# Adaptive Neural Network Control of AUVs With Control Input Nonlinearities Using Reinforcement Learning

Rongxin Cui, *Member, IEEE*, Chenguang Yang, *Senior Member, IEEE*, Yang Li *Student Member, IEEE* and Sanjay Sharma

***Abstract*— In this paper, we investigated the trajectory tracking problem for a fully actuated autonomous underwater vehicle (AUV) which moves in the horizontal plane. External disturbances, control input nonlinearities and model uncertainties are considered in our control design. Based on the dynamic model derived in the discrete time domain, two neural networks (NN) including a critical and an action one are integrated into our adaptive control design. The critical network is introduced to evaluate the long time performance of the designed control in current time step, and the action one is used to compensate for the unknown dynamics. To eliminate the AUV's control input nonlinearities, a compensation item is also designed in the adaptive control. Rigorous theoretical analysis has been performed to prove stability and performance of the proposed control law. Moreover, robustness and effectiveness of the proposed control method has been tested and validated by extensive numerical simulation results.**

***Index Terms*— Autonomous underwater vehicle (AUV); Trajectory tracking; Neural network; Adaptive control**

## I. Introduction

Nowadays, underwater vehicles, including AUVs, ROVs, and gliders have been widely applied in various underwater tasks [1]–[3]. In civil applications, they have been widely used in seafloor mapping, pipeline checking for oil and gas industry, and to find missing airplanes' wreckage in the air rescue operation, etc. In military applications, they have been extensively applied to surveillance and reconnaissance mission, mine countermeasures, oceanography, payload delivery, and other time-critical strike. AUVs have also been involved in scientific investigation of the ocean, ocean floor and lakes. Precise trajectory control of an AUV is crucial while performing the underwater tasks. However, it is a challenge because of the model nonlinearity, coupling and time-varying hydrodynamics coefficients of the dynamics, which need to be further studied.

AUVs usually move in 3D space with 6-DOF, which leads to coupled dynamics between its planar and diving motion. In most studies, the model of AUVs are always decoupled enabling possible application of various control methods [3]–[5]. There are several approaches that have been proposed for AUV trajectory tracking in 3D space, specifically for planner motion or diving. The nonlinear AUV model is usually linearized firstly which and then the controller can be designer based on this linear model [6], [7]. With decoupled model, the diving control of AUVs is concerned. A differentiator was utilized to enhance the noise attenuation performance so that active disturbance rejection can be achieved in [4]. By decoupling the depth and course motion, a fuzzy depth PD controller was designed in [8]. Meanwhile, an output feedback control was proposed for the AUVs which move in vertical plane by transforming the path following errors into Serret-Frenet frame and linearizing the error dynamics in [7]. For planner motion control of AUVs, a nonlinear control for both fully actuated and underactuated configuration was proposed in [5]. It analyzed the effectiveness of side-slip angle of AUVs in detail. Moreover a tilting thruster configuration was proposed in [3], and a selective switching control was designed separately for the two decoupled three-DOF (degree-of-freedom) subsystems. In [9], both a current induced vessel model and a general vehicle model were considered. where the former one accounted for the main current loads. Cascaded system theory and observer backstepping were then employed to deign a nonlinear Luenberger observer and a controller for AUVs. Besides, the results showed that the model-based controller performed better than conventional PD control. In this case, the model dynamics in the controller should be revised in case of the divergence.

Optimal control was also studied in [10]–[12] based on AUVs dynamical model. In [10], an optimal control is designed to control the AUV trajectory in kinematics level, and the cost function is described as kinetic energy cost. Appropriate Hamiltonian was then achieved based on the maximum principle and optimal solution is finally obtained. A nonlinear suboptimal control is presented for a non-affine AUV model, and the state-dependent Riccati equation controller is applied to the point-to-point tracking of NPS II AUV [11]. Treating uncertainty bounds as one item in the cost function, an optimal control problem could be obtained by transforming original robust control problem, then an indirect robust depth control was presented [12].

The hydrodynamic parameters of AUVs are always obtained by computational fluid dynamics (CFD) method or towed experiment identification. However, due to time-varying environmental and state changing while performing underwater tasks, the obtained hydrodynamic parameters are not fixed. Thus, both external disturbance and model parameter uncertainties should be considered in designing an appropriate controller [13]–[16]. To resolve the model parameter uncertainties, Mamdani Fuzzy rules based PID parameters adjustment was employed, and then the control design was decoupled in two

R. Cui (Corresponding author) and Y. Li are with the College of Marine Engineering, Northwestern Polytechnical University, Xi'an 710072, China. (email: r.cui@nwpu.edu.cn).

C. Yang is with the College of Engineering, Swansea University, UK, and was with the School of Computing and Mathematics, Plymouth University, UK. (email: cyang.ac@gmail.com).

S. Sharma is with the School of Marine Science and Engineering, Plymouth University. (email: sanjay.sharma@plymouth.ac.uk).

channels of heading and depth [17]. A discrete time-delay control was presented in [18], in which the dynamics of an AUV is estimated directly. Model uncertainties were also compensated for by the time-delay estimation in [19].

The velocity of an AUV can be measured by a Doppler velocity log (DVL), which always has a slow update rate of new data. To enhance the robustness of unmodeled dynamics and external disturbances for an AUV that uses a DVL, an integral sliding-mode control is introduced [20]. In [21], a novel method to compensate for bounded external disturbances and model uncertainties, in which the integral of the error sign control structure is presented, and semiglobal asymptotic tracking tracking performance can be established through Lyapunov stability analysis. Sliding mode control with backstepping were combined to design a trajectory tracking controller for an AUV with parameter uncertainties and external disturbances in [22].

To deal with external disturbances, a disturbance force measurement method was introduced to measure the forces/moments acting on AUVs, then a feed-forward control was employed on the vehicle based on the predicted response of the dynamic models [2]. Disturbance observer was another main method which had been involved to compensate for the unknown external disturbances [9], [15], [23], [24]. Low frequency motion and wave frequency motion of AUVs were estimated by nonlinear observers, and nonlinear tracking control was then designed for the AUVs motion in shallow wave disturbance in [15]. For the purpose of controlling AUVs in near space, a kind of sliding mode tracking control was applied in [24] based on disturbance observer. Meanwhile, adaptive tracking control for fully actuated AUVs employing a disturbance observer was designed in [25].

Due to the function approximation ability of neural networks, fuzzy approximators, neural networks and fuzzy control based algorithms had been widely studied to compensate for the environmental disturbances and model uncertainties of AUVs [26]–[30]. Neural network approximation was employed to compensate for unknown model parameters, external disturbances, which were induced by ocean currents and waves, and uniform ultimate boundedness of tracking errors was achieved in [26]. Neural networks were used to handle model uncertainties of AUVs, and dynamic surface control was also applied in the control design in [27]. The nonlinear uncertainties in the AUV dynamics were approximated by a two-layer neural network in [28]. To control the diving of AUVs, an adaptive control based on a stable neural network was proposed in [31]. Neural network adaptive control was presented for multiple AUVs, and the unmeasured states were estimated by a local observer in [32]. The RBF neural network was presented to derive the adaptive controller for systems that subject to external disturbances and unknown hysteresis in [33]. In a recent work [34], the nonaffine pure-feedback discrete-time nonlinear system subjected to input dead-zone was considered. To compensate for the dead-zone, an adaptive compensative term and an n-step-ahead predictor was constructed by transforming original system.

Practical control system of AUVs is usually implemented on an embedded computer in a digital manner with samplers. Thus, the continues-time controller needs to be transformed into a discrete-time version [35]. By using the discrete-time model directly, we develop a trajectory tracking control in the presence of external disturbances, model parameter uncertainties, and control input nonlinearities. It should be noted that there already exist a number of methods to deal with input nonlinearities problems, such as input dead-zone and saturation [36]–[38]. Based on the back-stepping method and Lyapunov analysis, an adaptive trajectory tracking controller was designed to overcome the model parameter uncertainties in [37], where a saturation function was utilized to resolve the actuator saturation problem. To prevent the velocity constraint violation, a robust adaptive controller was proposed for an remotely operated vehicle, and the barrier Lyapunov function was used in the Lyapunov syntheses in [36]. In [38], a novel dynamic surface control (DSC) was proposed for the pure-feedback system with unknown input dead zone. The complexity reduces obviously due to the dynamic surface control policy. Motivated by the work in [34], [39], [40], we propose reinforcement learning technique to achieve optimal trajectory tracking for AUVs by employing two neural networks. The unknown nonlinearities and disturbances are approximated by the action neural network, while in the mean time, the tracing evaluation of the tracking performance is approximated by the critical neural network. In addition, an adaptive compensation for the control input nonlinearities are considered. Preliminary results of this work were presented in [41], and extension have been made by taken into consideration of not only actuator deadzone and saturation, but also the nonlinear relationship between the nominal and actual force/moment. Moreover, a compensation policy for this nonlinearity is proposed, which will be discussed later.

The reminder of the paper is organized as follows. We present the nonlinear model of AUVs in Section II. The two adaptive neural networks will be designed in Section III. Simulation studies and conclusions are drawn in Section IV and V, respectively.

## II. Problem Formulation

### A. Motion Equations of an AUV

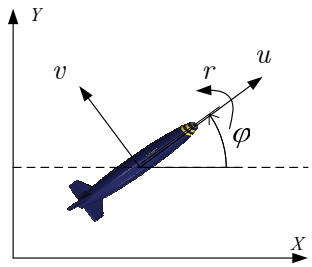


Fig. 1. AUV moves in horizontal plane.

As described in Section I, an AUV usually moves in a 3D space with 6-DOF and leads to coupled dynamics in its planner and diving motion. To facilitate control design, the model is usually decoupled, while the designed control will be validated using the coupled nonlinear dynamics. We consider the planar

motion of an AUV with 3 DOF as shown in Fig. 1. Let us denote the position coordinate of an AUV as $(x,y)$, the yaw as $(\psi)$ in the inertial frame, the velocity as $(u)$ in surge, $v$ in sway and $r$ in yaw in AUV body coordinate. Furthermore, let us denote the matrix of inertia of the AUV as $M$, the matrices for Coriolis and centripetal acceleration and damping as $C(\nu)$, and $D(\nu)$, respectively. In addition, we denote the forces and moments generated by gravitational and buoyancy as $g(\eta)$. Consider that the unknown external disturbances and model parameter uncertainties exists, then the AUV dynamics can be given as below:

$$\dot{\eta} = R(\psi)\nu \tag{1}$$

$$M\dot{\nu} + C(\nu)\nu + D(\nu)\nu + g(\eta) + \Delta(\eta,\nu) = \tau \tag{2}$$

where $\Delta(\eta,\nu)$ is model uncertainties vector which is induced by unmodelled dynamics and external disturbances $\Delta(\eta,\nu)$, and $R(\psi)$ describes the rotation from the AUV body coordinate to earth coordinate in 3 DOF. The control inputs of the AUV are defined by $\tau \in \mathbb{R}^3$.

The elements in $M$, $m_{ij},\ i,j=1,2,3$, the functions in $D(\nu)$ $d_{ij}(\nu),\ i,j=1,2,3$, and the element of the disturbance vector $\Delta_i(\eta,\nu),\ i=1,2,3$ are all unavailable for control design. To facilitate the control design, we assume that their is a nominal value of the unknown mass matrix $M$, which is defined by $M_0$. In addition, $M_0$ is known *a priori*. This assumption is feasible because mass and added mass of an AUV are mainly determined by its physical shape. Control design in this work is focused on the 3 DOF model. However, we could extend the control policy into 6 DOF conveniently due to the fully actuated model of AUVs used in this paper. In other words, the controller designed in this paper can also be applied into the vertical planer.

### *B. Dynamics Model in Discrete-Time Domain*

In this subsection, we transform the continuous-time model into discrete time for subsequent control design. Eqs. (1) and (2) can be rewritten as following equations.

$$\begin{aligned}\dot{\nu} =& -M^{-1}\left[(C(\nu)+D(\nu))\nu + g(\eta) + \Delta(\eta,\nu)\right] \\ &+ M^{-1}\tau \\ \dot{\eta} =& R(\psi)\nu\end{aligned} \tag{3}$$

If the sampling time of the embedded computer for the AUV control is selected as $T_s$, through the first-order Taylor expansion, we could obtain an approximative discrete-time model calculated from (3) as

$$\begin{aligned}\nu(k+1) &= \nu(k) + f_2\big(\eta(k),\nu(k)\big) + M^{-1}\tau(k) \\ \eta(k+1) &= \eta(k) + f_1\big(\eta(k)\big)\nu(k)\end{aligned} \tag{4}$$

where $\eta(k)$, $\nu(k)$ and $\tau(k)$ are the $k-$th sampling time sampled value of $\eta$, $\nu$ and $\tau$, respectively. The nonlinear functions

$$\begin{aligned}f_1(\eta) =& T_s R(\psi) \in \mathbb{R}^{3\times 3} \\ f_2(\eta,\nu) =& -T_s M^{-1}\big[(C(\nu)+D(\nu))\nu \\ &+ g(\eta) + \Delta(\eta,\nu)\big] \in \mathbb{R}^3\end{aligned} \tag{5}$$

Following definitions are introduced for convenience of future design.

$$\begin{aligned}f_{11}\big(\bar{x}(k)\big) &\triangleq f_1\big(\eta(k) + f_1(\eta(k))\nu(k)\big) \\ &= f_1(\eta(k+1)) \in \mathbb{R}^3\end{aligned} \tag{6}$$

where $\bar{x}(k) \triangleq \left[\eta^\top(k), \nu^\top(k)\right]^\top$.

By using the procedure presented in [42], we could drive the following equations from (4).

$$\begin{aligned}\eta(k+2) =& \eta(k+1) + f_1(\eta(k+1))\nu(k+1) \\ =& \eta(k) + f_1(\eta(k))\nu(k) + f_{11}(\bar{x}(k))\nu(k) \\ &+ f_{11}(\bar{x}(k))f_2(\bar{x}(k)) + f_{11}(\bar{x}(k))M^{-1}\tau\end{aligned} \tag{7}$$

It is easy to check that $\frac{1}{T_s^2} f_{11}^\top(\bar{x}(k)) f_{11}(\bar{x}(k)) = R(\psi(k+1))R^\top(\eta(k+1)) = I$. Now we define

$$\begin{aligned}f(\bar{x}(k)) \triangleq& f_{11}(\bar{x}(k))f_2(\bar{x}(k)) \in \mathbb{R}^3 \\ h(\bar{x}(k)) \triangleq& x_1(k) + f_1(\eta(k))\nu(k) \\ &+ f_{11}(\bar{x}(k))\nu(k) \in \mathbb{R}^3 \\ M_f(\bar{x}(k)) =& f_{11}(\bar{x}(k))M^{-1}f_{11}^\top(\bar{x}(k)) \in \mathbb{R}^{3\times 3} \\ \tau_f =& \frac{1}{T_s^2} f_{11}(\bar{x}(k))\tau \in \mathbb{R}^3\end{aligned}$$

Then (7) can be written as

$$\eta(k+2) = h(\bar{x}(k)) + f(\bar{x}(k)) + M_f(\bar{x}(k))\tau_f \tag{8}$$

It is noted that $f_{11}(\bar{x}(k)) = f_1\big(\eta(k) + f_1(\eta(k))\nu(k)\big)$ is known, and then at each time instant $k$, $R(\psi(k+1)) = f_{11}(\bar{x}(k))$ can be calculated, so that $h(\bar{x}(k))$ is also known. The function $f(\bar{x}(k))$ is not available, so that it has to be well considered in our design. Furthermore, the property of the $M$ matrix results in a positive definite matrix $M_f(\bar{x}(k))$ which is also unknown.

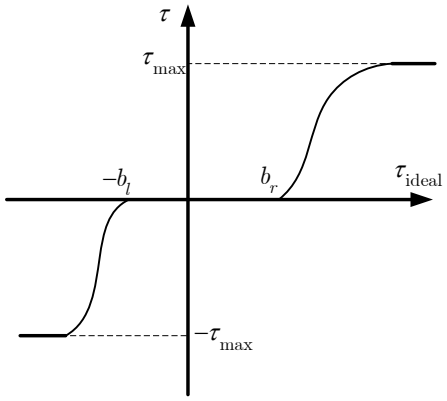


Fig. 2. Illustration of actuator nonlinearity.

Actuator dead-zone and saturation are inevitable existed in any physical system. In this work, we consider that the actuator has the nonlinearities including both saturation and dead-zone, as shown in Fig. 2. Let's define $\tau_{ideal}(k)$ as the ideal control input generated by the proposed controller, and the nominal force/moment acting on the vehicle can be described as

$$\overrightarrow{D}(\tau_{ideal}(k)) = \overrightarrow{m}(\tau_{ideal}(k))\tau_{ideal}(k) + \overrightarrow{b}(k) \tag{9}$$

where $\overrightarrow{D}(\tau_{ideal}(k))$ is a continuous first-order derivable function. This formulation (9) has been widely used to describe actuator nonlinearities with both saturation and dead-zone [43], [44].

The objective of this work is, based on the discrete-time model (8) and the control input nonlinearities (9), to develop a torque control input $\tau$ that makes trajectory of an AUV, $\eta = [x, y, \psi]^\top$ follows the user defined trajectory $\eta_d = [x_d, y_d, \psi_d]^\top$ asymptotically, i.e., $\lim_{k\to\infty}(\eta - \eta_d) = 0$.

### C. Neural Network Function Approximation

Neural networks and fuzzy system are typical function approximators widely used in the control community. We choose RBF neural network to compensate for unknown items in dynamics for our control design and following RBF neural networks are introduced. An unknown function $h(z) : \mathbb{R}^m \to \mathbb{R}$ can be approximated by

$$\phi(W, z) = W^\top S(z) \tag{10}$$

where the weight of the neural network is denoted as $W = [w_1, w_2, \cdots, w_l]^\top \in \mathbb{R}^l$, the input vector is denoted as $z \in \Omega_z \subset \mathbb{R}^m$, the node number of the neural network is denoted as $l$.

In this work, the basis function of a neural network is defined as $S(z) = [s_1(z), \cdots, s_l(z)]^\top$ and its element $s_i(z),\ i = 1, \ldots, l$ is selected as a Gaussian function.

Let us define $\mu_i = [\mu_{i1}, \cdots, \mu_{im}]^\top$ as centers of the receptive domain of the neural network and define $\sigma_i$ as width of the Gaussian function, then it can be written as

$$s_i(z) = \exp\left[\frac{-(z-\mu_i)^\top (z-\mu_i)}{\sigma_i^2}\right], i = 1, 2, ..., l \tag{11}$$

Let us define $W^*$ as the ideal constant weights of a neural network. In [45], it has been established that any continuous function can be arppximated by a RBF neural network (10) over a compact set $\Omega_z \subset \mathbb{R}^m$ as

$$\phi(z) = W^{*\top} S(z) + \varepsilon_z,\ \forall z \in \Omega_z \tag{12}$$

where $\varepsilon_z$ is the approximation error.

The ideal weight $W^*$ is required for stability analysis, and it minimizes $|\varepsilon_z|$ for all $z \in \Omega_z$, i.e.,

$$W^* \stackrel{def}{=} arg \min_{W' \in \mathbb{R}^l} \left\{ \sup |h(z) - W'^\top S(z)| \right\}, \quad z \in \Omega_z$$

In this work, the user defined trajectories for an AUV, $\eta_d$ and $v_d$, are assumed to be suitably defined such that if $z_d = [\eta_d^\top, v_d^\top]^\top$ is presented as the input to a RBF neural network, then $S(z_d)$ will satisfy persistent excitation condition [46], i.e.,

$$\alpha_{\min} I_{l\times l} \le \int_{k_0}^{k_0+k_f} S(\pi) S^\top(\pi) \mathrm{d}\pi \le \alpha_{\max} I_{l\times l}, \quad \forall k_0 \tag{13}$$

where $\alpha_{\min}, \alpha_{\max}$ and $k_f$ are positive constants, and $I$ is an identity matrix.

## III. Adaptive Neural Network Control Design

We have derived the discrete-time model of an AUV in Section II, and we noted that there are some unknown functions in the dynamics model. Thus, the technical challenges in the control design of an AUV include the external disturbances, the partial unknown dynamics, and the input nonlinearities. In this section, we propose the trajectory control for an AUV using NN. Two NNs are employed. The first critical NN is used to evaluate the long time performance of the control in current time step. Note that only $f(\bar{x}(k))$ and $M_f$ are unknown and unavailable. Thus, the action neural network is employed to approximate $f(\bar{x}(k))$ and $M_f$. In the other words, action neural network is used to compensate for the effect caused by the unknown dynamics. We will design the control $\tau$ to adjust $\eta(k)$ to track the desired trajectory $\eta_d(k)$. Define

$$e(k) \triangleq \eta(k) - \eta_d(k) \in \mathbb{R}^3 \tag{14}$$

and $e_v(k) = v(k) - v_d(k)$ as the position tracking error and velocity tracking error, respectively.

Define

$$\varepsilon(k) \triangleq \sum_{i=1}^{3} w_i \left|\lambda_i e_i(k) + e_{v,i}(k)\right| \tag{15}$$

as a new weighted tracking error, where superscript $i$ denotes the $i$-the element of the vector, $\lambda_i$ is a constant weighted the position tracking error and velocity error, and $w_i$ is a constant weight associated with the error in each channel. $w_i$ can also be viewed as the weights of impacts by $e_i$ and $e_{v,i}$. Generally, it is chosen to achieve the normalization of $e_i$ and $e_{v,i}$.

The definition in (15) is motivated by the results in [39], [40], while both position and velocity errors are taken into account in this work. The control objective can be then described $\lim_{k\to\infty} \varepsilon(k) = 0$. Thus $\varepsilon(k)$ can be viewed as a strategic utility function to observe the instant tracking performance. Then the long term performance measure with a future time horizon $N$ can be defined as [39]

$$Q(k) = \sum_{i=1}^{N-k} \alpha^{N+1-i} \varepsilon(k+i) \tag{16}$$

where the scaling factor $\alpha$ is defined by user and it satisfies $0 < \alpha < 1$ .

The long term performance $Q(k)$ was firstly introduced in [39] to stand for the tracking performance including all history information. It utilized the binary system-performance index $p_i(k) \in \mathbb{R}$. And $\varepsilon(k) = 0$ when tracking error is within the limits of a given boundary, $\varepsilon(k) = 1$ otherwise. In this paper, however, we use the weighted-error to stand for the performance other than the binary utility function. Moreover, this measure is also similar to the standard Bellman equation [47], [48].

In the following, we design the tracking controller with two NNs used. The utility function $Q(k)$ (16) is approximated by one critic neural network and the unknown item $f(\bar{x}(k))$ in (8) is approximated by another neural network.

### A. Critic Neural Network Design

As mentioned, two NNs are employed in our control design. In this subsection, we use a critic neural network to approximate the unknown strategic utility function $Q(k)$. Following the techniques used in [49], we calculate its estimation. Rewrite $Q(k)$ using NN and then it can be formulated as below

$$Q(k) = W_c^* S_c(k) + \mu_Q(k) \tag{17}$$

where $W_c^* \in \mathbb{R}^l$ is optimal neural weight, $\mu_Q(k)$ denotes the neural network approximation error, and $S_c(z(k)) \in \mathbb{R}^l$ denotes the activation vector. Then the estimation of $Q(k)$ can be given by

$$\hat{Q}(k) = \hat{W}_c^\top(k) S_c(k) \tag{18}$$

where $\hat{W}_c(k)$ is the neural weight.

The input vector $z(k)$ is given by

$$\begin{aligned} z(k) \quad = \quad & [\bar{x}^\top(k), \eta_d^\top(k), , \cdots, \\ & \eta_d^\top(k+N), \nu_d^\top(k), \cdots, \nu_d^\top(k+N)]^\top \end{aligned} \tag{19}$$

Now we could obtain the prediction error

$$e_c(k) = -\alpha \hat{Q}(k-1) + \hat{Q}(k) + \alpha^{N+1} \varepsilon(k) \tag{20}$$

Then, a critic neural network is designed to minimize objective function

$$E_c(k) = \frac{1}{2} e_c^2(k) \tag{21}$$

A simple method to update the critic neural network is to use the conventional gradient based adaption as follows:

$$\hat{W}_c(k+1) = \hat{W}_c(k) + \Delta \hat{W}_c(k) \tag{22}$$

In (22), the recurrence of $\hat{W}_c(k)$ is given by

$$\begin{aligned} & \hat{W}_c(k+1) = \hat{W}_c(k) - \alpha_c S_c(k) \\ & \times \left[ \hat{W}_c^\top(k) S_c(k) - \alpha \hat{W}_c^\top(k-1) S_c(k-1) + \alpha^{N+1} p(k) \right] \end{aligned} \tag{23}$$

where $\alpha_c \in \mathbb{R}$ denotes the parameter gain of the neural network. Eq. (23) shows that weights are adjusted in accordance with the reinforcement learning signal and past critic neural network output values with discount.

### B. Action Neural Networks Design

A critic neural network has been used to approximate the performance evaluation function. In this subsection, an action neural network based adaptive control considering the mentioned technical challenges is presented for the AUV as follows:

$$\tau_{ideal}(k) = \tau_f(k) + \hat{\xi}(k) + \beta(k) e(k) \tag{24}$$

where $\bar{\beta}$ is a user defined scaling factor satisfying $|\beta(k)| \leq \bar{\beta} < 1$, $\tau_f$ is the ideal control input defined as

$$\tau_f = \hat{W}_a^\top(k) \langle \cdot \rangle S_a(\bar{z}(k)) - M_i(\bar{x}(k)) [h(\bar{x}(k)) + y_d(k+2)] \tag{25}$$

where $\xi(k)$ is the input compensation signal that will be introduced later, and we denote $\hat{\xi}, \xi^*$ and $\tilde{\xi}$ as the real compensation, the optimal compensation and the error of $\xi$, i.e., $\tilde{\xi} = \xi^* - \hat{\xi}$. $\hat{W}_a^\top(k) \langle \cdot \rangle S_a(\bar{z}(k))$ is the neural network based approximation of the unknown item $f_n(\bar{x}(k)) \in \mathbb{R}^3$.

$$\begin{aligned} f_n(\bar{x}(k)) = & - \left[ M_f^{-1}(\bar{x}(k)) - M_i(\bar{x}(k)) \right] h(\bar{x}(k)) \\ & - M_f^{-1}(\bar{x}(k)) f(\bar{x}(k)) \end{aligned} \tag{26}$$

where $M_i(\bar{x}(k))$ is an estimation of $M_f^{-1}(\bar{x}(k))$ and

$$M_i(\bar{x}(k)) = f_{11}(\bar{x}(k)) M_0 f_{11}^\top(\bar{x}(k)) \tag{27}$$

If the nominal value of $M_0$ equals $M$ exactly, then $M_i(\bar{x}(k)) M_f(\bar{x}(k)) = I$ and the first item of $f_n(\bar{x}(k))$ in (26) equals to zero. In practice, $M_0$ is obtained according to the designer's experience. Even though $M_0$ is never achieved precisely equivalent to $M$, the distance between $M$ and $M_0$ will be compensated by the action neural networks. And $M_0$ is more closer to $M$, the less convergence time is needed.

According to universal approximation theory, there are $S(\bar{x}(k)) \in \mathbb{R}$ and ideal weights $W_a^{*\top}$ satisfying

$$f(\bar{x}(k)) = \mu(\bar{x}(k)) + W_a^{*\top} \langle \cdot \rangle S(\bar{x}(k)) \tag{28}$$

$$S_a(\bar{x}(k)) \in \mathbb{R}^{l \times 3}, \ \forall \bar{x}(k) \in \Omega_{\bar{x}} \tag{29}$$

where $\mu(\bar{x}(k))$ denotes the approximation error by the designed neural network.

The update law of the compensation item $\xi(k)$ is designed as

$$\hat{\xi}(k+2) = \hat{\xi}(k) - \left[ \gamma_\xi \hat{\xi}(k) + \alpha_\xi e(k+2) \right] \tag{30}$$

where $\alpha_\xi$ and $\gamma_\xi$ are parameters to be specified by the designer.

Substituting the designed control (24) into dynamics (7) results in

$$e_a(k+2) = M_f(\bar{x}(k)) \left[ \tilde{D}(v(k)) + \tilde{W}_a^\top(k) S_a(k) \right] + d_s^*(k) \tag{31}$$

where $d_s^*(k) = M_f \mu(k)$, and the augmented error of $\tilde{D}(v(k))$ is defined as $\tilde{D}(v(k)) = D^*(v(k)) - D(v(k))$ with superscript $*$ standing for the ideal value.

Now, we need to minimize the following objective function.

$$E_a(k) = \frac{1}{2} e_a^\top(k) e_a(k) + \frac{1}{2} \hat{Q}(k)^2 \tag{32}$$

Define $k_1 = k - n$ and now we could design the update law for the action neural network.

$$\hat{W}_a(k+1) = \Delta \hat{W}_a(k_1) + \hat{W}_a(k_1) \tag{33}$$

where $\Delta \hat{W}_a$ can be calculated using gradient decent method which results in

$$\hat{W}_a(k+1) = \hat{W}_a(k_1) - \alpha_a S_a(k_1) [-\text{sign}(e(k)) \hat{Q}(k) \alpha^{N+1} + e(k)] \tag{34}$$

where $\alpha_a \in \mathbb{R}$ is the adaption gain of the neural network.

*Lemma 1:* [44] Let $V(k) = \sum_{i=1}^n V_i(k)$, and $V_i(k) \geq 0, k \in \mathbb{Z}^+$. If $V(k+1) \leq \sum_{i=1}^n p(k-n+1) V_i(k-n+1) + q(k-n+1)$, $|p(k-n+1)| \leq \bar{p} < 1$ and $|q(k-n+1)| \leq \bar{q}$. Then, we have

$$V(k) \leq \bar{V}(0) + \frac{\bar{q}}{1-\bar{p}} \tag{35}$$

Moreover, we have

$$\limsup_{k \leftarrow \infty} V(k) \leq \frac{\bar{q}}{1-\bar{p}} \tag{36}$$

where $\bar{V}(0) = \max_{0\le i\le n-1}\{V(j)\}$

Now we can arrive at the following theorem which summarize the stability result of the developed control.

*Theorem 1:* If the dynamics of the vehicle can be described by (7), using the adaptive control (24), compensation parameter adaptation law (30), and neural network weights update law (22) and (33), the AUVs is able to follow the desired trajectory with bounded error when design parameters of the control satisfy $\alpha_c\|S_c(k)\|^2 < 1$, $\alpha_a\|S_a(k)\|^2 < 1$, $\gamma_\xi < 1/2$, $0 < \alpha < \frac{\sqrt{2}}{2}$ and $\bar{\beta} + 3\alpha_\xi < \frac{2}{M_m(\bar{x}(k))}$.

**Proof.** Using similar techniques employed in [34], [39]–[41], we choose a positive definite

$$V(k) = \sum_{i=1}^{4} V_i(k) \tag{37}$$

where

$$\begin{aligned} V_1(k) &= \frac{1}{\alpha_c}\mathrm{tr}\left[\tilde{W}_c^\top(k)\tilde{W}_c(k)\right] \\ V_2(k) &= \frac{1}{\gamma_c}\|\zeta_c(k-1)\|^2, \text{ and} \\ V_3(k) &= \frac{1}{\gamma_a\alpha_a}\sum_{j=0}^{n}\mathrm{tr}\left[\tilde{W}_a^\top(k-n+j)\tilde{W}_a(k-n+j)\right] \end{aligned} \tag{38}$$

where $\zeta_c(k) = \tilde{W}_c^\top(k)S_c(k)$ and parameters $\gamma_c, \gamma_a > 0$. The important item $V_4(k)$ defined in (38) is introduced to handle the control input nonlinearity items, which is selected as

$$V_4(k) = \frac{1}{\alpha_\xi}\tilde{\xi}^\top(k)\tilde{\xi}(k) \tag{39}$$

According to the results in [41], we can get

$$\Delta V_i(k+1) = V_i(k+1) - V_i(k) < 0,\ i = 1,2,3 \tag{40}$$

The details about the proof can be found in [41]. Based on (38), we obtain

$$\begin{aligned} V_4(k+2) =& \frac{\tilde{\xi}^\top(k+2)\tilde{\xi}(k+2)}{\alpha_\xi} \\ =& \frac{\tilde{\xi}^\top(k)\tilde{\xi}(k)}{\alpha_\xi} - 2\tilde{\xi}(k)^\top e(k+2) - \frac{2\gamma_\xi\tilde{\xi}^\top(k)\hat{\xi}(k)}{\alpha_\xi} \\ &+ 2\gamma_\xi e(k+2)^\top\hat{\xi}(k) + \frac{\gamma_\xi^2\hat{\xi}(k)}{\alpha_\xi} \\ &+ \alpha_\xi e^\top(k+2)e(k+2) \end{aligned} \tag{41}$$

Because

$$\begin{aligned} & M_f(\bar{x}(k))D(v(k)) \\ =& M_f(\bar{x}(k))\left\{-b(k) + m(k)\left[\tau_f(k) + \hat{\xi}(k) + \beta e(k)\right]\right\} \\ =& -M_f(\bar{x}(k))b(k) + M_f(\bar{x}(k))m(k)\tau_f(k)\hat{\xi}(k) \\ &+ M_f(\bar{x}(k))m(k) + M_f(\bar{x}(k))m(k)\beta e(k) \\ =& -M_f(\bar{x}(k))b(k) + M_f(\bar{x}(k))m(k)\tau_f(k) \\ &- M_f(\bar{x}(k))m(k)\tilde{\xi}(k) + M_f(\bar{x}(k))m(k)\xi^* \\ &+ M_f(\bar{x}(k))m(k)\beta e(k) \end{aligned}$$

Now we can rewrite (31) as

$$\begin{aligned} e(k+2) =& M_f(\bar{x}(k))(m(k)-1)\tau_f(k) + M_f(\bar{x}(k))\tilde{W}_a^\top(k)S_a(k) \\ &- M_m(\bar{x}(k))\tilde{\xi}(k) + M_m(\bar{x}(k))\beta e(k) + M_m(\bar{x}(k))\xi^* \\ &- M_f(\bar{x}(k))b(k) + d_s^*(k) \\ =& M_m(\bar{x}(k))\left[\tilde{\xi}(k) + \beta e(k)\right] + H(k) \end{aligned} \tag{42}$$

where

$$\begin{aligned} H(k) =& M_f(\bar{x}(k))(m(k)-1)\tau_f(k) + M_m(\bar{x}(k))\xi^* \\ &- M_f(\bar{x}(k))b(k) + M_f(\bar{x}(k))\tilde{W}_a^\top(k)S_a(k) + d_s^*(k) \end{aligned} \tag{43}$$

Let us multiply by $e(k+2)$ on both sides of (42) and to obtain $e^\top(k+2)e(k+2) = e^\top(k+2)H(k) + M_m(\bar{x}(k))[e^\top(k+2)\tilde{\xi}(k) + e^\top(k+2)\beta e(k)]$. Thus, we have

$$\begin{aligned} -2e^\top(k+2)\tilde{\xi}(k) =& -\frac{2e^\top(k+2)e(k+2)}{M_m(\bar{x}(k))} + \frac{2e^\top(k+2)H(k)}{M_m(\bar{x}(k))} \\ &+ 2\beta e^\top(k+2)e(k) \end{aligned} \tag{44}$$

It is easy to have the following equations or inequations.

$$\begin{aligned} & 2\tilde{\xi}^\top(k)\hat{\xi}(k) = -\xi^{*\top}\xi^* + \hat{\xi}^\top(k)\hat{\xi}(k) + \tilde{\xi}^\top(k)\tilde{\xi}(k) \\ & 2\gamma_\xi e^\top(k+2)\hat{\xi}(k) \le \alpha_\xi e^\top(k+2)e(k+2) + \frac{\gamma_\xi^2\hat{\xi}^\top(k)\hat{\xi}(k)}{\alpha_\xi} \\ & \frac{2H^\top(k)e(k+2)}{M_m(\bar{x}(k))} \le \alpha_\xi e^\top(k+2)e(k+2) + \frac{\bar{H}^\top(k)\bar{H}(k)}{\alpha_\xi} \\ & 2\beta e^\top(k+2)e(k) \le \bar{\beta}\left[e^\top(k+2)e(k+2) + e^\top(k)e(k)\right] \end{aligned} \tag{45}$$

where $\bar{H}(k)$ and $\bar{\beta}$ are the upper bounds of $H(k)$ and $\beta$, respectively. It is noted that the first equation is a direct conclusion of $2a^\top b < a^\top a + b^\top b$.

Substituting (45) into (41), we obtain

$$\begin{aligned} V_4(k+2) \le& \frac{1}{\alpha_\xi}\tilde{\xi}^\top(k)\tilde{\xi}(k) - \frac{2}{M_m(\bar{x}(k))}e^\top(k+2)e(k+2) \\ &+ \alpha_\xi e^\top(k+2)e(k+2) + \frac{1}{\alpha_\xi}\bar{H}^\top(k)\bar{H}(k) + \bar{\beta}e^\top(k)e(k) \\ &+ \bar{\beta}e^\top(k+2)e(k+2) - \frac{\gamma_\xi}{\alpha_\xi}\tilde{\xi}^\top(k)\tilde{\xi}(k) - \frac{\gamma_\xi}{\alpha_\xi}\hat{\xi}^\top(k)\hat{\xi}(k) \\ &+ \frac{\gamma_\xi}{\alpha_\xi}\xi^{*\top}\xi^* + 2\alpha_\xi e^\top(k+2)e(k+2) + \frac{2\gamma_\xi^2\hat{\xi}^\top(k)\hat{\xi}^\top(k)}{\alpha_\xi} \\ =& (1-\gamma_\xi)\frac{\tilde{\xi}^\top(k)\tilde{\xi}(k)}{\alpha_\xi} + \gamma_\xi(2\gamma_\xi - 1)\frac{\hat{\xi}^\top(k)\hat{\xi}(k)}{\alpha_\xi} \\ &+ \left[\bar{\beta} + 3\alpha_\xi - \frac{2}{M_m(\bar{x}(k))}\right]e^\top(k+2)e(k+2) + \bar{\beta}e^\top(k)e(k) \\ &+ \frac{\bar{H}^\top(k)\bar{H}(k)}{\alpha_\xi} + \frac{\gamma_\xi}{\alpha_\xi}\xi^{*\top}\xi^* \\ \le& (1-\gamma_\xi)V_4(k) + \gamma_\xi(2\gamma_\xi - 1)\frac{\hat{\xi}^\top(k)\hat{\xi}(k)}{\alpha_\xi} \\ &+ \left[\bar{\beta} + 3\alpha_\xi - \frac{2}{M_m(\bar{x}(k))}\right]e^\top(k+2)e(k+2) + \bar{q} \end{aligned} \tag{46}$$

where $\bar{q} = \frac{\bar{H}^\top(k)\bar{H}(k)}{\alpha_\xi} + \frac{\gamma_\xi}{\alpha_\xi}\xi^{*\top}\xi^* + \bar{\beta}e^\top(k)e(k)$. If the parameters are selected to satisfy $\gamma_\xi < 1/2$, and $\bar{\beta} + 3\alpha_\xi < 2/M_m(\bar{x}(k))$, then we could obtain

$$V_4(k+2) \leq (1-\gamma_\xi)V_4(k) + \bar{\beta}e^\top(k)e(k) + \bar{q} \tag{47}$$

In addition, because $\gamma_\xi > 0$, $\bar{\beta} < 1$, and $\Delta V_i(k) < 0$, $i = 1,2,3$, we obtain $V(k+2) \leq \beta_{v1}V_1(k) + \beta_{v2}V_2(k) + \beta_{v3}V_3(k) + (1-\gamma_\xi)V_4(k) + \bar{\beta}e^\top(k)e(k) + \bar{q}$. Thus we could conclude that $V(k) \leq V(0) + \frac{\bar{q}}{1-\bar{p}}$ according to Lemma 1, where the constant $\bar{p} = \max\{1-\gamma_\xi, \beta_{v1}, \beta_{v2}, \beta_{v3}\bar{\beta}\}$. This completes the proof.

## IV. Simulation Studies

In order to evaluate the performance of the proposed control, we perform a numerical simulation based on the model of a fully actuated AUV in this section, and which has been used in [13] with success. The parameters of the model are adapted from [13] as below:

$$M = \begin{bmatrix} 25.8 & 0 & 0 \\ 0 & 24.6612 & 0 \\ 0 & 0 & 2.76 \end{bmatrix} \tag{48}$$

$d_{11} = (0.7225 + 1.3274|u| + 5.8664v^2)kg/s$, $d_{22} = (0.8612 + 36.2823|v| + 8.05|r|)kg/s$, $d_{23} = (-0.1079 + 0.845|v| + 3.45|r|)kg/s$, $d_{32} = (-0.1052 - 5.0437|v| - 0.13|r|)kg/s$, and $d_{33} = (1.9 - 0.08|v| + 0.75|r|)kg/s$. The desired trajectory is given by

$$\begin{aligned} x_d(k) &= k \\ y_d(k) &= 4\sin(k/7) \end{aligned} \tag{49}$$

Thus, the desired yaw of the vehicle can be calculated as $\psi d(k) = \arctan(\frac{4}{7}\cos(k/7))$ for a smooth trajectory tracking. For convenience, we set the initial condition as $\eta(0) = [-2, 10, -\frac{\pi}{8}]^\top$ and $\nu(0) = 0_{3\times 1}$ for the AUV in the simulation. We introduce the time-varying and state-dependent disturbance in the earth coordinate as below:

$$f_c(k) = \begin{bmatrix} 0.1\sin(\nu(3)) \\ 0.0\nu(1)\eta(1) + 0.5 \\ -0.1\nu(2)\cos(\eta(3)) + 0.1\sin(v(2)) \end{bmatrix}$$

Accordingly, the disturbance acting on the AUV in the body-fixed frame can be written as

$$w(k) = R^\top(\psi)f_c(k)$$

The constant parameters in the controller are selected as $\alpha = 0.6$, $\alpha_0 = 0.5$, $\alpha_\alpha = 0.2$, $\alpha_\eta = 0.25$, $\alpha_c = 0.8$, $\gamma_\eta = 0.25$, and $T_s = 0.01$, which are chosen empirically. The sampling time $T_s$ is chosen to satisfy Shannon's law. The basis functions of the neural network are defined as

$$S_i(Z) := \frac{\mu_i(Z)}{\sum_{j=1}^{3^8}\mu_j(Z)},\ \mu_i := \prod_{j=1}^{8}\nu_j,\ j = 1,...,3^8 \tag{50}$$

where the function $\nu_j$ can be selected from the $j-$th sets of $1/\left(1+e^{(-a_{3j}(Z_j-b_{3j}))}\right)$, $1/\left(1+e^{a_{1j}(Z_j+b_{1j})}\right)$, and $e^{-a_{2j}|Z_j-b_{2j}|^2}$. The maximum control signal in each channel is set as $|\tau_1| \leq 35N$, $|\tau_2| \leq 35N$, and $|\tau_3| \geq 7Nm$.

Primary simulation results are provided in Figs. 3−9, which show the performance with our proposed control is reasonable

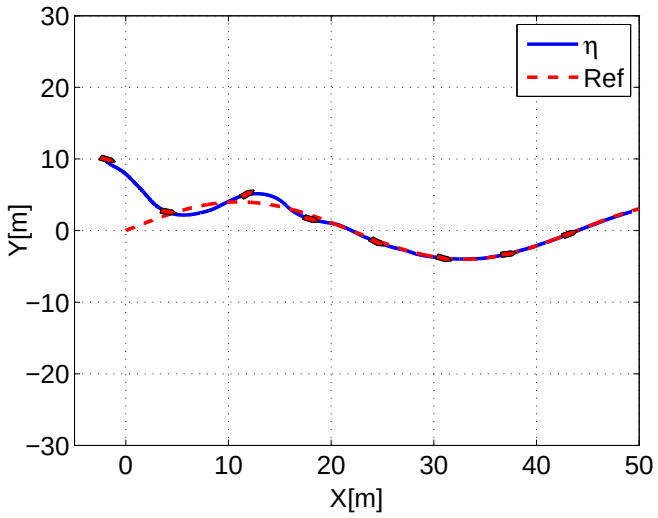


Fig. 3. Trajectory of an AUV in simulation.

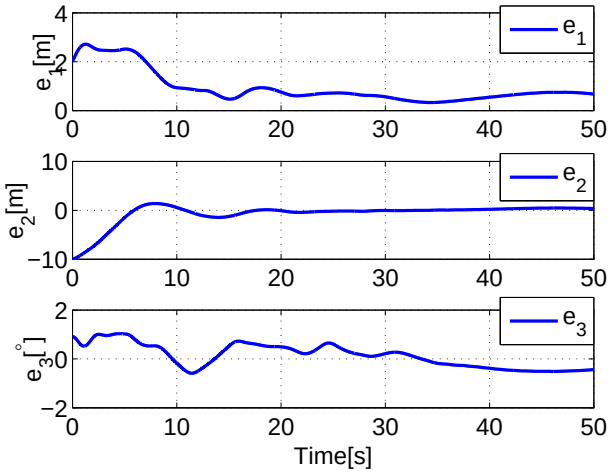


Fig. 4. Tracking error along the trajectory.

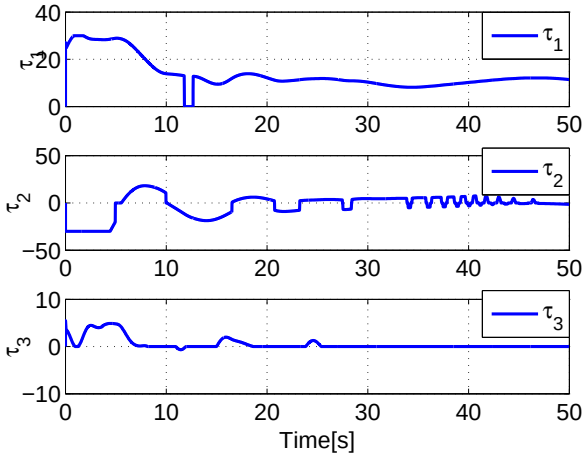


Fig. 5. Control inputs of the AUV.

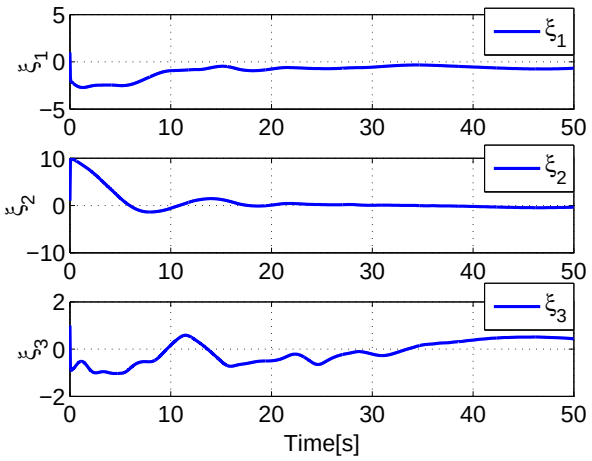


Fig. 6. Compensation control inputs $\xi$.

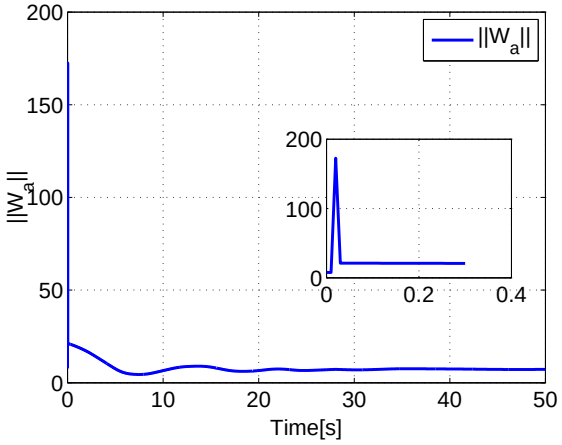


Fig. 7. Norm of action neural weights $\|\hat{W}_a\|$.

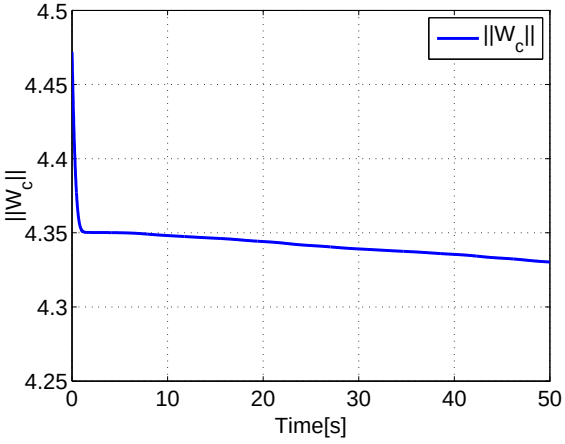


Fig. 8. Norm of critical neural weights $\|\hat{W}_c\|$.

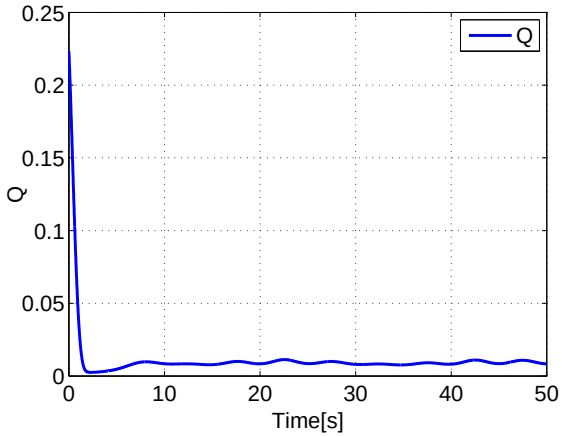


Fig. 9. Reinforcement learning signal.

good. It can be seen that the reference trajectory is tracked in 20 sec in Fig. 3. The tracking error is shown in Fig. 4 with a small boundary. It is clear that the norms of the NN weights are bounded as well as the control inputs in Figs. 7 and 6. In addition, Fig. 9 provides the reinforcement learning signal. It can be seen that $Q$ is bounded near zero, which means the weighted tracking error is also bounded to zero. From the figures, we can find that the performance of the AUV trajectory tracking is satisfactory, despite the unknown dynamics, control input nonlinearities and time-varying disturbance.

### A. *Compared with General NN control*

In order to evaluate our proposed adaptive control with reinforcement learning, we also provide the comparison between the general neural networks control and PD control. The results are shown in Fig. 10-Fig. 13. In Fig. 10, we can see that AUVs track the desired trajectory in a more effective manner with our proposed adaptive reinforcement learning control, i.e., the reference trajectroy is tracked well enough when $x = 15$ with our control, however it is almost $x = 25$ with general NN control. Fig. 11 shows the comparison of the error, from which we can see that our control has faster convergence. It means the learning time needed by the NNs can be reduced by our control.

### B. *Compared with PD control*

Results by PD controller are also given in Fig. 12 and Fig. 13. We can see that $\eta$ has steady state error because of the absence of the integrating although the convergence can also be achieved. However, the parameters are chosen empirically, which is generally difficult to be chosen in real applications. On the other hand, the parameters of PD control have great influence on the effectiveness of AUVs system. The parameters of PD control is set to be $Kp = [2,2,0.1]^\top$ $Kd = [10,10,5]^\top$ and $Kp = [10,10,0.5]^\top$ $Kd = [1,1,0.5]^\top$ respectively. The main difference is that the steady state error is smaller in PD2. However, they both have poorer performance compared with our adaptive control.

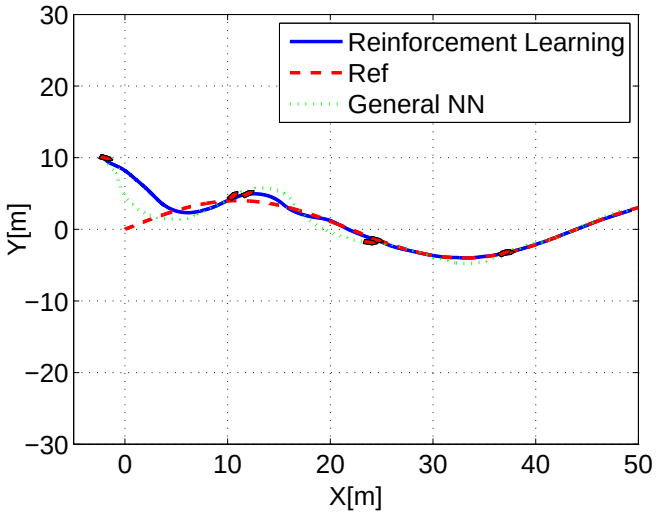


Fig. 10. Trajectory of an AUV compared with general NNs.

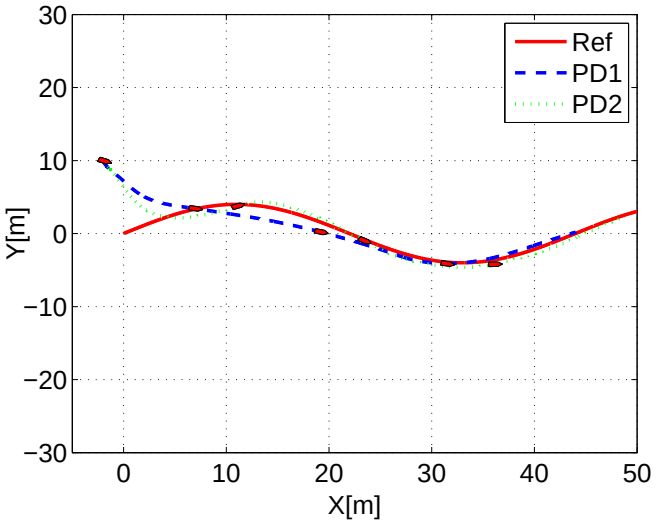


Fig. 12. Trajectory of an AUV compared with pd controller.

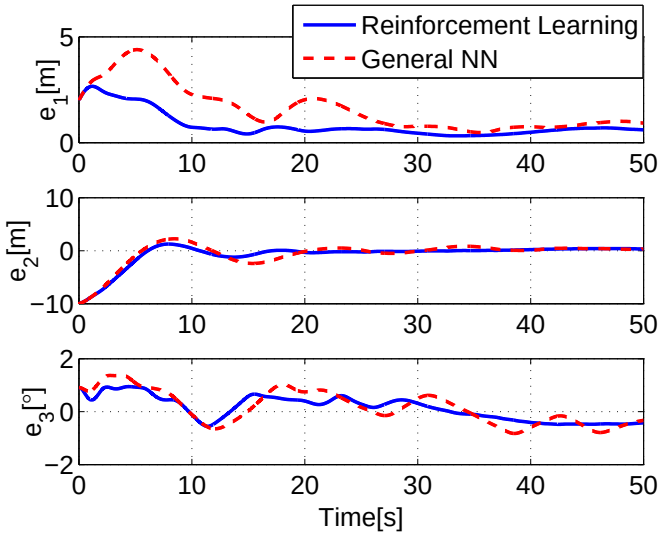


Fig. 11. Tracking error compared with general NNs.

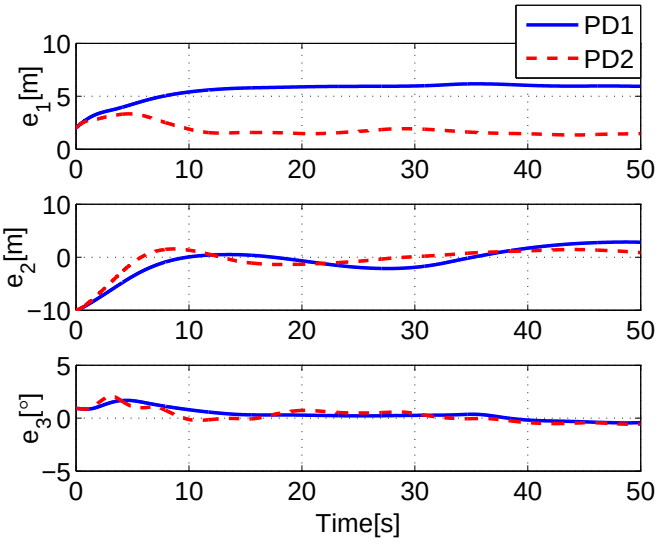


Fig. 13. Tracking error compared with pd controller. $Kp = [2,2,0.1]^\top$ $Kd = [10,10,5]^\top$ and $Kp = [10,10,0.5]^\top$ $Kd = [1,1,0.5]^\top$ respectively

## V. Conclusion

In this work, an adaptive trajectory tracking control law using NN approximation for a fully actuated AUV has been developed in discrete-time domain. An NN based reinforcement learning algorithm has been involved to overcome unknown disturbances, parameter uncertainties and control input nonlinearities. Two NNs are embeded in the proposed controller, whereas the first critical NN is used to evaluate the long time performance of the control in current time step, and the second action NN is used to compensate for the unknown dynamics. Rigorous theoretical analysis and extensive simulation studies have been performed to show the robustness and effectiveness of the proposed approach. One of the future research directions is to apply the proposed control onto the practical systems.

## References


[1] R. Kobayashi and S. Okada, "Development of hovering control system for an underwater vehicle to perform core internal inspections," *Journal of Nuclear Science and Technology*, vol. 53, pp. 566–573, APR 2 2016.

[2] J. M. Selvakumar and T. Asokan, "Station keeping control of underwater robots using disturbance force measurements," *Journal of Marine Science and Technology*, vol. 21, no. 1, pp. 70–85, 2016.

[3] S. Jin, J. Kim, J. Kim, and T. Seo, "Six-degree-of-freedom hovering control of an underwater robotic platform with four tilting thrusters via selective switching control," *IEEE/ASME Transactions on Mechatronics*, vol. 20, pp. 2370–2378, Oct 2015.

[4] Y. Shen, K. Shao, W. Ren, and Y. Liu, "Diving control of autonomous underwater vehicle based on improved active disturbance rejection control approach," *Neurocomputing*, vol. 173, Part 3, pp. 1377 – 1385, 2016.

[5] X. Xiang, C. Yu, Q. Zhang, and G. Xu, "Path-Following Control of an AUV: Fully Actuated Versus Under-actuated Configuration," *Marine Technology Society Journal*, vol. 50, pp. 34–47, JAN-FEB 2016.

[6] B. M. Ferreira, A. C. Matos, and N. A. Cruz, "Modeling and control of trimares auv," *Robotica*, pp. 57–62, 2012.

[7] B. Subudhi, K. Mukherjee, and S. Ghosh, "A static output feedback control design for path following of autonomous underwater vehicle in vertical plane," *Ocean Engineering*, vol. 63, pp. 72–76, 2013.

[8] I. S. Akkizidis, G. N. Roberts, P. Ridao, and J. Batlle, "Designing a Fuzzy-like PD controller for an underwater robot," *Control Engineering Practice*, vol. 11, no. 1, pp. 471–480, 2003.

[9] J. E. Refsnes, A. J. Sørensen, and K. Y. Pettersen, "Model-Based Output Feedback Control of Slender-Body Underactuated AUVs: Theory and Experiments," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 5, pp. 930–946, 2008.

[10] J. Biggs and W. Holderbaum, "Optimal Kinematic Control of an Autonomous Underwater Vehicle," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1623–1626, 2009.

[11] B. Geranmehr and S. R. Nekoo, "Nonlinear suboptimal control of fully coupled non-affine six-DOF autonomous underwater vehicle using the state-dependent riccati equation," *Ocean Engineering*, vol. 96, pp. 248–257, 2015.

[12] H. Pan and M. Xin, "Depth control of autonomous underwater vehicles using indirect robust control method," *International Journal of Control*, vol. 85, no. 1, p. 98C113, 2012.

[13] R. Cui, S. S. Ge, V. E. B. How, and Y. S. Choo, "Leader-follower formation control of underactuated autonomous underwater vehicles," *Ocean Engineering*, vol. 37, no. 17–18, pp. 1491–1502, 2010.

[14] J. Li and P. Lee, "Design of an adaptive nonlinear controller for depth control of an autonomous underwater vehicle," *Ocean Engineering*, vol. 32, pp. 2165–2181, 2005.

[15] S. Liu, D. Wang, and E. Pho, "Non-linear output feedback tracking control for AUVs in shallow wave disturbance condition," *International Journal of Control*, vol. 81, no. 11, pp. 930–946, 2008.

[16] S.-L. Dai, M. Wang, C. Wang, and L. Li, "Learning from adaptive neural network output feedback control of uncertain ocean surface ship dynamics," *International Journal of Adaptive Control and Signal Processing*, 2012.

[17] M. H. Khodayari and S. Balochian, "Modeling and control of autonomous underwater vehicle (AUV) in heading and depth attitude via self-adaptive fuzzy PID controller," *Journal of Marine Science and Technology*, vol. 20, no. 3, pp. 559–578, 2015.

[18] R. P. Kumar, A. Dasgupta, and C. S. Kumar, "Robust trajectory control of underwater vehicles using time delay control law," *Ocean Engineering*, vol. 34, pp. 842–849, 2007.

[19] R. P. Kumar, C. S. Kumar, D. Sen, and A. Dasgupta, "Discrete time-delay control of an autonomous underwater vehicle: Theory and experimental results," *Ocean Engineering*, vol. 36, no. 1, pp. 74–81, 2009.

[20] J. Kim, H. Joe, S. C. Yu, J. S. Lee, and M. Kim, "Time-delay controller design for position control of autonomous underwater vehicle under disturbances," *IEEE Transactions on Industrial Electronics*, vol. 63, pp. 1052–1061, Feb 2016.

[21] N. Fischer, D. Hughes, P. Walters, E. M. Schwartz, and W. E. Dixon, "Nonlinear RISE-Based Control of an Autonomous Underwater Vehicle," *IEEE Transactions on Robotics*, vol. 30, pp. 845–852, AUG 2014.

[22] J. Xu, M. Wang, and L. Qiao, "Dynamical sliding mode control for the trajectory tracking of underactuated unmanned underwater vehicles," *Ocean Engineering*, vol. 105, pp. 54 – 63, 2015.

[23] M. Chen, P. Shi, and C.-C. Lim, "Robust constrained control for mimo nonlinear systems based on disturbance observer," *IEEE Transactions on Automatic Control*, vol. 60, no. 12, pp. 3281–3286, 2015.

[24] M. Chen and J. Yu, "Disturbance observer-based adaptive sliding mode control for near-space vehicles," *Nonlinear Dynamics*, vol. 82, no. 4, pp. 1671–1682, 2015.

[25] N. Wang, C. Qian, J.-C. Sun, and Y.-C. Liu, "Adaptive robust finite-time trajectory tracking control of fully actuated marine surface vehicles," 2015.

[26] K. Shojaei and M. M. Arefi, "On the neuro-adaptive feedback linearising control of underactuated autonomous underwater vehicles in three-dimensional space," *IET Control Theory and Applications*, vol. 9, pp. 1264–1273, MAY 15 2015.

[27] B. S. Park, "Neural Network-Based Tracking Control of Underactuated Autonomous Underwater Vehicles With Model Uncertainties," *Journal Of Dynamic Systems Measurement and Control-Transactions of The ASME*, vol. 137, FEB 2015.

[28] J.-H. Li, P.-M. Lee, and S.-J. Lee, "Motion control of an auv using a neural network adaptive controller," in *Proceedings of the 2002 International Symposium on Underwater Technology, 2002*, pp. 217–221, 2002.

[29] Y.-J. Liu, L. Tang, S. Tong, C. Chen, and D.-J. Li, "Reinforcement learning design-based adaptive tracking control with less learning parameters for nonlinear discrete-time mimo systems," *IEEE Transactions on Neural Networks and Learning Systems,*, vol. 26, no. 1, pp. 165–176, 2015.

[30] Z. Peng, D. Wang, Y. Shi, H. Wang, and W. Wang, "Containment control of networked autonomous underwater vehicles with model uncertainty and ocean disturbances guided by multiple leaders," *Information Sciences*, vol. 316, pp. 163–179, 2015.

[31] J. H. Li, P. M. Lee, and B. H. Jun, "Application of a robust adaptive controller to autonomous diving control of an auv," in *30th IEEE Annual Conference on Industrial Electronics Society, 2004. IECON 2004*, vol. 1, pp. 419–424, Nov 2004.

[32] Z. Peng, D. Wang, H. Zhang, and Y. Lin, "Cooperative output feedback adaptive control of uncertain nonlinear multi-agent systems with a dynamic leader," *Neurocomputing*, vol. 149, pp. 132–141, 2015.

[33] M. Chen and S. S. Ge, "Adaptive neural output feedback control of uncertain nonlinear systems with unknown hysteresis using disturbance observer," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 12, pp. 7706–7716, 2015.

[34] Y.-J. Liu and S. Tong, "Adaptive nn tracking control of uncertain nonlinear discrete-time systems with nonaffine dead-zone input," *IEEE Transactions on Cybernetics*, vol. 45, no. 3, pp. 497–505, 2015.

[35] Y.-J. Liu, C. P. Chen, G.-X. Wen, and S. Tong, "Adaptive neural output feedback tracking control for a class of uncertain discrete-time nonlinear systems," *IEEE Transactions on Neural Networks*, vol. 22, no. 7, pp. 1162–1167, 2011.

[36] Z. Li, C. Yang, N. Ding, S. Bogdan, and T. Ge, "Robust adaptive motion control for underwater remotely operated vehicles with velocity constraints," *International Journal of Control, Automation and Systems*, vol. 10, no. 2, pp. 421–429, 2012.

[37] F. Rezazadegan, K. Shojaei, F. Sheikholeslam, and A. Chatraei, "A novel approach to 6-DOF adaptive trajectory tracking control of an AUV in the presence of parameter uncertainties," *Ocean Engineering*, vol. 107, pp. 246 – 258, 2015.

[38] T. P. Zhang and S. S. Ge, "Adaptive dynamic surface control of nonlinear systems with unknown dead zone in pure feedback form ," *Automatica*, vol. 44, no. 7, pp. 1895–1903, 2008.

[39] P. He and S. Jagannathan, "Reinforcement Learning Neural-Network-Based Controller for Nonlinear Discrete-Time Systems With Input Constraints," *IEEE Transactions on System, Man and Cybernetics, Part B*, vol. 37, no. 2, pp. 425–436, 2007.

[40] B. Xu, C. Yang, and Z. Shi, "Reinforcement learning output feedback nn control using deterministic learning technique," *IEEE Transactions on Neural Networks and Learning Systems*.

[41] R. Cui, C. Yang, Y. Li, and S. Sharma, "Neural network based reinforcement learning control of autonomous underwater vehicles with control input saturation," in *2014 UKACC International Conference on Control (CONTROL)*, pp. 50–55, 2014.

[42] C. Yang, S. S. Ge, C. Xiang, T. Y. Chai, and T. H. Lee, "Output feedback nn control for two classes of discrete-time systems with unknown control directions in a unified approach," *IEEE Transactions on Neural Networks*, vol. 19, no. 11, pp. 1873–1886, 2008.

[43] S. Ibrir, F. Xie, and Su, "Adaptive tracking of nonlinear systems with non-symmetric dead-zone input," *Automatica*, vol. 43, no. 3, pp. 522–530, 2007.

[44] Y. J. Liu and S. Tong, "Adaptive nn tracking control of uncertain nonlinear discrete-time systems with nonaffine dead-zone input," *Cybernetics IEEE Transactions on*, vol. 45, no. 3, pp. 497–505, 2014.

[45] S. S. Ge, T. H. Lee, and C. J. Harris, *Adaptive Neural Network Control of Robotic Manipulators*. London: World Scientific, 1998.

[46] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive System*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[47] D. V. Prokhorov and D. C. Wunsch, "Adaptive critic designs.," *IEEE Transactions on Neural Networks*, vol. 8, no. 5, pp. 997–1007, 1997.

[48] J. Si and Y. T. Wang, "On-line learning control by association and reinforcement," *IEEE Transactions on Neural Networks*, vol. 12, no. 2, pp. 264–76, 2001.

[49] P. He and S. Jagannathan, "Discrete-time neural network control of nonlinear systems in non-strict feedback form," *Proceedings of the 42nd IEEE Conference on Decision and Control*.