2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

# Robust Adaptive Dynamic Programming for Dual-redundant Actuation Systems in Aircrafts

Jiaxing Wang, Yue Guo, and Weiguo Xia

*Abstract*— **Redundancy actuation systems are widely applied in the fields of aviation and aerospace. This paper establishes a nonlinear dynamic model of the dual-redundancy actuation system driven by electro-mechanical actuators, involving control constraints and unmatched bounded perturbation. We next focus on the reset control of the dual-redundancy actuation system and raise a robust control problem, which can be transformed into a specified optimal controller design problem theoretically. To acquire the optimal controller, we propose a modified robust adaptive dynamic programming approach. Using multiple critic neural networks, a simultaneous approximation of the value function and the controller are achieved, and all the weight parameters can be estimated through an optimization framework, which is feasible guaranteed via well-defined basis functions. Finally, a numerical example from the dual-redundancy actuation system is presented to illustrate the derived controller.**

*Index Terms*— **Dual-redundancy actuation systems, Optimal control, Robust adaptive dynamic programming, Neural network, Optimization.**

## I. INTRODUCTION

In recent years, the growing and diversified requirements of aviation and aerospace missions have put forward higher performance for automatic flight control systems [1]–[3]. To enhance the reliability of automatic flight control systems, redundancy design is adopted in different actuation systems. For example, a control surface on ailerons is broadly driven by multiple actuators in parallel, and that forms redundancy actuation systems (RASs). As pivotal power units, actuators determine directly the operation performance of RASs. In contrast to the hydraulic actuator (HA) and electro-hydraulic actuator (EHA), electro-mechanical actuator (EMA), known for compact size, no oil leakage, and high power-to-weight ratio, has been treated as a crucial component in the development of multi-electric aircrafts and all-electric aircrafts.

Numerous works have paid attention to the modeling and control of the RASs in aircrafts. In [4], a linear model of the RASs driven by HAs is established. Based on the fractional order approach, a cascade controller is proposed in [5] and compensates the force signals to the position feedback loop, which synchronizes the dynamics of the EHAs in different channels. To reduce the force fight between the actuators, a cross-coupling control synthesis with force difference compensation is provided in [6]. Nevertheless, fewer works have focused on the RASs driven by EMAs. Furthermore, linear models are generally employed to simplify the dynamics of actuators, neglecting the nonlinear friction, the external load characteristics, and the stiffness of mechanical components, and this model inaccuracy results in the poor robustness of the designed controller.

In this paper, a nonlinear model of the dual-redundancy actuation system (DRAS) driven by EMAs is established, involving the control constraint and the unmatched bounded perturbation. In practice, the above issues stem from the space and power restrictions and operation environment, respectively, which are inherent in the dynamics of DRASs. Based on the nonlinear model, we focus on the reset control of DRASs and raise a robust control problem. In view of the control constraint and unmatched perturbation, it is difficult to design a controller directly. Inspired by the works in [7]–[10], the above control synthesis can be converted to solving the Hamilton-Jacobi-Bellman (HJB) equation. However, the closed-form solution to the HJB equation is almost unavailable due to the nonlinear property and dimensionality curse. In order to obtain the numerical solution, adaptive dynamic programming (ADP) is first proposed in [11], which can be applied both on-line and off-line. By means of the nonlinear mapping capacity of neural networks (NNs), ADP is high-performance and forward optimal [12]–[14].

In the literature, numerous ADP schemes are developed to address the optimal control problems of dynamical systems. A simultaneous policy iteration (PI) algorithm is provided in [9] to design a robust controller for nonlinear systems with unmatched perturbation. In [14], an on-line procedure is improved for uncertain systems, and the uniform stability is guaranteed. Based on the least square technique, a novel PI implementation with the persistent excitation is proposed in [15]. In this paper, we propose a modified robust ADP approach for dynamical systems subject to the control constraint and unmatched perturbation simultaneously. Using multiple critic NNs, the simultaneous approximation of the value function and the constrained controller are achieved. The NN weight parameters can be estimated by an optimization framework. Based on the characteristics of the DRAS model, we choose a sequence of well-defined basis functions to guarantee the feasibility of the designed optimization problem. Furthermore, the proposed approach is suitable for both on-line and off-line implementations, and its efficacy is verified by the reset control of DRASs in an off-line setting.

The rest of this paper is organized as follows. In section

This work was supported in part by the National Natural Science Foundation of China (62122016) and LiaoNing Science and Technology Program (2023JH2/101700361).

J, Wang, Y, Guo and W, Xia are with Key Laboratory of Intelligent Control and Optimization for Industrial Equipment of Ministry of Education and School of Control Science and Engineering, Dalian University of Technology, Dalian, China (wanna@mail.dlut.edu.cn; yueg@mail.dlut.edu.cn; wgxiaseu@dlut.edu.cn).

II, we establish a nonlinear model of the DRAS driven by EMAs, and then an optimal control synthesis is formulated. Section III shows a modified robust ADP methodology for nonlinear systems subject to both control constraint and unmatched perturbation. The derived results are demonstrated in Section IV. Conclusions are drawn in Section V.

*Notations:* Throughout this paper, let $\mathbb{R} := (-\infty, +\infty)$ and $\mathbb{R}^+ := [0, +\infty)$. Given a vector $x \in \mathbb{R}^n$, $\|x\|$ denotes the Euclidean norm of $x$. $I_n$ denotes the $n$-dimensional identity matrix. $\mathbf{0}_{n\times m} \in \mathbb{R}^{n\times m}$ is a matrix with all its elements equal to 0. For a matrix $A \in \mathbb{R}^{n\times m}$, $A^+ \in \mathbb{R}^{m\times n}$ denotes the Moore-Penrose inverse of $A$.

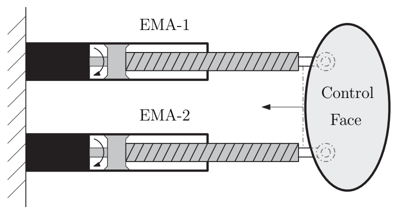


Fig. 1: The schematic diagram of the DRAS driven by two identical EMAs.

## II. Dual-redundancy actuation systems

In this section, a nonlinear model of the dual-redundancy actuation system driven by two identical electro-mechanical actuators is established. To achieve the specified control objectives of dual-redundancy actuation systems, we formulate a robust control problem, and then convert it to an optimal control problem.

### A. Mathematical Modelling

Hereinafter, a dual-redundancy actuation system (DRAS) driven by two identical electro-mechanical actuators (EMAs) is considered, and its schematic diagram is shown in Fig. 1. By synthesizing the stiffness and friction of the mechanical components and the air resistance, a nonlinear dynamics model is established as follows:

$$\begin{cases} J\ddot{\theta}_1 = T_1 - T_l(\theta_1, l_d) - T_f(\dot{\theta}_1), \\ J\ddot{\theta}_2 = T_2 - T_l(\theta_2, l_d) - T_f(\dot{\theta}_2), \\ M\ddot{l}_d = F_l(\theta_1, l_d) + F_l(\theta_2, l_d) - F_a(\dot{l}_d), \end{cases} \tag{1}$$

where $\theta_i$ and $T_i$ are respectively the angular rotation of the screw and the output torque of the motor in the EMA-$i$ for $i = 1, 2$, and $l_d$ is the linear displacement of the load. Due to the restrictions of safety operation and power supply, state constraints are presented as $\|\theta_i\| \leq \vartheta$ and $\|l_d\| \leq L$, and the input complies with $\|T_i\| \leq \lambda$, where $\vartheta, L, \lambda > 0$. $J$ is the rotational inertia of the rotating component; $M$ is the mass of the load. Let $\sigma = \frac{p}{2\pi}$, where $p$ is the screw lead. For $i = 1, 2$, $F_l(\theta_i, l_d) := k(\sigma\theta_i - l_d)$ denotes the force between the EMA-$i$ and the load, and its equivalent torque is defined as $T_l(\theta_i, l_d) := \frac{\sigma}{\eta} F_l(\theta_i, l_d)$, where $k$ is the bending stiffness of the load surface, and $\eta$ is the efficiency coefficient of the screw. The compound friction is formulated as

$$T_f(\dot{\theta}_i) := \gamma_1 \Big(\tanh(\gamma_2\dot{\theta}_i) - \tanh(\gamma_3\dot{\theta}_i)\Big) + \gamma_4 \tanh(\gamma_5\dot{\theta}_i) + \gamma_6\dot{\theta}_i, \quad \forall i = 1, 2,$$

where $\gamma_j > 0$ for $j \in \{1, 2, \ldots, 6\}$; see [16] for more details regarding this friction model. $F_a(\dot{l}_d) := \frac{1}{2}\Sigma_a \dot{l}_d^2$ denotes the air resistance on the load, where $\Sigma_a > 0$ is associated with the air density and drag coefficient and the effective frontal area of the load.

We denote $x = [\,\dot{\theta}_1, \theta_1, \dot{\theta}_2, \theta_2, \dot{l}_d, l_d\,]^\top$ and $u = [T_1, T_2]^\top$. In addition, $d = [T_f(\dot{\theta}_1), T_f(\dot{\theta}_2), F_a(\dot{l}_d)]^\top$ is regarded as the perturbation term, For simplicity, let $a = \frac{\sigma k}{\eta J}$ and $b = \frac{k}{M}$, and the dynamics of the DRAS (1) can be rewritten as

$$\dot{x} = F(x) + Gu + Hd(x), \tag{2}$$

where

$$F(x) := \begin{bmatrix} -a\sigma x_2 + a x_6 \\ x_1 \\ -a\sigma x_4 + a x_6 \\ x_3 \\ b\sigma x_2 + b\sigma x_4 - 2b x_6 \\ x_5 \end{bmatrix}, \quad G := \begin{bmatrix} \frac{1}{J} & 0 \\ 0 & 0 \\ 0 & \frac{1}{J} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$H := \begin{bmatrix} -\frac{1}{J} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{J} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{M} & 0 \end{bmatrix}^\top.$$

The control input $u$ takes values from the set $\mathbb{U} = \{u \in \mathbb{R}^2 : \|u_i\| \leq \lambda,\ i = 1, 2\}$. Note that the perturbation $d$ is state-dependent and bounded. In addtion, $d$ vanishes at the origin, that is, $d(0) = 0$. Hence, it can be inferred that there exists a continuous function $\bar{d} : \mathbb{R} \to \mathbb{R}^+$ such that $\|d(x)\| \leq \bar{d}(x)$ and $\bar{d}(0) = 0$. We define the set $\mathbb{D} = \{d \in \mathbb{R}^3 : \|d(x)\| \leq \bar{d}(x)\}$, and $\bar{d}$ is determined empirically via sufficient real-world experiments.

### B. Problem Statement and Control Objectives

We next investigate the following robust control problem of the system (2).

*Problem 1:* Design a feedback control policy $u(x) \in \mathbb{U}$ such that the system (2) is asymptotically stable (AS) with the perturbation $d \in \mathbb{D}$.

Problem 1 describes a reset control requirement for the DRAS. Since the system (2) suffers from both the control constraint $u \in \mathbb{U}$ and the nonlinear perturbation $d \in \mathbb{D}$. In particular, $d$ is unmatched due to $G \neq H$, thus it is generally difficult to design a robust controller directly. According to [7]–[9], such a robust control problem can be transformed into a specified optimal control problem, which provides an effective method for handling the robust control synthesis via optimal techniques.

Without loss of generality, a nonlinear control system is considered as follows:

$$\dot{x} = f(x) + g(x)\bar{u} + k(x)w(x), \tag{3}$$

where $x \in \mathcal{X} \subset \mathbb{R}^n$ is the system state, and $\bar{u} \in \mathcal{U} \subset \mathbb{R}^m$ is the control input. $\mathcal{X}$ and $\mathcal{U}$ are compact sets containing the origin. $w \in \mathcal{W} \subset \mathbb{R}^q$ is an external perturbation satisfying $\mathcal{W} := \{w \in \mathbb{R}^q : \|w(x)\| \leq \bar{w}(x)\}$, where $\bar{w} : \mathbb{R}^n \to \mathbb{R}^+$ is continuous with $\bar{w}(0) = 0$. Assume that $f : \mathbb{R}^n \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ and $k : \mathbb{R}^n \to \mathbb{R}^{n\times q}$ are Lipschitz continuous. Let $f(0) = 0$ and $\psi(x) := g^+(x)k(x)w(x)$, and assume that there exists a continuous function $\overline{\psi} : \mathbb{R}^n \to \mathbb{R}^+$ such that $\|\psi(x)\| \leq \overline{\psi}(x)$ satisfying $\overline{\psi}(0) = 0$.

We next construct an auxiliary system of the system (3). To this end, the perturbation term $k(x)w(x)$ is divided into matched and unmatched components as follows:

$$\begin{aligned} k(x)w(x) = g(x)g^+(x)k(x)w(x) \\ + (I_n - g(x)g^+(x))k(x)w(x). \end{aligned}$$

Hence, the system (3) can be reformulated as

$$\dot{x} = f(x) + g(x)u + \big(I_n - g(x)g^+(x)\big)\, k(x)w(x),$$

where $u = \bar{u} + g^+(x)k(x)w(x)$, and $w$ is regarded as an auxiliary control input. Let $\mathcal{M} : \mathcal{U} \times \mathcal{W} \to \mathbb{R}^{m+q}$ and $\mu := [u^\top, w^\top]^\top \in \mathcal{M}$, and then an auxiliary system is written as

$$\dot{x} = f(x) + M(x)\mu, \tag{4}$$

where $M(x) = \big[g(x)\ \ (I_n - g(x)g^+(x))k(x)\big]$.

We define the value function of the system (4) as

$$V(x,\mu) = \int_t^\infty \big(q(x) + U(\mu) + P(x) + W(x)\big)\, d\tau, \tag{5}$$

where $q(x) = x^\top Qx$ with $Q \geq 0$, and

$$U(\mu) = \int_0^u 2\lambda \tanh^{-1}(\frac{s}{\lambda})^\top R\, \mathrm{d}s + \rho^2\|w\|^2 \tag{6}$$

with $R \geq 0$ and $\rho > 0$. $U(\mu)$ is a generalized non-quadratic function used to address control constraints [9]. In addition, $P(x) = \|\overline{\psi}(x)\|^2$ and $W(x) = \rho^2\|\bar{w}(x)\|^2$. Hereinafter, an optimal control problem is established.

*Problem 2:* Design a feedback control policy $\mu(x) \in \mathcal{M}$ to minimize $V(x,\mu)$ for the system (4).

We next verify that the solution to Problem 2 includes a robust control policy for Problem 1. Assume that $\mu^* \in \mathcal{M}$ is the solution to Problem 2, and its corresponding optimal value function is defined as

$$V^*(x) = \min_{\mu\in\mathcal{M}} V(x,\mu) = V(x,\mu^*). \tag{7}$$

Let $\Gamma(x,\mu) := q(x) + U(\mu) + P(x) + W(x)$. According to the Bellman optimal principle, $V^*$ satisfies the Hamilton-Jacobi-Bellman (HJB) equation

$$\min_{\mu\in\mathcal{M}} H(x,V_x,\mu) = H(x,V_x^*,\mu^*) = 0, \tag{8}$$

where $V_x = \frac{\partial V(x)}{\partial x}$, $V_x^* = \frac{\partial V^*(x)}{\partial x}$ and

$$H(x,V_x^*,\mu^*) = \Gamma(x,\mu^*) + V_x^{*\top}(f(x) + M(x)\mu^*).$$

Utilizing the stationarity conditions $\frac{\partial H}{\partial u} = 0$ and $\frac{\partial H}{\partial w} = 0$ to (8), we have

$$\begin{aligned} u^*(x) &= -\lambda \tanh\left(\frac{1}{2\lambda}R^{-1}g(x)^\top V_x^*\right), \\ w^*(x) &= -\frac{1}{2\rho^2}k(x)^\top(I_n - g(x)g^+(x))^\top V_x^*. \end{aligned} \tag{9}$$

The following theorem indicates that $u^*$ ensures the system (4) to be AS.

*Theorem 1:* Consider the systems (3) and (4). Let $V^*(x)$ be an optimal value function of the system (4) and $\varphi^*(x) = \frac{1}{2\lambda}R^{-1}g(x)^\top V_x^*$. If $u^*$ and $w^*$ in (9) satisfy that

$$2\rho^2\|w^*(x)\|^2 + \lambda^2\|\varphi^*(x)\|^2 \leq q(x), \tag{10}$$

then $u^*$ guarantees the system (3) to be AS.

*Proof:* Consider a Lyapunov function candidate

$$\mathcal{V}(x) := V^*(x).$$

We have that $\mathcal{V}(x) \geq 0$ for all $x \in \mathcal{X}$ and $\mathcal{V}(x) = 0$ if and only if $x = 0$. Let $\mathcal{V}_x = \frac{\partial\mathcal{V}(x)}{\partial x}$. Taking the time derivative of $\mathcal{V}(x)$ along the system (3) under $u^*$, we obtain

$$\begin{aligned} \dot{\mathcal{V}}(x) &= \mathcal{V}_x^\top \dot{x} \\ &= \mathcal{V}_x^\top(f(x) + M(x)\mu^*(x)) + \mathcal{V}_x^\top g(x)\psi(x) \\ &\quad + \mathcal{V}_x^\top(I_n - g(x)g^+(x))k(x)\big(w(x) - w^*(x)\big). \end{aligned}$$

Combining (8)–(9), it can be derived that

$$\begin{aligned} \dot{\mathcal{V}}(x) &= -\Gamma(x,\mu^*) + 2\lambda\varphi^*(x)^\top\psi(x) \\ &\quad - 2\rho^2 w^*(x)^\top\big(w(x) - w^*(x)\big), \quad \forall x \in \mathcal{X}. \end{aligned}$$

Let

$$\begin{aligned} \Delta_1 &= -q(x) - U(\mu^*), \\ \Delta_2 &= -P(x) + 2\lambda\varphi^*(x)^\top\psi(x), \\ \Delta_3 &= -W(x) - 2\rho^2 w^*(x)^\top\big(w(x) - w^*(x)\big). \end{aligned}$$

One has

$$\begin{aligned} \Delta_2 &= -P(x) + 2\lambda\varphi^*(x)^\top\psi(x) \\ &= -\|\lambda\varphi^*(x) - \psi(x)\|^2 + \lambda^2\|\varphi^*(x)\|^2 \leq \lambda^2\|\varphi^*(x)\|^2. \end{aligned}$$

It follows from the Cauchy-Schwarz inequality that

$$-2w^*(x)^\top w(x) \leq \|w^*(x)\|^2 + \|w(x)\|^2,$$

which yields that

$$\begin{aligned} \Delta_3 &\leq -\rho^2\left(\|\bar{w}(x)\|^2 - \|w(x)\|^2\right) + 3\rho^2\|w^*(x)\|^2 \\ &\leq 3\rho^2\|w^*(x)\|^2. \end{aligned}$$

Summarizing the above analysis, one has

$$\begin{aligned} \dot{\mathcal{V}}(x) &= \Delta_1 + \Delta_2 + \Delta_3 \\ &\leq -q(x) + 2\rho^2\|w^*(x)\|^2 + \lambda^2\|\varphi^*(x)\|^2 \leq 0, \end{aligned}$$

where $\dot{\mathcal{V}}(x) = 0$ if and only if $x = 0$. Hence, we conclude that $u^*$ ensures the system (3) to be AS. ■

Based on Theorem 1, we can infer that Problem 1 can be transformed into Problem 2. In other words, we can design an optimal controller to achieve the reset control objective of the DRAS in (1).

## III. Robust Adaptive Dynamic Programming for Optimal Control Design

In this section, a robust adaptive dynamic programming approach is developed to solve the optimal control problem with the control constraint and unmatched bounded perturbation. For this purpose, we provide a policy iteration procedure to establish an optimization framework to approximate the numerical solution of the HJB equation.

**Algorithm 1:** Policy Iteration Algorithm

**Input:** An initial admissible control input $\mu_0(x)$ and a threshold $\varepsilon > 0$.
**Output:** The approximated optimal value function $V^*(x)$ and the approximated optimal control input $u^*(x)$.
Set $i = 0$.
**while** $i \geq 0$ **do**
  Solve the value function $V_i(x)$ by
$$\Gamma(x, \mu_i) + V_{xi}^\top \big(f(x) + M(x)\mu_i\big) = 0$$
  Update the control policy $\mu_{i+1}(x)$ via
$$u_{i+1}(x) = -\lambda \tanh\left(\frac{1}{2\lambda} R^{-1} g(x)^\top V_{xi}\right),$$
$$w_{i+1}(x) = -\frac{1}{2\rho^2} k(x)^\top (I_n - g(x)g^+(x))^\top V_{xi}.$$
  **if** $\|\mu_{i+1}(x) - \mu_i(x)\| \leq \varepsilon$ **then**
    **break;**
  **else**
    Set $i = i + 1$.
  **end**
**end**
**Return:** $V^*(x) = V_{i+1}(x)$, $u^*(x) = u_{i+1}(x)$.

### A. Policy Iteration Procedure

Due to the nonlinear property of the HJB equation in (8), it is generally difficult to obtain its closed-form solution. To approximate its numerical solution, we need to introduce a definition of the admissible control input.

*Definition 1 ( [17]):* If a controller $\mu(x) \in \mathcal{M}$ is continuous on $\mathcal{X}$, makes $V(x, \mu)$ in (5) finite for any $x \in \mathcal{X}$, and ensures the system (4) to be AS with $\mu(0) = 0$, then $\mu(x)$ is called an *admissible control input*.

Next, a modified policy iteration (PI) procedure is formulated as Algorithm 1. Note that an admissible control input $\mu_0$ is necessary for the initialization of the PI procedure. To obtain a proper $\mu_0$, many techniques can be applied, such as pole-placement, feedback linearization and linear quadratic regulator (LQR).

### B. Robust Adaptive Dynamic Programming

We next develop a robust ADP approach based on Algorithm 1 to execute its corresponding PI procedure. For this purpose, we consider the following system:

$$\dot{x}(t) = f(x(t)) + M(x(t))(\mu_0(x(t)) + \zeta(t)), \tag{11}$$

where $\mu_0 : \mathbb{R}^n \to \mathbb{R}^{m+q}$ is an admissible control input for the system (4). $\zeta : \mathbb{R}^+ \to \mathbb{R}^{m+q}$ is a suitable exploration noise to further ensure the persistence excitation (PE) [18], which is a commonly-used condition to guarantee that iteration parameters can be continuously updated and then the functions can converge to their optimal values eventually.

Let $v = \mu_0 - \mu + \zeta$. The system (11) can be written as

$$\dot{x} = f(x) + M(x)\mu + M(x)v. \tag{12}$$

For each integer $i \geq 0$, taking the time derivative of $V_i(x)$ along the trajectory of the system (12) yields that

$$\dot{V}_i(x) = -\Gamma(x, \mu_i) + V_{xi}^\top M(x) v_i. \tag{13}$$

By integrating (13) over an arbitrary time interval $[r, s]$ with $0 \leq r < s$, we have

$$\begin{aligned} &V_i(x(s)) - V_i(x(r)) \\ &= -\int_r^s \Gamma(x, \mu_i)\,\mathrm{d}t + \int_r^s V_{xi}^\top M(x) v_i\,\mathrm{d}t. \end{aligned} \tag{14}$$

According to the universal approximation property provided in [19], smooth functions on compact sets can be approximated by infinite series of basis functions. Consequently, for each integer $i \geq 0$, we can approximate $V_i(x)$, $V_{xi}(x)$ and $\mu_i(x)$ by

$$\begin{aligned} \hat{V}_i(x) &= C_{v|i}^\top \phi_v(x), \\ \hat{V}_{xi}(x)^\top &= C_{p|i}^\top \phi_p(x), \\ \hat{\mu}_i(x)^\top &= C_{u|i}^\top \phi_u(x). \end{aligned} \tag{15}$$

$\phi_v \in \mathbb{R}^{N_v}$, $\phi_p \in \mathbb{R}^{N_p \times n}$ and $\phi_u \in \mathbb{R}^{N_u \times (m+q)}$ in (15) are the linearly independent basis functions defined on $\mathcal{X}$ with $\phi_v(0) = \mathbf{0}_{N_v \times 1}$, $\phi_p(0) = \mathbf{0}_{N_p \times n}$ and $\phi_u(0) = \mathbf{0}_{N_u \times (m+q)}$, where $N_v > 0$, $N_p > 0$ and $N_u > 0$ are sufficiently large integers. $C_{v|i} \in \mathbb{R}^{N_v}$, $C_{p|i} \in \mathbb{R}^{N_p}$ and $C_{u|i} \in \mathbb{R}^{N_u}$ are three weight vectors. Substituting (15) into (14) yields that

$$\begin{aligned} &C_{v|i}^\top \left(\phi_v(x(s)) - \phi_v(x(r))\right) - C_{p|i}^\top \int_r^s \phi_p(x) M(x) \hat{v}_i\,\mathrm{d}t \\ &= -\int_r^s \big(q(x) + U(\hat{\mu}_i) + P(x) + W(x)\big)\mathrm{d}t, \end{aligned}$$

where $\hat{\mu}_0 = \mu_0$ and $\hat{v}_i = \mu_0 - \hat{\mu}_i + \zeta$. Let $z = 1, 2, ..., Z$. Denote $\mathcal{T}(x, \mu_0, \zeta, [r_z, s_z])$ as $Z$ feasible trajectories of the system (11) over a time interval $[r_z, s_z]$ with $0 \leq r_z < s_z$. Hence, we can define $Z$ error functions for each $i \geq 0$ as

$$\begin{aligned} &C_{v|i}^\top \big(\phi_v(x(s_z)) - \phi_v(x(r_z))\big) - C_{p|i}^\top \int_{r_z}^{s_z} \phi_p(x) M(x) \hat{v}_i\,\mathrm{d}t \\ &+ \int_{r_z}^{s_z} \left(q(x) + U(\hat{\mu}_i) + P(x) + W(x)\right)\mathrm{d}t =: e_{iz}, \end{aligned}$$

and hence $C_{v|i}$ and $C_{p|i}$ can be solved by minimizing the term $\sum_{z=1}^{Z} \|e_{iz}\|^2$. Combined with (9) and (15), we have

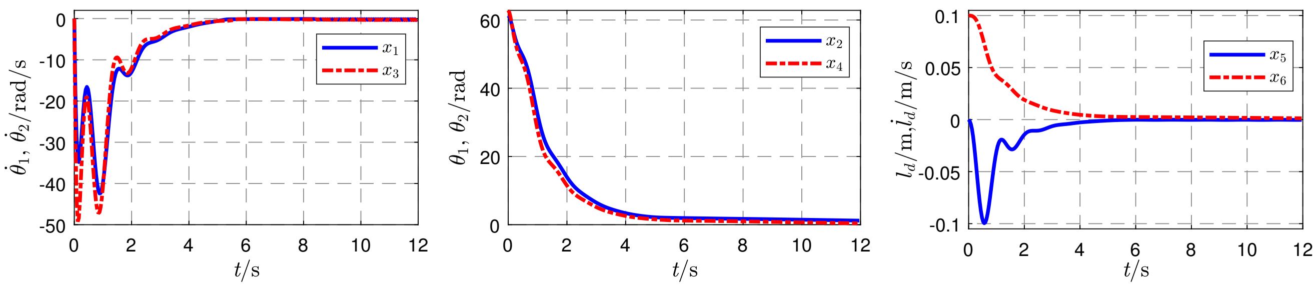
$$\begin{aligned} \hat{u}_i &= -\lambda \tanh\left(\frac{1}{2\lambda} R^{-1} g(x)^\top \phi_p(x)^\top C_{p|i}\right), \\ \hat{w}_i &= -\frac{1}{2\rho^2} k(x)^\top (I_n - g(x)g^+(x))^\top \phi_p(x)^\top C_{p|i}, \end{aligned} \tag{16}$$

$$C_{u|i}^\top \phi_u(x) = [\hat{u}_i^\top, \hat{w}_i^\top]. \tag{17}$$

Furthermore, there exist integers $i^* > 0$, $N_v^* > 0$, $N_p^* > 0$ and $N_u^* > 0$ for an arbitrary threshold $\epsilon > 0$, if $N_v > N_v^*$, $N_p > N_p^*$, and $N_u > N_u^*$, then

$$\begin{aligned} \|C_{v|i^*}^\top \phi_v(x) - V^*(x)\| &\leq \epsilon, \\ \|C_{p|i^*}^\top \phi_p(x) - V_x^*(x)^\top\| &\leq \epsilon, \\ \|C_{u|i^*}^\top \phi_u(x) - \mu^*(x)^\top\| &\leq \epsilon \end{aligned}$$

are satisfied for all $x \in \mathcal{X}$, and the convergence derives the approximate optimal functions $\hat{V}(x)$, $\hat{V}_x(x)$ and $\hat{\mu}(x)$. Note that the convergence verification is similar to Theorem 2 in




(a) (b) (c)

Fig. 2: The evolution of the dynamics of the DRAS. (a) the angular velocity of two EMAs; (b) the angular rotation of two EMAs; (c) the linear velocity and displacement of the load.

[17], and is omitted here due to space limitation. Next, we establish an optimization framework as follows:

$$\min \|A_i^\top C_i - b_i\|^2, \tag{18}$$

where for $i = 0, 1, ...,$

$$A_i = \begin{bmatrix} \phi_v(x(s_1)) - \phi_v(x(r_1)) \\ \vdots \\ \phi_v(x(s_Z)) - \phi_v(x(r_Z)) \\ -\int_{r_1}^{s_1} \phi_p(x) M(x)(\mu_0 + \zeta - \hat\mu_i)\,\mathrm{d}t \\ \vdots \\ -\int_{r_Z}^{s_Z} \phi_p(x) M(x)(\mu_0 + \zeta - \hat\mu_i)\,\mathrm{d}t \end{bmatrix},$$

$$C_i = \left[C_{v|i},\, C_{p|i}\right], \quad b_i = \begin{bmatrix} -\int_{r_1}^{s_1} \Gamma(x, \hat\mu_i)\,\mathrm{d}t \\ \vdots \\ -\int_{r_Z}^{s_Z} \Gamma(x, \hat\mu_i)\,\mathrm{d}t \end{bmatrix},$$

$\hat\mu_i^\top = C_{u|i}^\top \phi_u(x)$ and $\hat\mu_0 = \mu_0$. Hence, $C_{u|i}$ is derived from (17) via $Z$ collected data. Base on the above analysis, we summarize the modified robust ADP approach in Algorithm 2. From Algorithm 2, we can obtain the approximate optimal control policy $\hat\mu^*$. Based on Theorem 1, the component $\hat u^*$ in $\hat\mu^*$ can ensure the system (3) to be AS.

## IV. Simulation Verification

In this section, we show a simulation verification for the proposed robust ADP approach applied to the reset control of the aforementioned DRAS model.

The parameters of the DRAS (1) are provided in Table. I, and the weight matrices in the value function are designed as follows: $Q = I_6$, $R = I_5$, and $\rho = 1$. Let $\lambda = 60$ and $\varepsilon = 0.0001$. Because $w$ is an auxiliary input, the exploration noise is set as

$$\zeta(t) = \begin{bmatrix} \sin(50t) + 0.5e^{-0.04t}\sin(2t) + 0.1\sin(10t) \\ 0.1\sin(50t) + 0.7e^{-0.04t}\sin(2t) + 0.1\sin(10t) \end{bmatrix},$$

and the initial admissible control input is chosen as $u_0(x) = -Kx$, where

$$K = \begin{bmatrix} 1.1491,\, 3.6542,\, 0.1098,\, 2.7729,\, 0.3198,\, -3.0940 \\ 0.1098,\, 2.7729,\, 1.1491,\, 3.6542,\, 0.3198,\, -3.0940 \end{bmatrix}.$$

**Algorithm 2:** Robust ADP Algorithm

**Input:** An initial admissible control input $\mu_0(x)$; a proper exploration noise $\zeta(t)$; and a threshold $\epsilon > 0$.
**Output:** The approximate optimal value function $\hat V^*(x)$ and the approximate optimal control input $\hat\mu^*(x)$.
Apply $\mu(x) = \mu_0(x) + \zeta(t)$ as the control input to the system (4) during a sufficiently long time interval to collect necessary data.
Set $i = 0$.
**while** $i \geq 0$ **do**
  Generate $A_i$ and $b_i$ in (18).
  Obtain $C_i$ by solving the optimization problem (18).
  Calculate $\hat u_i$ and $\hat w_i$ by (16) using the collected data.
  Obtain $C_{u|i}$ by solving (17).
  **if** $i \geq 1$ *and* $\|C_i - C_{i-1}\|^2 \leq \epsilon$ **then**

$$\hat V^*(x) = C_{v|i}^\top \phi_v(x),$$
$$\hat\mu^*(x) = \phi_u(x)^\top C_{u|i}$$

  **else**
    Set $i = i + 1$.
  **end**
**end**

TABLE I: Parameters of the DRAS

| Symbol | Value | Unit | Symbol | Value | Unit |
|---|---|---|---|---|---|
| $J$ | $4.55 \times 10^{-2}$ | $\mathrm{kg \cdot m^2}$ | $\gamma_1$ | 0.5 | – |
| $M$ | 3710 | kg | $\gamma_2$ | 10 | – |
| $l$ | 0.01 | m | $\gamma_3$ | 1 | – |
| $k$ | $4.5 \times 10^4$ | N/m | $\gamma_4$ | 0.5 | – |
| $\Sigma_a$ | $4 \times 10^4$ | $\mathrm{kg \cdot m^2/s^2}$ | $\gamma_5$ | 10 | – |
| $\eta$ | 0.9 | – | $\gamma_6$ | 0.1 | – |

Note that $K$ is obtained by the LQR approach with $Q = I_6$ and $R = I_5$, and hence $u(x) = u_0(x) + \zeta(t)$. Because the matrix $M(x)$ is not full-rank ($rank = 3$), a dimensionality reduction is needed. We reset the control input as $\mu(x) = [u(x)^\top, w_3(x)]^\top$, and basis functions $\phi_v(x)$, $\phi_p(x)$ and $\phi_u(x)$ are chosen as

$$\phi_v^\top(x) = [x_1^2, x_1x_2, x_1x_3, x_1x_4, x_1x_5, x_1x_6, x_2^2, x_2x_3, x_2x_4, x_2x_5, x_2x_6, x_3^2, x_3x_4, x_3x_5, x_3x_6, x_4^2, x_4x_5, x_4x_6, x_5^2, x_5x_6, x_6^2].$$

$$\phi_p^\top(x) = \phi_u^\top(x) = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & \mathbf{0}_{1\times6} & \mathbf{0}_{1\times6} \\ \mathbf{0}_{1\times6} & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & \mathbf{0}_{1\times6} \\ \mathbf{0}_{1\times6} & \mathbf{0}_{1\times6} & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{bmatrix},$$

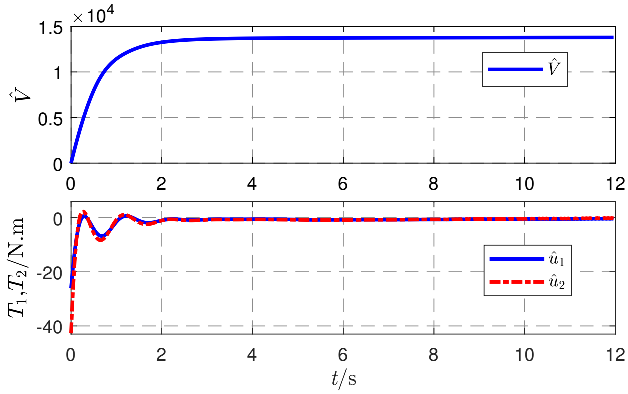


Fig. 3: The evolution of the approximated value function and control inputs.

which is ensured that $\phi_v$ and $\phi_p$ are full column rank and then the optimization problem (18) is feasible.

System trajectories are collected off-line with the total time $t_s = 8\,\mathrm{s}$, the sampled time $\Delta t = 0.04\,\mathrm{s}$, and the initial state $x_0 = [0,\ 19.9\pi,\ 0,\ 20\pi,\ 0,\ 0.1]^\top$. According to Algorithm 2, three weight vectors are calculated as $C_v^\top = [C_{v1}, C_{v2}, C_{v3}, C_{v4}]$, $C_p^\top = [C_{p1}, C_{p2}, C_{p3}, C_{p4}]$ and $C_u^\top = [C_{u1}, C_{u2}, C_{u3}, C_{u4}]$, where

$$
\begin{aligned}
C_{v1} &= \left[0.0724\ 0.3960\ 0.0048\ 0.4190\ 0.0552\ -0.3793\right],\\
C_{v2} &= \left[4.8869\ 0.5502\ 7.0477\ 1.4294\ -7.6902\ 0.0631\right],\\
C_{v3} &= \left[-0.1466\ 0.0118\ -0.1942\ 2.6494\ 0.6357\right],\\
C_{v4} &= \left[-6.0503\ 0.2902\ -0.7490\ 4.9597\right],
\end{aligned}
$$

$$
\begin{aligned}
C_{p1} &= \left[-0.0790\ -0.2934\ -0.0268\ -0.2506\ -0.0266\right],\\
C_{p2} &= \left[0.2621\ 0.0537\ -0.1224\ -0.1098\ -0.1482\right],\\
C_{p3} &= \left[-0.0142\ 0.1323\ 5.7972\ 10.0510\ -1.6816\right],\\
C_{p4} &= \left[10.8945\ 1.7375\ -9.1733\right],
\end{aligned}
$$

$$
\begin{aligned}
C_{u1} &= \left[-0.4106\ -1.1366\ 0.0383\ -0.7500\ -0.0956\right],\\
C_{u2} &= \left[0.9073\ -0.2858\ -1.2658\ -0.2871\ -1.7205\right],\\
C_{u3} &= \left[-0.1458\ 1.4292\ 0.2885\ 0.5887\ -0.0805\right],\\
C_{u4} &= \left[0.6303\ 0.1366\ -0.5199\right],
\end{aligned}
$$

with the number of iterations $i = 52$, and thus

$$\hat{V}(x) = C_v^\top \phi_v(x), \quad \hat{\mu}(x)^\top = C_u^\top \phi_u(x).$$

The simulation time is set as $t = 12\,\mathrm{s}$, and the derived results are shown in Fig. 2-Fig. 3. From Fig. 2, it is verified that the proposed approach can achieve the reset control of the DRAS. From Fig. 3, we have $\|\hat{u}_1\|_{max} = 26.02 < \lambda$ and $\|\hat{u}_2\|_{max} = 42.73 < \lambda$, which means that all the control constraints are satisfied, and $\hat{V}(x)$ is guaranteed to be finite and convergent eventually.

## V. CONCLUSIONS

This paper established a nonlinear dynamics of the dual-redundancy actuation systems driven by electro-mechanical actuators, involving both control constraint and unmatched bounded perturbation. To deal with the reset control of this system, we transformed the robust control problem into an optimal control problem, and then a modified robust adaptive dynamic programming approach is applied to approximate a numerical solution of the Hamilton-Jacobi-Bellman equation. Finally, simulation results illustrated the efficacy of the derived controller. Future work will aim at investigating control synthesis of reducing the force fight in redundancy actuation systems via adaptive dynamic programming.

## REFERENCES


[1] F. Golnaraghi and B. C. Kuo, *Automatic Control Systems*. McGraw-Hill Education, 2017.

[2] R. M. Murray, K. J. Astrom, S. P. Boyd, R. W. Brockett, and G. Stein, “Future directions in control in an information-rich world,” *IEEE Control Systems Magazine*, vol. 23, no. 2, pp. 20–33, 2003.

[3] D. Ye and G.-H. Yang, “Adaptive fault-tolerant tracking control against actuator faults with application to flight control,” *IEEE Transactions on Control Systems Technology*, vol. 14, no. 6, pp. 1088–1096, 2006.

[4] J. Zhang, J. Mu, and Y. Gao, “Modeling and simulation analysis of redundant actuator system of aircraft’s control surface,” *Computer Simulation*, vol. 33, no. 7, pp. 153–157, 2016.

[5] H. Qi and Y. Teng, “Force equalization control for dual-redundancy electro-hydrostatic actuator,” *Journal of Beijing University of Aeronautics and Astronautics*, vol. 43, no. 2, pp. 270–276, 2017.

[6] I. Salman, Y. Lin, and M. T. Hamayun, “Fractional order modeling and control of dissimilar redundant actuating system used in large passenger aircraft,” *Chinese Journal of Aeronautics*, vol. 31, no. 5, pp. 1141–1152, 2018.

[7] F. Lin, “An optimal control approach to robust control design,” *International Journal of Control*, vol. 73, no. 3, pp. 177–186, 2000.

[8] J. Zhao, J. Na, and G. Gao, “Adaptive dynamic programming based robust control of nonlinear systems with unmatched uncertainties,” *Neurocomputing*, vol. 395, pp. 56–65, 2020.

[9] X. Yang, H. He, and X. Zhong, “Adaptive dynamic programming for robust regulation and its application to power systems,” *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5722–5732, 2017.

[10] D. M. Adhyaru, I. Kar, and M. Gopal, “Bounded robust control of nonlinear systems using neural network–based hjb solution,” *Neural Computing and Applications*, vol. 20, pp. 91–103, 2011.

[11] P. J. Werbos, “Approximate dynamic programming for real-time control and neural modeling,” *Handbook of Intelligent Control Neural Fuzzy and Adaptive Approaches*, 1992.

[12] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*. John Wiley & Sons, 2012.

[13] R. Kamalapurkar, H. Dinh, S. Bhasin, and W. E. Dixon, “Approximate optimal trajectory tracking for continuous-time nonlinear systems,” *Automatica*, vol. 51, pp. 40–48, 2015.

[14] S. Bhasin, R. Kamalapurkar, M. Johnson, K. G. Vamvoudakis, F. L. Lewis, and W. E. Dixon, “A novel actor-critic-identifier architecture for approximate optimal control of uncertain nonlinear systems,” *Automatica*, vol. 49, no. 1, pp. 82–92, 2013.

[15] X. Xing and D. E. Chang, “The adaptive dynamic programming toolbox,” *Sensors*, vol. 21, no. 16, p. 5609, 2021.

[16] C. Makkar, G. Hu, W. G. Sawyer, and W. E. Dixon, “Lyapunov-based tracking control in the presence of uncertain nonlinear parameterizable friction,” *IEEE Transactions on Automatic Control*, vol. 52, no. 10, pp. 1988–1994, 2007.

[17] X. Han, X. Zhao, H. R. Karimi, D. Wang, and G. Zong, “Adaptive optimal control for unknown constrained nonlinear systems with a novel quasi-model network,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 7, pp. 2867–2878, 2021.

[18] J. Huang, Z. Zhang, F. Cai, and Y. Chen, “Optimized formation control for multi-agent systems based on adaptive dynamic programming without persistence of excitation,” *IEEE Control Systems Letters*, vol. 6, pp. 1412–1417, 2022.

[19] K. Hornik, M. Stinchcombe, and H. White, “Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks,” *Neural Networks*, vol. 3, no. 5, pp. 551–560, 1990.