, as reported in works like [11] and [12], these efforts have largely overlooked the critical aspects of detecting and isolating such attacks. Moreover, state estimation methods, including the EKF [13], fuzzy filtering [14], and UKF [15], rely heavily on accurate system models, which may not always be feasible in complex, nonlinear environments. In contrast, this paper advances the field by introducing auxiliary filters on both the C&C and plant sides, which effectively detect and isolate threats without relying on state estimation. This method is further supported by formal stability analysis, ensuring robustness and reliability. Additionally, the flexibility of the proposed approach, as highlighted in the Theorems 1 and 2, allows it to detect and isolate a wide range of cyber-attacks without specific assumptions about their nature or characteristics, as noted in Remark 2. The proposed methodology has been rigorously validated using a high-fidelity UAV nonlinear model, demonstrating its practical applicability and effectiveness in real-world scenarios, as will be elaborated in the next section.

V. Simulation

In this section, we present a case study involving a mini UAV, specifically the commercial off-the-shelf Radio-controlled drone known as the Ultra Stick 25E [19]. This UAV, as shown in Figure 2, was developed by a group of researchers from the University of Minnesota. The UAV features a conventional design with horizontal and vertical tail surfaces, and is equipped with three primary control surfaces: the elevator, rudder, and ailerons.

All the above control surfaces are actuated by using Hitec servos. The UAV is further outfitted with a comprehensive set of avionics to support research and development in flight control, and is equipped with a range of sensors, including those for measuring angular positions, rates, and linear accelerations. In this paper, the controller's design is accomplished through utilization of backstepping method, as proposed in [20], involving the control of the pitch, roll, and yaw angles by using a total of six Lyapunov functions.

Taking a moment around the aerodynamics center of the UAV, we can express the angular rate dynamics as follows:

$$\begin{aligned}
\dot{p}-\frac{I_{xz}}{I_{xx}}\dot{r} &= \frac{p_d S b}{I_{xx}}c_l-\frac{I_{zz}-I_{yy}}{I_{xx}}qr+\frac{I_{xz}}{I_{xx}}qp \qquad (14)\\
\dot{q} &= \frac{p_d S\bar{c}}{I_{yy}}c_m-\frac{I_{xx}-I_{zz}}{I_{yy}}pr-\frac{I_{xz}}{I_{yy}}(p^2-r^2)+\frac{I_p}{I_{yy}}w_p r\\
\dot{r}-\frac{I_{xz}}{I_{zz}}\dot{p} &= \frac{P_d S b}{I_{zz}}c_n-\frac{I_{yy}-I_{xx}}{I_{zz}}pq-\frac{I_{xz}}{I_{zz}}qr-\frac{I_p}{I_{zz}}w_p q.
\end{aligned}$$

In the above equations, p (rad/s), q (rad/s), and r (rad/s) represent the body axis angular rates, while ϕ (rad), θ (rad), and ψ (rad) denote the body attitude angles. Additionally, P_d (Pa) denotes the dynamic pressure, and α (rad) and β (rad) correspond to the angle of attack and sideslip angle of the UAV in the wind axis, w_p (rad/s) represents the propeller rotation speed, I_{xx}, I_{yy}, I_{zz}, I_{xz}, and I_p denote the moment of inertia coefficients for the UAV and propulsion system, expressed in kgm^2. Additionally, c_l, c_m, and c_n denote the non-dimensional moment coefficients [19].

The kinematics of the rotational motion, which links the body angular rates p, q, and r, the Euler angles ϕ, θ, and ψ, and the aerodynamic angles α, β, and γ (the slope of the flight trajectory), can be expressed as follows:

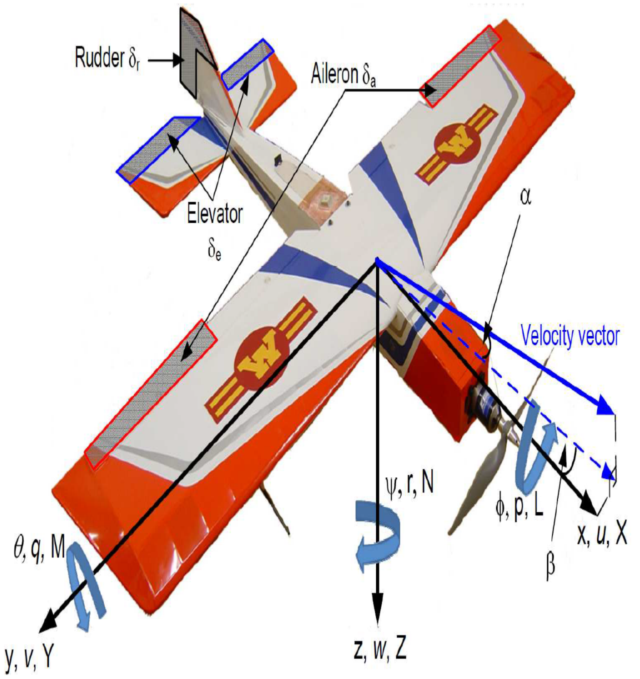

Fig. 2: Forces and moments in the UAV body axis [19]

$$\begin{aligned}
\dot{\phi} &= p+tan\theta(qsin\phi+rcos\phi)\\
\dot{\theta} &= qcos\phi-rsin\phi\\
\dot{\psi} &= \frac{qsin\phi+rcos\phi}{cos\theta}\\
\theta &= \gamma+\alpha cos\phi+\beta sin\phi \qquad (15)
\end{aligned}$$

Moreover, as mentioned above, a backstepping controller has been integrated into the plant side to stabilize the roll, pitch, and yaw angles. The desired values for these angles are denoted as $\bar{\phi}$, $\bar{\theta}$, and $\bar{\psi}$, while $\bar{p}$, $\bar{q}$, and $\bar{r}$ are determined based on the system parameters. The corresponding errors are denoted as e_x, with x representing each of the system states. The backstepping controller leads to computation of the desired aileron, elevator, and rudder deflections, denoted by δa, δe, and δr [20]. The designed control system has been implemented in the Matlab/Simulink environment, and its effectiveness has been evaluated through a series of numerical simulations.

Within the framework of the CPS, as depicted in Figure 1, we assume that the UAV functions as the plant. Meanwhile, the C&C center receives the system's outputs and determines reference values for the corresponding actuators on the UAV. First, we assume the UAV is in the healthy condition. Figure 3 depicts the reference value, the actual values, and the received signals at the C&C under this condition. Consequently, the communication networks for sensors and actuators, as well as the auxiliary filters, are assumed to be compromised by different FDI and covert attacks.

We consider two FDI attacks in the actuator network, affecting $\bar{\theta}$ and $\bar{\psi}$, and two FDI attacks in the actuator

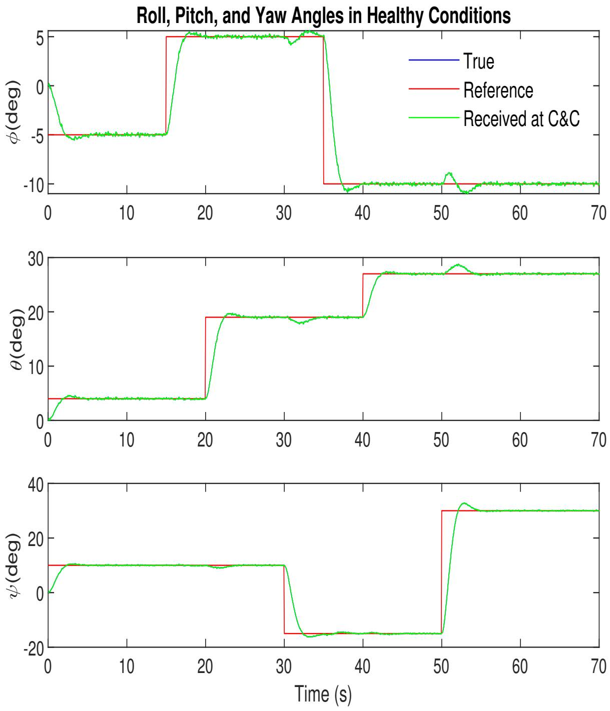

Fig. 3: The UAV's measurement outputs, reference values, and the received signals at the C&C center in healthy condition.

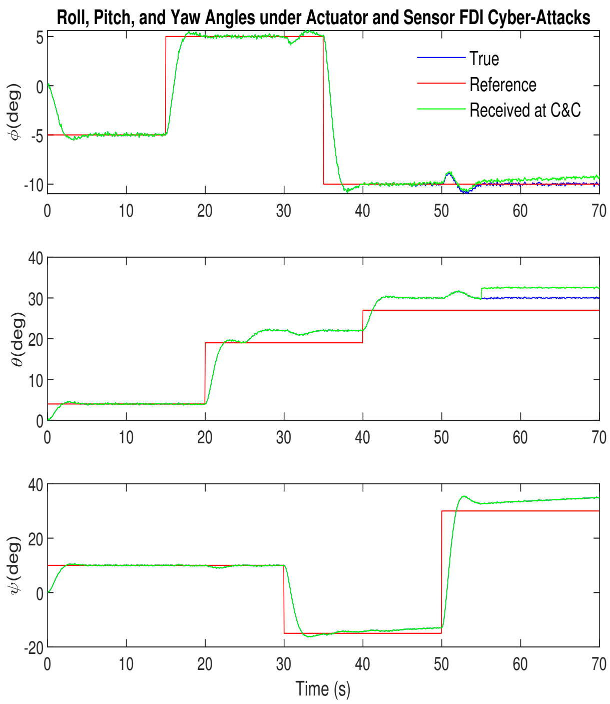

Fig. 4: The UAV's measurement outputs, reference values, and the received signals at the C&C center under actuator and sensor FDI attacks.

networks, corresponding to ϕ and θ. Additionally, two cover attacks affect $\bar{\theta}$ & $\bar{\psi}$ (at the actuator network) and θ & ψ (at the actuator network) simultaneously, such that sensor cyber attacks compensate for the effect of actuator attacks on the signal received at C&C. In summary, this work addresses the detection and isolation of six cyber-attacks, all of which are tackled by using our proposed methodology.

Even though according to the proposed theorems, residuals converge to zero in the absence of cyber-attacks, we still need to establish thresholds for the detection procedure due to the inclusion of noise in the simulation. These thresholds are determined by running the system multiple times under different conditions, ensuring that each residual remains lower than the corresponding threshold in the absence of cyber-attacks and is exceeded in the presence of the corresponding cyber-attack.

Initially, the CPS operates under normal, healthy conditions and adheres to reference values issued by the C&C center, as shown in the first 25 seconds of Figure 4. Subsequently, actuator FDI attacks corresponding to $\bar{\theta}$ and $\bar{\psi}$ occur at times 25 seconds and 35 seconds, respectively. Additionally, sensor cyber-attacks ϕ and θ are assumed to occur at times 45 seconds and 55 seconds, respectively. The impact of these cyber-attacks on system outputs is depicted in Figure 4. The models of actuator and sensor cyber-attacks, including auxiliary filters cyber-attacks, are represented in the first two subfigures of Figures 5 and 6, respectively. Actuator cyber-attacks can cause the system states (true values) to diverge from the reference values issued by the C&C center, while sensor cyber-attacks have the potential to mislead the monitoring system in the C&C by altering the received signals. By leveraging the proposed cyber-attack detection methodology, we are able to recognize the cause of the divergence and the alteration.

To detect and isolate actuator cyber-attacks, the auxiliary filters (3) and (4) are employed. The designed parameters of these auxiliary filters are determined by following the Algorithm 1, as follows:

$$\begin{aligned}
A_{Au} &= [-2\,0\,0; 0\,\,-0.5\,0; 0\,0\,\,-0.5], F_{Au}(x) = sin(x)\mathbf{I}_3 \\
T_{Au} &= \mathbf{I}_3,\ L^p_{Au} = [1.5\,0\,0; 0\,0\,0; 0\,0\,0] \\
K_{Au} &= 0,\ D^a_{Au} = B^a = [0\,0\,0; 0\,1\,0; 0\,0\,1] \\
B^a &= [0\,0\,0; 0\,1\,0; 0\,0\,1].
\end{aligned}$$

Similar auxiliary filters are designed for sensor cyber-attacks. Figure 5 illustrates the FDI attacks occurring on the communication networks of $\bar{\theta}$ and $\bar{\psi}$, along with the auxiliary filters and the residuals of actuator cyber-attacks. Moreover, Figure 6 presents the FDI attacks of the communication networks of ϕ and θ, along with the auxiliary filters and the residuals of sensor cyber-attacks. Each residual is defined particularly for each network channel, and they are only sensitive to the corresponding cyber-attack and are decoupled from others. Consequently, isolating among different cyber-attacks is possible; this can be illustrated in

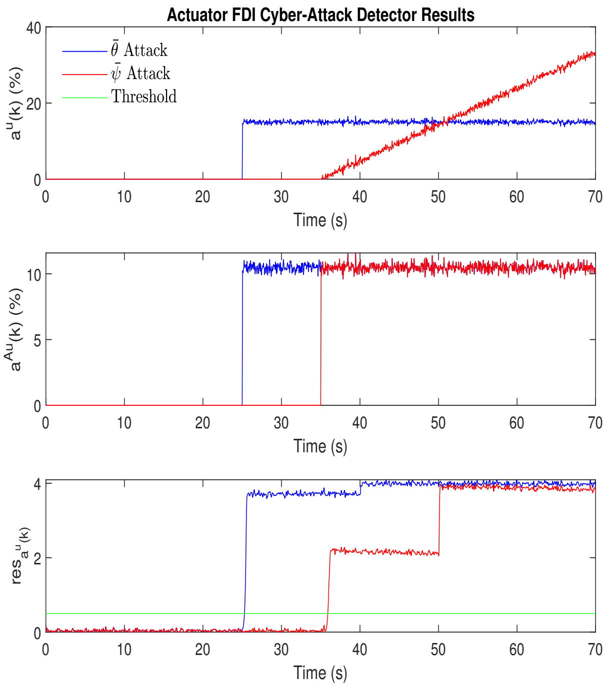

Fig. 5: Actuator cyber-attack detection results, including actuator and auxiliary filter cyber-attacks, and residuals.

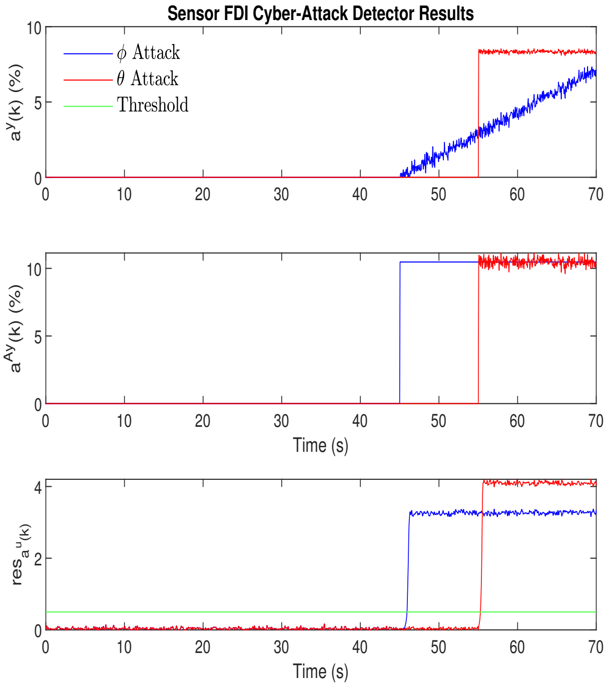

Fig. 6: Sensor cyber-attack detection results, including sensor and auxiliary filter cyber-attacks and residuals.

the last subfigures of Figures 5 and 6.

In another scenario, we consider two cover attacks occuring on the CSP communication networks. The first one impacts $\bar{\theta}$ and θ simultaneously at time 25 seconds, and the second addresses $\bar{\psi}$ and ψ simultaneously at time 45 seconds. Figure 7 illustrates the impact of these cyber-attacks on the system output. Given that attackers have system knowledge, they aim to generate stealthy cyber-attacks by compensating for the effect of actuator cyber-attacks on the sensor communication network.

The covert attacks signals are represented in the first two subfigures of Figure 8. Consequently, as shown in Figure 7, the signals received at C&C follow the issued references. However, the actual values for θ and ψ alter significantly on the plant side, contrary to the C&C's desire. Due to our proposed methodology, the actuator auxiliary filter can detect and isolate these covert attacks, as depicted in the last subfigure of Figure 8. These findings demonstrate that each proposed detector effectively diagnoses the corresponding cyber-attacks, and other anomalies do not affect their detection performance. Consequently, all cyber-attacks are successfully detected and isolated.

VI. CONCLUSIONS

In conclusion, this paper presents a novel framework tailored to bolster security in discrete-time nonlinear cyber-physical systems (CPS) by addressing vulnerabilities to FDI actuator and sensor attacks. Through the deployment of auxiliary filters on both the Command and Control (C&C) and plant sides, the framework effectively detects and isolates FDI and deceptive attacks such as covert attacks, which operate based on FDI attacks, without necessitating state estimation. The simulation studies conducted by using a high-fidelity Unmanned Aerial Vehicle (UAV) model demonstrate the framework's efficacy in detecting and isolating actuator and sensor cyber-attacks. These findings underscore the framework's potential to significantly enhance CPS security and resilience against cyber threats, representing a valuable contribution to the field of CPS security.

REFERENCES

[1] S. Tan, J. M. Guerrero, P. Xie, R. Han, and J. C. Vasquez, "Brief survey on attack detection methods for cyber-physical systems," *IEEE Systems Journal*, vol. 14, no. 4, pp. 5329–5339, 2020.

[2] W. Lucia, G. Franzè, and B. Sinopoli, "A supervisor-based control architecture for constrained cyber-physical systems subject to network attacks," *IEEE Transactions on Control of Network Systems*, 2022.

[3] H. Chen, "Applications of cyber-physical system: a literature review," *Journal of Industrial Integration and Management*, vol. 2, no. 03, p. 1750012, 2017.

[4] H. Fawzi, P. Tabuada, and S. Diggavi, "Security for control systems under sensor and actuator attacks," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pp. 3412–3417, IEEE, 2012.

[5] M. Kordestani and M. Saif, "Observer-based attack detection and mitigation for cyberphysical systems: A review," *IEEE Systems, Man, and Cybernetics Magazine*, vol. 7, no. 2, pp. 35–60, 2021.

[6] P. Zhu, S. Jin, X. Bu, Z. Hou, and C. Yin, "Model-free adaptive control for a class of mimo nonlinear cyberphysical systems under false data injection attacks," *IEEE Transactions on Control of Network Systems*, vol. 10, no. 1, pp. 467–478, 2022.

[7] M. Taheri, K. Khorasani, I. Shames, and N. Meskin, "Cyberattack and machine-induced fault detection and isolation methodologies for cyber-physical systems," *IEEE Transactions on Control Systems Technology*, 2023.

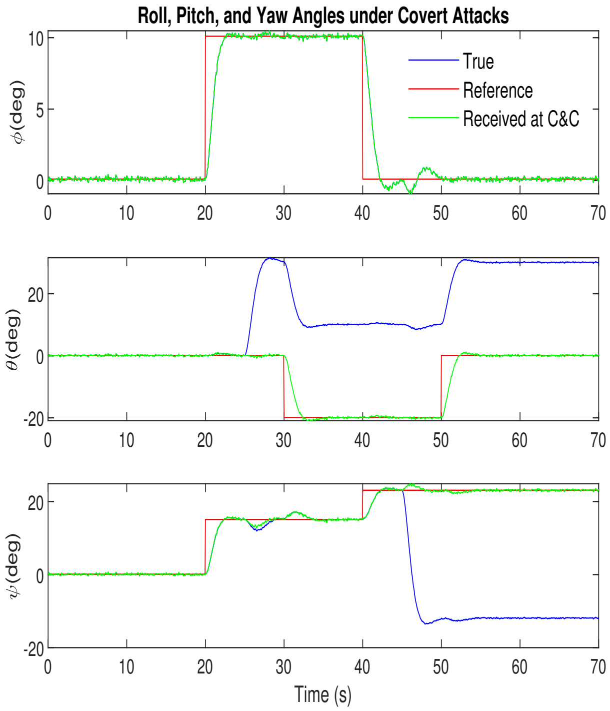

Fig. 7: The UAV's measurement outputs, reference values, and the received signals at the C&C center under covert attacks.

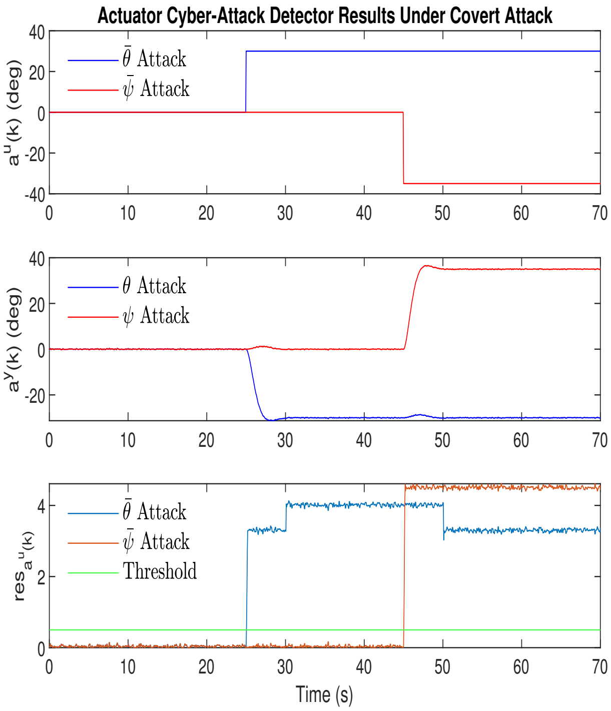

Fig. 8: Covert attacks signals and the detection results.

[8] H. S. Sánchez, D. Rotondo, T. Escobet, V. Puig, and J. Quevedo, "Bibliographical review on cyber attacks from a control oriented perspective," *Annual Reviews in Control*, vol. 48, pp. 103–128, 2019.

[9] A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, S. Sastry, *et al.*, "Challenges for securing cyber physical systems," in *Workshop on future directions in cyber-physical systems security*, vol. 5, Citeseer, 2009.

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

[11] A. Khazraei and M. Pajic, "Resiliency of nonlinear control systems to stealthy sensor attacks," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 7109–7114, IEEE, 2022.

[12] 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.

[13] K. Manandhar, X. Cao, F. Hu, and Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using kalman filter," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 4, pp. 370–379, 2014.

[14] M. Rahimifard, A. M. M. Sizkouhi, and R. R. Selmic, "Cyberattack detection for a class of nonlinear multiagent systems using set-membership fuzzy filtering," *IEEE Systems Journal*, 2024.

[15] R. M. Asl, Y. S. Hagh, S. Simani, and H. Handroos, "Adaptive square-root unscented kalman filter: An experimental study of hydraulic actuator state estimation," *Mechanical Systems and Signal Processing*, vol. 132, pp. 670–691, 2019.

[16] Y. Yu and Y. Yuan, "Event-triggered active disturbance rejection control for nonlinear network control systems subject to dos and physical attacks," *ISA Transactions*, vol. 104, pp. 73–83, 2020.

[17] Z. Cao, Y. Niu, and J. Song, "Finite-time sliding-mode control of markovian jump cyber-physical systems against randomly occurring injection attacks," *IEEE Transactions on Automatic Control*, vol. 65, no. 3, pp. 1264–1271, 2019.

[18] Z. Zhao, Y. Yang, Y. Li, and R. Liu, "Security analysis for cyber-physical systems under undetectable attacks: A geometric approach," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 11, pp. 4359–4370, 2020.

[19] Y. C. Paw, *Synthesis and validation of flight control for UAV*. University of Minnesota, 2009.

[20] M. Lungu and R. Lungu, "Adaptive backstepping flight control for a mini-uav," *International Journal of Adaptive Control and Signal Processing*, vol. 27, no. 8, pp. 635–650, 2013.