2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Neural Network-based Optimal Control for Trajectory Tracking of a Helicopter UAV

David Nodland, H. Zargarzadeh, and S. Jagannathan

Abstract— **Helicopter unmanned aerial vehicles (UAVs) may be widely used for both military and civilian operations. Because these helicopters are underactuated nonlinear mechanical systems, high-performance controller design for them presents a challenge. This paper presents an optimal controller design for trajectory tracking of a helicopter UAV using a neural network (NN). The state-feedback control system utilizes the backstepping methodology, employing kinematic and dynamic controllers. The online approximator-based dynamic controller learns the infinite-horizon Hamilton-Jacobi-Bellman (HJB) equation in continuous time and calculates the corresponding optimal control input to minimize the HJB equation forward-in-time. Optimal tracking is accomplished with a single NN utilized for cost function approximation. The overall closed-loop system stability is demonstrated using Lyapunov analysis, with the position, orientation, angular and translational velocity tracking errors, and NN weight estimation errors uniformly ultimately bounded (UUB) in the presence of bounded disturbances and NN functional reconstruction errors.**

I. INTRODUCTION

Due to their versatility and maneuverability, unmanned helicopters are invaluable for applications where human intervention may be restricted. For unmanned helicopter control [1], it is essential to produce moments and forces on the helicopter such that the desired regulated state is achieved and so that the helicopter can track a desired trajectory. The dynamics of the helicopter UAV are nonlinear, coupled with each other, and underactuated, which makes the control design very challenging.

In order to develop the controllers for such unmanned helicopters, Koo and Sastry [1] have utilized an approximate linearization-based control scheme [1] that transforms the system into linear form. Mettler et al. [2] have introduced a model for the helicopter independent of an accompanying control scheme [2]. Hovakimyan et al. [3] have implemented an output feedback control scheme with a neural network (NN)-based controller using feedback linearization [3]. Johnson and Kannan [4] have employed an inner and outer loop control using pseudo-control hedging [4], and Ahmed et al. [5] have introduced a backstepping-based controller for the helicopter. Frazzoli [6] and Mahoney [7] have both generated control schemes for Lyapunov-based control of helicopter UAVs. However, none of these schemes [1]-[7] present the optimal control of the unmanned helicopter.

Although optimal control of linear systems can be achieved by solving the Riccati equation [15], optimal control of nonlinear systems often requires solving the nonlinear Hamilton-Jacobi-Bellman (HJB) equation, which does not have a closed-form solution. Therefore, Enns and Si [8] have used neural network dynamic programming-based optimal control of a helicopter UAV. This optimal controller uses offline training. Stability of the control scheme is not included. Lee et al. [9] introduced a robust command augmentation system using a NN, but inversion errors can lead to problems [9].

Recently, Dierks and Jagannathan [10] introduced an optimal controller for nonlinear discrete-time systems in affine form. Here, the discrete-time Hamilton-Jacobi-Bellman (HJB) equation is solved online. An online approximator (OLA) such as a NN learns the HJB equation, with a second OLA utilized to minimize the cost (HJB) function. Dierks and Jagannathan [11] have extended this scheme to continuous-time systems by using a single online approximator (SOLA). The present work is partially derived from a modified form of this approach.

Therefore, a SOLA-based scheme for the optimal tracking control of a helicopter's nonlinear continuous-time feedback system has been considered in this paper. The dynamic controller learns the continuous-time HJB equation and then calculates the optimal control input to minimize the HJB equation forward-in-time. The proposed tracking controller consists of a single NN for approximating the cost function with the NN weights tuned online. Lyapunov analysis is utilized to demonstrate the stability of the closed-loop system.

II. BACKGROUND

Consider the helicopter shown in Figure 1 with six degrees of freedom (DOF) defined in the inertial coordinate frame $\mathcal{Q}^a$, where its position coordinates are given by $\rho = [x, y, z] \in \mathcal{Q}^a$ and its orientation described as yaw, pitch, and roll, respectively, is given by $\Theta = [\phi, \theta, \psi] \in \mathcal{Q}^a$. The equations of motion are expressed in the body fixed frame $\mathcal{Q}^b$ which is associated with the helicopter's center of mass. The ${}^b x$-axis is defined parallel to the helicopter's

David Nodland, H. Zargarzadeh, and S. Jagannathan are with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO 65409. Contact author e-mail: sarangap@mst.edu.

The authors wish to thank Arpita Ghosh for her assistance with this project. Arpita Ghosh is with the National Metallurgical Laboratory, Jamshedpur-831007, India.

Acknowledgement: Research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-10-2-0077. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

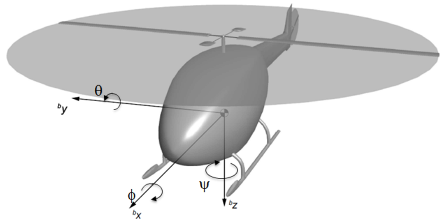

Fig. 1. Helicopter Dynamics

direction of travel and the ^{b}y-axis is defined perpendicular to the helicopter's direction of travel, while the ^{b}z-axis is defined as projecting orthogonally downwards from the xy-plane of the helicopter. The dynamics of the helicopter is given by the Newton-Euler equation in the body fixed frame and can be written as [7]

$$\begin{bmatrix} mI & 0 \\ 0 & \mathcal{J} \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{\omega} \end{bmatrix} + \begin{bmatrix} 0 \\ \omega \times \mathcal{J}\omega \end{bmatrix} = \begin{bmatrix} F \\ \tau \end{bmatrix} \tag{1}$$

where $m \in \mathbb{R}$ is a positive scalar denoting the mass of the helicopter, $F \in \mathbb{R}^{3\times1}$ is the body force applied to the helicopter's center of mass, $\tau \in \mathbb{R}^{3\times1}$ is the body torque applied about the helicopter's center of mass, $v = [v_x, v_y, v_z] \in \mathbb{R}^{3\times1}$ represents the translational velocity vector, $\omega = [\omega_x, \omega_y, \omega_z] \in \mathbb{R}^{3\times1}$ represents the body angular velocity vector, $I \in \mathbb{R}^{3\times3}$ is the identity matrix, and $\mathcal{J} \in \mathbb{R}^{3\times3}$ is the positive-definite inertia matrix. The kinematics of the helicopter are given by

$$\dot{\rho} = Rv \tag{2}$$

and

$$\dot{\Theta} = T\omega \tag{3}$$

The translational rotation matrix used to relate a vector in the body fixed frame to the inertial coordinate frame is defined as [12]

$$R(\Theta) = \begin{bmatrix} c_\theta c_\psi & s_\phi s_\theta c_\psi - c_\phi s_\psi & c_\phi s_\theta c_\psi + s_\theta s_\psi \\ c_\theta s_\psi & s_\phi s_\theta s_\psi + c_\phi c_\psi & c_\phi s_\theta s_\psi - s_\phi c_\psi \\ -s_\theta & s_\phi c_\theta & c_\phi c_\theta \end{bmatrix}$$

where $s_\bullet$ and $c_\bullet$ denote the $\sin(\bullet)$ and $\cos(\bullet)$ functions, respectively. The rotational transformation matrix from the body fixed frame to the inertial coordinate frame is defined as

$$T(\Theta) = \begin{bmatrix} 1 & s_\phi t_\theta & c_\phi t_\theta \\ 0 & c_\phi & -s_\phi \\ 0 & \frac{s_\phi}{c_\theta} & \frac{c_\phi}{c_\theta} \end{bmatrix}$$

where $t_\bullet$ has been used to represent $tan(\bullet)$. The transformation matrix is bounded according to $\|T\|_F < T_{max}$ for a known constant T_{max}, provided $-\pi/2 < \phi < \pi/2$ and $-\pi/2 < \theta < \pi/2$ such that the helicopter trajectory does not pass through any singularities [1]. Here, it is necessary to mention that $\|R\|_F = R_{max}$ for a known constant R_{max} and $R^{-1} = R^T$. Let the mass-inertia matrix M be defined as $M = diag\{mI, \mathcal{J}\}$.

Now, (1) can be rewritten in the form given in [12], but with dynamics as given in [7] as

$$M \begin{bmatrix} \dot{v} \\ \dot{\omega} \end{bmatrix} = \bar{S}(\omega) + \begin{bmatrix} 0^{3\times1} \\ \mathcal{N}_2 \end{bmatrix} + \begin{bmatrix} G(R) \\ 0^{3\times1} \end{bmatrix} + U + \tau_d \tag{4}$$

where $\bar{S}(\omega) = [0^{3\times1}, -\omega \times J\omega]^T$, $\mathcal{N}_2 \in \mathbb{R}^{3\times1}$ represents nonlinear aerodynamic effects, $G(R) \in \mathbb{R}^{3\times1}$ represents the gravity vector and is defined as $G(R) = m\bar{g}e_3$ with $\bar{g}$ the gravitational acceleration and m the helicopter's mass, $U \in \mathbb{R}^{6\times1}$ is the control input vector, with u providing the thrust in the z-direction, w_1, w_2, and w_3 providing the rotational torques in $x-$, $y-$ and $z-$ directions, respectively, and $\tau_d = [\tau_{d1}^T, \tau_{d2}^T]^T$ represents unknown bounded disturbances such that $\|\tau_d\| < \tau_M$ for all time t, with τ_M a known positive constant. Note that $(\times)$ denotes the vector cross product. The nonlinear aerodynamic effects taken into consideration for modeling of the helicopter are given by $\mathcal{N}_2 = Q_M e_3 - Q_T e_2$, with Q_M and Q_T aerodynamic constants originally found in [7]. Note that e_1, e_2, and e_3 are unit vectors directed along the x-, y-, and z-axes, respectively, in the inertial reference frame.

$$U = \begin{bmatrix} E_3^a & 0^{3\times3} \\ 0^{3\times1} & diag([p_{11}\ p_{22}\ p_{33}]) \end{bmatrix} \begin{bmatrix} \hat{u} \\ \hat{w}_1 \\ \hat{w}_2 \\ \hat{w}_3 \end{bmatrix}$$

where p_{ii} are positive definite constants, and with $E_3^a = [0\ 0\ 1]^T$. Defining the new augmented variables $X = [\rho^T\ \Theta^T]^T \in \mathbb{R}^{6\times1}$ and $V = [v^T\ \omega^T]^T \in \mathbb{R}^{6\times1}$, (4) can be rewritten employing the backstepping technique in the form given by

$$\dot{X} = AV + \xi \tag{5}$$
$$\dot{V} = f(V) + \bar{U} \tag{6}$$

where $f(V) = M^{-1}(\bar{S}(\omega) + [0^{3\times1}\ \mathcal{N}_2]^T) + \bar{G}$ with $\bar{G} = M^{-1}[G(R)\ 0^{3\times1}]^T \in \mathbb{R}^{6\times1}$, $\bar{U} = M^{-1}U$, with $\xi \in \mathbb{R}^{6\times1}$ the bounded sensor measurement noise such that $\|\xi\| \leq \xi_M$ for a known constant ξ_M. Equation (5) is in the body fixed frame, with equation (6) bringing the dynamics back to the earth frame. Note that these last two equations take the form

$$\dot{x_1} = f_1(x_1) + g_1(x_1)x_2 + \xi$$
$$\dot{x_2} = f_2(x_2) + g_2(x_2)u$$

with $f_1(x_1) = 0$. This system is a candidate for backstepping control [13]. Also,

$$A = \begin{bmatrix} R & 0^{3\times3} \\ 0^{3\times3} & -R^x \end{bmatrix}$$

where R^x denotes a skew-symmetric representation of the rotation matrix. In this section, the dynamic model of the helicopter with six degrees-of-freedom (DOF) and four inputs has been presented. The methodology for the controller design will now be considered.

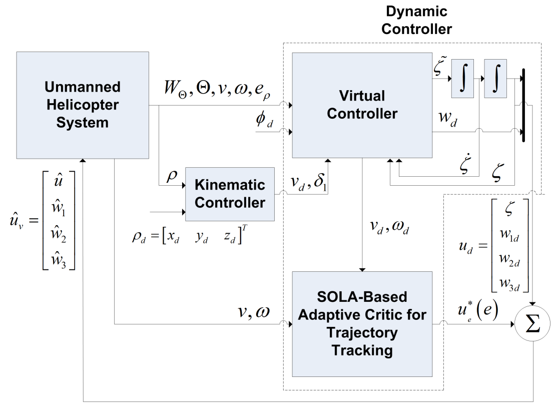

Fig. 2. Control Scheme for Optimal Tracking

III. METHODOLOGY

A. Nonlinear Optimal Tracking of the Unmanned Helicopter

The overall control objective for the unmanned helicopter is to track a desired trajectory $X_d(t)$ and a desired heading (yaw) while maintaining stable flight. The universal approximation property of NNs may be used in the design of the dynamic controller for tracking the desired trajectory in an optimal manner. In Figure 2, the entire NN-based control scheme for optimal tracking of the desired trajectory by the helicopter is illustrated. Note that the dynamic controller is comprised of the items within the dashed boundary and that the virtual controller will be addressed later as part of the dynamic controller.

B. Kinematic Controller

To design the kinematic controller for the unmanned helicopter, the tracking error for the position must first be defined. The position tracking error is given by

$$\delta_1 = \rho_d - \rho \ \in \ \mathcal{Q}^a \tag{7}$$

Also, it is essential to define $v = \dot{\rho}$, which then yields the desired velocity, v_d as in [7] as $v_d = v - \delta_1/m$. In addition, it is important to note that there exist desired trajectories which may reach unstable operating regions as the orientation about the x- and y- axes approaches $\pm\pi/2$.

C. Hamilton-Jacobi-Bellman Equation

In this section the optimal control input u_e^* is designed to ensure that the unmanned helicopter system in (4) tracks a desired trajectory $X_d(t)$. For optimal tracking, the desired dynamics are defined as

$$\dot{V}_d = f(V_d) + gu_d(V_d) \tag{8}$$

where $f(V_d) \in \mathbb{R}^{6\times1}$ is the internal dynamics of the helicopter system rewritten in terms of the desired state $V_d \in \mathbb{R}^{6\times1}$, g is such that $gu_V = M^{-1}U \in \mathbb{R}^{6\times1}$ is bounded satisfying $g_{min} \leq \|g\|_F \leq g_{max}$, and $u_d(V_d)$ is the desired control input corresponding to the desired states. It has been assumed that the system is observable and controllable, with $e = 0$ a unique equilibrium point on compact set $\Upsilon \subset \mathbb{R}^{6\times1}$. Under these conditions, the optimal control input for the unmanned helicopter system given in (8) can be determined [15]. Next, the state tracking error is defined as

$$e = V - V_d \tag{9}$$

and considering the actual dynamics $\dot{V} = f(V) + gu_V$, the tracking error dynamics in (9) can be written as

$$\dot{e} = f(V) + gu_V - \dot{V}_d = f_e(e) + gu_e \tag{10}$$

where $f_e(e) = f(V) - f(V_d)$ and $u_e = u_V - u_d$. In order to control (10) in an optimal manner, the control policy u_e should be selected such that it minimizes the cost function given by

$$W_T(e(t)) = \int_t^{\infty} r(e(\tau), u_e(\tau))d\tau \tag{11}$$

where $r(e(\tau), u_e(\tau)) = Q(e) + u_e^T B_e u_e$, $Q(e) > 0$ is the penalty on the states, with $B_e \in \mathbb{R}^{6\times6}$ a positive semi-definite matrix. After this, the Hamiltonian for the HJB tracking problem is defined as

$$H_T(e, u_e) = r(e, u_e) + W_{Te}^T(e)(f_e(e) + gu_e) \tag{12}$$

where $W_{Te}(e)$ is the gradient of $W_T(e)$ with respect to e. The basis function used for the neural network law is $\Phi(e) = [\nabla e \ \nabla e^2 \ \nabla e^3 \ \nabla sin(e) \ \nabla sin(2e) \ \nabla tanh(e) \ \nabla tanh(2e)]^T$. Now applying stationary condition $\partial H(e, u_e)/\partial u_e = 0$, the optimal control input is found to be

$$u_e^*(e) = -B_e^{-1} g^T W_{Te}^*(e)/2 \tag{13}$$

with $u_e^*(e) \in \mathbb{R}^4$. Substituting the optimal control input from (13) into the Hamiltonian (12) generates the HJB equation for the tracking problem as

$$\begin{aligned} 0 = Q_e(e) + W_{Te}^{*T}(e) f_e(e) \\ -W_{Te}^{*T}(e) g(e) B_e^{-1} g^T W_{Te}^*(e)/4 \end{aligned} \tag{14}$$

with $W_T^*(0) = 0$. The control input must be selected such that the cost function in (11) is finite, or u_e must be admissible [10]. At this point, Lemma 1 will be introduced.

Lemma 1 (Boundedness of system state errors). Given the unmanned helicopter system with cost function (11) and optimal control input (13), let $J_1(e)$ be a continuously differentiable, radially unbounded Lyapunov candidate function such that $\dot{J}_1(e) = J_{1e}^T(e)\dot{e} = J_{1e}^T(e)(f_e(e) + gu_e^*) < 0$ with $J_{1e}(e)$ the partial derivative of $J_1(e)$. In addition, let $\bar{Q}(e) \in \mathbb{R}^{6\times6}$ be a positive definite matrix satisfying $\|\bar{Q}(e)\| = 0$ only if $\|e\| = 0$ and $\bar{Q}_{min} \leq \|\bar{Q}(e)\| \leq \bar{Q}_{max}$ for $e_{min} \leq \|e\| \leq e_{max}$ for positive constants $\bar{Q}_{min}$, $\bar{Q}_{max}$, e_{min}, and e_{max}. Also, let $\bar{Q}(e)$ satisfy $\lim_{e\to\infty} \bar{Q}(e) = \infty$ as well as

$$W_e^{*T} \bar{Q}(e) J_{1e} \quad = \quad r(e, u_e^*) \ = \ Q(e) + u_e^{*T} B u_e^* \tag{15}$$

then the following relation is true

$$J_{1e}^T(f_e(e) + gu_e^*) = -J_{1e}^T \bar{Q}(e) J_{1e} \tag{16}$$

Proof: Applying the optimal control input, the cost function becomes $\dot{W}^*(e) = W_e^{*T}(e)\dot{e} = W_e^{*T}(e)(f_e(e)+gu_e^*) = -Q_e(e) - u_e^{*T}B_e u_e^*$. Because

$$(f_e(e) + gu_e^*) = -(W_e^* W_e^{*T})^{-1} W_e^* (Q_e(e) + u_e^{*T} B_e u_e^*) = -(W_e^* W_e^{*T})^{-1} W_e^* W_e^{*T} \overline{Q}_e(e) J_{1e} = -\overline{Q}_e(e) J_{1e}$$

One then has $J_{1e}^T(f_e(e)+gu_e^*) = -J_{1e}^T \bar{Q}_e(e) J_{1e}$, concluding the proof for Lemma 1. It is apparent that an expression including the optimally augmented control input in (13) can be written as

$$\hat{u}_V = u_d - B_e^{-1} g^T W_{Te}^*(e)/2 \tag{17}$$

and the desired feedforward control input u_d is obtained from [7]. Note that this $\hat{u}_V$ becomes the input U which is used as the system input, with the $(\hat{\bullet})$ notation here used to denote an estimate. Next, the SOLA is introduced.

D. Single Online Approximator (SOLA)-Based Optimal Control of Helicopter

In this paper, the adaptive critic for optimal control of a helicopter is realized online using only one OLA. For the SOLA to learn the cost function, the cost function is rewritten using the OLA representation as

$$W(e) = \Gamma^T \Phi(e) + \varepsilon(e) \tag{18}$$

where $\Gamma \in \mathbb{R}^{\mathbb{L}}$ is the constant target OLA vector, $\Phi(e) : \mathbb{R}^n \to \mathbb{R}^L$ is a linearly independent basis vector which satisfies $\Phi(e) = 0$, and $\varepsilon(e)$ is the OLA reconstruction error. The basis vector used in this case is the same as in the previous section. The target OLA vector and reconstruction errors are assumed to be upper bounded according to $\|\Gamma\| \leq \Gamma_M$ and $\|\varepsilon(e)\| \leq \varepsilon_M$, respectively [14]. The gradient of the OLA cost function in (18) is written as

$$\partial W(e)/\partial e = W_e(e) = \nabla_e^T \Phi(e)\Gamma + \nabla_e \varepsilon(e) \tag{19}$$

Using (19), the optimal control input in (13) and the HJB equation in (14) can be written as

$$u_e^* = -B^{-1} g^T \nabla_e^T \Phi(e)\Gamma/2 - B^{-1} g^T \nabla_e \varepsilon(e)/2$$

$$H^*(e, \Gamma) = Q(e) + \Gamma^T \nabla_e \Phi(e) f_e(e) \tag{20}$$
$$-\Gamma^T \nabla_e \Phi(e) C \nabla_e^T \Phi(e)\Gamma/4 + \varepsilon_{HJB} = 0$$

where $C = gB^{-1}g^T > 0$ is bounded such that $C_{min} \leq \|C\| \leq C_{max}$ for known constants C_{min} and C_{max} and
$\varepsilon_{HJB} = \nabla_e \varepsilon^T (f_e(e) - \frac{1}{2} g B^{-1} g^T (\nabla_e^T \Phi(e)\Gamma + \nabla_e \varepsilon)) + \frac{1}{4} \nabla_e \varepsilon^T g B^{-1} g^T \nabla_e\ \varepsilon$
$= \nabla_e \varepsilon^T (f_e(e) + gu_e^*) + \frac{1}{4} \nabla_e \varepsilon^T C \nabla_e \varepsilon$
is the OLA reconstruction error. The OLA estimate of (18) is

$$\hat{W}(e) = \hat{\Gamma}^T \Phi(e) \tag{21}$$

with $\hat{\Gamma}$ the OLA estimate of the target vector Γ. In the same way, the estimate for the optimal control input and the approximate Hamiltonian in (20) in terms of $\hat{\Gamma}$ can be expressed as

$$\hat{u}_e^* = -B^{-1} g^T \nabla_e^T \Phi(e)\hat{\Gamma}/2 \tag{22}$$

Employing (20) and (21), the approximate Hamiltonian may now be written as

$$\hat{H}^*(e, \hat{\Gamma}) = Q(e) + \hat{\Gamma}^T \nabla_e \Phi(e) f_e(e)$$
$$-\hat{\Gamma}^T \nabla_e \Phi(e) C \nabla_e^T \Phi(e)\hat{\Gamma}/4 \tag{23}$$

Recollecting the HJB equation in (12), the OLA estimate $\hat{\Gamma}$ should be tuned to minimize $\hat{H}^*(e, \hat{\Gamma})$. However, merely tuning $\hat{\Gamma}$ to minimize $\hat{H}^*(e, \hat{\Gamma})$ does not ensure the stability of the nonlinear helicopter system during the OLA learning process.

Therefore, the OLA tuning algorithm is designed to minimize (23) while considering the system stability and is given below

$$\begin{aligned} \dot{\hat{\Gamma}} = &-(\alpha_1 \hat{\beta}/(\hat{\beta}^T \hat{\beta} + 1)^2)(Q(e) + \hat{\Gamma}^T \nabla_e \Phi(e) f_e(e) \\ &-\hat{\Gamma}^T \nabla_e \Phi(e) C \nabla_e^T \Phi(e)\hat{\Gamma}/4) \\ &+\Sigma(e, \hat{u}_e)\frac{\alpha_2}{2} \nabla_e \Phi(e) g B^{-1} g^T J_{1e}(e) \end{aligned} \tag{24}$$

where $\hat{\beta} = \nabla_e \Phi(e) f_e(e) - \nabla_e \Phi(e) C \nabla_e^T \Phi(e)\hat{\Gamma}/2$, $\alpha_1 > 0$ and $\alpha_2 > 0$ are design constants, $J_{1e}(e)$ is defined in *Lemma 1*, and the operator $\Sigma(e, \hat{u}_e)$ is given by

$$\begin{aligned} &\Sigma(e, \hat{u}_e) = 0 \quad if \quad J_{1e}^T(e)\dot{e} = J_{1e}^T(e) \\ &(f_e(e) - gB^{-1} g^T \nabla_e^T \Phi(e)\hat{\Gamma}/2) < 0 \\ &1 \quad otherwise \end{aligned} \tag{25}$$

The first term in (24) is the portion of the tuning law which tries to minimize (23) and has been derived using a normalized gradient descent scheme with the auxiliary HJB error defined as below

$$E_{HJB} = (\hat{H}^*(e, \hat{\Gamma}))^2/2 \tag{26}$$

The second term in the OLA tuning law in (24) ensures that the system states remain bounded while the SOLA scheme learns the optimal cost function.

The dynamics of the OLA parameter estimation error is considered as $\tilde{\Gamma} = \Gamma - \hat{\Gamma}$. Since this yields $Q(e) = -\Gamma^T \nabla_e \Phi(e) f_e(e) + \Gamma^T \nabla_e \Phi(e) C \nabla_e^T \Phi(e)\Gamma/4 - \varepsilon_{HJB}$ from (20), the approximate HJB equation in (23) can be expressed in terms of $\tilde{\Gamma}$ as

$$\begin{aligned} \hat{H}(e, \hat{\Gamma}) = &-\tilde{\Gamma}^T \nabla_e \Phi(e) f_e(e) + \frac{1}{2} \tilde{\Gamma}^T \nabla_e \Phi(e) C \nabla_e^T \Phi(e)\Gamma \\ &-\frac{1}{4} \tilde{\Gamma}^T \nabla_e \Phi(e) C \nabla_e^T \Phi(e)\tilde{\Gamma} - \varepsilon_{HJB} \end{aligned} \tag{27}$$

Then, since $\dot{\tilde{\Gamma}} = \dot{\hat{\Gamma}}$ and $\hat{\beta} = \nabla_e \Phi(e)(\dot{e}^* + C\nabla_e \varepsilon/2) + \nabla_e \Phi(e) C \nabla_e^T \Phi(e)\tilde{\Gamma}/2$, where $\dot{e} = f_e(e) + gu_e^*$, the error dynamics of (24) are

$$\begin{aligned} \dot{\tilde{\Gamma}} \quad = \quad &\frac{\alpha_1}{\rho_1^2}(\nabla_e \Phi(e)(\dot{e}_1^* + \frac{C\nabla_e \varepsilon}{2}) + \frac{\nabla_e \Phi(e) C \nabla_e^T \Phi(e)\tilde{\Gamma}}{2}) \\ &(\tilde{\Gamma}^T \nabla_e \Phi(e)(\dot{e}_1^* + \frac{C\nabla_e \varepsilon}{2}) + \frac{\tilde{\Gamma}^T \nabla_e \Phi(e) C \nabla_e^T \Phi(e)\tilde{\Gamma}}{2} \\ &+\varepsilon_{HJB}) - \Sigma(e, \hat{u}_e)\frac{\alpha_2}{2} \nabla_e \Phi(e) g B^{-1} g^T J_{1e}(e) \end{aligned} \tag{28}$$

where $\rho_1 = (\hat{\beta}^T \hat{\beta} + 1)$. Next, it is necessary to examine the stability of the SOLA-based adaptive scheme for optimal control along with the stability of the helicopter system.

Definition: An equilibrium point e_e is said to be uniformly ultimately bounded (UUB) if there exists a compact set $S \subset \mathbb{R}^n$ such that for every $e_0 \in S$ there exists a bound D and time $T(D, e_0)$ such that $\|e(t) - e_e\| \leq D$ for all $t \geq t_0 + T$.

Theorem 2 (SOLA-based scheme for convergence to the HJB function and system stability). Given the unmanned helicopter system with target HJB equation (14), let the tuning law for the SOLA be given by (24). Then there exist constants b_{Je} and b_Γ such that the OLA approximation error $\tilde{\Gamma}$ and $\|J_{1e}(e)\|$ are UUB for all $t \geq t_0 + T$ with ultimate bounds given by $\|J_{1e}(e)\| \leq b_{Je}$ and $\|\tilde{\Gamma}\| \leq b_\Gamma$. Further, OLA reconstruction error $\|W^* - \hat{W}\| \leq \varepsilon_{r1}$ and $\|u_e^* - \hat{u}_e\| \leq \varepsilon_{r2}$ for small positive constants ε_{r1} and ε_{r2}, respectively. Proof will be provided later for the tracking case.

E. NN Control Scheme for the Dynamic Controller

The next step will be to consider how to obtain $u_d = [\zeta \; w_{1d} \; w_{2d} \; w_{3d}]^T$. This is done by obtaining $\tilde{w}_{1d}$, $\tilde{w}_{2d}$, $\tilde{w}_{3d}$, and $\tilde{\zeta}$ (with $\tilde{\zeta}$ obtained recursively) with the equations below [7]

$$\begin{bmatrix} \tilde{w}_{1d} \\ \tilde{w}_{2d} \\ \tilde{\zeta} \end{bmatrix} = \begin{bmatrix} 0 & \zeta & 0 \\ -\zeta & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} R(\Theta)^T(\dot{Y}_d - 2\dot{\zeta}R(\Theta)skew(\omega)e_3 + \delta_3 + \delta_4)$$

and

$$\tilde{w}_{3d} = \tfrac{c_\theta}{c_\psi}(\ddot{\hat{\phi}} - \epsilon_4 - \epsilon_3 + e_1^T W_\Theta^{-1}\dot{W}_\Theta W_\Theta^{-1}\omega - \tfrac{s_\psi}{c_\theta}\tilde{w}_{2d})$$

where Y_d is defined as

$$Y_d = \delta_2 + \delta_3 + \tfrac{d}{dt}(m\bar{g}e_3 - m\dot{v}_d + \delta_2 + \tfrac{1}{m}\delta_1)$$

with $\delta_1 = \rho_d - \rho$, $\delta_2 = m(v - v_d)$, $\epsilon_3 = \phi_d - \phi$, $\epsilon_4 = \dot{\phi} - \dot{\phi}_d$, as well as

$$\delta_3 = m\bar{g}e_3 - m\dot{v}_d + \delta_2 + \tfrac{1}{m}\delta_1 - \zeta R(\Theta)e_3$$

$$\delta_4 = Y_d - (\dot{\zeta}R(\Theta)e_3 + \zeta R(\Theta)skew(\omega)e_3)$$

$$W_\Theta = \begin{bmatrix} -s_\theta & 0 & 1 \\ c_\theta s_\psi & c_\psi & 0 \\ c_\theta c_\psi & -s_\psi & 0 \end{bmatrix}$$

and from the kinematic controller

$$v_d = v - \tfrac{1}{m}\delta_1$$

Now the real inputs must be obtained. To do this, first restate a portion of the dynamics to obtain w_d from

$$w_d = P^{-1}(\mathcal{J}\tilde{w}_d + \omega \times \mathcal{J}\omega - Q_M e_3 + Q_T e_2)$$

with $P = diag([p_{11} \; p_{22} \; p_{33}]^T)$ and then obtain ζ by double-integrating from

$$\ddot{\zeta} = \tilde{\zeta}$$

by using the value that has just been obtained for $\tilde{\zeta}$. Combining the preceding results yields

$$u_d = [\zeta \; w_{1d} \; w_{2d} \; w_{3d}]^T \tag{29}$$

from the values that have just been obtained for ζ, w_{1d}, w_{2d}, and w_{3d}. Proof that the inputs generated by these equations will assure convergence is provided in [7]. The proofs to be introduced shortly will be built on the basis of the work of [7] and [11]. It is found that the control input consists of a predetermined feedforward term, u_d, and an optimal feedback term. In order to implement the optimal control in (11), the SOLA based control law is used to learn the optimal feedback tracking control, such that the OLA tuning algorithm is able to minimize the Hamiltonian while maintaining the system stability.

Lemma 1 has been introduced and gives the boundedness of $||J_{1e}||$ and therefore the system state errors, which is necessary for Theorem 2. Theorem 2 was also introduced and reveals that the SOLA convergence to the HJB function is UUB for regulation of the states. Theorem 3, to be provided next, establishes the optimality of the SOLA-based adaptive critic controller feedback term. Lemma 4 will then provide a stability condition needed for the proof for Theorem 5, which establishes the stability of the entire closed-loop system.

Theorem 3 (Optimality and convergence of the SOLA-based adaptive critic controller feedback term). Given the nonlinear system defined in (4), with target HJB equation (14), let the SOLA tuning law be given by (24) and the control input be given by (6). Then the velocity tracking error and NN parameter estimation errors of the cost function term are UUB for all $t \geq t_0 + T$, and the tracking error feedback system is controlled in a near optimal manner. That is, $\|u_e^* - \hat{u}_e\| \leq \varepsilon_u$ for a small positive constant ε_u.
Theorems 3 and 5 are proven the same way as Theorem 2, with proof to follow shortly for Theorem 5.

Lemma 4 (Stability condition). If an affine nonlinear system is asymptotically stable and the cost function given in [10] is smooth, then the closed-loop dynamics are asymptotically stable [10].

Theorem 5 (Overall system stability). Given the unmanned helicopter system with target HJB equation (14), let the tuning law for the SOLA be given by (24), and let the feedforward control input be as in (29). Then there exist constants b_{Je} and b_Γ such that the OLA approximation error $\tilde{\Gamma}$ and $\|J_{1e}(e)\|$ are UUB for all $t \geq t_0 + T$ with ultimate bounds given by $\|J_{1e}(e)\| \leq b_{Je}$ and $\|\tilde{\Gamma}\| \leq b_\Gamma$. Further, OLA reconstruction error $\|W^* - \hat{W}\| \leq \varepsilon_{r1}$ and $\|u_e^* - \hat{u}_e\| \leq \varepsilon_{r2}$ for small positive constants ε_{r1} and ε_{r2}.

Proof: First, begin with the positive definite Lyapunov function candidate

$$J = \alpha_2 J_1(e) + \tilde{\Gamma}^T\tilde{\Gamma}/2 + \tfrac{1}{2}\delta_1^T\delta_1 + \tfrac{1}{2}\delta_2^T\delta_2 + \tfrac{1}{2}\delta_3^T\delta_3 + \tfrac{1}{2}\epsilon_3^T\epsilon_3 + \tfrac{1}{2}\delta_4^T\delta_4 + \tfrac{1}{2}\epsilon_4^T\epsilon_4$$

The proof may then be divided into steps, with the first part of the Lyapunov function candidate considered first.

Step 1: consider the optimal control Lyapunov function candidate $J_{HJB} = \alpha_2 J_1(e) + \tilde{\Gamma}^T\tilde{\Gamma}/2$. Differentiating, one obtains $\dot{J}_{HJB} = \alpha_2 J_{1e}^T(e)\dot{e} + \tilde{\Gamma}^T\dot{\tilde{\Gamma}}$. Using the nonlinear system, the optimal control input, and the tuning law's error dynamics along with the derivative of the Lyapunov candidate function, then completing the square, simplifying, and using Cauchy-Schwartz yields $\dot{J}_{HJB} \leq \alpha_2 J_{1e}^T(e)(f_e(e) - \frac{1}{2}gB_e^{-1}g^T\nabla_e^T\Phi(e)\hat{\Gamma}) - \Sigma(e,\hat{u}_e)\frac{\alpha_2}{2}\tilde{\Gamma}^T\nabla_e\Phi(e)gB_e^{-1}g^TJ_{1e}^T(e) - \frac{\alpha_1}{\rho^2}||\tilde{\Gamma}||^4\beta_1 + \frac{\alpha_1}{\rho^2}\eta(\varepsilon) + \frac{\alpha_1}{\rho^2}\beta_2\delta^4(e)$ where $\beta_1 = \nabla\Phi_{min}^4 C_{min}^2/64$, $\beta_2 = 1024/C_{min}^2 + 1.5$, $\eta(\varepsilon) = 64/C_{min}^2 + 1.5(\varepsilon_M^{'4} + \varepsilon_M^{'4}C_{max}^2)$, $\varepsilon_M^{'}$ is an upper bound on the OLA reconstruction error, and $0 < \nabla\Phi_{min} \leq ||\nabla\Phi(e)||$. Now it is necessary to consider the case $\Sigma(e,\hat{u}_e) = 0$: $\dot{J}_{HJB} \leq -(\alpha_2\dot{e}_{min} - \alpha_1\beta_2K^*)||J_{1e}(e)|| - \frac{\alpha_1||\tilde{\Gamma}||^4\beta_1}{\rho^2} + \frac{\alpha_1\eta(\varepsilon)}{\rho^2}$. This is less than zero if $\alpha_2/\alpha_1 > \beta_2K^*/\dot{e}_{min}$, $||J_{1e}(e)|| > \frac{\alpha_1\eta(\varepsilon)}{(\alpha_2\dot{e}_{min} - \alpha_1\beta_2K^*)} \equiv b_{Je0}$, or $||\tilde{\Gamma}|| > \sqrt[4]{\eta(\varepsilon)/\beta_1} \equiv b_{\Gamma 0}$. Next, to consider the case $\Sigma(e,\hat{u}_e) = 1$: $\dot{J}_{HJB} \leq \alpha_2 J_{1e}^T(e)(f_e(e) - \frac{1}{2}C(\nabla_e^T\Phi(e)\Gamma + \nabla_e\varepsilon)) + \frac{\alpha_2}{2}J_{1e}^T(e)C\nabla_e\varepsilon - \frac{\alpha_1}{\rho^2}||\tilde{\Gamma}^T||^4\beta_1 + \frac{\alpha_1\eta(\varepsilon)}{\rho^2} + \frac{\alpha_1}{\rho^2}\beta_2\delta^4(e) = \alpha_2 J_{1e}^T(e)(f_e(e) + gu^*) + \frac{\alpha_2}{2}J_{1e}^T(e)C\nabla_e\varepsilon - \alpha_1\frac{\alpha_1\beta_1}{\rho^2}||\tilde{\Gamma}||^4 + \frac{\alpha_1}{\rho^2}\beta_2K^*||J_{1e}||$. Lemma 4 yields $\dot{J}_{HJB} \leq -\frac{\alpha_2\overline{Q}_{e,min}||J_{1e}(e)||^2}{2} - \frac{\alpha_1||\tilde{\Gamma}||^4\beta_1}{\rho^2} + \frac{\alpha_1\eta(\varepsilon)}{\rho^2} + \frac{\alpha_2C_{max}^2\varepsilon_M^{'2}}{(4\overline{Q}_{e,min})} + \frac{\alpha_1^2\beta_2^2K^{*2}}{(\alpha_2\rho^4\overline{Q}_{e,min})}$ with $0 < \bar{Q}_{e,min} \leq ||Q_e(e)||$. The second part of the Lyapunov function candidate will be considered next.

Step 2: consider the feedforward control Lyapunov function candidate $J_{feedforward} = S_1 + S_2 + S_3 + S_4$ with $S_1 = \frac{1}{2}\delta_1^T\delta_1$, $S_2 = \frac{1}{2}\delta_2^T\delta_2$, $S_3 = \frac{1}{2}\delta_3^T\delta_3 + \frac{1}{2}\epsilon_3^T\epsilon_3$, and $S_4 = \frac{1}{2}\delta_4^T\delta_4 + \frac{1}{2}\epsilon_4^T\epsilon_4$. It has been shown that this selection of Lyapunov candidate will guarantee stability in [7]. Differentiating, $\dot{J}_{feedforward} = \dot{S}_1 + \dot{S}_2 + \dot{S}_3 + \dot{S}_4 = -\delta_1^T\delta_1/m - \delta_2^T\delta_2 - \delta_3^T\delta_3 - \delta_4^T\delta_4 - \epsilon_3^T\epsilon_3 - \epsilon_4^T\epsilon_4$ so $\dot{J}_{feedforward} < 0$.

Step 3: consider the stability of the entire system. Combining

$$\dot{J}_{HJB} + \dot{J}_{feedforward} = -0.5\alpha_2\overline{Q}_{e,min}||J_{1e}(e)||^2$$
$$-\frac{\alpha_1||\tilde{\Gamma}||^4\beta_1}{\rho^2} + \frac{\alpha_1\eta(\varepsilon)}{\rho^2} + \frac{\alpha_2C_{max}^2\varepsilon_M^{'2}}{(4\overline{Q}_{e,min})} + \frac{\alpha_1^2\beta_2^2K^{*2}}{(\alpha_2\rho^4\overline{Q}_{e,min})}$$
$$-\frac{1}{m}\delta_1^T\delta_1 - \delta_2^T\delta_2 - \delta_3^T\delta_3 - \delta_4^T\delta_4 - \epsilon_3^T\epsilon_3 - \epsilon_4^T\epsilon_4$$

Lemma 1 and Lemma 4 will then ensure $\dot{J}_{HJB} < 0$ given that

$$||J_{1e}(e)|| > \sqrt{C_{max}^2\varepsilon_M^{\prime 2}/(2\overline{Q}_{e,min}^2)} \equiv b_Je1' \tag{30}$$

and

$$||\tilde{\Gamma}|| > \sqrt[4]{\eta(\varepsilon)/\beta_1 + \alpha_1\beta_2^2K^{*2}/(\beta_1\alpha_2\overline{Q}_{e,min})} \equiv b_{\Gamma 1} \tag{31}$$

which allows the conclusion that $||W^*(e) - \hat{W}(e)|| \leq ||\tilde{\Gamma}||||\Phi(e)|| + \varepsilon_M \leq b_\Gamma\Phi_M + \varepsilon_M \equiv \varepsilon_{r1}$ and $||u_e^*(e) - \hat{u}_e(e)|| \leq \lambda_{max}(B_e^{-1})g_Mb_\Gamma\Phi_M^{\prime}/2 + \lambda_{max}(B_e^{-1})g_M\varepsilon_M^{\prime}/2 \equiv \varepsilon_{r2}$. Then $\dot{J}_{HJB} + \dot{J}_{feedforward} < 0$ provided that (30) and (31) hold. In other words, the overall system is UUB with the bounds from (30) and (31), completing the proof.□

IV. Conclusions

A NN-based optimal control law has been proposed, which uses a single online approximator for optimal regulation and tracking control of a helicopter UAV having a dynamic model in backstepping form. The SOLA-based approach is designed to learn the infinite horizon continuous-time HJB equation, and the optimal control input that minimizes the HJB equation is calculated forward-in-time. A feedforward controller has been introduced to compensate for the helicopter's weight and requirement for rotor thrust when in hover, and to permit trajectory tracking. Further, Theorem 2 illustrates that the estimated control input approaches the target optimal control input with a small bounded error. A kinematic control structure has been used to obtain the desired velocity such that the desired position is achieved. The stability of the system has been analyzed, and the unmanned helicopter is capable of regulation and trajectory tracking.

References

[1] T.J. Koo and S. Sastry, "Output tracking control design of a helicopter model based on approximate linearization", *in Proceedings of the 37th IEEE Conference on Decision and Control*, Tampa, FL, 1998, pp. 3635-3640.

[2] B. F. Mettler, M. B. Tischlerand, and T. Kanade, System identification modelling of a small-scale rotorcraft for flight control design, *International Journal of Robotics Research*, Vol. 20, 2000, pp. 795-807.

[3] N. Hovakimyan, N. Kim, A. J. Calise, J.V.R. Prasad, "Adaptive output feedback for high-bandwidth control of an unmanned helicopter", *in Proceedings of AIAA Guidance, Navigation, and Control Conference*, Montreal, Canada, 2001, pp. 1-11.

[4] E.N. Johnson and S.K. Kannan, Adaptive trajectory control for autonomous helicopters, *Journal of Guidance, Control and Dynamics*, vol. 28, 2005, pp. 524-538.

[5] B. Ahmed, H.R. Pota and M. Garratt, Flight control of a rotary wing UAV using backstepping, *International Journal of Robust and Nonlinear Control*, vol. 20, 2010, pp. 639-658.

[6] E. Frazzoli, M.A. Dahleh and E. Feron, "Trajectory tracking control design for autonomous helicopters using a backstepping algorithm", *in Proceedings of American Control Conference*, Chicago, IL, 2000, pp. 4102-4107.

[7] R. Mahoney and T. Hamel, Robust trajectory tracking for a scale model autonomous helicopter, *International Journal of Robust and Nonlinear Control*, vol. 14, 2004, pp. 1035-1059.

[8] R. Enns and J. Si, Helicopter trimming and tracking control using direct neural dynamic programming, *IEEE Transactions on Neural Networks*, vol. 14, 2003, pp. 929-939.

[9] S. Lee, C. Ha and B.S. Kim, Adaptive nonlinear control system design for helicopter robust command augmentation, *Aerospace Science and Technology*, vol. 9, 2005, pp. 241-251.

[10] T. Dierks and S. Jagannathan, "Optimal control of affine nonlinear discrete-time systems", *in Proceedings of the Mediterranean Conference on Control and Automation*, Thessaloniki, Greece, 2009, pp. 1390-1395.

[11] T. Dierks and S. Jagannathan, "Optimal control of affine nonlinear continuous-time systems", *in Proceedings of American Control Conference*, Baltimore, MD, 2010, pp. 1568-1573.

[12] T. Dierks and S. Jagannathan, Output feedback control of a quadrotor UAV using neural networks, *IEEE Transactions on Neural Networks*, vol. 21, 2010, pp. 50-66.

[13] H.K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice-Hall, Upper Saddle River, NJ; 2002.

[14] F.L. Lewis, S. Jagannathan, and A. Yesilderek, *Neural Network Control of Robot Manipulators and Nonlinear Systems*, Taylor & Francis, London, U.K.; 1999.

[15] F. L. Lewis and V. L. Syrmos, *Optimal Control*, 2nd ed. Wiley, Hoboken, NJ; 1995.