

基于 ADP 的非对称约束系统动态事件触发方法

薛珊^{1,2,3} 赵宁^{1,3} 张卫东^{1,3,4}

摘要 本文提出一种基于自适应动态规划的动态事件触发方法 (DEM), 用于解决具有状态与控制双重非对称约束的非线性连续时间系统最优控制问题. 首先, 利用非线性映射函数将非对称约束系统的控制问题转化为无约束形式. 然后, 设计一种静态事件触发方法 (SEM), 其中触发条件仅与当前状态相关. 进一步, 开发一种依赖额外内部动态变量的 DEM, 其触发条件也与系统历史信息相关. 事实上, DEM 是 SEM 的进阶方法. 理论分析证实 DEM 在确保系统性能的情况下, 能够进一步节省计算和网络资源. 最后, 介绍基于神经网络的实现方法. 在无人水面艇仿真实验环境下, 该方法的有效性得到了验证.

关键词 自适应动态规划; 自适应评价设计; 神经网络; 事件触发方法; 动态事件触发方法

引用格式 薛珊, 赵宁, 张卫东. 基于 ADP 的非对称约束系统动态事件触发方法. 自动化学报, 2026, 52(6): 1209–1220

DOI 10.16383/j.aas.c250544 **CSTR** 32138.14.j.aas.c250544

ADP-based Dynamic Event-triggering Method for Asymmetric Constrained Systems

XUE Shan^{1,2,3} ZHAO Ning^{1,3} ZHANG Wei-Dong^{1,3,4}

Abstract In this paper, an adaptive dynamic programming-based dynamic event-triggering method (DEM) is developed to solve the optimal control problem of nonlinear continuous-time systems with asymmetric constraints for both state and control. First, a nonlinear mapping function is used to transform the control problem of asymmetric constrained systems into an unconstrained form. Then, a static event-triggering method (SEM) is designed, where triggering conditions are only associated with the current state. Based on the SEM, a DEM that relies on an additional internal dynamic variable is developed, whose triggering condition is also related to the system historical information. In fact, the DEM is an advanced method of the SEM. Theoretical analysis proves that the DEM can further save computational and network resources while ensuring system performance. Finally, the neural network-based implementation is presented. The effectiveness of this method has been verified in the simulation experiment environment of the unmanned surface vehicle.

Keywords adaptive dynamic programming; adaptive critic designs; neural networks; event-triggering method; dynamic event-triggering method

Citation Xue Shan, Zhao Ning, Zhang Wei-Dong. ADP-based dynamic event-triggering method for asymmetric constrained systems. *Acta Automatica Sinica*, 2026, 52(6): 1209–1220

收稿日期 2025-10-14 录用日期 2025-12-19

Manuscript received October 14, 2025; accepted December 19, 2025

国家自然科学基金 (U2141234, U24A20260, 62403173), 海南省自然科学基金 (725RC724), 海南省高等教育科研资助项目 (Hnky2025ZD-2), 中国国家科技重大专项 (2022ZD0119900), 海南省科技专项 (ZDYF2024GXJS003), 海南大学科研基金 (XJ2400000440) 资助

Supported by National Natural Science Foundation of China (U2141234, U24A20260, 62403173), Hainan Provincial Natural Science Foundation (725RC724), Hainan Higher Education Scientific Research Funding Project (Hnky2025ZD-2), National Science and Technology Major Project of China (2022ZD0119900), Hainan Provincial Science and Technology Special Fund (ZDYF2024 GXJS003), and Scientific Research Fund of Hainan University (XJ2400000440)

本文责任编辑 罗彪

Recommended by Associate Editor LUO Biao

1. 海南大学信息与通信工程学院 海口 570228 2. 安徽大学人工智能学院 合肥 230601 3. 海洋智能系统教育部工程研究中心海口 570228 4. 上海交通大学自动化与智能感知学院 上海 200240

1. School of Information and Communication Engineering, Hainan University, Haikou 570228 2. School of Artificial Intelligence, Anhui University, Hefei 230601 3. Engineering Research Center of Marine Intelligent Systems, Ministry of Educa-

近年来, 强化学习 (reinforcement learning, RL) 受到广泛关注^[1-2], 这种人工智能方法强调在特定环境中采取行动以实现最大效益. 在控制领域, 受 RL 思想启发, 提出一种用于求解最优控制问题的近似方法——自适应动态规划 (adaptive dynamic programming, ADP) 或自适应评价设计 (adaptive critic design, ACD)^[3-4]. ADP 是一种类似于人类大脑智能的结构, 通常包含三个组成部分: 评价网络、执行网络和模型网络. 近似动态规划、神经动态规划和自适应评价学习均与 ADP 或 ACD 具有相似的学习模式^[5-6]. 启发式动态规划 (heuristic dynamic programming, HDP) 是 ADP 中最早且应用最广泛的算法^[7-9], 双启发式动态规划 (dual heuristic dynamic programming, DHP)、全局化 DHP、动作依赖型 HDP、动作依赖型 DHP、动作依赖型全

tion, Haikou 570228 4. School of Automation and Intelligent Sensing, Shanghai Jiao Tong University, Shanghai 200240

局化 DHP 及类似算法也引起广泛关注^[10-12]. 在处理实际控制问题时, 系统的输入需满足用户定义的边界条件, 该条件的设定考虑了设备的物理特性和安全性. 为解决污水处理厂溶解氧浓度与硝酸盐氮浓度的控制问题, 文献 [13] 提出一种基于数据驱动的 DHP 跟踪控制算法. 在文献 [14] 中, 针对未知约束离散时间非线性系统, 采用全局化 ADP 技术获得最优控制器. 针对四旋翼图像视觉伺服控制中的角度约束, 文献 [15] 解决了横向速度系统的最优鲁棒保证成本控制问题. 在微电网电池能源管理系统中, 为解决功率约束问题, 文献 [16] 设计一种非二次成本函数. 然而, 大多数研究往往只关注控制约束问题, 却忽视同样关键的状态约束问题. 实际上, 确保系统始终保持在用户定义的边界内, 也是使系统符合安全标准的重要方面^[17-18]. 根据小增益定理, 文献 [19] 提出一种适用于连续时间系统的鲁棒控制策略, 并阐明其渐近稳定性. 基于文献 [19], 一种非仿射多智能体系统的共识控制在文献 [20] 中进行研究. 在文献 [21] 中, 通过为带约束的分段仿射系统设计惩罚函数, 将非凸多面体联合状态约束和输入约束融合到 ADP 中. 现有的输入与输出约束研究大多基于时间触发方法 (time-triggering method, TTM), TTM 根据采样间隔定期更新控制策略. 然而, 当计算资源和网络带宽受限时, 这种方法的效果并不理想.

相较于 TTM, 事件触发方法 (event-triggering method, ETM) 仅在系统状态变化满足预设触发条件时才激活信息交互与控制更新, 不仅能最大程度减少非必要的数据传输与计算开销, 显著节省网络带宽, 而且能精准捕捉系统关键动态变化, 避免周期性触发带来的响应滞后. 在保持所需系统性能的情况下, 减少了不必要的信息传输. 在文献 [22] 中, 针对资源有限的微处理器, 提出一种基于 ETM 的新型实时调度算法, 该算法能够增强嵌入式设备的功能. 在文献 [23] 中, 给出一种事件控制方法的概述, 该方法在需要时实施相应操作. 在 ADP 和 RL 领域, 已有关于 ETM 的研究. 在文献 [24] 中, ETM 采用非周期性设计, 当状态偏差超过设定阈值时, 会触发传感器进行采样和执行控制动作. 针对离散型 Takagi-Sugeno 模糊系统, 文献 [25] 提出一种事件触发实时调度方法. 在事件触发机制下系统的稳定性、最优性、收敛性和鲁棒性已在文献 [26] 中进行分析. 基于 ETM 的控制策略被应用于多智能体系统的协同控制, 相关研究参考文献 [27-29]. 最近, 文献 [30] 提出一种基于事件的有限时间控制方法, 用于解决非线性多智能体系统的跟踪控制问题. 文献 [31] 针对切换网络研究了未知

多智能体系统的事件触发协同最优控制器. 然而, 上述工作中提出的 ETM 规则仅与当前状态相关, 这被称为静态事件触发方法 (static event-triggering method, SEM)^[32-33].

动态事件触发方法 (dynamic event-triggering method, DEM) 源自文献 [34], 该文献证明了系统的渐近稳定性. “动态”一词与内部动态变量密切相关. 事实上, 该变量通常具有非单调特性, 它是经过滤波的触发信号变体, 包含系统的历史信息^[35-37]. 在文献 [38] 中, 明确阐述静态系统与动态系统之间的区别: 后者能够基于可用的系统知识, 随时间动态调节自身. DEM 确保所需的闭环性能, 有效降低了资源传输的使用率^[34-37], 其存在一个严格的正下界限定执行间隔时间, 有效避免了芝诺现象的发生. 对于离散时间系统, 基于 DEM 的容错最优控制方法在文献 [39] 中进行研究. 考虑到扰动的存在, 文献 [40] 给出一种基于 DEM 的方法. 在文献 [41] 中, 针对多智能体系统控制问题, 采用一种基于 DEM 的分布式控制方法. 关于多智能体系统的进一步研究可参考文献 [42]. 文献 [43] 基于两个网络 (执行网络和评价网络) 的框架提出一种 DEM. 随后, 该方法被应用于工业系统^[44], 包括关节机械臂与电力系统, 实验结果表明, 该方法在实现预期效益的情况下, 能够降低信息传输速率. 文献 [45] 将 DEM 应用于复杂互联系统, 进一步减轻了计算负担. 在文献 [46] 中, 设计一种针对未知非线性零和博弈最优跟踪的事件触发控制器, 提高了目标轨迹对应稳定控制的精度. 在文献 [47] 中, 提出一种基于零和微分博弈的最优容错控制方法, 该方法被应用于通信链路故障的航天器编队控制系统. 研究发现, 在 DEM 领域中, 对于状态变量与控制变量均受非对称边界约束的连续时间非线性系统, 其最优控制问题尚未受到充分研究, 这一现状启发了我们的工作. 为解决具有非对称状态边界与控制边界的连续时间非线性系统最优控制问题, 本文采用非线性映射函数将原始约束问题转化为无约束形式. 首先, 提出静态事件触发条件; 然后, 基于引入的内部动力学设计动态触发条件, 理论分析动态变量的非负性; 最后, 采用一种基于神经网络 (neural network, NN) 的实现方法, 该方法使用基于积分强化学习的标准化梯度下降对权重进行学习, 降低了对漂移动力学知识的依赖性. 与文献 [9, 21, 24, 31, 43, 46-47] 相比, 本研究的主要贡献体现在以下三个方面:

- 1) 对于状态与控制均受非对称边界约束的最优控制问题, 本文采用非线性映射函数将原始受约束控制问题转化为无约束控制问题. 通过结合积分强化学习, 该算法显著降低对模型的依赖性, 提升

了方法的实际应用价值.

2) 为确保在预期性能的情况下缩短控制器执行时间, 本文提出两种 ETM. 首先提出静态触发条件, 进一步提出其动态形式. 通过无人水面艇 (unmanned surface vehicle, USV) 仿真实验系统验证了 ETM 方法的有效性.

3) 通过对比非映射方法^[43]、TTM^[9]、SEM^[24] 及 DEM, 验证了本文提出的映射函数方法能有效地将轨迹限制在预定义边界内. DEM 具有最长的平均执行间隔时间和最少的采样点数, 减轻了网络传输负担.

本文其余部分组织结构如下: 第 1 节对问题进行描述; 第 2 节设计 SEM 与 DEM, 对这些方法进行理论分析; 第 3 节描述基于神经网络的评价器实现方法: 第 3.1 节给出最优性能指标的近似方法, 第 3.2 节给出控制策略的近似方法, 第 3.3 节进行理论分析; 第 4 节展示仿真结果; 第 5 节阐述结论, 并对未来进行展望. 本文符号命名法详见表 1.

表 1 命名法
Table 1 Nomenclature

| 符号 | 含义 |
|---------------------------|--------------------|
| \mathbf{N} | 正整数集合 |
| \mathbf{R} | 实数集合 |
| \mathbf{R}^m | m 维欧氏空间 |
| $\mathbf{R}^{m \times n}$ | $m \times n$ 维矩阵空间 |
| \mathbf{T} | 转置 |
| $\nabla J(s)$ | 梯度算子 |
| $\ \cdot\ $ | 2 范数 |

1 问题建模

考虑以下类型的系统:

$$\dot{s}(t) = f(s(t)) + g(s(t))u(t) \quad (1)$$

其中, $u(t) \in \mathcal{A} \subset \mathbf{R}^n$ 表示约束控制策略, \mathcal{A} 为控制策略集合; $s(t) \in \mathcal{S} \subset \mathbf{R}^m$ 表示约束状态向量, \mathcal{S} 为状态集合; $f(s(t)) \in \mathbf{R}^m$ 表示漂移动力学; $g(s(t)) \in \mathbf{R}^{m \times n}$ 表示输入动力学, n 和 m 表示向量的维数. 对于 $u(t)$ 和 $s(t)$ 的各个分量, 它们分别满足

$$u_i \in (\mu_{ai}, \mu_{bi}) \quad (2)$$

$$s_i \in (\vartheta_{ai}, \vartheta_{bi}) \quad (3)$$

其中, $i \in \mathbf{N}$, μ_{ai} 、 μ_{bi} 、 ϑ_{ai} 和 ϑ_{bi} 是代表边界的常数.

假设 1. 假设系统 (1) 是可控的, 即存在一个作用于 $\Omega \subset \mathbf{R}^n$ 的控制策略能够使系统稳定. 此外, $f(s) + g(s)u$ 在 Ω 上具有利普希茨连续性.

在假设 1 中, 系统可控性是非线性系统稳定控制研究的核心前提, 若系统不满足可控性, 则无法通过设计控制策略调整状态轨迹以实现稳定. $f(s) + g(s)u$ 在 Ω 上具有利普希茨连续性, 该条件可确保系统状态解的存在唯一性, 是相关稳定性理论有效应用的必要前提.

为解决约束问题, 引入了一个非线性映射函数:

$$\mathcal{M}(z, \underline{z}, \bar{z}) = \ln \frac{\bar{z}(z - \underline{z})}{z(\bar{z} - z)} \quad (4)$$

其中, \underline{z} 和 \bar{z} 分别表示 z 的下界和上界. 对于 $\mathcal{M}(z, \underline{z}, \bar{z})$ 的反函数 $\mathcal{N}(\xi, \underline{z}, \bar{z})$ (ξ 为无约束变量), 其满足

$$\mathcal{N}(\xi, \underline{z}, \bar{z}) = \underline{z}\bar{z} \frac{e^{\frac{\xi}{2}} - e^{-\frac{\xi}{2}}}{z e^{\frac{\xi}{2}} - \bar{z} e^{-\frac{\xi}{2}}} \quad (5)$$

对于具有非对称约束的控制策略, 效用函数 $\mathcal{Y}_u(u)$ 可以表述为

$$\mathcal{Y}_u(u) = 2 \sum_{i=1}^m \int_0^{u_i} \mathcal{M}(\tau_i, \mu_{ai}, \mu_{bi}) r_i d\tau_i \quad (6)$$

其中, r_i 为常数. 进一步, 式 (1) 的性能指标可以表述为

$$\mathcal{V}(s(t)) = \int_t^\infty [\mathcal{Y}_s(s(v)) + \mathcal{Y}_u(u(v))] dv \quad (7)$$

其中, \mathcal{Y}_s 为状态效用函数.

本文的目标是设计控制策略 $u^*(s)$ 以最小化性能指标 (7), 确保系统 (1) 的稳定性, 同时满足状态与控制策略的约束条件.

对函数 $\mathcal{N}(\xi, \underline{z}, \bar{z})$ 求导, 可以得到

$$\frac{d}{d\xi} \mathcal{N}(\xi, \underline{z}, \bar{z}) = \frac{\bar{z}\underline{z}^2 - \underline{z}\bar{z}^2}{z^2 e^\xi - 2z\bar{z} + \bar{z}^2 e^{-\xi}} \quad (8)$$

根据式 (5), 对于具有非对称约束的状态 s_i , 其满足

$$s_i = \mathcal{N}(x_i, \vartheta_{ai}, \vartheta_{bi}) = \vartheta_{ai}\vartheta_{bi} \frac{e^{\frac{x_i}{2}} - e^{-\frac{x_i}{2}}}{\vartheta_{ai} e^{\frac{x_i}{2}} - \vartheta_{bi} e^{-\frac{x_i}{2}}} \quad (9)$$

其中, x_i 为无约束状态. 于是, 我们得到

$$\dot{s}_i = \frac{ds_i}{dx_i} \frac{dx_i}{dt} \quad (10)$$

根据式 (8) 和式 (10), \dot{x}_i 满足

$$\begin{aligned} \dot{x}_i &= \frac{\dot{s}_i dx_i}{ds_i} = \\ &= \frac{\vartheta_{ai}^2 e^{x_i} - 2\vartheta_{ai}\vartheta_{bi} + \vartheta_{bi}^2 e^{-x_i}}{\vartheta_{bi}\vartheta_{ai}^2 - \vartheta_{ai}\vartheta_{bi}^2} \times \\ &= [f_i(\mathcal{N}(x, \vartheta_a, \vartheta_b)) + g_i(\mathcal{N}(x, \vartheta_a, \vartheta_b))u(t)] \end{aligned} \quad (11)$$

因此, 约束系统 (1) 可按以下方式变换为无约

束形式, 表示为

$$\dot{x} = F(x) + G(x)u(t) \quad (12)$$

其中,

$$x = [x_1, \dots, x_i, \dots, x_m]^T$$

$$F(x) = [F_1(x), \dots, F_i(x), \dots, F_m(x)]^T$$

$$F_i(x) = \frac{\vartheta_{ai}^2 e^{x_i} - 2\vartheta_{ai}\vartheta_{bi} + \vartheta_{bi}^2 e^{-x_i}}{\vartheta_{bi}\vartheta_{ai}^2 - \vartheta_{ai}\vartheta_{bi}^2} f_i(\mathcal{N}(x, \vartheta_a, \vartheta_b))$$

$$G(x) = [G_1(x), \dots, G_i(x), \dots, G_m(x)]^T$$

$$G_i(x) = \frac{\vartheta_{ai}^2 e^{x_i} - 2\vartheta_{ai}\vartheta_{bi} + \vartheta_{bi}^2 e^{-x_i}}{\vartheta_{bi}\vartheta_{ai}^2 - \vartheta_{ai}\vartheta_{bi}^2} g_i(\mathcal{N}(x, \vartheta_a, \vartheta_b))$$

在上式中,

$$x_i = \mathcal{M}(s_i, \vartheta_{ai}, \vartheta_{bi}) \quad (13)$$

是单调递增的, 并且满足以下条件

$$\begin{cases} \lim_{s_i \rightarrow \vartheta_{ai}} x_i = -\infty \\ \lim_{s_i \rightarrow \vartheta_{bi}} x_i = +\infty \\ \lim_{s_i \rightarrow 0} x_i = 0 \end{cases}$$

通过变换, 状态 s_i 满足约束条件 (3), 其稳定性可由 x_i 的稳定性推导 (定理 1) 得出. 因此, 式 (12) 的最优控制问题在于设计一个最优控制策略 $u^*(s)$, 使下列性能指标最小化:

$$\mathcal{V}(x(t)) = \int_t^\infty [\mathcal{Y}_x(x(v)) + \mathcal{Y}_u(u(v))] dv \quad (14)$$

其中, $\mathcal{Y}_x(x)$ 可以表示为

$$\mathcal{Y}_x(x) = x^T Q x$$

其中, Q 为一个对称半正定矩阵.

在式 (7) 中, 考虑了对控制变量和状态变量的非对称约束. 通过引入适当的映射函数, 可以有效处理这些约束. 在式 (14) 中, 变换后的无约束条件被表示为状态变量的二次型, 这使得控制策略能够利用现有方法求解. 对式 (14) 求解关于式 (12) 的导数

$$\nabla \mathcal{V}^T(x)[F(x) + G(x)u(t)] + \mathcal{Y}_x(x) + \mathcal{Y}_u(u) = 0 \quad (15)$$

可以将哈密顿函数描述为

$$H(\nabla \mathcal{V}(x), x, u) = \mathcal{Y}_x(x) + \mathcal{Y}_u(u) +$$

$$\nabla \mathcal{V}^T(x)[F(x) + G(x)u(t)] \quad (16)$$

基于最优性原理, 可以得到以下哈密顿-雅可比-贝尔曼 (Hamilton-Jacobi-Bellman, HJB) 方程:

$$\min_u H(\nabla \mathcal{V}^*(x), x, u) = 0 \quad (17)$$

其中, $\mathcal{V}^*(x)$ 表示最优性能指标, 且满足

$$\mathcal{V}^*(x(t)) = \min_u \int_t^\infty [\mathcal{Y}_x(x(v)) + \mathcal{Y}_u(u(v))] dv \quad (18)$$

利用平稳性条件, 可以得到

$$\frac{\partial H(\nabla \mathcal{V}^*(x), x, u)}{\partial u} = 0$$

进一步, 基于 TTM 的控制策略可以表示为

$$u_i^*(x) = -\mathcal{N}\left(\frac{1}{2}r_i^{-1}G_i^T(x)\nabla\mathcal{V}^*(x), \mu_{ai}, \mu_{bi}\right) \quad (19)$$

设 $v_i^*(x)$ 为

$$v_i^*(x) = \frac{1}{2}r_i^{-1}G_i^T(x)\nabla\mathcal{V}^*(x) \quad (20)$$

由上式可知, 基于 TTM 的控制策略可以重新表示为

$$u_i^*(x) = -\mathcal{N}(v_i^*(x), \mu_{ai}, \mu_{bi}) \quad (21)$$

对于区间 T , (17) 的解可以转化为方程

$$\mathcal{V}^*(x(t+T)) - \mathcal{V}^*(x) + \int_t^{t+T} [\mathcal{Y}_x(x(v)) + \mathcal{Y}_u(u^*(v))] dv = 0 \quad (22)$$

基于贝尔曼最优性原理, 本文将最优控制策略的求解转化为求解一个 HJB 方程, 然而, 直接获得其解析解极其困难. 而 ADP 是一种高效近似求解最优控制问题的方法, 其在连续时间系统中的收敛性、稳定性和最优性已在文献 [5] 中得到验证. 因此, 本文采用 ADP 方法来求解非对称约束问题的最优控制策略.

2 事件触发方法

2.1 静态事件触发方法

对于 SEM, 将采样序列定义为 $\{l_j\}_{j=0}^\infty$, 其中, $l_0 = 0$, $l_j < l_{j+1}$, $j \in \mathbf{N}$. 基于 SEM 的状态向量和控制策略被分别描述为 $x(l_j)$ 和 $u(x(l_j))$. 基于 SEM 的执行间隔时间 $\Delta_j l$ 可以表示为

$$\Delta_j l = l_{j+1} - l_j \quad (23)$$

状态误差可以表述为

$$\Delta_j x(t) = x(l_j) - x(t), t \in [l_j, l_{j+1}) \quad (24)$$

在本文中, 静态事件触发条件被定义为

$$\|\Delta_j x(t)\| \leq \frac{\sqrt{2(1-\lambda_1)Q}}{\mathcal{L}_u} \|x\| \quad (25)$$

其中, Q 是最小特征值, $\mathcal{L}_u > 0$ 和 $0 < \lambda_1 < 1$ 是正常数. 为事件触发条件选择合适的参数是非常困难的, 目前尚无标准化方法或参考指南. 通常, 这些参

数是基于经验和试错法来确定的. 当静态事件条件 (25) 被触发时, 将根据下式来重新生成控制策略

$$u_i^*(x(l_j)) = -\mathcal{N}(v_i^*(x(l_j)), \mu_{ai}, \mu_{bi}) \quad (26)$$

其中, $v_i^*(x(l_j))$ 被定义为

$$v_i^*(x(l_j)) = \frac{1}{2}r_i^{-1}G_i^T(x(l_j))\nabla\mathcal{V}^*(x(l_j)) \quad (27)$$

假设 2. $u^*(x)$ 和 $G(x)$ 分别是具有系数 \mathcal{L}_u 和 \mathcal{L}_G 的利普希茨连续函数. $G(x)$ 有一个由常数 \bar{G} 确定的正上界.

在假设 2 中, $G(x)$ 有一个由常数 \bar{G} 确定的正上界, 这是对实际系统固有物理约束的数学抽象, 该假设避免了 $G(x)$ 无界引发的控制增益饱和、系统性能发散等问题, 为稳定性分析和参数优化设计提供了明确边界.

定理 1. 令假设 1 和假设 2 成立, 考虑基于 SEM 的控制策略 $u^*(x(l_j))$, 系统 (12) 在 $u^*(x(l_j))$ 作用下可以表示为

$$\dot{x} = F(x) + G(x)u^*(x(l_j)) \quad (28)$$

然后, 静态事件触发条件 (25) 确保了系统状态的一致最终有界性 (uniform ultimate boundedness, UUB).

定理 1 的证明参考文献 [43]. 为避免重复, 此处省略证明.

2.2 动态事件触发方法

需要注意的是, 静态触发条件 (25) 仅考虑了当前系统信息. 相比 SEM, DEM 可以通过充分利用历史状态信息进一步延长平均触发间隔. 现引入以下内部动态机制

$$\dot{\theta}_1(t) = -\rho_1\theta_1(t) + \pi_1(t) \quad (29)$$

其中, $\theta_1(t)$ 为内部动态变量, $\theta_1(0) \geq 0$ 和 $\rho_1 > 0$ 为常数, $\pi_1(t)$ 可以定义为:

$$\pi_1(t) := (1 - \lambda_1)x^T Q x - \frac{1}{2}\mathcal{L}_u^2 \|\Delta_j x(t)\|^2 \quad (30)$$

内部动力学 (29) 实际是一个一阶滤波器, 内部动态变量 θ_1 是 π_1 的滤波值, ρ_1 是滤波系数.

DEM 中的采样序列定义为 $\{l_d\}_{d=0}^\infty$, 其中, $l_0 = 0$, $l_d < l_{d+1}$, $d \in \mathbf{N}$. 基于 DEM 的执行间隔时间

$$\Delta_d l = l_{d+1} - l_d \quad (31)$$

状态误差定义为 $\Delta_d x(t) = x(l_d) - x(t)$, $t \in [l_d, l_{d+1})$.

l_{d+1} 是根据以下动态事件触发条件生成

$$\theta_1(t) + \gamma_1 \pi_1(t) \geq 0 \quad (32)$$

其中, $\gamma_1 > 0$ 为常数. 通过观察式 (32) 可知, 当

$\gamma_1 \rightarrow \infty$ 时, 动态条件趋向于静态条件. 当条件 (32) 被触发时, 控制策略通过下列方程进行更新:

$$u_i^*(x(l_d)) = -\mathcal{N}(v_i^*(x(l_d)), \mu_{ai}, \mu_{bi}) \quad (33)$$

$$v_i^*(x(l_d)) = \frac{1}{2}r_i^{-1}G_i^T(x(l_d))\nabla\mathcal{V}^*(x(l_d)) \quad (34)$$

定理 2. 对于式 (29) 中定义的内部动态变量 $\theta_1(t)$ 以及包含 $u^*(x(l_d))$ 的系统 (12), 满足以下条件:

- 1) $\theta_1(t)$ 为非负值;
- 2) 动态条件 (32) 保证系统状态为 UUB;
- 3) 基于 DEM 的执行间隔时间不小于静态时间, 即 $\Delta_j l \leq \Delta_d l$.

定理 2 的证明参考文献 [43], 此处略.

3 基于神经网络的实现方法

神经网络由大量相互连接的神经元构成. 神经元接收输入信号后, 通过激活函数生成输出信号, 这些信号随后传递至下一层的神经元. 通过不断优化算法并调整权重, 网络输出将逐渐接近真实值. 神经网络最显著的特性之一是其强大的函数逼近能力. 根据广义逼近定理, 当神经网络包含足够多的隐藏层神经元时, 可精确地逼近任意连续函数, 这使得神经网络在处理复杂非线性问题时表现极为高效^[3]. 为简化结构并利用最优性能指标与控制策略之间的关联, 本文采用单网络自适应评价设计. 该方法通过单一逼近器估计最优价值函数, 进而推导最优控制律, 无需单独的执行网络. 相较于执行-评价双网络结构, 单网络架构避免了执行网络与评价网络之间的迭代循环训练, 降低了整体复杂度^[4-5]. 本文设计的单网络自适应评价器采用 3 层全连接结构: 输入层神经元数匹配系统状态维度; 隐藏层为 1 层, 含有 2 维神经元; 输出层为 1 维神经元.

3.1 最优性能指标的近似计算

选择如下形式的单隐藏层神经网络来逼近 $\mathcal{V}^*(x)$

$$\mathcal{V}^*(x) = w^T \phi(x) + \eta(x) \quad (35)$$

其中, $w \in \mathbf{R}^k$ 为权重, k 为神经网络的节点数; $\phi(x) \in \mathbf{R}^k$ 表示激活函数; $\eta(x)$ 表示重构误差. 由于 w 未知, 引入 $\hat{w} \in \mathbf{R}^k$ 来近似 w . 于是, $\mathcal{V}^*(x)$ 可以近似为

$$\hat{\mathcal{V}}^*(x) = \hat{w}^T \phi(x) \quad (36)$$

$\nabla\mathcal{V}^*(x)$ 及其估计量满足

$$\nabla\mathcal{V}^*(x) = w^T \nabla\phi(x) + \nabla\eta(x) \quad (37)$$

$$\nabla\hat{\mathcal{V}}^*(x) = \hat{w}^T \nabla\phi(x) \quad (38)$$

根据式 (22), 可以重新得到下式

$$w^T \phi(x(t+T)) + \eta(x(t+T)) - (w^T \phi(x) + \eta(x)) + \int_t^{t+T} [\mathcal{Y}_x(x) + \mathcal{Y}_u(u^*)] dv = 0 \quad (39)$$

而式 (35) 中的误差源自估计值

$$\begin{aligned} \mathcal{C} &= \hat{\mathcal{V}}^*(x(t+T)) - \hat{\mathcal{V}}^*(x) + \\ &\int_t^{t+T} [\mathcal{Y}_x(x) + \mathcal{Y}_u(u^*)] dv = \\ &\hat{w}^T [\phi(x(t+T)) - \phi(x)] + \\ &\int_t^{t+T} [\mathcal{Y}_x(x) + \mathcal{Y}_u(u^*)] dv \end{aligned}$$

为了简化描述过程, 令

$$\Delta\phi(x) := \phi(x(t+T)) - \phi(x)$$

$$\Delta\mathcal{Y}(x) := \int_t^{t+T} [\mathcal{Y}_x(x) + \mathcal{Y}_u(u^*)] dv$$

进一步, 误差 \mathcal{C} 可以简单描述为

$$\mathcal{C} = \Delta\mathcal{Y}(x) + \hat{w}^T \Delta\phi(x) \quad (40)$$

接下来, 目标函数可以被定义为:

$$E = \frac{1}{2} \mathcal{C}^T \mathcal{C} \quad (41)$$

根据目标函数 E 对权重 \hat{w} 求偏导数, 可以得到

$$\frac{\partial E}{\partial \hat{w}} = \Delta\phi(x) \mathcal{C} \quad (42)$$

而权重 \hat{w} 可以根据下式进行更新:

$$\dot{\hat{w}} = -\frac{\alpha}{(1 + \Upsilon(x))^2} \Delta\phi(x) \mathcal{C} \quad (43)$$

其中, $\Upsilon(x) := \Delta\phi^T(x) \Delta\phi(x)$, $0 < \alpha < 1$ 表示学习率. 权重的估计误差被定义为:

$$\tilde{w} = w - \hat{w} \quad (44)$$

因此, $\dot{\tilde{w}}$ 变为

$$\begin{aligned} \dot{\tilde{w}} &= \frac{\alpha}{(1 + \Upsilon(x))^2} \Delta\phi(x) [\Delta\mathcal{Y}(x) + \hat{w}^T \Delta\phi(x)] = \\ &\frac{\alpha}{(1 + \Upsilon(x))^2} \Delta\phi(x) \times \\ &[\Delta\mathcal{Y}(x) + (w - \tilde{w})^T \Delta\phi(x)] = \\ &\frac{\alpha}{(1 + \Upsilon(x))^2} \Delta\phi(x) \times \\ &[\Delta\mathcal{Y}(x) + w^T \Delta\phi(x) - \tilde{w}^T \Delta\phi(x)] \end{aligned} \quad (45)$$

根据式 (39), 可以得到

$$\Delta\eta(x) := \eta(x) - \eta(x(t+T)) = w^T \Delta\phi(x) + \Delta\mathcal{Y}(x) \quad (46)$$

根据式 (46), $\dot{\tilde{w}}$ 可以变换为

$$\dot{\tilde{w}} = \frac{\alpha}{(1 + \Upsilon(x))^2} [-\Upsilon^T(x) \tilde{w} + \Delta\phi(x) \Delta\eta(x)] \quad (47)$$

在权重更新方法中, 考虑到方法的收敛速度与稳定性的平衡, 学习率较小会导致收敛过慢, 较大易引发权重震荡. 因此, 学习率取值范围一般设定为 $\alpha \in [0.01, 0.50]$, 实际应用中优先选取 0.10 作为初始值, 再根据权重误差动态微调.

3.2 控制策略的近似计算

基于 TTM 的最优控制策略可以表示为

$$u_i^*(x) = -\mathcal{N}(v_i^*(x), \mu_{ai}, \mu_{bi}) \quad (48)$$

其中, $v_i^*(x)$ 定义为

$$v_i^*(x) = \frac{1}{2} r_i^{-1} G_i^T(x) [w^T \nabla\phi(x) + \nabla\eta(x)] \quad (49)$$

根据式 (36), $\hat{v}_i^*(x)$ 满足

$$\hat{v}_i^*(x) = \frac{1}{2} r_i^{-1} G_i^T(x) \hat{w}^T \phi(x) \quad (50)$$

当静态事件触发条件 (25) 不成立时, 基于 SEM 的控制策略可以根据以下方式更新:

$$\hat{u}_i^*(x(l_j)) = -\mathcal{N}(\hat{v}_i^*(x(l_j)), \mu_{ai}, \mu_{bi}) \quad (51)$$

其中, $\hat{v}_i^*(x(l_j))$ 被定义为

$$\hat{v}_i^*(x(l_j)) = \frac{1}{2} r_i^{-1} G_i^T(x(l_j)) \hat{w}^T \nabla\phi(x(l_j)) \quad (52)$$

同样地, 当动态事件触发条件 (32) 不成立时, 基于 DEM 的控制策略可以根据下式更新:

$$\hat{u}_i^*(x(l_d)) = -\mathcal{N}(\hat{v}_i^*(x(l_d)), \mu_{ai}, \mu_{bi}) \quad (53)$$

其中, $\hat{v}_i^*(x(l_d))$ 被表示为

$$\hat{v}_i^*(x(l_d)) = \frac{1}{2} r_i^{-1} G_i^T(x(l_d)) \hat{w}^T \nabla\phi(x(l_d)) \quad (54)$$

3.3 理论分析

为了进行稳定性分析, 采用以下假设.

假设 3. $\nabla\phi(x)$ 是利普希茨连续的, 且存在常数 $\mathcal{L}_\phi > 0$. $\nabla\eta(x)$ 、 $\nabla\phi(x)$ 和 w 具有正上界, 其常数分别为 $\bar{\eta}_\nabla$ 、 $\bar{\phi}_\nabla$ 和 \bar{w} .

在假设 3 中, 为了避免因梯度剧烈波动导致控制策略震荡或稳定性分析失效, 一般将 $\nabla\phi(x)$ 定义为利普希茨连续的. 考虑到实际工程应用中的固有物理约束, 假设 $\nabla\eta(x)$ 、 $\nabla\phi(x)$ 和 w 具有正上界, 可以避免变量无界导致的控制增益饱和、系统性能发

散等问题.

定理 3. 对于采用基于 SEM 控制策略 $\hat{u}^*(x(l_j))$ 或基于 DEM 控制策略 $\hat{u}^*(x(l_d))$ 的系统 (12), 令假设 1 至假设 3 成立. 式 (36) 基于式 (43) 进行学习, 以下结论成立:

1) 当下述静态事件触发条件

$$\|\Delta_j x(t)\| \leq \frac{\sqrt{(1-\lambda_2)Q}}{\mathcal{L}_m} \|x\| \quad (55)$$

不成立时, 将触发静态事件. 基于 SEM 的控制策略 $\hat{u}^*(x(l_j))$ 确保了系统状态和权重估计误差的 UUB 特性.

2) 当下述动态事件触发条件

$$\theta_2(t) + \gamma_2 \pi_2(t) \geq 0 \quad (56)$$

不成立时, 将触发动态事件. 上式中, 事件触发动态变量 $\theta_2(t)$ 和 $\pi_2(t)$ 满足

$$\dot{\theta}_2(t) = -\rho_2 \theta_2(t) + \pi_2(t) \quad (57)$$

$$\pi_2(t) = (1-\lambda_2)x^T Q x - \mathcal{L}_m^2 \|\Delta_j x(t)\|^2 \quad (58)$$

其中, $0 < \lambda_2 < 1$, $\mathcal{L}_m > 0$, $\gamma_2 > 0$ 及 $\rho_2 > 0$ 为常数. 基于 DEM 的控制策略 $\hat{u}^*(x(l_d))$ 仍能确保系统状态与权重估计误差的 UUB 特性.

3) 对于权重学习, 仍可得出结论 $\Delta_j l \leq \Delta_{dl}$.

证明. 1) SEM. 选择以下形式的李雅普诺夫函数

$$\mathcal{W}_j = \mathcal{V}^*(x) + \mathcal{V}^*(x(l_j)) + \mathcal{W}_w(\tilde{w}) \quad (59)$$

其中, $\mathcal{W}_w(\tilde{w})$ 可以表示为

$$\mathcal{W}_w(\tilde{w}) = \tilde{w}^T \tilde{w}$$

考虑到触发瞬间与非触发瞬间的差异, 稳定性分析需分为两类情形: 情形 1: $t \in [l_j, l_{j+1})$ 和情形 2: $t = l_{j+1}$.

情形 1. 当 $t \in [l_j, l_{j+1})$ 时, 由于此时静态事件触发条件 (55) 未被违反, 事件未被触发, 因此有 $\dot{\mathcal{V}}^*(x(l_j)) = 0$. $\dot{\mathcal{V}}^*(x)$ 满足

$$\begin{aligned} \dot{\mathcal{V}}^*(x) &= \nabla \mathcal{V}^{*\top}(x) [F(x) + G(x)\hat{u}^*(x(l_j))] = \\ &\nabla \mathcal{V}^{*\top}(x) G(x) [\hat{u}^*(x(l_j)) - u^*(x)] + \\ &\nabla \mathcal{V}^{*\top}(x) [F(x) + G(x)u^*(x)] \leq \\ &\frac{1}{2} (\bar{\mathcal{V}}_{\nabla}^2 \bar{G}^2 + \|\hat{u}^*(x(l_j)) - u^*(x)\|^2) - \\ &\mathcal{Y}_x(x) - \mathcal{Y}_u(u^*(x)) \end{aligned} \quad (60)$$

其中 $\bar{\mathcal{V}}_{\nabla}^2 = \|\tilde{w}\bar{\phi}_{\nabla} + \bar{\eta}_{\nabla}\|^2$. 对于式 (60) 的 $\|\hat{u}^*(x(l_j)) - u^*(x)\|^2$, 结合式 (48) 与式 (51), 可以得到

$$\begin{aligned} &\|\hat{u}^*(x(l_j)) - u^*(x)\|^2 \leq \\ &\mathcal{L}_n^2 \|\hat{v}^*(x(l_j)) - \hat{v}^*(x)\|^2 \leq \\ &\frac{\mathcal{L}_n^2}{4r^2} \|G^T(x(l_j))\hat{w}^T \nabla \phi(x(l_j)) - \\ &G^T w^T \nabla \phi + G^T \nabla \eta\|^2 \end{aligned} \quad (61)$$

其中, \mathcal{L}_n 是 $\mathcal{N}(\cdot)$ 的利普希茨常数. 利用 $\|a+b\|^2 \leq 2\|a\|^2 + 2\|b\|^2$, 可以得到如下不等式

$$\begin{aligned} &\frac{1}{2} (\|\hat{u}^*(x(l_j)) - u^*(x)\|^2) \leq \\ &\frac{\mathcal{L}_n^2}{4r^2} \|G^T(x(l_j))\hat{w}^T \nabla \phi(x(l_j)) - G^T w^T \nabla \phi\|^2 + \\ &\frac{\mathcal{L}_n}{4r^2} \|G^T \nabla \eta\|^2 \leq \\ &\frac{\mathcal{L}_n^2}{4r^2} \underbrace{\|G^T(x(l_j))(w - \tilde{w})^T \nabla \phi(x(l_j)) - G^T w^T \nabla \phi\|^2}_{p_1} + \\ &\frac{\mathcal{L}_n^2}{4r^2} \bar{G}^2 \bar{\eta}_{\nabla}^2 \end{aligned} \quad (62)$$

式 (62) 中的项 p_1 满足

$$\begin{aligned} p_1 &= \\ &\|G^T(x(l_j))\nabla \phi^T(x(l_j))(w - \tilde{w}) - G^T \nabla \phi^T w\|^2 = \\ &\|(G^T(x(l_j))\nabla \phi^T(x(l_j)) - G^T \nabla \phi^T)w - \\ &G^T(x(l_j))\nabla \phi^T(x(l_j))\tilde{w}\|^2 \leq \\ &2\| \underbrace{G^T(x(l_j))\nabla \phi^T(x(l_j)) - G^T \nabla \phi^T}_{p_2} \|^2 \|w\|^2 + \\ &2\|G^T(x(l_j))\nabla \phi^T(x(l_j))\|^2 \|\tilde{w}\|^2 \end{aligned} \quad (63)$$

式 (63) 中的项 p_2 满足以下关系

$$\begin{aligned} p_2 &= \|G^T(x(l_j))\nabla \phi^T(x(l_j)) - G^T(x(l_j))\nabla \phi^T + \\ &G^T(x(l_j))\nabla \phi^T - G^T \nabla \phi^T\|^2 \leq \\ &2\|G^T(x(l_j))\nabla \phi^T(x(l_j)) - G^T(x(l_j))\nabla \phi^T\|^2 + \\ &2\|G^T(x(l_j))\nabla \phi^T - G^T \nabla \phi^T\|^2 \leq \\ &2\bar{G}^2 \|\nabla \phi^T(x(l_j)) - \nabla \phi^T\|^2 + \\ &2\bar{\phi}_{\nabla}^2 \|G^T(x(l_j)) - G^T\|^2 \leq \\ &2(\bar{G}^2 \mathcal{L}_{\phi}^2 + \bar{\phi}_{\nabla}^2 \mathcal{L}_G^2) \|\Delta_j x(t)\|^2 \end{aligned} \quad (64)$$

根据式 (64), p_1 可以变换为

$$\begin{aligned} p_1 &\leq 4(\bar{G}^2 \mathcal{L}_{\phi}^2 + \bar{\phi}_{\nabla}^2 \mathcal{L}_G^2) \bar{w}^2 \|\Delta_j x(t)\|^2 + \\ &2\bar{G}^2 \bar{\phi}_{\nabla}^2 \|\tilde{w}\|^2 \end{aligned} \quad (65)$$

因此, 可以得到

$$\begin{aligned} \frac{1}{2}(\|\hat{u}^*(x(l_j)) - u^*(x)\|^2) \leq \\ \mathcal{L}_m^2 \|\Delta_j x(t)\|^2 + \frac{1}{2} \mathcal{L}_n^2 r^{-2} \bar{G}^2 \bar{\phi}_{\nabla}^2 \|\tilde{w}\|^2 + \\ \frac{\mathcal{L}_n^2}{4r^2} \bar{G}^2 \bar{\eta}_{\nabla}^2 \end{aligned} \quad (66)$$

其中 $\mathcal{L}_m^2 = \mathcal{L}_n^2 r^{-2} (\bar{G}^2 \mathcal{L}_\phi^2 + \bar{\phi}_{\nabla}^2 \mathcal{L}_G^2) \bar{w}^2$. 利用式 (66), 并考虑 $\mathcal{Y}_u(u^*(x)) \geq 0$, 可以将式 (60) 变换为

$$\begin{aligned} \dot{\mathcal{V}}^*(x) \leq -\lambda_2 \mathcal{Y}_x(x) - (1 - \lambda_2) \mathcal{Y}_x(x) + \frac{1}{2} \bar{\mathcal{V}}_{\nabla}^2 \bar{G}^2 + \\ \frac{\mathcal{L}_n}{2r^2} \bar{G}^2 \bar{\eta}_{\nabla}^2 + \mathcal{L}_m^2 \|\Delta_j x(t)\|^2 + \\ \frac{1}{2} \mathcal{L}_n^2 r^{-2} \bar{G}^2 \bar{\phi}_{\nabla}^2 \|\tilde{w}\|^2 \end{aligned} \quad (67)$$

根据 SEM, 可以推导出

$$\begin{aligned} \dot{\mathcal{V}}^*(x) \leq -\lambda_2 \mathcal{Y}_x(x) + \frac{1}{2} \bar{\mathcal{V}}_{\nabla}^2 \bar{G}^2 + \\ \frac{1}{2} \mathcal{L}_n^2 r^{-2} \bar{G}^2 \bar{\phi}_{\nabla}^2 \|\tilde{w}\|^2 + \frac{\mathcal{L}_n}{2r^2} \bar{G}^2 \bar{\eta}_{\nabla}^2 \end{aligned} \quad (68)$$

对于 $\mathcal{W}_w(\tilde{w})$, 可以表示为

$$\dot{\mathcal{W}}_w(\tilde{w}) = 2\tilde{w}^T \dot{\tilde{w}} \quad (69)$$

注意到式 (47) 中的 $\dot{\tilde{w}}$, 式 (69) 可以重新表示为

$$\begin{aligned} \dot{\mathcal{W}}_w(\tilde{w}) = \frac{\alpha}{(1 + \Upsilon(x))^2} \times \\ (-2\tilde{w}^T \Upsilon^T(x) \tilde{w} + 2\tilde{w}^T \Delta\phi(x) \Delta\eta(x)) \leq \\ \frac{\alpha}{(1 + \Upsilon(x))^2} \times \\ (-\tilde{w}^T \Upsilon^T(x) \tilde{w} + \Delta\eta^T(x) \Delta\eta(x)) \leq \\ -\alpha \Upsilon^T(x) \|\tilde{w}\|^2 + \alpha \bar{\eta}_{\nabla}^2 \end{aligned} \quad (70)$$

结合式 (68) 与式 (70), 可以得到

$$\dot{\mathcal{W}}^j \leq -\lambda_2 Q \|x\|^2 + \lambda_3 - \lambda_4 \|\tilde{w}\|^2 \quad (71)$$

上式中的

$$\lambda_3 = \frac{1}{2} \bar{\mathcal{V}}_{\nabla}^2 \bar{G}^2 + \frac{\mathcal{L}_n}{2r^2} \bar{G}^2 \bar{\eta}_{\nabla}^2 + \alpha \bar{\eta}_{\nabla}^2$$

$$\lambda_4 = \alpha \Delta\phi(x) \Delta\phi^T(x) - \frac{1}{2} \mathcal{L}_n^2 r^{-2} \bar{G}^2 \bar{\phi}_{\nabla}^2 > 0$$

通过选择适当的参数可确保上式成立, 因此当 x 在集合 \mathcal{B}_j 外部或 \tilde{w} 在集合 \mathcal{B}_w 外部时, 有 $\dot{\mathcal{W}}^j < 0$, 其中,

$$\begin{cases} \mathcal{B}_j = \left\{ x: \|x\| \leq \sqrt{\frac{\lambda_3}{\lambda_2 Q}} \right\} \\ \mathcal{B}_w = \left\{ \tilde{w}: \|\tilde{w}\| \leq \sqrt{\frac{\lambda_3}{\lambda_4}} \right\} \end{cases}$$

根据李雅普诺夫定理的推广, x 与 \tilde{w} 均为 UUB.

情形 2. 对于 $t = l_{j+1}$, 考虑式 (59) 的微分方程,

$$\Delta \mathcal{W}_j = \Delta \mathcal{V}^*(x) + \Delta \mathcal{V}^*(x(l_{j+1})) + \Delta \mathcal{W}_w(\tilde{w}) \quad (72)$$

其中,

$$\Delta \mathcal{V}^*(x) = \mathcal{V}^*(x^+) - \mathcal{V}^*(x)$$

$$\Delta \mathcal{V}^*(x(l_{j+1})) = \mathcal{V}^*(x(l_{j+1})) - \mathcal{V}^*(x(l_j))$$

$$\Delta \mathcal{W}_w(\tilde{w}) = (\tilde{w}^+)^T \tilde{w}^+ - \tilde{w}^T \tilde{w}$$

根据情形 1, 得到在 SEM 条件下, $\mathcal{V}^*(x)$ 和 $\mathcal{W}_w(\tilde{w})$ 均为递减函数. 因此, 对于连续状态 x 及其对应的 \tilde{w} , 满足

$$\begin{cases} x^+ = \lim_{\rho \rightarrow 0} x(t + \rho) \\ \tilde{w}^+ = \lim_{\sigma \rightarrow 0} \tilde{w}(t + \sigma) \end{cases}$$

因此, 有 $\Delta \mathcal{V}^*(x) \leq 0$ 且 $\Delta \mathcal{W}_w(\tilde{w}) \leq 0$. 对于式 (72) 中的第二项, 有

$$\mathcal{V}^*(x(l_{j+1})) - \mathcal{V}^*(x(l_j)) \leq -\kappa(\|x(l_{j+1}) - x(l_j)\|)$$

其中, $\kappa(\cdot)$ 表示 κ 函数类^[48].

根据情形 1 和情形 2, $x(t)$ 和 \tilde{w} 均为 UUB.

2) DEM. 选择如下形式的李雅普诺夫函数

$$\mathcal{W}^d = \mathcal{W}^j + \theta_2(t) \quad (73)$$

情形 1. 对于 $t \in [l_j, l_{j+1})$, 结合 SEM 的结果, 可以得到

$$\begin{aligned} \dot{\mathcal{W}}^d = \dot{\mathcal{W}}^j + \dot{\theta}_2(t) \leq \\ - (1 - \lambda_2) \mathcal{Y}_x(x) + \mathcal{L}_m^2 \|\Delta_j x(t)\|^2 - \\ \rho_2 \theta_2(t) + \pi_2(t) \leq -\rho_2 \theta_2(t) \end{aligned} \quad (74)$$

根据式 (56) 和式 (57) 可得

$$\dot{\theta}_2(t) + \rho_2 \theta_2(t) \geq -\frac{\theta_2(t)}{\gamma_2}$$

将上式进行推导后, 可以得到

$$\theta_2(t) \geq \theta_2(0) e^{-\left(\rho_2 + \frac{1}{\gamma_2}\right)t} \geq 0 \quad (75)$$

由此可得 $\dot{\mathcal{W}}^d \leq 0$.

情形 2. 当 $t = l_{j+1}$ 时, $\theta_2(t)$ 不发生变化. 因此, $\theta_2(t)$ 的导数为 0. 由此可知, 在 DEM 算法中, x 和 \tilde{w} 的 UUB 特性依然得到保证.

3) 由式 (56) 可得

$$\|\Delta_j x(t)\| \leq \frac{\sqrt{(1 - \lambda_2)Q}}{\mathcal{L}_m} \|x\| + \frac{\sqrt{\frac{\theta_2(t)}{\gamma_2}}}{\mathcal{L}_m} \quad (76)$$

通过比较式 (55) 与式 (76) 可知, 由于

$$\frac{\sqrt{\frac{2\theta_2(t)}{\gamma_2}}}{\mathcal{L}_m} > 0$$

触发动态事件需要相对较大的 $\Delta_j x(t)$, 因此有 $\Delta_j l \leq \Delta d$. \square

4 仿真结果

USV 被广泛应用于海洋资源探测、军事侦察、环境监测等实际场景, 其航行过程中面临推力输出上下限非对称约束、航速与转向角物理限制以及海洋风浪流扰动带来的非线性动力学特性, 是验证本文方法有效性的理想载体. USV 的示意图如图 1 所示, 其模型可表示为:

$$\begin{bmatrix} \dot{s}_1 \\ \dot{s}_2 \\ \dot{s}_3 \end{bmatrix} = \begin{bmatrix} \cos s_3 & -\sin s_3 & 0 \\ \sin s_3 & \cos s_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_4 \\ s_5 \\ s_6 \end{bmatrix} \quad (77)$$

$$\begin{bmatrix} M_4 & 0 & 0 \\ 0 & M_5 & 0 \\ 0 & 0 & M_6 \end{bmatrix} \begin{bmatrix} \dot{s}_4 \\ \dot{s}_5 \\ \dot{s}_6 \end{bmatrix} = \begin{bmatrix} u_1 \\ 0 \\ u_2 \end{bmatrix} - \begin{bmatrix} C_4 s_4 \\ C_5 s_5 \\ C_6 s_6 \end{bmatrix} + \begin{bmatrix} M_5 & 0 & 0 \\ 0 & M_4 & 0 \\ 0 & 0 & M_4 - M_5 \end{bmatrix} \begin{bmatrix} s_5 s_6 \\ s_4 s_6 \\ s_4 s_5 \end{bmatrix} \quad (78)$$

其中, s_1 和 s_2 表示位置, s_4 和 s_5 为线速度, s_3 和 s_6 分别为航向角和角速度; M 和 C 分别为惯性系数与水动力系数, 其设定值分别为: $M_4 = 1.9$, $C_4 = 0.4$, $M_5 = 0.3$, $C_5 = 0.4$, $M_6 = 1.0$, $C_6 = 0.1$. 边界条件为 $s \in [-0.5, 1.5]$, $u \in [-1.5, 1.0]$, $Q = 0.01I_5$, 且 $r_i = 1$. 初始权重在区间 $[-1, 1]$ 内随机选取, $T = 0.1$, $\lambda_2 = 0.98$, $\mathcal{L}_m = 5\sqrt{2}$, $\theta(0) = 3$, $\rho_2 = 4$, $\gamma_2 = 0.5$. ETM 中的参数设置为: $\lambda_2 = 0.90$, $\mathcal{L}_m = \sqrt{15}$, $\theta(0) = 1$, $\rho_2 = 1$, $\gamma_2 = 1$.

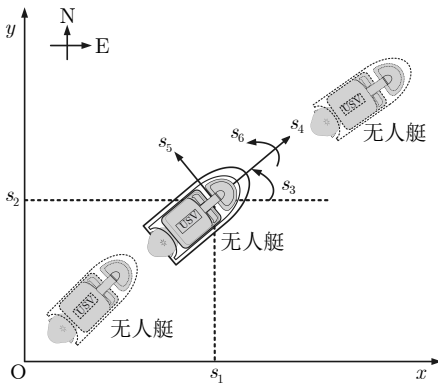


图 1 USV 示意图

Fig.1 The schematic diagram of USV

比较实验将分两个步骤进行. 第一步, 本文提出的方法与 TTM^[9] 和 SEM^[24] 进行了对比. 图 2 展示了 SEM^[24] 与 DEM 的权重学习过程, 两者呈现相似的学习轨迹. 图 3 展示了 SEM 与 DEM 的控制策略, 均为连续信号分割方法. 由图 4 可见 (Γ_j 和 Γ_d 为触发阈值), 无论采用静态还是动态方法, 误差均未超出设定阈值条件, 表明两种方法均有效. 图 5 比较了两种方法的执行间隔时间, 可见 DEM 的平均间隔时间大于 SEM. 图 6 比较了 TTM^[9]、SEM^[24] 和 DEM 的样本量, 相较于前两种方法, DEM 分别减少了 85% 和 38% 的样本量. TTM、SEM 与 DEM 的对比结果见表 2. 比较指标包括控制策略演化 (evolution of control policy, EOC)、误差是否超阈值 (error violates the threshold, EVT)、平均执行间隔时间 (average interexecution time, AET)、样本数量 (number of samples, NOS) 及执行神经网络更新次数 (number of actor NN updates, NOA). 可见 DEM 具有最长的平均执行间隔时间、最少的样本数量, 且无需更新执行网络. 第二

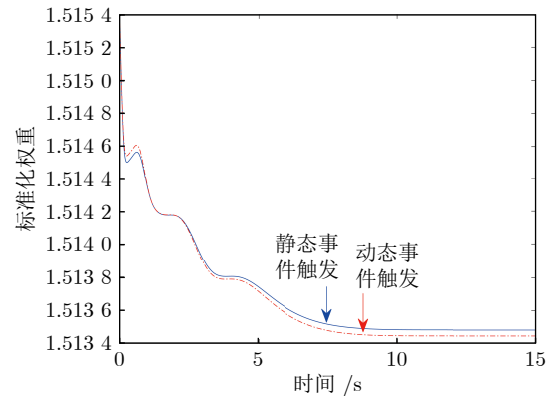


图 2 ETM 中 \hat{w} 的演化

Fig.2 Evolution of \hat{w} in ETM

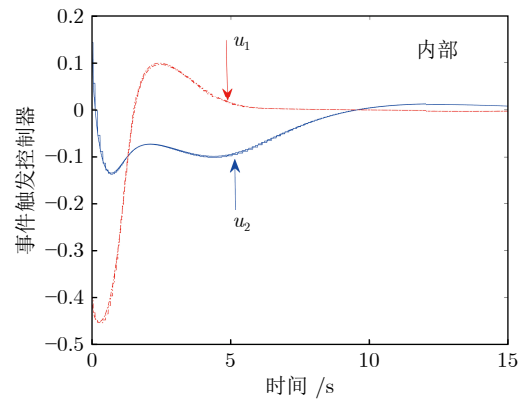


图 3 控制策略的演化

Fig.3 Evolution of control policy

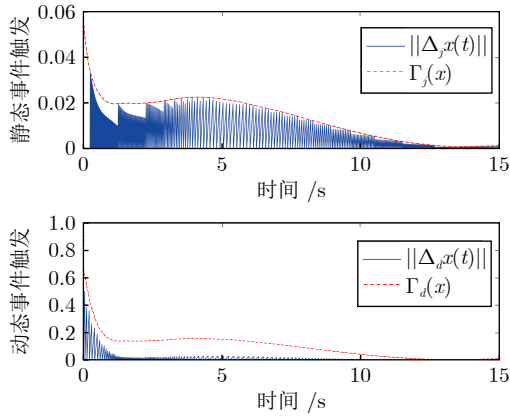


图 4 ETM 的轨迹
Fig.4 Trajectory of ETM

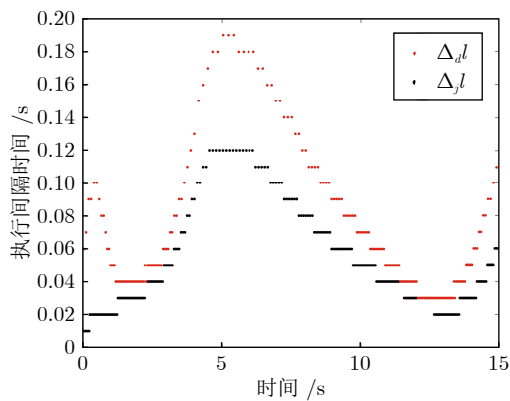


图 5 执行间隔时间的轨迹
Fig.5 Trajectory of the interexecution time

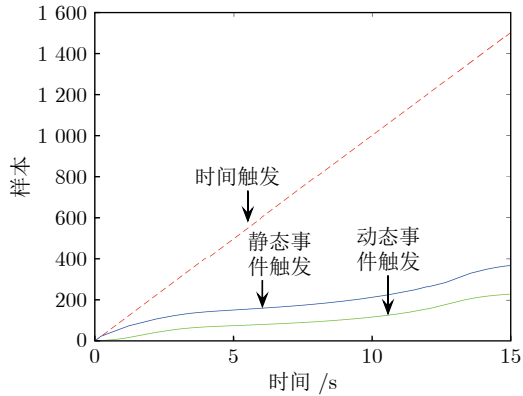


图 6 样本的演化
Fig.6 Evolution of samples

表 2 TTM, SEM 和 DEM 的比较
Table 2 Comparison of TTM, SEM and DEM

| 方法 | EOC | EVT | AET | NOS | NOA |
|---------------------|-----|-----|------|------|------|
| TTM ^[9] | 连续 | — | 0.01 | 1500 | 1500 |
| SEM ^[24] | 分割 | 否 | 0.04 | 368 | 1500 |
| DEM | 分割 | 否 | 0.07 | 229 | 0 |

步将提出的方法与无映射方法^[43]进行对比:从图 7 可见,无映射方法^[43]无法将轨迹限制在边界内;从图 8 可知,本文方法成功将轨迹限制在预设边界内.这证明了本文控制方法的有效性.

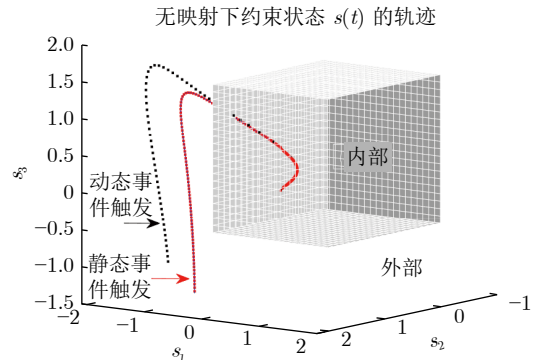


图 7 无映射下 s(t) 的轨迹
Fig.7 Trajectory of s(t) without mapping

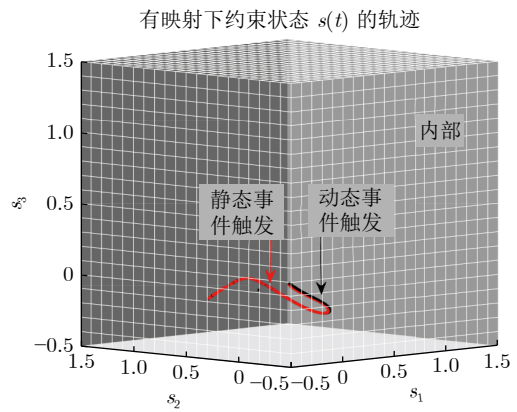


图 8 有映射下 s(t) 的轨迹
Fig.8 Trajectory of s(t) with mapping

5 结束语

在处理具有非对称边界的最优控制问题时,本研究采用了基于 ADP 的 SEM 与 DEM 方法.首先通过非线性映射函数将原始问题转化为无约束最优控制问题.其次基于静态事件触发条件,进一步提出动态事件触发条件,理论分析证明了动态触发间隔不小于静态触发间隔.进一步利用神经网络实现了所提方法.最后,仿真结果证明了该方法的可行性.在未来,本文方法在实际应用中可能面临环境的复杂扰动、传感器或执行器突发故障以及网络传输受阻等挑战.对此,可以在 ADP 评价网络中融入故障诊断与补偿项,实时辨识执行器效率衰减和传感器偏移等故障特征,自适应修正控制输入以抵消故障影响,以提升所提方法在复杂环境下的鲁棒性与实际场景中的适用性.

参考文献

- 1 Luo Biao, Hu Tian-Meng, Zhou Yu-Hao, Huang Ting-Wen, Yang Chun-Hua, Gui Wei-Hua. Survey on multi-agent reinforcement learning for control and decision-making. *Acta Automatica Sinica*, 2025, **51**(3): 510–539
(罗彪, 胡天萌, 周育豪, 黄廷文, 阳春华, 桂卫华. 多智能体强化学习控制与决策研究综述. 自动化学报, 2025, **51**(3): 510–539)
- 2 Du Cheng-Long, Han Jie, Li Fan-Biao, Gui Wei-Hua. Model-free policy gradient-based reinforcement learning algorithms for optimal control of unknown stochastic systems. *Acta Automatica Sinica*, 2025, **51**(10): 2245–2255
(杜城龙, 韩洁, 李繁飙, 桂卫华. 基于无模型策略梯度强化学习的未知随机系统最优控制. 自动化学报, 2025, **51**(10): 2245–2255)
- 3 Lewis F L, Liu D R. *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*. Hoboken: John Wiley & Sons, 2013.
- 4 Liu D R, Xue S, Zhao B, Luo B, Wei Q L. Adaptive dynamic programming for control: A survey and recent advances. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, **51**(1): 142–160
- 5 Liu D R, Wei Q L, Wang D, Yang X, Li H L. *Adaptive Dynamic Programming With Applications in Optimal Control*. Cham: Springer, 2017.
- 6 Wei Qing-Lai, Zhang Hua-Guang, Liu De-Rong, Zhao Yan. An optimal control scheme for a class of discrete-time nonlinear systems with time delays using adaptive dynamic programming. *Acta Automatica Sinica*, 2010, **36**(1): 121–129
(魏庆来, 张化光, 刘德荣, 赵琰. 基于自适应动态规划的一类带有时滞的离散时间非线性系统的最优控制策略. 自动化学报, 2010, **36**(1): 121–129)
- 7 Sun J Y, Zhang H G, Yan Y, Xu S, Fan X X. Optimal regulation strategy for nonzero-sum games of the immune system using adaptive dynamic programming. *IEEE Transactions on Cybernetics*, 2023, **53**(3): 1475–1484
- 8 Liu D R, Wei Q L. Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, **25**(3): 621–634
- 9 Zhao B, Zhang Y W, Liu D R. Adaptive dynamic programming-based cooperative motion/force control for modular reconfigurable manipulators: A joint task assignment approach. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, **34**(12): 10944–10954
- 10 Liu D R, Yang X, Wang D, Wei Q L. Reinforcement-learning-based robust controller design for continuous-time uncertain nonlinear systems subject to input constraints. *IEEE Transactions on Cybernetics*, 2015, **45**(7): 1372–1385
- 11 Kiumarsi B, Lewis F L. Actor-critic-based optimal tracking for partially unknown nonlinear discrete-time systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2015, **26**(1): 140–151
- 12 Zhu Y H, Zhao D B. Comprehensive comparison of online ADP algorithms for continuous-time optimal control. *Artificial Intelligence Review*, 2018, **49**(4): 531–547
- 13 Wang D, Zhao M M, Qiao J F. Intelligent optimal tracking with asymmetric constraints of a nonlinear wastewater treatment system. *International Journal of Robust and Nonlinear Control*, 2021, **31**(14): 6773–6787
- 14 Liu D R, Wang D, Yang X. An iterative adaptive dynamic programming algorithm for optimal control of unknown discrete-time nonlinear systems with constrained inputs. *Information Sciences*, 2013, **220**: 331–342
- 15 Yi X N, Luo B, Zhao Y Q. Neural network-based robust guaranteed cost control for image-based visual servoing of quadrotor. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, **35**(9): 12693–12705
- 16 Wei Q L, Liu D R, Liu Y, Song R Z. Optimal constrained self-learning battery sequential management in microgrid via adaptive dynamic programming. *IEEE/CAA Journal of Automatica Sinica*, 2017, **4**(2): 168–176
- 17 Yang Y L, Ding D W, Xiong H Y, Yin Y X, Wunsch D C. On-line barrier-actor-critic learning for H_∞ control with full-state constraints and input saturation. *Journal of the Franklin Institute*, 2020, **357**(6): 3316–3344
- 18 Kong L H, He W, Liu Z J, Yu X B, Silvestre C. Adaptive tracking control with global performance for output-constrained MIMO nonlinear systems. *IEEE Transactions on Automatic Control*, 2023, **68**(6): 3760–3767
- 19 Fan B, Yang Q M, Tang X Y, Sun Y X. Robust ADP design for continuous-time nonlinear systems with output constraints. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, **29**(6): 2127–2138
- 20 Fan B, Yang Q M, Jagannathan S, Sun Y X. Output-constrained control of nonaffine multiagent systems with partially unknown control directions. *IEEE Transactions on Automatic Control*, 2019, **64**(9): 3936–3942
- 21 He K H, Shi S L, van den Boom T, de Schutter B. Approximate dynamic programming for constrained piecewise affine systems with stability and safety guarantees. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2025, **55**(3): 1722–1734
- 22 Tabuada P. Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Transactions on Automatic Control*, 2007, **52**(9): 1680–1685
- 23 Heemels W P M H, Johansson K H, Tabuada P. An introduction to event-triggered and self-triggered control. In: Proceedings of the IEEE 51st IEEE Conference on Decision and Control (CDC). Maui, USA: IEEE, 2012. 3270–3285
- 24 Vamvoudakis K G. Event-triggered optimal adaptive control algorithm for continuous-time nonlinear systems. *IEEE/CAA Journal of Automatica Sinica*, 2014, **1**(3): 282–293
- 25 Xie X P, Zhou Q, Yue D, Li H Y. Relaxed control design of discrete-time Takagi-Sugeno fuzzy systems: An event-triggered real-time scheduling approach. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018, **48**(12): 2251–2262
- 26 Zhao B, Liu D R. Event-triggered decentralized tracking control of modular reconfigurable robots through adaptive dynamic programming. *IEEE Transactions on Industrial Electronics*, 2020, **67**(4): 3054–3064
- 27 Liu J, Zhang Y L, Yu Y, Sun C Y. Fixed-time leader-follower consensus of networked nonlinear systems via event/self-triggered control. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, **31**(11): 5029–5037
- 28 Bai W W, Li T S, Long Y, Chen P C L. Event-triggered multi-gradient recursive reinforcement learning tracking control for multiagent systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, **34**(1): 366–379
- 29 Li Zhen-Xing, Zhuang Jiao-Jiao, Yang Cheng-Dong, Qiu Jian-Long, Cao Jin-De. Event-triggered state consensus of heterogeneous uncertain second-order nonlinear multi-agent systems. *Acta Automatica Sinica*, 2025, **51**(4): 804–812
(李振兴, 庄娇娇, 杨成东, 邱建龙, 曹进德. 异构不确定二阶非线性多智能体系统事件触发状态趋同. 自动化学报, 2025, **51**(4): 804–812)
- 30 Li Y M, Li Y X, Tong S C. Event-based finite-time control for nonlinear multiagent systems with asymptotic tracking. *IEEE Transactions on Automatic Control*, 2022, **68**(6): 3790–3797
- 31 Zhao F Y, Luo S X Y, Gao W N, Wen C Y. Event-triggered cooperative adaptive optimal output regulation for multiagent systems under switching network: An adaptive dynamic programming approach. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2025, **55**(3): 1707–1721
- 32 Dong L, Tang Y F, He H B, Sun C Y. An event-triggered approach for load frequency control with supplementary ADP.

- IEEE Transactions on Power Systems*, 2017, **32**(1): 581–589
- 33 Wei Q L, Lu J W, Zhou T M, Cheng X, Wang F Y. Event-triggered near-optimal control of discrete-time constrained nonlinear systems with application to a boiler-turbine system. *IEEE Transactions on Industrial Informatics*, 2022, **18**(6): 3926–3935
- 34 Girard A. Dynamic triggering mechanisms for event-triggered control. *IEEE Transactions on Automatic Control*, 2015, **60**(7): 1992–1997
- 35 Dolk V S, Borgers D P, Heemels W P M H. Output-based and decentralized dynamic event-triggered control with guaranteed L_p -gain performance and zero-freeness. *IEEE Transactions on Automatic Control*, 2017, **62**(1): 34–49
- 36 Wang L J, Chen P C L. Reduced-order observer-based dynamic event-triggered adaptive NN control for stochastic nonlinear systems subject to unknown input saturation. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, **32**(4): 1678–1690
- 37 He Yi-Rui, Su Han-Guang, Zhang Hua-Guang, Luan Xin-Yang. Online decentralized dynamic event-triggered control of unknown large-scale interconnected systems. *Acta Automatica Sinica*, 2025, **51**(9): 2011–2026
(何怡睿, 苏涵光, 张化光, 栾鑫洋. 未知大规模互联系统在线分散式动态事件触发控制. 自动化学报, 2025, **51**(9): 2011–2026)
- 38 Ge X H, Han Q L, Ding L, Wang Y L, Zhang X M. Dynamic event-triggered distributed coordination control and its applications: A survey of trends and techniques. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020, **50**(9): 3112–3125
- 39 Zhang P, Yuan Y, Guo L. Fault-tolerant optimal control for discrete-time nonlinear system subjected to input saturation: A dynamic event-triggered approach. *IEEE Transactions on Cybernetics*, 2021, **51**(6): 2956–2968
- 40 Shi T, Shi P, Wu Z G. Dynamic event-triggered asynchronous MPC of Markovian jump systems with disturbances. *IEEE Transactions on Cybernetics*, 2022, **52**(11): 11639–11648
- 41 He W L, Xu B, Han Q L, Qian F. Adaptive consensus control of linear multiagent systems with dynamic event-triggered strategies. *IEEE Transactions on Cybernetics*, 2020, **50**(7): 2996–3008
- 42 Zhang H G, Li W H, Zhang J, Wang Y C, Sun J Y. Fully distributed dynamic event-triggered bipartite formation tracking for multiagent systems with multiple nonautonomous leaders. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, **34**(10): 7453–7466
- 43 Mu C X, Wang K, Qiu T. Dynamic event-triggering neural learning control for partially unknown nonlinear systems. *IEEE Transactions on Cybernetics*, 2022, **52**(4): 2200–2213
- 44 Mu C X, Wang K, Sun C Y. Learning control supported by dynamic event communication applying to industrial systems. *IEEE Transactions on Industrial Informatics*, 2021, **17**(4): 2325–2335
- 45 Yang X, Xu M M, Wei Q L. Dynamic event-sampled control of interconnected nonlinear systems using reinforcement learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, **35**(1): 923–937
- 46 Wang Ding, Hu Ling-Zhi, Zhao Ming-Ming, Ha Ming-Ming, Qiao Jun-Fei. Event-triggered control design for optimal tracking of unknown nonlinear zero-sum games. *Acta Automatica Sinica*, 2023, **49**(1): 91–101
(王鼎, 胡凌治, 赵明明, 哈明鸣, 乔俊飞. 未知非线性零和博弈最优跟踪的事件触发控制设计. 自动化学报, 2023, **49**(1): 91–101)
- 47 Ren Hao, Ma Ya-Jie, Jiang Bin, Liu Cheng-Rui. Fault-tolerant control for spacecraft formation with communication faults based on zero-sum differential game. *Acta Automatica Sinica*,

2025, **51**(1): 174–185

(任好, 马亚杰, 姜斌, 刘成瑞. 基于零和微分博弈的航天器编队通信链路故障容错控制. 自动化学报, 2025, **51**(1): 174–185)

- 48 Khalil H K. *Nonlinear Systems (Second edition)*. Upper Saddle River: Prentice Hall, 1996.



薛珊 海南大学信息与通信工程学院副教授. 2023 年获得华南理工大学计算机科学与工程学院博士学位. 主要研究方向为人工智能理论, 智能控制理论, 具身智能机器人.

E-mail: xueshan@hainanu.edu.cn

(**XUE Shan** Associate professor at the School of Information and Communication Engineering, Hainan University. She received her Ph.D. degree from the School of Computer Science and Engineering, South China University of Technology in 2023. Her research interests include theory of artificial intelligence, intelligent control theory, embodied intelligent robot.)



赵宁 海南大学信息与通信工程学院博士研究生. 主要研究方向为多智能体系统, 强化学习, 自适应动态规划, 海洋智能系统.

E-mail: znning@hainanu.edu.cn

(**ZHAO Ning** Ph.D. candidate at the School of Information and Communication Engineering, Hainan University. His research interests include multi-agent systems, reinforcement learning, adaptive dynamic programming, and intelligent marine systems.)



张卫东 海南大学信息与通信工程学院教授, 上海交通大学特聘教授. 1990 年、1993 年和 1996 年分别获得浙江大学学士、硕士和博士学位. 主要研究方向为控制理论, 机器学习理论及其在工业和机器人中的应用. 本文通信作者.

E-mail: wzhang@sjtu.edu.cn

(**ZHANG Wei-Dong** Professor at the School of Information and Communication Engineering, Hainan University, and distinguished professor at the Shanghai Jiao Tong University. He received his bachelor, master and Ph.D. degrees from Zhejiang University in 1990, 1993 and 1996, respectively. His research interests include control theory, machine learning theory, and their applications in industry and robotics. Corresponding author of this paper.)