Available online at www.sciencedirect.com

ScienceDirect

Journal of the Franklin Institute 355 (2018) 803–826

www.elsevier.com/locate/jfranklin

Trajectory tracking control for rotary steerable systems using interval type-2 fuzzy logic and reinforcement learning

Chi Zhang[a,b], Wei Zou[a,c], Ningbo Cheng[a,*], Junshan Gao[b]

[a] *Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, PR China*
[b] *School of Automation, Harbin University of Science and Technology, Harbin, 150080, PR China*
[c] *Tianjin Intelligent Technology Institute of CASIA Co. Ltd, Tianjin, 300309, PR China*

Received 12 September 2017; received in revised form 18 November 2017; accepted 4 December 2017
Available online 8 December 2017

Abstract

Rotary steerable system (RSS) is a directional drilling technique which has been applied in oil and gas exploration under complex environment for the requirements of fossil energy and geological prospecting. The nonlinearities and uncertainties which are caused by dynamical device, mechanical structure, extreme downhole environment and requirements of complex trajectory design in the actual drilling work increase the difficulties of accurate trajectory tracking. This paper proposes a model-based dual-loop feedback cooperative control method based on interval type-2 fuzzy logic control (IT2FLC) and actor-critic reinforcement learning (RL) algorithms with one-order digital low-pass filters (LPF) for three-dimensional trajectory tracking of RSS. In the proposed RSS trajectory tracking control architecture, an IT2FLC is utilized to deal with system nonlinearities and uncertainties, and an online iterative actor-critic RL controller structured by radial basis function neural networks (RBFNN) and adaptive dynamic programming (ADP) is exploited to eliminate the stick–slip oscillations relying on its approximate properties both in action function (actor) and value function (critic). The two control effects are fused to constitute cooperative controller to realize accurate trajectory tracking of RSS. The effectiveness of our controller is validated by simulations on designed function tests for angle building hole rate and complete downhole trajectory tracking, and by comparisons with other control methods.

* Corresponding author.
E-mail address: ningbo.cheng@ia.ac.cn (N. Cheng).

https://doi.org/10.1016/j.jfranklin.2017.12.001

1. Introduction

Deep-seated and offshore hydrocarbon exploration is of significance to obtain fossil energy resources while facing worldwide energy shortage. Rotary steerable system (RSS) is a directional drilling technique which has been expected to drill complex curved borehole trajectories accurately and smoothly in the complex stratum or deep ocean under extreme working conditions [1,2]. Nevertheless, in actual drilling experiments the trajectory tracking errors and stick–slip oscillations always exist. The borehole trajectories may even deviate from the desired drilling trajectory in overly complicated and extreme situations. The downhole environment usually contains different rock layer hardness, high/low temperature and pressure, ray interference and other unknown disturbances, which not only lead numerous intense nonlinearity and time-variation to systems, but also increase the difficulties of control for RSS. In order to guarantee convergence and improve the trajectory tracking accuracy, a number of advanced control algorithms have been studied and applied in RSS [3–8].

A great number of RSS control systems only consider an empirical model or a numerical model, which directly link the overall angle change rate to the directional force applied in RSS. For example, an empirical model is applied and the attitude control of bottom hole assembly (BHA) has been realized in [9]. Most oilfield technique companies also design and manufacture their RSS mechanisms via empirical models or numerical models and use open loop or traditional PID controller to realize downhole trajectory tracking. Empirical models could not fully reflect the dynamic behaviors and variations of the system, and they are difficult to be applied to simulate real drilling model in borehole propagation [10]. Other researches based on empirical models focus on stability control of a specific part of RSS including downhole stability platform, disc valve, servo motor or other downhole directional mechanisms. These works could not guarantee the stability of overall system and are hard to realize synergetic control of different parts of RSS to obtain desired borehole trajectory. In recent years, several analytical models are presented. In [11], according to three different simplifications and assumptions, several novel RSS models are established as linear uncertain borehole propagation equations with internal delays based on dynamic mechanical analysis by Downton. In [12–16], Perneder and Detournay formulate a three-dimensional directional drilling model using nonlinear delay differential equations, which is the benchmark model of our work and will be introduced in detail in Section 2. In the way of controller design based on analytical model, a L1 adaptive controller containing a state predictor, an adaptive law and a control law based on RSS rigid mathematical model [11] has been proposed in [4]. Based on two-dimensional simplification of RSS model [13], a robust output feedback control method for guaranteeing system stability and decreasing borehole oscillations has been presented in [3,6]. Based on a directional drilling model which combines the coupled axial-torsional drill string dynamics and a rate-independent bit-rock interaction law [16] including both cutting and frictional effects, Besselink et al. [17] propose a feedback control strategy using drag bits to overcome torsional stick–slip oscillations in vertical drilling systems. Using the same model with [17], a robust output-feedback control strategy based on skewed-μ DK-iteration is presented to eliminate stick–slip oscillations [18].

The interval type-2 fuzzy logic controllers (IT2FLC) and reinforcement learning (RL) are two main control strategies applied in our RSS controller design. IT2FLC with inference system and knowledge library is utilized to deal with the problem of trajectory tracking control for RSS in this paper. Fuzzy logic including type-1 and type-2 has been widely used in control systems [19–24], intelligent computation [25–27] and other information fields [28–29].

Compared with type-1 fuzzy logic controller (T1FLC), IT2FLC has advantages in dealing with nonlinearities, uncertainties and disturbances [30–33]. The downhole drilling work can be easily influenced by many nonlinearity and uncertainty elements, such as lithology, temperature, pressure, mechanical structure of BHA, rate of penetration (ROP) and others [34,35]. Therefore, these intense uncertainties in RSS and the downhole circumstances should be given consideration in controller design and may be solved by using IT2FLC to a certain extent. In IT2FLC, fuzzy sets consist of three-dimensional fuzzy membership functions (MFs) with footprint of uncertainty (FOU). The fuzzy inference outputs are not crisp values but MFs of type-1 fuzzy sets. Thus, in order to obtain crisp outputs as control input, IT2FLC needs to reduce the type-2 fuzzy outputs to type-1 fuzzy outputs, and further obtain crisp outputs by means of defuzzification. The increase of dimensions of fuzzy rules and MFs enhances abilities of fuzzy systems in dealing with complex uncertainty and nonlinear mapping, which is expected to decrease input nonlinearities, uncertainties and distance delays in RSS.

Reinforcement learning based on Markov decision processes (MDP) has been applied in robot control and decision [36,37], industrial process control [38,39], complex system programming and control [40–43], etc. In trajectory tracking control, including but not limited to directional drilling, the actor-critic algorithms in RL have been applied in UAV flight trajectory control [44], navigation behavior of mobile robots [45], AUVs trajectory tracking [46], etc. Differing from actor-only and critic-only algorithms, actor-critic algorithms are of importance on being able to obtain optimal control policies via low-variance gradient estimations and usually have better convergence properties [47]. Actor which can be regarded as action function approximator is utilized to search and output optimal control signal, and critic which can be regarded as value function approximator is designed to evaluate control performance of actor according to system states and reinforcement signals. Adaptive dynamic programming (ADP) is an important adaptive critic design method which can be applied to avoid curse of dimensionality in parameters storage and calculation, and establish online gradients update structure to approximate optimal Bellman equation. Application of actor-critic controller with ADP in our RSS trajectory tracking controller is to approximate reference trajectory for smooth tracking and stick–slip elimination.

This paper mainly focuses on the three-dimensional trajectory tracking controller design for a static push-the-bit RSS described in [12]. A model-based dual-loop feedback cooperative control architecture of RSS for the trajectory tracking consisting of IT2FLC and online actor-critic RL controller with one-order digital low pass filter (LPF) is proposed. In order to control the trajectories of inclination and azimuth of drill bit simultaneously, the directional force which is used to drive drill bit is decomposed to two force components to control inclination and azimuth respectively. The each control output signal component is a linear superposition of IT2FLC and actor-critic RL controller. The total control output is the vector sum of the two components.

The main contributions of this paper are described as follow:

(1) A model-based dual-loop feedback cooperative control method for RSS to realize accurate and smooth downhole three-dimensional trajectory tracking are presented. In this control method, a dual-loop feedback controller which is constructed by IT2FLC and actor-critic RL algorithm respectively is designed. Meanwhile, for the sake of optimizing the drilling trajectory further, a one-order low-pass filter is applied to reduce high-frequency noises in control output which is generated by cooperative controller.

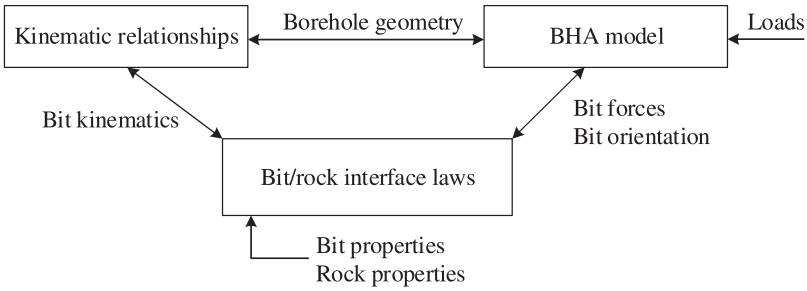

Fig. 1. Three components of RSS and their relationships [12].

(2) In order to realize accurate trajectory tracking control, an online iterative actor-critic RL control algorithm based on radial basis function neural networks (RBFNN) and ADP is proposed. The actor-critic control framework consists of critic network which is utilized to approximate value/cost function and actor network which is utilized to approximate action function.
(3) Considering actual drilling application environment, stable angle building hole rate validation and complex curve trajectory test are design in simulation studies. The practicability and effectiveness of proposed control method in actual application are validated by these desired trajectories.

This paper is organized as follows. Section 2 introduces mathematical model of RSS. Section 3 proposes the overall control architecture and introduces one-order digital low pass filter used in this paper. Section 4 discusses IT2FLC designed in this paper. In section 5, actor-critic controller which applied in RSS is presented. In Section 6, some trajectory tracking simulations are given and the performances of the controller proposed in this paper is verified by comparisons. Finally, several conclusions and expectations are discussed in Section 7.

2. Rotary steerable system model

Considering the directional propagation of the borehole in three-dimensional space, we decompose the drilling direction which are tangent to the borehole trajectory into inclination and azimuth. The inclination points to the direction of gravity, and the azimuth is orthogonal to the inclination and parallel to the horizontal plane. The RSS model developed by Perneder and Detournay [12–16] is static push-the-bit RSS system and consists of three components: (1) BHA model is simplified to Euler–Bernoulli beam and describes relationships between the penetration of drill bit into the rock and forces exerted on the bit; (2) bit/rock interface laws reflect the rock fragmentation and other dissipative processes between the cutter and the rock while drilling; (3) kinematics express the relationships of axial and lateral motions. Relationships of these three components are shown in Fig. 1 and details can be obtained in [3,13,48].

The geometry and its diagrammatic sketch of the RSS are shown in Fig. 2. $[\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z]^T$ in three-dimensional space represents the base coordinates of borehole trajectory. $\boldsymbol{e}_z$ points to the direction of gravity and is used to describe inclination. $\boldsymbol{e}_x$ is perpendicular to $\boldsymbol{e}_z$ (i.e. the direction with respect to azimuth). $\boldsymbol{e}_y$ is orthogonal to both $\boldsymbol{e}_x$ and $\boldsymbol{e}_z$. λ_i, $(i = 2, \ldots, n)$ is defined as the segment length of BHA between ith and $(i + 1)$th stabilizer, and λ_1

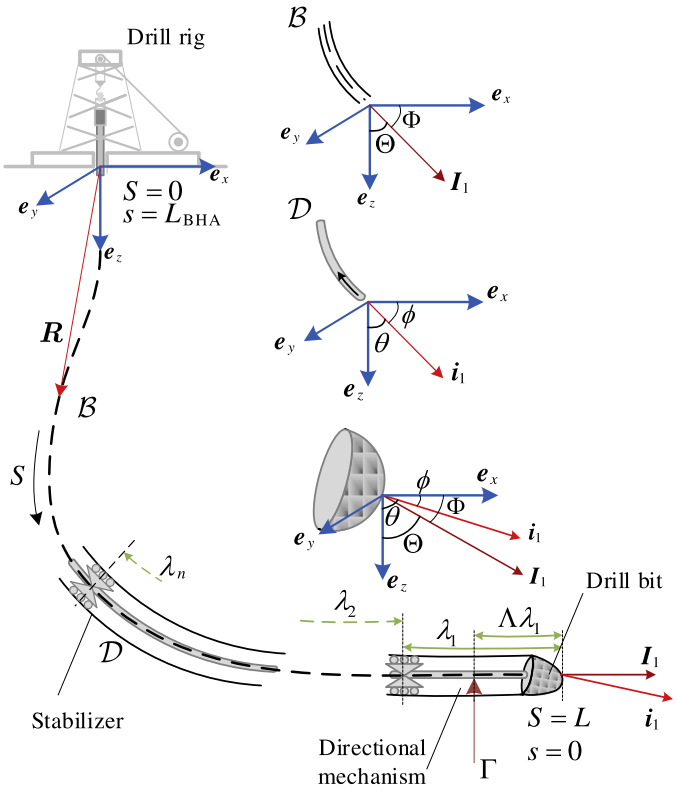

Fig. 2. Geometry of RSS [13].

is the length from the bit to the first stabilizer. $\Lambda\lambda_1$, $\Lambda\epsilon(0,1)$ is the distance between the bit and the directional mechanism where the directional force is exerted by the wellbore. $\mathcal{B}$ is the borehole axis defined by a set of reference points on the bit, and $\mathcal{D}$ is the BHA axis which may deviate from the borehole axis $\mathcal{B}$ because of bit rotation. In most cases the deviation between $\mathcal{B}$ and $\mathcal{D}$ can be regarded as zero for simplification. The borehole trajectory is described by its inclination $\Theta(S)$ and azimuth $\Phi(S)$ ($S\in$ [0, L] is a curvilinear coordinate of actual borehole trajectory). L is the length of drilling in the borehole and expressed by bit inclination $\theta(L,s)$ and bit azimuth $\phi(L,s)$ ($s\in$ [0, L_{BHA}] is the curvilinear coordinate of drill bit while drilling). L_{BHA} is the length of BHA measured by position sensor. A dimensionless drilling distance $\xi = L/\lambda_1$ as an independent variable is used to formulate the dynamics of borehole propagation according to the rate independent of bit/rock interface laws [16]. When ξ is applied to replace L, the bit inclination and azimuth can be described as $\theta(\xi,s)=\theta(\xi\lambda_1,s)$ and $\phi(\xi,s)=\phi(\xi\lambda_1,s)$, respectively.

According to [12], the RSS can be modeled as an Euler–Bernoulli beams with a constant bending stiffness EI. Define a scaled dimensionless characteristic force as $F_*=3\mathrm{EI}/\lambda 21$, then the directional forces and moments can be transformed into dimensionless parameters that depend on ξ and F_*.

The three-dimensional borehole trajectory evolution equations with zero walk angle has been expressed in [12]. Considering that the gravity load term has little influence on drilling

direction, we ignore this term and transform the evolution equations into nonlinear differential equation with multiple constant delay terms as follows:

$$\begin{cases} \chi\Pi\Theta' = -\mathcal{M}_b(\Theta - \langle\Theta\rangle_1) + \frac{\chi}{\eta}\mathcal{F}_b(\Theta - \Theta_1) + \sum_{i=1}^{n-1}\left(\frac{\mathcal{F}_b\mathcal{M}_i - \mathcal{F}_i\mathcal{M}_b - \mathcal{M}_i\eta\Pi}{\eta\Pi}\right)\left(\langle\Theta\rangle_i - \langle\Theta\rangle_{i+1}\right) \\ -\frac{\chi}{\eta}\sum_{i=1}^{n-1}\mathcal{F}_i\left(\frac{\Theta_{i-1}-\Theta_i}{\kappa_i} - \frac{\Theta_i-\Theta_{i+1}}{\kappa_{i+1}}\right) + \frac{\mathcal{F}_b\mathcal{M}_r - \mathcal{F}_r\mathcal{M}_b - \mathcal{M}_r\eta\Pi}{\eta\Pi}\Gamma_2 - \frac{\chi}{\eta}\Gamma_2', \\ \chi\Pi\Phi' = -\mathcal{M}_b(\Phi - \langle\Phi\rangle_1) + \frac{\chi}{\eta}\mathcal{F}_b(\Phi - \Phi_1) + \sum_{i=1}^{n-1}\left(\frac{\mathcal{F}_b\mathcal{M}_i - \mathcal{F}_i\mathcal{M}_b - \mathcal{M}_i\eta\Pi}{\eta\Pi}\right)\left(\langle\Phi\rangle_i - \langle\Phi\rangle_{i+1}\right) \\ -\frac{\chi}{\eta}\sum_{i=1}^{n-1}\mathcal{F}_i\left(\frac{\Phi_{i-1}-\Phi_i}{\kappa_i} - \frac{\Phi_i-\Phi_{i+1}}{\kappa_{i+1}}\right) + \left(\frac{\mathcal{F}_b\mathcal{M}_r - \mathcal{F}_r\mathcal{M}_b - \mathcal{M}_r\eta\Pi}{\eta\Pi\sin\Theta} + \frac{\chi}{\eta}\frac{\cos\Theta}{\sin^2\Theta}(\Theta - \Theta_1)\right)\Gamma_3 \\ -\frac{\chi}{\eta}\frac{\mathcal{F}_r}{\sin\Theta}\Gamma_3'. \end{cases} \tag{1}$$

In (Eq. 1), η and χ stand for the lateral and angular steering resistance, respectively. Γ_2 and Γ_3 stand for the RSS forces along inclination and azimuth directions, respectively. Constant Π represents the active weight-on-bit. The terms $\mathcal{F}_b$ and $\mathcal{M}_b$ stand for the relative orientation of the bit with respect to the chord, which links the bit and the first stabilizer. $\mathcal{F}_r$ and $\mathcal{M}_r$ represent the dependence of forces and moments at the bit on the RSS force. $\mathcal{F}_i$ and $\mathcal{M}_i$ with subscript i represent the constraints imposed by the geometry of the borehole. These coefficient expressions above mentioned for BHA can be seen in [13]. $\Theta_i = \Theta(\xi_i)$ and $\Phi_i = \Phi(\xi_i)$ for $i = 2,\ldots, n$ represent the delayed location of the ith stabilizer of the inclination and azimuth. $\langle\Theta\rangle_1$ and $\langle\Phi\rangle_1$ stand for the average inclination and azimuth between drill bit and the first stabilizer. $\langle\Theta\rangle_i$ and $\langle\Phi\rangle_i$ for $i = 2,\ldots, n$ represent the average inclination and azimuth, and can be expressed as follow, respectively:

$$\langle\Theta\rangle_i = \frac{1}{\kappa_i}\int_{\xi_{i-1}}^{\xi_i}\Theta(\sigma)d\sigma \tag{2}$$

$$\langle\Phi\rangle_i = \frac{1}{\kappa_i}\int_{\xi_{i-1}}^{\xi_i}\Phi(\sigma)d\sigma \tag{3}$$

where the $\xi_i = \xi - \sum_{j=1}^{i}\kappa_j$ for $i = 1, 2,\ldots, n$ and $j = 1, 2,\ldots, i$. Coefficient $\kappa_i = \lambda_i/\lambda_1$ is the dimensionless length of the ith BHA segment.

2.1. Continuous state space model

According to the properties of (Eq. 1), the RSS model can be rewritten as first order continuous state space equation as follows:

$$\boldsymbol{x}'(\xi) = \boldsymbol{A}_0\boldsymbol{x}(\xi) + \sum_{i=1}^{n}\boldsymbol{A}_i\boldsymbol{x}(\xi_i) + \boldsymbol{B}_0(\xi)\boldsymbol{\Gamma}(\xi) + \boldsymbol{B}_1(\xi)\boldsymbol{\Gamma}'(\xi) \tag{4}$$

where $\boldsymbol{x} = [\boldsymbol{x}_\Theta, \boldsymbol{x}_\Phi]^{\mathrm{T}} = [\Theta, \langle\Theta\rangle_1, \ldots, \langle\Theta\rangle_n, \Phi, \langle\Phi\rangle_1, \ldots, \langle\Phi\rangle_n]^{\mathrm{T}}$ is the system state related to dimensionless distance ξ. $\boldsymbol{A}_i$, $i = 1$ and $i = 2,\ldots, n$ are system matrices representing Euler–Bernoulli beams between the bit and directional mechanism, and drill strings between two adjacent stabilizers, respectively. $\boldsymbol{B}_0$ and $\boldsymbol{B}_1$ are distance-varying input matrices depending on inclination $\Theta(\xi)$. Scaled force $\boldsymbol{\Gamma} = [\Gamma_2, \Gamma_3]^{\mathrm{T}}$ represents the decomposed directional forces of inclination and azimuth, respectively.

In this paper, we just consider the state space equation of RSS with two stabilizers, which is reasonable confirmed by Kremers et al. [3]. So, (Eq. 4) could be simplified with system matrices A_0, A_1, A_2 (see Appendix as:

$$\boldsymbol{x}'(\xi) = \boldsymbol{A}_0\boldsymbol{x}(\xi) + \sum_{i=1}^{2} \boldsymbol{A}_i\boldsymbol{x}(\xi_i) + \boldsymbol{B}_0(\xi)\boldsymbol{\Gamma}(\xi) + \boldsymbol{B}_1(\xi)\boldsymbol{\Gamma}'(\xi) \tag{5}$$

where $\xi_1 = \xi - \kappa_1$ and $\xi_1 = \xi - \kappa_1 - \kappa_2$.

It is assumed that the inclination and azimuth sensors are installed between the drill bit and directional mechanism [7,8]. Then the measured output $\boldsymbol{y} = [\theta_0(\xi), \phi_0(\xi)]^{\mathrm{T}}$ is influenced by the coefficients $\eta\Pi$, $\mathrm{F_b}$, $\mathrm{F_r}$, F_i and can be expressed as follow:

$$\boldsymbol{y} = \frac{1}{\eta\Pi - \mathcal{F}_b}\begin{bmatrix}\eta\Pi\Theta - \mathcal{F}_b\langle\Theta\rangle_1 + \mathcal{F}_r\Gamma_2 + \mathcal{F}_1(\langle\Theta\rangle_1 - \langle\Theta\rangle_2)\\ \eta\Pi\Phi - \mathcal{F}_b\langle\Phi\rangle_1 + \mathcal{F}_r\Gamma_3 + \mathcal{F}_1(\langle\Phi\rangle_1 - \langle\Phi\rangle_2)\end{bmatrix}. \tag{6}$$

Rewrite (Eq. 6) into output equation:

$$\boldsymbol{y} = \boldsymbol{C}x(\xi) + \boldsymbol{D}\boldsymbol{\Gamma}(\xi) \tag{7}$$

where $\boldsymbol{C}$ is output matrix and $\boldsymbol{D}$ is direct control matrix with respect to the configuration of BHA. The forms of $\boldsymbol{C}$ and $\boldsymbol{D}$ can be seen in Appendix A.

2.2. Discrete state space model

In order to realize online reinforcement learning and iterative computation, the continuous RSS model is converted into discretization form. Set the dimensionless sampling distance of RSS is S_d, and set $p_1 = \kappa_1/S_\mathrm{d}$, $L_1 = \lfloor\kappa_1/S_\mathrm{d}\rfloor$, $d_1 = p_1 - L_1$, $p_2 = (\kappa_1 + \kappa_2)/S_\mathrm{d}$, $L_2 = \lfloor(\kappa_1 + \kappa_2)/S_\mathrm{d}\rfloor$, $d_2 = p_2 - L_2$, $0 \le d_1, d_2 < 1$ ($\lfloor\rfloor$ represents fractions are rounded down). The borehole trajectories of sampling points are approximated by using zero order holder. Discrete state space equation and output equation of BHA is expressed as follows

$$\begin{cases}\boldsymbol{x}(k+1) = \boldsymbol{G}_0\boldsymbol{x}(k) + \boldsymbol{G}_{11}\boldsymbol{x}(k-L_1) + \boldsymbol{G}_{12}\boldsymbol{x}(k-1-L_1) + \boldsymbol{G}_{21}\boldsymbol{x}(k-L_2) + \boldsymbol{G}_{22}\boldsymbol{x}(k-1-L_2)\\ +\boldsymbol{H}_0(k)\boldsymbol{\Gamma}(k) + \boldsymbol{H}_1(k)\boldsymbol{\Gamma}'(k)\\ = \boldsymbol{G}_0\boldsymbol{x}(k) + \boldsymbol{G}_{11}\boldsymbol{x}(k-L_1) + \boldsymbol{G}_{12}\boldsymbol{x}(k-1-L_1) + \boldsymbol{G}_{21}\boldsymbol{x}(k-L_2) + \boldsymbol{G}_{22}\boldsymbol{x}(k-1-L_2)\\ +\boldsymbol{H}(k)\boldsymbol{u}(k)\\ \boldsymbol{y}(k) = \boldsymbol{C}\boldsymbol{x}(k) + \boldsymbol{D}\boldsymbol{\Gamma}(k)\end{cases} \tag{8}$$

where the $\boldsymbol{G}_0$, $\boldsymbol{G}_{11}$, $\boldsymbol{G}_{12}$, $\boldsymbol{G}_{21}$, $\boldsymbol{G}_{22}$, $\boldsymbol{H}_0(k)$, $\boldsymbol{H}_1(k)$ in (Eq. 8) can be seen in Appendix B, and $\boldsymbol{H}(k)\boldsymbol{u}(k) = \boldsymbol{H}_1(k)\boldsymbol{\Gamma}(k) + \boldsymbol{H}_2(k)\ \boldsymbol{\Gamma}'(k)$.

3. Trajectory tracking control architecture and one-order low pass filter

3.1. Trajectory tracking control architecture

The trajectory tracking control architecture of RSS is shown in Fig. 3 which is a model-based dual-loop control system composed by an interval type-2 fuzzy logic controller and an actor-critic RL controller. Simultaneous adoption of these controllers can guarantee tracking accuracy and smoothness, while eliminate slip–stick oscillations as well. Considering that general type-1 fuzzy logic controller is inadequate in dealing with violent nonlinearity and uncertainty, IT2FLC is designed based on system errors and their derivatives. From the controller characteristics, IT2FLC is similar to nonlinear proportional differential (PD) controller

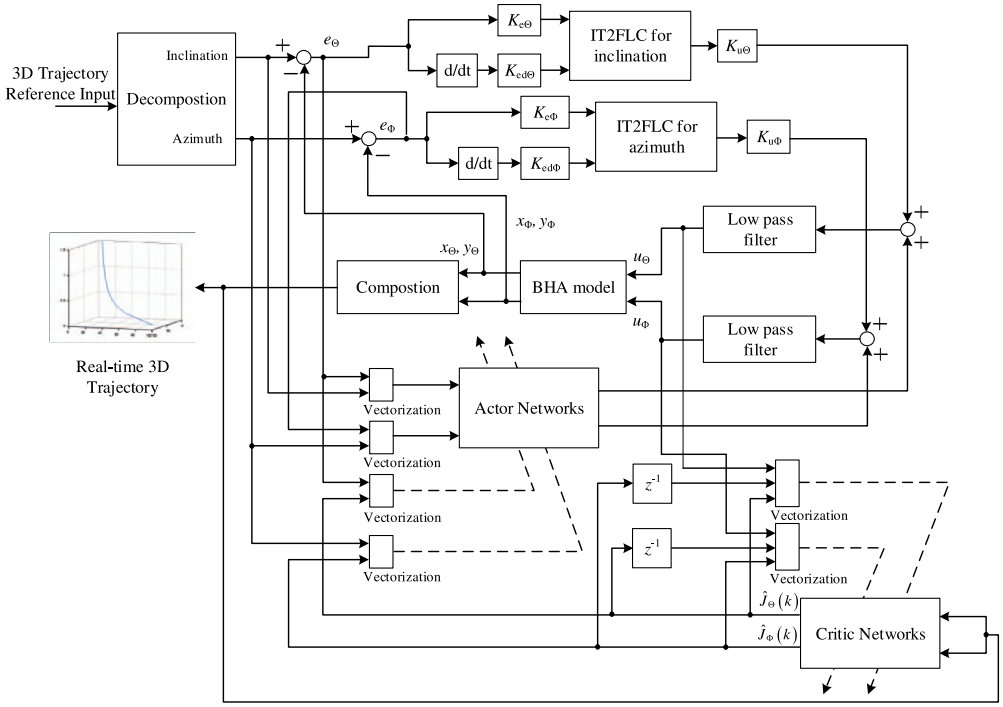

Fig. 3. Three-dimensional trajectory tracking control architecture of RSS.
The vectorization blocks mean that multiple vectors constitute a new vector, and the dimension of the new vector is sum of all the vectors in the instant k.

and it displays stronger robustness [49] which can be exploited to deal with more difficult uncertainties and nonlinearities in control systems. In addition, applications of IT2FLC for actual trajectory can narrow the search space of actor-critic controller in another one closed-loop to guarantee system convergence and reduce the time for online learning. These advantages are of great significance for practical drilling engineering applications. The actor-critic controller consists of two groups of actor and critic networks constituted by RBFNN related to inclination and azimuth respectively. The critic networks receive system states and evaluate system performances in the form of value functions. The actor networks provide control input to RSS model in the form of action functions according to the reference trajectory, system states and value functions. Networks parameters of actor and critic are updated based on the value functions, control effects and other indices which reflect system performances. Actor-critic controller in overall control architecture makes the drilling trajectories more accurate and smooth.

In order to eliminate further stick–slip oscillations and high-frequency noises in inclination and azimuth, and reduce frequent changes in azimuth while drilling in the initial state caused by inclination, the first-order digital LPF are applied to the control system to alleviate these problems.

3.2. One-order low pass filter

For the sake of reducing stick–slip oscillations further in RSS trajectory tracking control, LPF for inclination and azimuth respectively are designed as follows:

$$u_{\Theta}(k+1) = (1-\lambda_{\Theta})u_{\Theta}(k) + \lambda_{\Theta}\left(u_{a\Theta}(k) + u_{f\Theta}(k)\right) \tag{9}$$

$$u_{\Phi}(k+1) = (1-\lambda_{\Phi})u_{\Phi}(k) + \lambda_{\Phi}\left(u_{a\Phi}(k) + u_{f\Phi}(k)\right) \tag{10}$$

where $u_{a\Theta}$, $u_{a\Phi}$ and $u_{f\Theta}$, $u_{f\Phi}$ are control outputs of IT2FLC and actor-critic controller respectively. u_{Θ} and u_{Φ} are cooperative control outputs which are filtered by LPFs for inclination and azimuth respectively. $0 < \lambda_{\Theta}, \lambda_{\Phi} < 1$ represent coefficients of LPFs.

LPFs are used to further reduce high-frequency stick–slip oscillations, which makes drilling trajectory smoother. It is of significance on enhancing working life of directional mechanism.

4. Interval type-2 fuzzy logic controller

The interval type-2 fuzzy logic controller in RSS control architecture is utilized to track reference borehole trajectories while drilling. The antecedents and consequents of fuzzy inference engines are interval type-2 fuzzy sets and crisp numbers respectively (A2-C0) [50]. For a fuzzy controller with two input variables and an output variable, the n-th rule in rulebase R of type-2 fuzzy sets can be described as follows:

$$\tilde{R}^n \rightarrow IF\, x_1\, is\, \tilde{X}_1^n\, and\, x_2\, is\, \tilde{X}_2^n,\ THEN\, y\, is\, Y^n \tag{11}$$

where x_1, x_2 are the input variables of fuzzy antecedents. y is the output variable of a consequent.$\tilde{X}_i^n$ $(i = 1, 2)$ are type-2 fuzzy sets of antecedents. $Y^n = [\underline{y}^n, \bar{y}^n]$ is an interval of fuzzy set. The FOU of $\tilde{X}_i^n$ and Y^n which are formed by embedded type-1 fuzzy sets are bounded by upper and lower MFs. The boundaries of lower and upper MFs fuzzy MFs are represented as type-2 MFs, i.e. $\mu_{\tilde{X}^n} = [\underline{f}^n, \bar{f}^n]$. Similarly, $\underline{y}^n$ and $\bar{y}^n$ are lower and upper MF of Y^n, respectively.

The input variables of IT2FLC are set as the measured system output error $e = y - y_{\mathrm{r}}$ and its derivative $\dot{e} = \dot{y} - \dot{y}_r$, where y_{r} is the reference borehole trajectories of inclination and azimuth. In most cases, in order to make the error and error derivative ranges match with fuzzy logic fields, they need to be scaled by controller gains $K_{\mathrm{e}\Theta}$ and $K_{\mathrm{ed}\Theta}$ for inclination, and $K_{\mathrm{e}\Phi}$ and $K_{\mathrm{ed}\Phi}$ for azimuth, respectively. Similarly, output gains $K_{\mathrm{u}\Theta}$ and $K_{\mathrm{u}\Phi}$ are used for amplifying the crisp outputs of IT2FLC.

The error and error derivatives are fuzzified by seven triangle type-2 MFs into fuzzy sets denoted by seven linguistic variables shown in Fig. 4. The ranges of all the linguistic variables are determined from –6 to +6 based on design experience, and FOUs bounded by different upper and lower MFs are filled in different colors, respectively.

The fuzzy inference engine is driven by Takagi–Sugeno–Kang (TSK) fuzzy inference model in this paper which has 49 rules determined by the number of MFs and input variables (shown in the Table 1). Most of the fuzzy rules are obtained by prior knowledge and design experience, and part of them are adjusted based on controller performance. Fuzzy inference formed by the IF-THEN rules shown in (Eq. 11) is utilized to construct a nonlinear mapping from fuzzy inputs to outputs. Before defuzzification, the outputs of inference needs to be reduced from type-2 fuzzy sets to type-1 fuzzy sets, namely type reduction. Karnik–Mendel

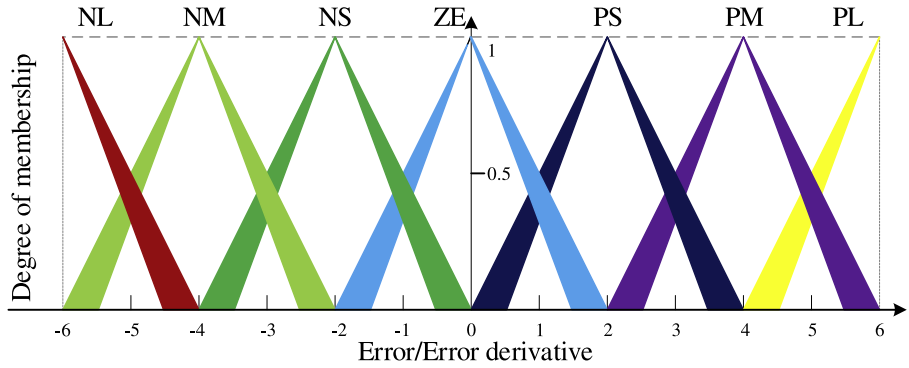

Fig. 4. MFs of type-2 fuzzy logic controller.
The linguistic variables of fuzzy sets are defined as follow: NL-negative large, NM-negative medium, NS-negative small, ZE-zero, PS-positive small, PM-positive medium, PL-positive large.

Table 1
Interval fuzzy rules of control architecture.

$e_d \backslash e$	NL	NM	NS	ZE	PS	PM	PL
NL	NL	NL	NL	NL	NM	NS	ZE
NM	NL	NL	NL	NM	NS	ZE	PS
NS	NL	NL	NM	NS	ZE	PS	PM
ZE	NL	NM	NS	ZE	PS	PM	PL
PS	NM	NS	ZE	PS	PM	PL	PL
PM	NS	ZE	PS	PM	PL	PL	PL
PL	ZE	PS	PM	PL	PL	PL	PL

(KM) algorithm [30] is commonly utilized for type reduction which has an advantage of lowering the computing cost. The basic form of KM algorithms can be represented as follow [30]:

$$y_{lt} = \min_{L\in[1,N-1]} \frac{\sum_{n=1}^{L} \underline{y}^n \bar{f}^n + \sum_{n=L+1}^{N} \underline{y}^n \underline{f}^n}{\sum_{n=1}^{L} \bar{f}^n + \sum_{n=L+1}^{N} \underline{f}^n} \tag{12}$$

$$y_{rt} = \max_{R\in[1,N-1]} \frac{\sum_{n=1}^{R} \bar{y}^n \underline{f}^n + \sum_{n=R+1}^{N} \bar{y}^n \bar{f}^n}{\sum_{n=1}^{R} \underline{f}^n + \sum_{n=R+1}^{N} \bar{f}^n} \tag{13}$$

where y_{lt} and y_{rt} are left and right end points of centroid interval, respectively. In order to reduce computing time further in real-time drilling fields, a type reduction method named enhanced opposite direction search (EODS) algorithm based on KM [51] is utilized in our IT2FLC, which has faster computing speed than conventional KM algorithm [52]. Solutions of y_{lt} and y_{rt} are solved by searching the positive process and negative process simultaneously via iterative computation until the two processes meet based on the fact that $\underline{y}^L \le y_{lt} \le \underline{y}^{L+1}$ and $\bar{y}^R \le y_{rt} \le \bar{y}^{R+1}$. Accordingly, the crisp output can be expressed as follows:

$$y = \frac{y_{lt} + y_{rt}}{2}. \tag{14}$$

The IT2FLC used in RSS trajectory tracking control can make the drill bit track the reference borehole trajectory roughly without the requirement of complex formula derivation

in control theory. In addition, it narrows search space of optimal control parameters and reduce time of online iterative learning for the following actor-critic networks, which is embedded in our control architecture to further improve and optimize tracking accuracy and stick–slip elimination.

5. Online actor-critic learning controller

The actor-critic algorithm in RL consists of actor and critic which are used to obtain action function and value/cost function, respectively. In this paper, we use the actor-critic framework with adaptive dynamic programming to construct an online reinforcement learning controller to realize more accurate trajectory tracking for RSS model which is established by nonlinear differential equation with multiple constant delay terms. Correlated system stability analysis of online RL algorithms indicates that the closed system tends to uniformly ultimate boundedness (UUB) while the persistent excitation (PE) condition requirement is satisfied [41,42].

5.1. Critic network design

In the actor-critic controller, the critic network is used to approximate the long-term cost function J. Before the system state $\boldsymbol{x}(k)$ transits to new state $\boldsymbol{x}(k+1)$, the control outputs from actor network and system outputs will generate an immediate reward or a utility function $r(k)$ to evaluate the system performance under the current step action, where $r(k)$ is defined as follows:

$$r(k) = (\boldsymbol{y}(k) - \boldsymbol{y}_r(k))^T \boldsymbol{Q}(\boldsymbol{y}(k) - \boldsymbol{y}_r(k)) + \boldsymbol{u}_a(k)^T \boldsymbol{R}\boldsymbol{u}_a(k) \tag{15}$$

where $\boldsymbol{y}_{\mathrm{r}}(k)$ is the desired system trajectory, and $\boldsymbol{u}_{\mathrm{a}}(k)$ is the control output of actor network in actor-critic controller. $\boldsymbol{Q}$ and $\boldsymbol{R}$ are positive semi-definite and positive definite matrices with appropriate dimensions, respectively.

The cost to go function in the Bellman equation $J(k)$ which is defined as follows [53]:

$$J(k) = \sum_{t=0}^{\infty} \gamma^t r(k+t) \tag{16}$$

where γ is a discount factor for the finite horizon problems ($0 < \gamma < 1$). So from (Eq. 16) it can be seen that the cost to go function is the accumulation of discount immediate reward function. In this paper, considering that the RBFNN with a clustering of nonlinear kernels can approximate any functions to an arbitrary degree of accuracy, it is used to construct the critic network to approximate the desired cost to go function for inclination and azimuth, respectively

$$\hat{J}_{\Theta}(k) = \hat{\omega}_{c\Theta}^T(k) h_c(\boldsymbol{x}_{\Theta}(k)), h_c(\boldsymbol{x}_{\Theta}(k)) \in \Re^{m_1 \times 1} \tag{17}$$

$$\hat{J}_{\Phi}(k) = \hat{\omega}_{c\Phi}^T(k) h_c(\boldsymbol{x}_{\Phi}(k)), h_c(\boldsymbol{x}_{\Phi}(k)) \in \Re^{m_1 \times 1} \tag{18}$$

where $\hat{\omega}_{c\Theta}(k) \in \Re^{m_1 \times 1}$ and $\hat{\omega}_{c\Phi}(k) \in \Re^{m_1 \times 1}$ ($m_1 = 3$ is the dimensions of input vectors $\boldsymbol{x}_{\Theta}$ and $\boldsymbol{x}_{\Phi}$) are the actual NN weights which are the estimations of optimal network parameters $W\mathrm{c}\Theta$ and $W\mathrm{c}\Phi$, respectively. $h_{\mathrm{c}}(\boldsymbol{x}_{\Theta}(k))$ and $h_{\mathrm{c}}(\boldsymbol{x}_{\Phi}(k))$ are outputs of NN hidden layers and represented by Gaussian basis functions. For convenient to estimate the inclination and

azimuth which RSS drills, we use two RBF networks to evaluate system performances of corresponding parts of BHA model respectively and approximate optimal cost to go functions. Accordingly, the optimal cost to go functions $J^*_{c\Theta}(k)$ and $J^*_{c\Phi}(k)$ are constructed with approximation errors $\varepsilon_c(\boldsymbol{x}_\Theta(k))$ and $\varepsilon_c(\boldsymbol{x}_\Phi(k))$ according to Weierstrass higher order approximation theorem [54] as follows:

$$J^*_{c\Theta}(k) = W^T_{c\Theta} h_c(\boldsymbol{x}_\Theta(k)) + \varepsilon_c(\boldsymbol{x}_\Theta(k)) \tag{19}$$

$$J^*_{c\Phi}(k) = W^T_{c\Phi} h_c(\boldsymbol{x}_\Phi(k)) + \varepsilon_c(\boldsymbol{x}_\Phi(k)). \tag{20}$$

According to adaptive critic designs theory, the prediction error of critic networks for inclination and azimuth are defined by cost to go functions and reward functions [53]

$$e_{c\Theta}(k) = \gamma \hat{J}_\Theta(k) - \hat{J}_\Theta(k-1) + r_\Theta(k) \tag{21}$$

$$e_{c\Phi}(k) = \gamma \hat{J}_\Phi(k) - \hat{J}_\Phi(k-1) + r_\Phi(k). \tag{22}$$

The quadratic objective functions to optimize critic networks are defined as

$$E_{c\Theta}(k) = \frac{1}{2} e^T_{c\Theta}(k) e_{c\Theta}(k) \tag{23}$$

$$E_{c\Phi}(k) = \frac{1}{2} e^T_{c\Phi}(k) e_{c\Phi}(k). \tag{24}$$

In order to obtain optimal critic NN weights, the standard gradient descent algorithms supervised by back propagation is used to update NN weights

$$\hat{\omega}_{c\Theta}(k+1) = \hat{\omega}_{c\Theta}(k) + \Delta\hat{\omega}_{c\Theta}(k) \tag{25}$$

$$\hat{\omega}_{c\Phi}(k+1) = \hat{\omega}_{c\Phi}(k) + \Delta\hat{\omega}_{c\Phi}(k) \tag{26}$$

where $\Delta\hat{\omega}_{c\Theta}(k)$ and $\Delta\hat{\omega}_{c\Phi}(k)$ are the weight correction terms which are given as follow:

$$\Delta\hat{\omega}_{c\Theta}(k) = -\alpha_{c\Theta} \frac{\partial E_{c\Theta}(k)}{\partial \hat{\omega}_{c\Theta}(k)} = -\alpha_{c\Theta} \gamma h_{c\Theta}(x_\Theta(k)) \left(\gamma \hat{J}_\Theta(k) - \hat{J}_\Theta(k-1) + r_\Theta(k) \right) \tag{27}$$

$$\Delta\hat{\omega}_{c\Phi}(k) = -\alpha_{c\Phi} \frac{\partial E_{c\Phi}(k)}{\partial \hat{\omega}_{c\Phi}(k)} = -\alpha_{c\Phi} \gamma h_{c\Phi}(x_\Phi(k)) \left(\gamma \hat{J}_\Phi(k) - \hat{J}_\Phi(k-1) + r_\Phi(k) \right) \tag{28}$$

where $\alpha_{c\Theta}$ and $\alpha_{c\Phi}$ are adaption NN gains. It is indicated that the weights of critic networks are updated online according to system states, reward signal, discount factor, cost to go function and its past value, etc.

5.2. Actor network design

First, define tracking error in current distance instant k as $\boldsymbol{e}(k) = \boldsymbol{x}(k) - \boldsymbol{x}_r(k)$. Then, the tracking error in distance instant $k+1$ is

$$\begin{aligned} \boldsymbol{e}(k+1) &= \boldsymbol{x}(k+1) - \boldsymbol{x}_r(k+1) \\ &= \boldsymbol{G}_0 \boldsymbol{e}(k) + \boldsymbol{G}_{11} \boldsymbol{e}(k - L_1) + \boldsymbol{G}_{12} \boldsymbol{e}(k - 1 - L_1) + \boldsymbol{G}_{21} \boldsymbol{e}(k - L_2) + \boldsymbol{G}_{22} \boldsymbol{e}(k - 1 - L_2) \\ &\quad + \boldsymbol{H}(k)(\boldsymbol{u}(k) - \boldsymbol{u}_r(k)) \end{aligned} \tag{29}$$

where $\boldsymbol{x}_{\mathrm{r}}(k)$ and its delay terms are desired system states. $\boldsymbol{u}_{\mathrm{r}}(k)$ is expected system control input.

Similar to the construction of the optimal cost to go function in critic networks, the desired control inputs are designed by using RBF networks as follow:

$$u_{r\Theta}(k) = W_{a\Theta}^{T} h_a(\boldsymbol{s}_{\Theta}(k)) + \varepsilon_a(\boldsymbol{s}_{\Theta}(k)) \tag{30}$$

$$u_{r\Phi}(k) = W_{a\Phi}^{T} h_a(\boldsymbol{s}_{\Phi}(k)) + \varepsilon_a(\boldsymbol{s}_{\Phi}(k)) \tag{31}$$

where $u_{\mathrm{r}\Theta}(k)$ and $u_{\mathrm{r}\Phi}(k)$ are desired control signals for inclination and azimuth, respectively, namely $\boldsymbol{u}_{\mathrm{r}}(k) = [u_{r\Theta}(k),\ u_{r\Phi}(k)]^{\mathrm{T}}$. $W_{\mathrm{a}\Theta}$ and $W_{\mathrm{a}\Phi}$ are optimal weights for $u_{\mathrm{r}\Theta}(k)$ and $u_{\mathrm{r}\Phi}(k)$. $\boldsymbol{s}_{\Theta}(k) = [\boldsymbol{x}_{\Theta}(k),\ y_{\Theta}(k),\ y_{r\Theta}(k),\ e_{\Theta}(k)]^{\mathrm{T}}$ for inclination and $\boldsymbol{s}_{\Phi}(k) = [\boldsymbol{x}_{\Phi}(k),\ y_{\Phi}(k),\ y_{r\Phi}(k),\ e_{\Phi}(k)]^{\mathrm{T}}$ for azimuth are actor networks input vectors, respectively. $y_{\Theta}(k)$ and $y_{\Phi}(k)$ are elements of system output vector $\boldsymbol{y}(k) = [y_{\Theta}(k),\ y_{\Phi}(k)]^{\mathrm{T}}$. Similarly, $y_{\mathrm{r}\Theta}(k)$, $y_{\mathrm{r}\Phi}(k)$ and $e_{\Theta}(k)$, $e_{\Phi}(k)$ are elements of reference input vector $\boldsymbol{y}_{\mathrm{r}}(k)$ and tracking error vector $\boldsymbol{e}(k)$. The actual actor networks outputs based on RBFNN to approximate desired control signals are designed as follow:

$$u_{\mathrm{a}\Theta}(k) = \hat{\omega}_{a\Theta}^{T}(k) h_a(\boldsymbol{s}_{\Theta}(k)),\ h_a(\boldsymbol{s}_{\Theta}(k)) \in \Re^{m_2 \times 1} \tag{32}$$

$$u_{a\Phi}(k) = \hat{\omega}_{a\Phi}^{T}(k) h_a(\boldsymbol{s}_{\Phi}(k)),\ h_a(\boldsymbol{s}_{\Phi}(k)) \in \Re^{m_2 \times 1} \tag{33}$$

where $\hat{\omega}_{a\Theta}(k) \in \Re^{m_2 \times 1}$ and $\hat{\omega}_{a\Phi}(k) \in \Re^{m_2 \times 1}$ (m_2 is the dimensions of input vectors $\boldsymbol{s}_{\Theta}$ and $\boldsymbol{s}_{\Phi}$) are the actual weights of actor networks. $u_{\mathrm{a}\Theta}(k)$ and $u_{\mathrm{a}\Phi}(k)$ are actual NN outputs which are used to control the downhole trajectory tracking of inclination and azimuth.

The approximate errors of the two networks in actor are defined

$$\delta_{a\Theta}(k) = u_{a\Theta}(k) - u_{r\Theta}(k) = \tilde{\omega}_{a\Theta}^{T}(k) h_a(\boldsymbol{s}_{\Theta}(k)) - \varepsilon_a(\boldsymbol{s}_{\Theta}(k)) \tag{34}$$

$$\delta_{a\Phi}(k) = u_{a\Phi}(k) - u_{r\Phi}(k) = \tilde{\omega}_{a\Phi}^{T}(k) h_a(\boldsymbol{s}_{\Phi}(k)) - \varepsilon_a(\boldsymbol{s}_{\Phi}(k)) \tag{35}$$

where $\tilde{\omega}_{a\Theta}(k) = \hat{\omega}_{a\Theta}(k) - W_{a\Theta}$ and $\tilde{\omega}_{a\Phi}(k) = \hat{\omega}_{a\Phi}(k) - W_{a\Phi}$.

Construct the networks error for actor networks based on cost to go function generated by critic networks and approximate errors in actor networks for inclination and azimuth, respectively:

$$e_{a\Theta}(k) = \sqrt{\|\boldsymbol{H}_{\Theta}\|}\,\delta_{a\Theta}(k) + \left(\sqrt{\|\boldsymbol{H}_{\Theta}\|}\right)^{-1}\left(\hat{J}_{\Theta}(k) - J_{c\Theta}^{*}(k)\right) \tag{36}$$

$$e_{a\Phi}(k) = \sqrt{\|\boldsymbol{H}_{\Phi}(k)\|}\,\delta_{a\Phi}(k) + \left(\sqrt{\|\boldsymbol{H}_{\Phi}(k)\|}\right)^{-1}\left(\hat{J}_{\Phi}(k) - J_{c\Phi}^{*}(k)\right) \tag{37}$$

where $\boldsymbol{H}_{\Theta}$ and $\boldsymbol{H}_{\Phi}(k)$ are vectors for inclination and azimuth in input matrix $\boldsymbol{H}(k)$ (The elements in input matrix of inclination are not changing with instant k).

Define the quadratic error functions to minimize the actor networks errors for control inputs

$$E_{a\Theta}(k) = \frac{1}{2} e_{a\Theta}^{T}(k) e_{a\Theta}(k) \tag{38}$$

$$E_{a\Phi}(k) = \frac{1}{2} e_{a\Phi}^{T}(k) e_{a\Phi}(k). \tag{39}$$

Similar to the update of critic NN weights, the optimal control signals generated by actor networks for inclination and azimuth are obtained by standard gradient descent

$$\hat{\omega}_{a\Theta}(k+1) = \hat{\omega}_{a\Theta}(k) + \Delta\hat{\omega}_{a\Theta}(k) \tag{40}$$

$$\hat{\omega}_{a\Phi}(k+1) = \hat{\omega}_{a\Phi}(k) + \Delta\hat{\omega}_{a\Phi}(k) \tag{41}$$

$\Delta\hat{\omega}_{a\Theta}(k)$ and $\Delta\hat{\omega}_{a\Phi}(k)$ can be expended by chain rules

$$\Delta\hat{\omega}_{a\Theta}(k) = -\alpha_{a\Theta}\frac{\partial E_{a\Theta}(k)}{\partial\hat{\omega}_{a\Theta}(k)} = -\alpha_{a\Theta}h_{\Theta}(s_{\Theta}(k))\Big(\|\boldsymbol{H}_{\Theta}\|\delta_{a\Theta}(k) + \hat{J}_{\Theta}(k) - J^{*}_{c\Theta}(k)\Big) \tag{42}$$

$$\Delta\hat{\omega}_{a\Phi}(k) = -\alpha_{a\Phi}\frac{\partial E_{a\Phi}(k)}{\partial\hat{\omega}_{a\Phi}(k)} = -\alpha_{a\Phi}h_{\Phi}(s_{\Phi}(k))\Big(\|\boldsymbol{H}_{\Phi}(k)\|\delta_{a\Phi}(k) + \hat{J}_{\Phi}(k) - J^{*}_{c\Phi}(k)\Big) \tag{43}$$

where $\alpha_{a\Theta}$ and $\alpha_{a\Phi}$ are the adaption NN gains. In (Eqs. 42) and (43), we use the system output errors $e_{\alpha\Theta}(k) = y_{\Theta}(k) - y_{r\Theta}(k)$ and $e_{\alpha\Phi}(k) = y_{\Phi}(k) - y_{r\Phi}(k)$ to replace approximate error terms $\|\boldsymbol{H}_{\Theta}(k)\|\delta_{a\Theta}(k)$ and $\|\boldsymbol{H}_{\Phi}(k)\|\delta_{a\Phi}(k)$ respectively in inclination and azimuth because $e_{\alpha\Theta}$ and $e_{\alpha\Phi}$ can reflect the convergence properties of approximate error terms, and desired control u_r is difficult to obtain due to the complex downhole environment in the actual drilling works.

5.3. Online learning framework for actor-critic controller

In every instant k of a complete trajectory tracking process, network parameters are adjusted via computing gradients and propagating backward iteratively until system converges to ideal states. It is an online real-time learning framework that uses reinforcement signal to realize system dynamic optimization. According to the critic networks and actor networks design above, the online reinforcement learning process can be regarded as policy evaluation and policy improvement respectively. Critic networks are utilized to evaluate system performances which are regulated by control policy that is fused by actor networks and IT2FLC, and actor networks are utilized to improve policy to realize accurate trajectory tracking effects. The key process of this framework is explained in details as follows.

6. Simulation studies

In this section, we use the control method proposed in this paper to realize the downhole trajectory tracking control. For the reason of process of project, we simulate the trajectory tracking effects by using MATLAB.

The chosen parameters of BHA model are shown in Table 2 [3,10].

6.1. Angle building hole simulation

Angle building hole rate is an important index to evaluate the directional performance of RSS. In order to realize the stationary control of angle building hole rate, we set relatively large values considering actual mechanical structures and construct linear trajectories for inclination and azimuth, respectively. Considering that the relationships between inclination and azimuth, the trajectories to test angle building hole rates are designed as follow:

Algorithm 1 Actor-critic controller with online learning.

1. Given system reference trajectories y_r. Pick initial state x_0, y_0, pick large positive number N as inner loop iterations, and small positive numbers β_Θ and β_Φ for inclination and azimuth respectively as the converge tolerances. Also, set discount factors γ_Θ, γ_Φ, adaption gains $\alpha_{c\Theta}$, $\alpha_{c\Phi}$, $\alpha_{a\Theta}$, $\alpha_{a\Phi}$, and choose appropriate kernel parameters $b_{c\Theta}$, $c_{c\Theta}$. $b_{c\Phi}$, $c_{c\Phi}$ for critic networks and $b_{a\Theta}$, $c_{a\Theta}$. $B_{a\Phi}$, $c_{a\Phi}$ for actor networks according to Gaussian kernel properties. Actor-critic network weights $\hat{\omega}_{c\Theta}$, $\hat{\omega}_{a\Theta}$, $\hat{\omega}_{c\Phi}$, $\hat{\omega}_{c\Phi}$ are initialized randomly.
2. Set iterative index $iter = 0$.
3. **begin**
4. **for** sampling distance $k \in (0, L)$
5. Compute system states $\boldsymbol{x}_\Theta$, $\boldsymbol{x}_\Phi$ and outputs $\boldsymbol{y}_\Theta$, $\boldsymbol{y}_\Phi$ with current control inputs u_Θ, u_Φ.
6. Update system errors $e_{\alpha\Theta}$ and $e_{\alpha\Phi}$.
7. Compute reward function r_Θ, r_Φ in current instant.
8. **begin** critic networks
9. Compute the cost to go function $\hat{J}_\Theta$, $\hat{J}_\Phi$ from Eqs. (17) and (18).
10. Update the critic networks weights $\hat{\omega}_{c\Theta}$, $\hat{\omega}_{c\Phi}$ from Eqs. (25) and (26).
11. **end**
12. **begin** actor networks
13. Compute control outputs $u_{a\Theta}$, $u_{a\Phi}$ of actor-critic controller from Eqs. (32) and (33).
14. Update the actor networks weights $\hat{\omega}_{a\Theta}$, $\hat{\omega}_{a\Phi}$ from Eqs. (40) and (41).
15. **end**
16. **end for**
17. $iter = iter+1$.
18. **if** $(\|e_{c\Theta}\| + \|e_{a\Theta}\| > \beta_\Theta$ **or** $\|e_{c\Phi}\| + \|e_{a\Phi}\| > \beta_\Phi)$ **and** $iter< N$, back to 4.
19. **end**

$$\Theta_r(\xi) = \begin{cases} 0, \xi \in [-(\kappa_1 + \kappa_2), 0] \\ 0.05\xi, \xi \in [0, \infty] \end{cases} \tag{44}$$

$$\Phi_r(\xi) = \begin{cases} \frac{\pi}{2}, \xi \in [-(\kappa_1 + \kappa_2), 0] \\ \frac{\pi}{2} - 0.05\xi, \xi \in [0, \infty] \end{cases}. \tag{45}$$

For the interval $[-(\kappa_1 + \kappa_2), 0]$, $\boldsymbol{x} \in \mathcal{C}([-(\kappa_1 + \kappa_2), 0], \Re^3)$, where $\mathcal{C}$ is the functions in Banach space mapping the interval $[-(\kappa_1 + \kappa_2), 0]$ to $\Re^3$ [10]. After parameters adjustment by trial and error, the relevant parameters in actor-critic controller are designed in Table 3.

Table 2
Geometry parameters of BHA.

Symbol	BHA model parameters	Value [Unit]
λ_1	Length from the bit to the first stabilizer	3.66 [m]
λ_2	Length from the first stabilizer to the second stabilizer	6.10 [m]
κ_1	Dimensionless length of the first BHA segment	1 [dimensionless]
κ_2	Dimensionless length of the second BHA segment	2/1.2 [dimensionless]
Λ	Proportion of the length from the bit to the directional mechanism in the length from the bit to first stabilizer	1/6 [dimensionless]
Ir	Inner radius of the BHA	0.053 [m]
Or	Outer radius of the BHA	0.086 [m]
η	Lateral steering resistance	30 [dimensionless]
Π	Scaled active weight-on-bit	30 [dimensionless]
χ	Angular steering resistance	0.1 [dimensionless]

Table 3
Parameters in actor-critic controller for trajectory (Eqs. 44) and (45).

	Q	R	γ	α_c	α_a
Θ	10	diag[1,1]	0.5	0.01	0.05
Φ	10	diag[1,1]	0.5	0.01	0.05

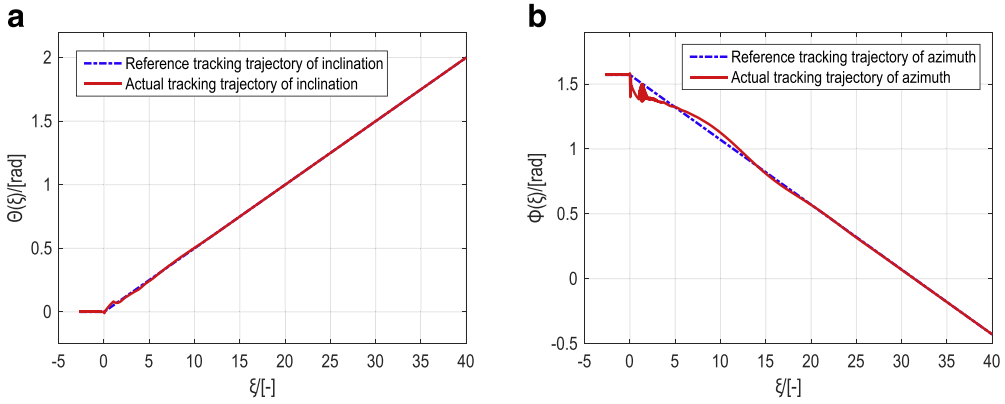

Fig. 5. Trajectories tracking effects of inclination and azimuth.

System initial values of inclination are set to small positive numbers to avoid azimuth direct divergence in the initial states according to (Eq. 1). The trajectories tracking effects for inclination and azimuth are shown in Fig. 5.

It is indicated that the drill bit can track the trajectories of inclination and azimuth accurately and smoothly from Fig. 5 through the cooperative control of IT2FLC and actor-critic controller based on online iterative learning with LPF. At the beginning of azimuth, there exist several frequent oscillations which could have undesirable effects on the smoothness of trajectory. According to analysis on system characteristics, it is mainly caused by the nonlinear amplification coefficient $1/\sin\Theta$ related to inclination in the input matrices of state space equation. This coefficient makes change of azimuth violently frequent when inclination is in transient process. One purpose of controller is to reduce the amplitude and frequency of oscillation in inclination and azimuth simultaneously. Theoretically, if the oscillations of azimuth can be reduced as we expect, the transient process of inclination must be smooth enough. Compared with other control methods, the proposed control method has better tracking effects in reducing the amplitude and frequency of azimuth oscillations just as shown in Fig. 6(b).

For the sake of representing the overall trajectory tracking control effects intuitively, digital PID, T1FLC, IT2FLC and IT2FLC with actor-critic controller (IT2FLC+AC in Fig. 6) are used to compare with proposed control method (IT2FLC+AC with LPF in Fig. 6). The trajectory tracking errors of them are shown as follow.

It is indicated that the digital PID algorithm is not effective enough both in system convergence and oscillation elimination. Similarly, T1FLC algorithm cannot eliminate system oscillation and has relatively large stability errors in azimuth. According to Fig. 6, RSS system is unstable both in inclination and azimuth by using above two control algorithms. IT2FLC algorithm is hard to eliminate the stick–slip oscillations in the transient process and stabilize angle building hole process. Both inclination and azimuth in the drilling process exist periodic oscillations with certain amplitudes. The IT2FLC and actor-critic cooperative

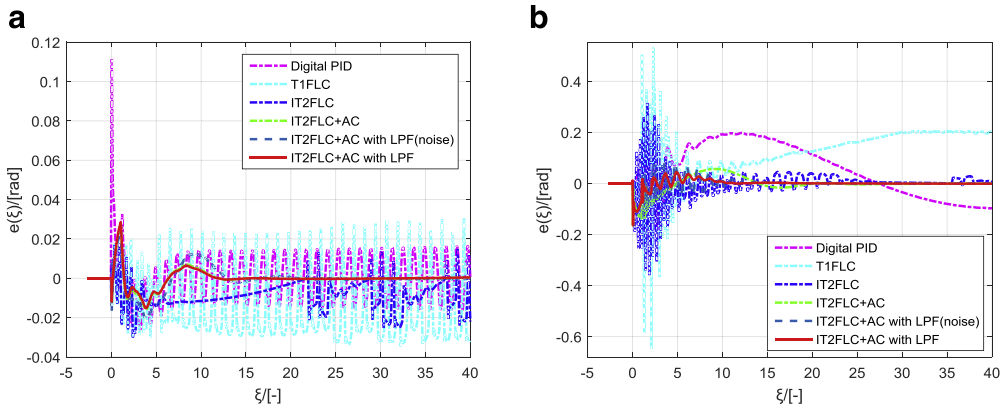

Fig. 6. Trajectories tracking errors of inclination and azimuth by using digital PID, T1FLC, IT2FLC, IT2F+ AC and presented method in this paper.

controller can eliminate the oscillations in the stable angle building hole process. The stick–slip oscillation frequency in the transient process is improved by using the online iterative learning and adjustment. After using low-pass filtering on the control input, the amplitude and frequency of oscillation are further improved, which is particularly evident in the control of azimuth trajectory. Also, some Gaussian white noises with reasonable power have been added to system model as uncertain disturbance. The power of noises is suppressed remarkably by cooperative control inputs. Thus convergence and stability system have been guaranteed by proposed control architecture to a certain extent. Nevertheless, due to the nonlinearity and distance-variation in azimuth bringing by inclination, the control accuracy and smoothness of azimuth are relatively worse compared to inclination.

In order to verify further the usefulness of IT2FLC, we utilize T1FLC algorithm instead of IT2FLC fusing with actor-critic controller to simulate trajectory tracking of RSS model. The results show that T1FLC could make online iterative learning process difficult to converge to ideal values because of system instability in azimuth control after repeated iterative learning training. In other words, T1FLC is not ideal enough to process system nonlinearity and uncertainty and is ineffective in narrowing search space and reducing online learning time. Therefore, IT2FLC, rather than T1FLC, is utilized as a part of cooperative control architecture in our control method.

Control inputs of RSS corresponding to system response which are the decomposed scaled RSS forces Γ_2 and Γ_3 for inclination and azimuth respectively are shown in Fig. 7. It is indicated that the scaled force of inclination becomes stable gradually in transient process. But the scaled force of azimuth needs longer distance to tend to be stable since it depends on inclination and its delay terms.

6.2. Complete trajectory tracking simulation

The complete three-dimensional reference trajectory in x-y-z plane is shown in Fig. 8, whose inclination and azimuth are generated by (Eqs. 46) and (47).

The designed distances are 915 m, and the dimensionless distances are 250 which are determined by dimensionless calculation $\xi = L/\lambda_1$. The simulated trajectory in three-dimensional underground space are aimed to drill horizontal well in the deep stratum where stores rich

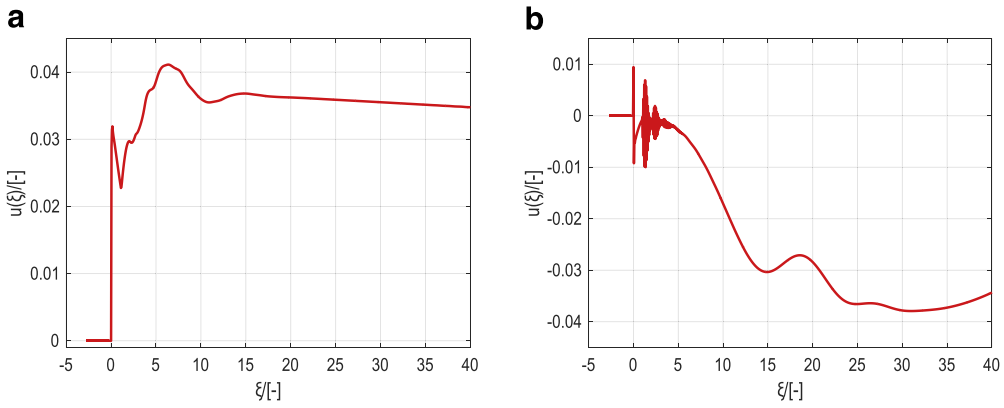

Fig. 7. Control inputs of inclination and azimuth.

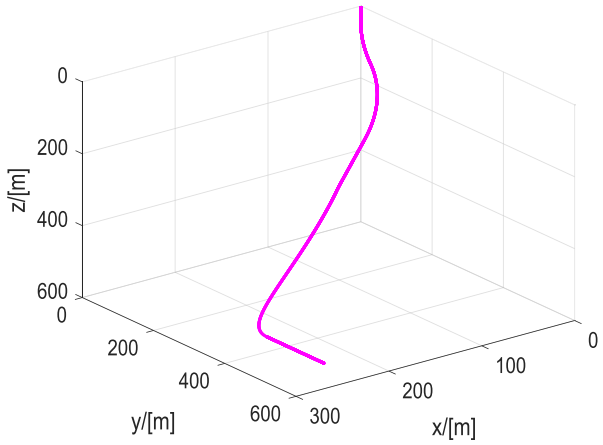

Fig. 8. Simulated drilling reference trajectory in x–y–z plane for (Eqs. 44) and (45).

Table 4
Parameters in actor-critic controller for trajectory (Eqs. 46) and (47).

	$\boldsymbol{Q}$	$\boldsymbol{R}$	γ	α_c	α_a
Θ	10	diag[1,1]	0.5	0.001	0.01
Φ	10	diag[1,1]	0.5	0.001	0.01

oil and gas. In the dimensionless distance interval of [50/λ_1, 150/λ_1], the directional drilling is limited in vertical plane. In other words, only the inclination is changing linearly as distance increases. In the interval of [150/λ_1, 300/λ_1], the inclination is fixed and the azimuth is changing linearly. And in the interval of [400/λ_1, 750/λ_1], we make the inclination and azimuth change simultaneously to drill three-dimensional trajectory in order to test the overall performance of directional drilling. The simulated parameters of actor-critic controller are

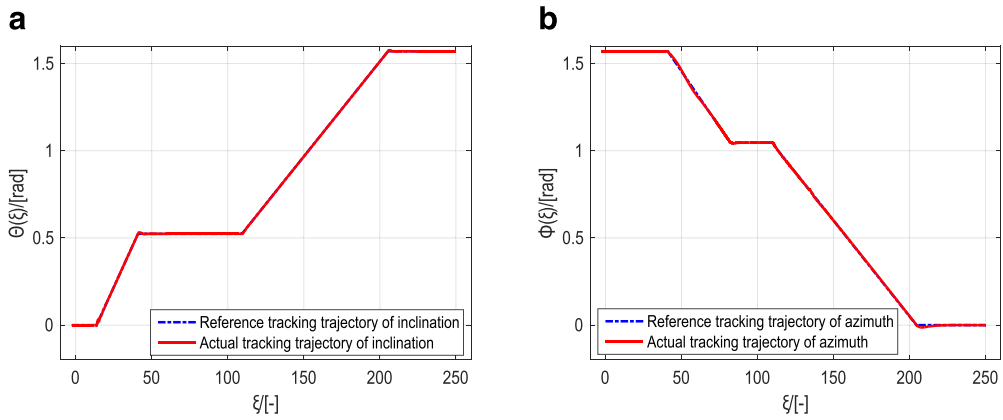

Fig. 9. Trajectories tracking effects of inclination and azimuth.

designed as Table 4 shows.

$$\Theta_r(\xi) = \begin{cases} 0, & \xi \in \left[-(\kappa_1+\kappa_2), \frac{50}{\lambda_1}\right] \\ 0.0192\xi - 0.2623, & \xi \in \left[\frac{50}{\lambda_1}, \frac{150}{\lambda_1}\right] \\ \frac{\pi}{6}, & \xi \in \left[\frac{150}{\lambda_1}, \frac{400}{\lambda_1}\right] \\ 0.0109\xi - 0.6677, & \xi \in \left[\frac{400}{\lambda_1}, \frac{750}{\lambda_1}\right] \\ \frac{\pi}{2}, & \xi \in \left[\frac{750}{\lambda_1}, \frac{915}{\lambda_1}\right] \end{cases} \tag{46}$$

$$\Phi_r(\xi) = \begin{cases} \frac{\pi}{2}, & \xi \in \left[-(\kappa_1+\kappa_2), \frac{150}{\lambda_1}\right] \\ -0.0128\xi + 2.0954, & \xi \in \left[\frac{150}{\lambda_1}, \frac{300}{\lambda_1}\right] \\ \frac{\pi}{3}, & \xi \in \left[\frac{300}{\lambda_1}, \frac{400}{\lambda_1}\right] \\ -0.0110\xi + 2.2494, & \xi \in \left[\frac{400}{\lambda_1}, \frac{750}{\lambda_1}\right] \\ 0, & \xi \in \left[\frac{750}{\lambda_1}, \frac{915}{\lambda_1}\right]. \end{cases} \tag{47}$$

The drilling curve constructed by (Eqs. 46) and (47) is extremely common in actual drilling experiments. However, it is difficult to drill this kind of borehole trajectories because of the lack of control performance. Aimed at resolving this problem, we use the proposed controller to simulate the complete drilling trajectories. The tracking effects are shown in Fig. 9.

Figs. 9(a) and 10(a) show the tracking effect of inclination, the reference trajectory is tracked accurately and smoothly by simulated drilling trajectory. Stick–slip oscillations are eliminated by actor-critic controller and LPF. Due to the relatively small angle building hole rate, overshoots in dynamic process are also smaller than the first simulation test. Fig. 9(b) shows the effective tracking performance in azimuth. Simulated drilling trajectory accurately tracks the reference drilling trajectory. Compared with drilling process in inclination, azimuth is still influenced by nonlinear amplification factor 1/sinΘ. Stick-slip oscillations are

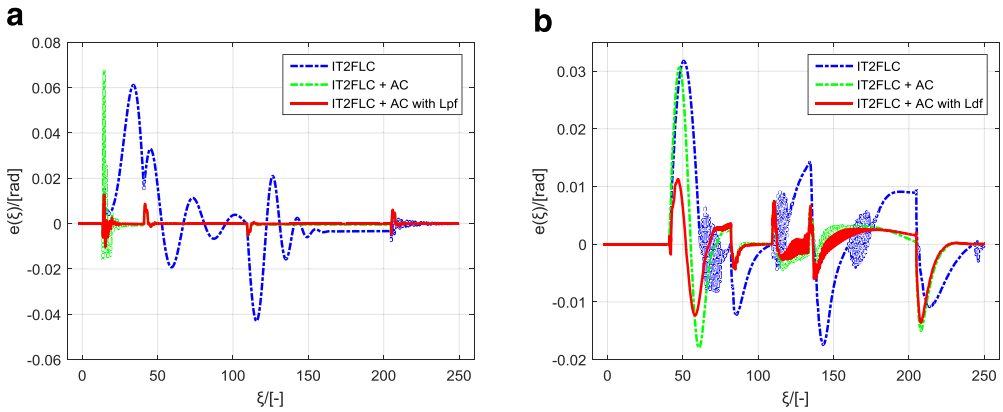

Fig. 10. Trajectories tracking errors of inclination and azimuth by using IT2FLC, IT2FLC+AC and presented method in this paper.

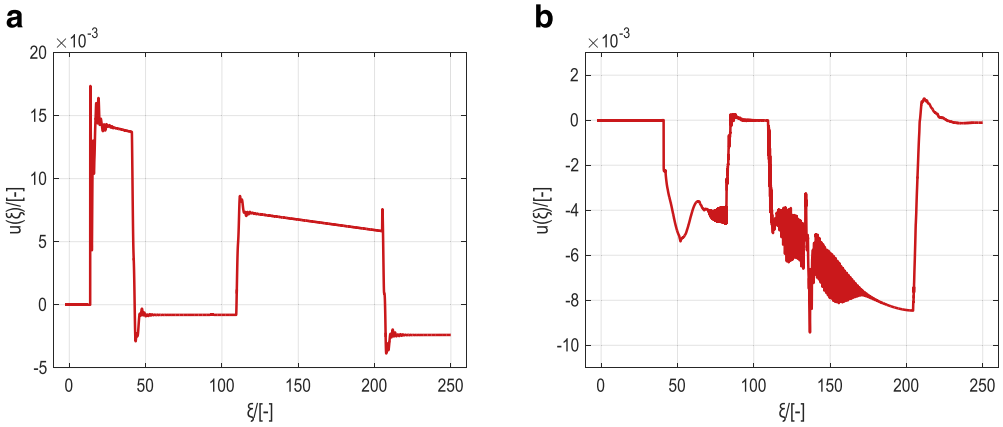

Fig. 11. Control inputs of inclination and azimuth.

eliminated to some extent but the final control effects in azimuth are relatively worse than inclination.

In order to show tracking control effects intuitively, IT2FLC and IT2FLC with actor-critic controller (IT2FLC+AC in Fig. 10) are also used to compare with our control method (IT2FLC+AC with LPF in Fig. 10) proposed in this paper. Digital PID and T1FLC algorithms are no longer used as comparisons considering that they are not effective enough in convergence and oscillation elimination. It is indicated that our controller is more excellent than two other controllers in oscillation elimination, along with trajectory accuracy and smoothness. In Fig. 10(b), when azimuth is converted from building hole process to holding deviation process or conversely, the stick–slip oscillations are changing violently but the amplitude is relatively small. Considering that the complex drilling environment and system nonlinearity, these amplitude and frequency of stick–slip oscillations fall within the acceptable boundaries.

The control inputs of inclination and azimuth are shown in Fig. 11. Although a small amount of stick–slip oscillation still exists in azimuth, it has little effects on the final control performances. The directional force of azimuth is smaller compared with inclination. It is also caused by nonlinear amplification factor $1/\sin\Theta$ which acts on input matrix to magnify directional force of azimuth. In addition, due to the mechanical constraints and small angle building rate compared with the first simulation test, directional force in inclination and azimuth are also smaller than the previous simulation.

The scaled RSS force is feasible in practice according to characteristic force calculation. For example, $\Gamma = 0.008$ is equivalent to actual RSS force 13,000N [3]. The rapid change of directional force in a small range has little effects on overall directional accuracy and smoothness. In addition, mechanical strength and working life of RSS are also guaranteed.

7. Conclusion

In this paper, a model-based trajectory tracking control method based on type-2 fuzzy logic controller and actor-critic controller is proposed for complex three-dimensional downhole curve drilled by static push-the-bit rotary steerable systems. The controller is designed to drive RSS to track the reference trajectory by adjusting directional force to change inclination and azimuth, respectively. We introduce nonlinearities and uncertainties processing of IT2FLC and approximate capabilities both in action function of actor network and value function in critic network to realize trajectory tracking and stick–slip oscillations elimination synchronously.

Simulation results show that the cooperative controller of IT2FLC and actor-critic with low pass filter lowers the tracking errors in inclination and azimuth control in angle building hole rate test and complex complete trajectory tracking. The stick–slip oscillations are eliminated in inclination and suppressed largely in azimuth. Therefore, the proposed controller provides satisfying control performances in realizing trajectory tracking accurately and smoothly of directional drilling.

In the future research, a dynamic reference trajectory will be studied to replace pre-defined trajectory. It is difficult to design the whole drilling trajectory in advance because the actual drilling environment is not completely observable relying on current measurement technique. Correspondingly, control methods with better performance will be designed to dynamic trajectory on real-time optimization, anti-disturbance and approximation. Another difficult challenge in RSS trajectory tracking is downhole location and navigation. Excellent location and navigation technique can help drill bit find the target hydrocarbon reservoir and avoid the complex stratum which is not suitable for drilling. In addition, deviation adjustment mechanism should be considered in control system. It must be guaranteed that the drill bit can return to correct direction if the drilling trajectory severely deviates from original reference trajectory.

Acknowledgment

This work was supported by National Natural Science Foundations of PR China (Grants Nos. 51405484, 61773374); and supported in part by Project of Development in Tianjin for Scientific Research Institutes Supported by Tianjin Government (Grant No. 16PTYJGX00050).

Appendix

(1) System and Output Matrices for Continuous Model

$$\boldsymbol{A}_0 = \begin{bmatrix} \boldsymbol{A}_{01} & 0 \\ 0 & \boldsymbol{A}_{02} \end{bmatrix},\ \boldsymbol{A}_1 = \begin{bmatrix} \boldsymbol{A}_{11} & 0 \\ 0 & \boldsymbol{A}_{12} \end{bmatrix},\ \boldsymbol{A}_2 = \begin{bmatrix} \boldsymbol{A}_{21} & 0 \\ 0 & \boldsymbol{A}_{22} \end{bmatrix}$$

The partitioned matrix blocks $\boldsymbol{A}_{01}, \boldsymbol{A}_{02}, \boldsymbol{A}_{11}, \boldsymbol{A}_{12}, \boldsymbol{A}_{21}, \boldsymbol{A}_{22}$ in $\boldsymbol{A}_0, \boldsymbol{A}_1, \boldsymbol{A}_2$, respectively can be found in [7], and the forms of $\boldsymbol{B}_0(\xi), \boldsymbol{B}_1(\xi)$, respectively, are

$$\boldsymbol{B}_0(\xi) = \frac{1}{\chi\Pi}\begin{bmatrix} E & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{E}{\sin\Theta} + \frac{\chi}{\eta}\frac{\cos\Theta}{\sin^2\Theta}(\Theta - \Theta_1) & 0 & 0 \end{bmatrix}^T$$

$$\boldsymbol{B}_1(\xi) = \frac{1}{\chi\Pi}\begin{bmatrix} \frac{\chi}{\eta}\mathcal{F}_r & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\chi}{\eta}\frac{\mathcal{F}_r}{\sin\Theta} & 0 & 0 \end{bmatrix}^T$$

where $E = (\mathcal{F}_b\mathcal{M}_1 - \mathcal{F}_1\mathcal{M}_b - \mathcal{M}_1\eta\Pi)/(\eta\Pi)$.

Output matrices $\boldsymbol{C}$ and $\boldsymbol{D}$ are expressed as follow:

$$\boldsymbol{C} = \frac{1}{\eta\Pi - \mathcal{F}_b}\begin{bmatrix} \eta\Pi & -\mathcal{F}_b + \mathcal{F}_1 & -\mathcal{F}_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \eta\Pi & -\mathcal{F}_b + \mathcal{F} & -\mathcal{F}_1 \end{bmatrix}$$

$$\boldsymbol{D} = \begin{bmatrix} \mathcal{F}_r & 0 \\ 0 & \mathcal{F}_r \end{bmatrix}.$$

(2) System Matrices for Discrete Model

$$\boldsymbol{G}_0 = e^{\boldsymbol{A}_0 S_d},$$

$$\boldsymbol{G}_{11} = \int_0^{S_d(1-d_1)} e^{\boldsymbol{A}_0\xi} d\xi * \boldsymbol{A}_1,\ \boldsymbol{G}_{12} = \int_{S_d(1-d_1)}^{S_d} e^{\boldsymbol{A}_0\xi} d\xi * \boldsymbol{A}_1,$$

$$\boldsymbol{G}_{21} = \int_0^{S_d(1-d_2)} e^{\boldsymbol{A}_0\xi} d\xi * \boldsymbol{A}_2,\ \boldsymbol{G}_{22} = \int_{S_d(1-d_2)}^{S_d} e^{\boldsymbol{A}_0\xi} d\xi * \boldsymbol{A}_2,$$

$$\boldsymbol{H}_0(k) = \int_0^{S_d} e^{\boldsymbol{A}_0\xi} d\xi * \boldsymbol{B}_0(k),$$

$$\boldsymbol{H}_1(k) = \int_0^{S_d} e^{\boldsymbol{A}_0\xi} d\xi * \boldsymbol{B}_1(k).$$

References

[1] J. Matheus, M. Ignova, P. Hornblower, A hybrid approach to close-loop directional drilling control using rotary steerable systems, in: Proceedings of the IFAC Workshop Automatic Control in Offshore Oil and Gas Production, vol. 45, 2012, pp. 84–89.

[2] C. Carpenter, Torsional dynamics and point-the-bit rotary steerable systems, J. Pet. Technol. 65 (12) (2013) 111–114.

[3] N. Kremers, E. Detournay, N. van de Wouw, Model-based robust control of directional drilling systems, IEEE Trans. Control Syst. Technol. 24 (1) (2016) 226–238.

[4] M. Bayliss, N. Panchal, J. Whidborne, Rotary steerable directional drilling stick/slip mitigation control, in: Proceedings of the IFAC Workshop Automatic Control Offshore Oil Gas Production., vol. 45, 2012, pp. 66–71.

[5] H. Sun, Z. Li, N. Hovakimyan, T. Basar, G. Downton, L1 adaptive controller for a rotary steerable system, in: Proceedings of the IFAC Workshop Autom. Control Offshore Oil Gas Production, vol.1, 2011, pp. 72–77.
[6] N. van de Wouw, N. Kremers, E. Detournay, Output-feedback inclination control of directional drilling systems, IFAC Papers Online 48 (6) (2015) 260–265.
[7] E. Pecht, M.P. Mintchev, Modeling of observability during in-drilling alignment for horizontal directional drilling, IEEE Trans. Instrum. Meas. 56 (5) (2007) 1946–1954 pp. 144-154.
[8] Q. Xue, H. Leung, R. Wang, B. Liu, Y. Wu, Continuous real-time measurement of drilling trajectory with new state-space models of Kalman filter, IEEE Trans. Instrum, Meas. 65 (1) (2016).
[9] N. Panchal, M. Bayliss, J. Whidborne, Attitude control system for directional drilling bottom hole assemblies, IET Control Theory Appl 6 (7) (May. 2012) 884–892.
[10] N. Kremers, Model-based robust control of a directional drilling system MS thesis, Department Mech Engineering, Eindhoven University of Tech, Eindhoven, Netherland, 2013.
[11] G. Downton, M. Ignova, Stability and response of closed loop directional drilling system using linear delay differential equations, in: Proceedings of the IEEE Conference On Control Applications (CCA), 2011, pp. 893–898.
[12] L. Perneder, E. Detournay, Anomalous behaviors of a propagating borehole, in: Proceedings of the SPE DDC, Galveston, TX, USA, 2012, pp. 352–358.
[13] L. Perneder, A three-dimensional mathematical model of directional drilling, Ph.D. dissertation, Department Civil Engineering, University Minnesota, Minneapolis, MN, USA, 2013.
[14] L. Perneder, E. Detournay, Equilibrium inclinations of straight boreholes, SPE J. 18 (3) (2013) 395–405.
[15] L. Perneder, E. Detournay, Steady-state solutions of a propagating borehole, Int. J. Solids Struct. 50 (9) (2013) 1226–1240.
[16] L. Perneder, E. Detournay, G.C. Downton, Bit/rock interface laws in directional drilling, Int. J. Rock Mech. Min. 51 (2012) 81–90.
[17] B. Besselink, T. Vromen, N. Kremers, N. van de Wouw, Analysis and control of stick–slip oscillations in drilling systems, IEEE Trans. Control Syst. Technol. 24 (5) (2016) 1582–1593.
[18] T. Vromen, C.-H. Dai, N. van de Wouw, T. Oomen, P. Astrid, H. Nijmeijer, Robust output-feedback control to eliminate stick–slip oscillations in drill-string systems, in: Proceedings of the Second IFAC Workshop Automatic Control Offshore Oil Gas Production, Florianopolis, Brazil, 2015, pp. 272–277.
[19] S. Tong, Y. Li, S. Sui, Adaptive fuzzy tracking control design for SISO uncertain nonstrict feedback nonlinear systems, IEEE Trans. Fuzzy Syst. 24 (6) (2016) 1441–1454.
[20] Y. Li, S. Tong, L. Liu, G. Feng, Adaptive output-feedback control design with prescribed performance for switched nonlinear systems, Automatica 80 (2017) 225–231.
[21] Y. Li, S. Sui, S. Tong, Adaptive fuzzy control design for stochastic nonlinear switched systems with arbitrary switchings and unmodeled dynamics, IEEE Trans. Cybern. 47 (2) (2017) 403–414.
[22] O. Castillo, L. Cervantes, J. Soria, M. Sanchez, J.R. Castro, A generalized type-2 fuzzy granular approach with applications to aerospace, Inf. Sci. 354 (2016) 165–177.
[23] L. Cervantes, O. Castillo, Type-2 fuzzy logic aggregation of multiple fuzzy controllers for airplane flight control, Inf. Sci. 324 (2015) 247–256.
[24] M.A. Sanchez, O. Castillo, J.R. Castro, Generalized type-2 fuzzy systems for controlling a mobile robot and a performance comparison with interval type-2 and type-1 fuzzy systems, Expert Syst. Appl. 42 (14) (2015) 5904–5914.
[25] S.M. Chen, N.Y. Chung, Forecasting enrollments of students using fuzzy time series and genetic algorithms, Int. J. Inf. Manage. Sci. 17 (3) (2006) 1–17.
[26] S.M. Chen, T.H. Chang, Finding multiple possible critical paths using fuzzy PERT, IEEE Trans. Syst. Man Cybern. B Cybern. 31 (6) (2001) 930–937.
[27] G. Wang, Y. Li, X. Li, Approximation performance of the nonlinear hybrid fuzzy system based on variable universe, Granul. Comput. 2 (2017) 73–84.
[28] S.M. Chen, P.Y. Kao, TAIEX forecasting based on fuzzy time series, particle swarm optimization techniques and support vector machines, Inf. Sci. 247 (2013) 62–71.
[29] M.A. Sanchez, J.R. Castro, O. Castillo, O. Mendoza, A. Rodriguez-Diaz, P. Melin, Fuzzy higher type information granules from an uncertainty measurement, Granul. Comput. 2 (2017) 95–103.
[30] J. Mendel, Type-2 fuzzy sets and systems: an overview, IEEE Comput. Intell. Mag. 2 (2) (May. 2007) 20–29.
[31] D. Wu, On the fundamental differences between interval type-2 and type-1 fuzzy logic controllers, IEEE Trans. Fuzzy Syst. 20 (5) (2012) 832–848.
[32] T. Kumbasar, Robust stability analysis and systematic design of single-input interval type-2 fuzzy logic controllers, IEEE Trans. Fuzzy Syst. 24 (3) (2016) 576–694.

[33] O. Castillo, L. Amador-Angulo, J.R. Castro, M. Garcia-Valdez, A comparable study of type-1 fuzzy logic systems, interval type-2 fuzzy logic systems and generalized type-2 fuzzy logic systems in control problems, Inf. Sci. 354 (2016) 257–274.
[34] C. Chatar, A. Stepnov, A. Mardyashov, J. Gonzalez, Remote directional drilling and logging while drilling operations in the Arctic, in: Proceedings of the IADC/SPE Drilling Conference and Exhibition, Fort Worth, TX, USA, March 1–3, 2016.
[35] H. Ebeltoft, Y. Majeed, E. Soergard, Hydrate control during deepwater drilling: overview and new drilling-fluids formulations, SPE Drilling Completion 16 (1) (2001) 19–26.
[36] D. Dong, C. Chen, J. Chu, T.J. Tarn, Robust quantum-inspired reinforcement learning for robot navigation, IEEE/ASME Trans. Mechatronics 17 (1) (2012) 86–97.
[37] F. Cruz, S. Magg, C. Weber, S. Wermter, Training agents with interactive reinforcement learning and contextual affordances, IEEE Trans. Cogn. Develop. Syst. 8 (4) (2016) 271–274.
[38] Q. Wei, D. Liu, Data-driven neuro-optimal temperature control of water-gas shift reaction using stable iterative adaptive dynamic programming, IEEE Trans. Ind. Electron. 61 (11) (2014) 6399–6408.
[39] R. Ganesan, T.K. Das, K.M. Ramachandran, A multiresolution analysis-assisted reinforcement learning approach to run-by-run control, IEEE Trans. Autom. Sci. Eng. 4 (2) (2007) 182–193.
[40] Q. Yang, S. Jagannathan, Reinforcement learning controller design for affine nonlinear discrete-time systems using online approximators, IEEE Trans. Syst. Man Cybern. B Cybern. 42 (2) (2012) 377–390.
[41] B. Xu, C. Yang, Z. Shi, Reinforcement learning output feedback nn control using deterministic learning technique, IEEE Trans. Neural Netw. Learn. Syst. 23 (3) (2014) 635–641.
[42] Q. Fan, G. Yang, Adaptive actor-critic design-based integral sliding-mode control for partially unknown nonlinear systems with input disturbances, IEEE Trans. Neural Netw. Learn. Syst. 27 (1) (2016) 165–177.
[43] L. Liu, Z. Wang, H. Zhang, Adaptvie fault-tolerant tracking control for MIMO discrete-time systems via reinforcement learning algorithm with less learning parameters, IEEE Trans. Autom. Sci. Eng. 14 (1) (Jan. 2017) 299–314.
[44] S. Hung, S.N. Givigi, A Q-learning approach to flocking with UAVs in a stochastic environment, IEEE Trans. Cybern. 47 (1) (Jan. 2017) 186–197.
[45] E. Antonelo, B. Schrauwen, On learning navigation behaviors for small mobile robots with reservoir computing architectures, IEEE Trans. Neural Netw. Learn. Syst. 26 (4) (2015) 763–780.
[46] R. Cui, C. Yang, Y. Li, S. Sharma, Adaptive neural network control of AUVs with control input nonlinearities using reinforcement learning, IEEE Trans. Syst., Man, Cybern., Syst. 47 (6) (2017) 1019–1029.
[47] I. Grondman, M. Vaandrager, L. Busoniu, R. Babuska, E. Schuitema, Efficient model learning methods for actor-critic control, IEEE Trans. Syst. Man Cybern. B Cybern. 42 (3) (2012) 591–602.
[48] V. Denoel, E. Detournay, Eulerian formulation of constrained elastic, Int. J. Solids Struct. 48 (3–4) (2011) 625–636.
[49] T. Kumbasar, Robust stability analysis and systematic design of single-input interval type-2 fuzzy logic controllers, IEEE Trans. Fuzzy Syst. 24 (3) (Jun. 2016) 675–694.
[50] M. Biglarbegian, W.W. Melek, J. Mendel, On the stability of interval type-2 TSK fuzzy logic control systems, IEEE Trans. Syst. Man Cybern. B Cybern. 40 (3) (Jun. 2010) 832–848.
[51] H. Hu, Y. Wang, Y. Cai, Advantages of the enhanced opposite direction searching algorithm for computing the centroid of an interval type-2 fuzzy set, Asian J. Control 14 (6) (2012) 1–9.
[52] D. Wu, Approaches for reducing the computational cost of interval type-2 fuzzy logic systems: overview and comparisons, IEEE Trans. Fuzzy Syst. 21 (1) (Feb. 2013) 80–99.
[53] D. Prokhorov, D. Wunsch, Adaptive critic designs, IEEE Trans. Neural Netw. 8 (5) (Sep. 1997) 997–1007.
[54] F. Lewis, D. Vrable, Reinforcement learning and adaptive dynamic programming for feedback control, IEEE Circuit Syst. Mag. 9 (3) (Third Quarter. 2009) 32–50.