

Markov 切换拓扑下非线性多智能体系统量化一致性控制

卢毅^{1,2}, 伍锡如^{1,2†}, 伍日立^{1,2}, 谢劼欣^{1,2}, 仲于海^{1,2}

(1. 桂林电子科技大学 电子工程与自动化学院, 广西 桂林 541004;

2. 智能综合自动化广西高校重点实验室, 广西 桂林 541004)

摘要: 针对受切换通信拓扑影响的非线性多智能体系统量化一致性问题, 提出一种学习型模型预测控制 (LMPC) 算法. 该算法利用神经网络实时逼近并优化 LMPC 代价函数, 在线预测最优控制增益矩阵, 有效减小通信缺陷对系统性能的影响. 同时, 结合迟滞量化器对控制输入进行量化, 缓解了网络资源受限对多智能体协同性能的限制. 为描述多智能体间的信息交换, 引入部分转移概率未知的 Markov 切换拓扑结构. 通过 Lyapunov 稳定性理论, 给出系统误差的指数一致性收敛. 最后, 通过非线性摆系统验证所提出方法的有效性和适用性.

关键词: 学习型模型预测控制; 非线性多智能体系统; Markov 切换拓扑; 神经网络; 量化一致性

中图分类号: O231.2 文献标志码: A

DOI: 10.13195/j.kzyjc.2024.1456

引用格式: 卢毅, 伍锡如, 伍日立, 等. Markov 切换拓扑下非线性多智能体系统量化一致性控制 [J]. 控制与决策, 2025, 40(10): 2933-2942.

Quantized consensus control for nonlinear multi-agent systems under Markov switching topologies

LU Yi^{1,2}, WU Xi-ru^{1,2†}, WU Ri-li^{1,2}, XIE Jie-xin^{1,2}, ZHONG Yu-hai^{1,2}

(1. School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China; 2. Key Laboratory of Intelligence Integrated Automation in Guangxi Universities, Guilin 541004, China)

Abstract: For the quantized consensus problem of nonlinear multi-agent systems affected by switching communication topologies, a learning model predictive control (LMPC) algorithm is proposed. The algorithm approximates and optimizes the LMPC cost function in real time through neural networks, and predicts the optimal control gain matrix online to effectively reduce the impact of communication defects on the system performance. Meanwhile, a hysteresis quantizer is combined to quantize the control inputs, alleviating the limitations of network resource constraints on the performance of multi-agent collaboration. To delineate the information exchange among multi-agents, the Markov switching topology with unknown partial transition probability is introduced. The exponentially consistent convergence of the systematic error is given using the Lyapunov stability theory. Finally, the effectiveness and applicability of the proposed method are verified by the nonlinear pendulum system.

Keywords: learning model predictive control; nonlinear multi-agent systems; Markov switching topology; neural network; quantized consensus

0 引言

近年来, 人们提出了非线性多智能体系统 (MASs) 来解决复杂控制系统中的协同任务, 并在移动机器人^[1]、无人机集群系统^[2]以及无线传感网络^[3]等领域得到了广泛应用. 作为 MASs 协同控制的核心问题之一, 一致性问题研究多智能体间如何通

过有限的信息交换实现状态的一致性, 这对于确保 MASs 的整体性能和稳定性至关重要. 为此, 学者们从不同角度对 MASs 一致性问题展开了深入研究, 包括有限/固定时间一致性^[4]、事件触发一致性^[5]以及 H_∞ 一致性^[6]等. 然而, 随着网络通信技术的快速发展, MASs 一致性问题在通信带宽、时延和拓扑结

收稿日期: 2024-12-16; 录用日期: 2025-04-11.

基金项目: 国家自然科学基金项目 (62263005); 广西高校人工智能与信息处理重点实验室 (河池学院) 开放基金项目 (2022GXZDSY004, 2024GXZDSY013); 桂林电子科技大学研究生教育创新计划项目 (2025 YCXS138).

责任编辑: 关新平.

†通信作者. E-mail: xiruwu@guet.edu.cn.

构动态变化等方面面临着新的挑战. 因此, 本文旨在进一步探讨通信网络受限情况下 MASs 的一致性控制问题, 以期对相关理论研究和实际应用提供新的思路和方法.

模型预测控制 (MPC) 凭借其预测特性和优化控制能力, 在实际工程应用中广受欢迎. 在 MPC 框架下, 控制问题通常被表述为动态优化问题. 近年来涌现出多种优化算法用于求解 MPC 问题^[7-8]. 在 MASs 一致性控制领域, MPC 方法已取得显著进展, 如针对轮式机器人系统, 文献 [1] 研究了具有拓扑切换结构的 MASs 一致性问题; 文献 [2] 提出了一种基于事件触发的 MPC 控制策略, 有效降低了具有干扰的非线性系统的网络资源消耗; 文献 [8] 采用交替方向乘子法求解 MPC 优化问题, 并通过多项物理实验验证了算法的有效性; 文献 [9] 开发了基于 Lyapunov 的 MPC 算法, 成功应用于非线性网络系统; 文献 [10] 则针对具有复杂动力学特性的非线性系统, 提出了一种数据驱动 MPC 策略, 实现了系统的实时控制. 然而, 尽管现有 MPC 算法在固定控制增益方面取得了显著成果, 但在处理系统未知参数和非线性函数估计方面仍存在局限性. 近年来, 神经网络技术因其在处理未知参数方面的灵活性和非线性函数逼近能力而受到广泛关注. 这促使神经网络与 MPC 结合应用于在线策略一致性问题, 如文献 [5] 采用并行学习技术更新神经网络权值, 解决了具有不确定性的非线性 MASs 鲁棒最优一致性问题; 文献 [11] 提出了一种受随机 Jacobi 交替方向乘子法启发的并行算法; 文献 [12] 成功将神经网络与 MPC 算法相结合, 有效解决了非线性 MASs 的一致性问题. 这些突破性研究为处理复杂 MASs 一致性问题提供了新的思路, 也构成了本文研究的重要动机.

在实际应用中, 网络通信的实时性、准确性和完整性往往受到带宽受限的严重影响, 这使得非线性 MASs 的量化一致性问题成为当前研究的热点, 并引起了学者们的广泛关注^[6,13]. 早期研究中, 文献 [14] 开创性地提出了一种整数值状态的平均一致性算法, 实现了多智能体状态向整数近似值的渐近收敛; 文献 [15] 深入分析了不同量化器对多机器人系统同步性能的影响; 文献 [16] 将二分实用一致性问题转化为一般实用一致性问题, 并提出了一种基于结构平衡拓扑假设的量化分布式控制策略; 文献 [17] 通过非光滑分析方法, 在 Filippov 意义下给出了受量化影响的 MASs 一致性问题的系统解; 文献 [18] 设计了具有标度函数的均匀量化器. 为了进一步降低通信负担, 文献 [19] 通过将量化控制与事件触发控制

相结合, 研究了具有外部干扰的 MASs 有界一致性问题, 并证明了通过参数调节可有效控制误差有界性; 文献 [20] 提出了一种可调节的自适应量化策略, 成功实现了异构 MASs 的一致性跟踪. 尽管量化策略研究已取得显著进展, 但在量化策略与 MPC 控制算法的融合方面仍存在研究空白. 现有方法在处理非线性 MASs 量化一致性问题时, 往往难以同时兼顾控制精度和通信效率. 因此, 本文旨在提出一种新型的融合量化策略的 MPC 控制算法, 以期在保证系统性能的同时显著降低通信成本, 为非线性 MASs 量化一致性问题的解决提供新的思路和方法.

值得注意的是, 上述研究中多智能体系统通常需要持续通信, 这使得考虑具有通信拓扑结构的 MASs 在实际应用中显得尤为重要. 文献 [21] 首次将 Markov 过程引入切换拓扑控制, 揭示了拓扑切换对一致性的影响与 Markov 过程正循环状态相关的拓扑网络密切相关; 文献 [22] 在随机切换拓扑下, 成功解决了具有扰动的二阶 MASs 固定时间一致性问题; 文献 [23] 进一步将研究扩展到切换拓扑下的非线性 MASs 固定时间一致性. 文献 [24] 则提出了一种适用于切换拓扑结构的自适应神经网络一致性控制策略, 展现了良好的控制性能. 然而, 上述大多数研究中的切换拓扑结构主要基于完全已知的转移概率信息. 在实际应用中, 获取所有的转移概率信息往往难以满足, 这一局限性严重制约了现有方法在实际工程中的应用. 因此, 本文针对部分转移概率未知的 Markov 切换拓扑下非线性 MASs 的量化一致性问题, 提出一种学习型 MPC (LMPC) 控制算法. 本文的主要工作如下:

- 1) 提出一种在线学习优化的 LMPC 控制算法. 该算法通过实时更新网络参数逼近 LMPC 代价函数, 显著增强了系统对不确定性和外部扰动的鲁棒性.

- 2) 设计一种集成迟滞量化器的反馈控制策略. 与文献 [12, 20] 中的方法相比, 不仅有效降低了多智能体间的通信负担, 还通过引入迟滞机制避免了信号颤振现象.

- 3) 将自适应神经网络引入 LMPC 框架, 显著提升系统对未知扰动的适应能力. 得益于神经网络的强大估计能力, 所提出的控制协议能够适用于更广泛的非线性系统, 为具有 Markov 切换拓扑的非线性 MASs 一致性研究开辟了新的方向.

- 4) 构建了模态依赖的 Lyapunov 函数, 并通过 Lyapunov 稳定性理论分析系统一致性误差的指数收敛性.

1 图论知识及问题描述

符号说明: $R^{N \times N}$ 表示 $N \times N$ 的矩阵, R^N 表示 N 维向量, \Leftrightarrow 为等值符号, \subseteq 是包含于符号, \in 是属于符号, $\|\cdot\|$ 为 L_2 范数, \otimes 为 Kronecker 直积, $\mathbf{1}_N$ 是 N 项 1 组成的向量, $|\cdot|$ 表示绝对值.

1.1 图论知识

MASs 的通信拓扑结构表示为一个有向图 $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$, 其中 $\mathcal{V} = \{v_1, \dots, v_N\}$, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ 和 $\mathcal{A} = [a_{ij}] \in R^{N \times N}$ 分别表示节点集、边集和邻接矩阵. 有向图 \mathcal{G} 对应的拉氏矩阵 $\mathcal{L} = [l_{ij}] \in R^{N \times N}$ 描述为

$$l_{ii} = \sum_{j \in N} a_{ij}, l_{ij} = -a_{ij}, i \neq j.$$

考虑到具有切换拓扑的系统, 该图可进一步描述为 $\mathcal{G}(\theta(k)) = (\mathcal{V}, \mathcal{E}(\theta(k)), \mathcal{A}(\theta(k)))$, 拉氏矩阵 $\mathcal{L}(\theta(k))$ 重构为 $l_{ii}(\theta(k)) = \sum_{j \in N} a_{ij}(\theta(k)), l_{ij}(\theta(k)) = -a_{ij}(\theta(k)), i \neq j$, 其中 $\theta(k) \in \mathfrak{S} = \{1, 2, \dots, S\}$ 表示 k 时刻的通信拓扑模式, 服从于马尔科夫链的切换模式. 部分转移概率未知的状态转换矩阵定义为

$$\Pi = \begin{bmatrix} ? & \pi_{12} & \dots & ? \\ ? & ? & \dots & ? \\ \vdots & \vdots & \ddots & \vdots \\ ? & ? & \dots & \pi_{SS} \end{bmatrix}. \quad (1)$$

其中: “?” 表示转移概率矩阵是未知的; π_{pq} 为从 p 转移到 q 的概率, 且满足 $\sum_{q=1}^S \pi_{pq} = 1, p, q \in \mathfrak{S}$.

1.2 问题描述

非线性 MASs 的控制结构如图 1 所示.

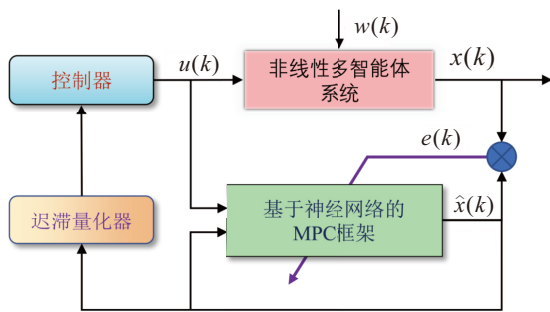


图1 非线性 MASs 量化一致性控制框图

考虑由 N 个子系统组成的非线性离散 MASs, 其动力学模型为

$$x_i(k+1) = Ax_i(k) + Bu_i(k) + f(k, x_i(k)) + Dw_i(k). \quad (2)$$

其中: A, B, D 是已知矩阵; $x_i(k) \in R^n$ 和 $u_i(k) \in R^m$ 分别为第 i 个智能体的系统状态和控制输入; $w(k) \in R^n$ 为定义在概率空间 $(\Omega, \mathcal{F}, \mathcal{P})$ 上的零均

值高斯白噪声序列, 其方差为 v ; $f(\cdot)$ 是关于 $(k, x(k))$ 的非线性函数, 且满足如下假设.

假设 1 非线性函数 $f(\cdot)$ 满足如下 Lipschitz 条件:

$$\|f(x) - f(y)\| \leq \alpha \|x - y\|. \quad (3)$$

其中: α 为非负常数; $x, y \in R^n$.

由于迟滞型量化器能有效减少振荡现象, 并确保量化值在切换过程中更加平稳, 本文考虑离散化的迟滞量化器^[6]

$$\Omega(u) = \begin{cases} u_j, & \frac{u_j}{1+\delta} < u \leq u_j, \Omega' \geq u_j, \text{ or} \\ & u_j \leq u < \frac{u_j}{1-\delta}, \Omega' \leq u_j; \\ (1+\delta)u_j, & u_j < u \leq \frac{u_j}{1-\delta}, \Omega' \geq (1+\delta)u_j, \text{ or} \\ & \frac{u_j}{1-\delta} \leq u < u_{j+1}, \Omega' \leq (1+\delta)u_j; \\ 0, & 0 \leq u \leq \frac{u_0}{1+\delta}, \text{ or } \frac{u_0}{1+\delta} < u < u_0, \Omega' = 0; \\ -\Omega(-u), & u < 0. \end{cases} \quad (4)$$

其中: $u_j = \rho^{1-j}u_0, j = 1, 2, \dots, n$, 参数 $\rho = \frac{1-\delta}{1+\delta}$ ($0 < \rho < 1$) 决定量化密度.

注 1 为了避免多智能体间网络传输容量的有限性, 本文对控制输入信号进行量化. 对于离散的输入信号, 迟滞量化器可以有效避免离散输入信号变化时的抖动. 鉴于此特性, 迟滞量化器更切合真实系统.

根据扇形有界法^[13], 可以用如下方式表示量化误差:

$$\Omega(\kappa) = (1 + \Delta_\kappa)\kappa, \quad (5)$$

其中 $\Delta_\kappa = \text{diag}\{\Delta_1(\kappa), \dots, \Delta_n(\kappa)\}$ 和 $\|\Delta_i(\kappa)\| \leq \delta, i \in \{1, 2, \dots, n\}$.

为了优化多智能体间的网络资源, 本文提出了一种量化反馈控制器

$$u_i(k) = -K_k \sum_{j=1}^N a_{ij}(\theta(k)) \epsilon_{ij}(k). \quad (6)$$

其中: $K_k \in R^m$ 为控制器增益矩阵, $\epsilon_{ij}(k) = \Omega(x_j(k)) - \Omega(x_i(k))$.

为了便于进一步讨论, 叠加单个控制信号得到控制输入向量 $u(k) = [u_1(k) \dots u_N(k)]^T$. 同理, $x(k) = [x_1(k) \dots x_N(k)]^T$. 结合式 (4) 和 (5), 控制输入可重新表述成如下向量形式:

$$u(k) = -(\mathcal{L}_{\theta(k)} \otimes K_k)\Omega(x(k)) = -(\mathcal{L}_{\theta(k)} \otimes K_k)(1 + \Delta_k)x(k). \quad (7)$$

令 $p = \theta(k)$, 则 MASs 的系统状态向量可描述为

$$\begin{aligned} x(k+1) = & Ax(k) + B(\mathcal{L}_p \otimes K_k)(1 + \Delta_k)x(k) + \\ & f(k, x(k)) + Dw(k). \end{aligned} \quad (8)$$

虚拟 MASs 的状态方程定义如下:

$$x_0(k+1) = Ax(0) + f(0, x(0)). \quad (9)$$

令 $e_i(k) = x_i(k) - x_0(k)$, 则误差的向量形式描述为

$$\begin{aligned} e(k+1) = & Ae(k) - B(\mathcal{L}_p \otimes K_k)(1 + \Delta_k)e(k) + \\ & f(k, e(k)) + Dw(k). \end{aligned} \quad (10)$$

定义 1 在零初始条件下, 如果存在标量 a 和 $\eta \in (0, 1)$, 使得

$$\|e(k, w(k))\|^2 \leq \eta^k \|e(0, w(0))\|^2 + \varpi \quad (11)$$

成立, 则称 MASs (2) 实现均方指数一致性, 其上界为 ϖ .

2 基于神经网络的学习型模型预测控制

首先, 给出非线性 MASs 实现指数一致性收敛的证明; 然后, 设计一种可在线学习的 LMPC 控制算法, 以找到一个最优的控制增益序列.

2.1 一致性分析

定理 1 假设存在控制增益序列 $\{K_k\}_{k=0}^\infty$ 和常数 $\alpha < 1$. 若存在常数 $\xi, \chi, \beta, \gamma > 1$ 以及正定矩阵 P 满足

$$\begin{bmatrix} P & A - B(\mathcal{L}_p \otimes K_k)(1 + \Delta_k) + a \\ * & -P^{-1} \end{bmatrix} < 0, \quad (12)$$

则在量化反馈控制器 (5) 的作用下, 非线性 MASs(2) 可实现指数一致性.

证明 给定如下的 Lyapunov 函数:

$$V(k) = e^T(k)P_p e(k), \quad (13)$$

其中 $P_p (p \in \mathfrak{S})$ 表示与 Markov 模态相关的正定对称矩阵. 为了量化分析通信拓扑结构对系统性能的影响, 设计如下矩阵:

$$P_p = \frac{1}{2}(\mathcal{L}(p) + \mathcal{L}^T(p)). \quad (14)$$

综合式 (9), 可得

$$\begin{aligned} \nabla V(k) = & V(k+1) - V(k) = \\ & e^T(k+1)P_p e(k+1) - e^T(k)P_p e(k) = \\ & (Ae(k) + B(\mathcal{L}_p \otimes K_k)(1 + \Delta_k)e(k) + \\ & f(k, e(k)) + Dw(k))^T P_p (Ae(k) + \end{aligned}$$

$$\begin{aligned} & (Ae(k) + B(\mathcal{L}_p \otimes K_k)(1 + \Delta_k)e(k) + \\ & f(k, e(k)) + Dw(k))^T P_p (Ae(k) + \\ & B(\mathcal{L}_p \otimes K_k)(1 + \Delta_k)e(k) + Dw(k) + \\ & f(k, e(k))) - e^T(k)P_p e(k). \end{aligned} \quad (15)$$

由假设 1 可得

$$\begin{aligned} \nabla V(k) \leq & (Ae(k) - B(\mathcal{L}_p \otimes K_k)(1 + \Delta_k)e(k) + \\ & \alpha e(k) + Dw(k))^T P_p (Ae(k) + Dw(k) - \\ & B(\mathcal{L}_p \otimes K_k)(1 + \Delta_k)e(k) + \alpha e(k)) - e^T(k)P_p e(k) \leq \\ & e^T(k)\Omega^T P_p \Omega e(k) + \lambda_{\max}(D^T P D)v - e^T(k)P_p e(k) = \\ & e^T(k)\Omega^T P_p \Omega e(k) - e^T(k)P_p e(k) + \xi v, \end{aligned} \quad (16)$$

其中 $\Omega = A - B(\mathcal{L}_p \otimes K_k)(1 + \Delta_k) + a$.

利用舒尔补条件和式 (12), 可以进一步得到

$$\nabla V(k) \leq -\rho e^T(k)e(k) + \xi v, \quad (17)$$

其中 $\rho > 0$.

根据式 (13), 可得

$$V(k) \leq \lambda_{\max}(P)e^T(k)e(k) = \beta e^T(k)e(k), \quad (18)$$

$$V(k) \geq \lambda_{\min}(P)e^T(k)e(k) = \chi e^T(k)e(k). \quad (19)$$

其中: $\beta = \lambda_{\max}(P), \chi = \lambda_{\min}(P)$. 结合式 (17) 和 (18), 对于任意 $\gamma > 1$, 可推导出如下等式成立:

$$\begin{aligned} \gamma^{k+1}V(k+1) - \gamma^k V(k) = & \gamma^{k+1}(V(k+1) - V(k)) - \gamma^k(\gamma - 1)V(k) = \\ & -\rho\gamma^{k+1}e^T(k)e(k) + \gamma^k(\gamma - 1)e^T(k)e(k) + \gamma^{k+1}\xi v = \\ & \gamma^k\alpha(\gamma)e^T(k)e(k) + \gamma^{k+1}\xi v, \end{aligned} \quad (20)$$

其中 $\alpha(\gamma) = -\rho\gamma + (\gamma - 1)\beta$. 将式 (20) 左右两边从 0 到 $T - 1$ 相加, 整理可得

$$\begin{aligned} \gamma^T V(k) - V(0) \leq & \alpha(\gamma) \sum_{k=0}^T \gamma^k \|e(k)\|^2 + \frac{\gamma(\gamma^T - 1)}{\gamma - 1} \xi v. \end{aligned} \quad (21)$$

根据式 (18), 可得

$$V(0) \leq \beta \|e(0)\|^2. \quad (22)$$

将式 (22) 和 (19) 代入 (21), 可得

$$\begin{aligned} \|e(k)\|^2 \leq & \frac{\alpha(\gamma)\gamma^{-T}}{\chi} \sum_{k=0}^T \gamma^k \|e(k)\|^2 + \\ & \frac{\beta\gamma^{-T}}{\chi} \|e(0)\|^2 + \frac{\gamma(1 - \gamma^{-T})}{(\gamma - 1)\chi} \xi v. \end{aligned} \quad (23)$$

显然, $\lim_{\gamma \rightarrow \infty} \alpha(\gamma) \rightarrow 0$. 因此, 可进一步得到

$$\begin{aligned} \|e(k)\|^2 \leq & \frac{\beta\gamma^{-T}}{\chi} \|e(0)\|^2 + \frac{\gamma(1 - \gamma^{-T})}{(\gamma - 1)\chi} \xi v \leq \end{aligned}$$

$$\frac{\beta\gamma^{-T}}{\chi} \|e(0)\|^2 + \frac{\gamma}{(\gamma-1)\chi} \xi v. \quad (24)$$

式 (24) 满足定义 1, 从而确保了 MASs 能够实现指数一致性. □

注 2 值得注意的是, 在这一类非线性 MASs 系统中, 寻求一个可信的控制增益序列并非易事. 因此, 接下来利用 LMPC 的在线学习算法来预测控制增益序列的最优方案, 从而使得假设 1 成立, 非线性 MASs 系统能够满足定理 1 并实现指数一致性.

2.2 模型预测控制

根据式 (13), 代价函数可定义为

$$J(e(k)) = \sum_{t=k+1}^{k+h+1} E\{e^T(k)P_p e(k)\}. \quad (25)$$

LMPC 的目标是找到使代价成本 $J(e(k))$ 最小的 J , 则最优的反馈增益矩阵 K_k^* 满足以下条件:

$$K_k^* = \arg \min_{K_k \dots K_{k+h}} J(e(k)). \quad (26)$$

注 3 考虑到通信拓扑结构的动态切换, 代价函数会受到 Markov 切换过程的影响, 这使得 LMPC 在求解 MASs 一致性问题时面临重大挑战. 鉴于这一难题, 本文提出一种基于神经网络的 LMPC 预测算法.

2.3 LMPC 算法

由于神经网络能够很好地逼近非线性函数, 本文采用神经网络来近似代价函数 $J(e(k))$ 与控制增益矩阵序列 $\{K_k \dots K_{k+h}\}$ 之间的关系.

本文旨在使用浅层神经网络解决优化问题 (26). 浅层神经网络输入层由长度为 $l (l = nm)$ 的序列 $K_k = [K_k^{ab}]$ 构成, 其中 $a \in \{1, 2, \dots, n\}$ 和 $b \in \{1, 2, \dots, m\}$. 为了训练得到最优的增益值 K_k , 选择一个具有 M 个神经元的单层隐藏层. 输入层由代价函数 $J(e(k))$ 的值组成.

令 $z = [1 \ K_k^{11} \ K_k^{1n} \ K_k^{m1} \ K_k^{nm}]^T \in R^{l+1}$, 隐藏层产生的信号由下式给出:

$$w = \sigma(Wz). \quad (27)$$

其中: σ 表示 sigmoid 函数, $W \in R^{M \times (l+1)}$ 表示神经元的权重矩阵. 对于给定的增益 K_k , 成本函数的输出层 \tilde{J} 可近似为

$$\tilde{J} = Yw, \quad (28)$$

其中 $Y \in R^{1 \times M}$ 表示输出层的权重矩阵. 神经网络在迭代学习的过程中不断优化权重 W 和 Y , 以便更好地拟合出控制增益.

算法 1 LMPC 控制算法. 初始阶段, 系统给定初始条件 $x(0)$ 、模态 $\theta(0)$ 和代价函数 J_0 . 为了方便起

见, 初始增益矩阵 K_0 设为空值, 与 J_0 共同构成神经网络的初始训练数据集. 神经网络通过迭代学习预测代价函数 \tilde{J} : 在每次迭代中, 基于上一次增益 K_k 生成不同的组合, 通过 $\pm\delta$ 微调 K_k^{nm} 以提高预测的精确度, 选取 \tilde{J} 最小值对应的增益 K_k 作为最优控制参数, 并计算控制输出 $u(k)$.

算法 1 LMPC 算法.

- 1) 初始化参数: $x(0), \theta(0), K_0^{nm} = 0, \hat{J} = J_0, \xi$
- 2) 初始化数据集: $[K_0, J_0]$
- 3) for $k = 0 : k_{\max}$ do
- 4) while $\hat{J} \leq \xi$ do
- 5) 训练数据集 $\rightarrow (W, Y)$
- 6) $\tilde{J} = Y\sigma(Wz) \forall K_k^{nm} + \delta_K$
- 7) $\tilde{K}_k^* = \arg \min_{K_k} \tilde{J}$
- 8) $u(k) = (\mathcal{L}_{\theta(k)} \otimes \tilde{K}_k^*)(1 + \Delta_k)x(k)$
- 9) $\hat{x}(k) = x(k), \hat{\theta}(k) = \theta(k)$
- 10) for $t = k : k + h + 1$ do
- 11) $p = \hat{\theta}(t)$
- 12) $\hat{u}(t) = (\mathcal{L}_p \otimes \tilde{K}_k^*)(1 + \Delta_t)\hat{x}(t)$
- 13) $\hat{x}(t+1) = F_p(\hat{x}(t))$
- 14) $x_0(t+1) = f(x_0(t), 0)$
- 15) $\hat{e}(t) = F_p(x(t)) - 1_N \otimes f(x_0(t), 0)$
- 16) $\hat{V}_t = \hat{e}^T(t)P_p\hat{e}(t)$
- 17) 随机更新模态信息 $\hat{\theta}(t) \rightarrow \hat{\theta}(t+1)$
- 18) end for
- 19) $\hat{J} = \sum_{t=k+1}^{k+h+1} \hat{V}_t$
- 20) 更新数据集: $[\tilde{K}_k^*, \hat{J}]$
- 21) end while
- 22) $K_k = \tilde{K}_k^*$
- 23) $x(k+1) = F_p(x(k))$
- 24) $x_0(k+1) = f(x_0(k), 0)$
- 25) end for

为了优化 \hat{V}_t 的值, 算法采用 LMPC 循环预测机制:

- 1) 根据前一时间状态及概率转移矩阵 Π 生成模态 $\theta(t)$;
- 2) 计算预测控制输入 $\hat{u}(t)$ 和系统状态 $\hat{x}(t)$;
- 3) 估算近似值 \hat{V}_t .

每轮迭代结束后, 累加 \hat{V}_t 得到成本函数近似值 \hat{J} , 并将 \hat{J} 与最优增益 \tilde{K}_k^* 存储在数据集中, 作为训练集以供下一步迭代. 重复上述步骤, 直至预测值 \tilde{J} 达

到预设阈值 ξ , 终止训练并输出优化结果.

注4 最速下降法作为机器学习中常用的一种迭代优化算法, 因其较低的复杂度和简单的操作而被应用, 这也适用于处理本文所提出的基于神经网络的算法应用. 因此, 本文选择最速下降法寻求局部最优点. 最速下降法步长 δ_K 的选择已有大量文献研究, 并取得了不错的成果. 但是, 本文所提出的基于神经网络的控制问题并不适用. 因此, 本文中步长 δ_K 采用固定值.

注5 空间复杂度主要来源于分布式稀疏存储策略的 $O(N \lg N)$ 和状态向量存储的 $O(nN)$, 其中通过限制平均邻居数为 $\lg N$ 避免了全连接图的 $O(N^2)$ 复杂度, 神经网络在线学习模块的梯度计算引入 $O(mnN)$ 的空间开销; 时间复杂度由神经网络计算的 $O(MT)$ 、矩阵运算的 $O(mnN^2)$ 和 MPC 的 $O(hN^2)$ 组成, 二次项源于多智能体间的协同计算.

3 仿真实例

为了验证所提出 LMPC 控制算法的有效性和实用性, 本文在不同场景下做出了模型比较.

对于每个系统, 隐藏层中的神经元数量是根据交叉验证过程选择的. 训练停止参数 ξ 选取为 0.01, LMPC 预测训练次数 h 为 100. 为了便于对比, 本文给出了稳定时间的具体评价指标, 即误差 SE_i^r 达到 2% 时的时间. 误差 SE_i^r 通过下式给出:

$$SE_i^r = \sum_{k=0}^N |e_i^r(k)|^2, \quad (29)$$

其中 $e_i^r(k)$ 是误差向量 $e(k)$ 的第 r 项.

设置不同的概率转移矩阵 $\Pi_s (s \in 1, 2, 3, 4, 5, 6)$, 进一步验证所提出模型的鲁棒性. 概率转移矩阵 $\Pi_s (s \in 1, 2, 3, 4, 5, 6)$ 如下所示:

$$\begin{aligned} \Pi_1 &= \begin{bmatrix} 0.05 & 0.95 \\ 0.98 & 0.02 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix}, \\ \Pi_3 &= \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}, \quad \Pi_4 = \begin{bmatrix} ? & ? \\ 0.5 & 0.5 \end{bmatrix}, \\ \Pi_5 &= \begin{bmatrix} 0.95 & 0.05 \\ ? & ? \end{bmatrix}, \quad \Pi_6 = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}. \end{aligned}$$

3.1 数值仿真

考虑如下的连续时间动力学系统^[12]:

$$\dot{x}_i(t) = Ax_i(t) + Bu_i(t) + f(t, x_i(t)) + Dw_i(t), \quad (30)$$

其中 $w_i(t)$ 是外部扰动, $i = 1, 2, 3, 4$. 其系统参数为 $f(t, x_i(t)) = 0.01 \sin(x_i(t))$, $A = \begin{bmatrix} 0 & 1 \\ 0 & -0.5 \end{bmatrix}$, $B = \begin{bmatrix} 0.8 \\ 1.2 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $w_i(t) = \begin{bmatrix} 0.02 \sin t \\ 0.02 \cos t \end{bmatrix}$.

系统初始状态 $x_0(0) = [3 \ -1]$, $x_1(0) = [3 \ -1]$, $x_2(0) = [1 \ -1]$, $x_3(0) = [1 \ -3]$, $x_4(0) = [3 \ -3]$. 采用 Euler 离散方法, 其采样周期为 $T = 0.01s$. 量化密度 ρ 为 0.3. 概率转移矩阵选取为 Π_5 . Laplacian 矩阵为

$$\begin{aligned} \mathcal{L}_1 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}, \\ \mathcal{L}_2 &= \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}. \end{aligned}$$

根据文献 [12], 神经元数量 M 取值为 10. 通过训练得到最优增益 \tilde{K}_k^* 为 $[13.25 \ 1.50]$. 图 2 展示了通信拓扑模式 $\theta(k)$ 的切换过程. 图 3 对比了不同控制方法下的系统状态均值 $x_i(k)$ 的演化过程 (10 次仿真). 其中方法 1 是文献 [12] 所提出的非量化控制方法, 方法 2 是采用文献 [20] 的对数量化器处理控制输入. 图 4 展示了控制输入 $u_i(k)$ 及其平均值的曲线 (10 次仿真). 所有方法均能使系统状态和控制输入在有限时间内收敛至稳态邻域. 显然, 在本文所提出方法下系统响应速度更快速. 图 5 为量化前后系统状态 $x_{21}(k)$ 与 $x_{22}(k)$ 的曲线对比 (以系统状态 $x_{21}(k)$)

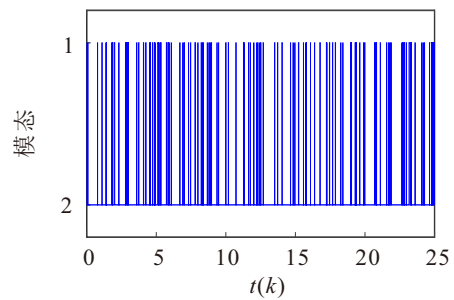


图2 通信拓扑模式 $\theta(k)$ 的切换过程

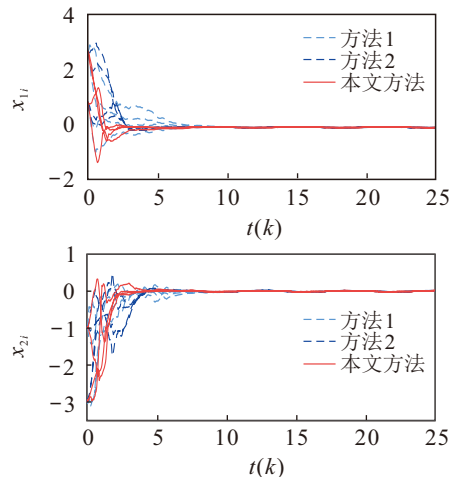


图3 不同控制方法下的系统状态 $x_i(k)$ 平均值对比 (10 次仿真)

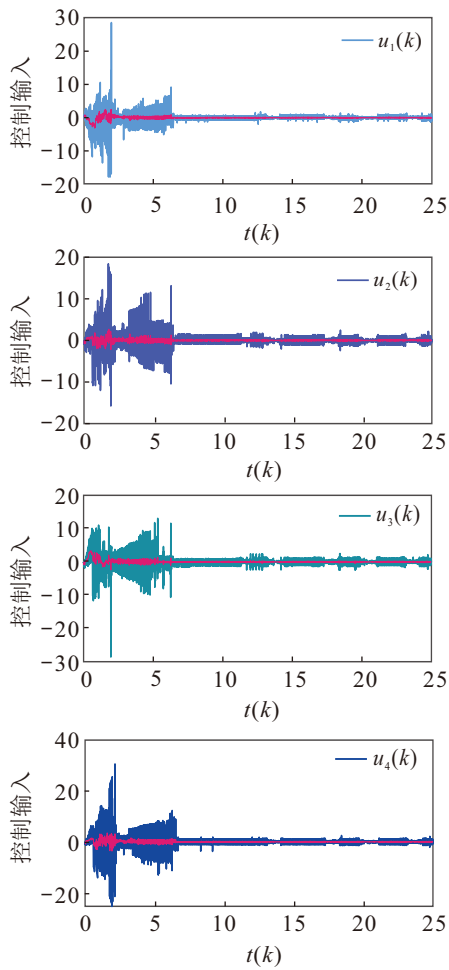


图4 控制输入 $u_i(k)$ 及其平均值 (10 次仿真)

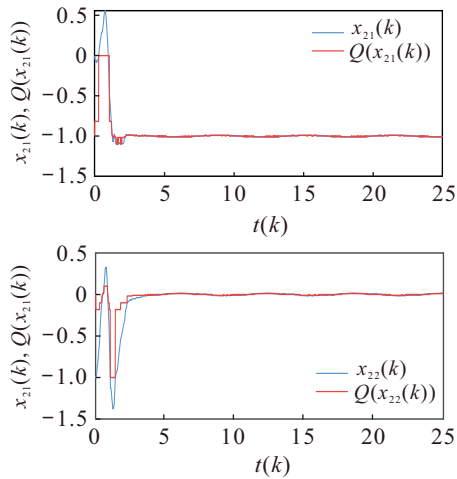


图5 量化前后系统状态 $x_{21}(k)$ 和 $x_{22}(k)$ 的对比曲线

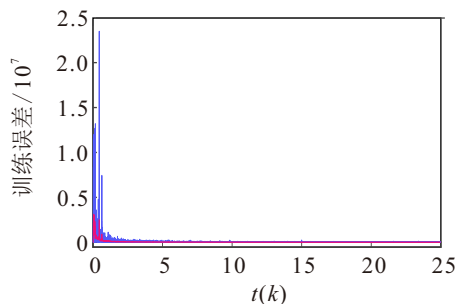


图6 训练误差曲线及其平均值 (10 次仿真)

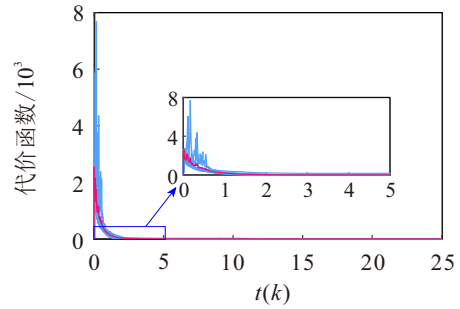


图7 代价函数 \hat{J} 的预测值及其平均值 (10 次仿真)

和 $x_{22}(k)$ 为例). 量化系统依然能够准确捕捉原系统的动态特征. 图 6 给出了神经网络训练过程中的误差曲线. 图 7 呈现了预测值 \hat{J} 及均值曲线. 综合上述分析, 本文提出的 LMPC 控制方法在系统响应速度、误差收敛性以及稳定性方面均展现出了显著的优势. 特别是在量化处理后, 系统状态的准确性得到了有效的保持, 同时预测值的快速衰减也验证了算法的高效性.

表 1 给出了不同状态转移概率下系统稳定时间 (误差 SE_i^r 在 2% 以内) 对比分析. 可以看出, 在不同的转移概率矩阵 $\Pi_s (s \in 1, 2, 3, 4, 5, 6)$ 下, 本文所提出的控制方法大多数情况下都优于其他两种方法, 尤其是在状态转移概率存在未知的情况下. 这表明, 本文提出的方法在量化控制输入方面更为有效, 能够在不同状态转移概率下迅速达到稳定状态, 从而实现更优的控制效果. 然而, 在某些状态转移概率下, 如 Π_2 , 方法 2 也能展现出优秀的效果. 总体而言, 本文提出的控制方法在处理含有部分状态转移概率未知的 Markov 切换拓扑时, 显示出了更强的鲁棒性和适应性.

表1 不同状态转移矩阵下系统稳定时间对比

	方法1	方法2	本文提出的控制方法
Π_1	2.72	1.65	1.23
Π_2	3.73	1.85	2.43
Π_3	3.23	2.73	1.42
Π_4	5.18	3.67	3.09
Π_5	2.87	1.95	2.73
Π_6	6.75	4.17	3.16

3.2 实例仿真

为了说明本文所提出方法的实用性, 将 LMPC 算法应用于非线性摆系统^[25]

$$\begin{aligned} \dot{x}_i(t) &= v_i(t), \\ \dot{v}_i(t) &= -\sin(x_i(t)) - 0.25v_i(t) + 0.15 \cos(2.5t) + \sin(v_i(t))v_i^2(t) + u_i(t). \end{aligned}$$

可以进一步表述为

$$\dot{\zeta}_i(t) = A\zeta_i(t) + Bu_i(t) + h(t, \zeta_i(t)).$$

其中

$$\begin{aligned} \zeta_i(t) &= [x_i(t), v_i(t)]^T, \\ h(t, \zeta_i(t)) &= \\ &[0, -\sin(x_i(t)) - 0.25v_i(t) + \\ &\sin(v_i(t))v_i^2(t) + 0.15\cos(2.5t)]^T, \\ B &= [0, 1]^T, A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

非线性摆系统的初始化参数为 $x_0(0) = 0, v_0(0) = 0, x_1(0) = 50\pi/13, v_1(0) = -4.5, x_2(0) = 12\pi/13, v_2(0) = 2.2, x_3(0) = -50\pi/13, v_3(0) = 4.5, x_4(0) = 24\pi/13$, 和 $v_4(0) = -2.2$. 摆系统的耦合拓扑结构矩阵参考实验 3.1, 概率转换矩阵选取为 Π_6 .

神经元数量 M 选取为 19, 量化密度 ρ 为 0.1. 通过训练得到最优增益 \tilde{K}_k^* 为 [57.00 57.75]. 图 8 展示了通信拓扑模式 $\theta(k)$ 的切换过程; 图 9 分别描述了非线性摆系统的位置和速度的状态曲线; 图 10 为非线性摆系统量化后的系统控制输入; 图 11 展示了非线性摆系统在 3 种不同控制方法下的系统误差 $e_i^T e_i$; 图 12 给出了非线性摆系统在神经网络训练过程中的误差曲线. 结果表明, 本文提出的控制方法在非线性的误差控制方面具有较好的性能, 显示出良好的稳定性和控制效果.

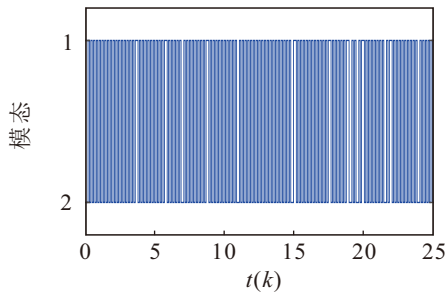


图8 非线性摆系统通信拓扑模式 $\theta(k)$ 的切换过程

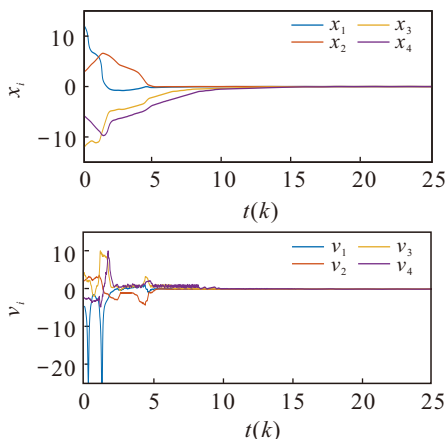


图9 非线性摆系统位置 x_i 和速度 v_i 的状态曲线

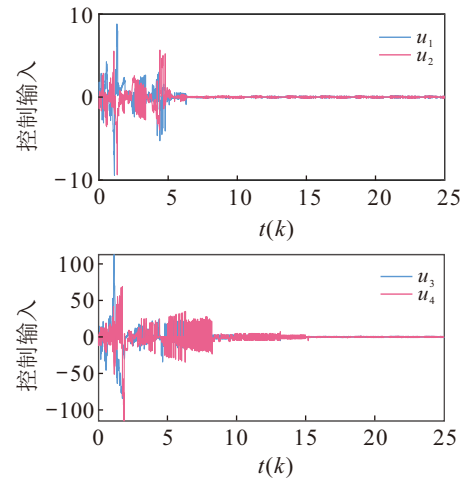


图10 非线性摆系统量化后的系统控制输入 u_i

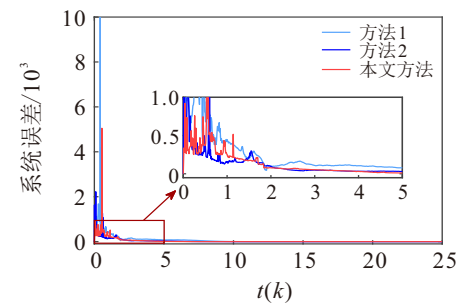


图11 非线性摆系统的系统误差 $e_i^T e_i$

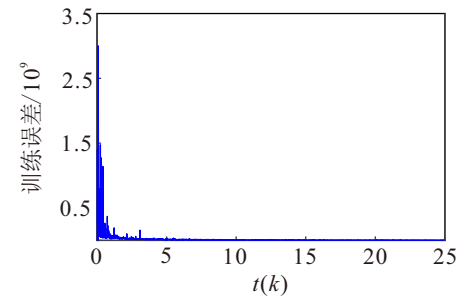


图12 非线性摆系统的训练误差曲线

对 3 种不同控制方法在不同状态转移矩阵 $\Pi_s (s \in 1, 3, 4, 5, 6)$ 下非线性摆系统的稳定时间 (误差 SE_i^T 在 4% 以内) 进行比较, 如表 2 所示. 可以观察到, 所有控制方法均能使非线性摆系统达到稳定状态. 本文提出的控制方法在不同状态转移矩阵下均展现出了优越的性能, 尤其是在状态转移矩阵中具有较多的未知元素时, 仍然具有较短的稳定时间, 显示出其对非线性系统的显著控制效能. 然而, 随着状态转移矩阵中未知元素的增多, 系统的稳定时间呈现出逐渐增长的趋势. 这是由于未知元素增加了系统的不确定性, 使得控制策略需要更多时间来适应和调整. 尽管如此, 本文提出的控制方法对于处理具有强非线性特性的系统仍然有效, 并且能够在面对不确定性和非线性挑战时保持鲁棒性.

表2 不同状态转移矩阵下非线性摆系统稳定时间对比

	方法1	方法2	本文提出的控制方法
Π_1	10.03	4.71	4.77
Π_2	11.16	8.42	7.27
Π_3	13.96	10.80	5.84
Π_4	13.45	9.21	7.16
Π_5	24.83	10.92	10.80
Π_6	17.21	12.69	12.20

4 结论

本文提出了一种可在线学习的 LMPC 算法, 针对非线性 MASs 的量化一致性问题进行了深入研究. 通过神经网络的实时自适应调整, LMPC 算法有效地实现了代价函数的最小化, 优化了反馈控制矩阵. 此外, 设计了一种量化反馈控制器, 显著提高了多智能体间的传输效率, 证明了非线性 MASs 可实现指数一致性收敛. 通过在不同非线性系统中的测试, 所提出的 LMPC 控制算法在响应速度和稳定时间方面均优于现有文献中的方法.

未来的研究将重点突破当前算法的局限性, 通过开发轻量化分布式学习架构提升大规模智能体系统的适用性, 结合时变拓扑与自适应机制增强动态环境下的鲁棒性, 同时研究非理想通信条件下的稳定性保障方法, 从而推动多智能体一致性控制在复杂场景中的实际应用.

参考文献 (References)

- [1] Sun F L, Li H, Zhu W, et al. Fixed-time formation tracking for multiple nonholonomic wheeled mobile robots based on distributed observer[J]. *Nonlinear Dynamics*, 2021, 106(4): 3331-3349.
- [2] Li P F, Wang T, Kang Y, et al. Event-based model predictive control for nonlinear systems with dynamic disturbance[J]. *Automatica*, 2022, 145: 110533.
- [3] Okine A A, Adam N, Naeem F, et al. Multi-agent deep reinforcement learning for packet routing in tactical mobile sensor networks[J]. *IEEE Transactions on Network and Service Management*, 2024, 21(2): 2155-2169.
- [4] 朱志斌, 刘忠信, 王付永, 等. 基于混合事件触发机制的多智能体系统有限时间包容控制[J]. *控制与决策*, 2025, 40(5): 1494-1502.
(Zhu Z B, Liu X X, Wang F Y, et al. Finite-time containment control for multi-agent systems via hybrid event-triggered mechanism[J]. *Control and Decision*, 2025, 40(5): 1494-1502.)
- [5] 孙琛, 林岩, 王雪松, 等. 复杂非线性和输出约束下一类非线性系统的自适应输出反馈事件触发跟踪控制[J]. *控制与决策*, 2025, 40(7): 2117-2124.
(Sun C, Lin Y, Wang X S, et al. Adaptive output feedback event-triggered tracking control for a class of nonlinear systems with severe nonlinearities and output constraints[J]. *Control and Decision*, 2025, 40(7): 2117-2124.)
- [6] Lu Y, Wu X R, Wang Y N, et al. Quantized event-triggered-based finite-time H_∞ control for interval type-2 fuzzy Markov jump systems with random coupling delays[J]. *Nonlinear Dynamics*, 2024, 112(21): 19125-19146.
- [7] Csomay-Shanklin N, Taylor A J, Rosolia U, et al. Multi-rate planning and control of uncertain nonlinear systems: Model predictive control and control Lyapunov functions[C]. *Proceedings of the 2022 IEEE 61st Conference on Decision and Control (CDC)*. Cancun, 2022: 3732-3739.
- [8] Aydinoglu A, Posa M. Real-time multi-contact model predictive control via ADMM[C]. *2022 International Conference on Robotics and Automation*. Philadelphia, 2022: 3414-3421.
- [9] Munoz de la Pena D, Christofides P D. Lyapunov-based model predictive control of nonlinear systems subject to data losses[J]. *IEEE Transactions on Automatic Control*, 2008, 53(9): 2076-2089.
- [10] Han H G, Fu S J, Sun H Y, et al. Data-driven model-predictive control for nonlinear systems with stochastic sampling interval[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, 53(5): 3019-3030.
- [11] Bai N, Duan Z S, Wang Q S. Distributed optimal consensus of multi-agent systems: A randomized parallel approach[J]. *Automatica*, 2024, 159: 111339.
- [12] Floriano B R O, Vargas A N, Ishihara J Y, et al. Neural-network-based model predictive control for consensus of nonlinear systems[J]. *Engineering Applications of Artificial Intelligence*, 2022, 116: 105327.
- [13] Fu M Y, Xie L H. The sector bound approach to quantized feedback control[J]. *IEEE Transactions on Automatic Control*, 2005, 50(11): 1698-1711.
- [14] Kashyap A, Başar T, Srikant R. Quantized consensus[J]. *Automatica*, 2007, 43(7): 1192-1203.
- [15] Liu H, Cao M, De Persis C. Quantization effects on synchronized motion of teams of mobile agents with second-order dynamics[J]. *Systems & Control Letters*, 2012, 61(12): 1157-1167.
- [16] 陈世明, 姜根兰, 张正. 通信受限的多智能体系统二分实用一致性[J]. *自动化学报*, 2022, 48(5): 1318-1326.
(Chen S M, Jiang G L, Zhang Z. Bipartite practical consensus control of multi-agent systems with communication constraints[J]. *Acta Automatica Sinica*, 2022, 48(5): 1318-1326.)
- [17] Zhu Y R, Li S L, Ma J Y, et al. Bipartite consensus in networks of agents with antagonistic interactions and quantization[J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2018, 65(12): 2012-2016.
- [18] Zhang J, Liu S, Zhang X F. Output-feedback distributed consensus for nonlinear multi-agent systems with quantization[J]. *Information Sciences*, 2022, 585: 246-

- 261.
- [19] Wu Z G, Xu Y, Pan Y J, et al. Event-triggered control for consensus problem in multi-agent systems with quantized relative state measurements and external disturbance[J]. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2018, 65(7): 2232-2242.
- [20] 马梓元, 万茹, 龚华军, 等. 异构多智能体输出调节量化自适应跟踪控制[J]. *宇航学报*, 2024, 45(3): 469-477. (Ma Z Y, Wan R, Gong H J, et al. Heterogeneous multi-agent output adjustment quantization adaptive tracking control[J]. *Journal of Astronautics*, 2024, 45(3): 469-477.)
- [21] You K Y, Li Z K, Xie L H. Consensus condition for linear multi-agent systems over randomly switching topologies[J]. *Automatica*, 2013, 49(10): 3125-3132.
- [22] 王加朋, 杨家宁, 罗家祥. 切换拓扑下二阶多智能体混合事件触发固定时间一致性[J]. *控制与决策*, 2025, 40(3): 775-784. (Wang J P, Yang J N, Luo J X. Hybrid event-triggered fixed-time consensus for second-order multi-agent systems under switching topologies[J]. *Control and Decision*, 2025, 40(3): 775-784.)
- [23] 孙梦薇, 任璐, 刘剑, 等. 切换拓扑下动态事件触发多智能体系统固定时间一致性[J]. *自动化学报*, 2023, 49(6): 1295-1305. (Sun M W, Ren L, Liu J, et al. Dynamic event-triggered fixed-time consensus control of multi-agent systems under switching topologies[J]. *Acta Automatica Sinica*, 2023, 49(6): 1295-1305.)
- [24] 王庆领, 王雪娆. 切换拓扑下非线性多智能体系统自适应神经网络一致性[J]. *控制理论与应用*, 2023, 40(4): 633-640. (Wang Q L, Wang X R. Adaptive NN consensus of nonlinear multi-agent systems under switching topologies[J]. *Control Theory & Applications*, 2023, 40(4): 633-640.)
- [25] Jin X, Lu S, Yu J. Adaptive NN-based consensus for a class of nonlinear multiagent systems with actuator faults and faulty networks[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2022, 33(8): 3474-3486.

作者简介

卢毅 (1997-), 男, 博士生, 主要研究方向为多机器人系统、多智能体系统, E-mail: yilu_1027@163.com;

伍锡如 (1981-), 男, 教授, 博士, 主要研究方向为机器人控制、智能控制, E-mail: xiruwu@guet.edu.cn;

伍日立 (1999-), 男, 博士生, 主要研究方向为多机器人系统、多智能体系统, E-mail: rili_wu@163.com;

谢劼欣 (1990-), 男, 博士, 主要研究方向为医疗机器人、深度学习, E-mail: jxxie20@fudan.edu.cn;

仲于海 (1989-), 男, 博士, 主要研究方向为仿生机器人、机器人系统, E-mail: yuhaizhong523@mail.nwpu.