Active Learning-Based Control for Resiliency of Uncertain Systems Under DoS Attacks

Sayan Chakraborty, *Graduate Student Member, IEEE*, Weinan Gao, *Senior Member, IEEE*, Kyriakos G. Vamvoudakis, *Senior Member, IEEE*, and Zhong-Ping Jiang, *Fellow, IEEE*

***Abstract*—In this letter, we present an active learning-based control method for discrete-time linear systems with unknown parameters under denial-of-service (DoS) attacks. For any DoS duration parameter, using switching systems theory and adaptive dynamic programming, an active learning-based control technique is developed. A critical DoS average dwell-time is learned from online input-state data, guaranteeing stability of the equilibrium point of the closed-loop system in the presence of DoS attacks with average dwell-time greater than or equal to the critical DoS average dwell-time. The effectiveness of the proposed methodology is illustrated via a numerical example.**

***Index Terms*—Learning-based control, resiliency, DoS attacks, output regulation.**

I. Introduction

CYBER-PHYSICAL systems (CPSs), which integrate network communication, control, and computing, have gained attention due to their benefits in energy efficiency, performance, and convenience. Unlike traditional networked control systems, CPSs - with both physical and network components - are more vulnerable to malicious network attacks [1], [2]. These attacks, which disrupt system functionality, can be classified as deception attacks (compromising data integrity) and denial-of-service (DoS) attacks (disrupting the availability of services or information exchange) [3], [4]. DoS attacks pose a serious threat due to their simplicity and the potential to cause communication breakdowns [5]. Thus, ensuring not only stability but also resilience against DoS attacks is critical. Researchers have studied resilient control strategies, with work such as [6] identifying attack frequency and duration to maintain system stability, alongside other studies [7], [8], [9] and references therein. However, these strategies often assume precise knowledge of system dynamics.

Given the growing complexity of modern systems, accurate modeling has become a challenge. In the era of big data, data-driven control methods have emerged as a promising alternative. Reinforcement learning (RL) and adaptive dynamic programming (ADP) have been used for optimal adaptive control in stabilization [10] and output regulation [11]. However, most ADP studies assume ideal communication channels, ignoring the risks of cyber-attacks. Recently, the authors in [12] and [13] proposed resilient RL methods for continuous-time systems under DoS attacks. Methods adopting data-driven predictive control [14], adaptive control [15], and ADP [16] have been proposed, but most of these methodologies cannot adapt to changes during the duration of DoS attacks. In contrast, this letter addresses the design problem, proposing an active learning-based controller methodology for discrete-time systems that adapts to varying DoS attacks.

In this letter, we tackle the learning-based optimal output regulation problem for linear discrete-time systems with unknown parameters under DoS attacks. Like in [13], the system is modeled as a switched system, alternating between stable closed-loop and unstable open-loop modes depending on DoS presence. However, a new condition on the DoS average dwell-time for discrete-time systems is proposed in this letter. We learn the divergence parameter λ_+ directly from input-state data without knowing the system parameters. λ_+ quantifies the divergence of the system during DoS attacks, while a convergence parameter λ_- is selected based on λ_+ and the DoS duration parameter T. A formula for choosing λ_- has been proposed in this letter, which differs from [13]. This convergence parameter helps to derive an optimal output regulation controller, ensuring closed-loop stability and resilience against DoS attacks. We use the policy iteration (PI) algorithm to learn the optimal controller. Unlike [13], the data matrix in PI may lack full column rank, which may lead to algorithmic divergence. This is addressed by reducing the linearly dependent columns of the data matrix. Additionally, we learn a lower bound (τ_{D_c}) for the average dwell time of DoS attacks, such that for any DoS attacks with an average dwell time $\tau_D \geq \tau_{D_c}$, closed-loop stability is guaranteed. As

Received 16 September 2024; revised 18 November 2024; accepted 5 December 2024. Date of publication 26 December 2024; date of current version 14 January 2025. This work was supported in part by NSF under Grant CNS-2148309, Grant EPCN-2210320, Grant CPS-2227185, Grant S&AS-1849198, Grant CPS-1851588, Grant SATC-2231651, and Grant CPS-2227153; and in part by the Army Research Office (ARO) under Grant W911NF-24-1-0174. Recommended by Senior Editor L. Menini. *(Corresponding author: Sayan Chakraborty.)*

Sayan Chakraborty and Zhong-Ping Jiang are with the CAN Lab, New York University, Brooklyn, NY 11201 USA (e-mail: sc8804@nyu.edu; zjiang@nyu.edu).

Weinan Gao is with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China (e-mail: gaown@mail.neu.edu.cn).

Kyriakos G. Vamvoudakis is with The Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: kyriakos@gatech.edu).

Digital Object Identifier 10.1109/LCSYS.2024.3522953

the impact of denial-of-service (DoS) is not taken into account in our previous work [17], [18], the developed methodologies are ineffective under DoS attack. This letter introduces a novel framework for designing resilient active learning-based optimal control policies adaptable to any duration of the DoS, an enhancement absent in our prior approaches. To our knowledge, active learning-based resilient control has not been explored before.

This letter is organized as follows. Section II provides the background information and problem formulation. Section III provides the learning framework using input-state data in DoS attacks. Sections IV and V discuss the divergence and convergence parameters, respectively. Section VI presents the simulation results. Concluding remarks are discussed in Section VII.

Notations: $\mathbb{R}_+$ denotes the set of nonnegative real numbers, $\mathbb{Z}_+$ denotes the set of nonnegative integers. σ_A denotes the complex spectrum of a square matrix A. $|.|$ denotes the Euclidean norm of a vector $x \in \mathbb{R}^n$ or the induced matrix norm for a matrix $A \in \mathbb{R}^{m\times n}$. For a real symmetric matrix A, $\lambda_m(A)$ and $\lambda_M(A)$ denote the minimum and maximum eigenvalues of A, respectively. For a quadratic Lyapunov function $V(x) = x^{\mathrm{T}}Px$, where $P \in \mathbb{R}^{m\times m}$ is a real symmetric and positive definite matrix, $x \in \mathbb{R}^m$, we have $\lambda_m(P)|x|^2 \leq V(x) \leq \lambda_M(P)|x|^2$. The Moore-Penrose pseudoinverse of an $m \times n$ matrix A is denoted as $A^\dagger$. The symbol $\otimes$ indicates the Kronecker product, $\mathrm{vec}(T) = \left[t_1^{\mathrm{T}}, t_2^{\mathrm{T}}, \ldots, t_m^{\mathrm{T}}\right]^{\mathrm{T}}$ with $t_i \in \mathbb{R}^r$ being the columns of $T \in \mathbb{R}^{r\times m}$. For a symmetric matrix $P \in \mathbb{R}^{m\times m}$, $\mathrm{vecs}(P) = [p_{11}, 2p_{12}, \ldots, 2p_{1m}, p_{22}, 2p_{23}, \ldots, 2p_{(m-1)m}, p_{mm}]^{\mathrm{T}} \in \mathbb{R}^{\frac{m(m+1)}{2}}$, for a column vector $v \in \mathbb{R}^n$, $\mathrm{vecv}(v) = [v_1^2, v_1v_2, \ldots, v_1v_n, v_2^2, v_2v_3, \ldots, v_{n-1}v_n, v_n^2]^{\mathrm{T}} \in \mathbb{R}^{\frac{n(n+1)}{2}}$. For any two sequences of vectors $a = \{a_i\}_{i=k_0}^{k_s}$, $b = \{b_i\}_{i=k_0}^{k_s}$, define $\Phi_a = [\mathrm{vecv}(a_{k_0+1}) - \mathrm{vecv}(a_{k_0}), \ldots, \mathrm{vecv}(a_{k_s}) - \mathrm{vecv}(a_{k_s-1})]^{\mathrm{T}}$, $J_{a,b} = [a_{k_0} \otimes b_{k_0}, \ldots, a_{k_s} \otimes b_{k_s}]^{\mathrm{T}}$, $J_a = [\mathrm{vecv}(a_{k_0}), \ldots, \mathrm{vecv}(a_{k_s})]^{\mathrm{T}}$, $\Xi_a = [a_{k_0} \otimes a_{k_1}, \ldots, a_{k_{s-1}} \otimes x_{k_s}]^{\mathrm{T}}$. I_n and 0_n are the identity and zero matrices of dimension $n \times n$, respectively.

II. Problem Formulation

Consider the following discrete-time cascade system:

$$x_{k+1} = Ax_k + Bu_k + Dw_k, \tag{1}$$
$$w_{k+1} = Ew_k, \tag{2}$$
$$e_k = Cx_k + Fw_k, \tag{3}$$

where $k \in \mathbb{Z}_+$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^n$ and $D \in \mathbb{R}^{n\times q}$ are unknown constant matrices, and $C \in \mathbb{R}^{1\times n}$, $E \in \mathbb{R}^{q\times q}$, $F \in \mathbb{R}^{1\times q}$ are known constant matrices, $e_k \in \mathbb{R}$ is the measurement output, $u_k \in \mathbb{R}$ the control input, $x_k \in \mathbb{R}^n$ is the state, $w_k \in \mathbb{R}^q$ is the exosystem state. The exosystem (2) generates the reference signal $y_{dk} = -Fw_k$ and the disturbance signal $d_k = Dw_k$ (see [19] for more details).

Assumption 1: The pair (A, B) is stabilizable and E has no eigenvalue with modulus smaller than one.

Assumption 2: $\mathrm{rank}\begin{bmatrix} A-\lambda I & B \\ C & 0 \end{bmatrix} = n+1$, $\lambda \in \sigma_E$.

Remark 1: Assumptions 1-2 are standard for solving the linear output regulation problem (LORP) [11], [19].

In the absence of DoS attacks, LORP is formulated by designing a controller of the form [11], [19]:

$$u_k = -Kx_k + Lw_k, \tag{4}$$

where $K \in \mathbb{R}^{1\times n}$ is the feedback control gain and $L \in \mathbb{R}^{1\times q}$ is the feedforward control gain. The feedforward gain L is obtained as $L = U + KX$, where $X \in \mathbb{R}^{n\times q}$ and $U \in \mathbb{R}^{1\times q}$ solves the following regulator equations

$$XE = AX + BU + D, \tag{5}$$
$$0 = CX + F. \tag{6}$$

For any given initial conditions x_0 and w_0, if the controller given in (4) solves the LORP, one has $\lim_{k\to\infty} \tilde{u}_k = 0$ and $\lim_{k\to\infty} \tilde{x}_k = 0$, where $\tilde{x}_k = x_k - Xw_k$ and $\tilde{u}_k = u_k - Uw_k$ are the errors between the actual state/input and their corresponding steady-state components Xw_k and Uw_k, respectively. By solving the LORP problem, we attempt to solve the problem of asymptotic tracking with disturbance rejection. The error system of (1) can be obtained as follows

$$\tilde{x}_{k+1} = A\tilde{x}_k + B\tilde{u}_k, \;\; e_k = C\tilde{x}_k. \tag{7}$$

Since K does not rely on (5)-(6), we can first design K such that $A - BK$ is Schur (i.e., its eigenvalues are inside the open unit disk) [11]. To obtain the optimal controller gain $K^\star$, the following optimization problem is solved.

Problem 1:

$$\min_{\tilde{u}} \sum_{k=0}^{\infty} \lambda_-^{2k}\left(\tilde{x}_k^{\mathrm{T}}Q\tilde{x}_k + \tilde{u}_k^2\right) \tag{8}$$
$$\text{s.t} \quad \tilde{x}_{k+1} = A\tilde{x}_k + B\tilde{u}_k, \tag{9}$$

where $Q = Q^{\mathrm{T}} \succ 0$ and $\lambda_- \geq 1$.
Defining $\tilde{x}_{k,\lambda_-} = \lambda_-^k\tilde{x}_k$, and $\tilde{u}_{k,\lambda_-} = \lambda_-^k\tilde{u}_k$, we obtained,

$$\tilde{x}_{k+1,\lambda_-} = \lambda_-^{k+1}\tilde{x}_{k+1} = \tilde{A}\tilde{x}_{k,\lambda_-} + \tilde{B}\tilde{u}_{k,\lambda_-}, \tag{10}$$

where $\tilde{A} = A\lambda_-$, and $\tilde{B} = B\lambda_-$. Then, Problem 1 can be converted to a standard discrete-time linear quadratic regulator problem as follows.

Problem 2:

$$\min_{\tilde{u}_{k,\lambda_-}} \sum_{k=0}^{\infty}\left(\tilde{x}_{k,\lambda_-}^{\mathrm{T}}Q\tilde{x}_{k,\lambda_-} + \tilde{u}_{k,\lambda_-}^2\right)$$
$$\text{s.t.} \quad \tilde{x}_{k+1,\lambda_-} = \tilde{A}\tilde{x}_{k,\lambda_-} + \tilde{B}\tilde{u}_{k,\lambda_-}. \tag{11}$$

The solution to Problem 2 is an optimal feedback controller of the form

$$\tilde{u}_{k,\lambda_-}^\star = -K^\star\tilde{x}_{k,\lambda_-}. \tag{12}$$

The following can be obtained from (12),

$$u_k^\star = -\lambda_-^{-k}K^\star\tilde{x}_{k,\lambda_-} + Uw_k = -K^\star x_k + L^\star w_k, \tag{13}$$

where $L^\star = U + K^\star X$, $K^\star = (1 + \tilde{B}^{\mathrm{T}}P^\star\tilde{B})^{-1}\tilde{B}^{\mathrm{T}}P^\star\tilde{A}$ and $P^\star = P^{\star T} \succ 0$ solves the following discrete-time algebraic Riccati equation

$$\tilde{A}^{\mathrm{T}}P^\star\tilde{A} - P^\star + Q - \tilde{A}^{\mathrm{T}}P^\star\tilde{B}\left(1 + \tilde{B}^{\mathrm{T}}P^\star\tilde{B}\right)^{-1}\tilde{B}^{\mathrm{T}}P^\star\tilde{A} = 0.$$

In this letter, we choose $\lambda_- = \delta\lambda_+^{\frac{1}{T-1}}$, where $\delta > 1$ and T is the DoS duration parameter. Section V gives a justification for

the choice of λ_-. The parameter $\lambda_+ > 1$ ensures that $\frac{A}{\lambda_+}$ is Schur, implying that the Lyapunov equation of the form

$$\frac{A^{\mathrm{T}}}{\lambda_+}P_+\frac{A}{\lambda_+} - P_+ = -\epsilon I_n,\ \epsilon > 0, \tag{14}$$

has a unique solution $P_+ = P_+^{\mathrm{T}} \succ 0$. Section IV illustrates a data-driven technique to learn λ_+ directly from input-state data when A is unknown.

Next, let $\{h_m\}_{m\in\mathbb{Z}_+}$ denote the sequence of off and on transitions of DoS, where $h_0 \geq 0$. They are the time instants at which DoS exhibits a transition from zero (transmissions are successful) to one (transmissions are not successful). The m^{th} DoS attack interval of length τ_m is represented as $\mathcal{J}_m \coloneqq [h_m, h_m + \tau_m)$. For each interval $[k_1, k_2]$, let $\Lambda_D(k_1, k_2) \coloneqq \bigcup_{m\in\mathbb{Z}_+} \mathcal{J}_m \bigcap [k_1, k_2]$ and $\Lambda_N(k_1, k_2) \coloneqq [k_1, k_2] \setminus \Lambda_D(k_1, k_2)$ denote the set of time instants where communication is denied and allowed, respectively. The following assumptions are now needed.

Assumption 3 (DoS Frequency): There exist constants $\eta > 1$ and $\tau_D > 0$ such that $\forall k_2 > k_1 \geq 0$,

$$n(k_1, k_2) \leq \eta + \frac{k_2 - k_1}{\tau_D}, \tag{15}$$

where $n(k_1, k_2)$ denotes the number of DoS off/on transitions occurring in the interval $[k_1, k_2]$.

Assumption 4 (DoS Duration): For any $k_2 > k_1 \geq 0$,

$$|\Lambda_D(k_1, k_2)| \leq \rho + \frac{k_2 - k_1}{T}, \tag{16}$$

where $|\Lambda_D(k_1, k_2)|$ denotes the Lebesgue measure of the set $\Lambda_D(k_1, k_2)$, and $T > 1$, $\rho > 0$ are chosen arbitrarily.

Remark 2: In Assumption 3, τ_D is the average dwell-time between DoS off and on transitions, and η is the chattering bound. Assumption 4 similarly constrains DoS durations, ensuring the average communication interruption duration does not exceed a certain fraction of time, as specified by $1/T$. Here, ρ serves as a regularization term, akin to η. These standard assumptions restrict DoS attacks by their average frequency and duration [6], [7].

III. Learning the Optimal Controller

In this section, we propose an online strategy to learn the optimal controller (13) while the system is under DoS attacks. We use policy iteration technique [20] to learn the optimal feedback gain matrix $K^\star$, which implements both the policy evaluation

$$A_j^{\mathrm{T}}P_jA_j - \lambda_-^{-2}P_j + \lambda_-^{-2}Q + \lambda_-^{-2}K_j^{\mathrm{T}}K_j = 0 \tag{17}$$

and policy improvement

$$K_{j+1} = \left(1 + \lambda_-^2 B^{\mathrm{T}}P_jB\right)^{-1}\lambda_-^2 B^{\mathrm{T}}P_jA, \tag{18}$$

where $A_j = A - BK_j$. To learn the optimal feedforward gain matrix $L^\star$, we first define the following Sylvester map

$$S_{X_i} = X_iE - AX_i. \tag{19}$$

Let $X_0 = 0$ and select X_1 such that $CX_1 + F = 0$, and each X_i for $i = 2, 3, \ldots, h+1$ is selected such that the vectors $\mathrm{vec}(X_i)$ form a basis for the null space of $(I_q \otimes C)$ with dimension $h = (n-1)q$. Then, $X = X_1 + \sum_{i=2}^{h+1}\alpha_iX_i$ gives a general solution to (6), where $\alpha_i \in \mathbb{R}$. Then, (5) implies $S_X = S_{X_1} + \sum_{i=2}^{h+1}\alpha_iS_{X_i} = BU + D$. Using S_{X_i}, for $i = 0, 1, \ldots, h+1$, the pair (X, U) that solves (5)-(6) can be obtained by solving Problem 3 in [11]. Then $L^\star = U + K^\star X$ can be learned. Next, we demonstrate a procedure for learning $K^\star$ and $L^\star$ using input-state data in two phases.

A. Phase 1: Learning the Optimal Feedback Gain

By defining $x_{k,i} = x_k - X_iw_k$, along the trajectories of (1)-(3) and using (19), the following can be obtained

$$x_{k+1,i} = A_jx_{k,i} + B\big(u_k + K_jx_{k,i}\big) + \Pi_iw_k, \tag{20}$$

where $\Pi_i = D - S_{X_i}$. Along the trajectories of (20) and using (17), the following can be obtained [17], [18]

$$\begin{aligned} x_{k+1,i}^{\mathrm{T}}P_jx_{k+1,i} - x_{k,i}^{\mathrm{T}}P_jx_{k,i} + \lambda_-^{-2}x_{k,i}^{\mathrm{T}}Q_jx_{k,i} &= 2u_k^{\mathrm{T}}\Gamma_{1j}x_{k,i} \\ &+ 2\big(K_jx_{k,i}\big)^{\mathrm{T}}\Gamma_{1j}x_{k,i} - \big(K_jx_{k,i}\big)^{\mathrm{T}}\Gamma_{2j}\big(K_jx_{k,i}\big) + u_k^{\mathrm{T}}\Gamma_{2j}u_k \\ &+ 2x_{k,i}^{\mathrm{T}}\Theta_{1ij}w_k + 2u_k^{\mathrm{T}}\Theta_{2ij}w_k + w_k^{\mathrm{T}}\Theta_{3ij}w_k \\ &+ \left(\lambda_-^{-2} - 1\right)x_{k,i}^{\mathrm{T}}P_jx_{k,i}, \end{aligned} \tag{21}$$

where $Q_j = Q + K_j^{\mathrm{T}}K_j$, $\Theta_{1ij} = A^{\mathrm{T}}P_j\Pi_i$, $\Theta_{2ij} = B^{\mathrm{T}}P_j\Pi_i$, $\Theta_{3ij} = \Pi_i^{\mathrm{T}}P_j\Pi_i$, $\Gamma_{1j} = B^{\mathrm{T}}P_jA$, $\Gamma_{2j} = B^{\mathrm{T}}P_jB$. By Assumption 4, there always exists a sequence $\{k_s\}_{s=0}^{\infty}$ such that communications are allowed. Following [17], [18], by collecting online data, the following linear equation can be obtained from (21)

$$\Psi_{1ij}\theta_{1ij} = -J_{x_i,x_i}\mathrm{vec}\big(Q_j\big), \tag{22}$$

$$\begin{aligned} &\text{where } \Psi_{1ij} = \lambda_-^2\Big[\Phi_{x_i} - \left(\lambda_-^{-2} - 1\right)J_{x_i}, -2J_{x_i,u} - 2J_{x_i,x_i}\left(I_n \otimes K_j^{\mathrm{T}}\right), J_{K_jx_i} - J_u, -2J_{w,x_i}, -2J_{w,u}, -J_w\Big], \\ &\theta_{1ij} = \Big[\mathrm{vecs}\big(P_j\big)^{\mathrm{T}}, \mathrm{vec}\big(\Gamma_{1j}\big)^{\mathrm{T}}, \mathrm{vecs}\big(\Gamma_{2j}\big)^{\mathrm{T}}, \mathrm{vec}\big(\Theta_{1ij}\big)^{\mathrm{T}}, \\ &\mathrm{vec}\big(\Theta_{2ij}\big)^{\mathrm{T}}, \mathrm{vecs}\big(\Theta_{3ij}\big)^{\mathrm{T}}\Big]^{\mathrm{T}}. \end{aligned}$$

For certain choices of E, the matrix J_w may lack full column rank. In such cases, the N number of linearly dependent columns of J_w are reduced to ensure that $\bar{\Psi}_{1ij}$ achieves full rank [17], where $\bar{\Psi}_{1ij}$ is constructed from $\bar{J}_w$, containing only the linearly independent columns of J_w. Similarly, let $\bar{\theta}_{1ij}$ be constructed from the reduced $\mathrm{vecs}(\Theta_{3ij})$ denoted as $\overline{\mathrm{vecs}(\Theta_{3ij})}$. Then, (22) can be solved as

$$\bar{\theta}_{1ij} = -\bar{\Psi}_{1ij}^{\dagger}J_{x_i,x_i}\mathrm{vec}\big(Q_j\big). \tag{23}$$

Assumption 5: There exists a $s^\star \in \mathbb{Z}_+$ such that for all $s > s^\star$, $i = 0, 1, \ldots, (n-1)q + 1$, and for any sequence $k_0 < k_1 < \cdots < k_s$ the following holds

$$\begin{aligned} \mathrm{rank}\Big(\big[J_{x_i}\ J_{x_i,u}\ J_u\ J_{w,x_i}\ J_{w,u}\ \bar{J}_w\big]\Big) &= \frac{n(n+1)}{2} + n \\ &+ 1 + nq + q + \frac{q(q+1)}{2} - N. \end{aligned} \tag{24}$$

Remark 3: Assumption 5, analogous to the persistency of excitation condition, is fundamental to adaptive control [21] and learning-based control [10], [11], [22]. It ensures a unique solution to (23) for all $j \in \mathbb{Z}_+$, and the sequences $\{P_j\}_{j=0}^{\infty}$ and $\{K_j\}_{j=0}^{\infty}$ from steps 10-14 of Algorithm 1 converge to the optimal values $P^\star$ and $K^\star$ [18].

Algorithm 1 Active Learning-Based Resilient Control

1: Input constants $\lambda_+ > 1$, $\alpha > 1$, $\delta > 1$, $T > 1$ and $\epsilon_0 > 0$.
2: Compute $X_0, X_1, \cdots, X_{h+1}$ and apply any locally essentially bounded input u_k on $[k_0, k_s]$. For $i = 0, 1, \cdots, h+1$ compute $J_{x_i}, J_{x_i,u}, J_u, J_{w,x_i}, J_{w,u}, J_w$ such that (24) holds.
3: Set $i \leftarrow 0$ and $j \leftarrow 0$.
4: **repeat**
5: $\lambda_+ \leftarrow \alpha\lambda_+$
6: **if** $\bar{\Psi}_3$ is full column rank **then**
7: Solve P_+ from (29)
8: **end if**
9: **until** $P_+ \succ 0$
10: Choose a stabilizing K_0 and set $\lambda_- \leftarrow \delta\lambda_+^{\frac{1}{T-1}}$.
11: **repeat**
12: Solve for P_j and $K_{j+1} = (1+\Gamma_{2j})^{-1}\Gamma_{1j}$ from (23)
13: $j \leftarrow j+1$
14: **until** $|P_k - P_{k-1}| < \epsilon_0$
15: Set $j\star \leftarrow j$, $X_0 = 0$. Then, $\Pi_0 = D$.
16: For $i = 0$, obtain B and D by solving (25).
17: **repeat**
18: $i \leftarrow i+1$
19: Solve for $S_{X_i} = D - \Pi_i$ from (25)
20: **until** $i = h+1$
21: Obtain X and U by solving Problem 3 in [11].
22: Return $K_{j\star}$ and $L_{j\star} = U + K_{j\star}X$.

B. Phase 2: Learning the Optimal Feedforward Gain

Similar to Phase 1, one can formulate (25) to learn the Sylvester maps by using the input-state data collected from Phase 1 [18]

$$\Psi_{2i}\theta_{2i} = \Xi_{x_i}\text{vec}(I_n), \tag{25}$$

where $\Psi_{2i} = [\frac{1}{2}J_{x_i}, J_{x_i,u}, J_{w,x_i}]$, $\theta_{2i} = [\text{vecs}(A^{\mathrm{T}} + A)^{\mathrm{T}}, \text{vec}(B)^{\mathrm{T}}, \text{vec}(\Pi_i)^{\mathrm{T}}]^{\mathrm{T}}$. Satisfying (24) ensures that (25) has a unique solution for each $i = 0, \ldots, h+1$. See Steps 15-22 of Algorithm 1 to learn $L^\star$ using (25). Then, the following resilient controller can be obtained for the closed-loop system

$$u_k = \begin{cases} -K_{j\star}x_k + L_{j\star}w_k, & \text{if } k \in \Lambda_N(k_1, k_2), \\ Uw_k, & \text{if } k \in \Lambda_D(k_1, k_2), \end{cases} \tag{26}$$

where $K_{j\star}$ and $L_{j\star}$ are obtained using Algorithm 1.

IV. Learning the Divergence Parameter λ_+

In this section, we develop a data-driven methodology to learn the divergence parameter from input-state data. Along the trajectories of (1), the following is obtained using (14)

$$\begin{aligned} x_{k+1}^{\mathrm{T}}P_+x_{k+1} - x_k^{\mathrm{T}}P_+x_k = &\left(\lambda_+^2 - 1\right)x_k^{\mathrm{T}}P_+x_k + 2u_k^{\mathrm{T}}\Lambda_1x_k \\ &+ 2x_k^{\mathrm{T}}\Lambda_3w_k + u_k^{\mathrm{T}}\Lambda_2u_k + 2u_k^{\mathrm{T}}\Lambda_4w_k + w_k^{\mathrm{T}}\Lambda_5w_k \\ &- \epsilon\lambda_+^2x_k^{\mathrm{T}}x_k, \end{aligned} \tag{27}$$

where, $\Lambda_1 = B^{\mathrm{T}}P_+A$, $\Lambda_2 = B^{\mathrm{T}}P_+B$, $\Lambda_3 = A^{\mathrm{T}}P_+D$, $\Lambda_4 = B^{\mathrm{T}}P_+D$, $\Lambda_5 = D^{\mathrm{T}}P_+D$. Noting that $x_k = x_{k,0}$ the input-state data collected in Section III-A can be used to obtain the following linear equation from (27)

$$\Psi_3\theta_3 = -\epsilon J_{x,x}, \tag{28}$$

where $\Psi_3 = \frac{1}{\lambda_+^2}[\Phi_x - (\lambda_+^2 - 1)J_x, -2J_{x,u}, -J_u, -2J_{w,x}, -2J_{w,u}, -J_w]$, $\theta_3 = [\text{vecs}(P_+)^{\mathrm{T}}, \text{vec}(\Lambda_1)^{\mathrm{T}}, \text{vecs}(\Lambda_2)^{\mathrm{T}}, \text{vec}(\Lambda_3)^{\mathrm{T}}, \text{vec}(\Lambda_4)^{\mathrm{T}}, \text{vecs}(\Lambda_5)^{\mathrm{T}}]^{\mathrm{T}}$. Since J_w may not have full column rank, following the discussion in Section III-A, (28) can be solved as:

$$\bar{\theta}_3 = -\epsilon\bar{\Psi}_3^{\dagger}J_{x,x}, \tag{29}$$

For clarity, the reduction process is demonstrated in the following example.

Example 1: Let $\Psi\theta = b$, where $\Psi = [a_1, a_2, a_3, a_4, a_5] \in \mathbb{R}^{m\times 5}$, $\theta = [x_1, x_2, x_3, x_4, x_5]^{\mathrm{T}}$, and $b = [b_1, b_2, \ldots, b_m]^{\mathrm{T}}$. If $a_5 = \alpha_1a_3 + \alpha_2a_4$, then $\Psi\theta = b \implies \bar{\Psi}\bar{\theta} = b$, where $\bar{\Psi} = [a_1, a_2, a_3, a_4]$, and $\bar{\theta} = [x_1, x_2, x_3 + \alpha_1x_5, x_4 + \alpha_2x_5]^{\mathrm{T}}$. Here, $\bar{\Psi}$ contains only the linearly independent columns of Ψ, allowing $\bar{\Psi}\bar{\theta} = b$ to be uniquely solved. Additionally, $\bar{\Psi} = \Psi M$ and $\bar{\theta} = S\theta$, where $M = \begin{bmatrix} I_2 & 0_{2\times 2} \\ 0_{3\times 2} & M_{2,2} \end{bmatrix}$, $S = \begin{bmatrix} I_2 & 0_{2\times 3} \\ 0_{2\times 2} & S_{2,2} \end{bmatrix}$, $M_{2,2} = \begin{bmatrix} I_2 \\ 0_{1,2} \end{bmatrix}$, $S_{2,2} = \begin{bmatrix} 1 & 0 & \alpha_1 \\ 0 & 1 & \alpha_2 \end{bmatrix}$.

The following lemma shows that the solution to (29) is unique if the rank condition (24) holds and $\frac{A}{\lambda_+}$ is Schur.

Lemma 1: For $i = 0$, suppose that the rank condition (24) holds and $\lambda_+ > 1$ is such that $\frac{A}{\lambda_+}$ is Schur, then $\bar{\Psi}_3$ has full column rank.

Proof: Select a vector $\bar{\beta}$ such that the following holds

$$\bar{\Psi}_3\bar{\beta} = -\epsilon J_{x,x}, \tag{30}$$

where $\bar{\beta} = [\text{vecs}(Y_1)^{\mathrm{T}}, \text{vec}(Y_2)^{\mathrm{T}}, \text{vecs}(Y_3)^{\mathrm{T}}, \text{vec}(Y_4)^{\mathrm{T}}, \text{vec}(Y_5)^{\mathrm{T}}, \overline{\text{vecs}(Y_6)}^{\mathrm{T}}]^{\mathrm{T}}$ with the matrices $Y_1 = Y_1^{\mathrm{T}} \in \mathbb{R}^{n\times n}$, $Y_2 \in \mathbb{R}^{1\times n}$, $Y_3 \in \mathbb{R}$, $Y_4 \in \mathbb{R}^{n\times q}$, $Y_5 \in \mathbb{R}^{1\times q}$, $Y_6 = Y_6^{\mathrm{T}} \in \mathbb{R}^{q\times q}$. We need to show that $\bar{\beta}$ can be uniquely determined. From (30), the following equation can be obtained

$$\begin{aligned} &J_x\text{vecs}(\Omega_1) + 2J_{x,u}\text{vec}(\Omega_2) + J_u\text{vecs}(\Omega_3) + 2J_{w,x}\text{vec}(\Omega_4) \\ &\quad + 2J_{w,u}\text{vec}(\Omega_5) + \bar{J}_w\overline{\text{vecs}(\Omega_6)} = 0, \\ &\text{where } \Omega_1 = A^{\mathrm{T}}Y_1A - \lambda_+^2Y_1 + \epsilon\lambda_+^2I_n,\ \Omega_2 = B^{\mathrm{T}}Y_1A - Y_2, \\ &\Omega_3 = B^{\mathrm{T}}Y_1B - Y_3,\ \Omega_4 = A^{\mathrm{T}}Y_1D - Y_4, \\ &\Omega_5 = B^{\mathrm{T}}Y_1D - Y_5,\ \Omega_6 = D^{\mathrm{T}}Y_1D - Y_6. \end{aligned} \tag{31}$$

The rank condition (24) implies that $\text{vecs}(\Omega_1) = 0$, $\overline{\text{vec}(\Omega_2)} = 0$, $\text{vecs}(\Omega_3) = 0$, $\text{vec}(\Omega_4) = 0$, $\text{vec}(\Omega_5) = 0$ and $\overline{\text{vecs}(\Omega_6)} = 0$. Thus, we have the following

$$\frac{A^{\mathrm{T}}}{\lambda_+}Y_1\frac{A}{\lambda_+} - Y_1 = -\epsilon I_n. \tag{32}$$

Since $\frac{A}{\lambda_+}$ is Schur stable, Y_1 can be uniquely determined. As shown in Example 1, one can obtain $M_{2,2}$ and $S_{2,2}$ for any given Ψ_3. Based on the linearity of the vecs operator, one can see that $\overline{\text{vecs}(\Omega_6)} = S_{2,2}\text{vecs}(D^{\mathrm{T}}Y_1D - Y_6) = 0 \implies \overline{\text{vecs}(Y_6)} = \overline{\text{vecs}(D^{\mathrm{T}}Y_1D)}$. Thus, every element of the $\bar{\beta}$ is uniquely determined due to the uniqueness of Y_1. The proof is thus complete. ■

The following lemma provides a condition for the stability of $\frac{A}{\lambda_+}$, that is verifiable using only input-state data.

Lemma 2: Suppose the rank condition (24) holds, then $\frac{A}{\lambda_+}$ is Schur if and only if the following holds

1) The matrix $\bar{\Psi}_3$ is full column rank.

2) The symmetric matrix P_+ obtained by solving (29) is positive definite.

Proof: ($\Rightarrow$) Let $\frac{A}{\lambda_+}$ is Schur. Then Property 1) can be inferred Lemma 1. From Lemma 1 we can further observe that P_+ can be uniquely obtained from (29). Furthermore, P_+ satisfying the Lyapunov equation (14) also satisfies (29). This implies that property 2) holds.

($\Leftarrow$) Suppose both the properties hold. Then, from (32), the matrix $P_+ \succ 0$ solves the Lyapunov equation (14). This implies $\frac{A}{\lambda_+}$ is Schur. The proof is thus complete. ■

Remark 4: Using Lemmas 1 and 2, λ_+ is learned in Steps 4-9 in Algorithm 1.

V. Selection of the Convergence Parameter λ_-

The following theorem shows that (26) acts as a resilient controller under a proper selection of λ_- and if the DoS frequency is small enough.

Theorem 1: Under Assumptions 1-4, for any DoS duration parameter $T > 1$ and $\lambda_+ > 1$, let $\lambda_- = \delta\lambda_+^{\frac{1}{T-1}}$, where $\delta > 1$. Then, the system (1)-(3) in closed-loop with the controller (26) achieves output regulation if the following condition holds

$$\tau_D \geq \frac{T}{T-1}\frac{\log\left(\sqrt{\frac{\lambda_M(P_{j\star})\lambda_M(P_+)}{\lambda_m(P_{j\star})\lambda_m(P_+)}}\right)}{\log\left(\delta^2\right)} := \tau_{D_c}. \tag{33}$$

Proof: Under the action of the controller (26), the system (1)-(3) evolves as a switched system as follows

$$\tilde{x}_{k+1} = \begin{cases} A_{j\star}\tilde{x}_k, & \text{if } k \in \Lambda_N(k_1,k_2), \\ A\tilde{x}_k, & \text{if } k \in \Lambda_D(k_1,k_2). \end{cases} \tag{34}$$

Consider the following piecewise-quadratic Lyapunov function

$$V_k = \begin{cases} \tilde{x}_k^{\mathrm{T}} P_{j\star}\tilde{x}_k := \bar{V}(x_k), & \text{if } k \in \Lambda_N(k_1,k_2), \\ \mu_1\tilde{x}_k^{\mathrm{T}} P_+\tilde{x}_k := \tilde{V}(x_k), & \text{if } k \in \Lambda_D(k_1,k_2), \end{cases} \tag{35}$$

where $\mu_1 > 0$ is to be determined. When $k \in \Lambda_N(k_1,k_2)$, the following can be obtained using (17) and (34)

$$\bar{V}(\tilde{x}_{k+1}) - \bar{V}(\tilde{x}_k) \leq \left(\lambda_-^{-2} - 1 - \lambda_-^{-2}c_1\right)\bar{V}(\tilde{x}_k), \tag{36}$$

where $0 < c_1 = \frac{\lambda_m(Q)}{\lambda_M(P_{j\star})} < 1$ (see [23]). When $k \in \Lambda_D(k_1,k_2)$, the following can be obtained using (14) and (34)

$$\tilde{V}(\tilde{x}_{k+1}) - \tilde{V}(\tilde{x}_k) \leq \left(\lambda_+^2 - 1\right)\tilde{V}(\tilde{x}_k). \tag{37}$$

For any $l \in \mathbb{Z}_+$, let the two intervals $[h_l+\tau_l, h_{l+1}]$ and $[h_l, h_l+\tau_l]$ be the two intervals where the communication is allowed and denied, respectively. Then, from (36) and (37), it can be obtained that the Lyapunov function V_k satisfies the following

$$V_k \leq \begin{cases} \omega^k V_{h_l+\tau_l}, & \text{if } k \in \left[h_l+\tau_l, h_{l+1}\right], \\ \lambda_+^{2k} V_{h_l}, & \text{if } k \in [h_l, h_l+\tau_l], \end{cases} \tag{38}$$

where, $\omega = \lambda_-^{-2}(1-c_1)$. Due to the fact that $\bar{V}(\tilde{x}_k)$ and $\tilde{V}(\tilde{x}_k)$ are quadratic with $P_{j\star}$ and P_+ being symmetric positive definite matrices, the following can be obtained

$$\bar{V}(\tilde{x}_k) \leq \mu_2\tilde{V}(\tilde{x}_k), \;\; \tilde{V}(\tilde{x}_k) \leq \mu_2\bar{V}(\tilde{x}_k), \tag{39}$$

where $\mu_1 = \sqrt{\frac{\lambda_M(P_{j\star})\lambda_m(P_{j\star})}{\lambda_M(P_+)\lambda_m(P_+)}}$ and $\mu_2 = \sqrt{\frac{\lambda_M(P_{j\star})\lambda_M(P_+)}{\lambda_m(P_{j\star})\lambda_m(P_+)}}$. Therefore, at all switching times k, we have that $V_k \leq \mu_2 V_{k-1}$. Thus, for all $k \in \mathbb{Z}_+$, the Lyapunov function V_k satisfies the following

$$V_k \leq \mu_2^{n(0,k)}\omega^{|\Lambda_N(0,k)|}\lambda_+^{2|\Lambda_D(0,k)|}V_0 \tag{40}$$

Using Assumptions 3-4 and the fact that $|\Lambda_N(0,k)| = k - |\Lambda_D(0,k)| \geq k-\rho-\frac{k}{T}$, the following can be obtained from (40)

$$V_k \leq \mu_2^{\eta}\left(\frac{\lambda_+^2}{\omega}\right)^{\rho}(1-c_1)^{\left(\frac{T-1}{T}\right)k}\left(\mu_2^{\frac{1}{\tau_D}}\left(\frac{1}{\lambda_-^2}\right)^{\frac{T-1}{T}}\lambda_+^{\frac{2}{T}}\right)^k V_0.$$

Since $0 < 1-c_1 < 1$, we need that

$$\mu_2^{\frac{1}{\tau_D}}\left(\frac{1}{\lambda_-^2}\right)^{\frac{T-1}{T}}\lambda_+^{\frac{2}{T}} \leq 1. \tag{41}$$

By taking log on both sides of (41) and choosing $\lambda_- = \delta\lambda_+^{\frac{1}{T-1}}$, one can obtain (33). Thus, (40) implies the following $\forall\tau_D \geq \tau_{D_c}$

$$V_k \leq \mu_2^{\eta}\left(\frac{\lambda_+^2}{\omega}\right)^{\rho}(1-c_1)^{\left(\frac{T-1}{T}\right)k}V_0. \tag{42}$$

Using (42), the following holds $\forall\tau_D > \tau_{D_c}$ and $\lambda_- = \delta\lambda_+^{\frac{1}{T-1}}$

$$|\tilde{x}_k| \leq c_2(1-c_1)^{\left(\frac{T-1}{2T}\right)k}|\tilde{x}_0|, \tag{43}$$

$$|e_k| \leq |C|c_2(1-c_1)^{\left(\frac{T-1}{2T}\right)k}|\tilde{x}_0|, \tag{44}$$

where $c_2 = \sqrt{\mu_2^{\eta}\left(\frac{\lambda_+^2}{\omega}\right)^{\rho}\frac{\max\{\lambda_M(P_{j\star}),\mu_1\lambda_M(P_+)\}}{\min\{\lambda_m(P_{j\star}),\mu_1\lambda_m(P_+)\}}}$. Thus, we have $\lim_{k\to\infty}(x_k - Xw_k) = 0$ and $\lim_{k\to\infty} e_k = 0$, which implies output regulation. The proof is thus complete. ■

Algorithm 1 summarizes the proposed active learning-based resilient control methodology.

VI. Simulation Results

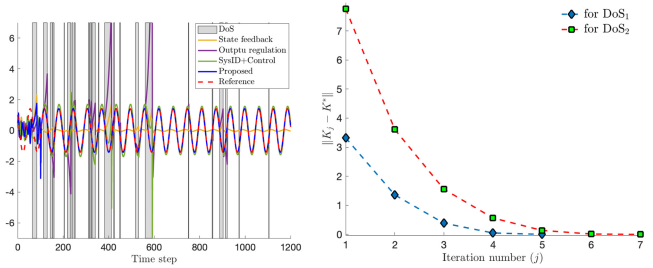
In this section, we show the efficacy of the proposed methodology by considering a numerical example. The system matrices are given as follows

$$A = \begin{bmatrix} 1 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, D = \begin{bmatrix} 0 & 0 \\ 0 & 0.1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \end{bmatrix},$$
$$E = \begin{bmatrix} \cos(0.1) & -\sin(0.1) \\ \sin(0.1) & \cos(0.1) \end{bmatrix}, F = \begin{bmatrix} -2 & 0 \end{bmatrix}.$$

The initial conditions are $x_0 = [0.5\ 0]^{\mathrm{T}}$, $w_0 = [0.5\ 0.5]^{\mathrm{T}}$, and Q is an identity matrix. The proposed methodology is tested with two DoS parameter sets: $\mathrm{DoS}_1 := \{\rho = 0.35, \tau_D = 3, T = 2, \eta = 1.1\}$ (applied for $k \in [0, 599]$) and $\mathrm{DoS}_2 := \{\rho = 0.35, \tau_D = 3, T = 20, \eta = 1.1\}$ (applied for $k \in [600, 1200]$). Input-state data are collected for $k \in [0, 100]$ using sinusoidal exploration signals. Algorithm 1 uses $\epsilon_0 = 0.1$ and $\delta = 1.1$, yielding $\lambda_+ = 2.2$. For DoS_1, $\lambda_- = 2.42$ is chosen, and the learned controller is applied after $k = 100$. When the attack changes at $k = 600$, $\lambda_- = 1.1466$ is used to relearn the resilient controller for DoS_2. The controller gains are compared in Table I and Fig. 1(b). From (33), the critical dwell-times are $\tau_{D_c} = 33.4827$ for DoS_1 and $\tau_{D_c} = 10.0376$ for DoS_2. While $\tau_D \geq \tau_{D_c}$ is

TABLE I
COMPARISON OF CONTROLLER GAIN VALUES

DOS_1			DOS_2		
Controller	1	2	Controller	1	2
$K^\star$	1.5165	2.2484	$K^\star$	0.81648	0.40568
K_5	1.5168	2.2494	K_7	0.81653	0.40581
$L^\star$	6.0679	2.4609	$L^\star$	2.0562	0.17888
L_5	6.0697	2.4621	L_7	2.0565	0.17904



(a) (b)

Fig. 1. (a) Tracking and disturbance rejection under DoS attacks, (b) Convergence of K_j to $K^\star$.

sufficient to ensure stability, the system handles frequent DoS attacks with $\tau_D = 3$. Fig. 1(a) shows the evolutions of the reference and output trajectories under the application of different controllers, with the DoS attacks represented as shaded areas. It can be seen that the proposed active controller provides superior performance compared to other controllers. We compared our methodology with an indirect approach based on system identification, using the dynamic mode decomposition method to identify system matrices. Conventional system identification techniques assume data is collected at regular intervals and require additional preprocessing to handle missing data. However, in the advent of DoS attacks, missing data is inevitable and may lead to inaccurate estimation of the system matrices. We have used interpolation technique to approximate the missing data points. As shown in the SysID+Control plot in Fig. 1(a), the performance of this indirect approach is less favorable than anticipated. In contrast, the proposed direct approach methodology guarantees convergence and uniqueness by utilizing online data collected over any sequence of time steps $k_0 < k_1 < \cdots < k_s$ as long as the rank condition in (24) is satisfied. The learned controller can track the reference signal even in the presence of varying DoS attacks.

VII. CONCLUSION

We have proposed an active learning-based controller design for discrete-time, linear, uncertain systems under DoS attacks. Leveraging switching systems theory and adaptive dynamic programming, the controller ensures closed-loop stability by learning an optimal control policy and a critical average dwell-time. Stability and resilience are guaranteed when the DoS average dwell-time exceeds the critical value. Divergence and convergence parameters, estimated from online input-state data, were central to learning the resilient controller. Numerical simulations demonstrate the effectiveness of the proposed methodology.

REFERENCES

[1] H.-T. Sun, C. Peng, and F. Ding, "Self-discipline predictive control of autonomous vehicles against denial of service attacks," *Asian J. Control*, vol. 24, no. 6, pp. 3538–3551, 2022.
[2] H. Mahvash, S. A. Taher, and J. M. Guerrero, "Modified backstepping control for cyber security enhancement of a wind farm based DFIG against false data injection, hijack and denial of service cyber attacks," *Electric Power Syst. Res.*, vol. 231, Jun. 2024, Art. no. 110357.
[3] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Security*, vol. 14, no. 1, pp. 1–33, 2011.
[4] H. Zhang, W. Han, X. Lai, D. Lin, J. Ma, and J. Li, "Survey on cyberspace security," *Sci. China Inf. Sci.*, vol. 58, pp. 1–43, Nov. 2015.
[5] S.-S. Sun, Y.-X. Li, and Z. Hou, "Prescribed performance-based resilient model-free adaptive control for CPSs against aperiodic DoS attacks," *Int. J. Robust Nonlinear Control*, vol. 34, no. 5, pp. 3335–3350, 2024.
[6] C. De Persis and P. Tesi, "Input-to-state stabilizing control under denial-of-service," *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 2930–2944, Nov. 2015.
[7] S. Feng and P. Tesi, "Resilient control under denial-of-service: Robust design," *Automatica*, vol. 79, pp. 42–51, May 2017.
[8] W. Liu, J. Sun, G. Wang, F. Bullo, and J. Chen, "Resilient control under quantization and denial-of-service: Codesigning a deadbeat controller and transmission protocol," *IEEE Trans. Autom. Control*, vol. 67, no. 8, pp. 3879–3891, Aug. 2022.
[9] R. Zhang, G. Li, and R. Yang, "Secure control for the discrete-time CPSs under DoS attacks via a switching strategy," in *Proc. IEEE 12th Data Driven Control Learn. Syst. Conf. (DDCLS)*, 2023, pp. 1678–1683.
[10] Y. Jiang and Z.-P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699–2704, 2012.
[11] W. Gao and Z.-P. Jiang, "Adaptive dynamic programming and adaptive optimal output regulation of linear systems," *IEEE Trans. Autom. Control*, vol. 61, no. 12, pp. 4164–4169, Dec. 2016.
[12] W. Gao, C. Deng, Y. Jiang, and Z.-P. Jiang, "Resilient reinforcement learning and robust output regulation under denial-of-service attacks," *Automatica*, vol. 142, Aug. 2022, Art. no. 110366.
[13] W. Gao, Z.-P. Jiang, and T. Chai, "Resilient control under denial-of-service and uncertainty: An adaptive dynamic programming approach," 2024, *arXiv:2411.06689*.
[14] W. Liu, J. Sun, G. Wang, F. Bullo, and J. Chen, "Data-driven resilient predictive control under denial-of-service," *IEEE Trans. Autom. Control*, vol. 68, no. 8, pp. 4722–4737, Aug. 2023.
[15] F. Li and Z. Hou, "Learning-based model-free adaptive control for nonlinear discrete-time networked control systems under hybrid cyber attacks," *IEEE Trans. Cybern.*, vol. 54, no. 3, pp. 1560–1570, Mar. 2024.
[16] X. Wang, D. Ding, X. Ge, and Q.-L. Han, "Neural-network-based control for discrete-time nonlinear systems with denial-of-service attack: The adaptive event-triggered case," *Int. J. Robust Nonlinear Control*, vol. 32, no. 5, pp. 2760–2779, 2022.
[17] S. Chakraborty, W. Gao, L. Cui, F. L. Lewis, and Z.-P. Jiang, "Learning-based adaptive optimal output regulation of discrete-time linear systems," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 10283–10288, 2023.
[18] S. Chakraborty, W. Gao, K. G. Vamvoudakis, and Z.-P. Jiang, "Adaptive optimal output regulation of discrete-time linear systems: A reinforcement learning approach," in *Proc. 62nd IEEE Conf. Decision Control (CDC)*, 2023, pp. 7950–7955.
[19] J. Huang, *Nonlinear Output Regulation: Theory and Applications*. Philadelphia, PA, USA: SIAM, 2004.
[20] G. Hewer, "An iterative technique for the computation of the steady state gains for the discrete optimal regulator," *IEEE Trans. Autom. Control*, vol. 16, no. 4, pp. 382–384, Aug. 1971.
[21] I. Mareels and J. W. Polderman, *Adaptive Systems: An Introduction*. Boston, MA, USA: Springer, 1996.
[22] K. G. Vamvoudakis and N.-M. T. Kokolakis, "Synchronous reinforcement learning-based control for cognitive autonomy," *Found. Trends® Syst. Control*, vol. 8, nos. 1–2, pp. 1–175, 2020.
[23] J. Garloff, "Bounds for the eigenvalues of the solution of the discrete Riccati and Lyapunov equations and the continuous Lyapunov equation," *Int. J. Control*, vol. 43, no. 2, pp. 423–431, 1986.