Fuzzy optimal tracking control of hypersonic flight vehicles via single-network adaptive critic design

Xiangwei Bu*, Qiang Qi

Abstract—Optimal performance is extremely important for hypersonic flight control. Different from the most existing methodologies which only consider basic control performance including stability, robustness and transient performance, this article deals with the design of nearly optimal tracking controllers for hypersonic flight vehicles (HFVs). Firstly, main controllers are developed for the velocity subsystem and the altitude subsystem of HFVs via concise fuzzy approximations. Then, optimal controllers are nearly implemented utilizing single-network adaptive critic design. Moreover, the stability of closed-loop systems and the convergence of optimal controllers are theoretically proved. Finally, compared simulation results are given to verify the superiority. The special contribution is the application of low-complex control structure owing to the critic-only network and advanced learning laws developed for fuzzy approximations, which is expected to guarantee satisfied real-time performance.

Index Terms—Hypersonic flight vehicles, Optimal performance, Fuzzy approximations, Single-network adaptive critic design, Real-time performance

I. INTRODUCTION

Hypersonic flight vehicles (HFVs) are potential to serve as long-range transports and carriers of rapid and accurate attack weapons [1]-[5]. The design of control systems for HFVs continues to be a topic of important research interest, and it is inherently difficult due to the fact that the vehicle dynamics is nonlinear, coupled and uncertain. Furthermore, hypersonic flight controllers must handle narrow flight envelopes, rapidly time-varying flight circumstances and notable flexible effects. In addition to ensuring control authority over the entire flight envelope, optimal index is also necessary for HFVs’ control systems to accomplish miscellaneous missions.

Hypersonic flight control has received worldwide attentions in recent years because of the interesting and hard-to-handle flight conditions connected with high Mach numbers. Thereby, considerable efforts have been made by researchers to exploit advanced controllers for HFVs for the purpose of attaining stable tracking of reference trajectories [6]-[9]. In the literature, stability performance of the closed-loop control system is a primarily considered index for HFVs by incorporating baseline controllers with additional terms [10]-[13]. In [12], a robust control method is studied for HFVs to maintain control stability and reject parametric perturbations. Firstly, dynamic inversion is combined with back-stepping to develop baseline controllers for the velocity subsystem and the altitude subsystem. And then, sliding mode switching terms together with neural estimators are applied to increase the tolerance of closed-loop control systems to external disturbances and model uncertainties. An alternative method which is available for resisting system uncertainties and external disturbances of HFVs is the active disturbance rejection control (ADRC) approach [14]. The difference from [12] is that it [14] only considers the attitude control issue. A common problem which arises in the hypersonic control domain is actuator faults/saturations [15]-[17]. This may result in tracking performance reducing even instability. A possible solution to this problem is to add auxiliary systems to baseline controllers. The compensation signals generated by auxiliary systems can stabilize hypersonic flight control systems in the presence of actuator faults/saturations. Unlike the above disturbance-compensation methodologies, a new offset-free control approach is proposed in [18] to make control system resistant to disturbances and unknown dynamics. Besides, other estimators such as disturbance observers [8] and intelligent approximations [19]-[21] also are usually used to resist disturbances. On the other hand, except for stabilization, transient performance is widely considered to be a very significant index for hypersonic flight control systems. Prescribed performance control (PPC) [9],[22] has been shown to efficiently guarantee transient performance. The key point of PPC is to devise performance functions which are used to impose funnel constraints on tracking errors. And then, the desired prescribed performance is realized owing to the boundedness of transformed errors [9],[22].

Most of the methodologies mentioned above however are aimed at guaranteeing steady-state performance and transient performance for hypersonic flight control systems under different actual conditions including uncertainties, disturbances and actuator saturations/faults. Unfortunately, only such basic performance isn’t enough for hypersonic flight control and instead we must further seek some optimal performance indexes [23]-[27]. Adaptive critic design (ACD) is a newly emerging methodology to solve optimal control problems, and

Manuscript received XX, 2020; revised XX, 202X; accepted XX XX, 202X. This journal was supported by Young Talent Support Project for Military Science and Technology (Grant No. 18-JCJQ-QT-007). (Corresponding author: Xiangwei Bu).

X. Bu and Q. Qi are with the Air and Missile Defense College, Air Force Engineering University, Xi’an 710051, China (e-mail: buxiangwei1987@126.com; qiqiang1225@126.com).
it has many possible uses in the optimal control of simple dynamic systems. Its typical structure is the dual-network framework containing a critic network and an action network. This is also called the critic-actor structure inspired by reinforcement learning. The main superiority of ACD in comparison with traditional optimal controllers is that its critic network can generate strength signals based on the current control quality to further improve the action movement [28]-[34]. This finally minimizes cost functions and enhances control performance. Despite this advantage, a well-known problem with dual-network structure is that it results in high-computation burdens caused by the fact that both the critic network and the action network should be approximately estimated by neural or fuzzy approximations. However, another popular framework called single-network ACD (SACD) exhibits a better application prospect because it has simpler structure compared with dual-network ACD. Even so, the existing SACD cannot be directly employed to hypersonic flight control because of the following serious defects: (1) most of SACD strategies only focus on stabilization problem while hypersonic flight control belongs to a trajectory tracking control issue and it enables system outputs to track given reference commands. Then, the control objective is to devise the control input must eventually converge to zero, which is unrealistic for hypersonic flight control systems.

This article considers nearly optimal trajectory tracking control designing for HFVs with uncertain dynamics. The main contributions are summarized as:

1) The control structure complexity is low via concise SACD for the sake of guaranteeing real-time performance. Different from the existing studies [27],[28],[33], this paper proposes a single-adaptive-parameter-based strategy to construct the critic network which has only one adaptive parameter.

2) The previous methods [29],[34] focus on the stabilization problem, that is, system outputs must converge to zero. The addressed optimal controller in this paper is extended to the tracking control issue and it enables system outputs to track given reference commands.

3) The optimal tracking methodologies developed in [30]-[32] are only applicable to special dynamic systems whose control inputs must eventually converge to zero. Such restriction is released in this paper and the control inputs are allowed to converge to non-zero constants.

II. PROBLEM FORMULATION AND PRELIMINARIES

A. Vehicle model

The considered vehicle model consists of five rigid-body states (velocity \( V \), altitude \( h \), flight-path angle \( \gamma \), pitch angle \( \phi \), and pitch rate \( \dot{\phi} \)) and two flexible states \( \eta_1 \) and \( \eta_2 \). The motion equations are described as [35]:

\[
V = \frac{1}{m}[\cos(\theta - \gamma)T - D] - g \sin \gamma
\]

\[
h = V \sin \gamma
\]

\[
\dot{\gamma} = \frac{1}{m}[\sin(\theta - \gamma)T - m - g \cos \gamma]
\]

\[
\dot{\phi} = \theta = O
\]

\[
\dot{\theta} = Q
\]

\[
\dot{\eta}_1 = 2\zeta_1 \omega \eta_1 - \frac{\eta_1 \omega}{k_1} - \frac{N_1}{k_1} \psi_2 (M + \psi_2 \eta_2)
\]

\[
\dot{\eta}_2 = 2\zeta_2 \omega \eta_2 - \frac{\eta_2 \omega}{k_2} - \frac{N_2}{k_2} \psi_2 (M + \psi_2 \eta_2)
\]

\[
\dot{\phi} = \eta_1 + \eta_2
\]

\[
\dot{\psi}_1 = \frac{2\zeta_3 \omega \psi_1 - \psi_1 \omega}{k_3} - \frac{N_3}{k_3} \psi_2 (M + \psi_2 \eta_2)
\]

\[
\dot{\psi}_2 = \frac{2\zeta_4 \omega \psi_2 - \psi_2 \omega}{k_4} - \frac{N_4}{k_4} \psi_2 (M + \psi_2 \eta_2)
\]

\[
\dot{\psi}_3 = \frac{2\zeta_5 \omega \psi_3 - \psi_3 \omega}{k_5} - \frac{N_5}{k_5} \psi_2 (M + \psi_2 \eta_2)
\]

\[
\dot{\psi}_4 = \frac{2\zeta_6 \omega \psi_4 - \psi_4 \omega}{k_6} - \frac{N_6}{k_6} \psi_2 (M + \psi_2 \eta_2)
\]

\[
\dot{\psi}_5 = \frac{2\zeta_7 \omega \psi_5 - \psi_5 \omega}{k_7} - \frac{N_7}{k_7} \psi_2 (M + \psi_2 \eta_2)
\]

where \( T, D, L, M, N_1 \) and \( N_2 \) stand for thrust force, drag force, lift force, pitching moment, and generalized forces, respectively. These forces are functions of system states and control inputs, given by

\[
T = a_1 (\beta_1 \Phi + \beta_2) + a_2 (\beta_3 \Phi + \beta_4) + a_3 (\beta_5 \Phi + \beta_6) + \beta_7 \Phi + \beta_8
\]

\[
D = 0.5 \rho V^2 S \left( C_{D0} + C_{D1} \delta + C_{D2} \delta^2 + C_{D3} \delta + C_{D4} \right)
\]

\[
L = 0.5 \rho V^2 S \left( C_{L0} + C_{L1} \delta + C_{L2} \right)
\]

\[
M = z_i T + 0.5 \rho V^2 S T \left[ C_{M0} \alpha^2 + C_{M1} \alpha + C_{M2} \delta + C_{M3} \delta^2 + C_{M4} \right]
\]

\[
N_1 = N_{10} \alpha^2 + N_{11} \alpha + N_{12} \delta + N_{13} \delta^2 + N_{14}
\]

\[
N_2 = N_{20} \alpha^2 + N_{21} \alpha + N_{22} \delta + N_{23} \delta^2 + N_{24}
\]

The control inputs include the elevator angular deflection \( \delta_e \) and the fuel equivalence ratio \( \Phi \). All the definitions of coefficients and variables in the above equations can be seen in [35]. We should mention that only rigid-body states can be actually measured such that they are available for state feedback. While, the flexible states are often suppressed as disturbances. Then, the control objective is to devise \( \Phi \) and \( \delta_e \) such that \( V \to V_{ad} \) and \( h \to h_{ad} \), where \( V_{ad} \) and \( h_{ad} \) are chosen reference trajectories for velocity and altitude.

B. Fuzzy approximation

In this subsection, we recall the basic principle of fuzzy approximations which will be applied to stabilize the unknown dynamics of HFVs’ model via an estimation-compensation approach.

The fuzzy approximation of a continuous function \( f(x) \) is formulated as the following input-to-output mapping [36]

\[
f(x) = \zeta f(x) + \epsilon
\]

where \( f(x) \) is the function of \( x \) and \( x = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n \) is the input vector of fuzzy system (8). \( \zeta = [\zeta_1, \zeta_2, \ldots, \zeta_n]^T \in \mathbb{R}^n \) is a weight vector. \( \zeta f(x) = [\zeta_1 f_1(x), \zeta_2 f_2(x), \ldots, \zeta_n f_n(x)]^T \) is a basis function vector with \( \zeta_i f_i(x) = \prod_{j=1}^n \mu_{\zeta_i}(x_j) \left( \sum_{i=1}^n \prod_{j=1}^n \mu_{\zeta_i}(x_j) \right) \), \( i=1,2,\ldots,n; f=1,2,\ldots,N \), where \( \mu_{\zeta_i}(x_j) \) is a Gaussian-function-based fuzzy membership function. \( \epsilon \) means the fuzzy approximation error. It has been proved that there exists a constant \( \epsilon^* \in \mathbb{R}^n \) such that

\[
\sup_{x \in \Omega} |f(x) - \zeta f(x)| = \epsilon \leq \epsilon^*
\]

where \( \Omega \) is a compact set.

Remark 1. Though \( \epsilon \) can be made infinitesimal when choosing sufficiently large dimensions for \( \zeta \) and \( \zeta f(x) \),
the unknown term $\varepsilon_f$ still leads to a problem related to the implementations of control laws, which is usually tackled using the adaptive strategy by online updating $\zeta_f$. This also results in high computational costs owing to too many elements of $\zeta_f$. For an arbitrarily unknown function, we can obtain a low-computational approximation by directly turning $||\zeta_f||^2$ instead of its elements. This is very significant for HFVs to guarantee their control systems with excellent real-time performance.

Based on SACD, function approximations are necessary for critic network designs. Unfortunately, the above fuzzy system (8) isn’t suitable for HFVs to devise critic networks. According to the gradient descend method, we must adjust all the elements of $\zeta_f$ via developing adaptive laws based on Lyapunov theory. Undoubtedly, the control real-time performance cannot meet the requirement of hypersonic flight control if we use fuzzy formulation (8) to develop critic networks. For this reason, we give another form of fuzzy approximations, called Fuzzy Hyperbolic Model (FHM) [37], to construct the critic network subsequently.

FHM can be utilized to reconstruct an arbitrary dynamic system

$$\phi_i = W_{i}^T \tanh(K_{i} \phi) + W_{i}^T \tanh(K_{i} \psi),$$

where $\phi = [\phi_1, \phi_2, \ldots, \phi_n]^T \in \mathbb{R}^n$ is a system state vector and $\psi = [\psi_1, \psi_2, \ldots, \psi_n]^T \in \mathbb{R}^n$ an input vector. $W_{i}^T \in \mathbb{R}^{m \times n}$ and $W_{i}^T \in \mathbb{R}^{m \times n}$ are weight vectors. $\tanh(K_{i} \phi) = [$tanh($k_1 \phi_1$), tanh($k_2 \phi_2$), ..., tanh($k_n \phi_n$)] and $\tanh(K_{i} \psi) = [$tanh($k_1 \psi_1$), tanh($k_2 \psi_2$), ..., tanh($k_n \psi_n$)] are Hyperbolic basis function vectors.

**Remark 2.** In what follows, we use FHM to design critic networks, which helps to reduce computation loads and ensure real-time performance because $\phi_i$ and $W_{i}^T \phi$ degenerate into scalars in the design process of each critic network.

### III. CONTROLLER DESIGN

**A. Vehicle controller design**

This subsystem presents the design process of a fuzzy optimal controller for velocity subsystem (1) to make $V \rightarrow V_{ref}$ and minimize cost function (20).

We represent velocity subsystem (1) as

$$\dot{V} = a_v + \Phi + \Phi^*$$

(10)

where $\Phi = \Phi + \Phi^*$. $\Phi$ is a main controller and $\Phi^*$ is an optimal controller which is used to optimize performance index (20). $a_v = T \cos(\theta - \gamma) / m - D / m - g \sin \gamma - \Phi$ is an unknown but continuous function. Note that $a_v$ is a function of states and control inputs, and control inputs are functions of states since they are computed based on state-feedback controllers. Hence, we can use fuzzy system (8) to estimate $a_v$,

$$a_v = \zeta_v^T \varphi_i (x) + \varepsilon_v,$$

(11)

where $\zeta_v = [\zeta_{v1}, \zeta_{v2}, \ldots, \zeta_{vn}]^T \in \mathbb{R}^n$ is a weight vector and $X = [V, h, \gamma, \theta, Q]^T \in \mathbb{R}^n$ is an input vector. The estimation error $\varepsilon_v$ satisfies $|\varepsilon_v| \leq \varepsilon_v^M$ with a positive constant $\varepsilon_v^M$.

The solution $\varphi_i (x)$ has the same formulation as $\varphi_i (x)$. Define $\tilde{V} = V - V_{ref}$. Invoking (10), we obtain

$$\dot{\tilde{V}} = a_v + \Phi + \Phi^* - \tilde{V}$$

(13)

We design $\Phi$ as

$$\Phi = -k_1 \tilde{V} - k_2 \int_{0}^{t} \tilde{V} dt - 0.5 \tilde{V} \hat{\varphi}_i \varphi_i^T (x) \varphi_i (x) + \tilde{V}_{ref}$$

(14)

where $k_{11} \in \mathbb{R}^n$ and $k_{12} \in \mathbb{R}^n$ are constants, and $\hat{\varphi}_i$ is the estimation of $\varphi_i = ||\zeta_f||^2$. We devise the following learning law for $\hat{\varphi}_i$.

$$\dot{\hat{\varphi}}_i = 0.5 \hat{\varphi}_i \varphi_i^T (x) \varphi_i (x) - 2 k_i \hat{\varphi}_i,$$

(15)

where $k_i \in \mathbb{R}^n$ is a constant.

Define Lyapunov function

$$L_{ij} = 0.5 \hat{\varphi}_i^2 + 0.5 k_{12} \int_{0}^{t} \hat{\varphi}_i^2 dt + 0.5 \psi_i^2 / L_i$$

(16)

with $\psi_i = \hat{\varphi}_i - \varphi_i$.

Utilizing (11) and (13)-(15), $\dot{L}_{ij}$ is reduced as

$$\dot{L}_{ij} = -k_{11} \tilde{V}^2 + 0.5 \dot{\varphi}_i \varphi_i^T (x) \varphi_i (x)$$

(17)

$$+ 0.5 \hat{\varphi}_i \varphi_i^T (x) \varphi_i (x) - 2 k_i \hat{\varphi}_i \varphi_i / L_i + 0.5 \psi_i^2 + \tilde{V}^2$$

Because of $2 k_{11} \psi_i \varphi_i \geq -k_{11} \varphi_i^2$, $\dot{L}_{ij} \leq 0.5 \varepsilon_v^2 + 0.5 \varepsilon_v^M$ and $\dot{\varphi}_i \varphi_i^T (x) \varphi_i (x) \leq 0.5 \varepsilon_v^2 + 0.5 \varepsilon_v^M + \varphi_i^T \varphi_i$.

$$\dot{L}_{ij} \leq -(k_{11} - 0.5) \varphi_i^2 + 0.5 k_{12} \psi_i \varphi_i^T (x) \varphi_i (x) + 0.5 \psi_i^2$$

(18)

The existing work [8] shows that $\Sigma_{ij}$ can be stabilized. For $\Sigma_{ij}$, we will develop an optimal controller based on SACD.

We give the following dynamics

$$\dot{\tilde{V}} = \Phi^*$$

(19)

Define the following cost function

$$J_r = \int_{0}^{T} [Q_r (\tilde{V}) + \partial_r (\Phi^*)] dt$$

(20)

where $Q_r (\tilde{V}) = q_r \tilde{V}^2$ and $\partial_r (\Phi^*) = r_1 (\Phi^*)^2$ with positive constants $q_r$ and $r_1$.

We further obtain Hamilton-Jacobi-Bellman (HJB) equation

$$H_r (\tilde{V}, \Phi^*) = Q_r (\tilde{V}) + \partial_r (\Phi^*) + \nabla J_r \Phi^*$$

(21)

with $\nabla J_r = \nabla \partial_r / \partial \tilde{V}$. Design $\Phi^*$ as

$$\phi_r = -0.5 k_{r1} \nabla J_r$$

(22)

Since $\nabla J_r$ is an unknown term, we use FHM to construct the critic network via adaptively estimating $J_r$.

$$J_r = A_t \tanh (K_t \tilde{V}) + e_i$$

(23)

where $A_t \in \mathbb{R}$, $K_t \in \mathbb{R}^n$ and $e_i$ is an estimation error.
From (23), we have
\[ \nabla J_v = \nabla J_v = A_p \dot{A}_p + \delta e_{\Sigma} / \partial \dot{V} \]
with
\[ \dot{A}_p = \hat{e} \tanh(K_v \dot{V}) / \partial \dot{V}. \]

Substituting (24) into (21), we have
\[ H_v (\dot{V}, \Phi^*, J_v) = -e_{\Sigma} \dot{A}_p + Q_v (\dot{V}) + \delta e_{\Sigma} (\Phi^* - e_{\Sigma}^{\text{bound}}). \]
where \( \sigma_v = A_p \Phi^* \) and \( e_{\Sigma}^{\text{bound}} = \Phi^* \delta e_{\Sigma} / \partial \dot{V} \) means the residual error of HJB. The sustained incentive condition can ensure the boundedness of \( e_{\Sigma}^{HJB} \), that is, \( ||e_{\Sigma}^{HJB}|| \leq e_{\Sigma}^{\text{bound}} \), where \( e_{\Sigma}^{M} \) is a positive constant.

Then, the optimal control \( \Phi^* \) becomes
\[ \Phi^* = -0.5 \sigma_v^{-1} \left( A_p A_v + \delta e_{\Sigma} / \partial \dot{V} \right) \]

Because \( A_p \) is an unknown variable, we define its estimation as \( \dot{A}_p \). Further, we use \( \dot{A}_p \) to replace \( A_p \), leading to
\[ \dot{J}_v = \dot{A}_p \tanh(K_v \dot{V}) \]
\[ \nabla J_v = \nabla J_v = A_p \dot{A}_p \]

Finally, we obtain
\[ \Phi^* = -0.5 \sigma_v^{-1} \left[ K_v - K_v \tanh^2(K_v \dot{V}) \right] \dot{A}_p \]
The HJB equation becomes
\[ e_v \dot{t} \equiv H_v (\dot{V}, \Phi^*, J_v) = \Sigma_v + Q_v (\dot{V}) + \delta e_{\Sigma} (\Phi^*) \]
with \( \Sigma_v = A_p \dot{A}_p \).

Define \( \dot{E}_r = e_v^2 / 2 \). We develop the following regulation law to minimize \( E_r \)
\[ \dot{A}_p = -\alpha_v \sigma_v \left( \sigma_v \dot{A}_p + \frac{\dot{E}_r}{2} \right) \]

Choosing Lyapunov function candidate
\[ L_{v,2} = \dot{A}_p^2 / (2 \alpha_v) + V^2 + 2 \dot{V}, \]
\[ \dot{L}_{v,2} \] is derived as
\[ \dot{L}_{v,2} = -\sigma_v J_v^2 - \frac{2 \alpha_v}{\sqrt{2 \alpha_v}} \sigma_v \dot{J}_v \sqrt{2 \alpha_v} e_{\Sigma}^{\text{bound}} + 2V \Phi^* \]
\[ + 2 \dot{V} \left[ -Q_v (\dot{V}) + \delta e_{\Sigma} (\Phi^*) \right] \]
Using \( \frac{2 \alpha_v}{\sqrt{2 \alpha_v}} \sigma_v \dot{J}_v \sqrt{2 \alpha_v} e_{\Sigma}^{\text{bound}} \leq \frac{2 \alpha_v}{\sqrt{2 \alpha_v}} \Sigma_v^2 + \frac{1}{2 \alpha_v} \left( e_{\Sigma}^{\text{bound}} \right)^2 \),
(34) becomes
\[ \dot{L}_{v,2} \leq -\left( \sigma_v J_v^2 - \frac{2 \alpha_v}{\sqrt{2 \alpha_v}} \sigma_v \dot{J}_v \sqrt{2 \alpha_v} e_{\Sigma}^{\text{bound}} \right) \]
\[ + [1 - 2 \dot{V} \lambda_{\text{min}} (Q_v)] V^2 + [1 - 2 \dot{V} \lambda_{\text{max}} (\delta e_{\Sigma})] (\Phi^*) \]
Combining (16) and (33), we have the total Lyapunov function
\[ L_v = L_{v,2} + L_{v,2} \]
We further have
\[ \dot{L_v} \leq -k_v \dot{V}^2 - \left( \sigma_v J_v^2 - \frac{2 \alpha_v}{\sqrt{2 \alpha_v}} \sigma_v \dot{J}_v \sqrt{2 \alpha_v} e_{\Sigma}^{\text{bound}} \right) \dot{V}^2 + [15 - 2 \dot{V} \lambda_{\text{min}} (\delta e_{\Sigma})] (\Phi^*)^2 \]
with \( \Psi_v = 0.5 + k_v \sigma_v^2 / l_v + 0.5 (e_v^2) + 0.5 (e_{\Sigma}^{\text{bound}}) / \alpha_v \).

Let
\[ \Gamma_v > \frac{1}{2 \lambda_{\text{max}} (Q_v)} + \frac{3}{4 \lambda_{\text{min}} (\delta e_{\Sigma})} \]
, \( k_v > 0 \) and \( \sigma_v^2 < 2 \). Then, (36) becomes
\[ L_v \leq -k_v \dot{V}^2 - \left( \sigma_v J_v^2 - \frac{2 \alpha_v}{\sqrt{2 \alpha_v}} \sigma_v \dot{J}_v \sqrt{2 \alpha_v} e_{\Sigma}^{\text{bound}} \right) \dot{V}^2 + [15 - 2 \dot{V} \lambda_{\text{min}} (\delta e_{\Sigma})] (\Phi^*)^2 \]

Remark 3. The sustained incentive condition leads to that
\[ \left| e_{\Sigma} / \partial \dot{V} \right| \leq \lambda_{\text{max}} (e_{\Sigma}^{\text{bound}}) \]
with a constant \( \lambda_{\text{max}} (e_{\Sigma}^{\text{bound}}) \). From (22) and (26), we know
\[ \left| \dot{\Phi}^* - \Phi^* \right| \leq 0.5 \sigma_v^{-1} \left( A_p \dot{A}_p - \delta e_{\Sigma} / \partial \dot{V} \right) \leq 0.5 \sigma_v^{-1} \left( A_p \dot{A}_p - \delta e_{\Sigma} / \partial \dot{V} \right) \]

B. Altitude controller design
In this subsystem, we will develop a fuzzy optimal controller for altitude subsystem (2)-(5) via back-stepping such that \( h \rightarrow h_{\text{ref}} \) and makes cost function (64) minimal.

Define \( \gamma_d = \arcsin \left( -k_v h / k_{\Sigma} \right) / \left[ h_{\dot{h}}/2 \right] \), where
\[ h = h_{\text{ref}} \] is the altitude tracking error, \( k_v \in \mathbb{R}^+ \) and \( k_{\Sigma} \in \mathbb{R}^+ \) are constants, and \( \gamma_d \) is the command of \( \gamma \). Then, we can conclude that \( \gamma \rightarrow 0 \) when \( t \rightarrow \infty \) if \( \gamma \rightarrow \gamma_d \) since
\[ \dot{h} = -k_v h^2 - k_{\Sigma} \left[ h_{\dot{h}}/2 \right] \leq 0 \] and \( \dot{h} = 0 \) only when \( h = 0 \). The subsequent design goal becomes \( \gamma \rightarrow \gamma_d \).

We formulate (3)-(5) as
\[ \dot{e}_v = x_v - \gamma_d \]
(39) with
\[ a_v = L_v \left( \gamma_v \right) + V \left( \gamma_v \right) \]
\[ a_v = 0 \]
Similarly, we use fuzzy system (8) to approximate unknown functions \( a_v \) and \( a_{\dot{v}} \).

\[ \left| a_v - \xi_v \left( V \right) \right| \leq \xi_v \]
\[ \left| a_{\dot{v}} - \xi_{\dot{v}} \left( V \right) \right| \leq \xi_{\dot{v}} \]

where \( \xi_v = [\xi_v^*, \xi_{\dot{v}}^*, \ldots, \xi_v^*] \in \mathbb{R}^n \) and \( \xi_{\dot{v}} = [\xi_{\dot{v}}^*, \xi_{\dot{v}}^*, \ldots, \xi_{\dot{v}}^*] \in \mathbb{R}^n \) are weight vectors; \( \xi_v^* \) and \( \xi_{\dot{v}}^* \) are the upper bounds of \( e_v \) and \( e_{\dot{v}} \), respectively.

Step 1. Define
\[ \dot{e}_v = x_v - \gamma_d \]
Utilizing (39), \( \dot{e}_v \) is given by
\[ \dot{e}_v = a_v + e_v + s_v + \delta \gamma - \gamma_d \]
(42) where \( \delta \gamma \) is a main virtual controller and \( \delta \gamma \) is an optimal virtual controller. \( e_v \) and \( s_v \) will be defined subsequently.

We define \( \delta \gamma \) as
\[ \delta \gamma = -k_v e_v - k_{\Sigma} \left[ e_v \delta \gamma / 2 \right] - e_v \phi \left( V \right) - \gamma_d \]
(43)
where \( k_{1j} \in \mathbb{R}^+ \) and \( k_{2j} \in \mathbb{R}^+ \) are design constants. \( \phi_j \) denotes the estimate of \( \phi_j = \| \zeta_j \| \). We select the following adaptive law for \( \phi_j \):

\[
\dot{\phi}_j = 0.5l_j e_j^2 \mathbf{w}^T(\mathbf{x}) \mathbf{w}(\mathbf{x}) - 2k_{2j} \phi_j
\]

(44)

where \( l_j \in \mathbb{R}^+ \) is a design parameter.

Define \( x_\theta^d \) as the estimate of \( \theta \), and introduce the following filter

\[
\tau_\theta x_\theta^d = \theta - x_\theta^d, \bar{\theta}(0) = x_\theta^d(0)
\]

(49)

where \( \tau_\theta \in \mathbb{R}^+ \) is a constant.

**Step 2.** Define

\[
e_\theta = x_0 - s_\theta^d \quad \text{and} \quad \epsilon_\theta \text{ is given by}
\]

\[
e_\theta = s_\theta^d - x_0 + \bar{\theta}, \epsilon_\theta = \epsilon_\theta^d(0)
\]

(47)

where \( e_\theta \) and \( s_\theta^d \) will be defined subsequently. \( \bar{\theta} \) is a main virtual controller and \( \epsilon_\theta \) is an optimal virtual controller.

Define \( \tilde{\theta}_j \) as

\[
\tilde{\theta}_j = -k_{3j} e_\theta - k_{3j} \int_0^t e_\theta dt - e_\theta + s_\theta^d
\]

(48)

where \( k_{3j} \in \mathbb{R}^+ \) and \( k_{3j} \in \mathbb{R}^+ \) are chosen constants.

To obtain the time derivative of \( \tilde{\theta}_j \), we give the following filter

\[
\tilde{\tau}_\theta \tilde{x}_\theta^d = \tilde{\theta}_j - x_\theta^d, \bar{\tilde{\theta}}(0) = x_\theta^d(0)
\]

(49)

where \( x_\theta^d \) is the estimate of \( \tilde{\theta}_j \).  

**Step 3.** Define

\[
e_\theta = x_0 - x_\theta^d
\]

(50)

Invoking (39), \( \epsilon_\theta \) is formulated as

\[
\dot{\epsilon}_\theta = a_0 + \delta_\theta - s_\theta^d = a_0 + \delta_\theta - \dot{s}_\theta^d
\]

(51)

where \( \delta_\theta \) is a main altitude controller and \( \dot{s}_\theta^d \) is an optimal altitude controller.

We design \( \delta_\theta^d \) as

\[
\delta_\theta^d = -k_{4j} e_\theta - k_{4j} \int_0^t e_\theta dt - 0.5e_\theta \phi_0 \mathbf{w}^T(\mathbf{x}) \mathbf{w}(\mathbf{x}) - e_\theta + s_\theta^d
\]

(52)

where \( k_{4j} \in \mathbb{R}^+ \) and \( k_{4j} \in \mathbb{R}^+ \) are constants. \( \phi_0 \) is the estimate of \( \phi_0 = \| \zeta_0 \| \), and its learning law is developed as

\[
\dot{\phi}_0 = 0.5l_0 e_\theta^2 \mathbf{w}^T(\mathbf{x}) \mathbf{w}(\mathbf{x}) - 2k_{0j} \phi_0
\]

(53)

with the constant \( l_0 \in \mathbb{R}^+ \).

Define

\[
s_\theta = \theta - \bar{\theta}, \quad s_\theta = x_\theta^d - \bar{\theta}
\]

According to (45), (49) and (54), we get

\[
\dot{s}_\theta = -s_\theta / \tau_\theta - \bar{\theta}
\]

(55)

The previous study [8] proves that there exist constants \( B^M \) and \( B^U \) such that \( -\bar{\theta} \leq B^M \) and \( -\bar{\theta} \leq B^U \).

We further define

\[
\tilde{\phi}_j = \phi_j - \phi_0, \tilde{\phi}_0 = \phi_0 - \phi_0
\]

(56)

Select Lyapunov function

\[
L_{\theta,j} = L_{\theta,j} + L_{\theta,j} + L_{\theta,j}
\]

(57)
where $A_{Z_k}$ is a matrix with unknown elements and $e_{Z_k}$ is the estimate error. It is derived from (67) that
\begin{equation}
\nabla L_{A_{Z_k}} = A_{Z_k}^* A_{Z_k} + \partial e_{Z_k}/\partial Z_k
\end{equation}
with $e_{Z_k} = \tanh(K Z_k Z_k)$. The total Lyapunov function is formulated as
\begin{equation}
L_{A_{Z_k}} = -\frac{1}{2} (k_{1} - \frac{1}{2} e_{Z_k} + \frac{1}{2} e_{Z_k} - (k_{1} - \frac{1}{2} e_{Z_k} + \frac{1}{2} e_{Z_k}^* e_{Z_k}^*) e_{Z_k}^* - \frac{1}{2} \frac{1}{2} B_{Z_k}^* B_{Z_k}^*) e_{Z_k}^* + \Psi_x
\end{equation}
\begin{equation}
+ [2 - 2 \Gamma_{\lambda} (\lambda_x)] ||Z_k|| + \frac{3}{2} 2 \Gamma_{\lambda} (\lambda_x (\lambda_x)) ||Z_k||
\end{equation}
with $\Psi_x = \Xi + (e_{Z_k}^*)^2 / (2 \alpha_{Z_k})$. Let $k_{1} > 0$, $\alpha_{Z_k} < 2$ and $\Gamma_x > \max \{1/\lambda_x (\lambda_x), 3/(4 \lambda_x (\lambda_x))\}$. Then, (81) becomes
\begin{equation}
L_{A_{Z_k}} \leq -\frac{1}{2} (k_{1} - \frac{1}{2} e_{Z_k} - \frac{1}{2} e_{Z_k} - (k_{1} - \frac{1}{2} e_{Z_k} - \frac{1}{2} e_{Z_k}^*) e_{Z_k}^*) e_{Z_k}^* - \frac{1}{2} \frac{1}{2} B_{Z_k}^* B_{Z_k}^*) e_{Z_k}^* + \Psi_x
\end{equation}
\begin{equation}
\leq [2 - 2 \Gamma_{\lambda} (\lambda_x)] ||Z_k|| + \frac{3}{2} 2 \Gamma_{\lambda} (\lambda_x (\lambda_x)) ||Z_k||
\end{equation}
with $\Psi_x = \Xi + (e_{Z_k}^*)^2 / (2 \alpha_{Z_k})$. We conclude that $e_{Z_k} \to \Omega_{Z_k}$, $e_{Z_i} \to \Omega_{Z_i}$, $s_0 \to \Omega_{s_0}$ and $\hat{A}_{Z_k} \to \Omega_{\hat{A}_{Z_k}}$ when $t \to \infty$, where
\begin{equation}
\Omega_{Z_k} = \{e_{Z_k} ||e_{Z_k}|| \leq \sqrt{\kappa} \lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x)))\}} - 0.5 \lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x)))\})
\end{equation}
\begin{equation}
\Omega_{s_0} = \{e_{s_0} ||e_{s_0}|| \leq \sqrt{\kappa} \lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x)))\}) - 0.5 \lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x)))\})
\end{equation}
\begin{equation}
\Omega_{\hat{A}_{Z_k}} = \{\hat{A}_{Z_k} \leq \sqrt{\kappa} \lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x)))\}) - 0.5 \lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x)))\})
\end{equation}
Remark 4. The sustained incentive condition guarantees that $||e_{Z_k}/(\lambda_x)|| \leq \Delta e_{Z_k,m}$ with the constant $\Delta e_{Z_k,m} \in \mathbb{R}$. Based on (70) and (73), we have $||e_{Z_k}/(\lambda_x)|| = 0.5 (r_k^*)^2 (A_{Z_k}^* A_{Z_k} - \partial e_{Z_k}/\partial Z_k) <= 0.5 (r_k^*)^2 (A_{Z_k}^* A_{Z_k} - (\lambda_x (\lambda_x (\lambda_x (\lambda_x (\lambda_x)))\})^2$, where $A_{Z_k}$ is the upper bound of $|| \hat{A}_{Z_k} ||$.

Remark 5. In (11) and (40), each fuzzy approximation only contains one online learning parameter. Undoubtedly, the computational load is low. Moreover, the developed critic network also has only one adaptive parameter. This guarantees satisfied real-time performance.

Remark 6. Traditionally, to ensure that the cost function is finite, the steady-state value of control input must be zero [30]-[32]. Unfortunately, most of actual systems don't satisfy this condition. In this paper, we decompose the controller into two parts namely the main controller and the optimal controller. Only the optimal controller ultimately converges to zero, while the total control input converges to a constant instead of zero.

Remark 7. In this paper, the vehicle dynamics is assumed to be unknown and it is approximated by a fuzzy system. Hence, the addressed method has the possibility of extending it to time-varying dynamic systems since the unknown vehicle dynamics can be time-varying or time-invariant.
IV. SIMULATION STUDY

The proposed controller is compared with a neural-approximation-based back-stepping control (NBC) [38] to validate its effectiveness and superiority. The values of model coefficients and parameters adopted in the simulation are referenced from [35]. The initial trim condition for HFVs is chosen as: \( V = 7700 \text{ ft/s}, h = 85000 \text{ ft}, \gamma = 0 \text{ deg}, \theta = 1.62 \text{ deg}, Q = 0 \text{ deg}/\text{s}^2, \eta_1 = 0.97 \) and \( \eta_2 = 0.80 \). Design parameters are chosen as: \( k_{v_1} = 10, k_{v_2} = 5.5, l_1 = 0.05, q_v = 5, r_r = 1, K_{\beta} = 0.2, \alpha_{\beta} = 0.5, k_{\beta_1} = 35, k_{\beta_2} = 1.5, k_{\gamma_1} = 1.9, k_{\gamma_2} = 0.2, l_2 = 0.05, r_\phi = 0.2, k_{\phi_1} = 25, k_{\phi_2} = 1.2, r_\theta = 0.2, k_{Q_1} = 45, k_{Q_2} = 15, l_Q = 0.05, \mathbf{K}_\mathbf{q} = [0.2, 0.2, 0.2], \mathbf{q}_\mathbf{q} = \text{diag}(5, 5, 5) \) and \( \tau_r = \text{diag}(1, 1, 1) \).

The obtained simulation results are presented in Figs. 1-8. Velocity tracking performance and altitude tracking performance, depicted in Figs. 1 and 2, reveal that the proposed controller can provide better tracking of reference commands in comparison with NBC (See Table I). Especially, the responses of \( \int_0^t |\dot{V}| \, dt \) and \( \int_0^t |\dot{h}| \, dt \) obviously show that the developed control approach’s tracking precision is higher compared with NBC (See Figs. 3 and 4). Moreover, it can be seen from Figs. 5 and 6 that the control inputs of both methods are smooth and their values are reasonable. Finally, all the critic weights are convergent, as shown in Figs. 7 and 8.

| Tracking error | NBC | Proposed method |
|----------------|-----|-----------------|
| \( \dot{V} \)  | \(-1 \text{ ft/s} < \dot{V} < 0.4 \text{ ft/s} \) | \(-1 \text{ ft/s} < \dot{V} < 0.35 \text{ ft/s} \) |
| \( \dot{h} \)  | \(-6 \text{ft} < \dot{h} < 10 \text{ft} \) | \(-2 \text{ft} < \dot{h} < 3 \text{ft} \) |

TABLE I

| Tracking error | NBC | Proposed method |
|----------------|-----|-----------------|
| \( \int_0^t |\dot{V}| \, dt \) | | |
| \( \int_0^t |\dot{h}| \, dt \) | | |

Fig. 1. Velocity tracking performance.

Fig. 2. Altitude tracking performance.

Fig. 3. The response of \( \int_0^t |\dot{V}| \, dt \).

Fig. 4. The response of \( \int_0^t |\dot{h}| \, dt \).
Fig. 5. The response of $\Phi$.

Fig. 6. The response of $\delta_e$.

Fig. 7. The responses of $|\dot{A}_p|$ and $|\dot{A}_g(1)|$.

Fig. 8. The responses of $|A_e(2)|$ and $|A_e(3)|$.

V. CONCLUSIONS

This article investigates an optimal tracking control problem of HFVs subject to unknown dynamics. The vehicle dynamics is consisted of the velocity subsystem and the altitude subsystem. Fuzzy approximations are applied to develop a robust tracking controller for the velocity subsystem via SACD. Furthermore, a back-stepping-based nearly optimal controller is devised for altitude dynamics. Advanced regulation algorithms are exploited for fuzzy approximations to construct a low-complexity control framework. Finally, the effectiveness and advantage of the addressed method are verified by simulation results.

REFERENCES

[1] J. Liu, H. An, Y. Gao, C. Wang and L. Wu, “Adaptive Control of Hypersonic Flight Vehicles With Limited Angle-of-Attack,” IEEE/ASME Transactions on Mechatronics, Vol. 23, no. 2, pp.883-894, April 2018.

[2] J. Sun, C. Li, Y. Guo, C. Wang and P. Li, “Adaptive fault tolerant control for hypersonic vehicle with input saturation and state constraints,” Acta Astronautica, vol. 167, pp.302-313, February 2020.

[3] X. Bu, Y. Xiao and H. Lei, “An Adaptive Critic Design-Based Fuzzy Neural Controller for Hypersonic Vehicles: Predefined Behavioral Nonaffine Control,” IEEE/ASME Transactions on Mechatronics, Vol. 24, no. 4, pp.1871-1881, Aug. 2019.

[4] J. Sun, Z. Pu, J. Yi and Z. Liu, “Fixed-Time Control with Uncertainty and Measurement Noise Suppression for Hypersonic Vehicles via Augmented Sliding Mode Observers,” IEEE Transactions on Industrial Informatics, vol. 16, no. 2, pp. 1192-1203, October 2019.

[5] Z. Guo, J. Guo, J. Zhou and J. Chang, “Robust Tracking for Hypersonic Reentry Vehicles via Disturbance Estimation-Triggered Control,” IEEE Transactions on Aerospace and Electronic Systems, vol. 56, no. 2, pp.1279-1289, April 2020.

[6] X. Yu, P. Li and Y. Zhang, “The Design of Fixed-Time Observer and Finite-Time Fault-Tolerant Control for Hypersonic Gliding Vehicles,” IEEE Transactions on Industrial Electronics, vol. 65, no. 5, pp. 4135-4144, May 2018.

[7] X. Shao and H. Wang, “Active disturbance rejection based trajectory linearization control for hypersonic reentry vehicle with bounded uncertainties,” ISA Transactions, vol. 54, pp. 27-38, 2015.
Xiangwei Bu was born in 1987. He received his B.S., M.S. and Ph.D. degrees from Air Force Engineering University in 2010, 2012 and 2016, respectively. He is currently an associate professor in Air and Missile Defense College, Air Force Engineering University.

His research interests include advanced hypersonic flight control. Dr. Bu is Supported by Young Talent Support Project for Military Science and Technology in 2018-2021. Dr. Bu is the expert of Chinese Association of Automation (CAA), and also is an Elsevier Chinese Highly Cited Scholar of aerospace engineering in 2019. Dr. Bu is an Editorial Board Member (EBM) of Advances in Mechanical Engineering.

Qiang Qi is currently pursuing his Master degree in Air and Missile Defense College, Air Force Engineering University. His research interests include advanced control theories with their applications to flight vehicles.