Adaptive Critic Autopilot Design of Bank-to-Turn Missiles Using Fuzzy Basis Function Networks

Chuan-Kai Lin

***Abstract*—A new adaptive critic autopilot design for bank-to-turn missiles is presented. In this paper, the architecture of adaptive critic learning scheme contains a fuzzy-basis-function-network based associative search element (ASE), which is employed to approximate nonlinear and complex functions of bank-to-turn missiles, and an adaptive critic element (ACE) generating the reinforcement signal to tune the associative search element. In the design of the adaptive critic autopilot, the control law receives signals from a fixed gain controller, an ASE and an adaptive robust element, which can eliminate approximation errors and disturbances. Traditional adaptive critic reinforcement learning is the problem faced by an agent that must learn behavior through trial-and-error interactions with a dynamic environment, however, the proposed tuning algorithm can significantly shorten the learning time by online tuning all parameters of fuzzy basis functions and weights of ASE and ACE. Moreover, the weight updating law derived from the Lyapunov stability theory is capable of guaranteeing both tracking performance and stability. Computer simulation results confirm the effectiveness of the proposed adaptive critic autopilot.**

***Index Terms*—Adaptive critic design, adaptive robust control, bank-to-turn missiles, fuzzy basis function networks, reinforcement learning.**

NOMENCLATURE

A_x, A_y, A_z	Acceleration along the directions b_x, b_y, and b_z at center of mass, respectively.
$\{b_x, b_y, b_z\}$	Right-handed orthonormal basis of body coordinate frame, which is attached to the center of mass, C, of the missile, where b_x, b_y are on the longitudinal and lateral axis, respectively.
$C_{F_x}, C_{F_y}, C_{F_z}$	Aerodynamic force coefficients corresponding to the axes b_x, b_y, and b_z, respectively.
$C_{M_x}, C_{M_y}, C_{M_z}$	Moment coefficients corresponding to the axes b_x, b_y, and b_z, respectively.
δ_p, δ_q, δ_r	Effective deflation angle, elevator deflation angle and rudder deflation angle, respectively, (rad).
$\{e_x, e_y, e_z\}$	Right-handed orthonormal basis of inertial coordinate frame.
F_x, F_y, F_z	External forces along the axes b_x, b_y, and b_z, respectively, (N).
I_{xx}, I_{yy}, I_{zz}	Moment of inertia of the missile body about the axes b_x, b_y, and b_z, respectively, $(\mathrm{kg}\cdot\mathrm{m}^2)$.
l, m	Missile length, (m) and missile mass, (kg).
M_m	Mach number.
M_x, M_y, M_z	Total moment of inertia about the axes b_x, b_y, and b_z, respectively, $(\mathrm{N}\cdot\mathrm{m})$.
Q_s	Dynamic pressure, $(\mathrm{N}/\mathrm{m}^2)$.
P, Q, R	Roll rate, pitch rate, and yaw rate corresponding to the axes b_x, b_y, and b_z, respectively (clockwise), (rad/s).
S_{ref}	Aerodynamic reference area, (m^2).
T	Thrust, (kg).
V_s, V_m	Velocity of sound, (m/s), and velocity of missile, (m/s).
ω_a	Bandwidth of actuator, (rad/s).
$[U, V, W]^T$	Velocity vector of the missile transformed with respect to the body frame, (m/s).
$[X, Y, Z]^T$	Position vector of the center of mass of the missile transformed with respect to the inertial frame, (m).
α, β	attack angle and sideslip angle, respectively, (rad)
u_p, u_q, u_r	Actuator inputs for aileron deflation angle, elevator deflation angle and rudder deflation angle, respectively, (rad).
Φ	Body-axis roll angle measured from the downward vertical to b_z about the axis b_x, (rad).
Θ	Body-axis pitch angle measured from the projection of b_x onto the horizontal plane to b_x, (rad).
Ψ	Body-axis yaw angle measured between a fixed compass bearing and the projection of b_x onto the horizontal plane, (rad).

Manuscript received July 15, 2004; revised December 6, 2004. This paper was recommended by Associate Editor F. Hoffman.

The author is with the Department of Electrical Engineering, Chinese Naval Academy, Kaohsiung 813, Taiwan, R.O.C. (e-mail: cklin@mail.cna.edu.tw).

Digital Object Identifier 10.1109/TSMCB.2004.842246

I. INTRODUCTION

IN RECENT years, the autopilot design for bank-to-turn (BTT) missiles has been widely studied by researchers according to BTT missiles gain an advantage over skid-to-turn missiles, which have lower maneuverability and aerodynamic acceleration [1]. To obtain desired direction of aerodynamic normal force, BTT missiles should be capable of rapidly changing the orientation of acceleration by the roll motion [2]. However, the dynamics of high performance BTT missiles produces such a high roll rate combined with the asymmetric geometry will introduce strong coupling between pitch and yaw motions. Moreover, the nonminimum phase phenomenon

and highly nonlinear aerodynamics make the autopilot design of BTT missiles to be a severe challenge.

Several control methodologies have been applied to the autopilot design for BTT missiles including sliding mode control [3], gain-scheduling control [4], H^∞ control [5], [6], neural-network-based control [6]–[8], and others [9]–[11]. Most approaches are based on input/output (I/O) feedback linearization technique to obtain the model of BTT missiles and then apply control methods to design advanced autopilots. Currently, exploiting neural networks for designing autopilots of missiles has been widely studied [6]–[8], [12], [13]. The neural-network-based autopilots [5], [8] can operate in many flight conditions with the same controller, and therefore are superior to gain-scheduling autopilots. Furthermore, neural networks can be used to approximate the unknown dynamics of the BTT model, and robust control law with weight updating law derived by Lyapunov theory can achieve the satisfactory tracking performance and guaranteed stability. In summary, designing autopilots using neural networks has become an attractive and interesting approach.

Unlike the above neural-network-based autopilots, a different intelligent autopilot which uses an adaptive critic [14] is proposed, in the form of the adaptive heuristic critic (AHC) design [15]. The AHC design consists of an associative search element (ASE) and an adaptive critic element (ACE). The ACE receives a reinforcement learning signal to generate an internal reinforcement learning signal for tuning the ASE, which is employed to generate control action. Therefore, the learning of adaptive critic neural networks is performed by the signal from a critic neural network instead of gradient information. In other words, the adaptive critic neural network can learn from a dynamic environment with less information than supervised learning. However, recent adaptive critic designs [16] including heuristic dynamic programming (HDP) [17], dual heuristic programming (DHP) and globalized DHP need many epochs for learning due to trial-and-error. A new adaptive critic control [18]–[20] design is proposed which does not produce the whole control. Instead, an ASE is used to approximate the nonlinear plant dynamics, and its output is used to augment the final control action. Based on the approximation capability of adaptive critic neural networks, such a control law with online weight updating laws derived by Lyapunov theory can provide a fast learning scheme.

The architecture of adaptive critic controller in [18] and [19] is slightly different from the original adaptive critic controller [15], the former has two independent multilayer neural networks: one for ACE (critic module) and another for ASE (action module), however, the latter has only one decoder (the same input and hidden layers or the same premise part of fuzzy rules) for ACE and ASE. Besides, the adaptive critic controllers in [18] and [19] need an additional integrator to accumulate an auxiliary critic signal. Thus, the structure of the original adaptive critic controllers is simpler. Another attractive feature of this adaptive critic controller [20] is that it uses the reinforcement directly instead from the ACE, however, a more informative internal reinforcement signal evaluated by the critic module will also be omitted. Inspired from [14], we proposed an autopilot with adaptive critic structure as the Barto-Sutton-Anderson's model [15] to improve the slow trial-and-error learning process and guarantee the stability of the closed-loop system as neural-network-based autopilots [5], [8]. In order to provide the local learning capability and preserve the simpler structure, we adopt fuzzy basis functions (FBFs) as the decoder and the ACE and ASE are linear combinations of the same FBFs. The difference between the proposed adaptive critic autopilot and the adaptive critic controller in [21] is that the proposed adaptive critic autopilot can tune centers and widths of FBFs to avoid designing a sparse and large fuzzy rule base or heuristically determining centers and widths of FBFs. Therefore, the purpose of this paper is to develop an adaptive critic autopilot to adapt all parameters of ACE, ASE and decoder (FBFs) online so as to ensure the tracking performance and stability for the BTT missile.

The rest of this paper is organized as follows. Section II describes the adaptive critic controller structure and some mathematic fundamentals. Section III states the model of the BTT missiles and the control objective. Section IV presents the adaptive critic autopilot and stability analysis. In Section V, simulation results illustrate the availability and effectiveness of the proposed adaptive critic autopilot for BTT missiles. Finally, Section VI concludes the paper.

II. Preliminaries

A. Fuzzy Basis Function Network

Generally speaking, a fuzzy system comprises a fuzzification unit, a fuzzy rulebase, an inference engine and a defuzzification unit. The fuzzy system can be viewed as performing a real (nonfuzzy) and nonlinear mapping from $\boldsymbol{U} \in \Re^n$ to $\Re^m$. The interfaces between real world and fuzzy world are a fuzzifier and a defuzzifier; the former maps real inputs to their corresponding fuzzy sets and the latter performs in the opposite way to map from fuzzy sets of output variables to the corresponding real outputs. The fuzzy rule base consists of fuzzy rules, which use linguistic IF-THEN sentences to describe the relationship between inputs and outputs. Consider a fuzzy rule consisting of N rules as follows.

> R_j (jth rule): If x_1 is A_{j1} and x_2 is A_{j2} and … and x_n is A_{jn} then f_1 is B_{j1} and f_2 is B_{j2} and … and f_m is B_{jm}

where $j = 1, 2, \ldots, N$, $x_i (i = 1, 2, \ldots, n)$ are the input variables of the fuzzy system, f_k $(k = 1, 2, \ldots, m)$ are the output variables of the fuzzy system, and input linguistic terms A_{ji} and output linguistic terms B_{kj} are characterized by their corresponding fuzzy membership functions $\mu_{A_{ji}}(x_i)$ and $\mu_{B_{kj}}(f_k)$, respectively. Each rule R_j can be viewed as a fuzzy implication determined by the inference engine.

The above fuzzy system with center-average defuzzifier, product inference, and singleton fuzzifier is of the following form:

$$f_k = \frac{\sum_{j=1}^{N} w_{kj} \left(\prod_{i=1}^{n} \mu_{A_{ji}}(x_i) \right)}{\sum_{j=1}^{N} \left(\prod_{i=1}^{n} \mu_{A_{ji}}(x_i) \right)}, \quad k = 1, \ldots, m \qquad (1)$$

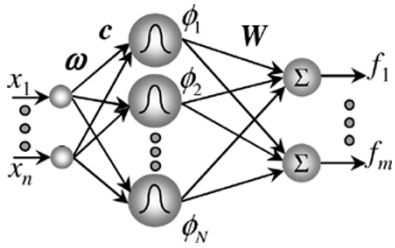

Fig. 1. FBFN structure.

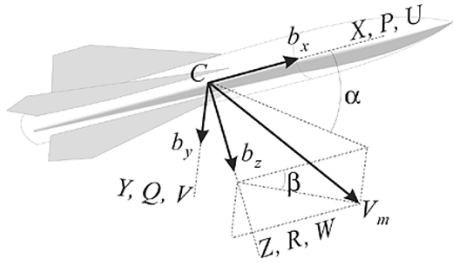

Fig. 2. BTT missile diagram.

where $\mu_{B_{kj}}(w_{kj}) = 1$. Equation (1) can be written as sum of firing strength of rules

$$f_k = \sum_{j=1}^{N} w_{kj}\phi_j, \quad k = 1, \ldots, m \tag{2}$$

where the firing strength $\phi_j = \prod_{i=1}^{n} \mu_{A_{ji}(x_i)} / \sum_{j=1}^{N} \left(\prod_{i=1}^{n} \mu_{A_{ji}(x_i)}\right)$. For ease of representing in the network architecture, the firing strength ϕ_j is rewritten as the product of $\mu_{A_{ji}}(x_i)$

$$\phi_j = \prod_{i=1}^{n} \mu_{A_{ji}}(x_i), \quad j = 1, \ldots, N. \tag{3}$$

The shape of each membership function $\mu_{A_{ji}}(x_i)$ is chosen as the Gaussian function, i.e., $\mu_{A_{ji}}(x_i) = e^{-\omega_{ji}^2 (x_i - c_{ji})^2}$ with inverse variance ω_{ji} and center c_{ji}. From the viewpoint of geometry [22], the universe of discourse can be covered by well-defined fuzzy rules and each fuzzy rule can be considered as a cluster with center c_{ji} $(i = 1, 2, \ldots, n)$. Therefore, the architecture of the fuzzy basis function network, which is shown in Fig. 1, can be represented by a three-layer network with Gaussian functions as its activation functions in the hidden layer and weights w_{kj}'s connecting hidden layer and output layer. Thus, the output vector of fuzzy basis function network (FBFN) can be expressed as

$$\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{\omega}, \boldsymbol{W}) = \boldsymbol{W}^T \boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{\omega}) \tag{4}$$

where $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T \in Rn$, $\boldsymbol{c} = [\boldsymbol{c}_1^T \boldsymbol{c}_2^T \cdots \boldsymbol{c}_N^T]^T \in \Re^{nN}$, $\boldsymbol{\omega} = [\boldsymbol{\omega}_1^T \boldsymbol{\omega}_2^T \cdots \boldsymbol{\omega}_N^T]^T \in \Re^{nN}$, $\boldsymbol{c}_j = [c_{j1} c_{j2} \cdots c_{jn}]^T \in \Re^n$, $\boldsymbol{\omega}_j = [\omega_{j1} \omega_{j2} \cdots \omega_{jn}]^T \in \Re^n$, $\boldsymbol{W}^T = [w_{kj}]$ is an $m \times N$ matrix and $\boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{\omega}) = [\,\phi_1 \quad \phi_2 \quad \ldots \quad \phi_N\,]^T$

$$= \left[\, e^{-[\omega_1^T (x - c_1)]^2} \quad e^{-[\omega_2^T (x - c_2)]^2} \quad \ldots \quad e^{-[\omega_N^T (x - c_N)]^2} \,\right]^T.$$

By the Stone–Weierstrass theorem, the FBFN can be proved that it is capable of uniformly approximating any real continuous function $\boldsymbol{f}(\boldsymbol{x})$ on a compact set $\boldsymbol{U}$ to any arbitrary accuracy b_ϵ [23], i.e., there exists an ideal FBFN, $\boldsymbol{W}^T \boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{\omega})$, with ideal parameters $\boldsymbol{c}$, $\boldsymbol{\omega}$, and $\boldsymbol{W}$ such that $\sup_{x \in U} \| \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{W}^T \boldsymbol{\phi}(x, \boldsymbol{c}, \boldsymbol{\omega}) \| < b_\varepsilon$. Therefore, $\boldsymbol{f}(\boldsymbol{x})$ can be represented as

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{W}^T \boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{\omega}) + \boldsymbol{\varepsilon}_f \tag{5}$$

where $\|\boldsymbol{\varepsilon}_f\| \leq b_\varepsilon$. In this paper, all norms of vectors and matrices adopt Frobenius norm. The norms of the ideal parameters should satisfy the following assumption.

Assumption 1: The norms of ideal parameters, $\|\boldsymbol{c}\|$, $\|\boldsymbol{\omega}\|$, and $\|\boldsymbol{W}\|$, are bounded by positive real values, i.e., $\|\boldsymbol{c}\| \leq b_c$, $\|\boldsymbol{\omega}\| \leq b_\omega$, and $\|\boldsymbol{W}\| \leq b_W$.

Clearly, we need to estimate the ideal FBFN by an estimate FBFN $\hat{\boldsymbol{f}} = \hat{\boldsymbol{W}}^T \boldsymbol{\phi}(\boldsymbol{x}), \hat{\boldsymbol{c}}, \hat{\boldsymbol{\omega}})$. The weight updating law will be stated in the following section.

B. Mathematical Notation

The norm of a vector and the Frobenius norm used in the following sections are defined previously. Let $\Re$ denote the real numbers, $\Re^n$ denote real n-vectors, and $\Re^{m \times n}$ be real $m \times n$ matrices. The norm of a vector $\boldsymbol{x} \in \Re^n$ is defined as

$$\|\boldsymbol{x}\| = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2} \tag{6}$$

and the Frobenius norm of $\boldsymbol{C} = [c_{ij}] \in \Re^{m \times n}$ is defined by

$$\|\boldsymbol{C}\|_F = [tr(\boldsymbol{C}^T \boldsymbol{C})]^{1/2} = \sqrt{\sum_{i,j} \boldsymbol{c}_{ij}^2}. \tag{7}$$

It should be noted that the Frobenius norm is also compatible with the 2-norm, i.e., $\|\boldsymbol{C}\boldsymbol{x}\| \leq \|\boldsymbol{C}\|_F \|\boldsymbol{x}\|$.

The trace of a square matrix $\boldsymbol{T} \in \Re^{n \times n}$, $tr(\boldsymbol{T})$, is defined as the sum of diagonal elements of $\boldsymbol{T}$. Clearly, the trace of a square matrix and the trace of the transpose of the square matrix are the same, i.e., $tr(\boldsymbol{T}) = tr(\boldsymbol{T}^T)$. Given $\boldsymbol{b} \in \Re^n$, $\boldsymbol{c} \in \Re^n$, $\boldsymbol{B} \in \Re^{n \times m}$ and $\boldsymbol{C} \in \Re^{m \times n}$, we have the following properties:

$$tr(\boldsymbol{B}\boldsymbol{C}) = tr(\boldsymbol{C}\boldsymbol{B}) \tag{8}$$

$$tr(\boldsymbol{b}\boldsymbol{c}^T) = tr(\boldsymbol{c}\boldsymbol{b}^T). \tag{9}$$

III. Problem Statement

All variables of the BTT missile used in the following equations are listed in Nomenclature. Assuming the BTT missile shown in Fig. 2 is a rigid body, the complete six-degree-of-freedom dynamics of BTT missiles can be given by

$$\dot{P} = \frac{I_{yy} - I_{zz}}{I_{xx}} QR + \frac{M_x}{I_{xx}} \tag{10}$$

$$\dot{Q} = \frac{I_{zz} - I_{xx}}{I_{yy}} RP + \frac{M_y}{I_{yy}} \tag{11}$$

$$\dot{R} = \frac{I_{xx} - I_{yy}}{I_{zz}} PQ + \frac{M_z}{I_{zz}} \tag{12}$$

$$\dot{\Phi} = P + (Qs\Phi + Rc\Phi)t\Theta \tag{13}$$

$$\dot{\Theta} = Qc\Phi - Rs\Phi \tag{14}$$

$$\dot{\Psi} = \frac{(Qs\Phi - Rc\Phi)}{c\Theta} \tag{15}$$

$$\dot{U} = RV - QW + \frac{1}{m} F_x \tag{16}$$

$$\dot{V} = -RU + PW + \frac{1}{m} F_y \tag{17}$$

$$\dot{W} = QU - PV + \frac{1}{m} F_z \tag{18}$$

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} = \begin{bmatrix} c\Psi c\Theta & c\Psi s\Theta s\Phi - s\Psi c\Phi & c\Psi s\Theta c\Phi + s\Psi s\Phi \\ s\Psi c\Theta & s\Psi s\Theta s\Phi + c\Psi c\Phi & s\Psi s\Theta c\Phi - c\Psi s\Phi \\ -s\Theta & c\Theta s\Phi & c\Theta c\Phi \end{bmatrix} \times \begin{bmatrix} U \\ V \\ W \end{bmatrix} \tag{19}$$

where $c\Phi = \cos\Phi$, $s\Phi = \sin\Phi$, $c\Theta = \cos\Theta$, $s\Theta = \sin\Theta$, $t\Theta = \tan\Theta$, $c\Psi = \cos\Psi$, $s\Psi = \sin\Psi$, $g = 9.82\ \mathrm{m/s^2}$, $F_x = C_{F_x} Q_s S_{\mathrm{ref}} + T - mg\sin\Theta$, $F_y = C_{F_y} Q_s S_{\mathrm{ref}} + T + mg\cos\Theta\sin\Phi$, $F_z = C_{F_z} Q_s S_{\mathrm{ref}} + T + mg\cos\Theta\cos\Phi$, $M_y = C_{M_y} Q_s S_{\mathrm{ref}} l$, $M_x = C_{M_x} Q_s S_{\mathrm{ref}} l$, and $M_z = C_{M_z} Q_s S_{\mathrm{ref}} l$. The actuator is modeled by a first-order system as follows: $\dot{\delta}_p = -\omega_a \delta_p + \omega_a u_p$, $\dot{\delta}_q = -\omega_a \delta_q + \omega_a u_q$ and $\dot{\delta}_r = -\omega_a \delta_r + \omega_a u_r$. The above 6-DOF missile dynamic equations and detailed process of derivation can be referred to [2].

In general, the objective of the autopilot design is to drive the BTT missile to track the commands generated from guidance system including desired rolling angle Φ_c, and desired accelerations A_{yc} and A_{zc}. Such an output, i.e., (Φ, A_y, A_z) for the BTT missile system will cause a nonminimum phase phenomenon, thus I/O feedback linearization techniques can not be applied directly for acceleration control of BTT missiles [2]. Much research surmounts the difficulty by adopting redefined outputs as (Φ, α, β) [9], [10], (Φ, V, W) [2], $(P\cos\alpha + R\sin\alpha, \alpha, \beta)$ [3] or (μ, α, β) (μ is an aerodynamic bank angle) [12] so as to apply I/O feedback linearization technique to BTT missiles. According to α and β derived from V and W ($\alpha = \tan^{-1}(W/U)$ and $\beta = \tan^{-1}(V/U)$), our new output signals are assigned as (Φ, V, W).

The nominal plant of BTT missiles can be rewritten by the input-output feedback linearization technique in the following form:

$$\begin{bmatrix} y_1^{(n_1)} \\ y_2^{(n_2)} \\ y_3^{(n_3)} \end{bmatrix} = \begin{bmatrix} f_1(\boldsymbol{x}) \\ f_2(\boldsymbol{x}) \\ f_3(\boldsymbol{x}) \end{bmatrix} + \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \text{ or}$$
$$\boldsymbol{y}^{(n)} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{G}(\boldsymbol{x})\boldsymbol{u} \tag{20}$$

where $\boldsymbol{y} = [y_1\ y_2\ y_3]^T = [\Phi\ V\ W]^T$, $\boldsymbol{u} = [u_p\ u_q\ u_r]^T$ and $\boldsymbol{x} = [P\ Q\ R\ \Phi\ \Theta\ \Psi\ U\ V\ W\ X\ Y\ Z]^T$. The assumed strong relative-degree (n_1, n_2, n_3) for the three input-output channels is found to be (3, 2, 2) [2], [8]. The zero dynamics are assumed to be exponentially stable. In the above equation, $\boldsymbol{f}(\boldsymbol{x})$ is an unknown function vector, $\boldsymbol{u}$ is the input vector, and the nine elements of the known $\boldsymbol{G}(\boldsymbol{x})$ are as follows:

$$\begin{aligned} g_{11} &= Q_s S_{\mathrm{ref}} l \left(\frac{C_1}{I_{xx}} + \frac{C_2}{I_{yy}} t\Theta s\Phi + \frac{C_3}{I_{zz}} t\Theta c\Phi \right) \\ g_{12} &= Q_s S_{\mathrm{ref}} l \frac{C_4}{I_{yy}} t\Theta s\Phi \\ g_{13} &= Q_s S_{\mathrm{ref}} l \left(\frac{C_5}{I_{xx}} + \frac{C_6}{I_{zz}} t\Theta c\Phi \right) \\ g_{21} &= Q_s S_{\mathrm{ref}} l \frac{C_7}{m} \\ g_{22} &= 0 \\ g_{23} &= Q_s S_{\mathrm{ref}} l \frac{C_8}{m} \\ g_{31} &= Q_s S_{\mathrm{ref}} l \frac{C_9}{m} \\ g_{32} &= Q_s S_{\mathrm{ref}} l \frac{C_{10}}{m} \\ g_{33} &= 0 \end{aligned} \tag{21}$$

where C_i's are aerodynamic coefficients, which are complex functions of M_m and $\boldsymbol{x}$.

The commands now are rewritten in the form of

$$(\Phi_c, V_c, W_c) \tag{22}$$

where Φ_c remains the same, while $V_c = 0$ and $W_c = K_w(A_z - A_{zc})$ with a constant K_w. In practice, the desired trajectories Φ_d (for Φ_d) and W_d (for W), which should be smooth, are generated by two first-order filters with inputs Φ_c and W_c. The desired trajectory for V, V_d, still remains zero for keeping small sideslip angle β due to $\beta = \tan^{-1}(V/U)$.

Then, the tracking errors are defined as follows:

$$\begin{aligned} e_1 &= y_1 - y_{1d} = \Phi - \Phi_d \\ e_2 &= y_2 - y_{2d} = V - V_d \\ e_3 &= y_3 - y_{3d} = W - W_d \end{aligned} \tag{23}$$

where y_{1d}, y_{2d}, and y_{3d} are the desired trajectories of y_1, y_2, and y_3, respectively. Hence, a sliding-surface vector $\boldsymbol{s}(t)$ is defined as

$$\begin{aligned} \boldsymbol{s}(t) &= [\, s_1(t) \quad s_2(t) \quad s_3(t)\,]^T \\ &= [\, \ddot{e}_1 + 2k_1\dot{e}_1 + k_1^2 e_1 \quad \dot{e}_2 + k_2 e_2 \quad \dot{e}_3 + k_3 e_3\,]^T \end{aligned} \tag{24}$$

where each k_j is a strictly positive constant for $1 \le j \le 3$. Differentiating $\boldsymbol{s}(t)$ with respect to time, we can get

$$\dot{\boldsymbol{s}}(t) = \begin{bmatrix} y_1^{(3)} \\ y_2^{(2)} \\ y_3^{(2)} \end{bmatrix} + \begin{bmatrix} -y_{1d}^{(3)} + 2k_1\ddot{e}_1 + k_1^2\dot{e}_1 \\ -y_{2d}^{(2)} + k_2\dot{e}_2 \\ -y_{3d}^{(2)} + k_3\dot{e}_3 \end{bmatrix} \text{ or}$$
$$\dot{\boldsymbol{s}}(t) = \boldsymbol{y}^{(n)} + \boldsymbol{q}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{G}(\boldsymbol{x})\boldsymbol{u} + \boldsymbol{q}(\boldsymbol{x}) \tag{25}$$

where

$$\boldsymbol{q}(\boldsymbol{x}) = [-y_{1d}^{(3)} + 2k_1\ddot{e}_1 + k_1^2\dot{e}_1 - y_{2d}^{(2)} + k_2\dot{e}_2 - y_{3d}^{(2)} + k_3\dot{e}_3]^T.$$

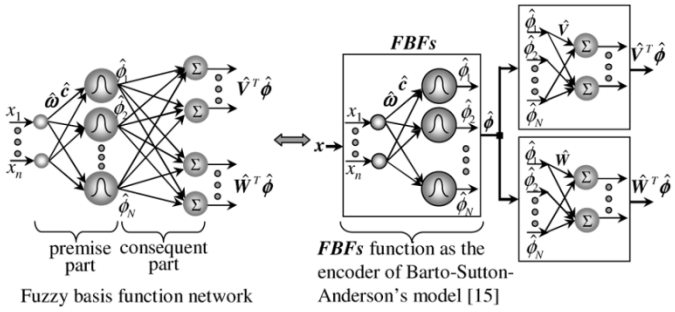

Fig. 3. Barto-Sutton-Anderson's model [15] is represented by an FBFN.

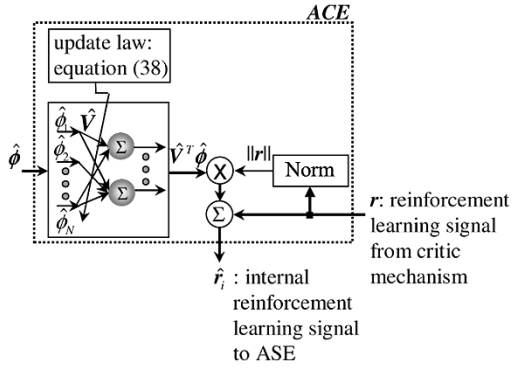

Fig. 4. ACE receives the reinforcement learning signal from critic mechanism to generate an internal reinforcement signal.

Assume the inverse of known matrix $\boldsymbol{G}(\boldsymbol{x})$, $\boldsymbol{G}^{-1}(\boldsymbol{x})$, exists and the unknown $\boldsymbol{f}(\boldsymbol{x})$ is approximated by $\hat{\boldsymbol{f}}(\boldsymbol{x})$, then the control law can be selected as

$$\boldsymbol{u} = \boldsymbol{G}^{-1}(\boldsymbol{x})[-\hat{\boldsymbol{f}}(\boldsymbol{x}) - \boldsymbol{q}(\boldsymbol{x}) - \boldsymbol{K}_s\boldsymbol{s}(t) - \hat{\boldsymbol{d}}] \tag{26}$$

where $\boldsymbol{K}_s = \mathrm{diag}\{K_{s1}, K_{s2}, K_{s3}\} > 0$ and $\hat{\boldsymbol{d}}$ is used to attenuate uncertainties including approximation errors and external disturbances. Substituting (26) into (25) yields

$$\dot{\boldsymbol{s}} = -\boldsymbol{K}_s\boldsymbol{s} + \boldsymbol{f}(\boldsymbol{x}) - \hat{\boldsymbol{f}}(\boldsymbol{x}) - \hat{\boldsymbol{d}}. \tag{27}$$

The estimation $\hat{\boldsymbol{f}}(\boldsymbol{x})$, which can approximate $\boldsymbol{f}(\boldsymbol{x})$ accurately, is implemented by the ASE. Due to approximation errors, our control goal is to design an adaptive critic autopilot for the BTT missile system which reduces approximation errors and tracks the desired trajectory while maintaining a small sideslip angle.

IV. Autopilot Design

A. Adaptive Critic Controller Structure

Based on the architecture of reinforcement learning control systems [15], we employ the premise part of FBFN as the decoder of states as shown in Fig. 3. The consequent part of FBFN is separated into two parts: one is corresponding to the ASE shown in Fig. 5 and another is corresponding to the ACE shown in Fig. 4. In the spirit of actor-critic reinforcement learning control, the weights of fuzzy approximator (ASE) will be tuned by the signal from ACE. During the online control process, the ASE establishes the associations between firing strengths of rules and nonlinear functions under the influence of reinforcement feedback. Moreover, the ACE in Fig. 4 produces a more informative signal than a single reinforcement signal as in [20].

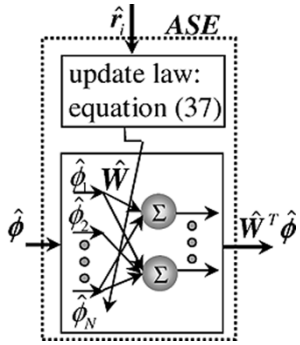

Fig. 5. ASE approximates nonlinear functions.

The firing strength of fuzzy rules can be denoted as the following vector form:

$$\hat{\boldsymbol{\phi}} = \boldsymbol{\phi}(\boldsymbol{x}, \hat{\boldsymbol{c}}, \hat{\boldsymbol{\omega}}) = [\,\hat{\phi}_1 \quad \hat{\phi}_2 \quad \ldots \quad \hat{\phi}_N\,]^T \tag{28}$$

where $\hat{\boldsymbol{c}}$ is the estimate of $\boldsymbol{c}$ and $\hat{\boldsymbol{\omega}}$ is the estimate of $\boldsymbol{\omega}$. Thus, the outputs of ASE and ACE can be represented as $\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}}$ and $\hat{\boldsymbol{V}}^T\hat{\boldsymbol{\phi}}$, respectively. Different from the original adaptive critic controller [15] and the reinforcement adaptive controllers [20], [21], the parameters of decoder (FBFs), $\hat{\boldsymbol{c}}$ and $\hat{\boldsymbol{\omega}}$, can be tuned. Besides self-tuning $\hat{\boldsymbol{c}}$ and $\hat{\boldsymbol{\omega}}$, there are two ways to determine $\hat{\boldsymbol{c}}$ and $\hat{\boldsymbol{\omega}}$: one is heuristically determined by experts and one is contains all rules but sparse and large. Therefore, it is worth making more computation effort to tune $\hat{\boldsymbol{c}}$ and $\hat{\boldsymbol{\omega}}$.

The process of generating the bounded reinforcement signal $\boldsymbol{r}$ is through the performance measurement unit measuring the system states $\boldsymbol{x}$ to provide an error metric signal vector $\boldsymbol{s}$ for the critic unit to generate the reinforcement signal $\boldsymbol{r} = [r_1 r_2 \ldots r_m]^T$. Thus, a special critic function is defined as [20]

$$r_k(t) = \frac{\chi_k}{1 + e^{-m_k s_k(t)/\chi_k}} - \frac{\chi_k}{1 + e^{m_k s_k(t)/\chi_k}} \qquad k = 1, \ldots, m \tag{29}$$

where m_k is a positive constant and the value of r_k is bounded in the interval $[-\chi_k, \chi_k]$ with $\chi_k > 0$. In [18] and [19], although the reinforcement signal, which directly employs $\boldsymbol{s}$, can be proved to be bounded, such a reinforcement signal with unknown bound is different from that in [15]. The critic function outputs a zero reward when the error metric is zero, and a larger reward magnitude for larger error metric gains. The difference between the reinforcement signal in [15] and this paper is that the reinforcement signal in the former is only "0" or "1", however, r_k in the latter is a real value and attempts to decay to zero by Lyapunov theory. After receiving the reinforcement signal $\boldsymbol{r}$, the adaptive critic learning scheme will manipulate an internal reinforcement signal $\hat{\boldsymbol{r}}_i$ for ACE to tune the ASE. The internal reinforcement signal is of the following form:

$$\hat{\boldsymbol{r}}_i = \boldsymbol{r} + \|\boldsymbol{r}\|\hat{\boldsymbol{V}}^T\hat{\boldsymbol{\phi}}, \tag{30}$$

which is an estimate of ideal internal reinforcement signal $\boldsymbol{r}_i$. It should be noted that if we can prove that $\boldsymbol{r}$ (or $\boldsymbol{s}$) will approach zero, then the reinforcement signal $\hat{\boldsymbol{r}}_i$ will also approach zero.

Specifically, the control input $\boldsymbol{u}$ in Fig. 6 mainly consists of an ASE, a fixed gain controller $\boldsymbol{K}\boldsymbol{s}$ and a robustifying term $\hat{\boldsymbol{d}}$,

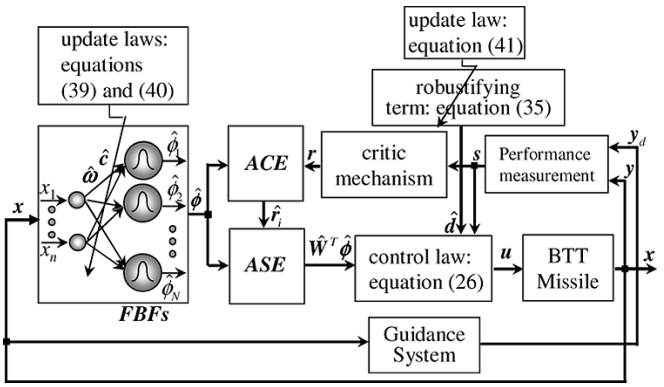

Fig. 6. Overall adaptive critic autopilot structure.

which corresponds to reject approximation errors of ASE and disturbances. From the above description, we know the basic structure of the proposed adaptive critic autopilot is the same as the Barto-Sutton-Anderson's model [15], however, there are still some differences including that the former employs fuzzy basis function networks to implement the decoder, ACE and ASE, and the evaluation of $\boldsymbol{r}(t)$ and $\hat{\boldsymbol{r}}_i$. Besides, the major differences are the control law, weight updating laws of the proposed adaptive critic autopilot can guarantee the stability of the closed-loop system and guarantee the tracking performance without trial-and-error learning process. In the following, the details of the control law, weight updating law, and stability analysis will be discussed.

B. Stability Analysis

It is assumed that there exist ideal weights $\boldsymbol{V}$ and $\boldsymbol{W}$ such that $\boldsymbol{r}_i = \boldsymbol{r} + \|\boldsymbol{r}\|\boldsymbol{V}^T\boldsymbol{\phi}$ is the ideal internal reinforcement signal and $\boldsymbol{W}^T\boldsymbol{\phi}$ is ideal approximator of $\boldsymbol{f}(\boldsymbol{x})$. According to the approximation capability of FBFN, the ideal $\boldsymbol{r}_i = \boldsymbol{r} + \|\boldsymbol{r}\|\boldsymbol{V}^T\boldsymbol{\phi}$ and $\boldsymbol{W}^T\boldsymbol{\phi}$ can be approximated by $\hat{\boldsymbol{r}}_i = \boldsymbol{r} + \|\boldsymbol{r}\|\hat{\boldsymbol{V}}^T\hat{\boldsymbol{\phi}}$ and $\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}}$, respectively. Thus, the objective of the autopilot design is to achieve good tracking performance and guaranteed stability with $\hat{\boldsymbol{c}}$, $\hat{\boldsymbol{\omega}}$, $\hat{\boldsymbol{V}}$ and $\hat{\boldsymbol{W}}$ tuned online. In order to accomplish this goal by Lyapunov theory extension, the FBFs are linearized. Thus, take the expansion of $\tilde{\boldsymbol{\phi}} = \boldsymbol{\phi} - \hat{\boldsymbol{\phi}}$ in a Taylor series to obtain

$$\tilde{\boldsymbol{\phi}} = \begin{bmatrix} \tilde{\phi}_1 \\ \vdots \\ \tilde{\phi}_N \end{bmatrix} = \left.\begin{bmatrix} \frac{\partial \phi_1}{\partial \boldsymbol{\omega}} \\ \vdots \\ \frac{\partial \phi_N}{\partial \boldsymbol{\omega}} \end{bmatrix}\right|_{\omega=\hat{\omega}} \tilde{\boldsymbol{\omega}} + \left.\begin{bmatrix} \frac{\partial \phi_1}{\partial \boldsymbol{c}} \\ \vdots \\ \frac{\partial \phi_N}{\partial \boldsymbol{c}} \end{bmatrix}\right|_{c=\hat{c}} \tilde{\boldsymbol{c}} + \boldsymbol{h} \text{ or}$$
$$\tilde{\boldsymbol{\phi}} = \boldsymbol{A}^T\tilde{\boldsymbol{\omega}} + \boldsymbol{B}^T\tilde{\boldsymbol{c}} + \boldsymbol{h} \tag{31}$$

where $\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega} - \hat{\boldsymbol{\omega}}$, $\tilde{\boldsymbol{c}} = \boldsymbol{c} - \hat{\boldsymbol{c}}$, $\boldsymbol{A} = [\partial\phi_1/\partial\boldsymbol{\omega} \ldots \partial\phi_N/\partial\boldsymbol{\omega}]|_{\omega=\hat{\omega}}$, $\boldsymbol{B} = [(\partial\phi_1/\partial c) \ldots (\partial\phi_N/\partial\boldsymbol{c})]|_{c=\hat{c}}$, $\boldsymbol{h}$ is a vector of higher order terms and $\partial\boldsymbol{\phi}_j/\partial\boldsymbol{\omega}$ and $\partial\boldsymbol{\phi}_j/\partial\boldsymbol{c}$ are defined as

$$\left[\frac{\partial\phi_j}{\partial\boldsymbol{\omega}}\right]^T = \begin{bmatrix} \underbrace{0\ldots0}_{(j-1)\times n} & \frac{\partial\phi_j}{\partial\omega_{j1}} & \cdots & \frac{\partial\phi_j}{\partial\omega_{jn}} & \underbrace{0\ldots0}_{(N-j)\times n} \end{bmatrix}$$
$$\left[\frac{\partial\phi_j}{\partial\boldsymbol{c}}\right]^T = \begin{bmatrix} \underbrace{0\ldots0}_{(j-1)\times n} & \frac{\partial\phi_j}{\partial c_{j1}} & \cdots & \frac{\partial\phi_j}{\partial c_{jn}} & \underbrace{0\ldots0}_{(N-j)\times n} \end{bmatrix}.$$

Assumption 2: The vector of higher order terms in (31), $\boldsymbol{h}$, is bounded, and A and B are also bounded, i.e., $\|\boldsymbol{h}\| \le b_h$, $\|\boldsymbol{A}\| \le b_A$, and $\|\boldsymbol{B}\| \le b_B$.

Using (31), the output of ACE can be represented as

$$\hat{\boldsymbol{V}}^T\hat{\boldsymbol{\phi}} = -\tilde{\boldsymbol{V}}^T(\hat{\boldsymbol{\phi}} - \boldsymbol{A}^T\hat{\boldsymbol{\omega}} - \boldsymbol{B}^T\hat{\boldsymbol{c}}) - \hat{\boldsymbol{V}}^T(\boldsymbol{A}^T\tilde{\boldsymbol{\omega}} + \boldsymbol{B}^T\tilde{\boldsymbol{c}}) - \boldsymbol{d}_1 \tag{32}$$

where $\boldsymbol{d}_1 = \boldsymbol{V}^T(\boldsymbol{A}^T\boldsymbol{\omega} + \boldsymbol{B}^T\boldsymbol{c} + \boldsymbol{h} - \boldsymbol{\phi}) - \hat{\boldsymbol{V}}^T(\boldsymbol{A}^T\boldsymbol{\omega} + \boldsymbol{B}^T\boldsymbol{c})$. On the other hand, the error between the output of ideal ASE and ASE can be derived as

$$\begin{aligned} \boldsymbol{f} - \hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}} &= \boldsymbol{\varepsilon}_f + \boldsymbol{W}^T\boldsymbol{\phi} - \hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}} \\ &= \boldsymbol{\varepsilon}_f + \boldsymbol{W}^T\tilde{\boldsymbol{\phi}} + \tilde{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}}. \end{aligned} \tag{33}$$

Therefore, error dynamics (27) can be rewritten as

$$\dot{\boldsymbol{s}} = -\boldsymbol{K}_s\boldsymbol{s} + \boldsymbol{\varepsilon}_f + \boldsymbol{W}^T\tilde{\boldsymbol{\phi}} + \tilde{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}} - \hat{\boldsymbol{d}}. \tag{34}$$

The robustifying term is defined as

$$\hat{\boldsymbol{d}} = \frac{\boldsymbol{r}}{\|\boldsymbol{r}\|}\hat{\boldsymbol{k}}_d^T\boldsymbol{\varphi} \tag{35}$$

with $\boldsymbol{\varphi} = [1 \quad \|\hat{\boldsymbol{V}}\| \quad \|\hat{\boldsymbol{W}}\| \quad \|\hat{\boldsymbol{V}}\|\|\hat{\boldsymbol{W}}\|]^T$.

The reinforcement signal plays an important role in adaptive critic learning system and we express it in the following vector form:

$$\boldsymbol{r} = \boldsymbol{\sigma}(\boldsymbol{s}) = [\sigma(s_1) \quad \sigma(s_2) \quad \sigma(s_3)]^T \tag{36}$$

where $\sigma(s_k) = (\chi_k/(1+e^{-m_k s_k/\chi_k})) - (\chi_k/(1+e^{m_k s_k/\chi_k}))$, $k = 1, 2, 3$. Next, we establish a theorem, which provides the stability analysis.

Theorem: Apply the control law (26) to the BTT missile system (20) with the robustifying term (35). The weight update rules for the FBFs, ASE, and ACE are defined as

$$\begin{aligned} \dot{\hat{\boldsymbol{W}}} &= \boldsymbol{K}_W\hat{\boldsymbol{\phi}}\hat{\boldsymbol{r}}_i^T - \eta\|\boldsymbol{r}\|\boldsymbol{K}_W\hat{\boldsymbol{W}} \\ &= \boldsymbol{K}_W\hat{\boldsymbol{\phi}}(\boldsymbol{r} + \|\boldsymbol{r}\|\hat{\boldsymbol{V}}^T\hat{\boldsymbol{\phi}})^T - \eta\|\boldsymbol{r}\|\boldsymbol{K}_W\hat{\boldsymbol{W}} \end{aligned} \tag{37}$$
$$\dot{\hat{\boldsymbol{V}}} = -\|\boldsymbol{r}\|\boldsymbol{K}_V(\hat{\boldsymbol{\phi}} - \boldsymbol{A}^T\hat{\boldsymbol{\omega}} - \boldsymbol{B}^T\hat{\boldsymbol{c}})(\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}})^T - \eta\|\boldsymbol{r}\|\boldsymbol{K}_V\hat{\boldsymbol{V}} \tag{38}$$
$$\dot{\hat{\boldsymbol{\omega}}} = -\|\boldsymbol{r}\|\boldsymbol{K}_\omega\boldsymbol{A}\hat{\boldsymbol{V}}\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}} - \eta\|\boldsymbol{r}\|\boldsymbol{K}_\omega\hat{\boldsymbol{\omega}} \tag{39}$$
$$\dot{\hat{\boldsymbol{c}}} = -\|\boldsymbol{r}\|\boldsymbol{K}_c\boldsymbol{B}\hat{\boldsymbol{V}}\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}} - \eta|\boldsymbol{r}\|\boldsymbol{K}_c\hat{\boldsymbol{c}} \tag{40}$$
$$\dot{\hat{\boldsymbol{k}}}_d = \|\boldsymbol{r}\|\boldsymbol{K}_d\boldsymbol{\varphi} \tag{41}$$

where $\eta > 0$, and $\boldsymbol{K}_W$, $\boldsymbol{K}_V$, $\boldsymbol{K}_\omega$, $\boldsymbol{K}_c$, and $\boldsymbol{K}_d$ are positive and diagonal constant matrices. Suppose that the desired trajectories are bounded and differentiable, and assumptions 1 and 2 are hold. And the reinforcement signal and internal reinforcement signal are provided by (36) and (30), respectively. Then, $\hat{\boldsymbol{W}}$, $\hat{\boldsymbol{V}}$, $\hat{\boldsymbol{\omega}}$, and $\hat{\boldsymbol{c}}$ remain bounded, and the error metric $\boldsymbol{s}(t)$ and the tracking errors will approach zero.

Proof: Define a Lyapunov function candidate

$$\begin{aligned} L(\boldsymbol{s}, \tilde{\boldsymbol{W}}, \tilde{\boldsymbol{V}}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{\omega}}, \tilde{\boldsymbol{k}}_d) = \rho(\boldsymbol{s}) &+ \frac{1}{2}tr\left\{\tilde{\boldsymbol{W}}^T\boldsymbol{K}_W^{-1}\tilde{\boldsymbol{W}}\right\} \\ &+ \frac{1}{2}tr\left\{\tilde{\boldsymbol{V}}^T\boldsymbol{K}_V^{-1}\tilde{\boldsymbol{V}}\right\} \\ &+ \frac{1}{2}tr\left\{\tilde{\boldsymbol{c}}^T\boldsymbol{K}_c^{-1}\tilde{\boldsymbol{c}}\right\} \\ &+ \frac{1}{2}tr\left\{\tilde{\boldsymbol{\omega}}^T\boldsymbol{K}_\omega^{-1}\tilde{\boldsymbol{\omega}}\right\} \\ &+ \frac{1}{2}tr\left\{\tilde{\boldsymbol{k}}_d^T\boldsymbol{K}_d^{-1}\tilde{\boldsymbol{k}}_d\right\} \end{aligned} \tag{42}$$

where $\rho(\boldsymbol{s}) = [\,\rho(s_1) \quad \rho(s_2) \quad \rho(s_3)\,]^T$ $(\rho(s_k)$

$= \dfrac{\chi_k^2}{m_k}\left[\ln(1+e^{m_k s_k/\chi_k}) + \ln(1+e^{-m_k s_k/\chi_k})\right], k = 1,2,3).$

Take the time derivative of L and use $\dot{\boldsymbol{\rho}}(\boldsymbol{s}) = \boldsymbol{\sigma}^T(\boldsymbol{s})\dot{\boldsymbol{s}} = \boldsymbol{r}^T\dot{\boldsymbol{s}}$ and (34) and (37); then (42) can be written as

$$\begin{aligned}\dot{L} = &-\boldsymbol{r}^T\boldsymbol{K}_s\boldsymbol{s} + \boldsymbol{r}^T(\varepsilon_f + \boldsymbol{W}^T\tilde{\boldsymbol{\phi}}) + \boldsymbol{r}^T\tilde{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}} - \boldsymbol{r}^T\hat{\boldsymbol{d}} \\ &- tr\{\tilde{\boldsymbol{W}}^T[\hat{\boldsymbol{\phi}}(\boldsymbol{r} + \|\boldsymbol{r}\|\hat{\boldsymbol{V}}^T\hat{\boldsymbol{\phi}})^T - \eta\|\boldsymbol{r}\|\hat{\boldsymbol{W}}]\} \\ &- tr\{\tilde{\boldsymbol{V}}^T\boldsymbol{K}_V^{-1}\dot{\hat{\boldsymbol{V}}}\} - tr\{\tilde{\boldsymbol{c}}^T\boldsymbol{K}_c^{-1}\dot{\hat{\boldsymbol{c}}}\} - tr\{\tilde{\boldsymbol{\omega}}^T K_{\omega}^{-1}\dot{\hat{\boldsymbol{\omega}}}\} \\ &- tr\{\tilde{\boldsymbol{k}}_d^T\boldsymbol{K}_d^{-1}\dot{\tilde{\boldsymbol{k}}}_d\}. \end{aligned} \tag{43}$$

Substituting (32) and (38)–(40) into (43) yields

$$\begin{aligned}\dot{L} = &-\boldsymbol{r}^T\boldsymbol{K}_s\boldsymbol{s} + \boldsymbol{r}^T\left(\varepsilon_f + \boldsymbol{W}^T\tilde{\boldsymbol{\phi}}\right) - \boldsymbol{r}^T\hat{\boldsymbol{d}} \\ &- \|\boldsymbol{r}\|tr\{\boldsymbol{W}^T\hat{\boldsymbol{\phi}}(\hat{\boldsymbol{V}}^T\hat{\boldsymbol{\phi}})^T\} \\ &+ \|\boldsymbol{r}\|tr\left\{\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}}[-\tilde{\boldsymbol{V}}^T(\hat{\boldsymbol{\phi}} - \boldsymbol{A}^T\hat{\boldsymbol{\omega}} - \boldsymbol{B}^T\hat{\boldsymbol{c}}) \right. \\ &\qquad \left. -\hat{\boldsymbol{V}}^T(\boldsymbol{A}^T\tilde{\boldsymbol{\omega}} - \boldsymbol{B}^T\tilde{\boldsymbol{c}}) - \boldsymbol{d}_1]^T\right\} \\ &+ \eta\|\boldsymbol{r}\|tr\{\tilde{\boldsymbol{W}}^T\hat{\boldsymbol{W}}\} \\ &+ \|\boldsymbol{r}\|tr\left\{\hat{\boldsymbol{V}}^T(\hat{\boldsymbol{\phi}} - \boldsymbol{A}^T\hat{\boldsymbol{\omega}} - \boldsymbol{B}^T\hat{\boldsymbol{c}})(\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}})^T\right\} \\ &+ \eta\|\boldsymbol{r}\|tr\{\tilde{\boldsymbol{V}}^T\hat{\boldsymbol{V}}\} + \|\boldsymbol{r}\|tr\left\{\tilde{\boldsymbol{c}}^T\boldsymbol{K}_c\boldsymbol{B}\hat{\boldsymbol{V}}\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}}\right\} \\ &+ \eta\|\boldsymbol{r}\|tr\{\tilde{\boldsymbol{c}}^T\hat{\boldsymbol{c}}\} + \|\boldsymbol{r}\|tr\left\{\tilde{\boldsymbol{\omega}}^T\boldsymbol{A}\hat{\boldsymbol{V}}\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}}\right\} \\ &+ \eta\|\boldsymbol{r}\|tr\{\tilde{\boldsymbol{\omega}}^T\hat{\boldsymbol{\omega}}\} - tr\left\{\tilde{\boldsymbol{k}}_d^T\boldsymbol{K}_d^{-1}\dot{\tilde{\boldsymbol{k}}}_d\right\}. \end{aligned} \tag{44}$$

Because $\boldsymbol{r}^T\boldsymbol{K}_s\boldsymbol{s} \le \boldsymbol{\chi}^T\boldsymbol{K}_s|\boldsymbol{s}|$ $(|\boldsymbol{s}| = [|s_1| \; |s_2| \; |s_3|]^T)$, (44) becomes

$$\begin{aligned}\dot{L} \le &-\chi^T\boldsymbol{K}_s|\boldsymbol{s}| + \boldsymbol{r}^T(\varepsilon_f + \boldsymbol{W}^T\tilde{\boldsymbol{\phi}}) - \boldsymbol{r}^T\hat{\boldsymbol{d}} \\ &- \|\boldsymbol{r}\|(\hat{\boldsymbol{V}}^T\hat{\boldsymbol{\phi}})^T\boldsymbol{W}^T\hat{\boldsymbol{\phi}} - \|\boldsymbol{r}\|\boldsymbol{d}_1^T\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}} \\ &- tr\left\{\tilde{\boldsymbol{k}}_d^T\boldsymbol{K}_d^{-1}\dot{\tilde{\boldsymbol{k}}}_d\right\} + \eta\|\boldsymbol{r}\|tr\{\tilde{\boldsymbol{W}}^T\hat{\boldsymbol{W}}\} \\ &+ \eta\|\boldsymbol{r}\|tr\{\tilde{\boldsymbol{V}}^T\hat{\boldsymbol{V}}\} + \eta\|\boldsymbol{r}\|tr\{\tilde{\boldsymbol{c}}^T\hat{\boldsymbol{c}}\} + \eta\|\boldsymbol{r}\|tr\{\tilde{\boldsymbol{\omega}}^T\hat{\boldsymbol{\omega}}\} \end{aligned} \tag{45}$$

where $\boldsymbol{\chi} = [\,\chi_1 \quad \chi_2 \quad \chi_3\,]^T$. The uncertainty term $\boldsymbol{d}$ multiplied by $\boldsymbol{r}$ is given by

$$\boldsymbol{r}^T\boldsymbol{d} = \boldsymbol{r}^T(\varepsilon_f + \boldsymbol{W}^T\tilde{\boldsymbol{\phi}}) - \|\boldsymbol{r}\|(\hat{\boldsymbol{V}}^T\hat{\boldsymbol{\phi}})^T\boldsymbol{W}^T\hat{\boldsymbol{\phi}} - \|\boldsymbol{r}\|\boldsymbol{d}_1^T\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}}. \tag{46}$$

According to $|\hat{\phi}_j| \le 1$ and $|\tilde{\phi}_j| \le 1, j = 1, \ldots, N$, it is easy to obtain

$$\begin{aligned}\|\boldsymbol{W}^T\tilde{\boldsymbol{\phi}}\| &\le \sqrt{N}\|\boldsymbol{W}\| \le \sqrt{N}b_W \\ \|\hat{\boldsymbol{V}}^T\hat{\boldsymbol{\phi}}\| &\le \sqrt{N}\|\hat{\boldsymbol{V}}\| \\ \|\boldsymbol{W}^T\hat{\boldsymbol{\phi}}\| &\le \sqrt{N}\|\boldsymbol{W}\| \le \sqrt{N}b_W \\ \|\hat{\boldsymbol{W}}^T\hat{\boldsymbol{\phi}}\| &\le \sqrt{N}\|\hat{\boldsymbol{W}}\|. \end{aligned} \tag{47}$$

Based on assumption 1 and 2, and (47), the norm of (46), $\boldsymbol{r}^T\boldsymbol{d}$, satisfies the following inequality:

$$\frac{\eta}{4}\|\boldsymbol{r}\|\left(b_W^2 + b_V^2 + b_\omega^2 + b_c^2\right) + \left\|\boldsymbol{r}^T\boldsymbol{d}\right\| \le \|\boldsymbol{r}\|\boldsymbol{k}_d^T\boldsymbol{\varphi} \tag{48}$$

with $\boldsymbol{k}_d^T = \left[b_{\varepsilon_f} + \sqrt{N}b_W + (\eta/4)(b_W^2 + b_V^2 + b_\omega^2 + b_c^2)Nb_W \right.$ $\left. \sqrt{N}b_V(b_A b_\omega + b_B b_c + b_h + \sqrt{N})\sqrt{N}(b_A b_\omega + b_B b_c)\right]$. Since $tr\{\tilde{\boldsymbol{V}}^T\hat{\boldsymbol{V}}\} \le b_V\left\|\tilde{\boldsymbol{V}}\right\| - \left\|\tilde{\boldsymbol{V}}\right\|^2$, $tr\{\tilde{\boldsymbol{W}}^T\hat{\boldsymbol{W}}\} \le b_W\left\|\tilde{\boldsymbol{W}}\right\| - \left\|\tilde{\boldsymbol{W}}\right\|^2$, $tr\{\tilde{\boldsymbol{c}}^T\hat{\boldsymbol{c}}\} \le b_c\|\tilde{\boldsymbol{c}}\| - \|\tilde{\boldsymbol{c}}\|^2$, and $tr\{\tilde{\boldsymbol{\omega}}^T\hat{\boldsymbol{\omega}}\} \le b_\omega\|\tilde{\boldsymbol{\omega}}\| - \|\tilde{\boldsymbol{\omega}}\|^2$, we have

$$\begin{aligned}\dot{L} \le &-\boldsymbol{\chi}^T\boldsymbol{K}_s|\boldsymbol{s}| + \|\boldsymbol{r}\|\boldsymbol{k}_d^T\boldsymbol{\varphi} - \boldsymbol{r}^T\hat{\boldsymbol{d}} - tr\left\{\tilde{\boldsymbol{k}}_d^T\boldsymbol{K}_d^{-1}\dot{\tilde{\boldsymbol{k}}}_d\right\} \\ &- \eta\|\boldsymbol{r}\|\left[\left(\|\tilde{\boldsymbol{W}}\| - \frac{b_W}{2}\right)^2 + \left(\|\tilde{\boldsymbol{V}}\| - \frac{b_V}{2}\right)^2 \right. \\ &\qquad \left. + \left(\|\tilde{\boldsymbol{\omega}}\| - \frac{b_\omega}{2}\right)^2 + \left(\|\tilde{\boldsymbol{c}}\| - \frac{b_c}{2}\right)^2\right]. \end{aligned} \tag{49}$$

Rewrite (49) using (35) and (41) as

$$\begin{aligned}\dot{L} \le &-\boldsymbol{\chi}^T\boldsymbol{K}_s|\boldsymbol{s}| + \|\boldsymbol{r}\|\boldsymbol{k}_d^T\boldsymbol{\varphi} - \|\boldsymbol{r}\|\hat{\boldsymbol{k}}_d^T\boldsymbol{\varphi} - \|\boldsymbol{r}\|\tilde{\boldsymbol{k}}_d^T\boldsymbol{\varphi} \\ &- \eta\|\boldsymbol{r}\|\left[\left(\|\tilde{\boldsymbol{W}}\| - \frac{b_W}{2}\right)^2 + \left(\|\tilde{\boldsymbol{V}}\| - \frac{b_V}{2}\right)^2 \right. \\ &\qquad \left. + \left(\|\tilde{\boldsymbol{\omega}}\| - \frac{b_\omega}{2}\right)^2 + \left(\|\tilde{\boldsymbol{c}}\| - \frac{b_c}{2}\right)^2\right]. \end{aligned} \tag{50}$$

Hence, $\dot{L} \le -\boldsymbol{\chi}^T\boldsymbol{K}_s|\boldsymbol{s}|$ for all $t \ge 0$, and if $\boldsymbol{s}(t)$, $\tilde{\boldsymbol{W}}$, $\tilde{\boldsymbol{V}}$, $\tilde{\boldsymbol{c}}$, $\tilde{\boldsymbol{\omega}}$, and $\tilde{\boldsymbol{k}}_d$ are bounded at initial time $t = 0$, they will remain bounded for all $t \ge 0$. According to $\boldsymbol{e}(0)$ and $\boldsymbol{s}(0)$ are bounded, $\boldsymbol{e}(t)$ and $\boldsymbol{s}(t)$ are bounded for all $t \ge 0$, and since $\boldsymbol{y}_d(t)$ is bounded as specified, $\boldsymbol{y}(t)$ is bounded as well. All $\tilde{\boldsymbol{W}}$, $\tilde{\boldsymbol{V}}$, $\tilde{\boldsymbol{c}}$, $\tilde{\boldsymbol{\omega}}$, and $\tilde{\boldsymbol{k}}_d$ are bounded indicates all weights $\hat{\boldsymbol{W}}$, $\hat{\boldsymbol{V}}$, $\hat{\boldsymbol{c}}$, $\hat{\boldsymbol{\omega}}$, and $\hat{\boldsymbol{k}}_d$ are bounded.

To prove $\boldsymbol{s}(t) \to 0$ as $t \to \infty$, it can be easily shown by applying Barbalat's lemma

$$L_b(\boldsymbol{s}(t)) = L(t) - \int_0^t [\dot{L}(\tau) + \boldsymbol{\chi}^T\boldsymbol{K}_s|\boldsymbol{s}|]d\tau$$

with

$$\dot{L}_b(\boldsymbol{s}(t)) = -\boldsymbol{\chi}^T\boldsymbol{K}_s|\boldsymbol{s}|. \tag{51}$$

We can observe that $\boldsymbol{s}(\cdot)$ denotes the solution of $\boldsymbol{s}(t;\boldsymbol{u},\boldsymbol{s}(0))$ to the system $\dot{\boldsymbol{s}}(t)$ given by (25) and every bounded term in (42) implies $\dot{L}_b$ is a uniformly continuous function of time. Since $L_b(\boldsymbol{s}(t))$ is nonnegative, and $\dot{L}_b(\boldsymbol{s}(t)) \le 0$ for all time t, using Barbalat's lemma can prove that $\dot{L}_b(\boldsymbol{s}(t)) \to 0$ and hence $\boldsymbol{s}(t) \to 0$ as $t \to \infty$. ■

Remark: Since $\boldsymbol{s}(t) \to 0$ as $t \to \infty$, the reinforcement signal $\boldsymbol{r}(t) \to 0$ as $t \to \infty$ and the internal reinforcement signal $\hat{\boldsymbol{r}}_i(t) \to 0$ as $t \to \infty$ as well.

V. Simulation Results

In this section, we will present simulation results of the proposed adaptive critic autopilot applied to BTT missiles. A comparison of original with extended adaptive critic controller [21] confirms that the latter with online tuning capability shortens the learning time. Moreover, a missile cannot be launched again and again for the trial-and-error learning algorithm, and therefore no comparison result with original adaptive critic controller

is presented. Results presented here only for the proposed autopilot is under the assumption that the aerodynamic pressure $= 0.5\rho||V_m||$, where ρ is the air density determined by the height of missile which can be derived from Q_s function (i.e., Q_s is a function of height and velocity of missile). Thus, we simulated an 18-s flight scenario which operated within 0.5 to 4.5 Mach and 0 to 35 km, velocity and height range, respectively. For traditional gain-scheduling autopilot designs, a controller is required for each operating point. The proposed adaptive critic autopilot design, on the other hand, does not change parameters under different operating regimes. The desired commands of the four principal flight conditions are

1) $\Phi_c = 0°$, $A_{yc} = 0$ G, and $A_{zc} = -50$ G ($G = 9.8\ \mathrm{m/s^2}$);
2) $\Phi_c = 135°$, $A_{yc} = 0$ G, and $A_{zc} = -20$ G;
3) $\Phi_c = 90°$, $A_{yc} = 0$ G, and $A_{zc} = -10$ G;
4) $\Phi_c = 0°$, $A_{yc} = 0$ G, and $A_{zc} = -5$ G.

In the following simulations, the BTT missile is subject to the following physical limitations:

1) attack angle α: $\alpha \geq 0°$;
2) sideslip angle β: $|\beta| \leq 5°$;
3) roll rate P: $|P| \leq 500°/\mathrm{s}$;
4) pitch rate Q: $|Q| \leq 150°/\mathrm{s}$;
5) yaw rate R: $|R| \leq 150°/\mathrm{s}$;
6) actuator position saturation: $|u_p| \leq 20°$, $|u_q| \leq 20°$ and $|u_r| \leq 20°$.

The number of FBFs (fuzzy rules) is chosen as 20 and each FBF specifies their initial centers and inverse radii as small random numbers. For traditional fuzzy system design, if each variable utilized is characterized by three fuzzy linguistic terms, then the number of fuzzy rules required for the BTT missile system with 12 input variables is 3^{12}. From the viewpoint of computation efficiency and real-time control, a huge and sparse fuzzy rule base is not adequate and, therefore pre-determined number of FBFs with self-tuned parameters of the proposed approach is a better design strategy. The initial state is assigned as $\boldsymbol{x}(0) = [0\,0\,0\,0\,0\,0\,1020\ (\mathrm{m/s})0\ -0.01\ (\mathrm{m/s})\,0\,0\ -5000\ \mathrm{m}]^T$. The parameters of the adaptive critic autopilot is as follows.

1) sliding-surface: $\boldsymbol{s}(t) = [\ddot{e}_1 + 26\dot{e}_1 + 169e_1\dot{e}_2 + 2e_2\dot{e}_3 + 10e_3]^T$;
2) fixed gain: $\boldsymbol{K}_s = \mathrm{diag}\{150, 60, 50\}$;
3) parameters of reinforcement signals: $\boldsymbol{\chi} = [50\ 100\ 500]^T$ and $[m_1\ m_2\ m_3] = [0.2\ 1\ 1]$;
4) parameters of updating law: $\eta = 0.1$, $\boldsymbol{K}_V = \boldsymbol{K}_W = \mathrm{diag}\{80, 5, 5\}$, $\boldsymbol{K}_c = \boldsymbol{K}_\omega = 20I_{nN}$ (I_{nN} is an $nN \times nN$ identity matrix), and $\boldsymbol{k}_d = [1\quad 1\quad 1\quad 1]^T$.

The simulation programs written in C language run on Pentium 4 PC with 128-MB RAM. The desired trajectories Φ_d and W_d are generated by $\dot{\Phi}_d = -10\Phi_d + 10\Phi_c$ and $\dot{W}_d = -10W_d + 10k_{A_z}(A_z - A_{zc})$, where k_{A_z} is dependent on the height of the missile and A_z. In the four nonzero commands intervals, BTT missile should roll to Φ_c and achieve the demand acceleration command A_{zc} by means of a high rolling rate P [as shown in Fig. 7(a)] which induces a cross-coupling effect resulting in a small transient in pitch acceleration A_y and a small overshoot in yaw acceleration A_z as shown in Fig. 7(b) and (c). Furthermore, Fig. 7 also confirms the cross-coupling effect can be overcome by our autopilot design. Such a continuous flight scenario only requires a single proposed autopilot design as oppose to many linear controllers in gain-scheduled autopilot design.

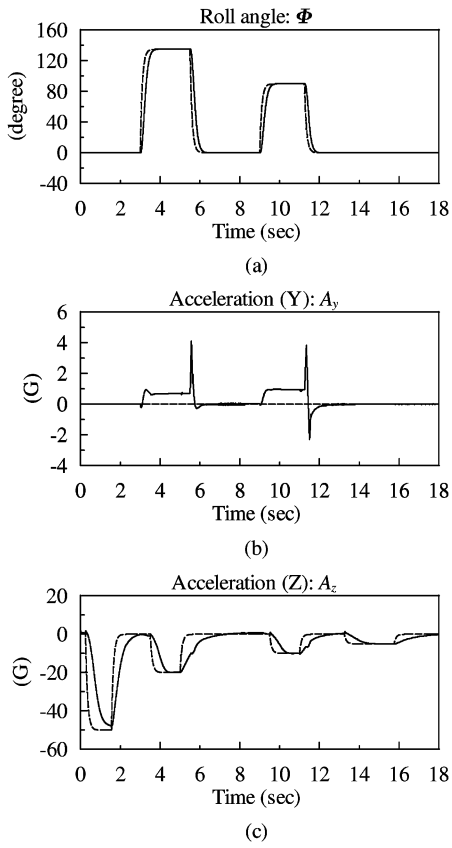

Fig. 7. Outputs of BTT missiles show the responses of an 18-s flight scenario: (a) roll angle Φ; (b) acceleration A_y; (c) acceleration A_z (dashed line: desired trajectory; solid line: actual trajectory).

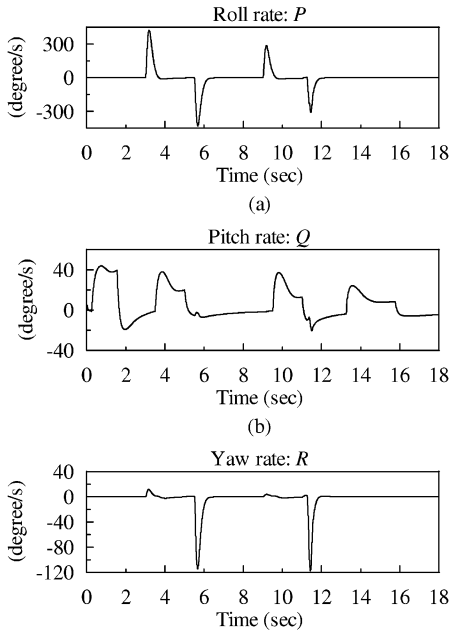

Fig. 8. Simulation results of roll rate, pitch rate, and yaw rate show that all meet the requirements: (a) $|P| \leq 500°/\mathrm{s}$; (b) $|Q| \leq 50°/\mathrm{s}$; and (c) $|R| \leq 150°/\mathrm{s}$.

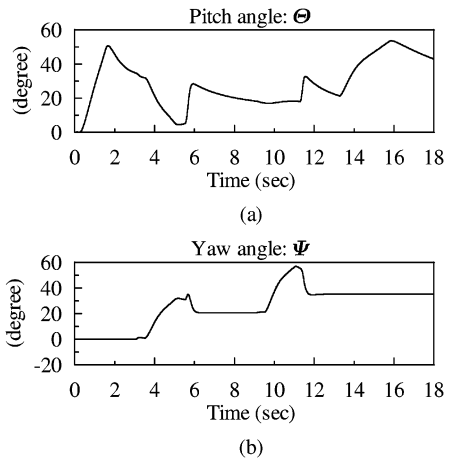

Fig. 9. Simulation results of pitch angle and yaw angle.

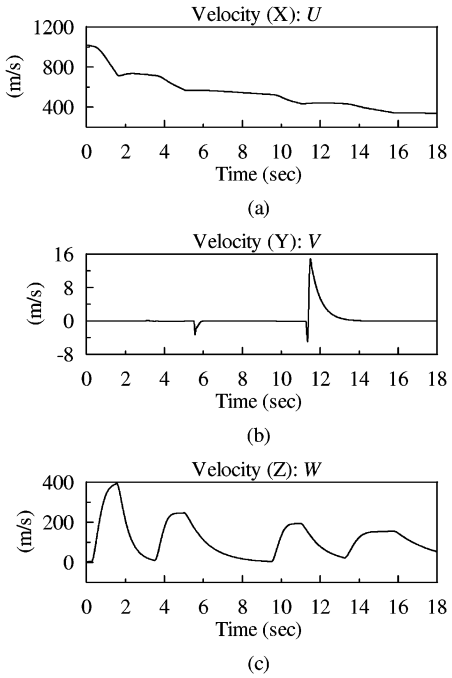

Fig. 10. Simulation results of body frame velocities of the BTT missile show that: (a) U and (c)W have significant magnitudes and (b) V keeps small for achieving the small A_y.

Figs. 8 and 12 show that not only the high roll rate P, pitch rate Q, and yaw rate R can meet the physical limitation, but also inputs of actuators can satisfy the physical requirements. Moreover, in Fig. 13, we have positive angle-of-attack α and small sideslip angle β as required. It should be noted that all training processes are online and all physical limitations are satisfied. Other states of BTT missiles including pitch angle, yaw angle, positions, and velocities are shown in Figs. 9–11. The reinforcement signal $\boldsymbol{r}(t)$, which will approach zero, is depicted in Fig. 14. From the simulation results, the robust stability and tracking performance of the proposed adaptive critic control scheme can be confirmed. Compared with traditional adaptive critic designs and reinforcement learning framework, the proposed adaptive critic control scheme can learn faster due to no time-consuming trial-and-error learning phase.

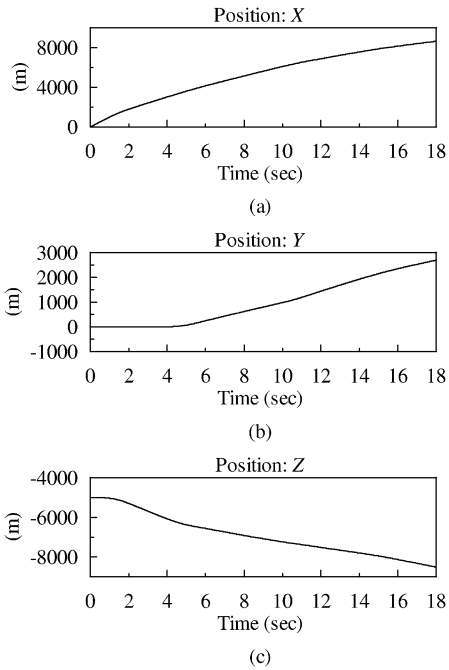

Fig. 11. Simulated positions of the BTT missile show the flight trajectories.

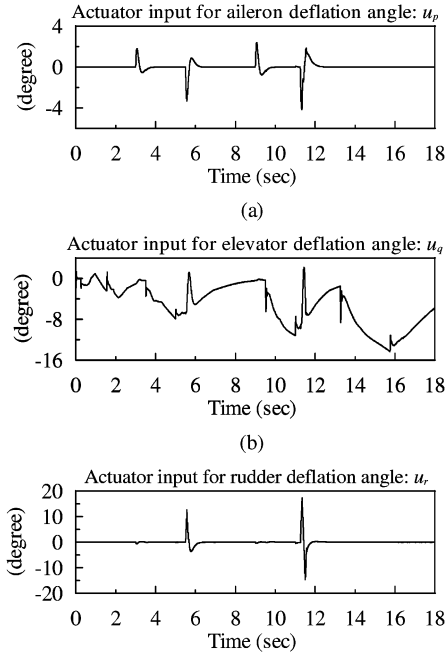

Fig. 12. Simulation results of actuator inputs of the BTT missile meet the specification, i.e.: (a) $|u_p| \leq 20°$; (b) $|u_q| \leq 20°$; and (c) $|u_r| \leq 20°$.

VI. Conclusion

In this paper, a new adaptive critic autopilot has been proposed to control BTT missiles. Unlike traditional reinforcement learning approaches, we adopted an FBFN-based ASE to approximate the nonlinear dynamics of a BTT missile which output is then used to augment the control action. Combined with robust adaptive control and Lyapunov theory, all parameters of FBFs, ACE and ASE can be online tuned with satisfactory tracking performance and guaranteed robust

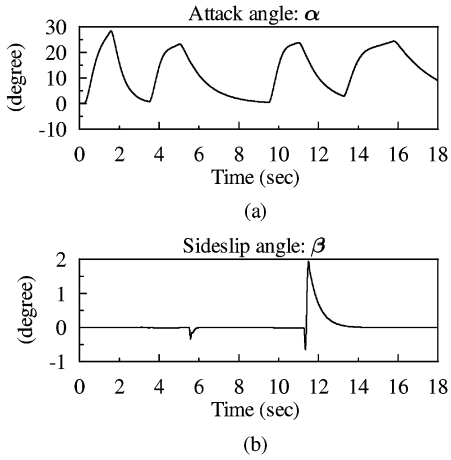

Fig. 13. Simulation results of attack angle α and sideslip angle β show that the attack angle keeps positive and the sideslip angle remains small.

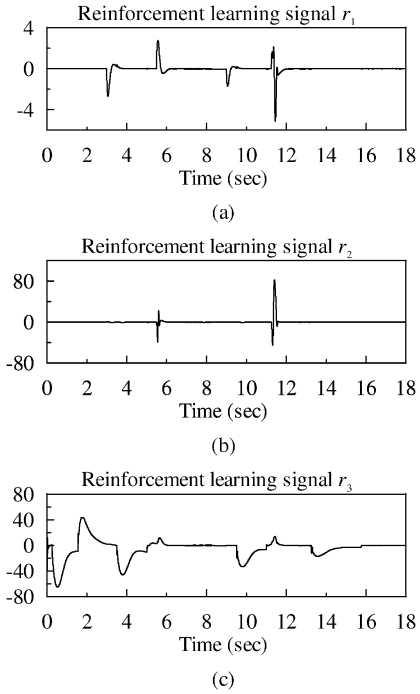

Fig. 14. Simulation results of reinforcement signals: (a) r_1, (b) r_2, and (c) r_3, which will decay to zero.

stability. Compared with traditional adaptive critic designs and reinforcement learning approaches, the proposed adaptive critic autopilot can prevent from time-consuming trial-and-error learning, which is not suitable for autopilot design. Furthermore, we can use only one autopilot through the entire flight process containing various flight conditions by adaptive control law and adaptive updating laws, and effectively reduce the size of rule base due to all tunable parameters of FBFs. Simulation results for the proposed autopilot applied to BTT missiles demonstrate that the control objectives can be achieved effectively and successfully.

ACKNOWLEDGMENT

The author would like to thank the Associate Editor and the anonymous referees for their constructive comments and suggestions.

REFERENCES

[1] S.-Y. Lee, J.-I. Lee, and I.-J. Ha, "Nonlinear autopilot for high maneuverability of bank-to-turn missiles," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 4, pp. 1236–1253, Oct. 2001.

[2] K.-Y. Lian, L.-C. Fu, D.-M. Chuang, and T.-S. Kuo, "Adaptive robust autopilot design for bank-to-turn aircraft," in *Proc. American Control. Conf.*, 1993, pp. 1746–1750.

[3] J. Huang and C. F. Lin, "Application of sliding mode control to bank-to-turn missile systems," in *Proc. 1st IEEE Regional Conf. Aerospace Control Systems*, 1993, pp. 569–573.

[4] L. H. Carter and J. S. Shamma, "Gain-schedule bank-to-turn autopilot design using linear parameter varying transform," *J. Guid., Contr., Dynam.*, vol. 19, no. 5, pp. 1056–1063, 1996.

[5] C.-K. Lin and S.-D. Wang, "An adaptive H^{∞} controller design for bank-to-turn missiles using ridge gaussian neural networks," *IEEE Trans. Neural Netw.*, vol. 15, no. 6, pp. 1507–1516, Nov. 2004.

[6] C.-F. Lin and J.-H. Ge, "Application of nonlinear H_{∞} control to autopilot design," in *Proc. IEEE Conf. Decision and Contol.*, 1996, pp. 1359–1364.

[7] D. M. McDowell, G. W. Irwin, G. Lightbody, and G. McConnell, "Hybrid neural adaptive control for bank-to-turn missiles," *IEEE Trans. Contr. Syst. Technol.*, vol. 5, no. 3, pp. 297–308, May 1997.

[8] L.-C. Fu, W.-D. Chang, J.-H. Yang, and T.-S. Kuo, "Adaptive robust bank-to-turn missile autopilot design using neural networks," *J. Guid., Contr., Dynam.*, vol. 20, no. 2, pp. 346–354, 1997.

[9] Z. Youan, Y. Huadong, and H. Yunan, "Robust nonlinear control for bank-to-turn missile," in *Proc. 4th World Congr. Intelligent Control and Automation*, 2002, pp. 2973–2977.

[10] M. C. Mickle and J. J. Zhu, "Bank-to-turn roll-yaw-pitch autopilot design using dynamic nonlinear inversion and PD-eigenvalue assignment," in *Proc. American Control Conf.*, 2000, pp. 1359–1364.

[11] C.-K. Lin and S.-D. Wang, "A self organizing fuzzy control approach for bank-to-turn missiles," *Fuzzy Sets Syst.*, vol. 96, no. 3, pp. 281–306, 1998.

[12] M. B. McFarland and A. J. Calise, "Adaptive nonlinear control of agile antiair missiles using neural networks," *IEEE Trans. Contr. Syst. Technol.*, vol. 8, no. 5, pp. 749–756, Sep. 2000.

[13] G. Lightbody and G. W. Irwin, "Neural model reference adaptive control and application to a BTT-CLOS guidance system," in *Proc. Int. Joint Conf. Neural Networks*, 1994, pp. 2429–2435.

[14] D. Han and S. N. Balakrishnan, "State-constrained agile missile control with adaptive-critic-based neural networks," *IEEE Trans. Contr. Syst. Technol.*, vol. 10, no. 4, pp. 481–489, Jul. 2002.

[15] A. G. Barto, R. S. Sutton, and W. Anderson, "Neuronlike adaptive elements can solve difficult learning control problems," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, no. 5, pp. 834–846, Sep./Oct. 1983.

[16] D. V. Prokhorov and D. C. Wunsch, "Adaptive critic designs," *IEEE Trans. Neural Netw.*, vol. 8, no. 5, pp. 997–1007, Sep. 1997.

[17] P. J. Werbos, "Consistency of HDP applied to a simple reinforcement learning problem," *Neural Netw.*, vol. 3, no. 2, pp. 179–189, 1990.

[18] S. Jagannathan and G. Galan, "Adaptive critic neural network-based object grasping control using a three-finger gripper," *IEEE Trans. Neural Netw.*, vol. 15, no. 2, pp. 395–407, Mar. 2004.

[19] O. Kuljaca and F. L. Lewis, "Adaptive critic design using nonlinear network structures," *Int. J. Adapt. Contr. Signal Process.*, vol. 17, no. 6, pp. 431–445, 2003.

[20] Y. H. Kim and F. L. Lewis, "Reinforcement adaptive learning neural-net-based friction compensation control for high speed and precision," *IEEE Trans. Contr. Syst. Technol.*, vol. 8, no. 1, pp. 118–126, Jan. 2000.

[21] C.-K. Lin, "A reinforcement learning adaptive fuzzy controller for robots," *Fuzzy Sets Syst.*, vol. 137, no. 3, pp. 339–352, 2003.

[22] J. Nie and D. A. Linkens, "Learning control using fuzzified self-organizing radial basis function network," *IEEE Trans. Fuzzy Syst.*, vol. 1, no. 1, pp. 280–287, 1993.

[23] L. X. Wang and J. M. Mendel, "Fuzzy basis functions, universal approximation, and orthogonal least-squares learning," *IEEE Trans. Neural Netw.*, vol. 3, no. 5, pp. 807–814, Sep. 1992.

[24] R. M. Sanner and J. -J. E. Slotine, "Gaussian networks for direct adaptive control," *IEEE Trans. Neural Netw.*, vol. 3, no. 6, pp. 837–863, Nov. 1992.

[25] P. J. Werbos, "Backpropagation through time: What it is and how to do it," *Proc. IEEE*, vol. 78, no. 10, pp. 1550–1560, Oct. 1990.

Chuan-Kai Lin was born in Kaohsiung, Taiwan, R.O.C., in 1967. He received the B.S. degree in electrical engineering from the National Sun Yat-Sen University, Kaohsiung, in 1989 and the M.S. and Ph.D. degrees in electrical engineering from National Taiwan University, Taipei, in 1991 and 1997, respectively.

From 1993 to 1998, he was with Chunghwa Telecom Co., Ltd. During 1998–2001, he served as an Assistant Professor at Southern Taiwan University of Technology. In August 2001, he joined the Department of Electrical Engineering, Chinese Naval Academy, Kaohsiung, where he is currently an Associate Professor. His research interests include soft computing and intelligent control with application to motor servo drives, robots, underwater vehicles, and flight control.