IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 28, 2019, accepted May 15, 2019, date of publication May 27, 2019, date of current version June 18, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2919249

Optimized Backstepping Design for Ship Course Following Control Based on Actor-Critic Architecture With Input Saturation

YUMING BAI, YUCHI CAO, AND TIESHAN LI
Navigation College, Dalian Maritime University, Dalian 116026, China

Corresponding authors: Yuchi Cao (caoyuchi@dlmu.edu.cn) and Tieshan Li (tieshanli@126.com)

This work was supported in part by the Natural Science Foundation of Liaoning Province under Grant 20170540093, in part by the Science and Technology Innovation Funds of Dalian under Grant 2018J11CY022, in part by the National Natural Science Foundation of China under Grant 61751202, Grant 61751205, Grant 61572540, and Grant 61803064, in part by the Fundamental Research Funds for the Central Universities under Grant 3132019124 and Grant 3132019126.

ABSTRACT This paper presents a course following control method for ships based on optimized backstepping (OB) technology. The backstepping technology is employed as the main control framework since the ship course can be modeled in the strict feedback form. Based on the actor-critic architecture and radial basis function (RBF) neural network (NN), the reinforcement learning (RL) strategy is introduced to avoid the difficulty in solving the traditional Hamilton-Jacobi-Bellman (HJB) equation directly. The actor NNs are used for carrying out the control law, while the critic NNs aim at evaluating the tracking performance. An auxiliary design system and Gaussian error function are employed to handle the practical problem of input saturation. The stability of the closed-loop system can be guaranteed via Lyapunov theory. Finally, simulation examples and comparison are provided to demonstrate and verify the superior performance and advantages on course following and energy saving of the control scheme proposed in this paper.

INDEX TERMS Actor-critic architecture, Gaussian error function, input saturation, optimized backstepping, ship course following.

I. INTRODUCTION

Due to the rising demands for higher safety level and energy saving, ship motion control has attracted ubiquitous attention for decades. As an important research branch in ship motion control, the ship course following control is also investigated in numerous literature [1]–[4]. The dynamic characteristics of ship rudder, which is the vital equipment for steering ship course, vary as soundings, navigational status, loading conditions, exogenous disturbances (wind, current and wave, etc.), so it is challenging to achieve satisfactory course following performance by means of ship steering control because of maneuverability difficulties [5] and high inherent nonlinearity, especially in combination with consideration of energy saving. Until now, backstepping has been one of the most powerful and popular control scheme in the context of lower triangular and strict feedback systems. Its basic principle is to view the state variables as ''virtual controls'' and then to systematically design the virtual control laws and final actual control law in a recursive design process. Backstepping has also been widely applied in ship course following control [6]–[8]. Although good tracking performance can be achieved, the optimization of backstepping is not involved.

Since the ship sailing and maneuvering are at the cost of massive energy consumption, which is more prominent and conspicuous for larger ships, it is vitally necessary to take the optimization and energy saving into consideration for ship course following control. The optimal control is usually achieved in the way of solving corresponding Hamilton-Jacobi-Bellman (HJB) equation. However, it is very difficult to solve the HJB equation directly due to its intractability and inherent nonlinearity. To conquer the difficulty in solving HJB equation, a novel optimized backstepping (OB) technology is firstly put forward in [9] to achieve optimized control by fusing optimization into backstepping technology based on the actor-critic architecture. OB is applied in the ship tracking control in [10]. However, the practical problem of input saturation is not considered in [9], [10].

The associate editor coordinating the review of this manuscript and approving it for publication was Zhong Wu.

Actuator saturation is a practical problem in ship course control system due to the rudder angle limitation The saturation constraints of actuator may deteriorate and degrade the control performance, or even make system lose stability if handled improperly [11]. On the other hand, violation of input saturation may result in system damage and failure. Thus, it is necessary to consider actuator saturation in the controller design. Input saturation is usually expressed by the sign and saturation functions [6]–[12], asymmetric non-smooth input saturation [13]. In order to overcome the disadvantages of sharp corner existing in general saturation functions when reaching the saturation limits, the saturation is approximated by a smooth hyperbolic functions [14]–[18] or Gaussian error function [19]. Moreover, an auxiliary design system [20], [21] is further employed for the controller design with input saturation and further stability analysis.

Model predictive control (MPC) is an important nonlinear control technique to handle input constraints while considering optimality simultaneously. MPC has been applied to a thrust allocation [22] and path following [23], [24] for marine vessels to minimize the power consumption. However, the main drawback of MPC is the computational burden of the optimization problem at each step [25].

Inspired by above-mentioned discussions, this manuscript constructs an OB control scheme for ship course following with actuator saturation. Two radial basis function (RBF) neural networks (NNs) are constructed to execute the reinforcement learning (RL) algorithm, and actor NNs are employed for executing the control law to obtain satisfied tracking performance, and critic NNs are used for evaluating the tracking performance by minimizing Bellman error. So the optimized control and Lyapunov stability can be guaranteed and balanced simultaneously, while the difficulty in solving HJB equation can be well avoided by applying RBF NNs with universal approximation ability [26].

Specifically, the main contributions of the paper are listed as follows.

(1) By integrating the actor-critic architecture into backstepping, the OB control technique adopted by this paper can minimize unnecessary energy losses and prolong the service life of rudder by optimizing the amplitude and operation frequency of rudder, which may produce tremendous benefit on energy saving and environment protection in maritime industry.
(2) The auxiliary design system and Gaussian error function are employed together to handle the physical problem of actuator saturation, which is more applicable and consistent to the actual ship manoeuvring situation.
(3) Numerical simulation is implemented to illustrate the effectiveness of the OB control technique. In addition, comparative simulations with direct NN control and MPC control are carried out to further demonstrate the advantages of proposed control scheme.

The rest sections of the manuscript are organized as follows. Problem formulation including the ship course model, RBF NNs and Gaussian error function are introduced briefly in the Section II. In Section III, ship course following controller is developed based on OB control approach, and the stability analysis is presented. A simulation example and comparisons are shown in Section IV to illustrate the satisfactory tracking performance and energy saving performance of the controller. Finally, the paper is concluded in Section V.

II. PROBLEM FORMULATION

A. NOTATIONS

Following notations will be used throughout the paper, $|\cdot|$ represents the absolute value of a scalar and $\|\cdot\|$ denotes the Euclidean norm of a vector or the Frobenius norm of a matrix. R^n represents the n-dimensional Euclidean space.

B. MATHEMATICAL MODEL OF SHIP COURSE CONTROL

The mathematical model with respect to ship steering control can be described as following [27]:

$$T\ddot{\psi} + \dot{\psi} + \alpha\dot{\psi}^3 = K\delta \tag{1}$$

The corresponding system parameters involved in equation (1) are listed in Table 1.

TABLE 1. Parameters of the ship course model.

Parameters	Physical	Units
T	Time constant	s
ψ	Ship heading angle	rad
$\dot{\psi}$	Yaw rate of ship heading angle	rad/s
α	Norrbin coefficient	-
K	rudder gain	1/s
δ	rudder angle	rad

By defining $x_1 = \psi$, $x_2 = \dot{\psi} = r$, $K\delta/T = u$, the ship steering control model (1) can be transformed into the following form:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f_2(x_2) + u \\ y = x_1 \end{cases} \tag{2}$$

where $f_2(x_2) = -x_2/T - \alpha x_2^3/T$.

Remark 1: Noted that the input controller u is developed for the ship course model (2), so the actual control law δ for system (1) can be derived by multiplying T/K.

C. SATURATION NONLINEAR MODEL

Definition 1[19]: Gaussian error function $erf(x)$ is a class of nonelementary function with sigmoid shape, it can be defined as

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{3}$$

Remark 2 [19]: Gaussian error function is real-valued and continuous differentiable, it has no singularities except at infinity, and its Taylor expansion always converges.

According to definition 1, a smooth saturation nonlinearity model can be expressed as following:

$$u(v) = u_M \times erf(av) \tag{4}$$

where $a = \sqrt{\pi}/(2u_M)$. The bounds of $u(v)$ can be easily adjusted by changing the value u_M in the equation (4) to match the input saturation.

To promote the successive controller design, construct a function as following:

$$\Delta v = u_M \times \mathrm{erf}(av) - v(t) \tag{5}$$

Then system (2) can be rewritten as

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f_2(x_2) + (v + \Delta v) \\ y = x_1 \end{cases} \tag{6}$$

D. RBF NN

RBF NN is usually utilized as a solution for modelling continuous nonlinear functions due to its universal approximation ability on a compact set [28]–[31].

In the manuscript, the following RBF NN [32], [33] is employed to approximate any continuous function $h(z)$: $R^q \to R$

$$h(z) = W^{*T} S(z) + \varepsilon, \quad \forall Z \in \Omega_z \tag{7}$$

where $Z = \left[z_1, z_2, \cdots, z_q\right]^{\mathrm{T}} \in \Omega_z \subset R^q$ is the input vector, $W^* = \left[w_1^*, w_2^*, \cdots, w_l^*\right]^{\mathrm{T}} \in R^l$ is the vector of ideal constant weight, $l > 1$ is the number of NN nodes; $\|\varepsilon\| < \varepsilon_M$ is the corresponding approximation error with a positive constant ε_M [34], [35]. Each $s_i(Z)$ in $S(Z) = [s_1(Z), s_2(Z), \cdots, s_l(Z)]^T$ is selected as the commonly used Gaussian functions with following form

$$s_i(Z) = \exp\left[\frac{-(Z-\mu_i)^T(Z-\mu_i)}{\eta_i^2}\right], \quad i = 1, 2, \cdots, 1 \tag{8}$$

where $\mu_i = \left[\mu_{i1}, \mu_{i2}, \cdots \mu_{iq}\right]^T$ is the center of the receptive field and η_i is the spread of the Gaussian function.

E. USEFUL LEMMAS

The following lemma will be used to simplify the stability analysis.

Lemma 1 [36]: $G(t) \in R$ is a continuous and positive function with bounded initial value $G(0)$. If the inequality $\dot{G}(t) \le -aG(t) + c$ holds, where a and c are constants, then following inequality can be obtained:

$$G(t) \le \mathrm{e}^{-at} G(0) + \frac{c}{a}(1 - \mathrm{e}^{-at}) \tag{9}$$

III. OB CONTROLLER DESIGN

An OB controller is designed in this section to track the reference course signal with superior performance and minimize the energy cost in the condition of input saturation.

Step 1: Define the error variable $z_1 = x_1 - y_d$, then differentiate z_1 with respect to time along (2), one has

$$\dot{z}_1 = \dot{x}_1 - \dot{y}_d = x_2 - \dot{y}_d \tag{10}$$

where y_d is the desired course signal, and x_2 is treated as the intermediate controller.

Construct the infinite horizon value function as following

$$V_1(z_1) = \int_t^\infty r_1(z_1, \alpha_1) ds \tag{11}$$

where $\alpha_1(z_1)$ is the virtual control for the subsystem, and $r_1(z_1, \alpha_1) = z_1^2 + \alpha_1^2$ is the cost function.

View $\alpha_1^*(z_1)$ as the optimal virtual control to obtain the following optimal value function.

$$\begin{aligned} V_1^*(z_1) &= \min_{\alpha_1 \in \Psi(\Omega_1)} \left(\int_t^\infty r_1(z_1, \alpha_1)\, ds \right) \\ &= \int_t^\infty r_1\left(z_1, \alpha_1^*\right) ds \end{aligned} \tag{12}$$

where $\Psi(\Omega_1)$ is the domain of admissible control policies over the compact set Ω_1.

The Hamiltonian function with respect to the value function (12) is

$$H_1(z_1, \alpha_1, \frac{\partial V_1}{\partial z_1}) = r_1(z_1, \alpha_1) + \frac{\partial V_1}{\partial z_1}\dot{z}_1 \tag{13}$$

where $\partial V_1/\partial z_1$ denotes the gradient of V_1 associating with z_1.

The following HJB equation is defined based on (12) and (13),

$$H_1(z_1, \alpha_1^*, \frac{\partial V_1^*}{\partial z_1}) = z_1^2 + \alpha_1^{*2} + \frac{\partial V_1^*}{\partial z_1}(\alpha_1^* - \dot{y}_d) = 0 \tag{14}$$

Under the assumption that the solution for equation (14) uniquely exists, the optimal virtual control α_1^* can be achieved by solving $\partial H_1(z_1, \alpha_1^*, \partial V_1^*/\partial z_1)/\partial \alpha_1^* = 0$.

$$\alpha_1^* = -\frac{1}{2}\frac{\partial V_1^*}{\partial z_1} \tag{15}$$

Substituting (15) into (14), the following HJB equation yields:

$$z_1^2 - \frac{\partial V_1^*}{\partial z_1}\dot{y}_d - \frac{1}{4}(\frac{\partial V_1^*}{\partial z_1})^2 = 0 \tag{16}$$

To facilitate the optimal virtual control design later, decompose the optimal value function into two items as:

$$V_1^*(z_1) = \beta_1 z_1^2 - \beta_1 z_1^2 + V_1^*(z_1) = \beta_1 z_1^2 + V_1^o(z_1) \tag{17}$$

where β_1 is a positive constant to be determined, and $V_1^o(z_1) = -\beta_1 z_1^2 + V_1^*(z_1)$.

Inserting (17) into (15), it can be achieved that

$$\alpha_1^* = -\frac{1}{2}\frac{\partial V_1^*}{\partial z_1} = -\beta_1 z_1 - \frac{1}{2}\frac{\partial V_1^o(z_1)}{\partial z_1} \tag{18}$$

Considering the difficulty in solving the equation due to the strong nonlinearities, the actor-critic architecture based RL is proposed by means of the excellent approximating ability of RBF NN as following:

$$\frac{\partial V_1^o(z_1)}{\partial z_1} = W_1^{*T} S_1(z_1) + \varepsilon_1(z_1) \tag{19}$$

where $W_1^* \in R^n$ is the ideal constant NN weight, n is the neuron number; $S_1(z_1) \in R^n$ is the Gaussian function vector; the approximation error $\varepsilon_1(z_1) \in R$ and its first derivative are required to be bounded.

According to (19), the optimal value function and optimal virtual control are rewritten as

$$\frac{\partial V_1^*(z_1)}{\partial z_1} = 2\beta_1 z_1 + W_1^{*T} S_1(z_1) + \varepsilon_1(z_1) \tag{20}$$

$$\alpha_1^* = -\beta_1 z_1 - \frac{1}{2}(W_1^{*T} S_1(z_1) + \varepsilon_1(z_1)) \tag{21}$$

Substituting (20) and (21) into the HJB equation (14), we have

$$\begin{aligned} H_1(z_1, \alpha_1^*, W_1^*) = &-(\beta_1^2 - 1)z_1^2 - 2\beta_1 z_1 \dot{y}_d \\ &- W_1^{*T} S_1(z_1)(\beta_1 z_1 + \dot{y}_d) \\ &- \frac{1}{4}(W_1^{*T} S_1(z_1))^2 + \rho_1 = 0 \end{aligned} \tag{22}$$

where $\rho_1 = \varepsilon_1(\alpha_1^* - \dot{y}_d) + (1/4)\varepsilon_1^2 \le \Psi_1$ is bounded with a positive constant Ψ_1.

Since W_1^* is unknown, the following critic and actor NNs are constructed to approximate the gradient term of optimal value function and the optimal virtual control respectively:

$$\frac{\partial \hat{V}_1^*(z_1)}{\partial z_1} = 2\beta_1 z_1 + \hat{W}_{1c}^T S_1(z_1) \tag{23}$$

$$\hat{\alpha}_1^* = -\beta_1 z_1 - \frac{1}{2}\hat{W}_{1a}^T S_1(z_1) \tag{24}$$

where $\hat{V}_1^*(z_1)$ and $\hat{\alpha}_1^*$ are the estimations of $V_1^*(z_1)$ and α_1^*, $\hat{W}_{1c}^T$ and $\hat{W}_{1a}^T$ are the critic and actor NN weights to estimate W_1^*, respectively.

Inserting (23) and (24) into (14), the approximated HJB equation can be derived as

$$\begin{aligned} H_1(z_1, \hat{\alpha}_1, \hat{W}_{1c}) &= z_1^2 + (\beta_1 z_1 + \frac{1}{2}\hat{W}_{1a}^T S_1(z_1))^2 \\ &\quad - (2\beta_1 z_1 + \hat{W}_{1c}^T S_1(z_1)) \\ &\quad \times (\beta_1 z_1 + \frac{1}{2}\hat{W}_{1a}^T S_1(z_1) + \dot{y}_d) \\ &= -(\beta_1^2 - 1)z_1^2 + \frac{1}{4}(\hat{W}_{1a}^T S_1(z_1))^2 \\ &\quad - 2\beta_1 z_1 \dot{y}_d + \sigma_1^T \hat{W}_{1c} \end{aligned} \tag{25}$$

where $\sigma_1 = -S_1(z_1)(\beta_1 z_1 + (1/2)\hat{W}_{1a}^T S_1(z_1) + \dot{y}_d)$.

According to (22) and (25), the Bellman residual error is yielded as

$$\begin{aligned} e_1 &= H_1(z_1, \hat{\alpha}_1, \hat{W}_{1c}) - H_1(z_1, \alpha_1^*, W_1^*) \\ &= H_1(z_1, \hat{\alpha}_1, \hat{W}_{1c}) \end{aligned} \tag{26}$$

Define a positive definite function with respect to the Bellman residual error as following:

$$E_1 = \frac{1}{2}e_1^2 \tag{27}$$

Based on the gradient descent algorithm, the following updating law for critic NN weight is yielded to minimize the Bellman residual error:

$$\begin{aligned} \dot{\hat{W}}_{1c} &= -\frac{\gamma_{1c}}{1 + \|\sigma_1\|^2}\frac{\partial \mathrm{E}_1}{\partial \hat{W}_{1c}} = -\frac{\gamma_{1c}}{1 + \|\sigma_1\|^2} e_1 \frac{\partial e_1}{\partial \hat{W}_{1c}} \\ &= -\frac{\gamma_{1c}}{1 + \|\sigma_1\|^2}\mathrm{H}_1(z_1, \hat{\alpha}_1, \hat{W}_{1c})\sigma_1 \\ &= -\frac{\gamma_{1c}}{1 + \|\sigma_1\|^2}\sigma_1(\sigma_1^T \hat{W}_{1c} \\ &\quad - (\beta_1^2 - 1)z_1^2 - 2\beta_1 z_1 \dot{y}_d + \frac{1}{4}\hat{W}_{1a}^T S_1(z_1) S_1^T(z_1)\hat{W}_{1a}) \end{aligned} \tag{28}$$

where the learning rate γ_{1c} is positive.

The updating law for actor NN weight is denoted as following:

$$\begin{aligned} \dot{\hat{W}}_{1a} = &\frac{1}{2}S_1(z_1)z_1 - \gamma_{1a}S_1(z_1)S_1^T(z_1)\hat{W}_{1a} \\ &+ \frac{\gamma_{1c}}{4(1 + \|\sigma_1\|^2)}S_1(z_1)S_1^T(z_1)\hat{W}_{1a}\sigma_1^T\hat{W}_{1c} \end{aligned} \tag{29}$$

where the learning rate γ_{1a} is positive.

Assumption 1 ([37] Persistence of Excitation (PE)): The signal of $\sigma_1\sigma_1^T$ should satisfy persistent excitation over the interval $[t, t + t_{\sigma 1}]$, with positive constants $\underline{k}_{\sigma 1}$, $\overline{k}_{\sigma 1}$, $t_{\sigma 1}$ for all t to satisfy:

$$\underline{k}_{\sigma 1} I_n \le \sigma_1\sigma_1^T \le \overline{k}_{\sigma 1} I_n \tag{30}$$

where $I_n \in R^{n\times n}$ is identity matrix.

Define the error variable $z_2 = x_2 - \hat{\alpha}_1^*$, the error dynamic (10) changes to

$$\dot{z}_1 = z_2 + \hat{\alpha}_1^* - \dot{y}_d \tag{31}$$

Design the positive definite Lyapunov function as following

$$L_1 = \frac{1}{2}z_1^2 + \frac{1}{2}\tilde{W}_{1a}^T\tilde{W}_{1a} + \frac{1}{2}\tilde{W}_{1c}^T\tilde{W}_{1c} \tag{32}$$

where $\tilde{W}_{1c} = \hat{W}_{1c} - W_1^*$, and $\tilde{W}_{1a} = \hat{W}_{1a} - W_1^*$. The time derivative of L_1 along (29), (30), and (32) is

$$\begin{aligned}\dot{L}_1 = {} & z_1(z_2 + \hat{\alpha}_1^* - \dot{y}_d) + \tilde{W}_{1a}^T \times (\frac{1}{2} S_1(z_1) z_1 \\ & - \gamma_{1a} S_1(z_1) S_1^T(z_1) \hat{W}_{1a} \\ & + \frac{\gamma_{1c}}{4(1 + \|\sigma_1\|^2)} S_1(z_1) S_1^T(z_1) \hat{W}_{1a} \sigma_1^T \hat{W}_{1c}) \\ & - \frac{\gamma_{1c}}{1 + \|\sigma_1\|^2} \tilde{W}_{1c}^T \sigma_1 \\ & \times (\sigma_1^T \hat{W}_{1c} - (\beta_1^2 - 1) z_1^2 - 2\beta_1 z_1 \dot{y}_d \\ & + \frac{1}{4} \hat{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a})\end{aligned} \tag{33}$$

Substituting (24) into (33), we have

$$\begin{aligned}\dot{L}_1 = {} & -\beta_1 z_1^2 + z_1 z_2 - z_1 \dot{y}_d - \frac{1}{2} z_1 \hat{W}_{1a}^T S_1(z_1) \\ & + \frac{1}{2} \tilde{W}_{1a}^T S_1(z_1) z_1 - \gamma_{1a} \tilde{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a} \\ & + \frac{\gamma_{1c}}{4(1 + \|\sigma_1\|^2)} \tilde{W}_{1a}^T S_1(z_1) S_1^T(z_1) \\ & \cdot \hat{W}_{1a} \sigma_1^T \hat{W}_{1c} - \frac{\gamma_{1c}}{1 + \|\sigma_1\|^2} \tilde{W}_{1c}^T \sigma_1 (\sigma_1^T \hat{W}_{1c} \\ & - (\beta_1^2 - 1) z_1^2 - 2\beta_1 z_1 \dot{y}_d \\ & + \frac{1}{4} \hat{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a})\end{aligned} \tag{34}$$

By applying $\tilde{W}_{1a} = \hat{W}_{1a} - W_1^*$, the following results yield:

$$\begin{aligned} & -\frac{1}{2} z_1 \hat{W}_{1a}^T S_1(z_1) + \frac{1}{2} \tilde{W}_{1a}^T S_1(z_1) z_1 \\ = {} & -\frac{1}{2} z_1 W_1^{*T} S_1(z_1) - \gamma_{1a} \tilde{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a} \\ = {} & -\frac{\gamma_{1a}}{2} \tilde{W}_{1a}^T S_1(z_1) S_1^T(z_1) \tilde{W}_{1a} \\ & - \frac{\gamma_{1a}}{2} \hat{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a} + \frac{\gamma_{1a}}{2} W_1^{*T} S_1(z_1) S_1^T(z_1) W_1^*\end{aligned} \tag{35}$$

Substituting above results into (34), one has

$$\begin{aligned}\dot{L}_1 = {} & -\beta_1 z_1^2 + z_1 z_2 - z_1 \dot{y}_d - \frac{1}{2} z_1 W_1^{*T} S_1(z_1) \\ & - \frac{\gamma_{1a}}{2} \tilde{W}_{1a}^T S_1(z_1) S_1^T(z_1) \tilde{W}_{1a} - \frac{\gamma_{1a}}{2} \hat{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a} \\ & + \frac{\gamma_{1a}}{2} W_1^{*T} S_1(z_1) S_1^T(z_1) W_1^* \\ & + \frac{\gamma_{1c}}{4(1 + \|\sigma_1\|^2)} \cdot \tilde{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a} \sigma_1^T \hat{W}_{1c} \\ & - \frac{\gamma_{1c}}{1 + \|\sigma_1\|^2} \tilde{W}_{1c}^T \sigma_1 (\sigma_1^T \hat{W}_{1c} - (\beta_1^2 - 1) z_1^2 \\ & - 2\beta_1 z_1 \dot{y}_d + \frac{1}{4} \hat{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a})\end{aligned} \tag{36}$$

The following facts can be obtained according to the Young's inequality:

$$\begin{aligned} z_1 z_2 \le {} & \frac{1}{2} z_1^2 + \frac{1}{2} z_2^2 \\ - z_1 \dot{y}_d \le {} & \frac{1}{2} z_1^2 + \frac{1}{2} \dot{y}_d^2 \\ - \frac{1}{2} z_1 W_1^{*T} S_1(z_1) \le {} & \frac{1}{4} z_1^2 + \frac{1}{4} (W_1^{*T} S_1(z_1))^2 \end{aligned} \tag{37}$$

Based on aforementioned inequalities, the equation (36) can be rewritten as:

$$\begin{aligned}\dot{L}_1 \le {} & \frac{1}{2} z_2^2 - (\beta_1 - \frac{5}{4}) z_1^2 - \frac{\gamma_{1a}}{2} \tilde{W}_{1a}^T S_1(z_1) S_1^T(z_1) \tilde{W}_{1a} \\ & - \frac{\gamma_{1a}}{2} \hat{W}_{1a}^T S_1(z_1) \cdot S_1^T(z_1) \hat{W}_{1a} \\ & + \frac{\gamma_{1c}}{4(1 + \|\sigma_1\|^2)} \tilde{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a} \sigma_1^T \hat{W}_{1c} \\ & - \frac{\gamma_{1c}}{1 + \|\sigma_1\|^2} \\ & \cdot \tilde{W}_{1c}^T \sigma_1 (\sigma_1^T \hat{W}_{1c} - (\beta_1^2 - 1) z_1^2 - 2\beta_1 z_1 \dot{y}_d \\ & + \frac{1}{4} \hat{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a}) \\ & + \frac{1}{2} \dot{y}_d^2 + \frac{1 + 2\gamma_{1a}}{4} (W_1^{*T} S_1(z_1))^2\end{aligned} \tag{38}$$

The following fact can be derived according to equation (22):

$$\begin{aligned} & -(\beta_1^2 - 1) z_1^2 - 2\beta_1 z_1 \dot{y}_d \\ = {} & -\sigma_1^T W_1^* - \frac{1}{2} \hat{W}_{1a}^T S_1(z_1) S_1^T(z_1) W_1^* \\ & + \frac{1}{4} W_1^{*T} S_1(z_1) S_1^T(z_1) W_1^* - \rho_1 \end{aligned} \tag{39}$$

Inserting (39) into (38) yields

$$\begin{aligned}\dot{L}_1 \le {} & \frac{1}{2} z_2^2 - (\beta_1 - \frac{5}{4}) z_1^2 - \frac{\gamma_{1a}}{2} \tilde{W}_{1a}^T S_1(z_1) S_1^T(z_1) \tilde{W}_{1a} \\ & - \frac{\gamma_{1a}}{2} \hat{W}_{1a}^T S_1(z_1) \cdot S_1^T(z_1) \hat{W}_{1a} \\ & + \frac{\gamma_{1a}}{4(1 + \|\sigma_1\|^2)} \tilde{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a} \sigma_1^T \hat{W}_{1c} \\ & - \frac{\gamma_{1c}}{1 + \|\sigma_1\|^2} \tilde{W}_{1c}^T \cdot \sigma_1 (\sigma_1^T \tilde{W}_{1c} - \frac{1}{2} \hat{W}_{1a}^T S_1(z_1) S_1^T(z_1) W_1^* \\ & + \frac{1}{4} W_1^{*T} S_1(z_1) S_1^T(z_1) W_1^* + \frac{1}{4} \hat{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a} - \rho_1) \\ & + \frac{1}{2} \dot{y}_d^2 + \frac{1 + 2\gamma_{1a}}{4} (W_1^{*T} S_1(z_1))^2\end{aligned} \tag{40}$$

Considering the following facts:

$$\begin{aligned} & -\frac{1}{2} \hat{W}_{1a}^T S_1(z_1) S_1^T(z_1) W_1^* + \frac{1}{4} W_1^{*T} S_1(z_1) S_1^T(z_1) W_1^* \\ & + \frac{1}{4} \hat{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a} = \frac{1}{4} \tilde{W}_{1a}^T S_1(z_1) S_1^T(z_1) \hat{W}_{1a} \\ & - \frac{1}{4} W_1^{*T} S_1(z_1) S_1^T(z_1) \tilde{W}_{1a} \end{aligned} \tag{41}$$

$$\frac{\gamma_{1c}}{1+\|\sigma_1\|^2}\tilde{W}_{1c}^T\sigma_1\rho_1 \le \frac{\gamma_{1c}}{2(1+\|\sigma_1\|^2)}\rho_1^2 + \frac{\gamma_{1c}}{2(1+\|\sigma_1\|^2)}\tilde{W}_{1c}^T\sigma_1\sigma_1^T\tilde{W}_{1c} \tag{42}$$

The inequality (40) can be expressed as

$$\begin{aligned}\dot{L}_1 \le{}& \frac{1}{2}z_2^2 - (\beta_1 - \frac{5}{4})z_1^2 - \frac{\gamma_{1a}}{2}\tilde{W}_{1a}^T S_1(z_1)S_1^T(z_1)\tilde{W}_{1a}\\ &- \frac{\gamma_{1a}}{2}\hat{W}_{1a}^T\cdot S_1(z_1)S_1^T(z_1)\hat{W}_{1a}\\ &+ \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1a}^T S_1(z_1)S_1^T(z_1)\hat{W}_{1a}\sigma_1^T\hat{W}_{1c}\\ &- \frac{\gamma_{1c}}{1+\|\sigma_1\|^2}\tilde{W}_{1c}^T\sigma_1\sigma_1^T\tilde{W}_{1c}\\ &- \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1c}^T\sigma_1\tilde{W}_{1a}^T S_1(z_1)S_1^T(z_1)\hat{W}_{1a}\\ &+ \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1c}^T\sigma_1 W_1^{*T} S_1(z_1)S_1^T(z_1)\tilde{W}_{1a}\\ &+ \frac{1+2\gamma_{1a}}{4}(W_1^{*T}S_1(z_1))^2 + \frac{\gamma_{1c}}{2(1+\|\sigma_1\|^2)}\rho_1^2\\ &+ \frac{\gamma_{1c}}{2(1+\|\sigma_1\|^2)}\tilde{W}_{1c}^T\sigma_1\sigma_1^T\tilde{W}_{1c} + \frac{1}{2}\dot{y}_d^2\\ ={}& \frac{1}{2}z_2^2 - (\beta_1\\ &- \frac{5}{4})z_1^2 - \frac{\gamma_{1a}}{2}\tilde{W}_{1a}^T S_1(z_1)S_1^T(z_1)\tilde{W}_{1a}\\ &- \frac{\gamma_{1c}}{2(1+\|\sigma_1\|^2)}\cdot\tilde{W}_{1c}^T\sigma_1\sigma_1^T\tilde{W}_{1c}\\ &- \frac{\gamma_{1a}}{2}\hat{W}_{1a}^T S_1(z_1)S_1^T(z_1)\hat{W}_{1a}\\ &+ \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1a}^T S_1(z_1)\cdot S_1^T(z_1)\hat{W}_{1a}\sigma_1^T\hat{W}_{1c}\\ &- \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1c}^T\sigma_1\tilde{W}_{1a}^T S_1(z_1)S_1^T(z_1)\hat{W}_{1a}\\ &+ \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1c}^T\sigma_1 W_1^{*T} S_1(z_1)S_1^T(z_1)\tilde{W}_{1a}\\ &+ \frac{1+2\gamma_{1a}}{4}(W_1^{*T}S_1(z_1))^2 + \frac{\gamma_{1c}}{2(1+\|\sigma_1\|^2)}\rho_1^2 + \frac{1}{2}\dot{y}_d^2\end{aligned} \tag{43}$$

Based on the following conclusion:

$$\begin{aligned}&\frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1a}^T S_1(z_1)S_1^T(z_1)\hat{W}_{1a}\sigma_1^T\hat{W}_{1c}\\ &- \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1c}^T\sigma_1\tilde{W}_{1a}^T S_1(z_1)\cdot S_1^T(z_1)\hat{W}_{1a}\\ ={}& \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1a}^T S_1(z_1)\hat{W}_{1c}^T\sigma_1 S_1^T(z_1)\hat{W}_{1a}\\ &- \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\cdot\tilde{W}_{1a}^T S_1(z_1)\tilde{W}_{1c}^T\sigma_1 S_1^T(z_1)\hat{W}_{1a}\\ ={}& \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1a}^T S_1(z_1)W_1^{*T}\sigma_1 S_1^T(z_1)\hat{W}_{1a}\end{aligned} \tag{44}$$

The inequality (43) becomes:

$$\begin{aligned}\dot{L}_1 \le{}& \frac{1}{2}z_2^2 - (\beta_1 - \frac{5}{4})z_1^2 - \frac{\gamma_{1a}}{2}\tilde{W}_{1a}^T S_1(z_1)S_1^T(z_1)\tilde{W}_{1a}\\ &- \frac{\gamma_{1c}}{2(1+\|\sigma_1\|^2)}\tilde{W}_{1c}^T\sigma_1\sigma_1^T\tilde{W}_{1c} - \frac{\gamma_{1a}}{2}\hat{W}_{1a}^T S_1(z_1)S_1^T(z_1)\hat{W}_{1a}\\ &+ \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1a}^T S_1(z_1)W_1^{*T}\sigma_1 S_1^T(z_1)\hat{W}_{1a}\\ &+ \frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1c}^T\sigma_1 W_1^{*T} S_1(z_1)S_1^T(z_1)\tilde{W}_{1a}\\ &+ \frac{1+2\gamma_{1a}}{4}(W_1^{*T}S_1(z_1))^2 + \frac{\gamma_{1c}}{2(1+\|\sigma_1\|^2)}\rho_1^2 + \frac{1}{2}\dot{y}_d^2\end{aligned} \tag{45}$$

Following inequalities yield by using Young's inequality:

$$\begin{aligned}&\frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1a}^T S_1(z_1)W_1^{*T}\sigma_1 S_1^T(z_1)\hat{W}_{1a}\\ &\le \frac{1}{32}\tilde{W}_{1a}^T S_1(z_1)W_1^{*T}\sigma_1\sigma_1^T W_1^* S_1^T(z_1)\tilde{W}_{1a}\\ &\quad + \frac{\gamma_{1c}^2}{2}\hat{W}_{1a}^T S_1(z_1)S_1^T(z_1)\hat{W}_{1a}\end{aligned} \tag{46}$$

$$\begin{aligned}&\frac{\gamma_{1c}}{4(1+\|\sigma_1\|^2)}\tilde{W}_{1c}^T\sigma_1 W_1^{*T} S_1(z_1)S_1^T(z_1)\tilde{W}_{1a}\\ &\le \frac{1}{32(1+\|\sigma_1\|^2)}\cdot\tilde{W}_{1c}^T\sigma_1 W_1^{*T} S_1(z_1)S_1^T(z_1)W_1^*\sigma_1^T\tilde{W}_{1c}\\ &\quad + \frac{\gamma_{1c}^2}{2}\tilde{W}_{1a}^T S_1(z_1)S_1^T(z_1)\tilde{W}_{1a}\end{aligned} \tag{47}$$

Adding above inequalities into (45), we have

$$\begin{aligned}\dot{L}_1 \le{}& \frac{1}{2}z_2^2 - (\beta_1 - \frac{5}{4})z_1^2 - (\frac{\gamma_{1a}}{2}\\ &- \frac{\gamma_{1c}^2}{2} - \frac{1}{32}W_1^{*T}\sigma_1\sigma_1^T W_1^*)\cdot\tilde{W}_{1a}^T S_1(z_1)S_1^T(z_1)\tilde{W}_{1a}\\ &- \frac{1}{1+\|\sigma_1\|^2}(\frac{\gamma_{1c}}{2} - \frac{1}{32}W_1^{*T}S_1(z_1)\\ &\cdot S_1^T(z_1)W_1^*)\tilde{W}_{1c}^T\sigma_1\sigma_1^T\tilde{W}_{1c}\\ &- (\frac{\gamma_{1a}}{2} - \frac{\gamma_{1c}^2}{2})\hat{W}_{1a}^T S_1(z_1)S_1^T(z_1)\hat{W}_{1a}\\ &+ \frac{1+2\gamma_{1a}}{4}(W_1^{*T}S_1(z_1))^2 + \frac{\gamma_{1c}}{2(1+\|\sigma_1\|^2)}\rho_1^2 + \frac{1}{2}\dot{y}_d^2\end{aligned} \tag{48}$$

Rewrite above inequality to following compact form

$$\begin{aligned}\dot{L}_1 \le{}& -\xi_1^T A_1\xi_1 + C_1 + \frac{1}{2}z_2^2\\ &- (\frac{\gamma_{1a}}{2} - \frac{\gamma_{1c}^2}{2})\hat{W}_{1a}^T S_1(z_1)S_1^T(z_1)\hat{W}_{1a}\end{aligned} \tag{49}$$

where

$$\xi_1 = [z_1, \tilde{W}_{1a}^T, \tilde{W}_{1c}^T]^T$$

$$A_1 = \begin{bmatrix}\beta_1 - \frac{5}{4} & 0 & 0\\ 0 & a22 & 0\\ 0 & 0 & a33\end{bmatrix}$$

$$a22 = (\frac{\gamma_{1a}}{2} - \frac{\gamma_{1c}^2}{2} - \frac{1}{32}W_1^{*T}\sigma_1\sigma_1^T W_1^*)S_1(z_1)S_1^T(z_1)$$
$$a33 = \frac{1}{1+\|\sigma_1\|^2}(\frac{\gamma_{1c}}{2} - \frac{1}{32}W_1^{*T}S_1(z_1)S_1^T(z_1)W_1^*)\sigma_1\sigma_1^T$$
$$C_1 = \frac{1+2\gamma_{1a}}{4}(W_1^{*T}S_1(z_1))^2 + \frac{\gamma_{1c}}{2(1+\|\sigma_1\|^2)}\rho_1^2 + \frac{1}{2}\dot{y}_d^2$$

According to assumption 1, the above matrix A_1 can be positive definite by choosing appropriate design parameters under following conditions:

$$\begin{aligned}\beta_1 &> \frac{5}{4}\\ \gamma_{1a} &> \gamma_{1c}^2 + \frac{\overline{k}_{\sigma 1}}{16}W_1^{*T}W_1^*\\ \gamma_{1c} &> \frac{1}{16}\sup_{t\geq 0}\left\{W_1^{*T}S_1(z_1)S_1^T(z_1)W_1^*\right\}\end{aligned} \tag{50}$$

Remark 3: Above requirements listed in (50) for design parameters are only for theoretical analysis to prove the existence of corresponding parameters. The actual value may be allowed to be smaller than above limits, especially for the parameter β_1. Better tracking performance may be obtained by setting smaller value for β_1 due to the huge inertia of ship. Based on above analysis and trial and error, the value of β_1 should be small enough to resist the overshoot, otherwise, the actual trajectory will fluctuate sharply around the reference signal.

Then (49) can be further rewritten as:

$$\dot{L}_1 \leq \frac{1}{2}z_2^2 - a_1\|\xi_1\|^2 + c_1 \tag{51}$$

where $a_1 = \inf_{t\geq 0}\{\lambda_{\min}\{A_1\}\}$, $c_1 = \sup_{t\geq 0}\{C_1\}$.

Step 2: Differentiate the error variable z_2 and obtain

$$\dot{z}_2 = \dot{x}_2 - \dot{\hat{\alpha}}_1^* = f_2(x_2) + v - \dot{\hat{\alpha}}_1^* \tag{52}$$

The optimal cost function is defined as

$$V_2^*(z_2) = \min_{v\in\Psi(\Omega_v)}(\int_t^\infty r_2(z_2, v)ds) = \int_t^\infty r_2(z_2, v^*)ds \tag{53}$$

where $r_2(z_2, v) = z_2^2 + v^2$, Ω_v is a compact set, v^* is the corresponding optimal control without considering saturation. Next, the HJB equation is constructed as

$$\begin{aligned}H_2(z_2, v^*, \frac{\partial V_2^*}{\partial z_2}) &= z_2^2 + v^{*2} + \frac{\partial V_2^*}{\partial z_2}\dot{z}_2 = z_2^2 + v^{*2}\\ &+ \frac{\partial V_2^*}{\partial z_2}(f_2(x_2) + v^* - \dot{\hat{\alpha}}_1^*) = 0\end{aligned} \tag{54}$$

The optimal control v^* can be calculated by solving $(\partial H_2/\partial v^*) = 0$.

$$v^* = -\frac{1}{2}\frac{\partial V_2^*}{\partial z_2} \tag{55}$$

Decompose the optimal cost function as

$$\begin{aligned}V_2^*(z_2) &= \beta_2 z_2^2 - \int_0^\infty 2e\dot{z}_2 dt - \beta_2 z_2^2 + \int_0^\infty 2e\dot{z}_2 dt\\ &+ V_2^*(z_2) = \beta_2 z_2^2 - \int_0^\infty 2e\dot{z}_2 dt + V_2^o(z_2)\end{aligned} \tag{56}$$

where β_2 is a positive constant to be designed, and $V_2^o(z_2) = -\beta_2 z_2^2 + \int_0^\infty 2e\dot{z}_2 dt + V_2^*(z_2)$.

Substituting (56) into (55), the optimal control can be transformed into following form:

$$\begin{aligned}\frac{\partial V_2^*}{\partial z_2} &= 2\beta_2 z_2 - 2e + \frac{\partial V_2^o(z_2)}{\partial z_2}\\ v^* &= -\beta_2 z_2 + e - \frac{1}{2}\frac{\partial V_2^o(z_2)}{\partial z_2}\end{aligned} \tag{57}$$

Approximate the uncertain term $\partial V_2^o(z_2)/\partial z_2$ by NNs as

$$\frac{\partial V_2^o(z_2)}{\partial z_2} = W_2^{*T}S_2(z_2) + \varepsilon_2(z_2) \tag{58}$$

where $W_2^* \in R^m$ is the ideal constant weight; and $S_2(z_2) \in R^m$ is the Gaussian function vector; $\varepsilon_2(z_2) \in R$ is approximation error.

The equation (57) can be rewritten as

$$\frac{\partial V_2^*(z_2)}{\partial z_2} = 2\beta_2 z_2 - 2e + W_2^{*T}S_2(z_2) + \varepsilon_2(z_2) \tag{59}$$

$$v^* = -\beta_2 z_2 + e - \frac{1}{2}(W_2^{*T}S_2(z_2) + \varepsilon_2(z_2)) \tag{60}$$

where ε_2 is bounded and satisfies $\|\varepsilon_2\| \leq \delta_2$, with the positive constant δ_2.

Substituting (52), (59) and (60) into (54), we have

$$\begin{aligned}&H_2(z_2, v^*, W_2^*) = z_2^2 + (-\beta_2 z_2 + e - \frac{1}{2}W_2^{*T}S_2(z_2)\\ &- \frac{1}{2}\varepsilon_2(z_2))v^* + (2\beta_2 z_2 - 2e + W_2^{*T}S_2(z_2) + \varepsilon_2(z_2))\\ &\cdot(f_2(x_2) + v^* - \dot{\hat{\alpha}}_1^*) = z_2^2 - \beta_2^2 z_2^2 - e^2\\ &- \frac{1}{4}(W_2^{*T}S_2(z_2))^2 - \frac{1}{4}\varepsilon_2(z_2)^2 + \beta_2 z_2 e - \frac{1}{2}\beta_2 z_2 W_2^{*T}\\ &S_2(z_2) - \frac{1}{2}\beta_2 z_2\varepsilon_2(z_2) + \beta_2 z_2 e + \frac{1}{2}W_2^{*T}S_2(z_2)e\\ &+ \frac{1}{2}\varepsilon_2(z_2)e - \frac{1}{2}\beta_2 z_2 W_2^{*T}S_2(z_2) + \frac{1}{2}W_2^{*T}S_2(z_2)e\\ &- \frac{1}{4}W_2^{*T}S_2(z_2)\varepsilon_2(z_2) - \frac{1}{2}\beta_2 z_2\varepsilon_2(z_2) + \frac{1}{2}\varepsilon_2(z_2)e\\ &- \frac{1}{4}W_2^{*T}S_2(z_2)\varepsilon_2(z_2) + (2\beta_2 z_2 - 2e + W_2^{*T}S_2(z_2)\\ &+ \varepsilon_2(z_2))(f_2(x_2) - \dot{\hat{\alpha}}_1^*) = -(\beta_2^2 - 1)z_2^2 + 2\beta_2 z_2\\ &(f_2(x_2) - \dot{\hat{\alpha}}_1^*) + W_2^{*T}S_2(z_2)(f_2(x_2) - \dot{\hat{\alpha}}_1^* - \beta_2 z_2 + e)\\ &- \frac{1}{4}(W_2^{*T}S_2(z_2))^2 + \varepsilon_2(z_2)(f_2(x_2) - \dot{\hat{\alpha}}_1^* + v^*)\\ &+ \frac{1}{4}\varepsilon_2(z_2)^2 + e(2\beta_2 z_2 - e - 2f_2(x_2) + 2\dot{\hat{\alpha}}_1^*)\\ &= -(\beta_2^2 - 1)z_2^2 + 2\beta_2 z_2(f_2(x_2) - \dot{\hat{\alpha}}_1^*) + W_2^{*T}S_2(z_2)\\ &(f_2(x_2) - \dot{\hat{\alpha}}_1^* - \beta_2 z_2 + e) - \frac{1}{4}(W_2^{*T}S_2(z_2))^2 + \rho_2(t)\\ &+ e(2\beta_2 z_2 - e - 2f_2(x_2) + 2\dot{\hat{\alpha}}_1^*) = 0\end{aligned} \tag{61}$$

where $\rho_2 = \varepsilon_2(f_2(x_2) - \dot{\hat{\alpha}}_1^* + v^*) + (1/4)\varepsilon_2^2$. Due to all terms included in ρ_2 are bounded, it is obviously that ρ_2 is also bounded, i.e., $|\rho_2| \leq \Psi_2$.

Considering the ideal weight W_2^* is not known, the gradient term of optimal value function and optimal control can be approximated by employing critic and actor NN respectively.

$$\frac{\partial \hat{V}_2^*(z_2)}{\partial z_2} = 2\beta_2 z_2 - 2e + \hat{W}_{2c}^T S_2(z_2) \tag{62}$$

$$v = -\beta_2 z_2 + e - \frac{1}{2}\hat{W}_{2a}^T S_2(z_2) \tag{63}$$

where $\hat{V}_1^*(z_1)$ and v are the estimations of $V_1^*(z_1)$ and v^*, $\hat{W}_{1c}^T$ and $\hat{W}_{1a}^T$ are the critic and actor NN weights to estimate W_1^*, respectively.

Inserting (62) and (63) into (61), the approximated HJB equation can be derived as

$$\begin{aligned} H_2(z_2, v, \hat{W}_{2c}) &= z_2^2 + (-\beta_2 z_2 \\ &+ e - \frac{1}{2}\hat{W}_{2a}^T S_2(z_2))^2 \\ &+ (2\beta_2 z_2 - 2e + \hat{W}_{2c}^T S_2(z_2)) \times (f_2(x_2) - \dot{\hat{\alpha}}_1^* - \beta_2 z_2 \\ &+ e - \frac{1}{2}\hat{W}_{2a}^T S_2(z_2)) = -(\beta_2^2 - 1)z_2^2 + 2\beta_2 z_2 (f_2(x_2) \\ &- \dot{\hat{\alpha}}_1^*) + \frac{1}{4}\hat{W}_{2a}^T S_2(z_2))^2 + \hat{W}_{2c}^T \sigma_2 \\ &+ e\left[2\beta_2 z_2 - e - 2(f_2(x_2) - \dot{\hat{\alpha}}_1^*)\right] \end{aligned} \tag{64}$$

where

$$\sigma_2 = S_2(z_2)(f_2(x_2) - \dot{\hat{\alpha}}_1^* - \beta_2 z_2 + e - \frac{1}{2}\hat{W}_{2a}^T S_2(z_2)) \in \mathrm{R}^{\mathrm{m}}$$

Similar with step 1, introduce a positive definite function as $E_2 = (1/2)e_2^2$, then construct the updating law for the critic NN weight by using gradient descent algorithm:

$$\begin{aligned} \dot{\hat{W}}_{2c} &= -\frac{\gamma_{2c}}{1 + \|\sigma_2\|^2} e_2 \frac{\partial e_2}{\partial \hat{W}_{2c}} \\ &= -\frac{\gamma_{2c}}{1 + \|\sigma_2\|^2}\sigma_2 \\ &\quad \left\{\sigma_2^T \hat{W}_{2c} - (\beta_2^2 - 1)z_2^2 + 2\beta_2 z_2 (f_1(x_2) - \dot{\hat{\alpha}}_1^*)\right. \\ &\quad + \frac{1}{4}\hat{W}_{2a}^T S_2(z_2) S_2^T(z_2)\hat{W}_{2a} \\ &\quad \left. + e\left[2\beta_2 z_2 - e - 2(f_2(x_2) - \dot{\hat{\alpha}}_1^*)\right]\right\} \end{aligned} \tag{65}$$

where the learning rate γ_{2c} is positive.

The updating law for actor NN weights

$$\begin{aligned} \dot{\hat{W}}_{2a} &= \frac{1}{2}S_2(z_2)z_2 - \gamma_{2a} S_2(z_2) S_2^T(z_2)\hat{W}_{2a} \\ &\quad + \frac{\gamma_{2c}}{4(1 + \|\sigma_2\|^2)} S_2(z_2) S_2^T(z_2)\hat{W}_{2a}\sigma_2^T \hat{W}_{2c} \end{aligned} \tag{66}$$

where the learning rate γ_{2a} is positive.

Construct the Lyapunov function for the whole system as follows:

$$L(t) = L_1 + \frac{1}{2}z_2^2 + \frac{1}{2}\tilde{W}_{2a}^T \tilde{W}_{2a} + \frac{1}{2}\tilde{W}_{2c}^T \tilde{W}_{2c} + \frac{1}{2}e^2 \tag{67}$$

where $\tilde{W}_{2c} = \hat{W}_{2c} - W_2^*$ and $\tilde{W}_{2a} = \hat{W}_{2a} - W_2^*$ are approximation errors for the critic and actor NN weights respectively. The e is an auxiliary design system to handle the practical problem of input saturation, it has following form:

$$\dot{e} = \begin{cases} -ke - \dfrac{|z_2 \Delta v| + 0.5\Delta v^2}{e^2} e + \Delta v, & |e| \geq \varepsilon \\ 0, & |e| < \varepsilon \end{cases} \tag{68}$$

The following results can be obtained:

$$e\dot{e} = \begin{cases} -ke^2 - |z_2 \Delta v| - \dfrac{1}{2}\Delta v^2 + e\Delta v \\ 0 \end{cases} \tag{69}$$

By using the Young's inequality, we have

$$e\Delta v \leq \frac{1}{2}e^2 + \frac{1}{2}\Delta v^2 \tag{70}$$

Expression (69) can be rewritten as

$$\begin{aligned} e\dot{e} &\leq \begin{cases} -ke^2 - |z_2 \Delta v| + \frac{1}{2}e^2 \\ 0 \end{cases} \\ &= \begin{cases} -(k - \frac{1}{2})e^2 - |z_2 \Delta v| \\ 0 \end{cases} \end{aligned} \tag{71}$$

Differentiate $L(t)$ along (52), (65), and (66), then obtain following result:

$$\begin{aligned} \dot{L}(t) &= \dot{L}_1 + z_2(f_2(x_2) - \dot{\hat{\alpha}}_1^* + v + \Delta v) + \tilde{W}_{2a}^T(\frac{1}{2}S_2(z_2)z_2 \\ &\quad - \gamma_{2a} S_2(z_2) S_2^T(z_2)\hat{W}_{2a} + \frac{\gamma_{2c}}{4(1 + \|\sigma_2\|^2)} S_2(z_2) S_2^T(z_2) \\ &\quad \cdot \hat{W}_{2a}\sigma_2^T \hat{W}_{2c}) - \frac{\gamma_{2c}}{1 + \|\sigma_2\|^2}\tilde{W}_{2c}^T \sigma_2(\sigma_2^T \hat{W}_{2c} - (\beta_2^2 - 1)z_2^2 \\ &\quad + 2\beta_2 z_2 (f_2(x_2) - \dot{\hat{\alpha}}_1^*) + \frac{1}{4}\hat{W}_{2a}^T S_2(z_2) S_2^T(z_2)\hat{W}_{2a} \\ &\quad + e(2\beta_2 z_2 - e - 2f_2(x_2) + 2\dot{\hat{\alpha}}_1^*)) + e\dot{e} \end{aligned} \tag{72}$$

Similar with the process from (34) to (48) in step 1, the following inequality can be ultimately derived.

$$\begin{aligned} \dot{L}(t) &\leq \dot{L}_1 - (\beta_2 - \frac{5}{4})z_2^2 \\ &\quad + \frac{1}{2}f_2(x_2)^2 + \frac{1 + 2\gamma_{2a}}{4}(W_2^{*T} S_2(z_2))^2 \\ &\quad + z_2 e + \frac{1}{2}\dot{\hat{\alpha}}_1^{*2} - (\frac{\gamma_{2a}}{2} - \frac{\gamma_{2c}^2}{2} \\ &\quad - \frac{1}{32}W_2^{*T}\sigma_2\sigma_2^T W_2^*)\tilde{W}_{2a}^T S_2(z_2) \\ &\quad \cdot S_2^T(z_2)\tilde{W}_{2a} - \frac{1}{1 + \|\sigma_2\|^2}(\frac{\gamma_{2c}}{2} - \frac{1}{32}W_2^{*T} S_2 S_2^T W_2^*) \\ &\quad \cdot \tilde{W}_{2c}^T \sigma_2\sigma_2^T \tilde{W}_{2c} - (\frac{\gamma_{2a}}{2} - \frac{\gamma_{2c}^2}{2})\hat{W}_{2a}^T S_2(z_2) S_2^T(z_2)\hat{W}_{2a} \\ &\quad + \frac{\gamma_{2c}}{2(1 + \|\sigma_2\|^2)}\rho_2^2 - (k - \frac{1}{2})e^2 \end{aligned} \tag{73}$$

The following results can be easily obtained according to the Young's inequality:

$$\frac{\gamma_{2c}}{1+\|\sigma_2\|^2}\tilde{W}_{2c}^T\sigma_2 2ef_2(x_2) \le \gamma_{2c}(f_2(x_2)\tilde{W}_{2c}^T\sigma_2)^2+\gamma_{2c}e^2 \tag{74}$$

$$-\frac{\gamma_{2c}}{1+\|\sigma_2\|^2}\tilde{W}_{2c}^T\sigma_2 2e\dot{\hat{\alpha}}_1^* \le \gamma_{2c}(\dot{\hat{\alpha}}_1^*\tilde{W}_{2c}^T\sigma_2)^2+\gamma_{2c}e^2 \tag{75}$$

The inequality (73) can be further transformed into the following form:

$$\begin{aligned}\dot{L}(t) &\le \dot{L}_1-(\beta_2-\frac{7}{4})z_2^2+\frac{1}{2}f_2(x_2)^2\\ &+\frac{1+2\gamma_{2a}}{4}(W_2^{*T}S_2(z_2))^2+\frac{1}{2}\dot{\hat{\alpha}}_1^{*2}\\ &-(\frac{\gamma_{2a}}{2}-\frac{\gamma_{2c}^2}{2}-\frac{1}{32}W_2^{*T}\sigma_2\sigma_2^TW_2^*)\tilde{W}_{2a}^TS_2(z_2)S_2^T(z_2)\tilde{W}_{2a}\\ &-(\frac{\gamma_{2c}}{2}-\frac{1}{32}W_2^{*T}S_2S_2^TW_2^*)\tilde{W}_{2c}^T\sigma_2\sigma_2^T\tilde{W}_{2c}\\ &-(\frac{\gamma_{2a}}{2}-\frac{\gamma_{2c}^2}{2})\hat{W}_{2a}^TS_2(z_2)S_2^T(z_2)\hat{W}_{2a}\\ &+\frac{\gamma_{2c}}{2(1+\|\sigma_2\|^2)}\rho_2^2-(k-1)e^2\\ &\le \dot{L}_1-(\beta_2-\frac{7}{4})z_2^2-(\frac{\gamma_{2a}}{2}-\frac{\gamma_{2c}^2}{2}\\ &-\frac{1}{32}W_2^{*T}\sigma_2\sigma_2^TW_2^*)\tilde{W}_{2a}^TS_2(z_2)\\ &S_2^T(z_2)\tilde{W}_{2a}-(\frac{\gamma_{2c}}{2}-\frac{1}{32}W_2^{*T}S_2S_2^TW_2^*)\tilde{W}_{2c}^T\sigma_2\sigma_2^T\tilde{W}_{2c}\\ &+\frac{\gamma_{2c}}{2(1+\|\sigma_2\|^2)}\rho_2^2+\frac{1}{2}f_2(x_2)^2\\ &+\frac{1}{2}\dot{\hat{\alpha}}_1^{*2}+\frac{1+2\gamma_{2a}}{4}(W_2^{*T}S_2(z_2))^2-(k-1)e^2\end{aligned} \tag{76}$$

Based on previous results, (76) can be denoted in a compact form as

$$\dot{L}(t) \le -a_1\|\xi_1\|^2+c_1-\xi_2^TA_2\xi_2+C_2 \tag{77}$$

where

$$\xi_2=[z_2,\tilde{W}_{2a}^T,\tilde{W}_{2c}^T,e]^T$$

$$A_2=\begin{bmatrix}\beta_2-\frac{9}{4} & 0 & 0 & 0\\ 0 & b22 & 0 & 0\\ 0 & 0 & b33 & 0\\ 0 & 0 & 0 & k-1\end{bmatrix}$$

$$b22=(\frac{\gamma_{2a}}{2}-\frac{\gamma_{2c}^2}{2}-\frac{1}{32}W_2^{*T}\sigma_2\sigma_2W_2^*)S_2(z_2)S_2^T(z_2)$$

$$b33=(\frac{\gamma_{2c}}{2}-\frac{1}{32}W_2^{*T}S_2(z_2)S_2^T(z_2)W_2^*)\sigma_2\sigma_2^T$$

$$C_2=\frac{1+2\gamma_{2a}}{4}W_2^{*T}S_2(z_2)S_2^T(z_2)W_2^*+\frac{1}{2}f_2^2(x_2)+\frac{\gamma_{2c}}{2(1+\|\sigma_2\|^2)}\rho_2^2+\frac{1}{2}\dot{\hat{\alpha}}_1^{*2}$$

The matrix A_2 can be guaranteed positive definite by adjusting the design parameters to satisfy the following conditions:

$$\begin{aligned}\beta_2 &> \frac{9}{4}\\ \gamma_{2a} &> \gamma_{2c}^2+\frac{\bar{k}_{\sigma 2}}{16}W_2^{*T}W_2^*\\ \gamma_{2c} &> \frac{1}{16}\sup_{t\ge 0}\left\{W_2^{*T}S_2(z_2)S_2^T(z_2)W_2^*\right\}\end{aligned} \tag{78}$$

where $\underline{k}_{\sigma 2}$ are $\bar{k}_{\sigma 2}$ are positive constants defined according to assumption 1 to satisfy the PE condition and $\underline{k}_{\sigma 2}I_3 \le \sigma_2\sigma_2^T \le \bar{k}_{\sigma 2}I_3$ for all t over the interval $[t, t+t_{\sigma 2}]$, $t_{\sigma 2}>0$.

Then (77) can become:

$$\dot{L}(t) \le -a_1\|\xi_1\|^2-a_2\|\xi_2\|^2+c_1+c_2 \tag{79}$$

where $a_2=\inf_{t\ge 0}\{\lambda_{\min}\{A_2\}\}$, $c_2=\sup_{t\ge 0}\{C_2\}$.

Based on above main results, the following theorem can be deduced.

Theorem 1: Considering the ship steering model (1) with bounded initial condition and desired course signals, the OB control scheme uses the critic and actor NNs weight updating laws (28), (29) for the virtual control (24), and (65), (66) for the actual control (4), (63), and the design parameters satisfy (50), (78), and PE conditions(Assumption 1) are satisfied, then:

1) all error signals of the OB control are semi-globally uniformly ultimately bounded (SGUUB);

2) the ship can track the reference course signal to desired accuracy.

Proof :

Taking $a=\min\{a_1,a_2\}$ and $c=c_1+c_2$, then (79) can be expressed as

$$\dot{L}(t) \le -aL(t)+c \tag{80}$$

Based on Lemma 1, the following inequality can be derived directly:

$$L(t) \le \mathrm{e}^{-at}L(0)+\frac{c}{a}(1-\mathrm{e}^{-at}) \tag{81}$$

It can be concluded obviously that all error signals, z_1, z_2, $\tilde{W}_{1a}$, $\tilde{W}_{1c}$, $\tilde{W}_{2a}$, $\tilde{W}_{2c}$, e included in L(t) are SGUUB.

Let $L_z(t)=\frac{1}{2}z_1^2+\frac{1}{2}z_2^2$, its time derivative along (31) and (52) is

$$\dot{L}_z(t)=z_1(z_2+\hat{\alpha}_1^*-\dot{y}_d)+z_2(f_2(x_2)+v+\Delta u-\dot{\hat{\alpha}}_1^*) \tag{82}$$

Substituting (24) and (63) into (82), we have

$$\begin{aligned}\dot{L}_z(t) &= -\beta_1z_1^2-\frac{1}{2}z_1S_1^T(z_1)\hat{W}_{1a}+z_1z_2-z_1\dot{y}_d+z_2f_2(x_2)\\ &-\beta_2z_2^2+z_2e-\frac{1}{2}z_2S_2^T(z_2)\hat{W}_{2a}+z_2\Delta u-z_2\dot{\hat{\alpha}}_1^*\end{aligned} \tag{83}$$

By applying the Young's inequality, one has following inequality

$$\dot{L}_z(t) \le -(\beta_1-\frac{5}{4})z_1^2-(\beta_2-\frac{11}{4})z_2^2+P(t) \tag{84}$$

where

$$P(t) = 1/2\dot{y}_d^2 + 1/2\dot{\hat{\alpha}}_1^{*2} + 1/2(f_2(x_2))^2 + 1/2e^2 + 1/2\Delta u^2 + 1/4\left\|S_1^T(z_1)\hat{W}_{1a}\right\|^2 + 1/4\left\|S_2^T(z_2)\hat{W}_{2a}\right\|^2.$$

Because$\tilde{W}_{1a}$, $\tilde{W}_{2a}$, e are SGUUB, $S_1^T(z_1)\hat{W}_{1a}$, $S_2^T(z_2)\hat{W}_{2a}$ and Δu are bounded. Since each term in $P(t)$ is bounded, it can be easily deduced that $P(t)$ is also bounded by a positive constant ζ, i.e., $|P(t)| < \zeta$.

Thus, the following inequality in compact form holds:

$$\dot{L}_z(t) \leq -\beta L_z(t) + \zeta \tag{85}$$

where $\beta = \min\left\{2(\beta_1 - 5/4), 2(\beta_2 - 11/4)\right\}$.

By applying Lemma 1 again, we have

$$L_z(t) \leq \mathrm{e}^{-\beta t}L_z(0) + \frac{\zeta}{\beta}(1 - \mathrm{e}^{-\beta t}) \tag{86}$$

Based on the above conclusion, the system tracking errors can be limited to arbitrary small extent by choosing appropriate β, it implies that the ship can follow the reference course signal to desired accuracy.

IV. SIMULATION AND COMPARISO

In this part, to further illustrate the effectiveness and tracking performance of the controller proposed in Section III, M.V. "YU LONG" is taken as the simulation plant of interest with following ship particulars: length between perpendiculars (L_{PP}) 126m, moulded breadth (B) 20.8m, summer draught 8.0m, block coefficient 0.681, forward speed 7.72m/s. According to above ship particulars, the ship nonlinear motion model parameters can be obtained $K = 0.478$, $T = 216$, $\alpha = 30$ [27].

The desired course signal is chosen by a representative practical mode as follows:

$$\ddot{\phi}_m(t) + 0.1\dot{\phi}_m(t) + 0.0025\phi_m(t) = 0.0025\phi_r(t) \tag{87}$$

where ϕ_m is the ideal ship course performance; $\phi_r(t)$ is command input signal, which varies between 0° and 30° with period 500s. In the following simulation, the sample time is 100ms.

For step 1, the critic and actor NNs both include 25 nodes, where centers μ_{i} are evenly distributed in the scope [− 7,7], the widths are $\phi_{\mathrm{i}} = 1$, $i = 1, \ldots, 25$. The initial conditions for critic and actor NNs weight are 0.01 with learning rates $\gamma_{1c} = 0.2$ and $\gamma_{1a} = 5$ respectively. The design parameter β_1 is 0.08, the initial conditions are $x_1(0) = 10°$, $x_2(0) = 0$.

For step 2, the critic and actor NNs are designed to include40 nodes, where centers μ_{i} are evenly spaced in [−7, 7], and the widths are $\phi_{\mathrm{i}} = 1$, $i = 1, \ldots, 40$. The learning rates for critic and actor NNs weights are $\gamma_{2c} = 0.3$ and $\gamma_{2a} = 5$, and initial weights are 0.03.The design parameter β_2 is 4.

To further demonstrate the superior performance on optimization of the proposed control law, a comparison study with the direct adaptive NN ship course control approach designed in [6], and MPC ship course control approach is conducted. The corresponding control parameters and RBF NN parameters included in [6] are adopted the same value as the above simulation example to enhance the persuasion of the comparison. The rest parameters mentioned in [6] are $\mathrm{c}_1 = 0.08$, $\mathrm{c}_2 = 100$, $\Gamma_1 = \Gamma_2 = 0.01$ and $\sigma_1 = \sigma_2 = 30$ respectively. For the MPC approach, the prediction horizon $Np = 130$, the control horizon $Nc = 3$.

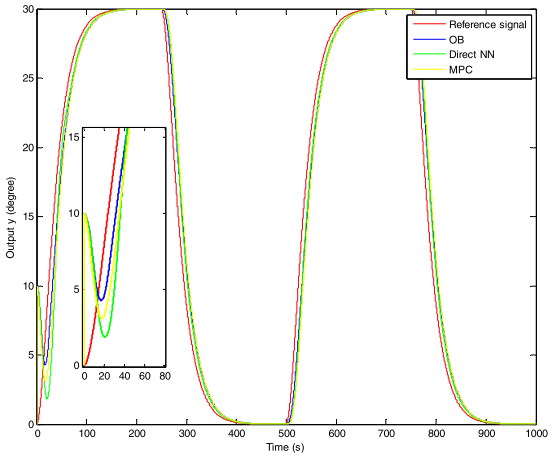

FIGURE 1. Tracking performance based on OB, direct NN and MPC.

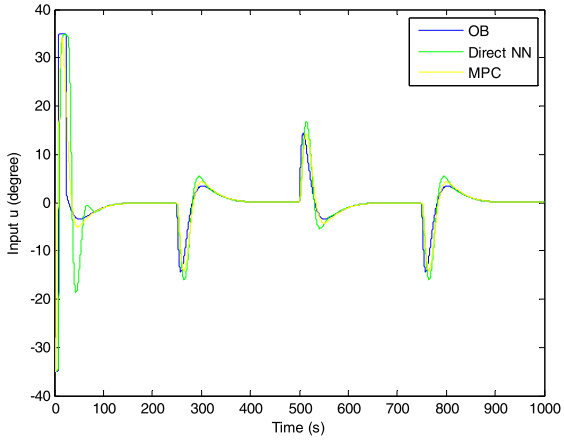

FIGURE 2. Input*u* with saturation based on OB, direct NN and MPC.

Figures 1–7 show the results of simulation based on above parameters and the OB technology utilized by this manuscript, and some comparisons with Direct NN and MPC control approaches are also presented in Figures 1-5. Figure 1 presents that satisfactory ship course tracking performance can be achieved based on above three control approaches. The OB can track the reference signal more quickly with less overshoot. The actual input with saturation is demonstrated in Figure2. The input saturation can be both well handled by the way adopted in this paper and MPC. Course tracking errors $z_1(t)$ are illustrated in Figure3. The OB control scheme can obtain a higher precision. The tracking errors in the second step based on backstepping are illustrated

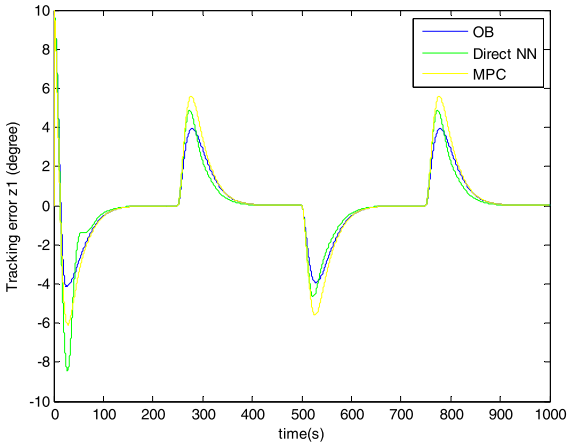

FIGURE 3. Tracking errors in course based on OB, direct NN and MPC.

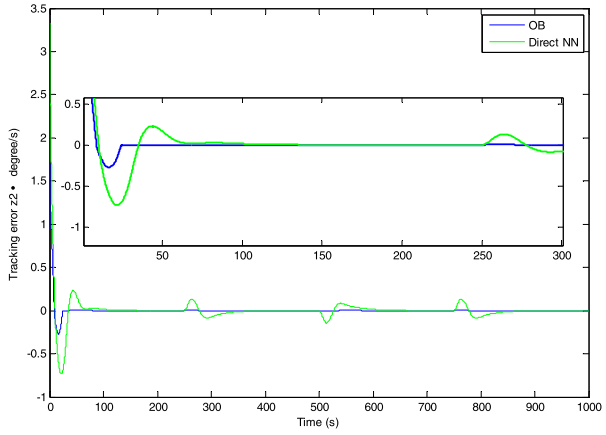

FIGURE 4. Tracking errors in the second step based on OB and direct NN.

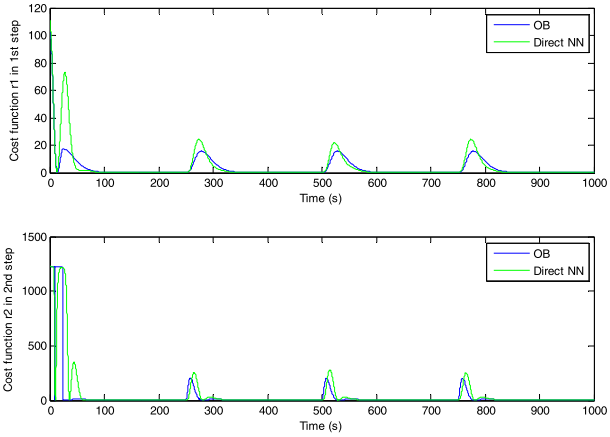

FIGURE 5. Cost functions in two step based on OB and direct NN.

in Figure 4. The error in OB is also smaller than that of Direct NN. The cost functions $r_1(z_1, \alpha_1)$ and $r_2(z_2, u)$ with saturation are presented in Figure5. It is obviously that the proposed control scheme in this paper is lower cost while having almost the same tracking performances. Figure 6 and Figure 7 show the L_2 norms of actor NN and critic NN weights in the first step and second step respectively.

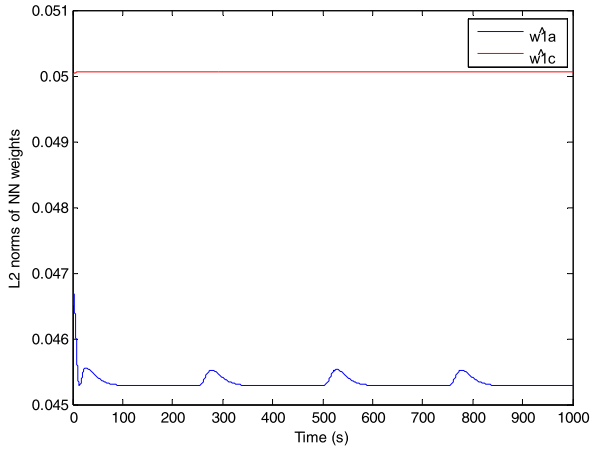

FIGURE 6. L_2 norms of the NN weights in the first step.

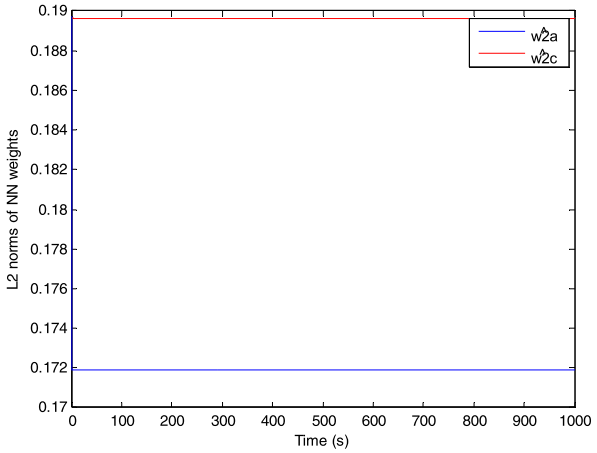

FIGURE 7. L_2 norms of the NN weights in the second step.

TABLE 2. Numerical comparison between the OB control and direct adaptive NN control.

		OB	Direct adaptive NN	Comparison (%)
Tracking error	1st step	9883.8	10080.0	-1.9
	2nd step	102.8	478.9	-78.5
Cost function	1st step	31591	40852	-22.7
	2nd step	390250	518070	-24.7
Total rudder（deg）		23571	30016	-21.5
Mean rudder（deg）		2.36	3.00	-21.5

It should be pointed out that although the optimized control is not considered in literature [6], the corresponding calculation in the same form of cost function $r_1(z_1, \alpha_1)$ and $r_2(z_2, \delta) = z_2^2 + \delta^2$ are still carried out to facilitate the comparison.

To demonstrate the optimizing advantage of the OB controller proposed in this paper more clearly and precisely, a further numerical statistic comparison with Direct NN is

performed and the corresponding results are listed in the Table 2. The total rudder, tracking error, total cost function are the sum of corresponding absolute value at each sample time, the mean rudder is the result of total rudder divided by the sample times.

From Table 2, all terms based on the OB are better than that of Direct NN. The course tracking error in first step z_1 is almost the same, so it implies that the tracking performance is similar. But the cost function and rudder based on OB are much lower. It is obvious and incontrovertible that the OB can achieve a superior control performance in combination of tracking accuracy and energy saving.

V. CONCLUSION

By employing the OB technique, the ship course following control is well developed and optimized. In the control scheme, the actor NNs are used to implement the control law, the critic NNs are used for evaluating the tracking performance and then feedback to actor NNs training for optimizing the control behavior further. Moreover, an auxiliary design system and Gaussian error function are utilized together to deal with the limit of rudder angle in marine practice. Combining above control schemes, the desired ship course can be tracked in high precision and the abrasion of the rudder machine can be reduced and optimized simultaneously. In the future work, the external disturbances including wind and wave and system uncertainty should be taken into consideration.

REFERENCES

[1] M. Wei and G. Chen, "Adaptive RBF neural network sliding mode control for ship course control system," in *Proc. 3rd Int. Conf. Intell. Hum.-Mach. Syst. Cybern.*, Aug. 2011, pp. 27–30.

[2] H. Sh, G. Chen, and L. Tieshan, "Ship course control based on reinforcement learning," in *Proc. ICIC Express Lett.*, 2016, pp. 1489–1495.

[3] X. Hu, X. Wei, H. Zhang, J. Han, and X. Liu, "Global asymptotic regulation control for MIMO mechanical systems with unknown model parameters and disturbances," *Nonlinear Dyn.*, vol. 95, no. 3, pp. 2293–2305, 2019.

[4] H. Xin, W. Xinjiang, and Z. Huifeng, "Robust adaptive tracking control for a class of mechanical systems with unknown disturbances under actuator saturation," *Int. J. Robust Nonlinear Control*, vol. 6, no. 29, pp. 1893–1908, 2019.

[5] N. E. Kahveci and P. A. Ioannou, "Adaptive control design for cargo ship steering with limited rudder angle," *IFAC Proc. Volumes*, vol. 42, no. 15, pp. 505–512, 2009.

[6] J. Li, T. Li, Z. Fan, R. Bu, Q. Li, and J. Hu, "Direct adaptive NN control of ship course autopilot with input saturation," in *Proc. 4th Int. Workshop Adv. Comput. Intell.*, Oct. 2011, pp. 655–661.

[7] Q. Zhang, N. Jiang, Y. Hu, and D. Pan, "Design of course-keeping controller for a ship based on backstepping and neural networks," *Int. J. e-Navigat. Maritime Economy*, vol. 7, pp. 34–41, Jun. 2017.

[8] J. Li, T. Li, Z. Fan, R. Bu, Q. Li, and J. Hu, "Robust adaptive backstepping design for course-keeping control of ship with parameter uncertainty and input saturation," in *Proc. IEEE Int. Conf. Soft Comput. Pattern Recognit.*, Oct. 2011, pp. 63–67.

[9] G. Wen, S. S. Ge, and F. Tu, "Optimized backstepping for tracking control of strict-feedback systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3850–3862, Aug. 2018.

[10] G. Wen, S. S. Ge, C. L. P. Chen, F. Tu, and S. Wang, "Adaptive tracking control of surface vessel using optimized backstepping technique," *IEEE Trans. Cybern.*, to be published.

[11] Z. Zheng, L. Sun, and L. Xie, "Error-constrained LOS path following of a surface vessel with actuator saturation and faults," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 10, pp. 1794–1805, Oct. 2018.

[12] Z. Chen, Q. Chen, X. He, and M. Sun, "Adaptive backstepping control design for uncertain rigid spacecraft with both input and output constraints," *IEEE Access*, vol. 6, pp. 60776–60789, 2018.

[13] H. Habibi, H. R. Nohooji, and I. Howard, "Backstepping Nussbaum gain dynamic surface control for a class of input and state constrained systems with actuator faults," *Inf. Sci.*, vol. 482, pp. 27–46, May 2019.

[14] C. Wen, J. Zhou, Z. Liu, and H. Su, "Robust adaptive control of uncertain nonlinear systems in the presence of input saturation and external disturbance," *IEEE Trans. Autom. Control*, vol. 56, no. 7, pp. 1672–1678, Jul. 2011.

[15] Y. Li, S. Tong, and T. Li, "Composite adaptive fuzzy output feedback control design for uncertain nonlinear strict-feedback systems with input saturation," *IEEE Trans. Cybern.*, vol. 45, no. 10, pp. 2299–2308, Oct. 2015.

[16] Y. Li, S. Tong, and T. Li, "Adaptive fuzzy output-feedback control for output constrained nonlinear systems in the presence of input saturation," *Fuzzy Sets Syst.*, vol. 248, pp. 138–155, Aug. 2014.

[17] L. Edalat, A. K. Sedigh, M. A. Shooredeli, and A. Moarefianpour, "Adaptive fuzzy dynamic surface control of nonlinear systems with input saturation and time-varying output constraints," *Mech. Syst. Signal Process.*, vol. 100, pp. 311–329, Feb. 2018.

[18] G. Ma, C. Chen, Y. Lyu, and Y. Guo, "Adaptive backstepping-based neural network control for hypersonic reentry vehicle with input constraints," *IEEE Access*, vol. 6, pp. 1954–1966, 2017.

[19] J. Ma, Z. Zheng, and P. Li, "Adaptive dynamic surface control of a class of nonlinear systems with unknown direction control gains and input saturation," *IEEE Trans. Cybern.*, vol. 45, no. 4, pp. 728–741, Apr. 2015.

[20] M. Chen, S. S. Ge, and B. Ren, "Adaptive tracking control of uncertain MIMO nonlinear systems with input constraints," *Automatica*, vol. 47, no. 3, pp. 452–465, Mar. 2011.

[21] B. Rui, Y. Yang, and W. Wei, "Nonlinear backstepping tracking control for a vehicular electronic throttle with input saturation and external disturbance," *IEEE Access*, vol. 6, pp. 10878–10885, 2017.

[22] S. Skjong and E. Pedersen, "Nonangular MPC-based thrust allocation algorithm for marine vessels—A study of optimal thruster commands," *IEEE Trans. Transport. Electrific.*, vol. 3, no. 3, pp. 792–807, Sep. 2017.

[23] Z. Li, J. Sun, and S. Oh, "Path following for marine surface vessels with rudder and roll constraints: An MPC approach," in *Proc. Amer. Control Conf.*, Jun. 2009, pp. 3611–3616.

[24] Z. Liu, C. Geng, and J. Zhang, "Model predictive controller design with disturbance observer for path following of unmanned surface vessel," in *Proc. IEEE Int. Conf. Mechatronics Automat. (ICMA)*, Aug. 2017, pp. 1827–1832.

[25] Y. Kim, T. H. Oh, T. Park, and J. M. Lee, "Backstepping control integrated with Lyapunov-based model predictive control," *J. Process Control*, vol. 73, pp. 137–146, Jan. 2019.

[26] X. Shi, Y. Cheng, C. Yin, X. Huang, and S.-M. Zhong, "Design of adaptive backstepping dynamic surface control method with RBF neural network for uncertain nonlinear system," *Neurocomputing*, vol. 330, pp. 490–503, Feb. 2019.

[27] J. Xinle and Y. Yansheng, *Ship Motion Mathematical Model*. Dalian, China: Dalian Maritime Univ. Press, 1999.

[28] E. J. Hartman, J. D. Keeler, and J. M. Kowalski, "Layered neural networks with Gaussian hidden units as universal approximations," *Neural Comput.*, vol. 2, no. 2, pp. 210–215, 1990.

[29] J. Park and I. W. Sandberg, "Universal approximation using radial-basis-function networks," *Neural Comput.*, vol. 3, no. 2, pp. 246–257, Mar. 1991.

[30] Q. Chen, X. Ren, J. Na, and D. Zheng, "Adaptive robust finite-time neural control of uncertain PMSM servo system with nonlinear dead zone," *Neural Comput. Appl.*, vol. 28, no. 12, pp. 3725–3736, 2017.

[31] J. Na, Q. Chen, X. Ren, and Y. Guo, "Adaptive prescribed performance motion control of servo mechanisms with friction compensation," *IEEE Trans. Ind. Electron.*, vol. 61, no. 1, pp. 486–494, Jan. 2014.

[32] Y.-J. Liu, L. Tang, S.-C. Tong, C. L. P. Chen, and D.-J. Li, "Reinforcement learning design-based adaptive tracking control with less learning parameters for nonlinear discrete-time MIMO systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 1, pp. 165–176, Jan. 2015.

[33] S. Haykin, *Neural Networks: A Comprehensive Foundation*. New York, NY, USA: Macmillan, 1994.

[34] S. S. Ge and C. Wang, "Direct adaptive NN control of a class of nonlinear systems," *IEEE Trans. Neural Netw.*, vol. 13, no. 1, pp. 214–221, Jan. 2002.

[35] W. Chen, S. S. Ge, J. Wu, and M. Gong, "Globally stable adaptive backstepping neural network control for uncertain strict-feedback systems with tracking accuracy known *a priori*," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 9, pp. 1842–1854, Sep. 2015.
[36] K. G. Vamvoudakis and F. L. Lewis, "Online actor critic algorithm to solve the continuous-time infinite horizon optimal control problem," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jun. 2009, pp. 3180–3187.
[37] G.-X. Wen, C. L. P. Chen, Y.-J. Liu, and Z. Liu, "Neural-network-based adaptive leader-following consensus control for second-order nonlinear multi-agent systems," *IET Control Theory Appl.*, vol. 9, no. 13, pp. 1927–1934, Aug. 2015.

YUMING BAI was born in Fushun, Liaoning, China, in 1981. He received the B.S. degree in marine technology and the M.S. degree in traffic information engineering and control from Dalian Maritime University (DMU), China, in 2003 and 2006, respectively. He is currently pursuing the Ph.D. degree in traffic information engineering and control.

He is currently a Principal Staff Member with China Maritime Safety Administration (China MSA), Beijing, China. His research interests include ship motion control, nonlinear control, and maritime safety technology.

YUCHI CAO was born in Weifang, Shandong, China, in 1981. He received the B.S. degree in marine technology and the M.S. degree in traffic information engineering and control from Dalian Maritime University (DMU), China, in 2004 and 2010, respectively, where he is currently pursuing the Ph.D. degree in traffic information engineering and control.

He is currently an Associate Professor with DMU, and also a Ship Captain. His research interests include ship motion control, nonlinear control, ship-mounted crane control, and marine technology.

TIESHAN LI received the B.S. degree from the Ocean University of Qingdao, China, in 1992, and the Ph.D. degree from Dalian Maritime University (DMU), China, in 2005. From 2007 to 2010, he was a Postdoctoral Scholar with Shanghai Jiao Tong University. He was with the City University of Hong Kong, as a Senior Research Associate (SRA), from 2008 to 2009.

He is currently a Professor with DMU. His research interests include adaptive control, fuzzy control and neural-network control for nonlinear systems, and their applications to marine control.