

# 控制与决策

Control and Decision

基于可变障碍函数和强化学习的预设性能最优安全跟踪控制

李小华, 刘莹, 邹嵩楠

引用本文:

李小华, 刘莹, 邹嵩楠. 基于可变障碍函数和强化学习的预设性能最优安全跟踪控制[J]. *控制与决策*, 2025, 40(3): 803–812.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2024.0326>

---

您可能感兴趣的其他文章

Articles you may be interested in

[基于移动传感器/执行器网络的时滞分布参数系统镇定控制](#)

Stabilization control for a class of distributed parameter systems with time–delay based on mobile sensor and actuator networks  
*控制与决策*. 2021, 36(8): 1955–1962 <https://doi.org/10.13195/j.kzyjc.2019.1309>

[输入约束不确定系统的点对点迭代学习控制与优化](#)

Point–to–point iterative learning control and optimization for uncertain systems with constrained input  
*控制与决策*. 2021, 36(6): 1435–1441 <https://doi.org/10.13195/j.kzyjc.2019.0908>

[四旋翼无人机抗干扰轨迹跟踪控制](#)

Anti–interference trajectory tracking control of quadrotor UAV  
*控制与决策*. 2021, 36(2): 379–386 <https://doi.org/10.13195/j.kzyjc.2019.0875>

[随机变批次长度的反馈辅助PD型量化迭代学习控制](#)

Feedback–assisted PD–type quantized iterative learning control with randomly iteration varying lengths  
*控制与决策*. 2021, 36(10): 2569–2576 <https://doi.org/10.13195/j.kzyjc.2020.0273>

[基于神经动态优化的非线性系统近似最优跟踪控制](#)

Approximate optimal tracking control for nonlinear systems based on neurodynamic optimization  
*控制与决策*. 2021, 36(1): 97–104 <https://doi.org/10.13195/j.kzyjc.2020.0056>

# 基于可变障碍函数和强化学习的预设性能 最优安全跟踪控制

李小华<sup>†</sup>, 刘莹, 邹嵩楠

(辽宁科技大学 电子与信息工程学院, 辽宁 鞍山 114051)

**摘要:** 研究一类具有未知初始跟踪条件的非线性系统预设性能最优安全跟踪控制问题. 首先, 开发一个基于可变障碍函数的性能约束控制设计的新方法, 并基于已有的安全边界保护法 (SBPM) 提出一个新的安全边界自调整规律 (SBSAL), 使其不仅可以处理实际输出约束发生突变的情况, 而且还可以解决突变解除后系统输出不能快速准确跟踪原期望轨迹的问题, 使得安全跟踪控制策略更为完善. 然后, 采用演员-评论家神经网络 (ACNNs) 强化学习 (RL) 算法优化系统的控制输入, 减少控制的能量消耗. 所设计预设性能最优安全跟踪控制器可保证系统在初始跟踪条件未知情况下的安全跟踪控制, 且系统输出具有预设有限时间控制性能. 最后, 通过仿真验证所提出方法的有效性.

**关键词:** 最优跟踪控制; 强化学习; 安全跟踪; 预设性能控制; 可变障碍函数; 输出约束

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2024.0326

引用格式: 李小华, 刘莹, 邹嵩楠. 基于可变障碍函数和强化学习的预设性能最优安全跟踪控制 [J]. 控制与决策, 2025, 40(3): 803-812.

## Optimal safety tracking control with prescribed performance based on variable barrier function and reinforcement learning

LI Xiao-hua<sup>†</sup>, LIU Ying, ZOU Song-nan

(School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan 114051, China)

**Abstract:** The optimal safety tracking control problem with prescribed performance is investigated for a class of nonlinear systems with unknown initial tracking condition. A new method for performance constraint control design is developed based on a variable barrier function. Based on the existing secure boundary protection method (SBPM), a novel secure boundary self-adjustment law (SBSAL) is proposed. It can not only handle the situations that the actual output constraints suddenly change, but also solve the problem that the system output is not able to quickly and accurately track the original expected trajectory after the mutation is relieved, so that the safety tracking control strategy is more consummate. Meanwhile, the reinforcement learning (RL) optimal method based on actor-critic neural networks (ACNNs) is adopted to optimize the control input of the system, and reduce the energy consumption for control. The designed optimal safety tracking controller with prescribed performance constraint can ensure the safe tracking control of the system with unknown initial tracking condition, and the output of the system has prescribed finite-time control performance. Finally, the effectiveness of the proposed method is verified by simulations.

**Keywords:** optimal tracking control; reinforcement learning; safe tracking; prescribed performance control; variable barrier function; output constraint

## 0 引言

实际工程中系统的控制性能、输出或状态往往需要满足特定的约束条件, 因此约束控制问题备受

关注, 并得到了广泛研究<sup>[1-4]</sup>. 这种约束控制主要包括性能约束和物理边界约束, 其中输出和状态约束属于物理边界约束. 所谓安全跟踪控制即为保证系统

收稿日期: 2024-03-28; 录用日期: 2024-09-26.

基金项目: 吉林大学汽车仿真与控制国家重点实验室开放基金项目 (20210219); 辽宁科技大学研究生科技创新项目 (LKDYC202313).

责任编辑: 吴立刚.

<sup>†</sup>通信作者. E-mail: lixiaohua6412@163.com.

输出不会违反实际物理边界约束的控制. 如当无人机通过狭窄隧道时, 飞行轨迹需要避免与其发生碰撞, 以确保飞机的安全飞行. 对系统安全跟踪控制的研究具有重要的实际意义.

目前, 在已有的约束控制研究中, 一般采用障碍李雅普诺夫函数方法进行设计<sup>[5-8]</sup>, 并假定输出的期望轨迹在实际约束边界内. 但是, 在系统运行期间实际输出约束边界可能会突变并导致期望轨迹与实际输出边界发生冲突. 此时, 传统控制方法将不能使用. 因此, 文献[9-12]研究了此类情况下系统的安全跟踪控制问题, 提出了一类确保系统输出不违反实际输出约束的安全跟踪控制方法. 该方法的思想是通过预测来重新设计新的期望轨迹, 使其不违反实际物理边界约束. 然而, 修改期望轨迹的过程非常繁琐, 同时, 在实际输出边界发生突变时, 此类方法将会造成控制输入的大幅抖动. 鉴于此, 文献[13]从另一个角度提出了一个安全边界保护法(SBPM), 用一个边界自调整规律生成安全边界, 并用双向滤波平滑机制(BFSM)抑制了控制输入的抖动. 然而, 文献[13]没有考虑实际输出约束突变解除后的跟踪控制情况, 对于解决突变解除后系统输出能够快速跟踪原期望轨迹的问题也是有实际价值的, 这是一个完整的SBPM必须解决的问题.

同时, 在此类安全跟踪控制中考虑其系统性能约束是非常有必要的. 为了约束系统的暂态和稳态性能, 文献[14]提出了预设性能控制(PPC)方法. 传统的预设性能控制方法和输出约束控制方法均需要系统满足初始约束条件. 然而, 在某些情况下约束变量的初始值可能不易获得或未知, 因此, 传统的性能约束控制方法将不再适用. 为了消除对初始条件的依赖性, 文献[15-16]通过使用一种变换函数将被约束量的初始值转换为0; 文献[17]构造了一个有界的映射函数, 以确保约束变量的初始条件转换到预设的范围内; 而文献[18]则通过使得预设函数的初始边界范围扩展到无穷大来解决这个问题. 然而, 上述方法只能针对性能函数为恒正或恒负的情况, 对于约束函数跨越零点的情况, 现有的约束控制方案均无法处理.

近年来, 在PPC技术中引入强化学习(RL)算法开始受到关注. 这可以显著降低控制输入的幅值, 优化控制性能. 目前, 一些新的研究成果已经发表, 如文献[19-21]. 在这些研究中, 同时考虑了PPC技术和最优控制策略, 获得了系统的预设性能最优跟踪控制器. 然而, 这些论文中的PPC设计大多采用了传统的预设性能设计方法, 很少有文献考虑未知初始

条件的PPC问题和安全跟踪的控制问题.

受上述分析的启发, 本文针对一类具有外部扰动的非线性系统, 研究其不依赖初始条件的预设性能最优安全跟踪控制问题. 本文的主要内容如下.

1) 本文提出一种基于可变障碍函数的无关初始跟踪条件的PPC设计方法, 与目前已有的此类方法相比, 不需要复杂的变量转换, 其更为简洁;

2) 针对文献[13]中的安全跟踪控制方法, 本文提出一个新的安全边界自调整规律, 使其不仅可以处理实际输出约束发生突变的情况, 而且还可以解决突变解除后系统输出不能快速准确跟踪原期望轨迹的问题, 使得安全跟踪控制策略更为完整;

3) 所提出可变障碍函数法放宽了对性能约束函数的选择范围, 这使得系统的状态约束和输出约束更为便捷;

4) 所提出PPC控制方法可同时做到跟踪误差的性能约束和安全输出约束, 这是目前已有PPC方法鲜有做到的.

## 1 系统描述与预备知识

### 1.1 系统描述

考虑以下非线性系统:

$$\begin{cases} \dot{x}_i = x_{i+1} + f_i(\bar{x}_i) + \omega_i(t), & i = 1, \dots, n-1; \\ \dot{x}_n = g_n(\bar{x}_n)u + f_n(\bar{x}_n) + \omega_n(t); \\ y = x_1. \end{cases} \quad (1)$$

其中:  $x_i (i=1, 2, \dots, n)$  为系统的状态;  $\bar{x}_i = [x_1, x_2, \dots, x_i]^T \in R^i$ ;  $u \in R$  和  $y \in R$  分别为控制输入和系统的输出;  $f_i(\bar{x}_i)$ 、 $g_n(\bar{x}_n)$  为未知非线性函数; 外部扰动  $|\omega_i(t)| < \bar{W}_i$ ,  $\bar{W}_i$  为正常数. 考虑系统的初始输出  $x_1(0)$  未知, 研究其最优安全跟踪控制问题.

**假设 1**<sup>[13]</sup> 期望轨迹  $y_r$  及其各阶导数  $y_r^{(i)}$  已知且连续有界,  $i = 1, 2, \dots, n$ .

**假设 2**<sup>[22]</sup> 函数  $g_n(\bar{x}_n)$  有界且函数符号已知, 并存在未知正常数  $\underline{g}_n$  使得  $0 < \underline{g}_n < |g_n(\bar{x}_n)| < \infty$ . 不失一般性, 假设  $g_n(\bar{x}_n) > 0$ .

控制目标: 在系统初始跟踪条件未知的情况下设计其预设性能最优安全跟踪控制器, 保证: 1) 当期望轨迹违反实际输出约束时, 安全边界可根据实际的输出约束边界自动调整, 系统输出按照安全边界约束运行, 不违反实际输出约束; 2) 当期望轨迹不违反实际输出约束时, 系统输出能够准确地跟踪期望轨迹, 并能够保证跟踪误差具有预设控制性能.

### 1.2 预备知识

**定义 1**<sup>[13]</sup> 若函数  $\vartheta(t)$  满足如下性质: 1)  $\vartheta(t) >$

0; 2)  $\dot{\vartheta}(t) \leq 0$ ; 3)  $\lim_{t \rightarrow T_f} \vartheta(t) = \vartheta_{T_f} > 0$ ; 4)  $\vartheta(t)$  为  $C^n$  函数, 且当  $t \geq T_f$  时, 有  $\vartheta(t) = \vartheta_{T_f}$ , 其中设计参数  $\vartheta_{T_f} > 0$  和  $T_f > 0$  分别为稳态误差和停息时间. 则  $\vartheta(t)$  为预设有限时间性能函数.

本文设计  $\vartheta_i(t)$  为

$$\vartheta_i(t) = \begin{cases} a_i \left( 0.5 \cos \frac{\pi t}{T_f} + 0.5 \right)^q + \vartheta_{T_f}, & 0 \leq t < T_f; \\ \vartheta_{T_f}, & t \geq T_f. \end{cases} \quad (2)$$

其中  $a_i > 0$  和  $q \geq n + 1$  为设计参数.

### 1.3 基于新 SBSAL 的安全边界保护法

对于系统的输出约束  $\rho_2(t) < x_1 < \rho_1(t)$ , 通常假设  $\rho_2(t) < y_r < \rho_1(t)$ , 但是, 在实际中输出约束边界可能会发生突变, 这将导致实际输出边界与期望轨迹发生冲突, 使得系统跟踪过程出现安全问题, 且当突变解除后, 系统输出能够快速跟踪期望轨迹. 因此, 提出一个新的 SBSAL 来处理这种从发生突变到突变解除的安全跟踪情况. 基于新 SBSAL 的 SBPM 的实施步骤可参见文献 [13] 的图 5.

所提出方法考虑如下式描述的输出约束突变及其解除的情况, 设  $\rho_1$ 、 $\rho_2$  分别为上、下约束边界:

$$\rho_1(t) = \begin{cases} \rho_{11}(t), & 0 \leq t < T_A; \\ \rho_{12}(t), & T_A \leq t < T_B; \\ \rho_{11}(t), & t \geq T_B; \end{cases} \quad (3)$$

$$\rho_2(t) = \begin{cases} \rho_{21}(t), & 0 \leq t < T_C; \\ \rho_{22}(t), & T_C \leq t < T_D; \\ \rho_{21}(t), & t \geq T_D. \end{cases} \quad (4)$$

其中:  $\rho_2 < \rho_1$ ,  $\rho_{11}$ 、 $\rho_{12}$ 、 $\rho_{21}$ 、 $\rho_{22}$  为平滑函数;  $T_A$ 、 $T_B$ 、 $T_C$ 、 $T_D$  分别为上、下约束的突变及其解除时刻.

考虑到系统需要进行预设性能约束且不依赖输出的初始条件, 本文假定在某个时间  $t \geq T$  后, 具有任意初值的系统输出均可进入约束范围, 其中  $T$  为一个时间参数. 该约束范围可被描述为

$$\bar{k}_d(t) < x_1(t) < \bar{k}_u(t), \quad t \geq T. \quad (5)$$

这里:  $\bar{k}_u(t)$ 、 $\bar{k}_d(t)$  分别为期望的上、下边界, 定义为

$$\bar{k}_u(t) = y_r(t) + \vartheta_1(t), \quad (6)$$

$$\bar{k}_d(t) = y_r(t) - \vartheta_2(t). \quad (7)$$

显然, 若满足式 (5), 则系统输出将被约束在期望边界内, 其控制性能可通过  $\vartheta_i(t)$  来调整. 当实际输出约束边界发生突变时, 为了避免冲突, 系统将通过 SBSAL 生成安全边界. 由于发生突变, 将会导致控制输入的大幅抖动. 为了防止这种情况, 采用文献 [13] 中的 BFSM, 使得生成的安全边界更平滑, 从而避免大幅抖动. BFSM 方法如下.

首先, 对输出约束采用反向滤波来构建虚拟输出约束  $\hat{\rho}_1$ 、 $\hat{\rho}_2$ , 有

$$\hat{\rho}_1 = \begin{cases} \sigma \dot{\rho}_{11} + \rho_{11} = \hat{\rho}_{11}, & 0 \leq t < T_A; \\ \sigma \dot{\rho}_{12} + \rho_{12} = \hat{\rho}_{12}, & T_A \leq t < T_B; \\ \sigma \dot{\rho}_{11} + \rho_{11} = \hat{\rho}_{11}, & t \geq T_B; \end{cases} \quad (8)$$

$$\hat{\rho}_2 = \begin{cases} \sigma \dot{\rho}_{21} + \rho_{21} = \hat{\rho}_{21}, & 0 \leq t < T_C; \\ \sigma \dot{\rho}_{22} + \rho_{22} = \hat{\rho}_{22}, & T_C \leq t < T_D; \\ \sigma \dot{\rho}_{21} + \rho_{21} = \hat{\rho}_{21}, & t \geq T_D. \end{cases} \quad (9)$$

其中  $\sigma$  为滤波常数. 同时, 使用反向滤波构建如下虚拟安全边界和虚拟期望轨迹:

$$\hat{k}_u = \sigma \dot{k}_u + \bar{k}_u, \quad (10)$$

$$\hat{k}_d = \sigma \dot{k}_d + \bar{k}_d, \quad (11)$$

$$\hat{y}_r = \sigma \dot{y}_r + y_r. \quad (12)$$

这里:  $\hat{k}_u$  和  $\hat{k}_d$  为虚拟安全边界,  $\hat{y}_r$  为虚拟期望轨迹. 经 SBSAL 调整后, 再对其进行正向滤波, 得到输出的安全边界  $k_u$ 、 $k_d$  和期望轨迹  $y_r$  为

$$\dot{k}_u = \frac{\hat{k}_u - k_u}{\sigma}, \quad (13)$$

$$\dot{k}_d = \frac{\hat{k}_d - k_d}{\sigma}, \quad (14)$$

$$\dot{y}_r = \frac{\hat{y}_r - y_r}{\sigma}. \quad (15)$$

下面给出新的 SBSAL. 为了便于描述, 只考虑上约束边界的突变和解除的情况. 对于实际输出约束发生突变的情况, 其自调整规律参见文献 [13], 这里主要介绍 SBSAL 中突变解除情况下的自调整规律, 具体分为以下两种情况.

**情况 1** 当  $t \geq T_B$  时, 若  $\hat{y}_r(t) \leq \hat{\rho}_1(t)$ , 则安全边界的 SBSAL 为

$$\begin{cases} \hat{k}_u = \hat{k}_u, \\ \hat{k}_d = \hat{k}_d. \end{cases} \quad (16)$$

**情况 2** 当  $\hat{y}_r(T_B) > \hat{\rho}_1(T_B)$  时, 为了能够更快且更准确地跟踪上期望轨迹, 定义一条直线  $\tilde{k}_u(t) = \hat{k}_u(T_B) + \varpi(t - T_B)$ , 其中  $\varpi$  为设计梯度参数. 为了尽快跟踪上期望轨迹, 可选择较大的  $\varpi$ . 假设系统输出跟踪上期望轨迹的时刻为  $t_n$ . 若  $\hat{y}_r(T_B) > \hat{\rho}_1(T_B)$ , 则 SBSAL 为

$$\begin{cases} \hat{k}_u = \tilde{k}_u(t) = \hat{k}_u(T_B) + \varpi(t - T_B), \\ \hat{k}_d = \hat{k}_u - \vartheta_1(t) - \vartheta_2(t), \\ T_B < t < t_n; \end{cases}$$

$$\begin{cases} \hat{k}_u = \hat{y}_r + \vartheta_1(t), \\ \hat{k}_d = \hat{k}_u - \vartheta_1(t) - \vartheta_2(t), \end{cases}$$

$$\tilde{k}_u(t) \geq \hat{y}_r + \vartheta_1(t), t \geq t_n. \quad (17)$$

#### 1.4 基于可变障碍函数的PPC新方法

一般的约束控制要求被约束量 $\bar{h}(t)$ 必须满足 $k_d(0) < \bar{h}(0) < k_u(0)$ ,而在实际应用中,初始条件 $\bar{h}(0)$ 可能难以获得.针对未知初始条件的情况,本文提出一个可变障碍函数,使得系统在任意初始条件下均能够满足输出约束和预设性能.

首先,定义一个调节函数,即

$$\psi(t) = \begin{cases} \left(0.5 \cos \frac{\pi t}{T} + 0.5\right)^r, & 0 \leq t < T; \\ 0, & t \geq T. \end{cases} \quad (18)$$

其中: $r \geq n + 1$ 为设计参数, $T$ 为 $\bar{h}$ 进入预设范围的时间, $\psi(t)$ 为 $C^n$ 函数.则可变障碍函数被构造为

$$z_1 = 0.5 \ln \frac{-k_d(1-\psi) + \psi(\bar{h}^2 + k) + \bar{h}}{k_u(1-\psi) + \psi(\bar{h}^2 + k) - \bar{h}}, \quad (19)$$

这里 $k \geq 1$ 为设计参数.

**注1** 函数(19)可在任意初始条件下实现被约束变量 $\bar{h}(t)$ 的性能约束设计.由于在 $t = 0$ 时,有 $z_1 = 0.5 \ln \frac{(\bar{h}^2 + k) + \bar{h}}{(\bar{h}^2 + k) - \bar{h}}$ ,其与 $k_u$ 和 $k_d$ 的值无关.因此,无需知道约束变量的初始条件.与文献[15]中不依赖初始条件的约束控制方法不同,所提出方法去掉了文献[15]中 $k_u(0) > 0$ 和 $k_d(0) < 0$ 的假设,即文献[15]的 $k_u$ 和 $k_d$ 必须是严格正和严格负的.而式(19)中的 $k_u$ 和 $k_d$ 并不总是正或负的函数,它们也可以通过零点来拓宽约束函数的选择范围.此外,该可变障碍函数还可同时处理预设性能约束和输出约束,并可根据需要任意选择 $k_u$ 和 $k_d$ ,只需要保证 $k_d < k_u$ 即可.

通过下文定理1给出新的无关初始条件PPC方法.

**定理1** 若 $z_1$ 有界,则当 $t > 0$ 时, $\bar{h}$ 有界,且不依赖初始条件 $\bar{h}(0)$ ,总有下式成立:

$$k_d < \bar{h} < k_u, t \geq T. \quad (20)$$

**证明** 由于 $k \geq 1$ , $\psi(0) = 1$ ,有

$$\psi(0)(\bar{h}^2(0) + k) > |\bar{h}(0)|. \quad (21)$$

由式(19)和(21),可得到

$$-k_d(0)(1-\psi(0)) + \psi(0)(\bar{h}^2(0) + k) + \bar{h}(0) > 0, \quad (22)$$

$$k_u(0)(1-\psi(0)) + \psi(0)(\bar{h}^2(0) + k) - \bar{h}(0) > 0. \quad (23)$$

因此,对于任意初始值 $\bar{h}(0)$ , $z_1(0)$ 均有定义.当且仅当 $-k_d(1-\psi) + \psi(\bar{h}^2 + k) + \bar{h} \rightarrow 0$ 或 $k_u(1-\psi) + \psi(\bar{h}^2 + k) - \bar{h} \rightarrow 0$ , $z_1 \rightarrow \infty$ .因此,当 $0 < t \leq T$ 时,

若 $z_1$ 有界,则

$$-k_d(1-\psi) + \psi(\bar{h}^2 + k) + \bar{h} > 0, \quad (24)$$

$$k_u(1-\psi) + \psi(\bar{h}^2 + k) - \bar{h} > 0. \quad (25)$$

由式(18)、(24)和(25)可知,当 $t \geq T$ 时, $\psi = 0$ ,有

$$k_d < \bar{h} < k_u, t \geq T. \quad (26)$$

当 $0 \leq t \leq T$ 时,由于 $\bar{h}$ 是连续的,可知 $\bar{h}$ 有界.□

## 2 控制器设计

首先,采用式(19)中的可变障碍函数对输出 $x_1(t)$ 进行转换,即取 $\bar{h}(t) = x_1(t)$ ,有

$$z_1 = 0.5 \ln \frac{-k_d(1-\psi) + \psi(x_1^2 + k) + x_1}{k_u(1-\psi) + \psi(x_1^2 + k) - x_1}. \quad (27)$$

构造如下坐标变换:

$$z_i = x_i - \alpha_{i-1}, i = 2, 3, \dots, n, \quad (28)$$

其中 $\alpha_{i-1}$ 为虚拟控制律.

step 1: 由式(1)、(27)和(28)可知

$$\begin{aligned} \dot{z}_1 = & \frac{\partial z_1}{\partial x_1}(f_1 + x_2 + \omega_1) + \frac{\partial z_1}{\partial k_d} \dot{k}_d + \frac{\partial z_1}{\partial k_u} \dot{k}_u + \frac{\partial z_1}{\partial \psi} \dot{\psi} = \\ & \bar{F}_1 + z_2 + \alpha_1 - \frac{1}{4} M^2 z_1 - z_1 + M \omega_1. \end{aligned} \quad (29)$$

其中: $\bar{F}_1 = M \left( f_1 + x_2 + \frac{1}{4} M z_1 \right) + \beta - x_2 + z_1$ , $M = \partial z_1 / \partial x_1$ , $\beta = (\partial z_1 / \partial k_d) \dot{k}_d + (\partial z_1 / \partial k_u) \dot{k}_u + (\partial z_1 / \partial \psi) \dot{\psi}$ .

利用径向基函数神经网络(RBFNN)对 $\bar{F}_1$ 逼近<sup>[23]</sup>,有

$$\bar{F}_1 = W_1^{*T} S_1(Z_1) + \varepsilon_1(Z_1), |\varepsilon_1(Z_1)| \leq \bar{\varepsilon}_1, \quad (30)$$

其中 $Z_1 = [x_1, x_2, \psi, \dot{\psi}, \vartheta_1, \dot{\vartheta}_1, \vartheta_2, \dot{\vartheta}_2, y_r, \dot{y}_r]^T$ .

定义最优性能指标函数 $J_1^*$ 为

$$J_1^*(z_1(t)) = \min_{\alpha_1 \in \Omega} \int_t^\infty (z_1^2(\tau) + \alpha_1^2(z_1(\tau))) d\tau. \quad (31)$$

其中: $\alpha_1$ 为虚拟控制输入, $\Omega$ 为包含所有控制输入的紧集.基于贝尔曼最优性原理,可得到哈密顿-雅可比-贝尔曼(HJB)方程为

$$H_1 \left( z_1, \alpha_1^*, \frac{\partial J_1^*}{\partial z_1} \right) = z_1^2 + \alpha_1^{*2} + \frac{\partial J_1^*}{\partial z_1} \dot{z}_1 = 0, \quad (32)$$

这里 $\alpha_1^*$ 为最优虚拟控制.由 $\frac{\partial H_1}{\partial \alpha_1^*} = 2\alpha_1^* + \frac{\partial J_1^*}{\partial z_1} = 0$ ,有

$$\alpha_1^* = -\frac{1}{2} \frac{\partial J_1^*}{\partial z_1}. \quad (33)$$

构造如下辅助最优性能指标函数:

$$J_1^0 = -\frac{1}{2} \frac{\partial J_1^*}{\partial z_1} + c_1 z_1 + z_1 + z_1 \theta_1^* S_1^T S_1, \quad (34)$$

其中 $\theta_1^* = \|W_1^*\|^2$ ,可得到

$$\alpha_1^* = -c_1 z_1 - z_1 - z_1 \theta_1^* S_1^T S_1 + J_1^0, \quad (35)$$

$$\frac{\partial J_1^*}{\partial z_1} = 2c_1 z_1 + 2z_1 + 2z_1 \theta_1^* S_1^T S_1 - 2J_1^0, \quad (36)$$

这里  $c_1 > 0$  为设计参数. 利用 RBFNN 逼近  $J_1^0$ , 有

$$J_1^0 = \varphi_1^{*T} \phi_1(X_1) + \tau_1(X_1), \quad |\tau_1(X_1)| \leq \bar{\tau}_1, \quad (37)$$

其中  $X_1 = [x_1, \psi, \dot{\psi}, \vartheta_1, \dot{\vartheta}_1, \vartheta_2, \dot{\vartheta}_2, y_r, \dot{y}_r]^T$ .

为了获得最优虚拟控制, 利用 ACNNs 在线训练  $\varphi_1^*$ . 优化的虚拟控制和性能指标函数的梯度估计分别为

$$\alpha_1 = -c_1 z_1 - z_1 - z_1 \hat{\theta}_1 S_1^T S_1 + \hat{\varphi}_{a1}^T \phi_1, \quad (38)$$

$$\frac{\partial \hat{J}_1}{\partial z_1} = 2c_1 z_1 + 2z_1 + 2z_1 \hat{\theta}_1 S_1^T S_1 - 2\hat{\varphi}_{c1}^T \phi_1. \quad (39)$$

其中:  $\hat{\theta}_1$  为  $\theta_1^*$  的估计值, 其自适应律为

$$\dot{\hat{\theta}}_1 = z_1^2 S_1^T S_1 - \lambda_1 \hat{\theta}_1. \quad (40)$$

这里:  $\lambda_1 > 0$  为设计参数, 设计 actor NN 的权更新律为

$$\dot{\hat{\varphi}}_{a1} = -\phi_1 \phi_1^T (\lambda_{a1} (\hat{\varphi}_{a1} - \hat{\varphi}_{c1}) + \lambda_{c1} \hat{\varphi}_{c1}). \quad (41)$$

同样地, 设计 critic NN 的权更新律为

$$\dot{\hat{\varphi}}_{c1} = -\lambda_{c1} \phi_1 \phi_1^T \hat{\varphi}_{c1}. \quad (42)$$

其中:  $\lambda_{a1} > 0, \lambda_{c1} > 0$  为设计参数.

结合式 (29)、(32)、(38) 和 (39), 得到 Bellman 残差  $\xi_1$  为

$$\begin{aligned} \xi_1 = & H_1(z_1, \alpha_1, \frac{\partial \hat{J}_1}{\partial z_1}) - H_1(z_1, \alpha_1^*, \frac{\partial J_1^*}{\partial z_1}) = \\ & z_1^2 + (-c_1 z_1 - z_1 - z_1 \hat{\theta}_1 S_1^T S_1 + \hat{\varphi}_{a1}^T \phi_1)^2 + \\ & (2c_1 z_1 + 2z_1 + 2z_1 \hat{\theta}_1 S_1^T S_1 - 2\hat{\varphi}_{c1}^T \phi_1) \times \\ & \left( W_1^{*T} S_1 + \varepsilon_1 + z_2 - \frac{1}{4} M^2 z_1 - z_1 + M\omega_1 - \right. \\ & \left. c_1 z_1 - z_1 - z_1 \hat{\theta}_1 S_1^T S_1 + \hat{\varphi}_{a1}^T \phi_1 \right). \end{aligned} \quad (43)$$

为了获得近似最优 actor NN 权向量, 使得残差  $\xi_1$  达到最小, 取  $\xi_1$  的梯度为

$$\frac{\partial \xi_1}{\partial \hat{\varphi}_{a1}} = 2\phi_1 \phi_1^T (\hat{\varphi}_{a1} - \hat{\varphi}_{c1}) = 0. \quad (44)$$

为了获得 RL 更新律来保证式 (44) 成立, 定义权向量误差  $\hat{\varphi}_{a1} - \hat{\varphi}_{c1}$  构造正函数为

$$L_{F_1}(t) = (\hat{\varphi}_{a1} - \hat{\varphi}_{c1})^T (\hat{\varphi}_{a1} - \hat{\varphi}_{c1}). \quad (45)$$

对式 (45) 求导, 可得到

$$\dot{L}_{F_1}(t) = \frac{\partial L_{F_1}(t)}{\partial \hat{\varphi}_{a1}^T} \dot{\hat{\varphi}}_{a1} + \frac{\partial L_{F_1}(t)}{\partial \hat{\varphi}_{c1}^T} \dot{\hat{\varphi}}_{c1}. \quad (46)$$

由于  $\frac{\partial L_{F_1}(t)}{\partial \hat{\varphi}_{a1}} = -\frac{\partial L_{F_1}(t)}{\partial \hat{\varphi}_{c1}} = 2(\hat{\varphi}_{a1} - \hat{\varphi}_{c1})$ , 由式 (45) 和 (46), 可知

$$\dot{L}_{F_1} = -\frac{1}{2} \lambda_{a1} \frac{\partial L_{F_1}}{\partial \hat{\varphi}_{a1}^T} \phi_1 \phi_1^T \frac{\partial L_{F_1}}{\partial \hat{\varphi}_{a1}} \leq 0. \quad (47)$$

因此, 式 (41) 和 (42) 可使得权向量误差  $\hat{\varphi}_{a1} - \hat{\varphi}_{c1}$  渐近收敛至 0, 换言之, 最终  $\hat{\varphi}_{a1}$  和  $\hat{\varphi}_{c1}$  使得式 (44) 成立.

选择李雅普诺夫函数为

$$L_1 = \frac{1}{2} z_1^2 + \frac{1}{2} \tilde{\theta}_1^2 + L_{E_1}. \quad (48)$$

其中:  $\tilde{\theta}_1 = \hat{\theta}_1 - \theta_1^*$  估计误差, 且

$$L_{E_1} = \frac{1}{2} \tilde{\varphi}_{a1}^T \tilde{\varphi}_{a1} + \frac{1}{2} \tilde{\varphi}_{c1}^T \tilde{\varphi}_{c1}. \quad (49)$$

这里:  $\tilde{\varphi}_{a1} = \hat{\varphi}_{a1} - \varphi_{a1}^*$ ,  $\tilde{\varphi}_{c1} = \hat{\varphi}_{c1} - \varphi_{c1}^*$  分别为 actor NN 和 critic NN 的权向量误差. 结合式 (41) 和 (42), 对式 (49) 求导, 可得到

$$\begin{aligned} \dot{L}_{E_1} = & -\tilde{\varphi}_{a1}^T \phi_1 \phi_1^T \lambda_{a1} \hat{\varphi}_{a1} + (\lambda_{a1} - \lambda_{c1}) \tilde{\varphi}_{a1}^T \phi_1 \phi_1^T \hat{\varphi}_{c1} - \\ & \lambda_{c1} \tilde{\varphi}_{c1}^T \phi_1 \phi_1^T \hat{\varphi}_{c1}, \end{aligned} \quad (50)$$

其中  $\lambda_{a1} > \lambda_{c1}$ . 考虑如下等式成立:

$$\begin{aligned} & -\tilde{\varphi}_{a1}^T \phi_1 \phi_1^T \lambda_{a1} \hat{\varphi}_{a1} = \\ & -\frac{1}{2} \tilde{\varphi}_{a1}^T \phi_1 \phi_1^T \lambda_{a1} \tilde{\varphi}_{a1} - \frac{1}{2} \hat{\varphi}_{a1}^T \phi_1 \phi_1^T \lambda_{a1} \hat{\varphi}_{a1} + \\ & \frac{1}{2} \varphi_{a1}^{*T} \phi_1 \phi_1^T \lambda_{a1} \varphi_{a1}^*, \end{aligned} \quad (51)$$

$$\begin{aligned} & -\tilde{\varphi}_{c1}^T \phi_1 \phi_1^T \lambda_{c1} \hat{\varphi}_{c1} = \\ & -\frac{1}{2} \tilde{\varphi}_{c1}^T \phi_1 \phi_1^T \lambda_{c1} \tilde{\varphi}_{c1} - \frac{1}{2} \hat{\varphi}_{c1}^T \phi_1 \phi_1^T \lambda_{c1} \hat{\varphi}_{c1} + \\ & \frac{1}{2} \varphi_{c1}^{*T} \phi_1 \phi_1^T \lambda_{c1} \varphi_{c1}^*. \end{aligned} \quad (52)$$

利用 Young's 不等式, 可得到

$$\begin{aligned} & (\lambda_{a1} - \lambda_{c1}) \tilde{\varphi}_{a1}^T \phi_1 \phi_1^T \hat{\varphi}_{c1} \leq \\ & \frac{1}{2} (\lambda_{a1} - \lambda_{c1}) (\tilde{\varphi}_{a1}^T \phi_1)^2 + \\ & \frac{1}{2} (\lambda_{a1} - \lambda_{c1}) (\phi_1^T \hat{\varphi}_{c1})^2. \end{aligned} \quad (53)$$

将式 (51) ~ (53) 代入 (50), 可得到

$$\begin{aligned} \dot{L}_{E_1} \leq & -\frac{1}{2} \lambda_{c1} \tilde{\varphi}_{c1}^T \phi_1 \phi_1^T \tilde{\varphi}_{c1} - \frac{1}{2} \lambda_{c1} \tilde{\varphi}_{a1}^T \phi_1 \phi_1^T \tilde{\varphi}_{a1} - \\ & \left( \lambda_{c1} - \frac{1}{2} \lambda_{a1} \right) \tilde{\varphi}_{c1}^T \phi_1 \phi_1^T \hat{\varphi}_{c1} + \frac{1}{2} (\lambda_{a1} + \lambda_{c1}) \times \\ & (\varphi_{a1}^{*T} \phi_1)^2 - \frac{1}{2} \lambda_{a1} (\hat{\varphi}_{a1}^T \phi_1)^2. \end{aligned} \quad (54)$$

对式 (48) 求导, 可得到

$$\begin{aligned} \dot{L}_1 = & z_1 \left( W_1^{*T} S_1(Z_1) + \varepsilon_1(Z_1) + z_2 + \alpha_1 - \right. \\ & \left. \frac{1}{4} M^2 z_1 - z_1 + M\omega_1 \right) + \tilde{\theta}_1 \dot{\hat{\theta}}_1 + \dot{L}_{E_1}. \end{aligned} \quad (55)$$

利用 Young's 不等式和文献 [22] 的引理 2, 有

$$z_1 M\omega_1 \leq \frac{1}{4} M^2 z_1^2 + \bar{W}_1^2, \quad (56)$$

$$z_1 \varepsilon_1 \leq z_1^2 + \frac{\bar{\varepsilon}_1^2}{4}, \quad (57)$$

$$z_1 W_1^{*T} S_1(Z_1) \leq |z_1| \|W_1^*\| \|S_1(\bar{Z}_1)\| \leq$$

$$z_1^2 \theta_1^* S_1^T(\bar{Z}_1) S_1(\bar{Z}_1) + \frac{1}{4}, \quad (58)$$

其中  $\bar{Z}_1 = [x_1, \psi, \dot{\psi}, \vartheta_1, \dot{\vartheta}_1, \vartheta_2, \dot{\vartheta}_2, y_r, \dot{y}_r]^T$ . 由式 (38)、(40) 和 (55) ~ (58), 可得到

$$\begin{aligned} \dot{L}_1 \leq & -c_1 z_1^2 + z_1 \hat{\varphi}_{a1}^T \phi_1 - \lambda_1 \tilde{\theta}_1 \hat{\theta}_1 + \bar{W}_1^2 + \\ & \frac{\bar{\varepsilon}_1^2}{4} + \frac{1}{4} - z_1^2 + z_1 z_2 + \dot{L}_{E_1}. \end{aligned} \quad (59)$$

由 Young's 不等式, 可得到

$$-\lambda_1 \tilde{\theta}_1 \hat{\theta}_1 \leq -\frac{\lambda_1}{2} \tilde{\theta}_1^2 + \frac{\lambda_1}{2} \theta_1^{*2}, \quad (60)$$

$$z_1 \hat{\varphi}_{a1}^T \phi_1 \leq z_1^2 + \frac{1}{4} (\hat{\varphi}_{a1}^T \phi_1)^2. \quad (61)$$

由式 (54) 和 (59) ~ (61) 可知

$$\begin{aligned} \dot{L}_1 \leq & -c_1 z_1^2 - \frac{\lambda_1}{2} \tilde{\theta}_1^2 + \frac{\lambda_1}{2} \theta_1^{*2} + \bar{W}_1^2 + \frac{\bar{\varepsilon}_1^2}{4} + \frac{1}{4} + \\ & z_1 z_2 - \frac{1}{2} \lambda_{c1} \lambda_{\phi_1}^{\min} \tilde{\varphi}_{c1}^T \tilde{\varphi}_{c1} - \frac{1}{2} \lambda_{c1} \lambda_{\phi_1}^{\min} \tilde{\varphi}_{a1}^T \tilde{\varphi}_{a1} - \\ & \left( \lambda_{c1} - \frac{1}{2} \lambda_{a1} \right) \hat{\varphi}_{c1}^T \phi_1 \phi_1^T \hat{\varphi}_{c1} + \frac{1}{2} (\lambda_{a1} + \lambda_{c1}) \times \\ & (\varphi_1^{*T} \phi_1)^2 - \frac{1}{2} \left( \lambda_{a1} - \frac{1}{2} \right) (\hat{\varphi}_{a1}^T \phi_1)^2, \end{aligned} \quad (62)$$

其中  $\lambda_{\phi_1}^{\min}$  为  $\phi_1 \phi_1^T$  的最小特征值. 若满足  $c_1 > 0$ ,  $\lambda_{a1} > \lambda_{c1} \geq \frac{1}{2} \lambda_{a1} \geq 0$ ,  $\lambda_{a1} \geq \frac{1}{2}$ , 则有

$$\dot{L}_1 \leq -A_1 L_1 + B_1 + z_1 z_2. \quad (63)$$

这里:  $A_1 = \min\{2c_1, \lambda_1, \lambda_{c1} \lambda_{\phi_1}^{\min}\}$ ,  $B_1 = \max\left\{\frac{\lambda_1}{2} \theta_1^{*2} + \bar{W}_1^2 + \frac{\bar{\varepsilon}_1^2}{4} + \frac{1}{4} + \frac{1}{2} (\lambda_{a1} + \lambda_{c1}) (\varphi_1^{*T} \phi_1)^2\right\}$ .

step  $i$  ( $2 \leq i \leq n$ ): 由式 (1) 和 (28), 有

$$\dot{z}_i = \dot{x}_i - \dot{\alpha}_{i-1} =$$

$$\begin{aligned} \bar{F}_i + x_{i+1} + \omega_i - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} \omega_j - \\ \frac{1}{4} z_i \sum_{j=1}^{i-1} \left( \frac{\partial \alpha_{i-1}}{\partial x_j} \right)^2 - \frac{1}{4} z_i - z_{i-1} - d_i. \end{aligned} \quad (64)$$

其中

$$\begin{aligned} \bar{F}_i = & f_i - \eta_{i-1} + \frac{1}{4} z_i \sum_{j=1}^{i-1} \left( \frac{\partial \alpha_{i-1}}{\partial x_j} \right)^2 + \\ & \frac{1}{4} z_i + z_{i-1} + d_i, \\ \eta_{i-1} = & \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} (x_{j+1} + f_j) + \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}_j} \dot{\hat{\theta}}_j + \\ & \sum_{j=0}^{i-1} \left( \frac{\partial \alpha_{i-1}}{\partial y_r^{(j)}} y_r^{(j+1)} + \frac{\partial \alpha_{i-1}}{\partial \psi^{(j)}} \psi^{(j+1)} \right) + \\ & \sum_{k=1}^2 \sum_{j=0}^{i-1} \frac{\partial \alpha_{i-1}}{\partial \vartheta_k^{(j)}} \vartheta_k^{(j+1)}. \end{aligned}$$

利用 RBFNN 对  $\bar{F}_i$  进行逼近, 即

$$\bar{F}_i = W_i^{*T} S_i(Z_i) + \varepsilon_i(Z_i), \quad |\varepsilon_i(Z_i)| \leq \bar{\varepsilon}_i. \quad (65)$$

这里:  $Z_i = [\bar{x}_i, y_r, \dots, y_r^{(i)}, \hat{\theta}_1, \dots, \hat{\theta}_{i-1}, \vartheta_1, \dots, \vartheta_1^{(i)}, \vartheta_2, \dots, \vartheta_2^{(i)}, \psi, \dots, \psi^{(i)}]^T$ , 当  $i < n$  时,  $x_{i+1} = \alpha_i + z_{i+1}$ ,  $d_i = z_i$ ; 当  $i = n$  时,  $x_{n+1} = g_n u$ ,  $d_n = g_n^2 z_n$ .

最优性能指标函数定义为

$$J_i^*(z_i(t)) = \min_{\alpha_i \in \Omega} \int_t^\infty (z_i^2(\tau) + \alpha_i^2(z_i(\tau))) d\tau, \quad (66)$$

其中  $u = \alpha_n$ . 因此, HJB 方程为

$$H_i\left(z_i, \alpha_i^*, \frac{\partial J_i^*}{\partial z_i}\right) = z_i^2 + \alpha_i^{*2} + \frac{\partial J_i^*}{\partial z_i} \dot{z}_i = 0. \quad (67)$$

这里:  $\alpha_i^*$  为最优控制,  $\alpha_n^* = u^*$ . 类似于 step 1, 可得到

$$\alpha_i = -c_i z_i - z_i \hat{\theta}_i S_i^T S_i - z_i + \hat{\varphi}_{ai}^T \phi_i. \quad (68)$$

其中:  $c_i > 0$  为设计参数,  $u = \alpha_n$ . 参考式 (40) ~ (42), 可得到更新律为

$$\dot{\hat{\theta}}_i = z_i^2 S_i^T S_i - \lambda_i \hat{\theta}_i, \quad (69)$$

$$\dot{\hat{\varphi}}_{ai} = -\phi_i \phi_i^T (\lambda_{ai} (\hat{\varphi}_{ai} - \hat{\varphi}_{ci}) + \lambda_{ci} \hat{\varphi}_{ci}), \quad (70)$$

$$\dot{\hat{\varphi}}_{ci} = -\lambda_{ci} \phi_i \phi_i^T \hat{\varphi}_{ci}, \quad (71)$$

这里  $\lambda_i > 0$ ,  $\lambda_{ai} > 0$ ,  $\lambda_{ci} > 0$  为设计参数.

选择李雅普诺夫函数为

$$L_i = L_{i-1} + \frac{1}{2} z_i^2 + \frac{g_i}{2} \tilde{\theta}_i^2 + L_{E_i}, \quad (72)$$

其中  $\tilde{\theta}_i = \hat{\theta}_i - \theta_i^*$  为估计误差, 当  $i < n$  时,  $g_i = 1$ , 且

$$L_{E_i} = \frac{1}{2} \tilde{\varphi}_{ai}^T \tilde{\varphi}_{ai} + \frac{1}{2} \tilde{\varphi}_{ci}^T \tilde{\varphi}_{ci}. \quad (73)$$

这里:  $\tilde{\varphi}_{ai} = \hat{\varphi}_{ai} - \varphi_{ai}^*$ ,  $\tilde{\varphi}_{ci} = \hat{\varphi}_{ci} - \varphi_{ci}^*$  分别为 actor NN 和 critic NN 的权向量误差. 类似于式 (51) ~ (53), 可得到

$$\begin{aligned} \dot{L}_{E_i} \leq & -\frac{1}{2} \lambda_{ci} \tilde{\varphi}_{ci}^T \phi_i \phi_i^T \tilde{\varphi}_{ci} - \frac{1}{2} \lambda_{ci} \tilde{\varphi}_{ai}^T \phi_i \phi_i^T \tilde{\varphi}_{ai} - \\ & \left( \lambda_{ci} - \frac{1}{2} \lambda_{ai} \right) \hat{\varphi}_{ci}^T \phi_i \phi_i^T \hat{\varphi}_{ci} + \\ & \frac{1}{2} (\lambda_{ai} + \lambda_{ci}) (\varphi_i^{*T} \phi_i)^2 - \frac{1}{2} \lambda_{ai} (\hat{\varphi}_{ai}^T \phi_i)^2. \end{aligned} \quad (74)$$

对式 (72) 求导并结合 (64), 可得到

$$\begin{aligned} \dot{L}_i \leq & -A_{i-1} L_{i-1} + B_{i-1} + z_i \left( W_i^{*T} S_i(Z_i) + \right. \\ & \left. \varepsilon_i(Z_i) + \omega_i + x_{i+1} - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_j} \omega_j - \right. \\ & \left. \frac{1}{4} z_i \sum_{j=1}^{i-1} \left( \frac{\partial \alpha_{i-1}}{\partial x_j} \right)^2 - \frac{1}{4} z_i - d_i \right) + \\ & g_i \tilde{\theta}_i \dot{\hat{\theta}}_i + \dot{L}_{E_i}. \end{aligned} \quad (75)$$

与 step 1 类似, 利用 Young's 不等式, 并由式 (68) 和 (69)、(74) 和 (75), 有

$$\begin{aligned}
\dot{L}_i &\leq -A_{i-1}L_{i-1} + B_{i-1} - \underline{g}_i c_i z_i^2 - \frac{g_i \lambda_i}{2} \tilde{\theta}_i^2 + \\
&\frac{g_i \lambda_i}{2} \theta_i^{*2} + \sum_{j=1}^i \bar{W}_j^2 + \frac{\bar{\varepsilon}_i^2}{4\underline{g}_i} + \frac{1}{4\underline{g}_i} + z_i z_{i+1} - \\
&\frac{1}{2} \lambda_{ci} \lambda_{\phi_i}^{\min} \tilde{\varphi}_{ci}^T \tilde{\varphi}_{ci} - \frac{1}{2} \lambda_{ci} \lambda_{\phi_i}^{\min} \tilde{\varphi}_{ai}^T \tilde{\varphi}_{ai} - \\
&\left( \lambda_{ci} - \frac{1}{2} \lambda_{ai} \right) \tilde{\varphi}_{ci}^T \phi_i \phi_i^T \tilde{\varphi}_{ci} + \\
&\frac{1}{2} (\lambda_{ai} + \lambda_{ci}) (\varphi_i^{*T} \phi_i)^2 - \\
&\frac{1}{2} (\lambda_{ai} - \frac{1}{2}) (\tilde{\varphi}_{ai}^T \phi_i)^2, \quad (76)
\end{aligned}$$

其中  $\lambda_{\phi_i}^{\min}$  为  $\phi_i \phi_i^T$  的最小特征值. 若满足  $c_i > 0$ ,  $\lambda_{ai} > \lambda_{ci} \geq \frac{1}{2} \lambda_{ai} \geq 0$ ,  $\lambda_{ai} \geq \frac{1}{2}$ , 则式 (76) 可写为

$$\dot{L}_i \leq -A_i L_i + B_i. \quad (77)$$

这里:  $z_{n+1} = 0$ ,  $A_i = \min\{A_{i-1}, 2\underline{g}_i c_i, \lambda_i, \lambda_{ci} \lambda_{\phi_i}^{\min}\}$ ,  $B_i = \max\left\{B_{i-1} + \frac{g_i \lambda_i}{2} \theta_i^{*2} + \sum_{j=1}^i \bar{W}_j^2 + \frac{\bar{\varepsilon}_i^2}{4\underline{g}_i} + \frac{1}{4\underline{g}_i} + \frac{1}{2} (\lambda_{ai} + \lambda_{ci}) (\varphi_i^{*T} \phi_i)^2\right\}$ .

基于上述推导过程, 得出如下定理.

**定理 2** 对于满足假设 1 和假设 2 的系统 (1), 若按照式 (38)、(68)、(40) ~ (42) 和 (69) ~ (71) 选取系统的虚拟控制律、实际控制律、自适应律和权学习律, 则: 1) 闭环系统中所有信号均是有界的; 2) 当期望轨迹不违反实际输出约束时, 跟踪误差在  $t \geq T$  后能够满足预设性能; 3) 当期望轨迹违反输出约束时, 系统输出可满足安全性能  $k_d < x_1 < k_u$ , 且满足输出约束  $\rho_2 < x_1 < \rho_1$ ,  $t \geq T$ .

**证明** 1) 有界性.

由式 (77) 可知,  $z_i$ 、 $\tilde{\theta}_i$ 、 $\tilde{\varphi}_{ai}$ 、 $\tilde{\varphi}_{ci}$  ( $i = 1, 2, \dots, n$ ) 是有界的. 由定理 1 可知, 若  $z_1$  有界, 则  $t > 0$  时  $x_1$  有界, 其次  $\alpha_1$ 、 $\hat{\theta}_1$ 、 $\hat{\varphi}_{a1}$ 、 $x_2$  也是有界的, 进而可得到  $x_3, x_4, \dots, x_n$ 、 $\hat{\theta}_2, \hat{\theta}_3, \dots, \hat{\theta}_n$ 、 $\hat{\varphi}_{a2}, \hat{\varphi}_{a3}, \dots, \hat{\varphi}_{an}$ 、 $\alpha_2, \alpha_3, \dots, \alpha_{n-1}$  和  $u$  是有界的.

2) 当期望轨迹不违反输出约束时, 跟踪误差在  $t \geq T$  时满足预设性能, 且满足  $\bar{k}_d(t) < x_1(t) < \bar{k}_u(t)$ .

由式 (10) ~ (15) 可知, 若期望的轨迹没有违反输出约束, 则  $k_d = \bar{k}_d$ ,  $k_u = \bar{k}_u$ , 有

$$\bar{k}_d(t) < x_1(t) < \bar{k}_u(t), \quad t \geq T. \quad (78)$$

由式 (6) 和 (7), 可得到

$$-\vartheta_2(t) < x_1(t) - y_r(t) < \vartheta_1(t). \quad (79)$$

显然, 当  $t \geq T$  时, 跟踪误差  $x_1 - y_r$  满足预设性能.

3) 当期望轨迹违反输出约束时, 系统输出满足安全性能和输出约束.

由于  $z_1$  有界, 由定理 1 可知, 当  $t > 0$  时,  $x_1$  有界, 且当  $t \geq T$  时, 满足安全性能约束  $k_d < x_1 < k_u$ , 由 SBPM 可知,  $k_u \leq \rho_1$ ,  $k_d \geq \rho_2$ , 满足输出约束  $\rho_2 \leq k_d < x_1 < k_u \leq \rho_1$ . 因此, 系统输出满足安全性能和输出约束.  $\square$

### 3 仿真分析

为了验证所提出控制策略的有效性, 考虑文献 [24] 中的单个智能车系统模型为

$$\begin{cases} \dot{x}_1 = x_2 + h_1(x_1), \\ \dot{x}_2 = x_3 + h_2(\bar{x}_2), \\ \dot{x}_3 = \frac{1}{\varsigma} u - \frac{1}{\varsigma} x_3 + \frac{1}{\varsigma} h_3(\bar{x}_3) + \frac{1}{\varsigma} \omega, \\ y = x_1. \end{cases} \quad (80)$$

其中:  $\bar{x}_i = [x_1, x_2, \dots, x_i]^T \in R^i$ ,  $i = 1, 2, 3$ ;  $x_1$ 、 $x_2$ 、 $x_3$  分别为位置、速度和加速度;  $h_1(x_1)$ 、 $h_2(\bar{x}_2)$  和  $h_3(\bar{x}_3)$  为由参数变化和建模误差引起的未知函数;  $\varsigma$  为时间常数;  $\omega$  为有界外部扰动;  $u$  为设计的输入信号 [24].

考虑实际输出约束为

$$\rho_1(t) = \begin{cases} 0.026 \sin(3t) + 1.1, & 0 \leq t < 7.7; \\ 0.026 \sin(3t) + 0.13, & 7.7 \leq t \leq 13.5; \\ 0.026 \sin(3t) + 1.1, & t > 13.5; \end{cases} \quad (81)$$

$$\rho_2(t) = \begin{cases} -0.02 \sin(3t) - 0.9, & 0 \leq t < 11.5; \\ -0.02 \sin(3t) - 0.2, & 11.5 \leq t \leq 14; \\ -0.02 \sin(3t) - 0.9, & t > 14. \end{cases} \quad (82)$$

智能车系统的设计参数为  $c_1 = 5$ ,  $c_2 = 3$ ,  $c_3 = 0.1$ ,  $\lambda_1 = \lambda_2 = \lambda_3 = 1$ ,  $T = 3$ ,  $\lambda_{aj} = 7$ ,  $\lambda_{cj} = 4$ ,  $h_j = 0$ ,  $j = 1, 2, 3$ ,  $\varsigma = 0.0001$ ,  $\sigma = 0.25$ ,  $\omega_0 = \varpi_0 = 0.1$ ,  $\omega_m = \varpi_m = 0.1$ ,  $\rho_m = 0.01$ .  $\vartheta_i(t)$  ( $i = 1, 2$ ) 中参数取  $a_1 = 0.693$ ,  $a_2 = 0.593$ ,  $\vartheta_{T_f} = 0.01$ ,  $q = 5$ ,  $T_f = 5$ . 外部扰动为  $\omega(t) = 2 \sin(\pi t)$ . 期望轨迹为  $y_r(t) = 0.2 \sin(t) + 0.2 \sin(2t)$ . 初始值为  $[x_1(0), x_2(0), x_3(0)]^T = [1, 0, 0]^T$ ,  $[\hat{\theta}_1(0), \hat{\theta}_2(0), \hat{\theta}_3(0)]^T = [1, 1, 1]^T$ .  $\hat{\varphi}_{ai}(0) = [4, \dots, 4]^T \in R^{7^{(6i+3)}}$ ,  $\hat{\varphi}_{ci}(0) = [3, \dots, 3]^T \in R^{7^{(6i+3)}}$  ( $i = 1, 2, 3$ ). RBF NNs  $\hat{W}_i^T S_i$ , actor NNs  $\hat{\varphi}_{ai}^T \phi_i$  和 critic NNs  $\hat{\varphi}_{ci}^T \phi_i$  各包括  $7^{(6i+3)}$  个节点, 中心  $\mu_i$  ( $i = 1, \dots, 7^{(6i+3)}$ ) 平均分布在  $\underbrace{[-3, 3] \times \dots \times [-3, 3]}_{6i+3}$ , 宽度为 2.

按照定理 2 设计系统控制器, 根据 SBSAL 生成安全边界. 为了验证所提出方法的有效性, 分别对是否采用最优跟踪控制 (OTC) 两种情况进行仿真研究, 控制效果如图 1 和图 2 所示. 图 1 为智能车系统

的输出跟踪效果, 图2为控制输入 $u$ .

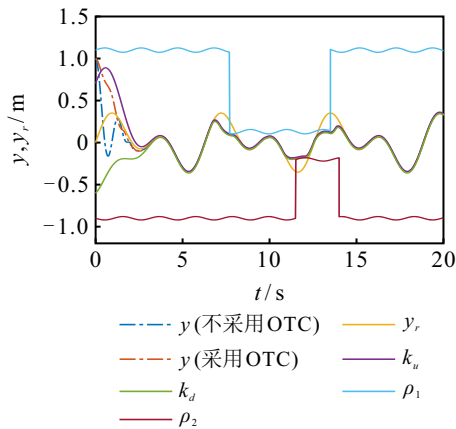


图1 输出跟踪效果

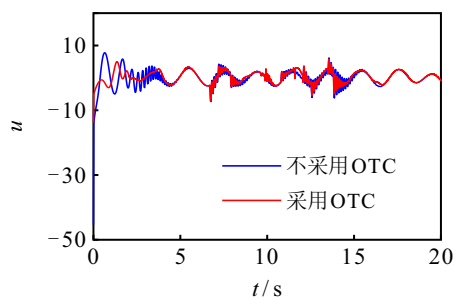


图2 控制输入 $u$

由图1可知, 尽管系统输出初值在约束边界外, 系统输出仍然在3s内进入安全边界约束范围. 当期望轨迹与输出约束不发生冲突时, 系统的输出在停息时间 $T_f$ 内被约束在期望轨迹附近的邻域内. 当期望轨迹即将与输出约束发生冲突时, 控制器根据 SBSAL 自动调整安全边界, 确保了系统的输出满足输出约束. 当突变解除时, 控制器按照 SBSAL 自动调整新的安全边界, 确保了系统的输出快速准确地跟踪上期望轨迹. 由图1可见, 采用 OTC 方法能够达到较好跟踪性能的同时具有更小的超调. 由图2可见, 由于 OTC 的引入, 控制输入大大减少, 系统的能耗降低, 即更小的控制输入也可以获得很好的跟踪结果.

为了验证所提出算法的优越性, 将所提出方法中是否有 BFSM 的仿真结果与文献 [13] 安全跟踪方法的结果进行对比. 控制器仿真参数均取为 $c_1 = 5$ ,  $c_2 = 12$ ,  $c_3 = 0.1$ , 其他参数保持不变. 仿真结果如图3和图4所示. 由图3可见, 在 $t = 13.5$ s时, 文献 [13] 的方法没有考虑到实际输出约束突变解除的情况, 导致系统输出不能快速地跟踪期望轨迹, 而所提出方法考虑了突变解除的情况, 可以快速地跟踪上. 由图4可知, 本文采用最优控制方法使得初始控制输入比文献 [13] 的小, 同时采用 BFSM 后减少了 $u$ 的

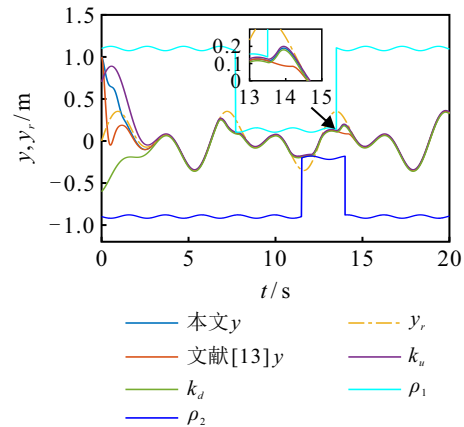


图3 输出曲线以及对比

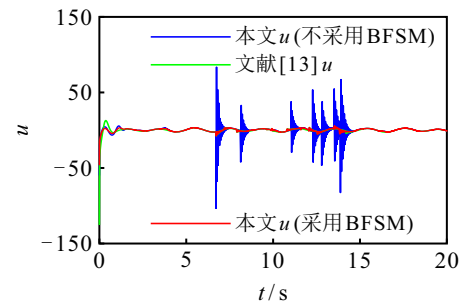


图4 控制输入 $u$ 以及对比

大幅抖动.

为了显示所提出强化学习算法的收敛过程, 根据本文最优性能指标定义每步的奖励函数为 step 1:  $b_1 = z_1^2 + \alpha_1^2$ ; step 2:  $b_2 = z_2^2 + \alpha_2^2$ ; step 3:  $b_3 = z_3^2 + u^2$ . 在上述仿真中, 获得其奖励函数的收敛曲线和 ACNNs 权值的收敛过程如图5~图7所示. 由图5~图7可知, 强化学习的过程是一个收敛过程. 注意到, 在图5中, 奖励函数的收敛过程中出现了抖振, 这主要是实际物理边界突变造成的. 随着突变恢复后收敛过程达到平稳.

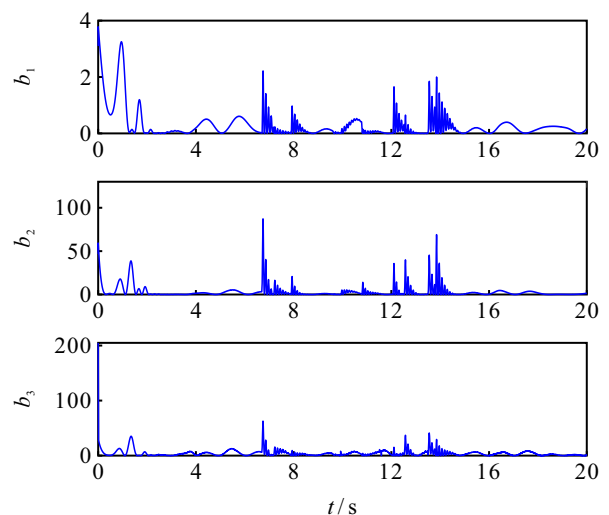


图5 奖励函数的收敛曲线

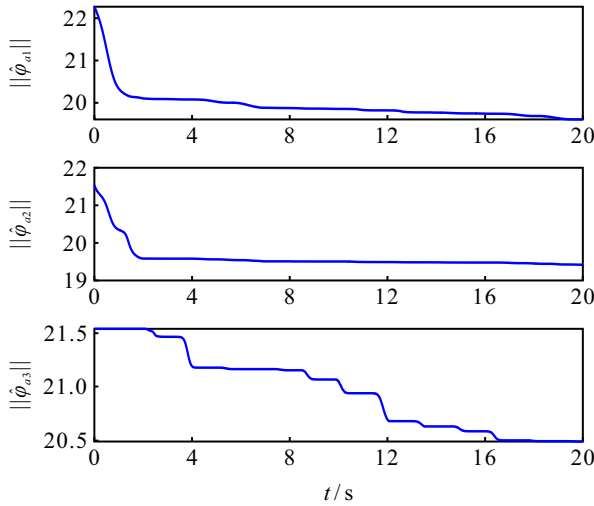


图6 actor NN 权值收敛过程

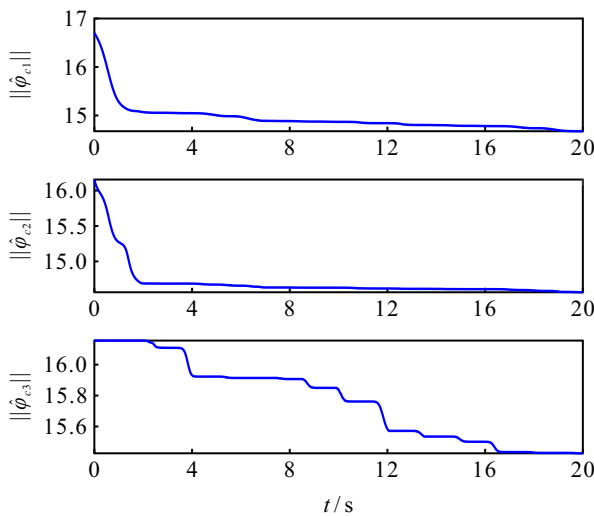


图7 critic NN 权值收敛过程

## 4 结论

本文针对系统实际输出约束的突变及其解除情况, 研究了一类不依赖初始条件的非线性系统预设性能最优安全跟踪控制问题; 提出了一种基于可变障碍函数的 PPC 方法来解决传统 PPC 依赖被约束变量初始值的问题; 同时开发了一个新的 SBSAL, 不仅可以处理实际输出约束发生突变的情况, 而且还可以解决突变解除后系统输出快速准确跟踪期望轨迹的问题, 使得已有的安全跟踪控制策略更为完整. 所提出方法通过引入 RL, 对控制输入进行了优化, 降低了系统的能量损耗. 所提出方法的思想可适用于各类非线性系统的安全跟踪控制, 特别是在无人机和多智能体系统协同控制等领域具有实际应用前景.

## 参考文献 (References)

[1] Wang X, Pang N, Xu Y W, et al. On state-constrained containment control for nonlinear multiagent systems using event-triggered input[J]. *IEEE Transactions on*

*Systems, Man, and Cybernetics: Systems*, 2024, 54(4): 2530-2538.

[2] 刘乐, 孟德宇, 常刘杰, 等. 考虑时变非对称输出约束的机械臂固定时间自适应优化控制[J]. *控制与决策*, DOI: 10.13195/j.kzyjc.2024.0417.

(Liu L, Meng D Y, Chang L J, et al. Fixed-time adaptive optimal control of a robotic arm considering time-varying asymmetric output constraints[J]. *Control and Decision*, DOI: 10.13195/j.kzyjc.2024.0417.)

[3] Cao Y M, Xu N, Wang H Q, et al. Neural networks-based adaptive tracking control for full-state constrained switched nonlinear systems with periodic disturbances and actuator saturation[J]. *International Journal of Systems Science*, 2023, 54(14): 2689-2704.

[4] 王宇梁, 李一平, 李良. 考虑状态约束与执行器饱和的水下机器人轨迹跟踪控制[J]. *控制与决策*, 2024, 39(6): 1778-1786.

(Wang Y L, Li Y P, Li L. Trajectory tracking control of underwater vehicle considering state constraint and actuator saturation[J]. *Control and Decision*, 2024, 39(6): 1778-1786.)

[5] Xu Z B, Xie N G, Shen H, et al. Extended state observer-based adaptive prescribed performance control for a class of nonlinear systems with full-state constraints and uncertainties[J]. *Nonlinear Dynamics*, 2021, 105(1): 345-358.

[6] Xie X J, Guo C, Cui R H. Removing feasibility conditions on tracking control of full-state constrained nonlinear systems with time-varying powers[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, 51(10): 6535-6543.

[7] Wang C X, Wu Y Q, Wang F H, et al. TABLEF-based adaptive control for uncertain nonlinear systems with time-varying asymmetric full-state constraints[J]. *International Journal of Control*, 2021, 94(5): 1238-1246.

[8] Liu B J, Hou M S, Ni J K, et al. Asymmetric integral barrier Lyapunov function-based adaptive tracking control considering full-state with input magnitude and rate constraint[J]. *Journal of the Franklin Institute*, 2020, 357(14): 9709-9732.

[9] Unnikrishnan S, Prasad J V R, Yavrucuk I. Flight evaluation of a reactionary envelope protection system for UAVs[J]. *Journal of the American Helicopter Society*, 2011, 56(1): 12009.

[10] Chen M, Ma H X, Kang Y, et al. Adaptive neural safe tracking control design for a class of uncertain nonlinear systems with output constraints and disturbances[J]. *IEEE Transactions on Cybernetics*, 2022, 52(11): 12571-12582.

[11] Ma H X, Chen M, Yang H, et al. Switched safe tracking control design for unmanned autonomous helicopter with disturbances[J]. *Nonlinear Analysis: Hybrid Systems*, 2021, 39: 100979.

[12] Yao Y G, Tan J Q, Wu J, et al. Decentralized adaptive neural safe tracking control for nonlinear systems with

- conflicted output constraints[J]. *ISA Transactions*, 2023, 137: 263-274.
- [13] 李小华, 邹嵩楠. 基于安全边界保护法的非线性系统自适应安全跟踪控制[J]. *控制与决策*, 2024, 39(9): 3023-3032.  
(Li X H, Zou S N. Adaptive safe tracking control of nonlinear systems based on a secure boundary protection method[J]. *Control and Decision*, 2024, 39(9): 3023-3032.)
- [14] Bechlioulis C P, Rovithakis G A. Robust adaptive control of feedback linearizable MIMO nonlinear systems with prescribed performance[J]. *IEEE Transactions on Automatic Control*, 2008, 53(9): 2090-2099.
- [15] Song Y D, Zhou S Y. Tracking control of uncertain nonlinear systems with deferred asymmetric time-varying full state constraints[J]. *Automatica*, 2018, 98: 314-322.
- [16] Zhang J X, Yang G H. Robust adaptive fault-tolerant control for a class of unknown nonlinear systems[J]. *IEEE Transactions on Industrial Electronics*, 2017, 64(1): 585-594.
- [17] Liu H, Li X H, Liu X P. A bounded-mapping-based prescribed constraint tracking control method without initial condition[J]. *Nonlinear Dynamics*, 2023, 111(4): 3451-3468.
- [18] Li D D, Dong J X. Performance-constrained fault-tolerant DSC based on reinforcement learning for nonlinear systems with uncertain parameters[J]. *Applied Mathematics and Computation*, 2023, 443: 127759.
- [19] Liu G D, Sun N, Yang T, et al. Reinforcement learning-based prescribed performance motion control of pneumatic muscle actuated robotic arms with measurement noises[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, 53(3): 1801-1812.
- [20] Li K W, Li Y M. Fuzzy adaptive optimization prescribed performance control for nonlinear vehicle platoon[J]. *IEEE Transactions on Fuzzy Systems*, 2024, 32(2): 360-372.
- [21] Yu Z Q, Li J X, Xu Y W, et al. Reinforcement learning-based fractional-order adaptive fault-tolerant formation control of networked fixed-wing UAVs with prescribed performance[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 35(3): 3365-3379.
- [22] Zou S N, Li X H, Liu Y. Safe tracking control based on a secure boundary protection method for nonlinear systems with unknown initial tracking condition[J]. *International Journal of Systems Science*, 2024, 55(8): 1655-1672.
- [23] 朱新峰, 丁文武, 张天平. 具有输入量化和全状态约束的非严格反馈随机非线性系统的有限时间动态面控制[J]. *控制与决策*, 2022, 37(10): 2575-2584.  
(Zhu X F, Ding W W, Zhang T P. Finite-time dynamic surface control for nonstrict-feedback stochastic nonlinear systems with input quantization and full-state constraints[J]. *Control and Decision*, 2022, 37(10): 2575-2584.)
- [24] Liu Y, Yao D Y, Li H Y, et al. Distributed cooperative compound tracking control for a platoon of vehicles with adaptive NN[J]. *IEEE Transactions on Cybernetics*, 2022, 52(7): 7039-7048.

#### 作者简介

李小华 (1964-), 女, 教授, 博士, 博士生导师, 主要研究方向为复杂系统结构与控制、非线性系统控制理论以及工业过程建模与控制, E-mail: [lixiaohua6412@163.com](mailto:lixiaohua6412@163.com);

刘莹 (1997-), 女, 硕士生, 主要研究方向为非线性系统预设性能控制、最优控制、安全跟踪控制, E-mail: [yingliu8739@163.com](mailto:yingliu8739@163.com);

邹嵩楠 (1998-), 男, 硕士生, 主要研究方向为非线性系统预设性能控制、有限时间控制、安全跟踪控制, E-mail: [zousongnan@163.com](mailto:zousongnan@163.com).