

ĐIỀU KHIỂN BẮM TỐI ƯU BỀN VỮNG DỰA TRÊN ADP CHO TAY MÁY ROBOT

ADP-BASED ROBUST OPTIMAL CONTROL OF ROBOT MANIPULATORS

Nguyễn Đức Điển^{1,*}, Lại Khắc Lãi²

DOI: <http://doi.org/10.57001/huih5804.2024.287>

TÓM TẮT

Bài báo đề xuất một sơ đồ điều khiển bám tối ưu bền vững dựa trên ADP (Adaptive Dynamic Programming) cho tay máy robot. Đầu tiên, luật điều khiển truyền thẳng được thiết kế để chuyển đổi bài toán điều khiển bám tối ưu bền vững cho tay máy robot thành bài toán điều khiển tối ưu bền vững cho hệ phi tuyến affine. Sau đó, luật điều khiển phản hồi được thiết kế để xác định luật điều khiển tối ưu và luật bù nhiễu. Thuật toán đảm bảo rằng các sai số bám là ổn định UUB (Uniformly Ultimately Bounded), trong khi hàm chi phí hội tụ đến giá trị tối ưu. Cuối cùng, hiệu quả bộ điều khiển đề xuất được kiểm chứng thông qua kết quả mô phỏng.

Từ khóa: Tay máy robot; học tăng cường; quy hoạch động thích nghi; điều khiển bám; điều khiển tối ưu bền vững.

ABSTRACT

This article proposes a robust optimal tracking control scheme for robot manipulators based on ADP (Adaptive Dynamic Programming). First, the feedforward control law is designed to convert the problem of robust optimal tracking control for the robot manipulator into a robust optimal control problem for an affine nonlinear system. Then, the feedback control algorithm is designed to determine the optimal control and disturbance compensation laws. The algorithm ensures that the tracking errors are UUB (Uniformly Ultimately Bounded) while the cost function converges to the optimal value. Finally, the effectiveness of the proposed controller is verified through simulation results.

Keywords: Robot manipulators; Reinforcement Learning (RL); Adaptive Dynamic Programming (ADP); Tracking control; Robust optimal control.

¹Trường Đại học Kinh tế Kỹ thuật - Công nghiệp

²Khoa Điện, Trường Đại học Kỹ thuật Công nghiệp Thái Nguyên

*Email: nddien@uneti.edu.vn

Ngày nhận bài: 15/4/2024

Ngày nhận bài sửa sau phản biện: 05/6/2024

Ngày chấp nhận đăng: 27/9/2024

CHỮ VIẾT TẮT

ADP	Quy hoạch động thích nghi
PD	Bộ điều khiển PD
NN	Mạng nơ-ron
RBF	Mạng RBF (Radial Basis Function)
HJB	Phương trình Hamilton-Jacobi-Bellman
HJI	Phương trình Hamilton-Jacobi-Issac
HOTC	Bộ điều khiển bám tối ưu H _∞
SRBF	Bộ điều khiển trượt thích nghi sử dụng mạng RBF

1. GIỚI THIỆU

Tay máy robot đã trở thành thiết bị quan trọng và mang lại hiệu quả vượt trội trong dây chuyền sản xuất, lĩnh vực y tế, và dịch vụ. Vì vậy, việc thiết kế bộ điều khiển để nâng cao chất lượng điều khiển cho tay máy robot luôn nhận được sự quan tâm của các nhà nghiên cứu [1]. Trong quá trình làm việc, tay máy robot bị ảnh hưởng bởi nhiễu bên ngoài, thay đổi trọng lượng tải, ma sát phi tuyến, những thay đổi không mong muốn về thông số mô hình của hệ thống. Do đó, thuật toán PD (Proportional Derivative) bù trọng trường [1] truyền thống không đảm bảo hiệu suất điều khiển. Các bộ điều khiển nâng cao đã được nghiên cứu và áp dụng cho tay máy robot [2-4]. Trong [2], bộ điều khiển trượt được thiết kế cho tay máy robot với các tham số bất định và nhiễu ngoài. Bộ điều khiển trượt đầu cuối dựa trên kỹ thuật cuốn chiếu đã được đề xuất trong [3,4]. Các bộ điều khiển thông minh dựa trên điều khiển mờ, điều khiển mạng nơ-ron (NN - Neural Network) kết hợp với bộ điều khiển trượt cũng đã được ứng dụng cho tay máy robot [5, 6]. Van và Ge [5] đã sử dụng bộ điều khiển trượt mờ thích nghi, Jie và các cộng sự [6] đã đề xuất bộ điều khiển trượt đầu cuối kết hợp với

mạng RBF (Radial Basis Function). Nói chung, các bộ điều khiển trên đã đảm bảo hiệu quả chất lượng bám quỹ đạo cho tay máy robot với các tham số bất định và nhiễu ngoài. Tuy nhiên, chúng không tối thiểu hàm bất kỳ hàm chi phí nào, tức là chúng không tối ưu.

Đối với bài toán điều khiển tối ưu, ta cần giải được phương trình HJB (Hamilton-Jacobi-Bellman). Một giải pháp cho điều khiển tối ưu bền vững là điều khiển tối ưu H_∞ , ta cần giải được phương trình HJI (Hamilton-Jacobi-Issac). Tuy nhiên, vấn đề giải phương trình HJB/HJI phi tuyến là một thách thức. Gần đây, ADP, một phiên bản của học tăng cường là một phương pháp hữu ích được sử dụng để xấp xỉ trực tuyến nghiệm của phương trình HJB/HJI [7, 8]. Trong [9], một bộ điều khiển bám tối ưu được thiết kế cho tay máy robot, trong đó bộ điều khiển sử dụng cấu trúc ADP với hai hàm xấp xỉ sử dụng 2 NN, nhưng nhiễu ngoài chưa được loại bỏ. Trong [10], một bộ điều khiển bám tối ưu bền vững đã được đề xuất, thuật toán bao gồm bộ điều khiển tối ưu và bộ ước lượng nhiễu, trong đó bộ điều khiển tối ưu sử dụng cấu trúc ADP với 2 NN. Trong [11], bộ điều khiển bám tối ưu H_∞ được xây dựng, trong đó bộ điều khiển sử dụng cấu trúc ADP với ba hàm xấp xỉ sử dụng 3 NN. Để giảm chi phí tính toán, trong [12], một thuật toán điều khiển bám tối ưu chỉ sử dụng một NN duy nhất đã được đề xuất, tuy nhiên nhiễu ngoài lại không được đề cập. Bài báo này giới thiệu một bộ điều khiển bám tối ưu H_∞ (HOCT - H_∞ optimal tracking controller) cho tay máy robot trên cơ sở ADP, luật điều khiển truyền thẳng mới được đề xuất để chuyển đổi bài toán điều khiển bám cho tay máy robot thành bài toán điều khiển tối ưu H_∞ cho một hệ phi tuyến affine và luật điều khiển tối ưu H_∞ được thiết kế trên cơ sở ADP, trong đó luật điều khiển chỉ sử dụng một NN duy nhất thay vì ba để giảm chi phí tính toán.

2. CƠ SỞ LÝ THUYẾT

2.1. Động lực học tay máy robot

Xem xét tay máy robot n bậc tự do với phương trình động lực học được trình bày như sau [1]:

$$M(\psi)\ddot{\psi} + C(\psi, \dot{\psi})\dot{\psi} + G(\psi) + F(\dot{\psi}) = \tau - \tau_0 \quad (1)$$

trong đó $\psi \in \mathbb{R}^{n \times 1}$ là vector vị trí góc khớp, $\dot{\psi} \in \mathbb{R}^{n \times 1}$ là vector vận tốc góc, $\ddot{\psi} \in \mathbb{R}^{n \times 1}$ là vector gia tốc góc, $M(\psi) \in \mathbb{R}^{n \times n}$ là ma trận quán tính đối xứng xác định dương, $C(\psi, \dot{\psi}) \in \mathbb{R}^{n \times n}$ là ma trận Coriolis và ly tâm, $G(\psi) \in \mathbb{R}^n$ là vector lực trọng trường, $F(\dot{\psi}) \in \mathbb{R}^n$ là vector ma sát, $\tau \in \mathbb{R}^{n \times 1}$ là vector mô-men tác động lên các khớp, $\tau_0 \in \mathbb{R}^{n \times 1}$ là vector nhiễu ngoài.

Thuộc tính 1: $M(\psi), C(\psi, \dot{\psi}), \|G(\psi)\| \leq b_G$ là bị chặn bởi $m_1 \leq \|M(\psi)\| \leq m_2, \|C(\psi, \dot{\psi})\| \leq b_C, \|G(\psi)\| \leq b_G$, trong đó m_1, m_2, b_C, b_G là hằng số dương.

Thuộc tính 2: τ_0 có năng lượng hữu hạn, nghĩa là $\tau_0 \in L_2[0, T], 0 < T < \infty$.

Để thuận lợi cho việc thiết kế bộ điều khiển, động lực học (1) được biến đổi thành một hệ thống phi tuyến phản hồi nghiêm ngặt như sau:

$$\begin{cases} \dot{\psi} = f_\psi(\psi) + g_\psi(\psi)u \\ \dot{u} = f_v(\psi, u) + g_v(\psi, u)\tau + k_v(\psi, u)\tau_0 \end{cases} \quad (2)$$

trong đó, u là vector vận tốc góc của biến khớp, $f_\psi(\psi) = 0_{n \times 1}, g_\psi(\psi) = I_n, f_v(\psi, u) = -M^{-1}(Cu + G + F) \in \mathbb{R}^{n \times 1}, g_v(\psi, u) = M^{-1} \in \mathbb{R}^{n \times n}, k_v(\psi, u) = -M^{-1} \in \mathbb{R}^{n \times n}$.

Thuộc tính 3: $f_v(\psi, u)$ bị chặn bởi $\|f_v(\psi, u)\| \leq -b_f \|u\|$, $g_v(\psi, u)$ và $k_v(\psi, u)$ là bị chặn, tức là $\|g_v(\psi, u)\| \leq m_{\min}^{-1}, \|k_v(\psi, u)\| \leq m_{\min}^{-1}$, trong đó b_f là một hằng số dương.

Giả thiết 1: Quỹ đạo vị trí tham chiếu $\psi_d(t)$ là trơn và bị chặn.

Mục tiêu chính của bài toán là thiết kế luật điều khiển bám tối ưu cho hệ thống (2), sao cho $\lim_{t \rightarrow \infty} \|\psi(t) - \psi_d(t)\| \rightarrow 0$ khi nhiễu ngoài bằng 0. Tuy nhiên, các nhiễu ngoài là khác không; do đó, mục tiêu là thiết kế luật điều khiển để làm cho các sai số bám bị chặn bởi độ lợi L_2 [13].

2.2. Thiết kế bộ điều khiển bám tối ưu bền vững

Bộ điều khiển bám tối ưu bền vững cho tay máy robot gồm hai thành phần, đó là luật điều khiển truyền thẳng và luật điều khiển tối ưu bền vững. Luật điều khiển truyền thẳng được thiết kế để xây dựng động lực học sai số bám cho tay máy robot, sau đó bài toán điều khiển bám tối ưu bền vững cho tay máy robot được chuyển thành bài toán điều khiển tối ưu bền vững cho hệ phi tuyến affine. Luật điều khiển tối ưu bền vững được thiết kế trên cơ sở ADP, bao gồm luật điều khiển tối ưu và luật bù nhiễu.

2.2.1. Thiết kế luật điều khiển truyền thẳng

Trong phần này trình bày các bước chuyển đổi hệ (2) sang hệ phi tuyến affine tương đương bằng cách áp dụng kỹ thuật cuốn chiếu [14]. Đầu tiên ta định nghĩa các biến mới như sau: $u_d = u_d^* + u_d^a, \tau = \tau^* + \tau^a$, trong đó u_d là vector đầu vào điều khiển ảo, u_d^* là vector đầu vào điều khiển ảo tối ưu, u_d^a là vector đầu vào điều khiển ảo truyền

thẳng, τ là vector đầu vào điều khiển thực, τ^* là vector đầu vào điều khiển thực tối ưu, τ^a là vector đầu vào điều khiển thực truyền thẳng. Định nghĩa các sai số là $e_\psi = \psi - \psi_d$, $e_u = u - u_d$. Lấy đạo hàm e_ψ và e_u chúng ta có:

$$\begin{cases} \dot{e}_\psi = -\dot{\psi}_d + g_\psi(\psi)u_d^* + g_\psi(\psi)u_d^a + g_\psi(\psi)e_u \\ \dot{e}_u = -\dot{u}_d + f_u(\psi, u) + g_u(\psi, u)\tau + g_u(\psi, u)\tau^a + k_u(\psi, u)\tau_0 \end{cases} \quad (3)$$

Các đầu vào điều khiển truyền thẳng được thiết kế như sau:

$$\begin{cases} u_{da} = g_\psi^{-1}(\psi)(\Lambda_1 e_\psi + \dot{\psi}_d + f_\psi(e_\psi)) \\ \tau_a = g_u^{-1}(\psi, u)[f_u(e_\psi, e_u) + \dot{u}_d - f_u(\psi, u) - g_u^T(\psi)e_\psi - \Lambda_2 e_u] \end{cases} \quad (4)$$

Thay (4) vào (3), động học sai số bám trở thành

$$\begin{cases} \dot{e}_u = \bar{f}_u(e_\psi, e_u) + g_u(\psi, u)\tau^* + k_u(\psi, u)\tau_0 - g_u^T(\psi)e_\psi \\ \dot{e}_\psi = \bar{f}_\psi(e_\psi, e_u) + g_\psi(\psi)u_d^* + g_\psi(\psi)e_u \end{cases} \quad (5)$$

trong đó $\bar{f}_\psi(e_\psi, e_u) = f_\psi(e_\psi) + \Lambda_1 e_u$,

$$\bar{f}_u(e_\psi, e_u) = f_u(e_\psi, e_u) - \Lambda_2 e_u$$

Bổ đề 1: Xem xét động lực sai số bám sau

$$\dot{z} = \bar{f}_{\psi u} + g_{\psi u}u^* + k_{\psi u}d \quad (6)$$

trong đó,

$$z = [e_\psi^T, e_u^T]^T \in \mathbb{R}^{2n \times 1}, \bar{f}_{\psi u} = [\bar{f}_\psi^T(e_\psi, e_u), \bar{f}_u^T(e_\psi, e_u)]^T \in \mathbb{R}^{2n \times 2n},$$

$$u^* = [u_d^{*T}, \tau^{*T}]^T \in \mathbb{R}^{2n \times 1}, u_a = [u_{da}^T, \tau_a^T]^T \in \mathbb{R}^{2n \times 1},$$

$$g_{\psi u} = \text{diag}[g_\psi(\psi), g_u(\psi, u)] \in \mathbb{R}^{2n \times 2n}, k_{\psi u} = \text{diag}[k_\psi(\psi), k_u(\psi, u)] \in \mathbb{R}^{2n \times 2n},$$

$d = [0_{1 \times n}, \tau_0^T]^T \in \mathbb{R}^{2n \times 1}$. Giả sử luật điều khiển tối ưu u^* được thiết kế ổn định hệ thống (6). Trong trường hợp này, bài toán điều khiển bám tối ưu H_∞ cho hệ thống (2) được chuyển đổi thành bài toán điều khiển tối ưu H_∞ cho hệ thống (6) là tương đương.

Chứng minh: Chọn một hàm Lyapunov cho hệ thống (2) như sau:

$$V_1 = \frac{1}{2}e_\psi^T e_\psi + \frac{1}{2}e_u^T e_u \quad (7)$$

Lấy đạo hàm (7) dọc theo (5), ta có được:

$$\begin{aligned} \dot{V}_1 &= e_\psi^T \bar{f}_\psi(e_\psi) + e_\psi^T g_\psi(\psi)u_d^* + e_\psi^T g_\psi(\psi)e_u + e_u^T \bar{f}_u(e_\psi, e_u) \\ &\quad + e_u^T g_u(\psi, u)\tau^* - e_u^T g_u^T(\psi)e_\psi + e_u^T k_u(\psi, u)\tau_0 \\ &= e_\psi^T (\bar{f}_\psi(e_\psi) + g_\psi(\psi)u_d^*) + e_u^T \begin{pmatrix} \bar{f}_u(e_\psi, e_u) \\ +g_u(\psi, u)\tau^* + k_u(\psi, u)\tau_0 \end{pmatrix} \\ &= z^T (\bar{f}_{\psi u} + g_{\psi u}u^* + k_{\psi u}d) \end{aligned} \quad (8)$$

Chọn một hàm Lyapunov cho hệ thống (6) như sau:

$$V_2 = \frac{1}{2}z^T z \quad (9)$$

Lấy đạo hàm (9) dọc theo (6), ta thu được:

$$\dot{V}_2 = z^T (\bar{f}_{\psi u} + g_{\psi u}u^* + k_{\psi u}d) \quad (10)$$

So sánh (8) và (10), có thể thấy rằng nếu luật điều khiển u^* làm cho hệ thống (10) ổn định, tức là $\dot{V}_1 < 0$ thì $\dot{V}_2 < 0$. Do đó, hệ thống (2) cũng ổn định. Hay nói cách khác, bài toán điều khiển bám tối ưu H_∞ cho hệ thống (2) và bài toán điều khiển tối ưu H_∞ cho hệ thống (6) là tương đương.

2.2.2. Thiết kế luật điều khiển tối ưu H_∞

Phần này trình bày vấn đề thiết kế luật điều khiển tối ưu H_∞ cho hệ thống (6). Luật điều khiển tối ưu H_∞ được thiết kế dựa trên phương pháp ADP kết hợp với lý thuyết trò chơi [11].

Định nghĩa 1 [15]: Hệ thống (6) có độ lợi $L_2 \leq \xi$ $\forall d \in L_2[0, \infty]$ nếu

$$\int_0^\infty (z^T Q z + u^T R u) dt \leq \xi^2 \int_0^\infty (d^T d) dt, \quad (11)$$

trong đó, $Q \in \mathbb{R}^{n \times n} > 0$, $R \in \mathbb{R}^{m \times m} > 0$ là ma trận đối xứng, u là xấp xỉ của u^* tại thời điểm t , $\xi > \xi^* > 0$ là mức suy giảm nhiễu, ξ^* là giá trị nhỏ nhất để (6) ổn định.

Trên cơ sở điều kiện (11), hàm chi phí được chọn như sau:

$$J(z, u, d) = \int_t^\infty r(\tau) dt = \int_t^\infty (z^T Q z + u^T R u - \xi^2 d^T d) dt \quad (12)$$

trong đó, $r(\tau) = z^T Q z + u^T R u - \xi^2 d^T d$. Định nghĩa hàm Hamilton cho hệ thống (10) như sau:

$$H(z, u, d, J_z) = r + J_z^T (\bar{f}_{\psi u} + g_{\psi u}u^* + k_{\psi u}d) \quad (13)$$

trong đó, $J_z = \partial J / \partial z$. Lý thuyết trò chơi kết hợp với ADP được sử dụng để xác định hàm giá trị tối ưu bền vững $J^*(z)$ thỏa mãn điều kiện Nash:

$$J^*(z) = \min_u \max_d \int_0^\infty r(\tau) dt = \max_d \min_u \int_0^\infty r(\tau) dt. \quad (14)$$

Khi đó, tồn tại điểm yên ngựa (u^*, d^*) , trong đó u^* là luật điều khiển tối ưu, d^* là luật bù nhiễu. Điểm yên ngựa (u^*, d^*) được xác định bằng cách áp dụng các điều kiện dừng cho (13).

Do đó, luật u^* và d^* được xác định như sau:

$$u^* = -\frac{1}{2}R^{-1}g_{\psi u}^T J_z^*, \quad (15)$$

$$d^* = \frac{1}{2\xi^2} k_{\psi}^T J_z^* \quad (16)$$

Thay (15) và (16) vào (13), ta có phương trình HJI như sau:

$$0 = z^T Qz + J_z^{*T} \bar{f}_{\psi} - \frac{1}{4} (J_z^*)^T g_{\psi} R^{-1} g_{\psi}^T J_z^* + \frac{1}{4\xi^2} (J_z^*)^T k_{\psi} R^{-1} k_{\psi}^T J_z^* \quad (17)$$

$$J^*(0) = 0.$$

Để tìm giải pháp điều khiển tối ưu H_{∞} , người ta cần giải phương trình HJI (17). Tuy nhiên, việc giải phương trình HJI phi tuyến là không thể. Hàm đánh giá $J^*(z)$ được xấp xỉ như sau:

$$J(z) = W^T \phi(z) + \varepsilon(z), \quad (18)$$

trong đó, $W \in \mathbb{R}^N$ là vector trọng số NN, $\phi(z) : \mathbb{R}^n \rightarrow \mathbb{R}^N$ là một vector của N hàm trơn, N số tế bào nơ-ron lớp ẩn, $\varepsilon(z)$ là sai số xấp xỉ hàm. Có thể chọn một tập cơ sở hoàn toàn độc lập $\phi(z)$ thỏa mãn Giả thiết 2.

Giả thiết 2 [16]: Có thể chọn $\phi(z)$ thỏa mãn $\|\phi(z)\| \leq b_{\phi}$, $\|\phi_z\| = \|\partial\phi(z)/\partial z\| \leq b_{\nabla\phi}$, $\|\varepsilon(z)\| \leq b_{\varepsilon}$, $\|\varepsilon_z\| = \|\partial\varepsilon(z)/\partial z\| \leq b_{\nabla\varepsilon}$, trong đó $b_{\phi}, b_{\nabla\phi}, b_{\varepsilon}, b_{\nabla\varepsilon}$ là các hằng số dương.

Lấy đạo hàm (18), ta thu được

$$J_z^* = W^T \frac{\partial\phi(z)}{\partial z} + \frac{\partial\varepsilon(z)}{\partial z} = W^T \phi_z + \varepsilon_z. \quad (19)$$

Thay thế (19) vào (13), phương trình Hamilton (13) trở thành:

$$H^*(z, u^*, d^*, W^T \phi_z) = r(z, u^*, d^*) + W^T \phi_z (\bar{f}_{\psi} + g_{\psi} u^* + k_{\psi} d^*) - \varepsilon_H = 0 \quad (20)$$

trong đó, $\varepsilon_H = -\varepsilon_z (\bar{f}_{\psi} + g_{\psi} u^* + k_{\psi} d^*)$. Trọng số lý tưởng xấp xỉ hàm (18) là chưa biết, do đó $J^*(z)$ được xấp xỉ bởi

$$\hat{J}(z) = \hat{W}^T \phi(z), \quad (21)$$

trong đó, $\hat{W} \in \mathbb{R}^N$ là vector trọng số xấp xỉ hàm. Luật điều khiển (15) và luật bù nhiễu (16) trở thành:

$$\hat{u} = -\frac{1}{2} R^{-1} g_{\psi}^T \phi_z^T \hat{W}, \quad (22)$$

$$\hat{d} = \frac{1}{2\xi^2} k_{\psi}^T \phi_z^T \hat{W}. \quad (23)$$

Định nghĩa e_1 là sai số gây ra bởi xấp xỉ hàm. Sử dụng (21), (22), (23) cho phương trình Hamilton (20), ta thu được:

$$\begin{aligned} & \hat{H}(z, \hat{u}, \hat{d}, \hat{W}^T \phi_z) \\ & = r(z, \hat{u}, \hat{d}) + \hat{W}^T \phi_z (\bar{f}_{\psi} + g_{\psi} \hat{u} + k_{\psi} \hat{d}) = e_1 \end{aligned} \quad (24)$$

Định nghĩa $\tilde{W} = W - \hat{W}$ là sai số xấp xỉ trọng số NN. Từ (20) và (24), ta có

$$e_1 = -\tilde{W}^T \phi_z (\bar{f}_{\psi} + g_{\psi} \hat{u} + k_{\psi} \hat{d}) + \varepsilon_H \quad (25)$$

Để $\hat{W} \rightarrow W$, ta cần điều chỉnh \hat{W} để tối thiểu sai số bình phương $E = (1/2)e_1^T e_1$. Sử dụng thuật toán suy giảm độ dốc chuẩn hóa, luật cập nhật trọng số \hat{W} được xác định như sau [17]:

$$\text{Nếu } z^T (\bar{f}_{\psi} + g_{\psi} \hat{u} + k_{\psi} \hat{d}) < 0 \quad \text{thì}$$

$$\dot{\hat{W}} = -\alpha_1 \frac{\sigma}{(\sigma^T \sigma + 1)^2} (\sigma^T \hat{W} + z^T Qz + \hat{u}^T R \hat{u} - \gamma^2 \hat{d}^T \hat{d}). \quad (26)$$

$$\text{Nếu } z^T (\bar{f}_{\psi} + g_{\psi} \hat{u} + k_{\psi} \hat{d}) \geq 0, \quad \text{thì}$$

$$\dot{\hat{W}} = \hat{W} + \alpha_2 \phi_z (G - K)z, \quad (27)$$

trong đó, $\sigma = \phi_z (\bar{f}_{\psi} + g_{\psi} \hat{u} + k_{\psi} \hat{d})$, $G = \frac{1}{2} g_{\psi} R^{-1} g_{\psi}^T$, $K = \frac{1}{2\xi^2} k_{\psi} k_{\psi}^T$, $\alpha_1 > 0$, $\alpha_2 > 0$.

Định lý 1: Xem xét động lực của sai số bám được xác định bởi (6). Giả thiết 1, 2 được thỏa mãn. Hàm giá trị được xác định bởi (21), luật điều khiển tối ưu được đưa ra bởi (22), luật bù nhiễu được đưa ra bởi (23), trong đó các trọng số NN được điều chỉnh trực tuyến bởi (26) và (27). Khi đó, HOCT đảm bảo rằng các sai số bám và sai số xấp xỉ là ổn định UUB.

Chứng minh: Chọn hàm Lyapunov cho hệ thống (6) như sau:

$$V_3 = \frac{1}{2} \alpha_2 z^T z + \frac{1}{2} \text{trace}(\tilde{W}^T \tilde{W}) \quad (28)$$

Thực hiện đạo hàm V_3 theo thời gian, ta có

$$\begin{aligned} \dot{V}_{31} & = \alpha_2 z^T \bar{f}_{\psi} - \alpha_2 z^T (G - K) \phi_z^T \hat{W} \\ & \quad + \tilde{W}^T (-\alpha_1 \bar{\sigma} (\bar{\sigma}^T \tilde{W} - \bar{\varepsilon}_H)) \end{aligned} \quad (29)$$

trong đó, $\bar{\sigma} = \sigma / (\sigma^T \sigma + 1)$, $\bar{\varepsilon}_H = \varepsilon_H / (\sigma^T \sigma + 1)$. Bởi

Thuộc tính 5, ta có $\alpha_2 z^T \bar{f}_{\psi}$ là bị chặn bởi $(-\beta_3 \|z\|^2)$, tức là, $\alpha_2 z^T \bar{f}_{\psi} \leq -\beta_1 \|z\|^2$, trong đó $\beta_1 = \alpha_2 b_{\bar{f}}$, $\|\bar{f}_{\psi}\| \leq -b_{\bar{f}}$. Áp dụng bất đẳng thức Young, (29) trở thành $\dot{V}_{31} \leq -\beta_1 \|z\|^2 - \beta_2 \|\tilde{W}\|^2 + \beta_3$, trong đó $\psi = \bar{\sigma} \bar{\sigma}^T$, $b_{\varepsilon_H} > 0$ là chặn trên của $\bar{\varepsilon}_H$, $\beta_2 = (\alpha_1 - 1) \lambda_{\min}(\psi)$,

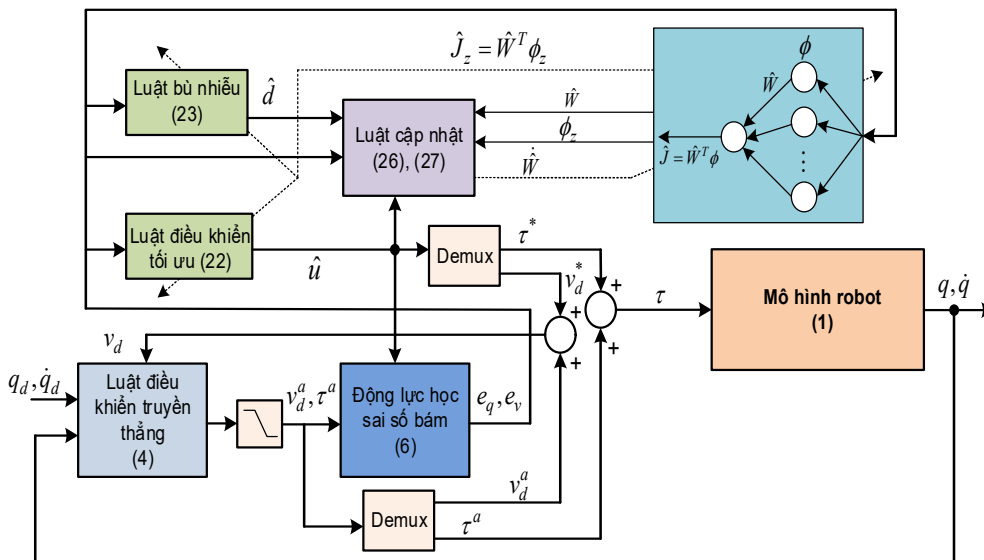
$\beta_3 = \alpha_2 (G_{\max} + K_{\min}) b_{\Delta\phi} b_w + \frac{\alpha_1^2}{4} b_{\varepsilon H}^2$, $\alpha_1 > 1$. Do đó $\dot{V}_3 \leq 0$, nếu và chỉ nếu $\|z\| \geq \sqrt{\beta_3/\beta_1} = b_z$ hoặc $\|\tilde{W}\| \geq \sqrt{\beta_3/\beta_2} = b_{\tilde{W}}$. Ta có thể thấy rằng $\|z\|$ hay $\|\tilde{W}\|$ vượt qua tập đóng b_z hay $b_{\tilde{W}}$, thì $\dot{V}_3 \leq 0$. Như vậy các sai số bám và sai số xấp xỉ là UUB.

Dựa trên Bổ đề 1 và Định lý 1, sơ đồ cấu trúc của thuật toán đề xuất có thể được trình bày như Hình 1. Bộ điều khiển gồm 2 thành phần: Luật điều khiển truyền thẳng và luật điều khiển tối ưu bền vững. Trước tiên, luật điều khiển truyền thẳng được xác định như (4), sau đó động lực học sai số bám được xây dựng như (6), hàm chi phí được xấp xỉ bởi NN với luật cập nhật trọng số như (26), (27) và luật điều khiển tối ưu được xác định như (22) và luật điều khiển bù nhiễu xác định như (23).

mong muốn $E_d = (X_d, Y_d)$ được chọn là $X_d = 1 + 0,5\sin(0,5t + \pi/2)$, $Y_d = 1 + 0,5\cos(0,5t + \pi/2)$. Quỹ đạo vị trí tham chiếu q_d được xác định bởi phương trình động học ngược $q_{d2} = \arccos(X_d^2 + Y_d^2 - l_1^2 - l_2^2)/(2l_1l_2)$, $\psi_{d1} = \arctan(Y_d/X_d - \arctan(l_2 \sin(\psi_{d2}))/l_1 + l_2 \cos(\psi_{d2}))$, trong đó $l_1 = l_2 = 1m$. Quỹ đạo thực tế của E được xác định bởi phương trình động học thuận $X = l_1 \cos(\psi_1) + l_2 \cos(\psi_1 + \psi_2)$, $Y = l_1 \sin(\psi_1) + l_2 \sin(\psi_1 + \psi_2)$. Chọn $\psi(0) = [0,5, -0,5]^T$, $\dot{\psi}(0) = [0, 0]^T$, τ_d có giá trị ngẫu nhiên trong khoảng $[-1, 1]Nm$. Chọn các tham số của HOTC như sau: $\Lambda_1 = \Lambda_2 = \text{diag}[1,1]$, $\phi(z) = [e_{\psi_1}^2, e_{\psi_1}, e_{\psi_2}^2, e_{\psi_2}, e_{v_1}^2, e_{v_1}, e_{v_2}^2, e_{v_2}, e_{\tau}^2, e_{\tau}, e_{v_1}, e_{v_1}, e_{v_2}, e_{v_2}, e_{\tau}^2, e_{\tau}]$, $Q = I_4, R = I_4, \alpha_1 = 50, \alpha_2 = 0,01, \xi = 0,1$.

Thực hiện mô phỏng trên Matlab với thời gian mô phỏng 50s, thời gian tắt nhiễu PE là 20s, thời gian lấy mẫu $T = 0,1s$.

Hình 2 là trình bày sự hội tụ của các trọng số, nó cho thấy rằng các trọng số hội tụ sau khoảng thời gian 12s. Hiệu suất điều khiển bám được thể hiện trên hình 3, 4 và các sai số bám được thể hiện trên hình 5. Chúng cho thấy rằng sau khi thuật toán hội tụ, HOTC cung cấp hiệu suất điều khiển bám tốt, với sai số bám không vượt quá $4.10^{-3}rad$. Hình 6 trình bày kết



Hình 1. Sơ đồ cấu trúc điều khiển bám tối ưu bền vững cho tay máy robot trên cơ sở ADP

3. KẾT QUẢ MÔ PHỎNG VÀ THẢO LUẬN

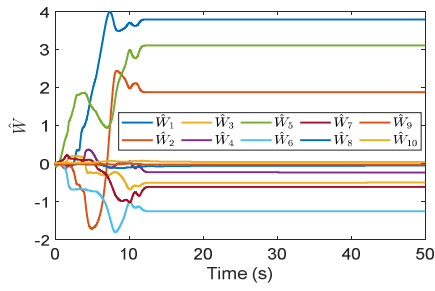
Xét một robot 2 bậc tự do [11] với các ma trận của phương trình động lực học là:

$$\begin{aligned}
 M(\psi) &= \begin{bmatrix} h_1 + 2h_3c_2 & h_2 + h_3c_2 \\ h_2 + h_3c_2 & h_2 \end{bmatrix}, \\
 C(\psi, \dot{\psi}) &= \begin{bmatrix} -h_3s_2\dot{\psi}_2 & -h_3s_2(\dot{\psi}_1 + \dot{\psi}_2) \\ h_3s_2\dot{\psi}_1 & 0 \end{bmatrix} \\
 G(\psi) &= [8,45 \tanh(\psi_1) \quad 2,35 \tanh(\psi_2)]^T
 \end{aligned} \tag{30}$$

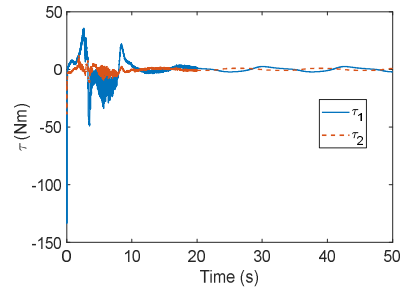
trong đó, $c_2 = \cos(\psi_2)$, $s_2 = \sin(\psi_2)$, $h_1 = 3,473 \text{ kgm}^2$, $h_2 = 0,196 \text{ kgm}^2$, $h_3 = 0,242 \text{ kgm}^2$. Gọi $E = (X, Y)$ là tọa độ của khâu cuối trong không gian làm việc. Quỹ đạo

quả của các đầu vào điều khiển tối ưu và hình 7 trình bày kết quả của mô-men điều khiển. Quỹ đạo bám trong không gian làm việc được thể hiện trên hình 8.

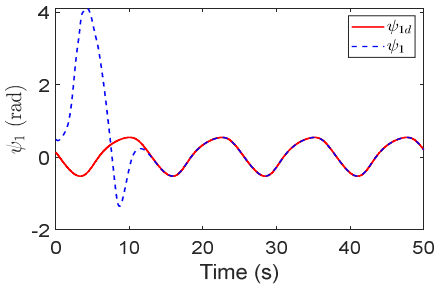
Thực hiện mô phỏng so sánh HOTC với bộ điều khiển trượt thích nghi sử dụng RBF (SRBF) [18], các kết quả mô phỏng so sánh được thể hiện trên các hình 9, 10 và 11. Quan sát các hình con trong hình 9, ta thấy các sai số bám của SRBF có giá trị lớn hơn HOTC. Quỹ đạo bám trong không gian làm việc của HOTC và SRBF được trình bày trên hình 10. Mô-men điều khiển của SRBF được trình bày trên hình 11, cho thấy rằng các giá trị mô-men điều khiển bị dao động mạnh, trong khi đó HOTC cung cấp các mô-men điều khiển trơn hơn, không bị dao động khi hội tụ (hình 7). Thông qua các kết quả mô phỏng, ta thấy rằng HOTC cung cấp hiệu suất điều khiển tốt hơn SRBF. Như vậy, hiệu quả của HOTC được xác minh.



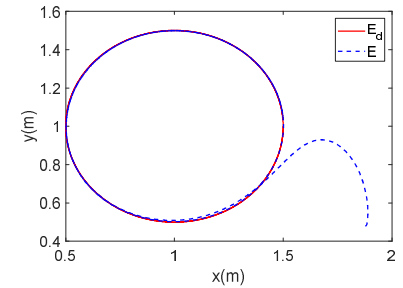
Hình 2. Sự hội tụ của các trọng số



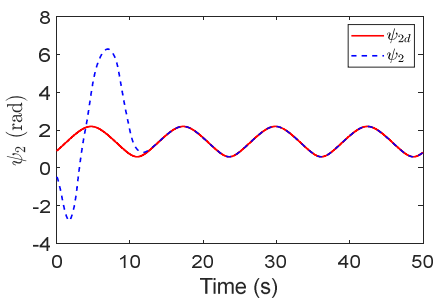
Hình 7. Mô-men điều khiển



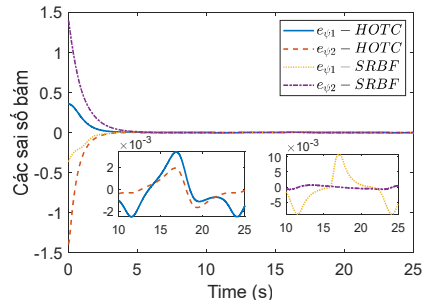
Hình 3. Quỹ đạo bám ψ_1



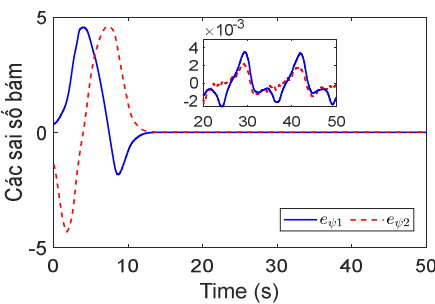
Hình 8. Quỹ đạo E trong không gian làm việc khi các trọng số hội tụ



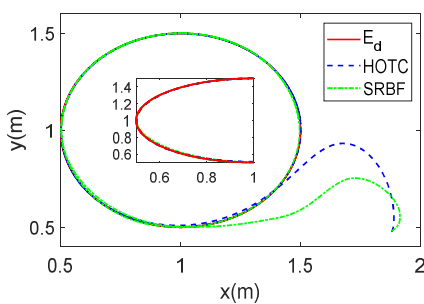
Hình 4. Quỹ đạo bám ψ_2



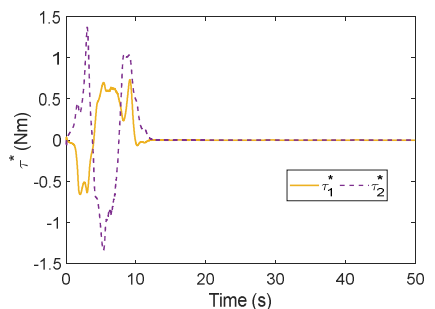
Hình 9. So sánh các sai số bám của HOTC và SRBF



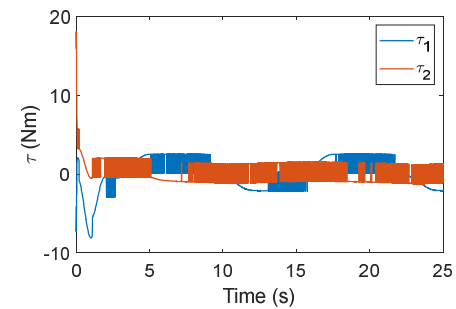
Hình 5. Các sai số bám



Hình 10. Quỹ đạo bám trong không gian làm việc của HOTC và SRBF



Hình 6. Đầu vào điều khiển tối ưu



Hình 11. Mô-men điều khiển của SRBF

4. KẾT LUẬN

Bài báo đã đề xuất một sơ đồ điều khiển bám tối ưu H_∞ dựa trên ADP cho tay máy robot với các nhiễu ngoài. Luật điều khiển truyền thẳng đã được thiết kế để chuyển đổi bài toán điều khiển bám tối ưu bền vững cho tay máy robot thành bài toán điều khiển tối ưu bền vững cho hệ phi tuyến affine. Sau đó, luật điều khiển phản hồi đã được đề xuất để xác định luật điều khiển tối ưu và luật bù nhiễu. Thuật toán đảm bảo rằng các sai số bám là ổn định UUB. Kết quả mô phỏng đã cho thấy bộ điều khiển đề xuất đảm bảo các yêu cầu đặt ra. Công việc tiếp theo, HOTC sẽ được cài đặt thử nghiệm trên tay máy robot thực.

TÀI LIỆU THAM KHẢO

- [1]. F. L. Lewis, D. M. Dawson, C. T. Abdallah, *Robot Manipulator Control Theory and Practice*. CRC Press, 2003.
- [2]. D. Chwa, H. Kwon, "Nonlinear Robust Control of Unknown Robot Manipulator Systems With Actuators and Disturbances Using System Identification and Integral Sliding Mode Disturbance Observer," *IEEE Access*, 10, 35410-35421, 2022.
- [3]. M. Van, M. Mavrouniotis, S. S. Ge, "An adaptive backstepping nonsingular fast terminal sliding mode control for robust fault tolerant control of robot manipulators," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49, 7, 1448-1458, 2018.
- [4]. T. N. Truong, A. T. Vo, H. J. Kang, "A backstepping global fast terminal sliding mode control for trajectory tracking control of industrial robotic manipulators," *IEEE Access*, 9, 31921-31931, 2021.
- [5]. M. Van, S. S. Ge, "Adaptive fuzzy integral sliding-mode control for robust fault-tolerant control of robot manipulators with disturbance observer," *IEEE Transactions on Fuzzy Systems*, 29, 5, 1284-1296, 2021.
- [6]. W. Jie, L. M. Cheol, K. Jaehyung, K. H. Hee, "Fast fractional-order terminal sliding mode control with rbfnn based sliding perturbation observer for 7-dof robot manipulator," *IEEE Access*, 9, 67117-67128, 2021.
- [7]. K. G. Vamvoudakis, F. L. Lewis, N. Control, "Online solution of nonlinear two-player zero-sum games using synchronous policy iteration," *International Journal of Robust*, 22, 13, 1460-1483, 2012.
- [8]. K. G. Vamvoudakis, F. L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, 46, 5, 878-888, 2010.
- [9]. R. Kamalapurkar, H. Dinh, S. Bhasin, Warren E Dixon, "Approximate optimal trajectory tracking for continuous-time nonlinear systems," *Automatica*, 51, 40-48, 2015.
- [10]. V. T. Vu, P. N. Dao, P. T. Loc, T. Q. Huy, "Sliding variable-based online adaptive reinforcement learning of uncertain/disturbed nonlinear mechanical systems," *Journal of Control, Automation Electrical Systems*, 32, 2, 281-290, 2021.
- [11]. H. Modares, F. L. Lewis, Z. P. Jiang, " H_∞ tracking control of completely unknown continuous-time systems via off-policy reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, 26, 10, 2550-62, 2015.
- [12]. X. Long, Z. He, Z. Wang, "Online optimal control of robotic systems with single critic NN-based reinforcement learning," *Complexity*, 1-7, 2021.
- [13]. Q. Jiao, H. Modares, S. Xu, F. L. Lewis, K. G. Vamvoudakis, "Multi-agent zero-sum differential graphical games for disturbance rejection in distributed control," *Automatica*, 69, 24-34, 2016.
- [14]. H. Zargarzadeh, T. Dierks, S. Jagannathan, "Adaptive neural network-based optimal control of nonlinear continuous-time systems in strict-feedback form," *International Journal of Adaptive Control and Signal Processing*, 28, 3-5, 305-324, 2014.
- [15]. A. J. Van Der Schaft, " L_2 -gain analysis of nonlinear systems and nonlinear state feedback H_∞ control," *IEEE Transactions on Automatic Control*, 37, 6, 770-784, 1992.
- [16]. L. N. Tan, T. C. Pham, "Optimal Tracking Control for PMSM With Partially Unknown Dynamics, Saturation Voltages, Torque, and Voltage Disturbances," *IEEE Transactions on Industrial Electronics*, 69, 4, 3481-3491, 2021.
- [17]. N. T. Luy, N. T. Thanh, H. M. Tri, "Reinforcement learning-based intelligent tracking control for wheeled mobile robot," *Transactions of the Institute of Measurement Control*, 36, 7, 868-877, 2014.
- [18]. J. Liu, *Intelligent Control Design and MATLAB Simulation*. Springer, 2018.

AUTHORS INFORMATION

Nguyen Duc Dien¹, Lai Khắc Lai²

¹University of Economics - Technology for Industries, Vietnam

²Faculty of Electrical Engineering, Thainguyen University of Technology, Vietnam