Model-Based Actor–Critic Learning for Secure Tracking Control of Vehicle Systems With a Dynamic Event-Triggered Mechanism

Jinliang Liu, *Senior Member, IEEE*, Bin Lv, Meng Yang, Engang Tian, *Senior Member, IEEE*, Chen Peng, *Senior Member, IEEE*, and Tingwen Huang, *Fellow, IEEE*

***Abstract*—This article proposes a synergistic model-based actor–critic (AC) learning framework for secure vehicle tracking under a dynamic event-triggered scheme (DETS) and subject to deception attacks. To explicitly resolve the conflict between communication conservation and learning convergence, the AC algorithm is structurally embedded within a discrete linear quadratic tracker (LQT) and backstepping architecture. This co-design utilizes the nominal linear structure as a rigorous prior to analytically derive update laws, ensuring stable Hamilton–Jacobi–Bellman approximation even when data are sparse and stochastically corrupted. By bounding the policy search space, it ensures a minimal number of learnable parameters and a clear physical interpretation. Lyapunov analysis rigorously proves the asymptotic convergence of the tracking errors to zero. Numerical simulations demonstrate superior tracking accuracy, faster convergence, and significant bandwidth savings compared to standard model-free reinforcement learning baselines.**

***Index Terms*—Deception attacks, discrete vehicle systems, dynamic event-triggered scheme (DETS), linear quadratic tracker (LQT), model-based actor–critic (AC) learning, optimal tracking control.**

I. Introduction

In RECENT years, vehicle platooning has emerged as a pivotal technology for intelligent transportation systems to alleviate congestion [1], [2]. However, its reliance on vehicle-to-vehicle networks introduces critical challenges: limited bandwidth and vulnerability to cyber threats. These challenges can lead to packet loss and data tampering, thereby severely degrading the system's stability. Ensuring precise tracking under these resource constraints and security risks is a fundamental problem requiring urgent attention. Recent research [3], [4] has explored a variety of advanced control strategies. For example, adaptive control [5], fuzzy control [6], and robust control [7] have each been applied to improve the performance of connected vehicle systems. An adaptive coordinated tracking control scheme was developed for autonomous ground vehicle formations under deception attacks and sensor failures [8]. Similarly, Wen et al. [9] proposed an optimized attitude control for quadrotor autonomous aerial vehicle system using an adaptive actor–critic (AC) reinforcement learning (RL) approach with fuzzy-logic approximation. Dong et al. [10] developed a cooperative quantized event-based fuzzy tracking control for nonlinear autonomous surface vehicles.

With the evolution of wireless communication technology, autonomous vehicles can perform control tasks via wireless data transmission. However, a growing body of research has demonstrated that networked vehicle control must contend with both limited communication bandwidth and increasingly sophisticated security threats. A discrete-time linear quadratic tracker (LQT) framework offers a compelling solution [11]. Zhao et al. [12] developed an adaptive cooperative tracking control scheme featuring resilient event-triggering to handle actuator faults and communication link failures in multiagent formations. Despite these advantages, it has yet to fully characterize performance trade-offs under combined high-frequency triggering and adversarial network conditions [13], nor has it explored extensions to non-Gaussian disturbance models—gaps that motivate the present investigation [14].

In the realm of RL [15], [16], [17], several key algorithms have been developed to address complex system constraints and communication limitations. For instance, a synchronization learning scheme utilizing hybrid order adaptive dynamic optimizations was developed in [18] to ensure secure communication in networked environments. Meanwhile, to handle unreliable communications in cyber-physical systems, a zero-sum optimal control strategy was investigated in [19] via secure decentralized policy iteration. Despite the success of RL in many areas, its application to optimal tracking control is

Received 8 January 2026; revised 27 March 2026 and 25 April 2026; accepted 26 April 2026. Date of publication 6 May 2026; date of current version 21 July 2026. This work was supported in part by the National Natural Science Foundation of China under Grant 62373252 and Grant 62173231 and in part by the Startup Foundation for Introducing Talent of Nanjing University of Information Science and Technology (NUIST) under Grant 2024r063. *(Corresponding author: Jinliang Liu.)*

Jinliang Liu and Bin Lv are with the School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: liujinliang@vip.163.com; lvbin2002@163.com).

Meng Yang is with College of Command and Control Engineering, Army Engineering University of PLA, Nanjing 210007, China (e-mail: mengyang9505@163.com).

Engang Tian is with the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China (e-mail: tianengang@163.com).

Chen Peng is with the School of Mechatronic Engineering and Automation, Department of Automation, Shanghai University, Shanghai 200444, China (e-mail: c.peng@shu.edu.cn).

Tingwen Huang is with the Faculty of Computer Science and Control Engineering, Shenzhen University of Advanced Technology, Shenzhen 518055, China (e-mail: huangtw2024@163.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSMC.2026.3688583.

Digital Object Identifier 10.1109/TSMC.2026.3688583

limited [20]. Recently, the authors introduced a sampled-data event-triggered secure estimation and backstepping tracking control framework for nonlinear systems under sensor and network attacks [21]. Moreover, a UKF-based multistep heuristic dynamic programming scheme was proposed in [22] to achieve optimal event-triggering control for nonlinear systems with asymmetric input constraints. The work in [23] employs memory event-triggered tracking differentiators to guarantee performance and stability in vehicle systems. Optimal tracking control design includes feedback and feedforward components. The authors of [24] converted a LQT into a linear quadratic regulator (LQR) using Q-learning by assuming the reference trajectory was generated by a linear command generator. Other articles focus on theoretical analysis over practical experimentation and do not consider practical issues, such as environmental disturbances and input saturation, which may occur in actual vehicle systems [25].

Securing networked nonlinear control systems under limited communication and active cyber threats adds a critical dimension to optimal control design [26]. Traditional optimal control aims to minimize a quadratic or general performance index while respecting resource constraints and system dynamics, but the Hamilton–Jacobi–Bellman (HJB) equation at its core remains intractable for most nonlinear systems [27]. The RL-based AC methods have shown promise in approximating the HJB solution and deriving near-optimal feedback laws without full model knowledge [28]. Recently, to further handle strict-feedback nonlinearities and reduce the number of tunable parameters, optimized backstepping techniques have been proposed that embed RL-based cost approximation within the backstepping recursion. Such techniques provide a unified framework for secure, resource-aware optimal control in complex networked environments [29].

This study investigates the complex cybernetic interplay between physical vehicle dynamics, resource-constrained communications, and intelligent cyber-defense. By framing the platoon as an integrated cyber-physical system, we explicitly resolve the nontrivial conflict between sparse data transmission enforced by dynamic event-triggered scheme (DETS) and the stable convergence of learning algorithms under stochastic deception attacks.

The primary contributions of this study are as follows.

1) In comparison with [15], [21], which treat learning and communication independently, this study synergistically resolves the fundamental conflict between sparse data transmission and the data-hungry nature of RL. By structurally embedding AC learning within the discrete LQT and backstepping framework, the nominal linear time invariant (LTI) physics acts as a rigorous prior. This synergy guarantees stable HJB approximation even when tracking data is sporadic and stochastically corrupted by deception attacks.
2) In contrast to the system model under deception attacks in [30] and [31], this study couples AC learning directly with the physical error dynamics to actively compensate for malicious data injections. It analytically guarantees bounded tracking errors against energy-constrained deception attacks without relying on standalone detection modules.
3) Unlike event-triggered control strategies [32], [33] that treat communication independently, the proposed DETS interacts constructively with the AC learning process. By dynamically adjusting the threshold, it filters out redundant benign data to optimize bandwidth while isolating anomalous attack signals, achieving a deeply synergistic co-design of communication efficiency and cyber-physical security.

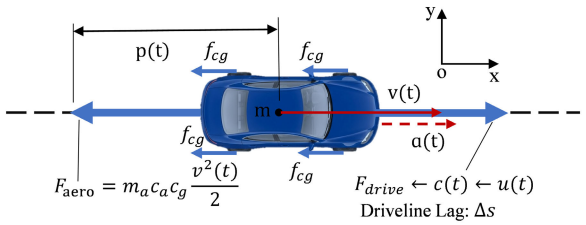


Fig. 1. Longitudinal vehicle dynamics model.

II. Preliminaries

A. Longitudinal Vehicle Dynamics

To explicitly illustrate the physical formulation of the vehicle tracking problem, the longitudinal vehicle dynamics model is established, as depicted in Fig. 1. The vehicle model is defined as follows [34]:

$$\begin{cases} p(t+1) = p(t) + v(t)\,\Delta t \\ v(t+1) = v(t) + a(t)\,\Delta t \\ a(t+1) = a(t) + (f(t) + q(t)\,c(t))\,\Delta t \end{cases} \tag{1}$$

where $p(t)$, $v(t)$, and $a(t)$ denote the position, velocity, and acceleration of the vehicle, respectively. The engine input $c(t)$ is the required torque for braking or driving and Δt represents the sampling interval. The functions $f(t)$ and $q(t)$ are defined as $f(t) = -1/\Delta s(a(t) + m_a c_a c_g(v^2(t))/(2m) + (f_{c_g})/(m)) - m_a c_a c_g(v(t)a(t))/(\Delta s m)$, and $q(t) = 1/(\Delta s m)$, where m_a, m, Δs, c_a, c_g, and f_{c_g} denote the specific air mass, vehicle mass, time lag, cross-sectional area, drag coefficient and resistance [34]. Based on the feedback linearization technique established in [35], the nonlinear longitudinal dynamics are strictly transformed into a nominal LTI structure. This deliberate decoupling strategy ensures that the core challenge addressed by our reinforcement learning framework is not identifying physical nonlinearities, but rather securing the optimal control policy against unmodeled, attack-induced stochasticity and severe data sparsity under event-triggered constraints. The engine input is described by

$$c(t) = mu(t) + m_a c_a c_g \frac{v^2(t)}{2} + f_{c_g} + m_a c_a c_g v(t)\,a(t) \tag{2}$$

where $u(t)$ denotes the control signal acting on the vehicle. The state of vehicle is represented by the vector $x(t) = [p(t), v(t), a(t)]^T$. Thus, the vehicle dynamics in (1) can be rewritten as

$$x(t+1) = Ax(t) + Bu(t) \tag{3}$$

where

$$A = \begin{bmatrix} 1 & \Delta t & 0 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 - \frac{\Delta t}{\Delta s} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \frac{\Delta t}{\Delta s} \end{bmatrix}.$$

B. Optimal Tracking Control

This command generator model introduces the reference trajectory that is described by $r(t+1) = Fr(t)$, where F is a constant Hurwitz matrix that ensures stability of the reference dynamics. The desired state is $r(t) \triangleq [p_r(t), v_r(t), a_r(t)]^T$. According to (3), the tracking error is

$$z(t) = x(t) - r(t). \tag{4}$$

Based on the methodology presented by [36], the control input is formulated as

$$u(t) = \sigma\left(u_b(t) + u_f(t)\right) \tag{5}$$

where the feedforward and feedback terms are specified as

$$\begin{aligned} u_f(t) &= -B^{-1}(A - F)\, r(t) \\ u_b(t) &= K_b z(t) \end{aligned} \tag{6}$$

where K_b is the feedback gain matrix. By minimizing the quadratic cost function via the stationarity condition of the Hamiltonian function, the optimal control law is derived. This LQR problem is solved in order to obtain

$$K_b = -\left(B^T P B + I\right)^{-1} B^T P A \tag{7}$$

where P is the positive definite solution to the discrete-time algebraic Riccati equation (ARE)

$$A^T P A - A^T P B \left(I + B^T P B\right)^{-1} B^T P A - P + Q = 0 \tag{8}$$

where Q represents the state weighting matrix. The resulting error dynamics are expressed as

$$z(t+1) = (A + B K_b)\, z(t). \tag{9}$$

Hence, when $A + BK_b$ is a Hurwitz matrix, the tracking error $z(t)$ converges to zero.

C. DETS Design

The DETS is designed to conserve scarce network resources. It regulates the transmission frequency of signals from sensor within the network. In this sensor network framework, the release of a sampler's current data packet to the controller is governed by its associated triggering condition. The proposed DETS condition is given by

$$\begin{cases} \mu \Phi_i(t) > \Upsilon_i(t) \\ \Phi_i(t) = e_i^T(t) R e_i(t) - \zeta z_i^T\left(T_i^t\right) R z_i\left(T_i^t\right) \end{cases} \tag{10}$$

where the measurement error $e_i(t)$ is explicitly defined as: $e_i(t) = z_i(T_i^t) - z_i(T_i^t + \Delta)$. Here, $z_i(T_i^t + \Delta)$ is an intermediate variable constructed as $z_i(T_i^t + \Delta) = \xi[z_i(t) - z_i(T_i^t)] + z_i(T_i^t)$, where $\xi \in (0, 1]$ serves as an adjustment factor. i represents different triggering moments. T_i^t denotes the sequence of releasing instants. μ and ζ are known constant values. R is the matrix that needs to be created for the DETS. $\Upsilon_i(k)$ is the auxiliary offset variable updated according to the following update law:

$$\Upsilon_i(t+1) = \varepsilon \Upsilon_i(t) - \Phi_i(t) \tag{11}$$

where $\Upsilon_i(t) \geq 0$, and $\varepsilon \in (0, 1)$ in which $\mu\varepsilon \geq 1$. The term $e_i(t) = z_i(T_i^t) - z_i(T_i^t + \Delta)$ is defined as the error with

$$z_i\left(T_i^t + \Delta\right) = \xi\left[z_i(t) - z_i\left(T_i^t\right)\right] + z_i\left(T_i^t\right). \tag{12}$$

If the DETS condition (10) is violated, the current measurement is transmitted over the wireless network. Abrupt signals are more likely than normal variations to trigger a violation of condition (10), which can lead to erroneous packet transmissions. Setting an adjustment factor $\xi \in (0, 1]$ can help to reduce this problem, when creating the error $e_i(t)$ for each moment i.

Remark 1: Unlike conventional DETS, the proposed mechanism introduces a smoothing factor ξ to filter high-frequency noise and reduce unnecessary transmissions. Specifically, a smaller ξ enhances robustness against anomalous spikes by prioritizing the previously transmitted value. Parameters μ and ζ determine the baseline sensitivity, while the decay factor ϵ (set to 0.4) governs the dynamic threshold's reset rate, where a smaller ϵ increases triggering frequency. This design effectively balances communication efficiency with tracking accuracy and system security.

Based on the definition of $e_i(t)$ and the expression given in (12), the signal transmitted through the networked channel is characterized as follows $z_i(T_i^t) = 1/\xi e_i(t) + z_i(t)$, then let $z(T^t) \triangleq \mathrm{col}_N\{z_i(T_i^t)\}$, $\hat{e}(t) \triangleq \mathrm{col}_N\{e_i(t)\}$, $\bar{z}(t) \triangleq \mathrm{col}_N\{z_i(t)\}$, the triggered signals are denoted as:

$$z\left(T^t\right) = \frac{1}{\xi}\hat{e}(t) + \bar{z}(t). \tag{13}$$

D. Deception Attacks

The error state can be stochastically tampered with by deception attacks. Under potential deception attacks, the actual error state is

$$z^*(t) = \beta(t)\,\bar{h}\left(T^t\right) + (1 - \beta(t))\, z\left(T^t\right) \tag{14}$$

where $\bar{h}(T^t)$ represents the attack signal. $\beta(t)$ is employed to characterize the occurrence of attacks. Given the constraints imposed by limited resources and the requirement for attack concealment, a novel hypothesis is proposed

$$\left\|\bar{h}\left(T^t\right)\right\|_2 \leq \left\|M z\left(T^t\right)\right\|_2 \tag{15}$$

where M is known matrix. When $\beta(t) = 0$, there are no deception attacks. $\beta(t) = 1$ signifies their presence. $\beta(t)$ assumes values in the set $\{0, 1\}$, and the probability distribution: $Pr\{\beta(t) = 0\} = 1 - \bar{\beta}$, $Pr\{\beta(t) = 1\} = \bar{\beta}$, $0 \leq \bar{\beta} \leq 1$ for any $\bar{\beta}$.

Assume that the vehicle controller input $u^*(t)$ may be subject to stochastic modification by injection attacks before reaching the agent. Based on (5), (6), and (14), modeling the actual received controller input $u^*(t)$ in light of potential injection threats

$$u^*(t) = \rho(t)\, u^a\left(T^t\right) + (1 - \rho(t))\, u\left(T^t\right) \tag{16}$$

where $u^a(T^t)$ is attack function of $u(T^t)$, as determined by the attacker. $\rho(t)$ represents a random variable which describes the occurrence of attacks. In consideration of resource limitations and the need for attack concealment, it is hypothesized that

$$\left\|u^a\left(T^t\right)\right\|_2 \leq \left\|Ou\left(T^t\right)\right\|_2 \tag{17}$$

where O is a known matrix of appropriate dimensions. $\rho(t) = 0$ denotes the absence of injection attacks. $\rho(t) = 1$ indicates their presence. $\rho(t)$ takes values in the set $\{0, 1\}$. Follows the distribution of probabilities: $Pr\{\rho(t) = 0\} = 1 - \bar{\rho}$, $Pr\{\rho(t) = 1\} = \bar{\rho}$, $0 \leq \bar{\rho} \leq 1$ for any $\bar{\rho}$.

Remark 2: The proposed deception attack models (14)–(17) reflect realistic vulnerabilities in vehicular networks, such as man-in-the-middle or false data injection attacks. Specifically, the bounded energy constraints in (15) and (17) capture the adversary's strategic need to maintain stealth and evade anomaly detection systems, while the probabilistic nature (14) accounts for the stochastic behavior of network intrusions.

After discussing the vehicle dynamics modeling, optimal tracking control, dynamic event-triggered control and deception attack mentioned above, the final vehicle error system state can be expressed as

$$\begin{aligned} z^*(t+1) &= A\left[\left((1-(t))z(T^t) + \beta(t)\bar{h}\left(T^t\right)\right] + Bu^*(t) + \epsilon(t) \\ &= Az^*(t) + Bu^*(t) + \epsilon(t) \end{aligned} \tag{18}$$

where $z^*(t) = x^*(t) - r^*(t)$ denotes the tracking error, and $x^*(t)$ represents the actual system state. Constant matrices with compatible dimensions are A and B. $r^*(t)$ signifies the actual reference trajectory at time t. The term $\epsilon(t)$ comprehensively encapsulates unmodeled high-order nonlinear aerodynamics and environmental stochasticity, which are strictly bounded. The actual received controller input is $u^*(t)$.

The purpose of this article is to achieve secure tracking control then the following inequality holds of discrete vehicle systems under the proposed model-based AC algorithm. The following lemma is presented for the aim of making the subsequent analysis easier.

Lemma 1: [37]: If $V(x(t), r(t))$ is positive continuous function, such that the initial value $V(0, 0)$ is bounded. When ignoring the Gaussian noise term, where $\psi(t) = [x^T(t), r^T(t)]^T$ and P is a symmetric positive-definite matrix, then the following inequality holds $V(x(t), r(t)) \approx 1/2\psi^T(t)P\psi(t)$.

Lemma 2: [37]: Based on the definition in Lemma 1 and ignoring the terms related to noise, the critic can be expressed as $V(\varphi(t+1)) + c(t) = 1/2(\varphi(t) \otimes \varphi(t))^T \mathrm{vec}(H)$. By calculating minimal value of critic, the actor can be determined. $\nabla_{u(t)} Q_v(x(t), r(t), u(t)) = 0$ is a stationary condition that is required for optimality.

III. Main Results

This section presents a model-based AC learning method, which is built using the LQT framework. This algorithm permits the system to follow the reference trajectory with optimal precision through the minimization of a predetermined performance metric. The LQT problem is reformulated as a LQR problem through a causal representation of the tracking control, which enables a solution without requiring preexisting information on the system parameters. By incorporating the error system as described in (18), the augmented system is given

$$\begin{bmatrix} x^*(t+1) \\ r^*(t+1) \end{bmatrix} = \bar{A}\begin{bmatrix} x^*(t) \\ r^*(t) \end{bmatrix} + \bar{B}u^*(t) + \bar{\epsilon}(t)$$
$$\bar{A} = \begin{bmatrix} A & 0 \\ 0 & F \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} B \\ 0 \end{bmatrix}, \quad \bar{\epsilon}(t) = \begin{bmatrix} \epsilon(t) \\ 0 \end{bmatrix}. \tag{19}$$

The LQT index is defined as follows:

$$\begin{aligned} \min c(t) = \frac{1}{2}\big[&u^{*T}(t)u^*(t) + \delta(x^*(t) - r^*(t))^T \\ &\times (x^*(t) - r^*(t))\big] \end{aligned} \tag{20}$$

where the parameter $\delta \in (0, \infty)$ functions as a low-gain tuning factor for the input gain, which ensures system stability. The value function is defined as the accumulated cost

$$\begin{aligned} V(x^*(t), r^*(t)) = \frac{1}{2}\sum_{j=t}^{\infty}\big[&u^{*T}(j)u^*(j) + \delta(x^*(j) - r^*(j))^T \\ &\times (x^*(j) - r^*(j))\big]. \end{aligned} \tag{21}$$

The actor is assumed to have the following form and correlates to the control law of augmented system (19) in order to develop a LQT controller:

$$u^*(t) = K_b x^*(t) + K_f r^*(t) \tag{22}$$

where $K = [K_b \ K_f]$, with K_b and K_f denoting the feedback and feedforward control gains, respectively. According to Lemma 1, for (19) with the actor defined in (22), (21) is approximated as

$$V(x^*(t), r^*(t)) \approx \frac{1}{2}\psi^T(t)P\psi(t) \tag{23}$$

where $\psi(t) = \left[x^{*T}(t), r^{*T}(t)\right]^T$, P denotes a positive definite matrix. By invoking Bellman's optimality principle, based on (21), the critic function is

$$Q_v(x^*(t), r^*(t), u^*(t)) = V(x^*(t+1), r^*(t+1)) + c(t). \tag{24}$$

Based on the quadratic representation of value function established in Lemma 1 with (24), the following content demonstrates that this LQT problem is reformulated as a LQR problem. Moreover, by expressing reference trajectory and state variables in a unified framework, the LQT control law is obtained by calculating the corresponding augmented ARE.

According to Lemma 2, (19) and (21), (24) is expressed as $Q_v = 1/2(\varphi(t) \otimes \varphi(t))^T \mathrm{vec}(H)$, where $\varphi(t) = [x^{*T}(t), r^{*T}(t), u^{*T}(t)]^T$. The LQT problem about optimal control is provided by

$$u^*(t) = -\left(\bar{B}^T P\bar{B} + I\right)^{-1}\bar{B}^T P\bar{A}\psi(t) \tag{25}$$

where P satisfies the following LQT ARE:

$$\delta\bar{I} - P + \bar{A}^T P\bar{A} - \bar{A}^T P\bar{B}\left(I + \bar{B}^T P\bar{B}\right)^{-1}\bar{B}^T P\bar{A} = 0. \tag{26}$$

Specifically, backstepping decomposes the high-order vehicle dynamics into lower-dimensional virtual subsystems to handle structural complexity. Within this framework, the LQT formulation is employed to optimize each step, thereby simultaneously ensuring tracking precision and energy efficiency.

A. Critic Update

The goal of the critic is to evaluate the performance of the current control policy. By iteratively solving the Bellman equation, the critic approximates the value function associated with the current tracking error and control input. Based on (24), the critic function can be evaluated recursively as follows:

$$\begin{aligned} &Q_v\left(x^*(t), r^*(t), u^*(t)\right) \\ &\quad - Q_v\left(x^*(t+1), r^*(t+1), u^*(t+1)\right) \\ &= \frac{1}{2}\left[u^{*T}(t)u^*(t) + \delta\left(x^*(t) - r^*(t)\right)^T\left(x^*(t) - r^*(t)\right)\right] - \eta \end{aligned} \tag{27}$$

where $\eta = 1/2[(\bar{x}^{*T}(t+1) - \bar{x}^{*T}(t+2))P\bar{\epsilon}(t+1) + (\bar{\epsilon}^T(t+1) - \bar{\epsilon}^T(t+1))P\bar{x}^*(t+1)$ is the noise term.

According to Lemma 2, the critic is expressed as

$$\begin{aligned} &\frac{1}{2}\left(\varphi(t)\otimes\varphi(t) - \varphi(t+1)\otimes\varphi(t+1)\right)^T \mathrm{vec}(H) \\ &= \frac{1}{2}u^{*T}(t)u^*(t) + \frac{1}{2}\delta\left(x^*(t) - r^*(t)\right)^T\left(x^*(t) - r^*(t)\right) - \eta. \end{aligned} \tag{28}$$

To address the term η, (28) can be expressed as

$$\left(\varphi(t)\otimes\varphi(t) - \gamma\varphi(t+1)\otimes\varphi^T(t+1)^T\right)\mathrm{vec}(H) \approx c(t) \tag{29}$$

where γ denotes discount factor. When there is no noise, γ is set to 1. Otherwise, when noise increases, γ falls. Equation (29) represents the update law for critic.

B. Actor Update

Based on the value function estimated by the critic, the actor updates the control gain K to minimize the cost function, achieving policy improvement.

Finding the maximization of the critic function is the standard method for identifying the actor

$$K = \arg\max_{u^*(t)} Q_v\left(x^*(t), r^*(t), u^*(t)\right). \tag{30}$$

However, because greedy policy improvement requires global maximum at every stage, it is difficult in continuous action spaces. An appealing option is to adjust the policy along the gradient of critic. According to the deterministic policy gradient paradigm, this strategy improvement update is given by

$$K^{j+1} = K^j + \alpha\mathbb{E}_{x^*(t),r^*(t)}\left[\nabla_K Q_v\left(x^*(t), r^*(t), u^*(t)\right)\right] \tag{31}$$

where $\nabla_K Q_v$ denotes the gradient of critic in relation to control scheme, $\mathbb{E}\nabla_{\bar{x}(t)}[\nabla_K Q_v(x^*(t), r^*(t).u^*(t))]$ represents the expectation of this gradient. α denotes learning rate.

The gradient of the critic with respect to the control input and the gradient of the control input with respect to K can be multiplied via the chain rule to determine the gradient of the critic with respect to K

$$\begin{aligned} &\mathbb{E}_{x^*(t),r^*(t)}\left[\nabla_K Q_v\left(x^*(t), r^*(t), u^*(t)\right)\right] \\ &= \mathbb{E}_{x^*(t),r^*(t)}\left[\nabla_{u_k} Q_v\left(x^*(t), r^*(t), u^*(t)\right)\nabla_K u^*(t)\right] \end{aligned} \tag{32}$$

where $\nabla_{u^*(t)}Q_v(x^*(t), r^*(t), u^*(t))\nabla_K u^*(t) = 2(H_{uxr}\psi(t) + H_{uu}K\psi(t) + B^T P\bar{\epsilon}(t))\psi^T(t)$ with $H_{uxr} = \begin{bmatrix} H_{ux} & H_{ur} \end{bmatrix}$.

It is possible to approximate the expectation function

$$\begin{aligned} &\mathbb{E}_{x^*(t),r^*(t)}\left[\nabla_K Q_v\left(x^*(t), r^*(t), u^*(t)\right)\right] \\ &\approx \frac{1}{N}\sum_{t=1}^{N}\left[2\left(H_{uxr}\psi(t) + H_{uu}K\psi(t) + B^T P\bar{\epsilon}(t)\right)\psi^T(t)\right]. \end{aligned} \tag{33}$$

Given that the noise term becomes closer to zero as N increases, the expectation function can be expressed as

$$\begin{aligned} &\mathbb{E}_{x^*(t),r^*(t)}\left[\nabla_K Q_v\left(x^*(t), r^*(t), u^*(t)\right)\right] \\ &\approx \frac{2}{N}\sum_{t=1}^{N}\left[\left(H_{uxr} + H_{uu}K\right)\left(\psi(t)\psi^T(t)\right)\right]. \end{aligned} \tag{34}$$

Equation (31) can be equivalently expressed as

$$K^{j+1} = K^j + \frac{2\alpha}{N}\sum_{t=1}^{N}\left[\left(H_{uxr} + H_{uu}K^j\right)\left(\psi(t)\psi^T(t)\right)\right]. \tag{35}$$

Using the deterministic policy gradient technique, the actor update rule is presented in (35). Without taking into account the prior action, the update of actor is solely dependent on the system state, critic function, and reference trajectory.

Crucially, (35) analytically captures the adversarial interplay within this cyber-physical co-design. Under DETS, the state updates $\psi(t)$ are severely sporadic, while deception attacks inject bounded adversarial gradients into the system. Rather than treating these as independent additives, this framework leverages the LTI structural priors H_{uxr}, H_{uu}. This prior rigorously constrains the policy search space, ensuring that the learning convergence remains stable and secure even when driven by attack-influenced, highly sparse data.

Theorem 1: Using the value function in (21), consider this LQT problem for this system in (19). Then, the control law derived from the solution of the actor update equation (35) asymptotically stabilizes $z^*(t)$ and achieves the lowest possible value of the cost functional (21) among all stabilizing controllers.

Proof: Based on the critic update law for (29), the stable critic is attained under the following conditions:

$$\begin{aligned} &\frac{1}{2}\left(\varphi(t)\otimes\varphi(t) - \gamma\varphi(t+1)\otimes\varphi(t+1)\right)^T \\ &\quad\times \mathrm{vec}(H) - c(t) \approx 0. \end{aligned} \tag{36}$$

Based on (35), the steady actor is achieved

$$\frac{1}{N}\sum_{t=1}^{N}\left[2\left(H_{uxr} + H_{uu}K\right)\left(\psi(t)\psi^T(t)\right)\right] = 0. \tag{37}$$

When $\psi(t)\psi^T(t) \neq 0$, it can be shown from (37) that $H_{uxr} + H_{uu}K = 0$ and

$$K = -H_{uu}^{-1}H_{uxr} = -\left(I + \bar{B}^T P\bar{B}\right)^{-1}\bar{B}^T P\bar{A}. \tag{38}$$

The similarity between (25) and (38) suggests that the obtained control aligns with the optimal LQT solution.

According to the evolution of $\psi(t)$, $z^*(t)$ converges to 0 as $t \to \infty$. The definition of Lyapunov function is

$$V\left(x^*(t), r^*(t)\right) = \frac{1}{2}\psi^T(t)P\psi(t) \geq 0. \tag{39}$$

Based on (26), its difference

$$V\left(x^*\left(t+1\right),r^*\left(t+1\right)\right)-V\left(x^*\left(t\right),r^*\left(t\right)\right)=\frac{1}{2}\psi^T\left(t\right)\left(-\delta\bar{I}-K^TK\right)\psi\left(t\right)\leq 0 \tag{40}$$

$x^*(t), r^*(t)$ converges to zero as $t \to \infty$. That is described as $\lim_{t\to\infty} z(t) = \lim_{t\to\infty}(x^*(t) - r^*(t)) = 0$.

The LQT control obtained from (35) asymptotically stabilizes the tracking error. The next section demonstrates that the learned controller yields the minimum value of (21) out of all stabilizing strategies. Note that

$$\frac{1}{2}\psi^T\left(t\right)P\psi\left(t\right)=\frac{1}{2}\sum_{j=t}^{\infty}\left[\psi^T\left(j\right)P\psi\left(j\right)-\psi^T\left(j+1\right)P\times\psi\left(j+1\right)\right]+\frac{1}{2}\psi_\infty^TP\psi_\infty. \tag{41}$$

When $\psi_\infty^T P\psi_\infty = 0$, by substituting (41) into the value function in (21), value function is determined by

$$\begin{aligned}V\left(\psi\left(t\right)\right)=&\frac{1}{2}\psi^T\left(t\right)P\psi\left(t\right)\\&+\frac{1}{2}\sum_{j=t}^{\infty}\left[u^*\left(j\right)+\left(I+\bar{B}^TP\bar{B}\right)^{-1}\times\bar{B}^TP\bar{A}\psi\left(j\right)\right]^T\left(I+\bar{B}^TP\bar{B}\right)\\&\times\left[u^*\left(j\right)+\left(I+\bar{B}^TP\bar{B}\right)^{-1}\bar{B}^TP\bar{A}\psi\left(j\right)\right].\end{aligned} \tag{42}$$

The equation (42) is minimized when $u^*(t) = K\psi(t)$, which is equivalent to the actor in (38). Thus, the value function in (21) is minimized by the learned control, completing the proof.

C. Optimized Position Control

Given the position reference trajectory $r(t) = [p_r(t), v_r(t), a_r(t)]^T \in \mathbb{R}^3$, the tracking error is defined as $z_1(t) = x^*(t) - r^*(t)$ and $z_2(t) = u^*(t) - \alpha(t)$, where $\alpha(t)$ denotes the actual control. The resulting error dynamics are expressed as

$$\begin{aligned}\Delta z_1\left(t\right)&=x^*\left(t+1\right)-x^*\left(t\right)+r^*\left(t+1\right)-r^*\left(t\right)\\\Delta z_2\left(t\right)&=u^*\left(t+1\right)-u^*\left(t\right)-\Delta\alpha\left(t\right).\end{aligned} \tag{43}$$

The following two-step backstepping is used to determine the ideal position control. During the backstepping stage, the AC RL is applied, with the critic responsible for enhancing control performance. Implementing the control action is the goal of actor, thereby yielding the optimized position tracking control input $\hat{u}^*$.

Step 1: The update error dynamics in (43) can be equivalently expressed as $z_1(t + 1) = \phi z_1(t) + B\, z_2(t)$, where $\phi = A - I - k_1 B$. The virtual control is written as

$$\alpha\left(t\right)=-k_1z_1\left(t\right)+\left(F-I\right)r^*\left(t\right) \tag{44}$$

where $k_1 > 1/2$ serves as the design parameter.

Consider the Lyapunov function candidate for the backstepping procedure as $L_1(t) = 1/2 z_1^T(t) z_1(t)$. By computing the time derivative $\Delta L_1(t)$ along the dynamics (43) and applying Young's inequality and the Cauchy–Buniakowsky–Schwarz, $ab \leq (a^2/2) + (b^2/2)$, $(A^TB)^2 \leq \|A\|^2\|B\|^2$, where $a, b \in \mathbb{R}$, $A, B \in \mathbb{R}^n$ [21], we obtain the following inequality:

$$\Delta L_1\left(t\right)\leq-\left(2k_1-1\right)L_1\left(t\right)+\frac{1}{2}z_2^T\left(t\right)z_2\left(t\right). \tag{45}$$

Step 2: The RL technique is used to determine optimal vehicle position control in the last backstepping step. In association with (43), the performance index is specified as

$$J_1\left(z_2\left(t\right)\right)=\sum_{s=t}^{\infty}V_1\left(z_2\left(s\right),u^*\left(s\right)\right) \tag{46}$$

where $V_1(z_2(t), u^*(t)) = z_2^T(t) z_2(t) + u^{*T}(t) u^*(t)$ is cost function. Let u_a^* denote actual optimal control input. An ideal performance index is represented by

$$\begin{aligned}J_1^*\left(z_2\left(t\right)\right)&=\min_{u\in\Psi(\Omega_U)}\left(\sum_{s=t}^{\infty}V_1\left(z_2\left(s\right),u^*\left(s\right)\right)\right)\\&=V_1\left(z_2\left(t\right),u^*\left(t\right)\right)+J_1^*\left(z_2\left(t+1\right)\right)\end{aligned} \tag{47}$$

where $\Omega_U \in \mathbb{R}^3$ denotes the compact set. The HJB equation corresponding to (47) is

$$H_1\left(z_2,u_a^*,\nabla_{z_2}J_1^*\right)=\left\|z_2\left(t\right)\right\|^2+\left\|u_a^*\left(t\right)\right\|^2+\nabla_{z_2}J_1^*\left(t\right)\Delta z_2\left(t\right)=0 \tag{48}$$

where $\Delta J_1^*(t) = J_1^*(t + 1) - J_1^*(t) = \nabla_{z_2} J_1^*(t) \Delta z_2$. By solving $\nabla_{u_a^*} H_1 = 0$, the optimal vehicle position control u_a^* is given by

$$u_a^*\left(t\right)=-\frac{1}{2}\nabla_{z_2}J_1^*\left(t\right). \tag{49}$$

The gradient term $\nabla_{z_2} J_1^*(t)$ can be decomposed into two components as follows:

$$\nabla_{z_2}J_1^*\left(t\right)=2k_2z_2\left(t\right)+J_1^o\left(z_2\right) \tag{50}$$

where $k_2 > 0$ is the positive parameter. $J_1^o(z_2) = -2k_2 z_2(t) + \nabla_{z_2} J_1^*(t) \in \mathbb{R}^3$. By substituting (50) into (49), the optimal control can be formulated by $u_a^* = -k_2 z_2(t) - 1/2 J_1^o(z_2)$. Since $J_1^o(z_2)$ is a continuous but uncertain term, NN is capable of approximating this mapping on the compact set Ω_U. One way to express the NN approximation is as

$$J_1^o\left(z_2\right)=W_1^{*T}S_1\left(z_2\right)+\omega\left(z_2\right) \tag{51}$$

where $W_1^* \in \mathbb{R}^{p_1\times 3}$ denotes an ideal weight associated with NN size p_1. The basis function vector is $S_1(z_2) \in \mathbb{R}^{p_1}$ and $\omega(z_2) \in \mathbb{R}^3$ represents the approximation error. By substituting (51) into (50) yields

$$\begin{aligned}\nabla_{z_2}J_1^*\left(t\right)&=2k_2z_2\left(t\right)+W_1^{*T}S_1\left(z_2\right)+\omega\left(z_2\right)\\u_a^*&=-k_2z_2\left(t\right)-\frac{1}{2}W_1^{*T}S_1\left(z_2\right)-\frac{1}{2}\omega\left(z_2\right).\end{aligned} \tag{52}$$

Since $W_1^* \in \mathbb{R}^{p_1\times 3}$ is unknown, the optimal vehicle position control in (52) cannot be obtained directly. Therefore, to derive a feasible optimized control law, the following actor and critic are constructed using the RL. The following is the design of the critic used to assess control performance:

$$\nabla_{z_2}\hat{J}_1^*\left(t\right)=2k_2z_2\left(t\right)+\hat{W}_{c1}^T\left(t\right)S_1\left(z_2\right) \tag{53}$$

where $\nabla_{z_2}\hat{J}_1^*(t)$ denotes the estimate of $\nabla_{z_2}J_1^*(t)$, and $\hat{W}_{c1}(t) \in \mathbb{R}^{p_1\times 3}$ represents the critic NN weights, which are updated from the following law:

$$\Delta\hat{W}_{c1}(t) = -\kappa_{c1}S_1(z_2)S_1^T(z_2)\hat{W}_{c1}(t) \tag{54}$$

where $\kappa_{c1} \in \mathbb{R}$ denotes a positive critic gain parameter. The actor responsible for executing control action is described

$$\hat{u}^*(t) = -k_2 z_2(t) - \frac{1}{2}\hat{W}_{a1}^T(t)S_1(z_2) \tag{55}$$

where $\hat{W}_{a1}(t) \in \mathbb{R}^{p_1\times 3}$ denotes the actor NN weights, which are updated based on the following law:

$$\Delta\hat{W}_{a1}(t) = -S_1(z_2)S_1^T(z_2) \times \left(\kappa_{a1}\left(\hat{W}_{a1}(t) - \hat{W}_{c1}(t)\right) + \kappa_{c1}\hat{W}_{c1}(t)\right) \tag{56}$$

where $\kappa_{a1} \in \mathbb{R}$ is a positive gain parameter for the actor.

The design parameters k_2, κ_{c1}, and κ_{a1} are chosen to satisfy $k_2 > 2, \kappa_{a1} > 1/2, \kappa_{a1} > \kappa_{c1} > 1/2\kappa_{a1}$.

Remark 3: The update laws (54) and (56) are derived based on the gradient descent rule to minimize the squared Hamiltonian error H_1^2, ensuring that the weights converge to their optimal values.

In this backstepping, Lyapunov candidate function for overall position control is chosen

$$L_2(t) = L_1(t) + \frac{1}{2}z_2^T(t)z_2(t) + \frac{1}{2}\mathrm{Tr}\left\{\tilde{W}_{a1}^T(t)\tilde{W}_{a1}(t)\right\} + \frac{1}{2}\mathrm{Tr}\left\{\tilde{W}_{c1}^T(t)\tilde{W}_{c1}(t)\right\} \tag{57}$$

where $\tilde{W}_{a1}(t) = \hat{W}_{a1}(t) - W_1^*$ and $\tilde{W}_{c1}(t) = \hat{W}_{c1}(t) - W_1^*$ denote the estimation errors of AC NN weights.

The time derivative of $L_2(t)$ through (43), (54), (55), and (56) can be computed using

$$\Delta L_2(t) = \Delta L_1(t) - \kappa_{c1}\mathrm{Tr}\left\{\tilde{W}_{c1}^T(t)S_1(z_2)S_1^T(z_2)\hat{W}_{c1}(t)\right\} - \mathrm{Tr}\left\{\tilde{W}_{a1}^T(t)S_1(z_2)S_1^T(z_2) \times \left(\kappa_{a1}\left(\hat{W}_{a1}(t) - \hat{W}_{c1}(t)\right) + \kappa_{c1}\hat{W}_{c1}(t)\right)\right\} \tag{58}$$

where $\Delta z_2(t) = \phi z_1(t) + Bu(t) - \Delta\alpha(t)$. By applying Young's inequality and standard trace properties to evaluate the cross terms, we can straightforwardly establish the following upper bound for $\Delta L_2(t)$:

$$\Delta L_2(t) \le -\frac{\kappa_{c1}}{2}\lambda_{S_1}^{\min}\left(\mathrm{Tr}\left\{\tilde{W}_{a1}^T(t)\tilde{W}_{a1}(t)\right\} + \mathrm{Tr}\left\{\tilde{W}_{c1}^T(t)\tilde{W}_{c1}(t)\right\}\right) + C_1(t) - (2k_1 - 1)L_1(t) - \left(k_2 - 1\frac{3}{4}\right)z_2^T(t)z_2(t) \tag{59}$$

where $\lambda_{S_1}^{\min}$ denotes the minimal eigenvalue of $S_1(z_2)S_1^T(z_2)$, $C_1(t) = [(\kappa_{c1} + \kappa_{a1})/2]\|W_1^{*T}S_1(z_2)\|^2 + (1/2)\|\Delta\alpha(t)\|^2 + \phi^2/2$. Since each component of $C_1(t)$ is bounded, one can identify a constant c_1 satisfying $|C_1(t)| \le c_1$. Let $a = \min\{2k_1 - 1, 2k_2 - 3(1/2), \kappa_{c1}\lambda_{S_1}^{\min}\}$. Then (59) can be rewritten as

$$\Delta L_2(t) \le -aL_2(t) + c_1. \tag{60}$$

Therefore, the ultimate bound c_1 and a explicitly capture the adversarial interplay. It analytically absorbs the approximation errors, the sporadic update deviations constrained by DETS, and the energy-bounded deception injections. This rigorously proves the framework's theoretical robustness against both coupled cyber-threats and model uncertainties.

D. Theorem With Proof

Based on Theorem 1, the next theorem specifies the optimal control solution for the vehicle error system.

Theorem 2: Consider the vehicle system, which is modeled by the translational dynamics in (18). If the optimized backstepping vehicle position control is implemented through the virtual control in (44) and (55), which employs the RL update laws in (54) and (56), and the conditions $k_1 > 1/2$ are met by the design parameters. Then, the optimal control can be guaranteed that:

1) all error signals of the optimized vehicle control are global uniform ultimate bounded (GUUB);
2) the vehicle system tracks the predefined reference trajectory with the desired accuracy.

Proof: Select the following Lyapunov candidate of the vehicle control. Based on the preceding result in (60), the time derivative of $L_2(t)$ satisfies:

$$\Delta L_2(t) \le -aL_2(t). \tag{61}$$

By applying lemma of [38] to (61), the following result is obtained

$$L_2(t) \le \frac{1}{a}\left(1 - e^{-at}\right) + L_2(0)e^{-at}. \tag{62}$$

The aforementioned inequality indicates that all error signals are GUUB and that by selecting sufficiently large k_1 and k_2, the tracking errors can be driven to the required accuracy.

IV. Simulation Example

In this section, simulation results are presented to demonstrate the performance of the proposed model-based AC approach for secure tracking control in discrete vehicle systems.

From a systems engineering perspective, the proposed cyber-physical co-design is highly applicable to modern cooperative adaptive cruise control (CACC) systems in intelligent connected vehicles. In practical deployments, vehicular networks suffer from severely limited shared bandwidth, rendering continuous high-frequency data transmission infeasible. Furthermore, practical electronic control units (ECUs) possess limited computational resources, which restricts the deployment of massive deep reinforcement learning networks. The proposed framework directly addresses these engineering bottlenecks. By structurally embedding AC learning into the LQT framework, the algorithm requires significantly fewer learnable parameters, ensuring computational feasibility for onboard ECUs. Simultaneously, the DETS mechanism dynamically throttles transmission rates to preserve bandwidth without sacrificing platoon safety during unpredictable cyber-attacks.

To explicitly illustrate the practical deployability of the proposed framework, the overall cyber-physical system architecture is depicted in Fig. 2. As shown, the architecture integrates the heterogeneous physical vehicle dynamics with

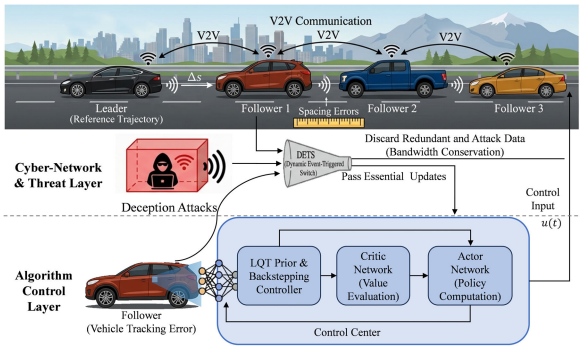
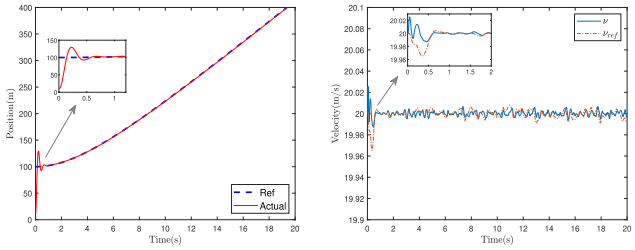


Fig. 2. Cyber-physical co-design of DETS and model-based AC for platoon control.



(a) (b)

Fig. 3. State of the leader and following vehicles. (a) Position. (b) Velocity.

the cyber-network constraints. It demonstrates how the DETS securely throttles the bandwidth, while the model-based AC algorithm, structurally guided by the LQT prior, computes the optimal physical control effort under deception attacks.

A. Baseline Theoretical Verification

Consider the following error system:

$$z^*(t+1) = Az^*(t) + Bu^*(t) + \epsilon(t)$$

where $A = \begin{bmatrix} 1 & 0.08 & 0 \\ 0 & 1 & 0.08 \\ 0 & 0 & 0.84 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 0 \\ 0.01 \end{bmatrix}$. $\epsilon(t)$ denotes unknown noise. The sampling interval is set to 0.08 s. Each training episode lasts 20 s, and the total training process comprises ten episodes. For DETS, the parameters are $h = 0.01$ s, $\mu = 3$, $\zeta = 0.01$, $\varepsilon = 0.4$, $\xi = 0.3$. Set the initial value $\Upsilon_1(0) = 50$. The probabilities $\bar{\beta}$ and $\bar{\rho}$ of deception attacks are selected as 0.3 and 0.1.

Fig. 3(a) illustrates the position tracking performance of the leader and follower vehicles, which shows that the vehicle locations effectively track the desired path. Fig. 3 depicts the evolution of the velocity of the leader and follower vehicles. Fig. 4 presents the trajectories of position error z_p, velocity error z_v, and acceleration error z_a, where $z_p(0) = 0$, $z_v(0) = 0.01$, $z_a(0) = 0.5$.

The event-triggering instants and the response of internal variable under the proposed DETS protocol are shown in Fig. 5. It is evident that redundant data failing to meet condition (10) is not transmitted. By leveraging the dynamic threshold imposed by the internal variable, further conservation of limited network resources can be achieved.

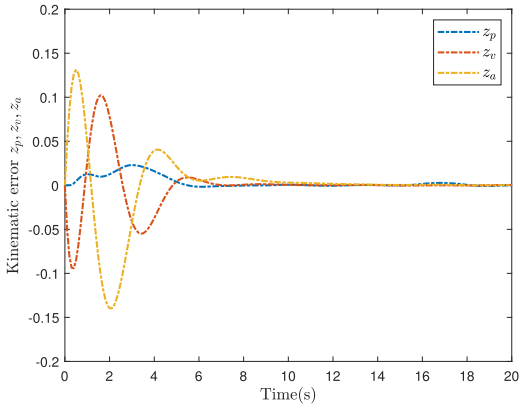


Fig. 4. Position, velocity, and acceleration tracking errors of the vehicle.

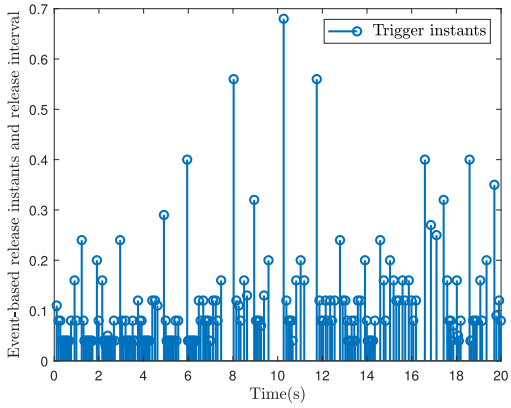


Fig. 5. Evolution of the dynamic event-triggered instants and intervals.

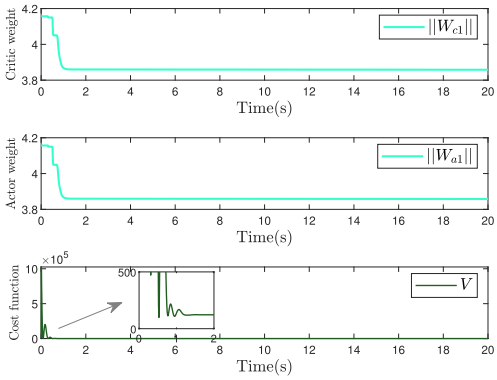
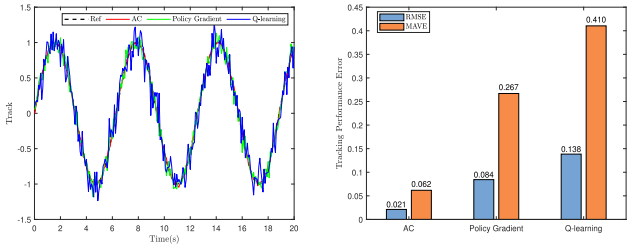


Fig. 6. AC NN weights and cost function $V(x^*(t), r^*(t))$ of position control.

The position reference trajectory is set as $r(t) = [100,\ 0,\ t]^T$ with an initial position $r(0) = [0,\ 0,\ 0.1]^T$, as depicted in Fig. 3(a). Following the procedure in Section III-C, the optimal position control strategy is derived via backstepping approach. For the first stage, the virtual control corresponding to (44) selects $k_1 = 18$. For the second stage, the virtual control corresponding to (55) uses $k_2 = 12$. The learning rates in the RL update rules (54) and (56) select $\kappa_{a1} = 15$ and $\kappa_{c1} = 12$. $z_1(0) = [0.3,\ 0.3,\ 0.3]^T$, $z_2(0) = [0.5,\ 0.5,\ 0.5]^T$, and $\hat{W}_{a1}(0) = \hat{W}_{c1}(0) = [0.6]_{20\times 3}$ establish the starting values.

Fig. 6 illustrates that the AC network weights remain bounded, and shows the evolution of the cost function



(a) (b)

Fig. 7. Tracking performance of vehicle. (a) Different learned control methods. (b) Root mean square and maximum absolute value error.

$V(x^*(t), r^*(t))$. Together, these results collectively verify that the proposed approach successfully achieves the intended control goals.

Fig. 7(a) compares the proposed method against baseline algorithms. To directly address the cyber-physical co-design challenges, we conceptually benchmarked the framework against a standard LQR and a fully model-free AC under identical DETS and attack conditions. While a standard LQR stabilizes the nominal system, it strictly diverges under energy-bounded deception attacks due to a lack of adaptive compensation. Conversely, fully model-free AC approaches suffer from severe gradient starvation under sparse DETS updates, failing to converge efficiently. As shown in Fig. 7, by leveraging the LQT structural prior to overcome these exact limitations, our model-based AC reduces the RMSE by 74.5% compared to generic policy gradient methods, actively mitigating the adversarial cyber-physical interplay.

B. Practical Drive Cycle Validation With Heterogeneous Platoon

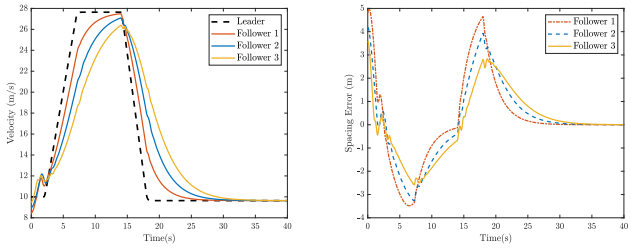
To further validate the practical engineering applicability of the proposed framework, an additional experiment was conducted utilizing real-world vehicle parameters and a standard driving dataset. In practical CACC systems, vehicle parameters are subject to significant uncertainties and heterogeneity, such as varying mass and longitudinal dynamics [39]. To rigorously evaluate the scalability and string stability against these parametric uncertainties, a heterogeneous vehicle platoon comprising one leader and three distinct followers (e.g., sedans, SUVs, and light trucks) is simulated.

As detailed in Table I, the followers possess heterogeneous physical characteristics, including significantly varying masses, aerodynamic profiles, and driveline time lags. This heterogeneity introduces severe unmodeled dynamics, making the tracking task significantly more challenging than standard homogeneous platoon simulations.

Furthermore, dynamic event-triggered mechanisms have been recently emphasized as essential components for networked autonomous vehicles to adapt to varying road conditions while substantially reducing communication frequency and bit rates. To validate the framework's transient robustness and communication conservation under such mechanisms, the leader follows the highly dynamic SFTP-US06 profile. As demonstrated in recent intelligent transportation studies [40], the SFTP-US06 cycle is a representative of aggressive, high-speed, and high-acceleration driving behaviors. By utilizing this rigorous benchmark along with the heterogeneous dynamics, we explicitly evaluate the string stability and the bandwidth-saving performance of the proposed DETS-based AC co-design under severe real-world traffic conditions.

TABLE I
PHYSICAL PARAMETERS OF THE HETEROGENEOUS VEHICLE PLATOON

Parameter	Leader	Follower 1	Follower 2	Follower 3
Vehicle Type	Sedan	SUV	Light Truck	Sedan
Mass m (kg)	1500	2100	3200	1650
Driveline Lag Δs (s)	0.20	0.35	0.50	0.25
Drag Coefficient c_g	0.30	0.35	0.45	0.32
Frontal Area c_a (m^2)	2.2	2.8	3.5	2.4
Rolling Resistance f_{cr}	0.015	0.015	0.020	0.015
Common Platoon Settings				
Sampling Time Δt	0.08 s			
Standstill Distance d_0	5.0 m			
Time Headway h	1.2 s			
Max Acceleration $a_{\max}$	3.0 m/s^2			
Max Deceleration $a_{\min}$	-5.0 m/s^2			



(a) (b)

Fig. 8. Tracking performance and kinematic errors under the practical drive cycle. (a) Velocity states of the leader and followers. (b) Spacing errors of the heterogeneous platoon.

As depicted in Fig. 8(a), despite the severe dynamic fluctuations of the SFTP-US06 cycle and the extreme powertrain heterogeneity, the proposed model-based AC controller rapidly adapts to the aggressive velocity changes. Crucially, Fig. 8 demonstrates that the spacing errors remain strictly bounded and exhibit rigorous string stability (i.e., $\|e_3(t)\|_\infty < \|e_2(t)\|_\infty < \|e_1(t)\|_\infty$). The errors attenuate smoothly along the platoon, successfully preventing the fluctuation amplification typically caused by parametric uncertainties and severe unmodeled dynamics.

Furthermore, this resilient kinematic performance is maintained even under the simultaneous constraints of deception attacks and DETS-induced data sparsity. By dynamically throttling redundant transmissions, the algorithm ensures bandwidth efficiency without sacrificing platoon stability. This explicitly demonstrates the practical robustness, scalability, and systems engineering value of the proposed co-design for real-world mixed-powertrain CACC deployments.

V. CONCLUSION

This article developed a synergistic model-based AC framework for secure vehicle tracking under deception attacks. By

integrating DETS with an LQT-backstepping structure, this co-design explicitly resolved the conflict between communication conservation and learning convergence. Lyapunov analysis rigorously guarantees asymptotic stability, and simulations confirm superior tracking accuracy and faster adaptation compared to baseline model-free RL approaches. However, current stability proofs inherently rely on idealized assumptions, including known attack bounds and linearized nominal dynamics. Future work must address these limitations by analyzing robustness against completely unknown uncertainties and extending validation to high-fidelity physical platforms to bridge the gap between theory and practice.

REFERENCES

[1] Y. Li, K. Li, T. Zheng, X. Hu, H. Feng, and Y. Li, "Evaluating the performance of vehicular platoon control under different network topologies of initial states," *Phys. A, Stat. Mech. Appl.*, vol. 450, pp. 359–368, May 2016.

[2] T. Tank and J.-P.-M. G. Linnartz, "Vehicle-to-vehicle communications for AVCS platooning," *IEEE Trans. Veh. Technol.*, vol. 46, no. 2, pp. 528–536, May 1997.

[3] J. Wang, X. Luo, X. Li, and X. Guan, "Robust predefined-time platoon control of networked vehicles with uncertain disturbances," *Int. J. Syst. Sci.*, vol. 52, no. 15, pp. 3128–3140, Nov. 2021.

[4] Z. Zhou, M. Zhang, D. Navarro-Alarcon, and X. Jing, "Predefined-time output feedback control for active vehicle suspension systems with beneficial couplings, disturbances, and nonlinearities," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 55, no. 11, pp. 7878–7889, Nov. 2025.

[5] J. Hu, P. Bhowmick, F. Arvin, A. Lanzon, and B. Lennox, "Cooperative control of heterogeneous connected vehicle platoons: An adaptive leader-following approach," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 977–984, Apr. 2020.

[6] J. Wei, Y.-J. Liu, H. Chen, and L. Liu, "Fuzzy adaptive control for vehicular platoons with constraints and unknown dead-zone input," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 4, pp. 4403–4412, Apr. 2023.

[7] H. Ma, Z. Wang, D. Wang, D. Liu, P. Yan, and Q. Wei, "Neural-network-based distributed adaptive robust control for a class of nonlinear multiagent systems with time delays and external noises," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 46, no. 6, pp. 750–758, Jun. 2016.

[8] R. Ma, Y. Yu, S. Zhang, and P. Shi, "Event-triggered fault-tolerant tracking control for autonomous ground vehicle via descriptor reduced-order observer," *IEEE Trans. Intell. Transp. Syst.*, vol. 26, no. 7, pp. 10308–10319, Jul. 2025.

[9] G. Wen, D. Yu, and Y. Zhao, "Optimized fuzzy attitude control of quadrotor unmanned aerial vehicle using adaptive reinforcement learning strategy," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 60, no. 5, pp. 6075–6083, Oct. 2024.

[10] S. Dong, Z. Lai, Z.-G. Wu, M. Liu, and G. Chen, "Cooperative quantized event-based fuzzy tracking control of nonlinear autonomous surface vehicles with prescribed performance," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 17594–17605, 2025.

[11] L. Zha, J. Miao, J. Liu, E. Tian, and C. Peng, "Privacy-preserving distributed estimation over sensor networks with multistrategy injection attacks: A chaotic encryption scheme," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 55, no. 7, pp. 4969–4978, Jul. 2025.

[12] N. Zhao, X. Zhao, M. Chen, G. Zong, and H. Zhang, "Resilient distributed event-triggered platooning control of connected vehicles under denial-of-service attacks," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 6, pp. 6191–6202, Jun. 2023.

[13] D. Wang, Z. Yuan, A. Liu, Q. Hu, and J. Qiao, "Static and dynamic event-triggered control for disturbed interconnected systems through adaptive critic designs," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 56, no. 2, pp. 878–892, Feb. 2026.

[14] J. Liu, Z. Liu, Y. Li, E. Tian, and C. Peng, "Privacy-protected formation control for unmanned surface vehicles: A neural predictor-based noncooperative game framework," *IEEE Trans. Veh. Technol.*, pp. 1–10, 2026, doi: 10.1109/TVT.2026.3651837.

[15] G. Wen, C. L. P. Chen, S. S. Ge, H. Yang, and X. Liu, "Optimized adaptive nonlinear tracking control using actor–critic reinforcement learning strategy," *IEEE Trans. Ind. Informat.*, vol. 15, no. 9, pp. 4969–4977, Sep. 2019.

[16] J. Liu, Y. Dong, L. Zha, X. Xie, and E. Tian, "Reinforcement learning-based tracking control for networked control systems with DoS attacks," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 4188–4197, 2024.

[17] H. Bao, Q. Pan, C.-M. Chew, L. Wang, and L. Gao, "An end-to-end framework for energy-efficient cascaded dual-shop collaborative scheduling with mating operations," *IEEE Trans. Cybern.*, vol. 55, no. 10, pp. 4929–4942, Oct. 2025.

[18] K. Zhang, H. Zhang, Y. Zhao, H.-N. Wu, and R. Su, "Synchronization learning scheme of hybrid order adaptive dynamic optimizations for secure communication," *IEEE Trans. Inf. Forensics Security*, vol. 20, pp. 4110–4120, 2025.

[19] K. Zhang, X. Dong, H. Zhang, and R. Su, "Zero-sum optimal control for a cyber–physical system in unreliable communications via secure decentralized policy iteration," *IEEE Internet Things J.*, vol. 12, no. 14, pp. 27339–27349, Jul. 2025.

[20] H. Bao, Q. Pan, C.-M. Chew, L. Wang, and L. Gao, "Energy-efficient distributed heterogeneous hybrid flow-shop scheduling using graph neural network and deep reinforcement learning," *IEEE Trans. Autom. Sci. Eng.*, vol. 23, pp. 3619–3635, 2026.

[21] G. Wen, W. Hao, W. Feng, and K. Gao, "Optimized backstepping tracking control using reinforcement learning for quadrotor unmanned aerial vehicle system," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 52, no. 8, pp. 5004–5015, Aug. 2022.

[22] K. Zhang, N. Liu, X. Xie, and D. Wang, "UKF-based multistep heuristic dynamic programming for optimal event-triggering control of nonlinear systems with asymmetric input constraints," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 55, no. 10, pp. 6986–6997, Oct. 2025.

[23] Z. Gu, X. Huang, X. Sun, X. Xie, and J. H. Park, "Memory-event-triggered tracking control for intelligent vehicle transportation systems: A leader-following approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 5, pp. 4021–4031, May 2024.

[24] J. Liu, E. Ma, L. Zha, and E. Tian, "Distributed energy management for microgrids: When privacy preservation meets multiple mechanisms," *IEEE Trans. Consum. Electron.*, vol. 72, no. 1, pp. 1514–1523, Feb. 2026.

[25] X. Wang, J. Na, B. Niu, X. Zhao, T. Cheng, and W. Zhou, "Event-triggered adaptive bipartite secure consensus asymptotic tracking control for nonlinear MASs subject to DoS attacks," *IEEE Trans. Autom. Sci. Eng.*, vol. 21, no. 3, pp. 3816–3825, Jul. 2024.

[26] D.-W. Zhang and G.-P. Liu, "Secure tracking control of cyber-physical systems against hybrid attacks via FAS terminal sliding-mode predictive control," *IEEE Trans. Cybern.*, vol. 55, no. 11, pp. 5177–5188, Nov. 2025.

[27] G. Zhu, Y. Liu, and J. Qiu, "Dynamic event-triggered adaptive neural practical predefined-time control for uncertain nonlinear systems with unknown control directions," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 55, no. 11, pp. 7629–7638, Nov. 2025.

[28] J. Liu, Z. Zhang, E. Tian, C. Peng, J. Cao, and T. Huang, "Reinforcement learning-based formation control for uncrewed surface vehicles under aperiodic DoS attacks: A stackelberg–Nash game approach," *IEEE Trans. Cybern.*, pp. 1–11, 2026, doi: 10.1109/TCYB.2026.3674952.

[29] G. Wen, L. Xu, and B. Li, "Optimized backstepping tracking control using reinforcement learning for a class of stochastic nonlinear strict-feedback systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 3, pp. 1291–1303, Mar. 2023.

[30] S. Liu, B. Jiang, Z. Mao, Y. Zhang, and J. Huang, "Adaptive fractional-order fault-tolerant coordinated tracking control of heterogeneous multiagent systems against multiple faults under deception attacks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 61, no. 2, pp. 1860–1870, Apr. 2025.

[31] Z. Gu, T. Yin, and Z. Ding, "Path tracking control of autonomous vehicles subject to deception attacks via a learning-based event-triggered mechanism," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 12, pp. 5644–5653, Dec. 2021.

[32] S. Yan, Z. Gu, J. H. Park, and M. Shen, "Fusion-based event-triggered H_∞ state estimation of networked autonomous surface vehicles with measurement outliers and cyber-attacks," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 7, pp. 7541–7551, Jul. 2024.

[33] L. Zha, R. Liao, J. Liu, X. Xie, E. Tian, and J. Cao, "Outlier-resistant distributed filtering over sensor networks under dynamic event-triggered schemes and DoS attacks," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 1152–1162, 2025.

[34] J. Wang, X. Li, J. H. Park, and G. Guo, "Distributed MPC-based string stable platoon control of networked vehicle systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 3, pp. 3078–3090, Mar. 2023.

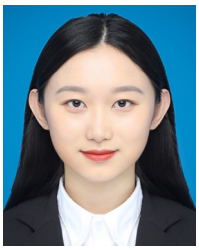
[35] Y. Liu, C. Pan, H. Gao, and G. Guo, "Cooperative spacing control for interconnected vehicle systems with input delays," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 10692–10704, Dec. 2017.
[36] H. Zhang, Q. Wei, and Y. Luo, "A novel infinite-time optimal tracking control scheme for a class of discrete-time nonlinear systems via the greedy HDP iteration algorithm," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 38, no. 4, pp. 937–942, Aug. 2008.
[37] X. Zhao, B. Tao, L. Qian, and H. Ding, "Model-based actor-critic learning for optimal tracking control of robots with input saturation," *IEEE Trans. Ind. Electron.*, vol. 68, no. 6, pp. 5046–5056, Jun. 2021.
[38] G.-X. Wen, C. L. P. Chen, Y.-J. Liu, and Z. Liu, "Neural-network-based adaptive leader-following consensus control for second-order non-linear multi-agent systems," *IET Control Theory Appl.*, vol. 9, no. 13, pp. 1927–1934, Aug. 2015.
[39] S. Cheng, L. Li, C.-Z. Liu, X. Wu, S.-N. Fang, and J.-W. Yong, "Robust LMI-based H-infinite controller integrating AFS and DYC of autonomous vehicles with parametric uncertainties," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 51, no. 11, pp. 6901–6910, Nov. 2021.
[40] J. Lan, D. Zhao, and D. Tian, "Data-driven robust predictive control for mixed vehicle platoons using noisy measurement," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 6, pp. 6586–6596, Jun. 2023.

Jinliang Liu (Senior Member, IEEE) received the Ph.D. degree in control theory and control engineering from the School of Information Science and Technology, Donghua University, Shanghai, China, in 2011.

He was a Post-Doctoral Research Associate with the School of Automation, Southeast University, Nanjing, China, from December 2013 to June 2016. He was a Visiting Researcher/Scholar with the Department of Mechanical Engineering, The University of Hong Kong, Hong Kong, from October 2016 to October 2017. He was a Visiting Scholar with the Department of Electrical Engineering, Yeungnam University, Gyeongsan, South Korea, from November 2017 to January 2018. From June 2011 to May 2023, he was an Associate Professor and then a Professor with Nanjing University of Finance and Economics, Nanjing. In June 2023, he joined Nanjing University of Information Science and Technology, Nanjing, where he is currently a Professor with the School of Computer Science. His research interests include cyber-physical systems, unmanned systems, privacy preservation, and game theory.

Bin Lv received the B.S. degree in software engineering from Wuxi Taihu University, Wuxi, China, in 2024, where he is currently pursuing the M.S. degree in computer technology with the College of Computer, Nanjing University of Information Science and Technology, Nanjing, China.

His research interests include Internet of Vehicles, reinforcement learning, and embedded development.



Meng Yang received the Ph.D. degree in control science and engineering from Southeast University, Nanjing, China, in 2025.

Her is currently a Lecturer with the College of Command and Control Engineering, Army Engineering University of PLA, Nanjing.

Engang Tian (Senior Member, IEEE) received the B.S. degree in mathematics from Shandong Normal University, Jinan, China, in 2002, the M.Sc. degree in operations research and cybernetics from Nanjing Normal University, Nanjing, China, in 2005, and the Ph.D. degree in control theory and control engineering from Donghua University, Shanghai, China, in 2008.

From 2015 to 2016, he was a Visiting Scholar with the Department of Information Systems and Computing, Brunel University London, Uxbridge, U.K. From 2008 to 2018, he was an Associate Professor and then a Professor with the School of Electrical and Automation Engineering, Nanjing Normal University. In 2018, he was appointed as an Eastern Scholar by the Municipal Commission of Education, Shanghai, and joined the University of Shanghai for Science and Technology, Shanghai, where he is currently a Professor with the School of Optical-Electrical and Computer Engineering.

Chen Peng (Senior Member, IEEE) received the B.S. and M.S. degrees in coal preparation, and the Ph.D. degree in control theory and control engineering from Chinese University of Mining Technology, Xuzhou, China, in 1996, 1999, and 2002, respectively.

From November 2004 to January 2005, he was a Research Associate with The University of Hong Kong, Hong Kong. From September 2010 to August 2012, he was a Post-Doctoral Research Fellow with Central Queensland University, Rockhampton, QLD, Australia. He is currently a Professor with the School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, China. His current research interests include networked control systems, multiagent systems, power systems, and interconnected systems.

Dr. Peng is an Associate Editor for a number of international journals, including IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, *Information Sciences*, and TRANSACTIONS OF THE INSTITUTE OF MEASUREMENT AND CONTROL. He was named a Highly Cited Researcher in 2020, 2021, and 2022 by Clarivate Analytics.

Tingwen Huang (Fellow, IEEE) received the B.S. degree in mathematics from Southwest University, Chongqing, China, in 1990, the M.S. degree in mathematics from Sichuan University, Chengdu, China, in 1993, and the Ph.D. degree in mathematics from Texas A&M University, College Station, TX, USA, in 2002.

He has been a Professor with the Faculty of Computer Science and Control Engineering, Shenzhen University of Advanced Technology, Shenzhen, China, since 2024. He has expertise in optimization and control, machine learning, privacy protection, and neural networks.

Prof. Huang currently serves as the Editor-in-Chief for *IEEE Systems, Man, and Cybernetics Magazine* and an Editorial Board Member for IEEE TRANSACTIONS ON INDUSTRIAL CYBER-PHYSICAL SYSTEMS, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON FUZZY SYSTEMS, IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE, IEEE TRANSACTIONS ON ARTIFICIAL INTELLIGENCE, and *Neural Networks*.