Optimized Data-Driven Prescribed Performance Attitude Control for Actuator Saturated Spacecraft

Haoyang Yang, *Graduate Student Member, IEEE,* Qinglei Hu, *Senior Member, IEEE,* Hongyang Dong, Xiaowei Zhao, Dongyu Li, *Member, IEEE*

***Abstract*—This article addresses the crucial requirements in spacecraft attitude control: prescribed performance guarantees under actuator saturation and real-time cost optimization. As an application-oriented study, an approximate optimal prescribed performance attitude control scheme is proposed for this objective. To be specific, the prescribed performance constraint is converted into the system dynamics and merged into the adaptive dynamic programming design philosophy. Subsequently, the online learning law is designed based on a special saturated HJB error, in which a dynamical scale is introduced to adjust the learning gain by measured data. It enhances learning efficiency and applicability. Then, uniformly ultimately bounded stability of the whole system is achieved with guaranteed convergence of optimization by the Lyapunov-based stability analysis. Finally, both numerical simulation and hardware-in-the-loop experiments demonstrate the superiority and effectiveness of the proposed method. These attributes and outcomes attained will promote the development of practical space missions.**

***Index Terms*—Prescribed Performance Control, Adaptive Dynamic Programming (ADP), Attitude Control, Actuator Saturation, Approximated Optimal Control**

I. Introduction

RIGID bodies attitude maneuver has elicited ever-increasing attention among academics scholars and industry engineers, owing to its great significance in aerospace missions, robotic systems, and underwater exportation. Various elegant efforts, such as proportional-derivative control [1], [2], sliding mode control [3], [4], H_∞ control [5], backstepping control [6], etc., have been devoted to this field. However, designing an attitude control scheme for the aforementioned mechanical systems that can achieve desired performance requirements is still a challenging task, due to the following three aspects: Firstly, the rotation kinematics and dynamics are characterized by nonlinearity and coupling. Secondly, the closed-loop transient-state performance (e.g., convergence rate and overshoot) is hard to satisfy under actuator saturation. Thirdly, the on-orbit control cost is difficult to optimize in real-time.

Traditionally, the previous methods for designing closed-loop transient-state performance rely on the parameter tuning of the devised controller, which implies that the performance assessment is a posteriori. Owing to the prescribed performance control (PPC) technique proposed in [7], which has emerged as an effective way to ensure the transient-state performance of closed-loop systems. The design philosophy of PPC is to construct a time decay function to prescribe the system's performance, and then transform the motion state into the feedback error with respect to performance [8]. Zhou *et al.* [9] design a robust prescribed performance attitude controller for rigid spacecraft in the presence of uncertainties and disturbances. A finite-time PPC method is devised in [10], in which Chebyshev neural network is introduced to approximate the lumped uncertainties. A serial of PPC methods [11]–[13] are proposed for rigid spacecraft attitude control in the case of actuator fault and input saturation. Note that these elegant results mainly focus on the stability of system and specific requirements for transient performance. Under the aforementioned design framework, it is inaccessible to consider control cost optimization, while this capability is critical for the efficiency and economy of spacecraft systems.

To minimize the on-orbit control cost, the explicit form of the optimized controller can be derived in case of the special simple case [14]. However, taking account of some complex situations, the analytical form of control torque cannot be worked out anymore. Facing these circumstances, a good deal of outstanding results are presented via solving optimization problem. To list a few, Lee *et al.* [15] propose a geometrical optimization method for attitude control considering pointing and reaction wheel constraints. In [16], a switch geometrical method control scheme is derived for energy-optimal reorientation under input saturation. Addressing the power limitation of CubeSat, a constrained optimal attitude control problem can be solved aiding by the *IPOPT* solver [17]. It can be noted that the above-considered constraints are static, and numerical optimization algorithms are still awkward to cope with them while satisfying real-time computation, not to mention the guarantee of prescribed performance (a dynamical system behavioral bound). As a consequence, there is few

This work was funded in part by the Natural Science Foundation of China under grants 61960206011 and 62103028, in part by the Zhejiang Provincial Natural Science Foundation under Grant LD22E050004, in part by the Industry-University-Research Foundation of China under Grant 2021ZYA02022. Qinglei Hu appreciates the support of Hangzhou Qianjiang Distinguished Expert program and Haoyang Yang appreciates support of China Scholarships Council program.

Haoyang Yang is with the School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China. Email: yanghaoyang8352@buaa.edu.cn.

Qinglei Hu (corresponding author) is with the School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China, and also with the Beihang Hangzhou Innovation Institute Yuhang, Hangzhou 310023, China. Email: huql_buaa@buaa.edu.cn.

Hongyang Dong and Xiaowei Zhao are with the School of Engineering, University of Warwick, Coventry, CV4 7AL, UK. Email: {hongyang.dong, xiaowei.zhao}@warwick.ac.uk.

Dongyu Li is with the School of Cyber Science and Technology, Beihang University, Beijing 100191, China, and also with Shanghai Institute of Satellite Engineering, Shanghai 201109, China. Email: dongyuli@buaa.edu.cn.

results address the optimal prescribed performance control via the traditional optimization method.

Therefore, the incompatibility between real-time cost optimization and the guarantee of prescribed performance has become a major impediment to the development of spacecraft attitude control systems. Fortunately, the adaptive/approximate dynamic programming (ADP) technique provides an effective way to tackle the optimal control problem. The basic principle of ADP is to enhance the control scheme by suitably valuing state feedback from the environment [18], in which the Hamilton–Jacobi–Bellman (HJB) equations are approximately solved in real-time to achieve near-optimal control. This powerful technique has aroused significant research interest recently in robotic systems [19], [20]. Therefore, a question spontaneously arises here: Can we develop an optimal prescribed performance control scheme for spacecraft reorientation based on this technique? In this regard, a constrained approximate optimal attitude controller is designed for spacecraft reorientation [21], in which both pointing and angular velocity constraints are handled by barrier function type reward shaping. This approach is unsuitable for the processing of dynamic performance constraints. Besides, the magnitude of torque can only be restrained by raising the weight. In [22] and [23], the ADP technique is introduced to achieve prescribed performance attitude tracking and formation flying under actuator saturation, respectively. It is noted that both of them use the ADP method as a supplementary of the traditional control scheme. Strictly speaking, such a manner is conservative and partial in optimization, unable to optimize the overall cost. Therefore, it needs to further investigate how to optimize the overall control cost while satisfying real-time computation in the design of prescribed performance controller under actuator saturation.

Motivated by the above discussion, this article seeks to remove technical barriers of real-time overall cost optimization and transient-state performance guarantee under actuator saturation simultaneously in spacecraft attitude control. Aiming at this objective, an approximate optimal prescribed performance attitude control scheme is proposed for actuator saturated spacecraft. By utilizing the mechanism of ADP, the controller achieves online improvement driven by the measured motion state. Then the bounded stability and optimality of the whole attitude system are guaranteed via Lyapunov-based method analysis. Per practical application concerns, a typical hardware-in-the-loop (HITL) experimental validation is presented on a satellite attitude control semi-physical test system to evaluate the effectiveness and practicality. To be specific, the major contributions of this paper can be summarized as follows:

1) Dispose of the incompatibility between real-time cost optimization and the guarantee of prescribed performance in the spacecraft control system. By merging the prescribed performance function design into the ADP design philosophies, the proposed method achieves that the prescribed performance controller has the capability of real-time optimizing the overall control cost.
2) The current and previous motion data are compositely utilized in the learning law design, which enables a simple controller to be improved to the desired characteristics controller. In this design, the online learning scheme is designed based on the special saturated HJB error, in which a variable scale is introduced to dynamically adjust the learning gain and release the persistent excitation (PE) condition dependence, warranting the convergence and efficiency of learning.

In the rest of this article, we introduce the mathematical preliminaries and formulate the attitude control problem in Sec. II. Subsequently, the data-driven prescribed performance control scheme is developed in Sec. III with related discussion and analysis. After that, the typical numerical examples and experimental demonstrations are illustrated in Sec. IV. Finally, this article is concluded in Sec. V.

II. Preliminaries and Problem Formulation

Notation: Throughout this paper, $\mathbb{R}^{n\times m}$ presents the set of $n \times m$ real matrix. For a vector $\boldsymbol{a} \triangleq [a_1, a_2, a_3]^{\mathrm{T}} \in \mathbb{R}^3$, the tangent function of it is defined as $\tanh(\boldsymbol{a}) \triangleq [\tanh(a_1), \tanh(a_2), \tanh(a_3)]^{\mathrm{T}}$, and $\boldsymbol{a}^{\times} \triangleq [0, -a_3, a_2; a_3, 0, -a_1; -a_2, a_1, 0] \in \mathbb{R}^{3\times 3}$ denotes the cross product matrix. Let $\|\cdot\|$ to be the Euclidean norm of a vector and its induced norm of matrix. Given any square matrix $\boldsymbol{X}$, $\lambda_{\min}(\boldsymbol{X})$ denotes its minimum eigenvalues. Matrix $\boldsymbol{I}_n$ represents the n-dimensional identity matrix, $n\times m$ matrix with all elements are one and zeros are denoted by $\mathbf{1}_{n\times m}$ and $\mathbf{0}_{n\times m}$, respectively.

A. Attitude Dynamics of Rigid Spacecraft

The attitude stabilization problem of a rigid spacecraft is considered in this section. Presume that full-motion information (both attitude and angular velocity of spacecraft) can be directly measured by the onboard sensors.

Then, the dynamics of rigid spacecraft can be represented in term of modified Rodriguez parameters (MRP) [24]:

$$\dot{\boldsymbol{\sigma}} = \boldsymbol{H}(\boldsymbol{\sigma})\boldsymbol{\omega} \tag{1}$$

$$\boldsymbol{J}\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega}^{\times}\boldsymbol{J}\boldsymbol{\omega} + \boldsymbol{\tau}; \tag{2}$$

in which, $\boldsymbol{\sigma} \triangleq [\sigma_1, \sigma_2, \sigma_3]^{\mathrm{T}} = \tan(\Phi/4)\boldsymbol{n} \in \mathbb{R}^3$ denotes the MRP, $\Phi \in \mathbb{R}$ denotes Euler angle, $\boldsymbol{n} \in \mathbb{R}^3$ the Euler axis, $\boldsymbol{\omega} \triangleq [\omega_1, \omega_2, \omega_3]^{\mathrm{T}} \in \mathbb{R}^3$ is the angular velocity, $\boldsymbol{J} \in \mathbb{R}^{3\times 3}$ represents moment of inertia, $\boldsymbol{\tau} \triangleq [\tau_1, \tau_2, \tau_3]^{\mathrm{T}} \in \mathbb{R}^3$ is the output torque from actuators, and $\boldsymbol{H}(\boldsymbol{\sigma})$ is given by

$$\boldsymbol{H}(\boldsymbol{\sigma}) \triangleq \frac{1+\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}}{4}\boldsymbol{I}_3 - \frac{1}{2}\boldsymbol{\sigma}^{\times} + \frac{1}{2}\boldsymbol{\sigma}^{\times}\boldsymbol{\sigma}^{\times} \tag{3}$$

In addition, using the geometric properties of MRP, it can map the MRP vector $\boldsymbol{\sigma}$ whenever $\|\boldsymbol{\sigma}\| > 1$ to its corresponding counterpart $\boldsymbol{\sigma}^s$ through [25]:

$$\boldsymbol{\sigma}^s = -\frac{\boldsymbol{\sigma}}{\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}} \tag{4}$$

Therefore, the above mapping can guarantee that MRP vector $\boldsymbol{\sigma}$ remains bounded within a unit sphere ($\|\boldsymbol{\sigma}\| \leq 1$).

B. Actuator Saturation

In practical mission, the actuators equipped on spacecraft (e.g., reaction wheel (RW), magnetic torquer, control moment gyro (CMG), reaction thruster) always have their physical limitation [26]. This may imply that the control torque command should not large than the maximum output of the actuators k_τ. This characteristic can be illustrated by chart of "command - actuator output" as shown in Fig. 1.

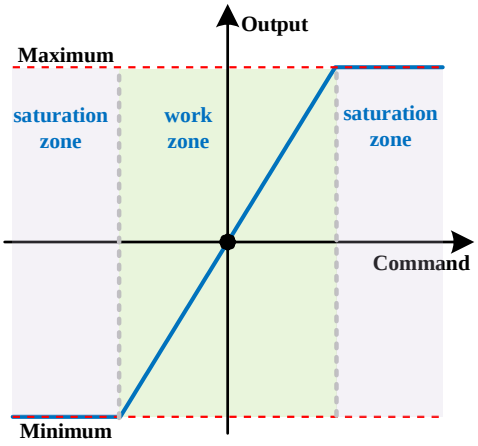

Fig. 1. Saturation characteristic of the actuators

This constraint can be represented as:

$$\|\tau_i\| \leq k_\tau, \ i = 1, 2, 3 \tag{5}$$

Thus, the control torque's constraint can be compactly rewritten by vector inequality.

$$-k_\tau \mathbf{1}_{3\times 1} \leq \boldsymbol{\tau} \leq k_\tau \mathbf{1}_{3\times 1} \tag{6}$$

C. Prescribed Performance Constraint and Augment State Transformation

In accordance with the design principle of prescribed performance control in [11], [12], this paper takes advantage of a typical continuous transformation function $\boldsymbol{\rho} \triangleq [\rho_1, \rho_2, \rho_3]^{\mathrm{T}} \in \mathbb{R}^3$ to restrict the attitude state $\boldsymbol{\sigma}$, which is defined as:

$$\rho_i(t) \triangleq \rho_{Li} + (\rho_{Hi} - \rho_{Li}) \exp(-k_\rho t), \quad i = 1, 2, 3 \tag{7}$$

where k_ρ is constants, ρ_{Hi} and ρ_{Li} are the i^{th} element of $\boldsymbol{\rho}_H$ and $\boldsymbol{\rho}_L$, with $0 < \rho_{Li} < \rho_{Hi}$ and $\|\sigma_i(0)\| < \|\rho_{Hi}\|$. The attitude state σ_i should satisfy that:

$$(s_i(\alpha_\rho - 1) - \alpha_\rho)\rho_i(t) \leq \sigma_i(t) \leq (1 + s_i(\alpha_\rho - 1))\rho_i(t) \tag{8}$$

in which, $\alpha_\rho \in (0, 1]$ is a constant, and s_i, defined in (9), is utilized to represent the sign of the initial attitude $\boldsymbol{\sigma}(0)$.

$$s_i = \begin{cases} 0, & \text{if } \sigma_i(0) \geq 0 \\ 1, & \text{if } \sigma_i(0) < 0 \end{cases}, \quad i = 1, 2, 3 \tag{9}$$

Note that, the inequation formulated in (8) depicts the admissible region for spacecraft's attitude $\boldsymbol{\sigma}$, as shown in Fig.2. According to the prescribed performance function (PPF), the

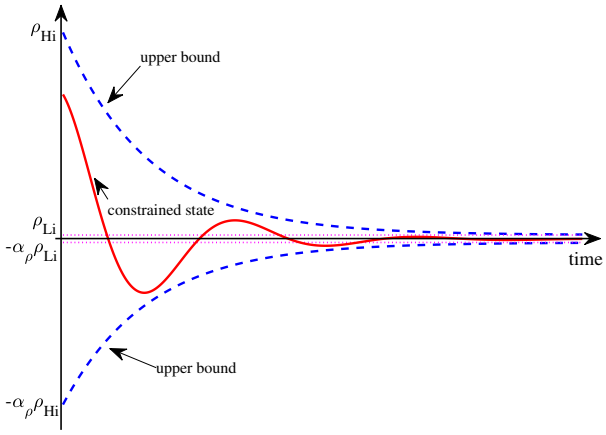

Fig. 2. Illustration of the prescribed performance constraint.

argument state transformation is defined as:

$$e_i \triangleq \log \frac{\alpha_\rho + (s_i(\alpha_\rho - 1) + 1)z_i}{\alpha_\rho + (s_i(\alpha_\rho - 1) - \alpha_\rho)z_i}, \quad i = 1, 2, 3 \tag{10}$$

where $z_i \triangleq \sigma_i(t)/\rho_i(t)$, e_i denotes the i^{th} element of augment state $\boldsymbol{e}$. Then, calculating the time derivative of (10), one has

$$\dot{e}_i = \psi_i(\dot{\sigma}_i - z_i\dot{\rho}_i) \tag{11}$$

with $\psi_i = (\alpha_\rho + \alpha_\rho^2)/[(\alpha_\rho + (s_i(\alpha_\rho - 1) + 1)z_i)(\alpha_\rho + (s_i(\alpha_\rho - 1) - \alpha_\rho)z_i)\rho_i]$. Recalling the attitude dynamics (1). Then, the augment state can be formulated by:

$$\dot{\boldsymbol{e}} = \boldsymbol{\Upsilon}(\boldsymbol{H}(\boldsymbol{\sigma})\boldsymbol{\omega} - \boldsymbol{\Gamma}\boldsymbol{\sigma}) \tag{12}$$

with

$$\boldsymbol{\Upsilon} \triangleq \begin{bmatrix} \psi_1 & & \\ & \psi_2 & \\ & & \psi_3 \end{bmatrix}, \quad \boldsymbol{\Gamma} \triangleq \begin{bmatrix} \frac{\dot{\rho}_1}{\rho_1} & & \\ & \frac{\dot{\rho}_2}{\rho_2} & \\ & & \frac{\dot{\rho}_3}{\rho_3} \end{bmatrix}$$

D. Control Problem Statement

Formally, the main problem addressed in this paper reads as follows:

Problem 1. *Consider the rigid spacecraft attitude dynamics defined by* (1)-(2)*, design a real-time saturated torque control scheme* $\boldsymbol{\tau}$ *to satisfy the convergence of prescribed performance* (8) *and the following optimization of specific control cost index* (13)*.*

$$\mathcal{J} = \int_0^\infty \underbrace{\mathfrak{Q}}_{\textit{motion term}} + \underbrace{\mathfrak{U}}_{\textit{torque term}} \, \mathrm{d}t \tag{13}$$

III. Data-Driven Optimal Prescribed Performance Controller Design

A. Optimal Control Formulation and Cost Function Shaping

To formulate the above-mentioned optimal control problem, we reorganize (1), (2), (7), (12) as the following compact formulation:

$$\dot{\boldsymbol{\chi}} = \boldsymbol{F}(\boldsymbol{\chi}) + \boldsymbol{G}(\boldsymbol{\chi})\boldsymbol{\tau} \tag{14}$$

in which, $\boldsymbol{\chi} \triangleq \left[\boldsymbol{\sigma}^{\mathrm{T}},\ (\boldsymbol{J}\boldsymbol{\omega})^{\mathrm{T}}, \boldsymbol{e}^{\mathrm{T}},\ \boldsymbol{\rho}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{12}$, and

$$\boldsymbol{F}(\boldsymbol{\chi}) \triangleq \begin{bmatrix} \boldsymbol{H}(\boldsymbol{\sigma})\boldsymbol{\omega} \\ -\boldsymbol{\omega}^{\times}\boldsymbol{J}\boldsymbol{\omega} \\ \boldsymbol{\Upsilon}(\boldsymbol{H}(\boldsymbol{\sigma})\boldsymbol{\omega} - \boldsymbol{\Gamma}\boldsymbol{\sigma}) \\ -\kappa_{\rho}\exp\left(-k_{\rho}t\right)(\boldsymbol{\rho}_H - \boldsymbol{\rho}_L) \end{bmatrix},\ \boldsymbol{G}(\boldsymbol{\chi}) \triangleq \begin{bmatrix} \boldsymbol{0}_{3\times 3} \\ \boldsymbol{I}_3 \\ \boldsymbol{0}_{3\times 3} \\ \boldsymbol{0}_{3\times 3} \end{bmatrix}$$

Then, based on the cost index (13), the following cost-to-go function is defined to formulation the cost from the current time instant.

$$\mathcal{V} = \int_t^{\infty} \mathfrak{Q}(\boldsymbol{\chi}) + \mathfrak{U}(\boldsymbol{\tau})\mathrm{d}t \tag{15}$$

where the cost related to state and torque are designed as the following form:

$$\mathfrak{Q}(\boldsymbol{\chi}) = \boldsymbol{e}^{\mathrm{T}}\boldsymbol{Q}_e\boldsymbol{e} + \boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{Q}_{\sigma}\boldsymbol{\sigma} + \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{Q}_{\omega}\boldsymbol{\omega} \tag{16}$$

$$\mathfrak{U}(\boldsymbol{\tau}) = 2k_{\tau}\int_0^{\boldsymbol{\tau}} \tanh^{-\mathrm{T}}(\frac{\boldsymbol{u}}{k_{\tau}})\mathrm{d}\boldsymbol{u} \tag{17}$$

in which, the weight matrices $\boldsymbol{Q}_e$, $\boldsymbol{Q}_{\sigma}$, $\boldsymbol{Q}_{\omega} \in \mathbb{R}^{3\times 3}$ are used to balance the cost. Matrix $\mathfrak{Q}$ is positive definite, $\mathfrak{Q} = 0$ only when $\boldsymbol{\sigma} = \boldsymbol{0}_{3\times 1}$ and $\boldsymbol{\omega} = \boldsymbol{0}_{3\times 1}$ rad/s. If the attitude state $\boldsymbol{\sigma}$ close to the prescribe performance boundary, $\boldsymbol{Q}_{\sigma}$ will approach infinity. Note that, the argument state $\boldsymbol{e}$ relates to both desired attitude and prescribed performance constraint, so the weight of it should be appropriately set. By such a cost shaping, the state constraint can be soft handled in the optimal control framework.

Then, the optimal control torque is defined by $\boldsymbol{\tau}^*$, and $\mathcal{V}^*$ represents the corresponding cost function. After that, the optimal cost $\mathcal{V}^*$ satisfies the following Hamiltonian function:

$$\mathcal{H}(\boldsymbol{\chi}, \boldsymbol{\tau}, \nabla_{\chi}\mathcal{V}) \triangleq \nabla_{\chi}^{\mathrm{T}}\mathcal{V}(\boldsymbol{F}(\boldsymbol{\chi}) + \boldsymbol{G}(\boldsymbol{\chi})\boldsymbol{\tau}) + \mathfrak{Q}(\boldsymbol{\chi}) + \mathfrak{U}(\boldsymbol{\tau}) \tag{18}$$

By taking the partial differential for (18), the explicit expression for optimal control scheme can be obtained as follows:

$$\boldsymbol{\tau}^* = -k_{\tau}\tanh(\frac{\boldsymbol{\mu}^*}{k_{\tau}}) \tag{19}$$

in which, $\boldsymbol{\mu}^* \triangleq [\mu_1^*, \mu_2^*, \mu_3^*] \triangleq \frac{1}{2}\boldsymbol{G}^{\mathrm{T}}(\boldsymbol{\chi})\nabla_{\chi}\mathcal{V}^*$. Substituting (19) into (18), then the HJB equation in presentation of $\nabla_{\chi}\mathcal{V}^*$ can be formulated as:

$$\begin{aligned}\mathcal{H}(\boldsymbol{\chi}, \boldsymbol{\tau}^*, \nabla_{\chi}\mathcal{V}^*) = \nabla_{\chi}^{\mathrm{T}}\mathcal{V}^*\boldsymbol{F}(\boldsymbol{\chi}) - 2k_{\tau}^2(\boldsymbol{\mu}^*)^{\mathrm{T}}\tanh(\frac{\boldsymbol{\mu}^*}{k_{\tau}}) \\ +\mathfrak{Q}(\boldsymbol{\chi}) + 2k_{\tau}\int_0^{-k_{\tau}\tanh(\frac{\boldsymbol{\mu}}{k_{\tau}})} \tanh^{-\mathrm{T}}(\frac{\boldsymbol{u}^*}{k_{\tau}})\mathrm{d}\boldsymbol{u} = 0\end{aligned} \tag{20}$$

Then using the property of inverse hyperbolic function [27] that

$$\begin{aligned}&2k_{\tau}\int_0^{-k_{\tau}\tanh(\frac{\boldsymbol{\mu}^*}{k_{\tau}})} \tanh^{-\mathrm{T}}(\frac{\boldsymbol{u}}{k_{\tau}})\mathrm{d}\boldsymbol{u} \\ &= 2k_{\tau}\sum_{i=1}^{3}\int_0^{-k_{\tau}\tanh(\frac{\mu_i^*}{k_{\tau}})} \tanh^{-\mathrm{T}}(\frac{u_i}{k_{\tau}})\mathrm{d}u_i \\ &= 2k_{\tau}^2\boldsymbol{\mu}^{*\mathrm{T}}\tanh(\boldsymbol{\mu}^*) + k_{\tau}^2\sum_{i=1}^{3}\log(1-\tanh^2(\mu_i^*))\end{aligned} \tag{21}$$

Then, the HJB equation (20) can be simplified as follows:

$$\mathfrak{Q}(\boldsymbol{\chi}) + \nabla_{\chi}^{\mathrm{T}}\mathcal{V}^*\boldsymbol{F}(\boldsymbol{\chi}) + k_{\tau}^2\sum_{i=1}^{3}\log(1-\tanh^2(\mu_i^*)) = 0 \tag{22}$$

It can be seen that although the form of (22) is much more concise than (20), it is still difficult to solve an analytical solution.

B. Data-Driven Approximation Solution

Conventional methods is hard to analytically solve the optimal control torque $\boldsymbol{\tau}^*$. Therefore, the ADP-framework provides an efficient way to handle this challenge by approximation method.

A neural network, which contains sufficient basis functions (neural function), can be used to approximate the optimal cost $\mathcal{V}^*$ [28]. Then, we employ following approximation to express $\mathcal{V}^*$ in a compact set $\mathbb{X} \subset \mathbb{R}^9$

$$\mathcal{V}^*(\boldsymbol{\chi}) = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{\alpha}(\boldsymbol{\chi}) + \epsilon(\boldsymbol{\chi}) \tag{23}$$

where $\boldsymbol{\alpha}(\boldsymbol{\chi}) \triangleq [\alpha_1(\boldsymbol{\chi}), \cdots, \alpha_m(\boldsymbol{\chi})]^{\mathrm{T}} \in \mathbb{R}^{m\times 12}$ is the network's basis function (m denotes the basis number), satisfies that:

$$\begin{aligned}\alpha_i(\boldsymbol{0}_{12\times 1}) &= 0; \\ \frac{\partial}{\partial t}\alpha_i(\boldsymbol{0}_{12\times 1}) &= 0;\end{aligned} \qquad i = 1, 2, \cdots, m$$

Constant vector $\boldsymbol{w} \in \mathbb{R}^m$ is the optimal weight but we are unable to know the value of it. The reconstruction error $\epsilon(\boldsymbol{\chi}) \in \mathbb{R}$ is bounded [28]. Substituting (23) into (19), the explicit expression of $\boldsymbol{\tau}^*$ can be reformulated as:

$$\boldsymbol{\tau}^* = -k_{\tau}\tanh(\frac{\boldsymbol{\mu}^*}{k_{\tau}}) \tag{24}$$

with

$$\boldsymbol{\phi}(\boldsymbol{\chi}) \triangleq \nabla_{\chi}\boldsymbol{\alpha}(\boldsymbol{\chi}),\ \boldsymbol{\varepsilon}(\boldsymbol{\chi}) \triangleq \nabla_{\chi}\epsilon(\boldsymbol{\chi})$$

Because the optimal approximation weight is unknown, the vector $\hat{\boldsymbol{w}}$ is utilized to estimate the value of $\boldsymbol{w}$. Accordingly, the approximation expression of (23) and (19) can be expressed by (25) and (26).

$$\mathcal{V}(\boldsymbol{\chi}) = \hat{\boldsymbol{w}}^{\mathrm{T}}\boldsymbol{\alpha}(\boldsymbol{\chi}) \tag{25}$$

$$\boldsymbol{\tau} = -k_{\tau}\tanh(\frac{\boldsymbol{\mu}}{k_{\tau}}) \tag{26}$$

where $\boldsymbol{\mu} = \frac{1}{2}\boldsymbol{G}^{\mathrm{T}}(\boldsymbol{\chi})\boldsymbol{\phi}(\boldsymbol{\chi})\hat{\boldsymbol{w}}$. If the attitude maneuver achieves the (sub)optimal scheme, the estimated weight $\hat{\boldsymbol{w}}$ should converge to the best value $\boldsymbol{w}$ during the attitude maneuver. Then, the online learning law of estimated weight $\hat{\boldsymbol{w}}$ is going to be introduced in the following part.

C. Online Learning Scheme Design

Consider the following Bellman error:

$$\Delta_b = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{\phi}(\boldsymbol{\chi})(\boldsymbol{F}(\boldsymbol{\chi}) + \boldsymbol{G}(\boldsymbol{\chi})\boldsymbol{\tau}) + \mathfrak{Q}(\boldsymbol{\chi}) + \mathfrak{U}(\boldsymbol{\tau}) \tag{27}$$

Further substitute (25)-(26) into (27), and combine (20) with (27), yield:

$$\begin{aligned}\Delta_b =& \boldsymbol{w}^{\mathrm{T}}\boldsymbol{\phi}^{\mathrm{T}}(\boldsymbol{\chi})(\boldsymbol{F}(\boldsymbol{\chi}) + \boldsymbol{G}(\boldsymbol{\chi})\boldsymbol{\tau}) + \mathfrak{Q}(\boldsymbol{\chi}) + \mathfrak{U}(\boldsymbol{\tau}) \\ & - H(\boldsymbol{\chi}, \boldsymbol{\tau}^*, \nabla_\chi \mathcal{V}^*) \\ =& \mathfrak{Q}(\boldsymbol{\chi}) + \boldsymbol{w}^{\mathrm{T}}\boldsymbol{\phi}^{\mathrm{T}}(\boldsymbol{\chi}) + k_\tau^2 \sum_{i=1}^{3} \log(1 - \tanh^2(\mu_i)) \\ & - H(\boldsymbol{\chi}, \boldsymbol{\tau}^*, \nabla_\chi \mathcal{V}^*) \\ =& \boldsymbol{\nu}^{\mathrm{T}}\tilde{\boldsymbol{w}} + \epsilon_\Delta\end{aligned} \tag{28}$$

For easy expression, we define $\boldsymbol{\nu} \triangleq \boldsymbol{\phi}^{\mathrm{T}}(\boldsymbol{\chi})(\boldsymbol{F}(\boldsymbol{\chi}) + k_\tau^2 \boldsymbol{G}(\boldsymbol{\chi})\mathrm{sgn}(\boldsymbol{\mu}))$, in which $\mathrm{sgn}(\cdot)$ is sign function. The estimated error and the bounded induced reconstruction error are denoted by $\tilde{\boldsymbol{w}} = \hat{\boldsymbol{w}} - \boldsymbol{w}$ and ϵ_Δ, respectively.

Before the learning law design, two necessary assumptions are listed as follows.

Assumption 1. *There exist $T > 0$ and $c_{L1}, c_{L2}, c_{L3} \geq 0$ satisfy the following finite excitation (FE) conditions.*

$$c_{L1}\boldsymbol{I}_m \leq \int_t^{t+T} \frac{\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}}{(\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\nu} + 1)^2}\mathrm{d}t, \forall t > 0 \tag{29}$$

$$c_{L2} \leq \inf_{t>0}\Big(\sum_{k=1}^{p} \lambda_{\min}\Big(\frac{\boldsymbol{\nu}(t_k)\boldsymbol{\nu}^{\mathrm{T}}(t_k)}{(\boldsymbol{\nu}(t_k)^{\mathrm{T}}\boldsymbol{\nu}(t_k) + 1)^2}\Big)\Big) \tag{30}$$

$$c_{L3} \leq \int_t^{t+T} \Big(\sum_{k=1}^{p} \frac{\boldsymbol{\nu}(t_k)\boldsymbol{\nu}^{\mathrm{T}}(t_k)}{(\boldsymbol{\nu}(t_k)^{\mathrm{T}}\boldsymbol{\nu}(t_k) + 1)^2}\Big)\mathrm{d}t, \forall t > 0 \tag{31}$$

Assumption 2. *For $\chi \in \mathbb{X}$ ($\mathbb{X}$ is a compact set), there exist positive constants b_ϕ, b_ϵ, b_ε, and b_Δ, satisfy that $\|\boldsymbol{\phi}\| \leq b_\phi$, $\|\epsilon\| \leq b_\epsilon$, $\|\boldsymbol{\varepsilon}\| \leq b_\varepsilon$, and $\epsilon_\Delta \leq b_\Delta$.*

Note that, the Assumption. 2 is a standard assumption [21], [28], and the initial policy is able to restrict the motion states of attitude system in a compact set.

It is worth noticing that the estimated weight error is contained within Eq. (28). Thereby, the online learning law is designed in term of the Δ_b as:

$$\dot{\hat{\boldsymbol{w}}} = -c_1\gamma\frac{\boldsymbol{\nu}\Delta_b}{(\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\nu} + 1)^2} - c_2\boldsymbol{\Psi} \tag{32}$$

where constant $c_1, c_2 > 0$, Ψ is defined as:

$$\begin{aligned}\boldsymbol{\Psi} =& \sum_{k=1}^{p} \frac{\boldsymbol{\nu}(t_k)\boldsymbol{\nu}^{\mathrm{T}}(t_k)\hat{\boldsymbol{w}} + (\mathfrak{Q}(t_k) + \mathfrak{U}(t_k))\boldsymbol{\nu}(t_k)}{(\boldsymbol{\nu}^{\mathrm{T}}(t_k)\boldsymbol{\nu}(t_k) + 1)^2} \\ =& \boldsymbol{\Theta}\tilde{\boldsymbol{w}} + \epsilon_\Theta\end{aligned} \tag{33}$$

in which, $\boldsymbol{\Theta} = 2\sum_{k=1}^{p}(\boldsymbol{\nu}(t_k)\boldsymbol{\nu}^{\mathrm{T}}(t_k))/(\boldsymbol{\nu}^{\mathrm{T}}(t_k)\boldsymbol{\nu}(t_k) + 1)^2$, and ϵ_Θ is the bounded residual error.

The gain γ in (32) is a time varying value, and its update law is shown as the following equation:

$$\dot{\gamma} = -(-g_\gamma\gamma + \gamma^2 c_1\frac{\|\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}\|}{(\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\nu} + 1)^2} + \gamma^2 c_2\lambda_{\min}(\boldsymbol{\Theta})) \tag{34}$$

in which, g_γ is a positive constant, initial value of γ is denoted by $\gamma_0 > 0$.

Remark 1. *Note that $\boldsymbol{\Psi}$ is an experience stack, which contains the motion information at previous time instants $t_1, t_2, \cdots, t_p$. In order to satisfy the FE condition, the experience stack recording algorithm, similar to [29], [30], is developed to maximize the minimum singular value of $\sum_{k=1}^{p}\boldsymbol{\nu}(t_k)\boldsymbol{\nu}^{\mathrm{T}}(t_k)$ by swapping the current data with every data in the stack and then comparing the corresponding minimum eigenvalues. The detailed procedure is introduced by Algorithm 1.*

Algorithm 1 Experience Stack Recording Algorithm

Set $k = 1$, $\boldsymbol{\nu}_{\mathrm{stock}} = \boldsymbol{0}_{m\times p}$, $\boldsymbol{w}_{\mathrm{stock}} = \boldsymbol{0}_{m\times p}$, $\mathfrak{Q}_{\mathrm{stock}} = \boldsymbol{0}_{1\times p}$, $\mathfrak{U}_{\mathrm{stock}} = \boldsymbol{0}_{1\times p}$, $s_{\mathrm{flag}} = 0$
if $k \leq p$ **then**
  $\boldsymbol{\nu}_{\mathrm{stock}}(:,k) = \boldsymbol{\nu}$, $\boldsymbol{w}_{\mathrm{stock}}(:,k) = \hat{\boldsymbol{w}}$, $\mathfrak{Q}_{\mathrm{stock}}(:,k) = \mathfrak{Q}$, $\mathfrak{U}_{\mathrm{stock}}(:,k) = \mathfrak{U}$
else
  $\boldsymbol{T}_{\mathrm{stock}} = \boldsymbol{\nu}_{\mathrm{stock}}$, $S_{\min} = \lambda_{\min}(\sum_{k=1}^{N}\boldsymbol{\nu}(t_k)\boldsymbol{\nu}^{\mathrm{T}}(t_k))$
  for $i = 1$ to p **do**
    $\boldsymbol{\nu}_{\mathrm{stock}}(:,j) = \boldsymbol{\nu}$, $S_{\mathrm{temp}} = \lambda_{\min}(\sum_{k=1}^{N}\boldsymbol{\nu}(t_k)\boldsymbol{\nu}^{\mathrm{T}}(t_k))$
    $\boldsymbol{\nu}_{\mathrm{stock}} = \boldsymbol{T}_{\mathrm{stock}}$
    if $S_{\mathrm{temp}} < S_{\min}$ **then**
      $S_{\min} = S_{\mathrm{temp}}$, $s_{\mathrm{flag}} = i$
    end if
  end for
  if $s_{\mathrm{flag}} \neq 0$ **then**
    $\boldsymbol{\nu}_{\mathrm{stock}}(:,s_{\mathrm{flag}}) = \boldsymbol{\nu}$, $\boldsymbol{w}_{\mathrm{stock}}(:,s_{\mathrm{flag}}) = \hat{\boldsymbol{w}}$, $\mathfrak{Q}_{\mathrm{stock}}(:,s_{\mathrm{flag}}) = \mathfrak{Q}$, $\mathfrak{U}_{\mathrm{stock}}(:,s_{\mathrm{flag}}) = \mathfrak{U}$
  end if
end if
Calculate $\boldsymbol{\Psi}$ by (33)

D. Discussion and Analysis

This part presents the relevant analysis and discussion. The following lemma establishing upper and lower bounds of the update gain γ for subsequent stability and convergence analysis.

Lemma 1. *The update law given in* (34)*, there exist two constants $\underline{\gamma}$, $\overline{\gamma}$ such that gain satisfies $0 < \underline{\gamma} \leq \gamma \leq \overline{\gamma}$ under Assumption 1.*

Proof. The proof procedures closely follow that in [31]. Divide both sides of the update law (34) by $-\gamma^2$ can be rewritten as:

$$\frac{\partial\gamma^{-1}}{\partial t} = -\frac{\dot{\gamma}}{\gamma^2} = -g_\gamma\frac{1}{\gamma} + c_1\frac{\|\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}\|}{(\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\nu} + 1)^2} + c_2\lambda_{\min}(\boldsymbol{\Theta}) \tag{35}$$

Hence,

$$\begin{aligned}\gamma^{-1} =& \exp(-g_\gamma t)\gamma_0^{-1} + c_1\int_0^t \exp(-g_\gamma(t-\iota))\frac{\|\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}\|}{(\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\nu} + 1)^2}\mathrm{d}\iota \\ & + c_2\int_0^t -g_\gamma(t-\iota)\lambda_{\min}(\boldsymbol{\Theta})\mathrm{d}\iota\end{aligned} \tag{36}$$

When $t < T$, one has:

$$\gamma^{-1} \geq \exp(-g_\gamma t)\gamma_0^{-1} \tag{37}$$

If $t \geq T$, using Assumption 1:

$$\begin{aligned}\gamma^{-1} \geq & \exp(-g_\gamma T)(c_1 \int_{t-T}^{t} \frac{\|\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}\|}{(\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\nu}+1)^2}\mathrm{d}\iota \\ & + c_2 \int_{t-T}^{t} \lambda_{\min}(\boldsymbol{\Theta})\mathrm{d}\iota) \\ \geq & \exp(-g_\gamma T)c_1 c_{L1} + 2c_2 \max\{c_{L2}T, c_{L3}\}\end{aligned} \tag{38}$$

Furthermore, using the facts that $\|\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}\|/(\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\nu}+1)^2 \leq \frac{1}{2}$, one has:

$$\begin{aligned}\gamma^{-1} \leq & \int_0^t \exp(-g_\gamma(t-\iota)(\frac{1}{2}c_1 + c_2 p))\mathrm{d}\iota + \exp(-g_\gamma t)\gamma_0^{-1} \\ \leq & \gamma_0^{-1} + \frac{c_1 + 2c_2 p}{2g_\gamma}\end{aligned} \tag{39}$$

Therefore, through the above analysis, the upper bound $\overline{\gamma} = (\exp(-g_\gamma T)\min\{\gamma_0^{-1}, c_1 c_{L1} + 2c_2 \max\{c_{L2}T, c_{L3}\}\})^{-1}$, and the lower bound $\underline{\gamma} = (\gamma_0^{-1} + \frac{c_1+2c_2 p}{2g_\gamma})^{-1}$. The proof is complete. □

Theorem 1. *Consider the attitude reorientation model with the the prescribed performance error presented in* (14)*, and the controller in* (25)-(26)*. Design the online learning law* (32)-(34)*. Then, under the Assumption.2, $\boldsymbol{\chi}$ and $\tilde{\boldsymbol{w}}$ are uniformly ultimately bounded.*

Proof. Consider the following candidate Lyapunov function:

$$\mathcal{L} = \mathcal{V}^*(\boldsymbol{\chi}) + \frac{\beta}{2\gamma}\tilde{\boldsymbol{w}}^{\mathrm{T}}\tilde{\boldsymbol{w}} \tag{40}$$

where $\beta > 0$ is positive constant which is introduced just for the sake of analysis. Obviously, the candidate Lyapunov function is positive definite. Then analyze the time derivative of it and recall (22), yield.

$$\begin{aligned}\dot{\mathcal{L}} = & \nabla_\chi^{\mathrm{T}}\mathcal{V}^*(\boldsymbol{F} + \boldsymbol{G}\boldsymbol{\tau}) + \frac{\beta}{\gamma}\tilde{\boldsymbol{w}}^{\mathrm{T}}\dot{\tilde{\boldsymbol{w}}} - \frac{\beta\dot{\gamma}}{2\gamma^2}\tilde{\boldsymbol{w}}^{\mathrm{T}}\tilde{\boldsymbol{w}} \\ = & -\mathfrak{Q} - k_\tau^2 \sum_{i=1}^{3}\log(1-\tanh^2(\mu_i^*)) - 2k_\tau^2\frac{\boldsymbol{\mu}^{\mathrm{T}}}{k_\tau}\tanh(\frac{\boldsymbol{\mu}}{k_\tau}) \\ & + k_\tau\tilde{\boldsymbol{w}}^{\mathrm{T}}\boldsymbol{\phi}^{\mathrm{T}}\boldsymbol{G}\tanh(\frac{\boldsymbol{\mu}}{k_\tau}) - k_\tau\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{G}\tanh(\frac{\boldsymbol{\mu}}{k_\tau}) + \frac{\beta}{\gamma}\tilde{\boldsymbol{w}}^{\mathrm{T}}\dot{\tilde{\boldsymbol{w}}} \\ & - \frac{\beta\dot{\gamma}}{2\gamma^2}\tilde{\boldsymbol{w}}^{\mathrm{T}}\tilde{\boldsymbol{w}}\end{aligned} \tag{41}$$

Then, substitute (32)-(34) into (41), then employing the facts that $(\frac{\boldsymbol{\mu}}{k_\tau})^{\mathrm{T}}\tanh(\frac{\boldsymbol{\mu}}{k_\tau}) \geq \tanh^{\mathrm{T}}(\frac{\boldsymbol{\mu}}{k_\tau})\tanh(\frac{\boldsymbol{\mu}}{k_\tau})$ and arithmetic–geometric average inequality, it can deduce that:

$$\begin{aligned}\dot{\mathcal{L}} \leq & -\mathfrak{Q} - k_\tau^2 \sum_{i=1}^{3}\log(1-\tanh^2(\mu_i^*)) - k_\tau^2\frac{\boldsymbol{\mu}^{\mathrm{T}}}{k_\tau}\tanh(\frac{\boldsymbol{\mu}}{k_\tau}) \\ & + \frac{k_\tau}{2}\tilde{\boldsymbol{w}}^{\mathrm{T}}\boldsymbol{\phi}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{\phi}\tilde{\boldsymbol{w}} - c_1\beta\frac{\tilde{\boldsymbol{w}}^{\mathrm{T}}\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}\tilde{\boldsymbol{w}}}{2(\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\nu}+1)^2} + \frac{c_2\beta}{2}\tilde{\boldsymbol{w}}^{\mathrm{T}}\boldsymbol{\Theta}\tilde{\boldsymbol{w}} \\ & - \frac{\beta g_\gamma}{\gamma}\tilde{\boldsymbol{w}}^{\mathrm{T}}\tilde{\boldsymbol{w}} + \frac{c_1\beta}{2}\tilde{\boldsymbol{w}}^{\mathrm{T}}\frac{\|\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}\|\boldsymbol{I}_m}{(\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\nu}+1)^2}\tilde{\boldsymbol{w}} \\ & + \frac{c_2\beta}{2}\tilde{\boldsymbol{w}}^{\mathrm{T}}\lambda_{\min}(\boldsymbol{\Theta})\boldsymbol{I}_m\tilde{\boldsymbol{w}} + \epsilon_L \\ \leq & -\boldsymbol{e}^{\mathrm{T}}\boldsymbol{Q}_e\boldsymbol{e} - \boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{Q}_\sigma\boldsymbol{\sigma} - \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{Q}_\omega\boldsymbol{\omega} - \tilde{\boldsymbol{w}}^{\mathrm{T}}\boldsymbol{D}\tilde{\boldsymbol{w}} + \epsilon_L\end{aligned} \tag{42}$$

where $\boldsymbol{D} = \frac{\beta g_\gamma}{\overline{\gamma}}\boldsymbol{I}_m - \frac{1}{2}\boldsymbol{\phi}^{\mathrm{T}}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{\phi}$, and recall Assumption 2, ϵ_L is defined to be the upper bound of $\frac{1}{2}b_\varepsilon^2 + \frac{\beta c_1}{4}b_\Delta^2 + \frac{\beta c_2}{2}\epsilon_\Theta^2$. Therefore, by appropriately setting the auxiliary parameters β to satisfy

$$\beta \geq \frac{\overline{\gamma}b_\phi^2}{2g_\gamma} \tag{43}$$

the ultimate boundedness of $\tilde{\boldsymbol{w}}$ and $\boldsymbol{\chi}$ are guaranteed. □

Remark 2. *The proof of Theorem 1 shows $\mathbf{e} \in \mathcal{L}_\infty$, which indicates that the prescribed performance of the attitude can be satisfied. What's more, the boundedness of $\tilde{\boldsymbol{w}}$ implies that the control scheme ultimately achieves the near-optimal controller. Hence, the stability and optimality can be guaranteed.*

Remark 3. *Compared with the existing learning-based prescribed performance spacecraft control methods [22], [23], the ADP-based controller merely plays as a supplementary of the traditional control scheme. In view of optimization, this design is conservative and partial, which cannot optimize the overall cost. Conversely, our design achieves the overall controller's cost optimization.*

According to the aforementioned design and analysis, the whole system's major framework of information flow is concluded in Fig. 3 to intuitively illustrate the proposed method.

IV. Numerical Simulation and Experimental Verification

A. Scenario Description

The scenario considered in this paper is a rigid body satellite, which is demanded to observe a specific target. The detection instrument is installed aligned with its Z-axis. The mission required to point the sightline of instrument to the target under preset convergence performance.

The detailed information of the rigid body satellite is given as: the inertial matrix $\boldsymbol{J} = [20.0, 1.2, 0.9; 1.2, 17.0, 1.4; 0.9, 1.4, 15.0]$ kg $\cdot$ m^2, the initial motion states $\boldsymbol{\sigma}(0) = [0.76, 0.12, -0.31]^{\mathrm{T}}$ and $\boldsymbol{\omega}(0) = [0.0, 0.0, 0.0]^{\mathrm{T}}$rad/s, respectively. The maximum output torque of actuators set to be $k_\tau = 0.1$Nm. The parameters of prescribed performance are listed in Table I.

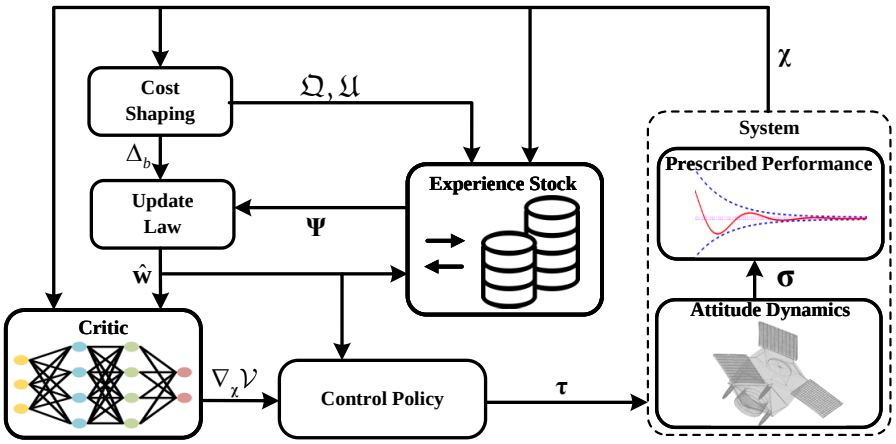

Fig. 3. The major framework and information flow.

TABLE I
PARAMETERS OF PRESCRIBED PERFORMANCE.

Parameter	Values/Ranges
α_ρ	1
k_ρ	0.2
ρ_{Li}	1.0×10^{-4}
ρ_{Hi}	$1.5\|\boldsymbol{\sigma}_i(0)\|$ $(i \in \{1,2,3\})$

B. Numerical Example

The numerical example runs on an Intel-i7 Core (1.80 GHz) Laptop computer under the *RK4* (4th Runge–Kutta method) numerical solver with 0.05s sample time, in this digital environment the effectiveness and advantages of the proposed method will be presented.

Before proceeding, let the learning parameters to be: $c_1 = 1$, $c_2 = 0.01$, $g_\gamma = 5.0 \times 10^{-3}$, and $\gamma_0 = 1$. The weight matrices follows: $\boldsymbol{Q}_e = 0.01\boldsymbol{I}_3$, $\boldsymbol{Q}_\sigma = \boldsymbol{I}_3$, $\boldsymbol{Q}_\omega = 5\boldsymbol{I}_3$. The basis function and corresponding initial weight are chosen as: $\boldsymbol{\alpha}(\boldsymbol{\chi}) = 0.5[2e_1\omega_1, 2e_2\omega_2, 2e_3\omega_3, 2\sigma_1\omega_1, 2\sigma_2\omega_2, 2\sigma_3\omega_3, \omega_1^2, \omega_2^2, \omega_3^2]^{\mathrm{T}}$, and $\hat{\boldsymbol{w}}(0) = [0.0, 0.0, 0.0, 0.2, 0.2, 0.2, 10, 10, 10]^{\mathrm{T}}$. Note that, the above combination of basis function and initial weight can be reconstructed as:

$$\boldsymbol{\tau} = -0.1 \tanh\left(\frac{0.1\boldsymbol{\sigma} + 5\boldsymbol{\omega}}{0.1}\right) \tag{44}$$

which ensures the initial stability of the attitude system.

Based on the above configurations, the time responses curves of control input and attitude system states under the proposed control method are shown in Figs. 4-7.

The numerical results present that the proposed control method achieves the attitude reorientation objective with the attitude accuracy is better than 1.0×10^{-4} under the actuators' constraint. In Fig. 4, the argument state converges to zero, which implies the controller meets the prescribed performance specification. Besides, Figs. 7-8 also verify this result. It can be seen from Fig. 8, it gives four screenshots of the Z-axis pointing admissible region (colored by green), which becomes gradually smaller with time (corresponds to the varying of the prescribed performance constraint) until the desired orientation. Meanwhile, it can be noted that the projection of the direction of Z-axis on the unit sphere is always located in the admissible region, which means that the prescribed performance constraint can be guaranteed.

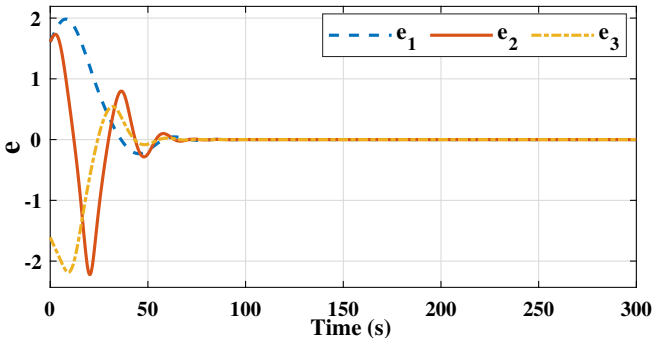

Fig. 4. Time responses of transformed error $\boldsymbol{e}$ under numerical simulation.

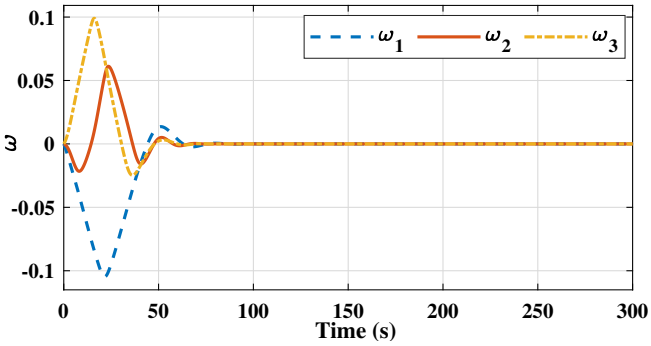

Fig. 5. Time responses of $\boldsymbol{\omega}$ under numerical simulation.

The online learning results of the proposed method are given in Figs. 9-10. Since the optimal weight is unknown, we can employ the Δ_b to evaluate the learning effectiveness. It can be observed that the Bellman error tends to zero, and the weight estimation also tends to be stable, correspondingly. Therefore, the control scheme improves to the near-optimal

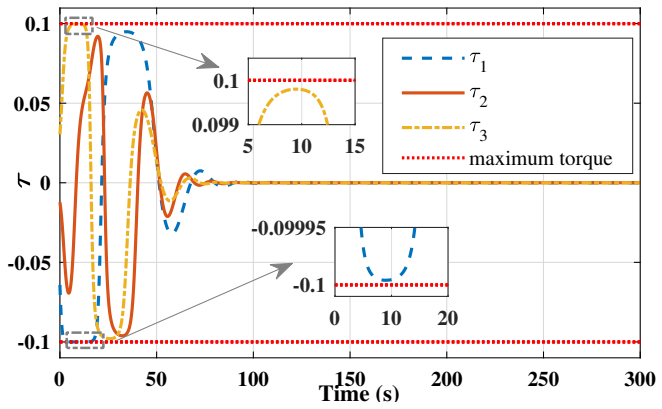

Fig. 6. Time responses of $\boldsymbol{\tau}$ under numerical simulation.

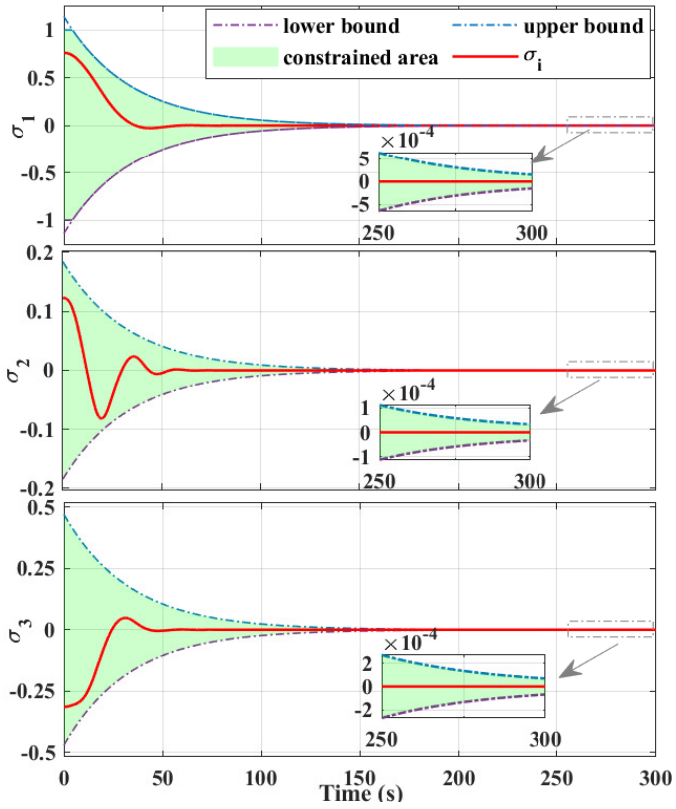

Fig. 7. Time responses of $\boldsymbol{\sigma}$ with prescribed performance function under numerical simulation.

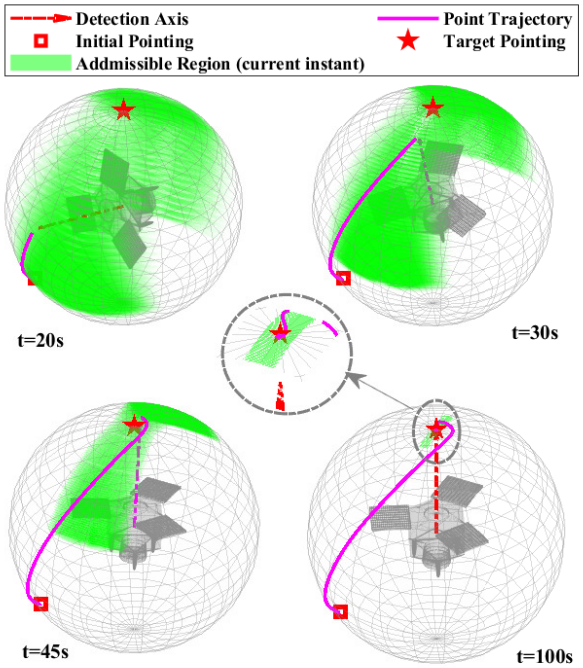

Fig. 8. 3D illustration of rotation at different time instance with admissible region.

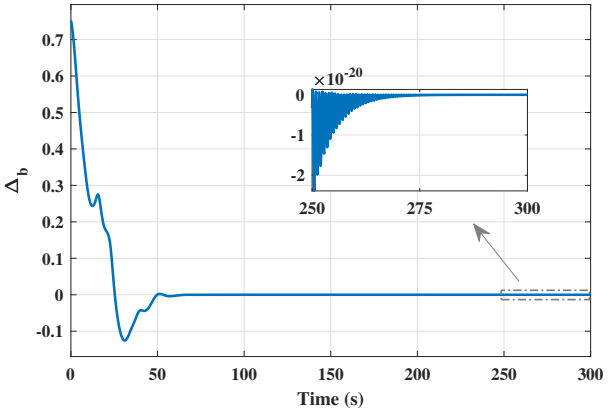

Fig. 9. Time responses of Bellman error Δ_b.

scheme from an initial controller. The related cost comparison with the initial controller is demonstrated in Fig. 11. For fair comparison, the performance index is employed as: $r_{\text{comp}} = \int_0^t (\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{Q}_\sigma\boldsymbol{\sigma}+\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{Q}_\omega\boldsymbol{\omega}+\boldsymbol{\tau}^{\mathrm{T}}\boldsymbol{\tau})\mathrm{d}t$. From the bar charts in Fig. 11, the proposed method presents a better performance and renders a lower overall cost. This also demonstrates the effectiveness of data-driven online improvement.

C. Experiment Validations

A practical experimental studies are conducted on a professional spacecraft attitude control HITL experimental system, as presented in Fig. 12, to validate the practicality of the proposed method. The HITL experimental system is composed of fiber-optic gyroscopes (FOGs), a three-axis turntable, reaction wheels assembly (RWA), and a real-time simulation computer. For more detailed introduction of this experimental system can be found in [32].

The experiment scenario is similar to the numerical simulation. Inevitably, the disturbances of control signals, and the measurement noise of motion state (angular velocity $\boldsymbol{\omega}$ and attitude $\boldsymbol{\sigma}$) will be introduced into the control-loop. Besides, the time duration is set to be 150s in order to improve experimental efficiency.

The HITL experimental results are given in Figs. 13-15.

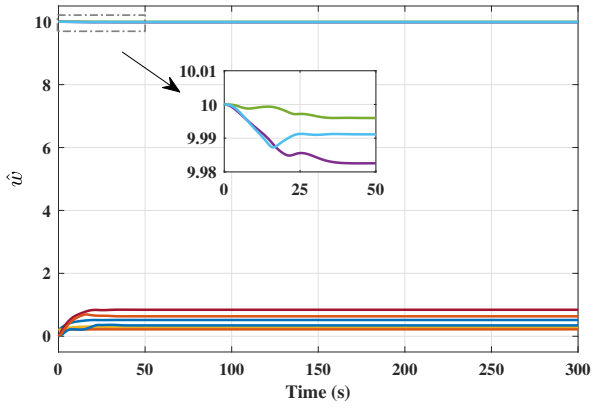

Fig. 10. Estimation of the weight $\hat{\boldsymbol{w}}$.

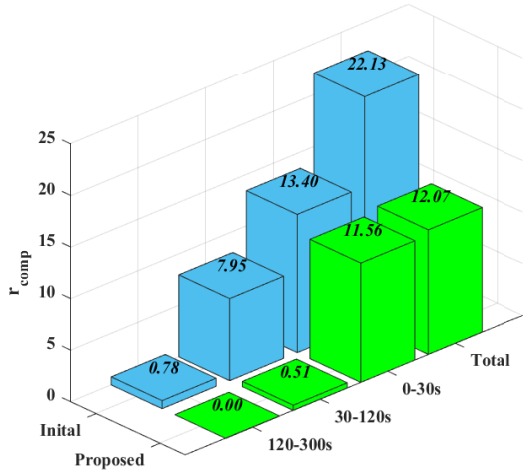

Fig. 11. Overall cost comparison.

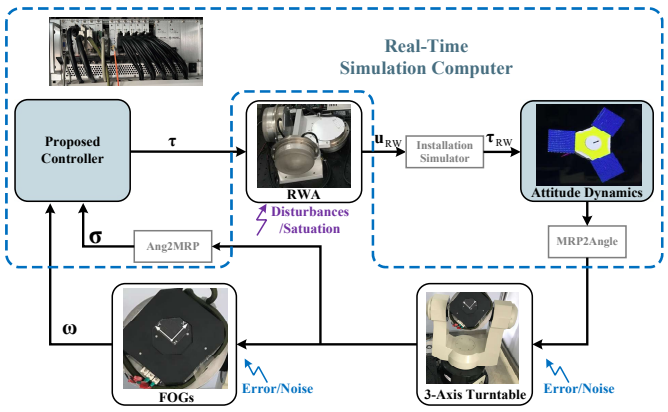

Fig. 12. Structure diagram of the spacecraft attitude control HITL experiment platform.

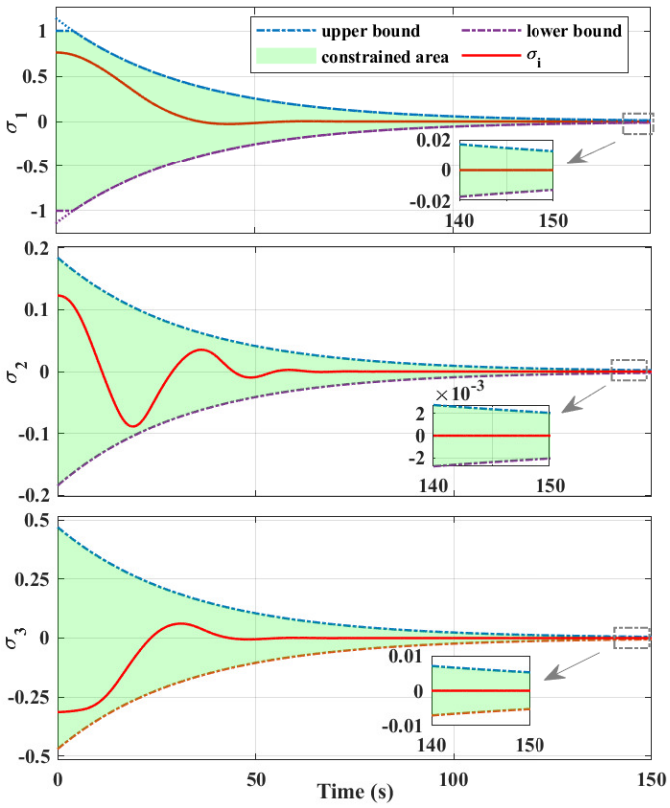

Fig. 13. Experiment result of $\boldsymbol{\sigma}$ with prescribed performance function.

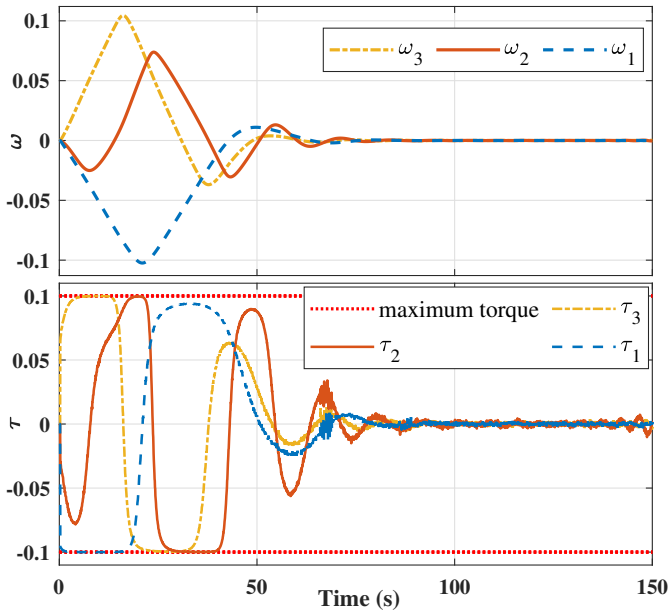

Fig. 14. Experiment result of $\boldsymbol{\omega}$ and $\boldsymbol{\tau}$.

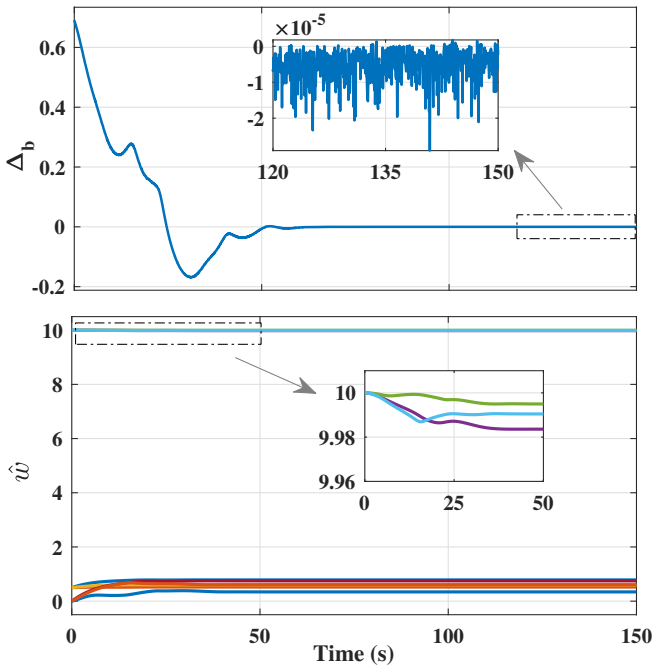

Fig. 15. Experiment result of online learning.

As shown in these figures, the HITL experiment demonstrates the proposed method is able to achieve the mission objective. The general trend of the experimental curves are similar to that in the simulation, which means that the proposed method generally arrives the performance of the simulation in the experiment. Note that, there is a remarkable feature in Fig. 14 that the visible fluctuations appears on the control torque curves. It is attributable to the hardware hardware characteristics of reaction wheels, its actual output torque is with irreducible disturbances. This feature also affects the reduction of learning accuracy in Fig.15 compared to the simulation results.

In general, both numerical example and HITL experiment results demonstrate that the data-driven controller proposed in this paper has the capability to simultaneously achieve the approximate optimal prescribed performance control under torque saturation.

V. Conclusion

A data-driven optimal prescribed performance controller is devoted to spacecraft attitude maneuvering under the actuator saturation. Specially, the nontrivial update law is presented, achieving real-time scheme improvement by utilizing the online measurable information. Then, the critical requirements in aerospace missions that prescribed performance guarantees, actuator saturation, and cost optimization can be simultaneously satisfied. In terms of theory, the ultimately bounded stability of the whole system and the convergence of optimization are proven by the Lyapunov-based method.

Notably, from the view of practical implementation, the representative numerical simulation and HITL experiment are carried out to evaluate the application value of the proposed method. It would be interesting to conduct in-depth practical research in real space-engineering systems in future work.

References

[1] H. Bang, M. J. Tahk, and H. D. Choi, "Large angle attitude control of spacecraft with actuator saturation," *Control Engineering Practice*, vol. 11, no. 9, pp. 989–997, 2003.

[2] S. P. Arjun Ram and M. R. Akella, "Uniform exponential stability result for the rigid-body attitude tracking control problem," *Journal of Guidance, Control, and Dynamics*, vol. 43, no. 1, pp. 39–45, 2020.

[3] R. Q. Dong, A. G. Wu, and Y. Zhang, "Anti-unwinding sliding mode attitude maneuver control for rigid spacecraft," *IEEE Transactions on Automatic Control*, vol. 67, no. 2, pp. 978–985, 2022.

[4] R. Q. Dong, A. G. Wu, Y. Zhang, and G. R. Duan, "Anti-unwinding sliding mode attitude control via two modified rodrigues parameter sets for spacecraft," *Automatica*, p. 109642, 2021.

[5] A. M. Zanchettin, A. Calloni, and M. Lovera, "Robust magnetic attitude control of satellites," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 4, pp. 1259–1268, aug 2013.

[6] X. Cao, P. Shi, Z. Li, and M. Liu, "Neural-network-based adaptive backstepping control with application to spacecraft attitude regulation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 9, pp. 4303–4313, 2018.

[7] C. P. Bechlioulis and G. A. Rovithakis, "Adaptive control with guaranteed transient and steady state tracking error bounds for strict feedback systems," *Automatica*, vol. 45, no. 2, pp. 532–538, 2009.

[8] C. Wei, Q. Chen, J. Liu, Z. Yin, and J. Luo, "An overview of prescribed performance control and its application to spacecraft attitude system," *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering*, vol. 235, no. 4, pp. 435–447, 2021.

[9] Z. G. Zhou, Y. A. Zhang, X. N. Shi, and D. Zhou, "Robust attitude tracking for rigid spacecraft with prescribed transient performance," *International Journal of Control*, vol. 90, no. 11, pp. 2471–2479, 2017.

[10] S. Gao, X. Liu, Y. Jing, and G. M. Dimirovski, "Finite-time prescribed performance control for spacecraft attitude tracking," *IEEE/ASME Transactions on Mechatronics*, 2021, doi:10.1109/TMECH.2021.3108558.

[11] Q. Hu, X. Shao, and L. Guo, "Adaptive fault-tolerant attitude tracking control of spacecraft with prescribed performance," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 1, pp. 331–341, feb 2018.

[12] X. Shao, Q. Hu, Y. Shi, and B. Jiang, "Fault-tolerant prescribed performance attitude tracking control for spacecraft under input saturation," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 2, pp. 574–582, 2020.

[13] M. Liu, X. Shao, and G. Ma, "Appointed-time fault-tolerant attitude tracking control of spacecraft with double-level guaranteed performance bounds," *Aerospace Science and Technology*, vol. 92, pp. 337–346, 2019.

[14] R. Sharma and A. Tewari, "Optimal nonlinear tracking of spacecraft attitude maneuvers," *IEEE Transactions on Control Systems Technology*, vol. 12, no. 5, pp. 677–682, 2004.

[15] D. Y. Lee, R. Gupta, U. V. Kalabić, S. Di Cairano, A. M. Bloch, J. W. Cutler, and I. V. Kolmanovsky, "Geometric mechanics based nonlinear model predictive spacecraft attitude control with reaction wheels," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 2, pp. 309–319, 2017.

[16] T. Sun, X.-M. Sun, X. Zhao, and H. Liu, "Attitude control of rigid bodies: An energy-optimal geometric switching control approach," *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 2, pp. 1162–1173, apr 2022.

[17] B. A. Kristiansen, J. T. Gravdahl, and T. A. Johansen, "Energy optimal attitude control for a solar-powered spacecraft," *European Journal of Control*, vol. 62, pp. 192–197, 2021.

[18] D. Liu, Q. Wei, D. Wang, X. Yang, and H. Li, *Adaptive Dynamic Programming with Applications in Optimal Control*, ser. Advances in Industrial Control. Cham, Switzerland: Springer, 2017.

[19] L. Kong, S. Zhang, and X. Yu, "Approximate optimal control for an uncertain robot based on adaptive dynamic programming," *Neurocomputing*, vol. 423, pp. 308–317, 2021.

[20] L. Kong, W. He, C. Yang, and C. Sun, "Robust neurooptimal control for a robot via adaptive dynamic programming," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 6, pp. 2584–2594, 2021.

[21] H. Dong, X. Zhao, and H. Yang, "Reinforcement learning-based approximate optimal control for attitude reorientation under state constraints," *IEEE Transactions on Control Systems Technology*, vol. 29, no. 4, pp. 1664–1673, 2020.

[22] Z. Yin, J. Luo, and C. Wei, "Novel adaptive saturated attitude tracking control of rigid spacecraft with guaranteed transient and steady-state performance," *Journal of Aerospace Engineering*, vol. 31, no. 5, p. 04018062, 2018.
[23] C. Wei, J. Luo, H. Dai, and G. Duan, "Learning-based adaptive attitude control of spacecraft formation with guaranteed prescribed performance," *IEEE Transactions on Cybernetics*, vol. 49, no. 11, pp. 4004–4016, 2019.
[24] M. D. Shuster, "A survey of attitude representations," *Journal of the Astronautical Sciences*, vol. 41, no. 4, pp. 439–517, 1993.
[25] H. Schaub and J. L. Junkins, "Stereographic orientation parameters for attitude dynamics: A generalization of the rodrigues parameters," *Journal of the Astronautical Sciences*, vol. 44, no. 1, pp. 1–19, 1996.
[26] R. Kristiansen and D. Hagen, "Modelling of actuator dynamics for spacecraft attitude control," *Journal of Guidance, Control, and Dynamics*, vol. 32, no. 3, pp. 1022–1025, 2009.
[27] I. N. Bronshtein, K. A. Semendyayev, G. Musiol, and H. Muehlig, *Handbook of Mathematics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.
[28] K. G. Vamvoudakis and F. L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878–888, 2010.
[29] G. Chowdhary, M. Mühlegg, and E. Johnson, "Exponential parameter and tracking error convergence guarantees for adaptive controllers without persistency of excitation," *International Journal of Control*, vol. 87, no. 8, pp. 1583–1603, 2014.
[30] H. Dong, Q. Hu, M. R. Akella, and H. Yang, "Composite adaptive attitude-tracking control with parameter convergence under finite excitation," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 6, pp. 2657–2664, 2020.
[31] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Mineola, New York: Courier Corporation, 2012.
[32] H. Yang, Q. Hu, H. Dong, and X. Zhao, "Adp-based spacecraft attitude control under actuator misalignment and pointing constraints," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 9, pp. 9342–9352, 2022.

Haoyang Yang (Graduate Student Member, IEEE) received the B.Eng. degree from the School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, China, in 2017. He is currently pursuing the Ph.D. degree in navigation, guidance, and control with Beihang University, Beijing, China. His current research interests include reinforcement learning-based control, intelligent control, and attitude and pose control. He is also working on the hardware-in-the-loop experiments for various nonlinear control systems.

Qinglei Hu (Senior Member, IEEE) received the B.Eng. degree in electrical and electronic engineering from Zhengzhou University, Zhengzhou, China, in 2001, and the Ph.D. degree in control science and engineering with the specialization in guidance and control from the Harbin Institute of Technology, Harbin, China, in 2006.

From 2003 to 2014, he was with the Department of Control Science and Engineering, Harbin Institute of Technology. He joined Beihang University, Beijing, China, in 2014, as a Full Professor. His current research interests include variable structure control and applications, and fault-tolerant control and applications. In these areas, he has authored or coauthored more than 80 technical articles. Prof. Hu serves as an Associate Editor for Aerospace Science and Technology.

Hongyang Dong is an Assistant Professor in the School of Engineering, University of Warwick, Coventry, UK. He obtained his PhD degree in Control Science and Engineering from Harbin Institute of Technology, Harbin, China, in 2018. His research interests include control theories and machine learning with their applications in offshore renewable energy and autonomous systems.

Xiaowei Zhao is Professor of Control Engineering and an EPSRC Fellow at the School of Engineering, University of Warwick, Coventry, UK. He obtained the PhD degree in Control Theory from Imperial College London in 2010. After that he worked as a postdoctoral researcher at the University of Oxford for three years before joining Warwick in 2013. His main research areas are control theory with applications on offshore renewable energy systems, local smart energy systems, and autonomous systems.

Dongyu Li (S'16-M'19) received the B.S. and Ph.D. degree from Control Science and Engineering, Harbin Institute of Technology, China, in 2016 and 2020. He was a joint Ph.D. student with the Department of Electrical and Computer Engineering, National University of Singapore from 2017 to 2019, and a research fellow with the Department of Biomedical Engineering, National University of Singapore, from 2019 to 2021. He is currently an Associate Professor with the School of Cyber Science and Technology, Beihang University, China. His research interests include networked system cooperation, adaptive systems, and spacecraft engineering.