Difan Tang, Lei Chen, Zhao F. Tian and Eric Hu

Adaptive nonlinear optimal control for active suppression of airfoil flutter via a novel neural-network-based controller

Journal of Vibration and Control, 2018; 24(22):5261-5272

Published version available via DOI: http://dx.doi.org/10.1177/1077546317750504

Original Manuscript

Corresponding Author:
Lei Chen, School of Mechanical Engineering, University of Adelaide, Adelaide, SA 5005, Australia
Email: lei.chen@adelaide.edu.au

Adaptive nonlinear optimal control for active suppression of airfoil flutter via a novel neural-network-based controller

Difan Tang, Lei Chen, Zhao F. Tian and Eric Hu

School of Mechanical Engineering, University of Adelaide, Adelaide, SA 5005, Australia

Abstract

This paper proposes a novel adaptive nonlinear controller based on neural networks (NNs) for active suppression of airfoil flutter (ASAF) from the optimal control perspective. A new form of NN-based value function approximation (VFA) is proposed for solving the Hamilton-Jacobi-Bellman equation online. A systematic procedure based on linear matrix inequalities is further proposed for designing a scheduled parameter matrix to generalize the new VFA to globally nonlinear systems to suit ASAF applications. Closed-loop stability is examined through Lyapunov stability analysis. Comparisons drawn with a linear-parameter-varying optimal controller in wind-tunnel experiments confirms the effectiveness and validity of the proposed control scheme.

Keywords

Aeroelasticity, Adaptive Control, Nonlinear Control, Optimal Control, Neural Network

1. Introcution

Aeroelastic systems are subjected to various nonlinearities and are generally prone to the instability known as “limit-cycle oscillation (LCO)”, which can cause serious damage to the airfoil. For active suppression of airfoil flutter (ASAF), practically feasible solutions primarily include embedded piezoelectric actuation (Song and Li, 2014; Fazelzadeh et al., 2017) as well as proper deployment of the existing airfoil control surfaces. The latter is to be discussed in detail next.

In terms of control algorithms for suppressing LCOs various methods apply (Chen et al., 2009; Saaed et al., 2017; Keyser et al., 2017). Due to the time-varying nature and nonlinear characteristics of an aeroelastic system (Nayfeh et al., 2012; Bichiou et al., 2016; Vasconcellos et al., 2016) and the increasing demand on a wider operation range beyond the flutter boundary, advanced methods in adaptive, nonlinear, and robust control have received more attention in recent ASAF studies, although conventional frequency-domain analysis remains a useful tool for control synthesis (Schmidt, 2016). These advanced methods include but are not limited to: online linear-quadratic-regulator (Pak et al., 1995), optimal control synthesized via time-domain finite elements method (Fazelzadeh et al., 2014), self-tuning regulator (Viswamurthy and Ganguli, 2008), linear-parameter-varying techniques (Prime et al., 2010; Chen et al., 2012), feedback linearization (Strganac et al., 2000; Platanitis and Strganac, 2004), model reference adaptive control (Ko et al., 2002), back-stepping-based adaptive output feedback control (Singh and Wang, 2002), robust output feedback control (Zhang and Behal, 2016), modular adaptive control (Singh and Brenner, 2003; Rao et al., 2006), modified filtered-X least-mean-square control (Carnahan and Richards, 2008), L_1 adaptive control (Lee and Singh, 2013), sliding-mode control (Wang et al., 2015; Luo et al., 2016), finite-time H_∞ adaptive fault-tolerant control (Gao and Cai, 2016; Gao et al., 2016) and neural-network-based adaptive control (Gujjula et al., 2005; Wang et al., 2011; Brillante and Mannarino, 2016), etc.

However, optimal controllers among the aforementioned methods are susceptible to un-modeled dynamics. Though some other controllers are designed to be more adaptive to the changing environments and tolerant to un-modeled dynamics, these methods do not provide nonlinear optimal control. These two problems, although of significance to improving ASAF performance, have nevertheless not been addressed. This paper, thus, aims to reduce the impact of these problems by proposing an approach that synthesizes nonlinear optimal control in real time for ASAF according to online updated dynamics.

Optimal control for nonlinear systems involves iteratively solving a nonlinear Hamilton-Jacobi-Bellman (HJB) equation for a neural network (NN) based value function approximation (VFA) via a Critic-Actor configuration. A review on recent developments in this field until 2015 can be found in Tang et al. (2015), with some associated limitations revealed. According to Tang et al. (2015), to guarantee closed-loop stability for unstable and marginally stable systems, the existing methods require an initial stabilizing control, a logic switching mechanism for the Critic NN, or an additional tuning loop for the Actor NN. These increase complexity, and can introduce more uncertainties into the system. Moreover, the logic-switch mechanism can also cause discontinuities in control. Despite advances since 2015 (Feng et al., 2015; Jiang and Jiang, 2015; Kiumarsi et al., 2015; Zhao et al., 2017; Cui et al., 2017), the limitations remain unsolved.

Besides the limitations discussed, the existing methods for online synthesis of nonlinear optimal control are also confined to locally nonlinear systems, which are a sub-class of globally nonlinear scenarios. Aeroelastic systems are nonlinear at a constant

freestream airspeed (i.e. locally nonlinear), and the dynamics also vary nonlinearly with the airspeed (i.e. globally nonlinear). This makes all these existing methods inapplicable to ASAF without modification and improvement.

Therefore, the study in this paper focuses on solving the aforementioned problems and delivers the following contributions:

(i) A new form of NN-based value function approximation (VFA) is proposed to eliminate the need for an initial stabilizing control, a logic-switch mechanism for the Critic NN, or a tuning loop for the Actor NN. This contributes to a compact NN controller configuration suitable for real-time implementation without jeopardizing closed-loop stability.
(ii) A systematic procedure based on linear matrix inequalities is further purposed for the design of a scheduled parameter matrix for the new form of NN-based VFA to generalize the proposed method to globally nonlinear cases, so that the proposed NN controller suits ASAF applications.
(iii) The proposed method successfully solves the dilemma in ASAF controller design, and wind-tunnel experiments were conducted to validate the proposed algorithm. To the best of our knowledge, it is the first experimentally validated approach in this regard.

2. Aeroelastic system

A typical rigid airfoil section featuring two-dimensional vibration modes (i.e. the first plunge/pitch-mode oscillations) is considered for its well-established theory basis and experimental validation (O'Neil and Strganac, 1998; Strganac et al., 2000; Ko et al., 2002; Platanitis and Strganac, 2004; Prime et al., 2010; Prime, 2010). Leading- and trailing-edge control surfaces are used to actively suppress flutter. Specifically in terms of the analytical model for control synthesis, a four-degrees-of-freedom (4-DOF) aeroelastic system as in Prime (2010) is considered, which includes not only the plunge and pitch DOFs but also the deflection angle of the leading- and trailing-edge control surfaces as another two DOFs. It models the lift and moment that act on the airfoil elastic axis using quasi-steady aerodynamics (Fung, 1955; Strganac et al., 2000), describes the coupled dynamics of the plunge and pitch DOFs, incorporates the inertial coupling of the leading- and trailing-edge control surfaces to the airfoil rigid-body dynamics, and also takes into account servo motors dynamics to capture control delay. This 4-DOF aeroelastic model is derived by the Lagrangian energy method, verified with a different modeling technique -- the Newton-Euler iteration, and validated in wind-tunnel experiments using the Nonlinear Aeroelastic Test Apparatus (NATA) at Texas A&M University. The model was proven accurate for low Strouhal numbers, which primarily accounts for cases in subsonic flow conditions. Nonlinear translational and torsional stiffness is introduced in a polynomial form up to second order, and all the trigonometric terms are retained. A schematic illustration of the 4-DOF aeroelastic system is shown in Figure 1.

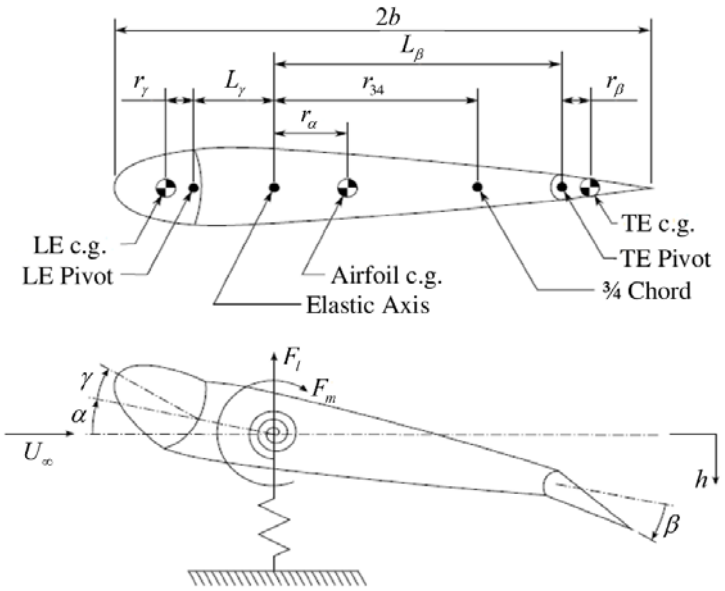

Figure 1. Schematic figure of the 4-DOF aeroelastic system ("LE": leading edge; "TE": trailing edge; "c.g." center of gravity)

The equation of motion of the 4-DOF aeroelastic system is:

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{E} = \mathbf{F}, \tag{1}$$

with

$$\ddot{\mathbf{q}}(t) = \left[\ddot{h}(t) \quad \ddot{\alpha}(t) \quad \ddot{\beta}(t) \quad \ddot{\gamma}(t)\right]^{\mathrm{T}},$$

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & I_\beta & 0 \\ M_{41} & M_{42} & 0 & I_\gamma \end{bmatrix},$$

$$\mathbf{E} = \left[E_1 \quad E_2 \quad E_3 \quad E_4\right]^{\mathrm{T}},$$

and

$$\mathbf{F} = \left[-F_l \quad F_m \quad k_\beta \beta_{cmd} \quad k_\gamma \gamma_{cmd}\right]^{\mathrm{T}},$$

where

$$M_{11} = m_\alpha + m_\beta + m_\gamma,$$

$$\begin{aligned} M_{12} = M_{21} = \; & (m_\alpha r_\alpha + m_\beta L_\beta + m_\gamma L_\gamma)\cos(\alpha) \\ & + m_\beta r_\beta \cos(\alpha+\beta) + m_\gamma r_\gamma \cos(\alpha+\gamma), \end{aligned}$$

$$M_{13} = M_{31} = m_\beta r_\beta \cos(\alpha+\beta),$$

$$M_{14} = M_{41} = m_\gamma r_\gamma cos(\alpha + \gamma),$$

$$\begin{aligned} M_{22} = \; & I_\alpha + I_\beta + I_\gamma + m_\beta L_\beta^2 + m_\gamma L_\gamma^2 \\ & + 2L_\beta m_\beta r_\beta cos(\beta) + 2L_\gamma m_\gamma r_\gamma cos(\gamma), \end{aligned}$$

$$M_{23} = M_{32} = I_\beta + L_\beta m_\beta r_\beta cos(\beta),$$

$$M_{24} = M_{42} = I_\gamma + L_\gamma m_\gamma r_\gamma cos(\gamma),$$

$$\begin{aligned} E_1 = \; & k_h h + c_h \dot{h} - (\dot{\alpha} + \dot{\beta})^2 m_\beta r_\beta sin(\alpha + \beta) \\ & - (\dot{\alpha} + \dot{\gamma})^2 m_\gamma r_\gamma sin(\alpha + \gamma) \\ & - \dot{\alpha}^2 sin(\alpha)(m_\alpha r_\alpha + m_\beta L_\beta + m_\gamma L_\gamma), \end{aligned}$$

$$\begin{aligned} E_2 = \; & k_\alpha \alpha + c_\alpha \dot{\alpha} - \dot{\beta}(\dot{\beta} + 2\dot{\alpha}) m_\beta r_\beta L_\beta sin(\beta) \\ & - \dot{\gamma}(\dot{\gamma} + 2\dot{\alpha}) m_\gamma r_\gamma L_\gamma sin(\gamma), \end{aligned}$$

$$E_3 = k_\beta \beta + c_\beta \dot{\beta} + \dot{\alpha}^2 m_\beta r_\beta L_\beta sin(\beta),$$

$$E_4 = k_\gamma \gamma + c_\gamma \dot{\gamma} + \dot{\alpha}^2 m_\gamma r_\gamma L_\gamma sin(\gamma),$$

$$\begin{aligned} F_l = \; & \rho U_\infty^2 b S C_{l\alpha} (\alpha + \frac{\dot{h}}{U_\infty} + r_{34} \frac{\dot{\alpha}}{U_\infty}) \\ & + \rho U_\infty^2 b S C_{l\beta} \beta + \rho U_\infty^2 b S C_{l\gamma} \gamma, \end{aligned}$$

$$\begin{aligned} F_m = \; & \rho U_\infty^2 b^2 S C_{m_{\alpha e}} (\alpha + \frac{\dot{h}}{U_\infty} + r_{34} \frac{\dot{\alpha}}{U_\infty}) \\ & + \rho U_\infty^2 b^2 S C_{m_{\beta e}} \beta + \rho U_\infty^2 b^2 S C_{m_{\gamma e}} \gamma, \end{aligned}$$

where m_α, m_β and m_γ are mass of the airfoil (without control surfaces), trailing and leading edges; I_α, I_β and I_γ are moment inertia of the wing without control surfaces (about the elastic axis), and trailing/leading edges (about their respective pivot); k_h, k_α, k_β, k_γ, c_h, c_α, c_β and c_γ are the stiffness and damping coefficients associated with each DOF; ρ is the air density; S is the airfoil planform area; $C_{l\alpha} = \partial C_l / \partial \alpha$, $C_{l\beta} = \partial C_l / \partial \beta$ and $C_{l\gamma} = \partial C_l / \partial \gamma$, with C_l being the lift coefficient of the airfoil section; $C_{m_{\alpha e}} = \frac{r_{34}}{b} C_{l\alpha} + 2\frac{\partial C_m}{\partial \alpha}$, $C_{m_{\beta e}} = \frac{r_{34}}{b} C_{l\beta} + 2\frac{\partial C_m}{\partial \beta}$ and $C_{m_{\gamma e}} = \frac{r_{34}}{b} C_{l\gamma} + 2\frac{\partial C_m}{\partial \gamma}$, with C_m being moment coefficient of the airfoil section at quarter-chord.

The aeroelastic system can be mathematically expressed in a state-space control-affine form as:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), U_\infty) + \mathbf{g}(\mathbf{x}(t))\mathbf{u}(t), \tag{2}$$

where $\mathbf{x}(t) \in \mathbb{R}^{n_x}$ denotes n_x system states; $\mathbf{u}(t) \in \mathbb{R}^{n_u}$ refers to n_u control inputs; $\mathbf{f}(\mathbf{x}(t), U_\infty) \in \mathbb{R}^{n_x}$ describes system internal dynamics that are dependent on the freestream airflow velocity U_∞; $\mathbf{g}(\mathbf{x}(t)) \in \mathbb{R}^{n_x \times n_u}$ represents control input dynamics.

Specifically,

$$\mathbf{x}(t) = [h \quad \alpha \quad \beta \quad \gamma \quad \dot{h} \quad \dot{\alpha} \quad \dot{\beta} \quad \dot{\gamma}]^{\mathrm{T}},$$

$$\mathbf{u}(t) = [\beta_{cmd} \quad \gamma_{cmd}]^{\mathrm{T}},$$

$$\mathbf{f}(\mathbf{x}(t), U_\infty) = \left[\dot{\mathbf{q}}^{\mathrm{T}} \quad (\mathbf{M}^{-1}\bar{\mathbf{F}})^{\mathrm{T}}\right]^{\mathrm{T}},$$

and

$$\mathbf{g}(\mathbf{x}(t)) = \left[\mathbf{0} \quad (\mathbf{M}^{-1}\bar{\mathbf{g}})^{\mathrm{T}}\right]^{\mathrm{T}},$$

where

$$\bar{\mathbf{F}} = \left[F_l - E_1 \quad F_m - E_2 \quad -E_3 \quad -E_4\right]^{\mathrm{T}},$$

and

$$\bar{\mathbf{g}} = \begin{bmatrix} 0 & 0 & k_\beta & 0 \\ 0 & 0 & 0 & k_\gamma \end{bmatrix}^{\mathrm{T}}.$$

3. Proposed controller

3.1 Continuous-time HJB equation

For a constant velocity U_∞, equation (2) is reduced to:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t))\big|_{U_\infty = U} + \mathbf{g}(\mathbf{x}(t))\mathbf{u}(t); \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{3}$$

which can be written in a compact form:

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), \mathbf{u}(t))\big|_{U_\infty = U}; \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{4}$$

where $U \in \mathbb{R}^{+}$ is any valid value of U_∞.

For convenience in discussion, the dynamics associated with a constant velocity U is hereafter written in a simpler form by omitting the notation of $U_\infty = U$.

Assumption 1: *The initial state of the system internal dynamics* $\mathbf{f}(0) = 0$ *; The system as in (4) is Lipschitz continuous on a set* $\boldsymbol{\Omega} \subseteq \mathbb{R}^{n_x}$ *with the origin included, and can be stabilized by* $\mathbf{u} \in \boldsymbol{\Psi}(\boldsymbol{\Omega})$ *that are admissible.*

Assumption 2: $\|\mathbf{f}(\mathbf{x}(t))\| \le b_f \|\mathbf{x}(t)\|$ *and* $\|\mathbf{g}(\mathbf{x}(t))\| \le b_g$, *where constants* $b_f \in ¡^{+}$ *and* $b_g \in ¡^{+}$.

The control problem is to determine a control law $\mathbf{u}(t)$ to minimize the following performance index (cost function):

$$V(\mathbf{x}_0) = \int_0^\infty [\bar{Q}(\mathbf{x}(\tau)) + \bar{R}(\mathbf{u}(\tau))]d\tau, \tag{5}$$

with $\bar{Q}(\mathbf{x}(t))$ and $\bar{R}(\mathbf{u}(t)) = \mathbf{u}^{\mathrm{T}}(t)\mathbf{R}\mathbf{u}(t)$ being positive-definite functions, in which $\mathbf{R} \in ¡^{n_u \times n_u}$ is a positive-definite weighting matrix.

Differentiating (5) yields its infinitesimal version that is a nonlinear Lyapunov equation (LE) (Abu-Khalaf and Lewis, 2005), written as:

$$\mathbf{V}_x^{\mathrm{T}}(\mathbf{x})\mathbf{F}(\mathbf{x},\mathbf{u}) + \bar{Q}(\mathbf{x}) + \bar{R}(\mathbf{u}) = 0; \quad V(0) = 0. \tag{6}$$

Let $V^*(\mathbf{x})$ denote the optimal (minimal) cost function, named as the "value function", and let $\mathbf{V}_x^*(\mathbf{x}) @\partial V^*(\mathbf{x}) / \partial \mathbf{x}$. Then the corresponding optimal control is:

$$\mathbf{u}^*(\mathbf{x}) = -\frac{1}{2}\mathbf{R}^{-1}\mathbf{g}^{\mathrm{T}}(\mathbf{x})\mathbf{V}_x^*(\mathbf{x}), \tag{7}$$

which satisfies the following Hamilton-Jacobi-Bellman (HJB) equation based on (6):

$$-\frac{1}{4}\left[\mathbf{V}_x^*(\mathbf{x})\right]^{\mathrm{T}} \mathbf{g}(\mathbf{x})\mathbf{R}^{-1}\mathbf{g}^{\mathrm{T}}(\mathbf{x})\mathbf{V}_x^*(\mathbf{x}) + \bar{Q}(\mathbf{x}) + \left[\mathbf{V}_x^*(\mathbf{x})\right]^{\mathrm{T}} \mathbf{f}(\mathbf{x}) = 0; \quad V^*(0) = 0. \tag{8}$$

By solving (8) for $V^*(\mathbf{x})$, the optimal control law can then be obtained as in (7).

3.2 NN-based value function approximation

The HJB equation is nonlinear and solving for $V^*(\mathbf{x})$ in a direct way is difficult. Instead, it can be solved recursively through a policy-iteration approach. To implement policy iteration, an appropriately structured representation of $V^*(\mathbf{x})$ is necessary, which can be a neural-network (NN) approximation. Different from existing methods, we propose a new form of value function approximation (VFA) as:

$$V^*(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{x} + \mathbf{W}^{\mathrm{T}}\mathbf{\Phi}(\mathbf{x}) + \varepsilon(\mathbf{x}), \tag{9}$$

where $\mathbf{\Phi}(\cdot) = [\phi_1(\mathbf{x}), \mathrm{L}\ , \phi_N(\mathbf{x})]^{\mathrm{T}} : ¡^{n} \to ¡^{N}$ contains N hidden-layer neurons, each of which is a nonlinear activation function; $\mathbf{W} \in ¡^{N}$ is a vector of ideal NN weights; $\mathbf{P} \in ¡^{n_x \times n_x}$ is a diagonal positive-definite matrix; $\varepsilon(\mathbf{x}) \in ¡$ is the approximation error.

Remark 1: *The new NN-based formula proposed for VFA in this study is different from all the existing methods in that it has an auxiliary term featuring a quadratic function of* $\mathbf{x}$ *in addition to the commonly adopted standard structure of a double-layer NN. This new NN-based VFA allows individual control of the coefficients for the second-order regressors in a determinate manner. That is* $\mathbf{W}$ *in the second part of (9) is to be determined via an online tuning algorithm, while* $\mathbf{P}$ *is a pre-selected parameter matrix not affected by online tuning. The advantages of using this new NN-based VFA are given in Theorem 1 in details.*

Remark 2: *The NN used in (9) is linear in its weights while the activation functions are of explicit nonlinear forms with respect to the NN inputs. This type of NN is generally known as a linear-in-the-parameters (LIP) NN (Lewis et al., 2003). The linearity in weights facilitates tuning while the activation functions can provide firm approximation results for specified accuracy, given that the functions are selected as a suitable basis. A systematic way to nominate the activation functions utilizes the high-order Weierstrass approximation theorem (Finlayson, 1972), which renders a power series of the NN inputs up to the specified order. A simpler formulation can take the form with the centers of the series being zero.*

In regard to (9), the derivative of $V^*(\mathbf{x})$ with respect to $\mathbf{x}$ is:

$$\mathbf{V}_x^*(\mathbf{x}) \triangleq \frac{\partial V^*(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{P}\mathbf{x} + \nabla\mathbf{\Phi}^{\mathrm{T}}(\mathbf{x})\mathbf{W} + \nabla\boldsymbol{\varepsilon}^{\mathrm{T}}(\mathbf{x}), \tag{10}$$

where $\nabla\mathbf{\Phi}(\mathbf{x}) \triangleq \left[\frac{\partial\mathbf{\Phi}(\mathbf{x})}{\partial\mathbf{x}}\right]^{\mathrm{T}}$ and $\nabla\varepsilon(\mathbf{x}) \triangleq \left[\frac{\partial\varepsilon(\mathbf{x})}{\partial\mathbf{x}}\right]^{\mathrm{T}}$.

Assumption 3: *There exist constants* $b_\phi \in \mathbb{R}^+$ *and* $b_\varepsilon \in \mathbb{R}^+$ *such that the inequalities* $\|\nabla\mathbf{\Phi}(\mathbf{x})\| \le b_\phi\|\mathbf{x}\|$ *and* $\|\nabla\varepsilon(\mathbf{x})\| \le b_\varepsilon\|\mathbf{x}\|$ *hold.*

The associated optimal control law under the proposed approximation scheme is:

$$\mathbf{u}(\mathbf{x}) = -\frac{1}{2}\mathbf{R}^{-1}\mathbf{g}^{\mathrm{T}}(\mathbf{x})\left[\mathbf{P}\mathbf{x} + \nabla\mathbf{\Phi}^{\mathrm{T}}(\mathbf{x})\mathbf{W}\right]. \tag{11}$$

This results in:

$$\left[\mathbf{x}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}} + \mathbf{W}^{\mathrm{T}}\nabla\mathbf{\Phi}(\mathbf{x})\right][\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}(\mathbf{x})] + \bar{Q}(\mathbf{x}) + \bar{R}(\mathbf{u}) = \varepsilon_H, \tag{12}$$

where ε_H is the difference caused by the approximation error $\varepsilon(\mathbf{x})$ in (9).

Let $\hat{\mathbf{W}}$ be the estimate of the ideal weights. In this case,

$$\hat{V}(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{P}\mathbf{x} + \hat{\mathbf{W}}^{\mathrm{T}}\mathbf{\Phi}(\mathbf{x}), \tag{13}$$

and

$$\hat{\mathbf{u}}(\mathbf{x}) = -\frac{1}{2}\mathbf{R}^{-1}\mathbf{g}^{\mathrm{T}}(\mathbf{x})\left[\mathbf{P}\mathbf{x} + \nabla\boldsymbol{\Phi}^{\mathrm{T}}(\mathbf{x})\hat{\mathbf{W}}\right]. \tag{14}$$

The resulting nonlinear LE then becomes:

$$\left[\mathbf{x}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}} + \hat{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}(\mathbf{x})\right][\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\hat{\mathbf{u}}(\mathbf{x})] + \bar{Q}(\mathbf{x}) + \bar{R}(\hat{\mathbf{u}}) = \varepsilon_H + \xi, \tag{15}$$

where ξ is the error of weights estimation during a tuning process.

The proposed algorithm is based on the synchronous policy-iteration framework (Vamvoudakis and Lewis, 2010) and basically involves two steps -- policy evaluation (using (13) and (15) as a "Critic") and policy improvement (using (14) as an "Actor"). Starting with $\hat{\mathbf{u}}(\mathbf{x})\big|_{\mathbf{x}=\mathbf{x}_0}$, the algorithm proceeds by minimizing ξ at every infinitesimal time step until convergence is reached at $V^*(\mathbf{x})$ and $\mathbf{u}^*(\mathbf{x})$ or their close proximity.

For uniform convergence of $\hat{\mathbf{W}}$ to the ideal $\mathbf{W}$ so that ξ is minimized, an extended Kalman filter (EKF) is used, and (15) can be rearranged as:

$$\begin{cases} \dot{\hat{\mathbf{W}}} = \mathbf{0} + \mathbf{w} \\ y = h(\mathbf{x}, \hat{\mathbf{W}}) - \varepsilon_H - \xi + v \end{cases}, \tag{16}$$

with

$$y = -\bar{Q}(\mathbf{x}) - \bar{R}(\hat{\mathbf{u}}),$$

and

$$h(\mathbf{x}, \hat{\mathbf{W}}) = \left[\mathbf{x}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}} + \hat{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}(\mathbf{x})\right]\left[\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\hat{\mathbf{u}}(\mathbf{x})\right],$$

where $\mathbf{0}$ is a null matrix, $\mathbf{w}$ and v are artificial white-noise inputs with covariance matrix $\mathbf{Q}_f$ f 0 and $\mathbf{R}_f$ f 0 , respectively.

Introducing an EKF into the system described by (16) yields:

$$\begin{cases} \dot{\hat{\mathbf{W}}} = \mathbf{K}_f(y - \hat{y}) \\ \hat{y} = h(\mathbf{x}, \hat{\mathbf{W}}) \end{cases}, \tag{17}$$

where $\hat{y}$ denotes the estimated output and $\mathbf{K}_f \in ¡^{N\times 1}$ is the EKF gain.

The EKF gain $\mathbf{K}_f$ can be computed from:

$$\mathbf{K}_f = \mathbf{S}\mathbf{H}^{\mathrm{T}}\mathbf{R}_f^{-1}, \tag{18}$$

with

$$\mathbf{H}^{\mathrm{T}} = \frac{\partial h(\mathbf{x}, \hat{\mathbf{W}})}{\partial \hat{\mathbf{W}}} = \nabla \mathbf{\Phi}(\mathbf{x})[\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\hat{\mathbf{u}}(\mathbf{x})], \tag{19}$$

and

$$\dot{\mathbf{S}} = \mathbf{Q}_f - \mathbf{S}\mathbf{H}^{\mathrm{T}}\mathbf{R}_f^{-1}\mathbf{H}\mathbf{S}, \tag{20}$$

where $\mathbf{H} \in \mathbb{R}^{N\times 1}$ is defined as in (19) and $\mathbf{S} \in \mathbb{R}^{N\times N}$ is a symmetrical positive-definite matrix with initial state $\mathbf{S}(0) = \mathbf{S}(0)^{\mathrm{T}} > 0$.

Remark 3: $\tilde{\mathbf{W}} = \mathbf{W} - \hat{\mathbf{W}}$ *is shown to be bounded (*$\|\tilde{\mathbf{W}}\| \le b_{\tilde{W}}$*) regardless of the VFA forms and tuning methods (Vamvoudakis and Lewis, 2010; Tang et al., 2015). However, this boundedness does not guarantee closed-loop stability during transient response in tuning (Tang et al., 2015), which necessitates a separate analysis on closed-loop stability.*

Theorem 1: *Given Assumptions 1 to 3 and the EFK estimation scheme provided by (17) to (20), there exists a scalar* $m_P \in \mathbb{R}^{+}$ *and a matrix* $\mathbf{P}$ *for* (9) *such that* $\|\mathbf{P}\| > m_P$ *and the nonlinear system as in (4) remains asymptotically stable during online tuning with the control law given by (14).*

Proof: Consider $L_V = \hat{V}$ as the Lyapunov function candidate. Its time derivative is:

$$
\begin{aligned}
\dot{L}_V &= \left(\frac{\partial \hat{V}}{\partial \mathbf{x}}\right)^{\mathrm{T}} (\mathbf{f}+\mathbf{g}\hat{\mathbf{u}}) \\
&= (\mathbf{P}\mathbf{x}+\nabla\boldsymbol{\Phi}^{\mathrm{T}}\hat{\mathbf{W}})^{\mathrm{T}}(\mathbf{f}+\mathbf{g}\hat{\mathbf{u}}) \\
&= (\mathbf{P}\mathbf{x}+\nabla\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{W}-\nabla\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\mathbf{W}})^{\mathrm{T}}(\mathbf{f}+\mathbf{g}\mathbf{u}^{*}-\mathbf{g}\tilde{\mathbf{u}}) \\
&= \left(\frac{\partial V^{*}}{\partial \mathbf{x}}\right)^{\mathrm{T}} (\mathbf{f}+\mathbf{g}\mathbf{u}^{*}) - \nabla\varepsilon(\mathbf{f}+\mathbf{g}\mathbf{u}^{*}) \\
&\quad + \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\boldsymbol{\Gamma}(\nabla\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\mathbf{W}}+\nabla\varepsilon^{\mathrm{T}}) \\
&\quad + \frac{1}{2}\mathbf{W}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\boldsymbol{\Gamma}(\nabla\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\mathbf{W}}+\nabla\varepsilon^{\mathrm{T}}) - \tilde{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\mathbf{f} \\
&\quad + \frac{1}{2}\tilde{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\boldsymbol{\Gamma}(\mathbf{P}\mathbf{x}+\nabla\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{W}+\nabla\varepsilon^{\mathrm{T}}) \\
&\quad - \frac{1}{2}\tilde{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\boldsymbol{\Gamma}(\nabla\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\mathbf{W}}+\nabla\varepsilon^{\mathrm{T}}) \\
&= -\bar{Q}-\mathbf{u}^{*\mathrm{T}}\mathbf{R}\mathbf{u}^{*}-\nabla\varepsilon\mathbf{f}-\tilde{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\mathbf{f}+\nabla\varepsilon\boldsymbol{\Gamma}\mathbf{P}\mathbf{x} \\
&\quad + \frac{1}{2}\nabla\varepsilon\boldsymbol{\Gamma}\nabla\varepsilon^{\mathrm{T}}+\nabla\varepsilon\boldsymbol{\Gamma}\nabla\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{W}+\tilde{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\boldsymbol{\Gamma}\mathbf{P}\mathbf{x} \\
&\quad + \tilde{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\boldsymbol{\Gamma}\nabla\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{W}-\frac{1}{2}\tilde{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\boldsymbol{\Gamma}\nabla\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\mathbf{W}} \\
&= -\bar{Q}-\frac{1}{4}\mathbf{x}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\boldsymbol{\Gamma}\mathbf{P}\mathbf{x}-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\boldsymbol{\Gamma}\nabla\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{W}-\nabla\varepsilon\mathbf{f} \\
&\quad + \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\boldsymbol{\Gamma}\nabla\varepsilon^{\mathrm{T}}-\frac{1}{4}\mathbf{W}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\boldsymbol{\Gamma}\nabla\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{W}+\frac{1}{4}\nabla\varepsilon\boldsymbol{\Gamma}\nabla\varepsilon^{\mathrm{T}} \\
&\quad + \frac{1}{2}\mathbf{W}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\boldsymbol{\Gamma}\nabla\varepsilon^{\mathrm{T}}-\tilde{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\mathbf{f}+\tilde{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\boldsymbol{\Gamma}\mathbf{P}\mathbf{x} \\
&\quad + \tilde{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\boldsymbol{\Gamma}\nabla\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{W}-\frac{1}{2}\tilde{\mathbf{W}}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\boldsymbol{\Gamma}\nabla\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\mathbf{W}}.
\end{aligned}
\tag{33}
$$

As $\bar{Q}(\mathbf{x})>0$, we have $m_Q\|\mathbf{x}\|^2 \le \bar{Q}(x) \le b_q\|\mathbf{x}\|^2$ for constants $m_Q, b_Q \in \mathbb{R}^{+}$. Based on Assumptions 2 and 3, there exist constants $m_G, b_G, m_{DW}, b_{DW} \in \mathbb{R}^{+}$ such that

$$m_G \le \left\|\mathbf{g}\mathbf{R}^{-1}\mathbf{g}^{\mathrm{T}}\right\| \le b_G,$$

and

$$m_{DW}\|\mathbf{x}\|^2 \le \left\|\mathbf{W}^{\mathrm{T}}\nabla\boldsymbol{\Phi}\mathbf{g}\mathbf{R}^{-1}\mathbf{g}^{\mathrm{T}}\nabla\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{W}\right\| \le b_{DW}\|\mathbf{x}\|^2.$$

Let $\|\mathbf{W}\| = n_W$. Since $0 \le \|\tilde{\mathbf{W}}\| \le b_{\tilde{W}}$, then (33) can be upper bounded as:

$$
\begin{aligned}
\dot{L}_V \le & -m_Q \|\mathbf{x}\|^2 - \frac{1}{4} m_G \|\mathbf{P}\|^2 \|\mathbf{x}\|^2 + \frac{1}{2} b_G b_\varepsilon \|\mathbf{P}\| \|\mathbf{x}\|^2 \\
& + \frac{1}{2} b_G b_\phi n_W \|\mathbf{P}\| \|\mathbf{x}\|^2 - \frac{1}{4} m_{DW} \|\mathbf{x}\|^2 + b_\varepsilon b_f \|\mathbf{x}\|^2 \\
& + \frac{1}{2} b_\phi b_G b_\varepsilon n_W \|\mathbf{x}\|^2 + \frac{1}{4} b_\varepsilon^2 b_G \|\mathbf{x}\|^2 + b_\phi b_f b_{\tilde{W}} \|\mathbf{x}\|^2 \\
& + b_\phi b_G b_{\tilde{W}} \|\mathbf{P}\| \|\mathbf{x}\|^2 + b_\phi^2 b_G n_W b_{\tilde{W}} \|\mathbf{x}\|^2 \\
= & -\|\mathbf{x}\|^2 \left(\frac{1}{4} m_G \|\mathbf{P}\|^2 - c_{p1} \|\mathbf{P}\| - c_{p2} \right).
\end{aligned}
\tag{34}
$$

where

$$
c_{p1} = \frac{1}{2} b_G b_\phi n_W + \frac{1}{2} b_G b_\varepsilon + b_\phi b_G b_{\tilde{W}},
$$

and

$$
\begin{aligned}
c_{p2} = & \frac{1}{2} b_\phi b_G b_\varepsilon n_W + \frac{1}{4} b_\varepsilon^2 b_G + b_\varepsilon b_f + b_\phi b_f b_{\tilde{W}} \\
& + b_\phi^2 b_G n_W b_{\tilde{W}} - m_Q - \frac{1}{4} m_{DW}.
\end{aligned}
\tag{35}
$$

Equation (34) shows that $\dot{L}_V$ is negative as long as

$$
\|\mathbf{P}\| \ge \frac{2c_{p1} + 2\sqrt{c_{p1}^2 + m_G c_{p2}}}{m_G} \triangleq m_P. \tag{36}
$$

It can be easily derived that the second-order derivative of L_V with respect to time is a function of $\mathbf{x}$ and $\tilde{W}$, i.e., $\ddot{L}_V = Z(\mathbf{x}, \tilde{W})$. Since $\mathbf{x}$ and $\tilde{W}$ are both bounded, $\ddot{L}_V$ is also bounded. Therefore, it can be concluded that the system states $\mathbf{x}$ are asymptotically stable. This completes the proof.

Remark 4: *Under the new form of NN-based VFA, this proposed control scheme maintains the stability of the closed-loop system during online tuning without the necessity of providing an initial stabilizing control, adding a stabilizing logic-switch mechanism to the Critic NN, or adding an additional stabilizing tuning loop to the Actor NN. This is one of the major differences of our approach from other existing methods. Moreover, to the best of our knowledge, the proposed algorithm is the first among the available methods to provide proven asymptotic stability rather than the relatively weaker UUB stability to the states of nonlinear systems as in (3) during online tuning. As shown by Theorem 1, the proposed new form of VFA shows significant advantages over the existing methods, and hence builds an important contribution of this paper.*

3.3 Generalization of the new form of NN-based value function approximation}

Note that the discussion in Subsections 3.1 and 3.2 are confined to locally nonlinear systems, with the parameter matrix $\mathbf{P}$ being constant, which is not suitable for a wider flight envelop with varying traveling speed U_∞ beyond the flutter boundary. As a second contribution in this paper, a systematic approach is proposed in the following for the selection of $\mathbf{P}$ to cope with U_∞ dependent dynamics as in (2), generalizing the new NN-based VFA to globally nonlinear cases so that it suits ASAF applications.

Linearizing (2) about $\mathbf{x} = 0$ gives:

$$\dot{\mathbf{x}} = \mathbf{A}_p(U_\infty)\mathbf{x} + \mathbf{B}_p\mathbf{u} + \mathbf{w}_p, \tag{37}$$

where

$$\mathbf{A}_p(U_\infty) \triangleq \left.\frac{\partial \mathbf{f}(\mathbf{x}, U_\infty)}{\partial \mathbf{x}}\right|_{\mathbf{x}=0},$$

$$\mathbf{B}_p \triangleq \left.\frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}=0},$$

and $\mathbf{w}_p$ is unit white-noise input.

With performance output $\mathbf{z}$ considered, there is:

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \mathbf{z} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_p(U_\infty) & \mathbf{I} & \mathbf{B}_p \\ \mathbf{C}_z & 0 & \mathbf{E}_z \\ \mathbf{C}_p & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{w} \\ \mathbf{u} \end{bmatrix}, \tag{38}$$

where $\mathbf{C}_p = \mathbf{I}$ for full-state feedback, $\mathbf{C}_z = [\mathbf{Q}^{\frac{1}{2}} \quad 0]^{\mathrm{T}}$, and $\mathbf{E}_z = [0 \quad \mathbf{R}^{\frac{1}{2}}]^{\mathrm{T}}$.

Let $\mathbf{P}(U_\infty)$ be a scheduled matrix which varies with the freestream airspeed U_∞. According to Theorem 1, a stable closed-loop system under the dynamically tuned control law as in (14) requires $\|\mathbf{P}\| > m_P$ where the value of the scalar $m_P \in \mathbb{R}^+$ depends on the system dynamics. In the case of ASAF, m_P is not constant but varies with U_∞. That is, $m_P = m_P^*(U_\infty)|_{U_\infty = U}$ for any practical airspeed U, where $m_P^*(U_\infty)$ is a generalized function. To find $\mathbf{P}(U_\infty)$ that satisfies the condition of $\|\mathbf{P}(U_\infty)\| > m_P^*(U_\infty)$, a Lyapunov matrix $\mathbf{X}(U_\infty) = \mathbf{X}^{\mathrm{T}}(U_\infty) > 0$ and an auxiliary parameter-dependent performance variable $\mathbf{Z}(U_\infty)$ are introduced to form the following linear matrix inequalities (LMIs):

$$\begin{bmatrix} \dot{\mathbf{X}} + \mathbf{A}_c^{\mathrm{T}}\mathbf{X} + \mathbf{X}\mathbf{A}_c & \mathbf{X} \\ \mathbf{X} & -\nu\mathbf{I} \end{bmatrix} \leq 0, \tag{39}$$

$$\begin{bmatrix} \mathbf{X} & \mathbf{C}_{cz}^{\mathrm{T}} \\ \mathbf{C}_{cz} & \mathbf{Z} \end{bmatrix} \geq 0, \tag{40}$$

and

$$\mathrm{Tr}(\mathbf{Z}) < \nu, \tag{41}$$

where

$$\mathbf{A}_c := \mathbf{A}_p(U_\infty) - \frac{1}{2}\mathbf{B}_p\mathbf{R}^{-1}\mathbf{B}_p^{\mathrm{T}}\mathbf{P}(U_\infty),$$

$$\mathbf{C}_{cz} := \mathbf{C}_z - \frac{1}{2}\mathbf{E}_z\mathbf{R}^{-1}\mathbf{B}_p^{\mathrm{T}}\mathbf{P}(U_\infty),$$

and ν is a performance index.

Let $\bar{\mathbf{B}}_c := \frac{1}{2}\mathbf{B}_p\mathbf{R}^{-1}\mathbf{B}_p^{\mathrm{T}}$, $\mathbf{G}(U_\infty) = \mathbf{P}(U_\infty)\mathbf{X}^{-1}(U_\infty)$, $\mathbf{Y}(U_\infty) = \mathbf{X}^{-1}(U_\infty)$. Then (39) and (40) can be transformed into:

$$-\dot{\mathbf{Y}} + \mathbf{A}_p\mathbf{Y} + \mathbf{Y}\mathbf{A}_p^{\mathrm{T}} + \bar{\mathbf{B}}_c\mathbf{G} + \mathbf{G}^{\mathrm{T}}\bar{\mathbf{B}}_c^{\mathrm{T}} \geq 0, \tag{42}$$

$$\begin{bmatrix} \mathbf{Y} & (\mathbf{C}_z\mathbf{Y} + \mathbf{E}_z\mathbf{G})^{\mathrm{T}} \\ \mathbf{C}_z\mathbf{Y} + \mathbf{E}_z\mathbf{G} & \mathbf{Z} \end{bmatrix} \geq 0. \tag{43}$$

In light of (1), $\mathbf{A}(U_\infty)$ can be structured as:

$$\mathbf{A}(U_\infty) = \mathbf{A}_1 + \mathbf{A}_2U_\infty + \mathbf{A}_3U_\infty^2. \tag{44}$$

Therefore, $\mathbf{Y}(U_\infty)$ and $\mathbf{G}(U_\infty)$ take the same structure as:

$$\mathbf{Y}(U_\infty) = \mathbf{Y}_1 + \mathbf{Y}_2U_\infty + \mathbf{Y}_3U_\infty^2, \tag{45}$$

and

$$\mathbf{G}(U_\infty) = \mathbf{G}_1 + \mathbf{G}_2U_\infty + \mathbf{G}_3U_\infty^2. \tag{46}$$

Solving for $\mathbf{Y}(U_\infty)$ and $\mathbf{G}(U_\infty)$ through (41), (42) and (43) gives $\|\mathbf{P}(U_\infty)\| > m_P^*(U_\infty)$ in the form of:

$$\mathbf{P}(U_\infty) = \mathbf{G}(U_\infty)\mathbf{Y}^{-1}(U_\infty). \tag{47}$$

with $m_P^*(U_\infty)$ implicitly embedded in the LMIs derived.

Remark 5: *As discussed in Section 1, existing methods capable of synthesizing optimal control laws for nonlinear systems are all limited to locally nonlinear cases and hence not suitable for direct implementation for ASAF without modification and improvements. The proposed procedure for designing the parameter matrix* $\mathbf{P}$ *yields* $\mathbf{P}(U_\infty)\| > m_P^*(U_\infty)$ *across* U_∞ *of interest, satisfying the condition of* $\mathbf{P}\| > m_P$ *in Theorem 1, and thus generalizes Theorem 1 to a more general scenario, the globally nonlinear case, with nonlinear dynamics described in (2). This makes the proposed method suitable for ASAF, and forms the second contribution of this paper.*

3.4 Online system identification

Note that the knowledge of $\mathbf{F}(\mathbf{x},\mathbf{u})$ and $\mathbf{g}(\mathbf{x})$ is required for real-time synthesis of nonlinear optimal control laws. Although $\mathbf{F}(\mathbf{x},\mathbf{u})$ and $\mathbf{g}(\mathbf{x})$ is analytically available, the presence of un-modeled dynamics or uncertainties can degrade controller performance as discussed in Section 1. To mitigate this problem, an NN-based identifier is proposed in the following form:

$$\dot{\mathbf{x}} = \mathbf{W}_s^{\mathrm{T}}\mathbf{\Phi}_s(\mathbf{x},\mathbf{u}) + \boldsymbol{\varepsilon}_s, \tag{48}$$

where $\mathbf{W}_s^{\mathrm{T}} \in \mathbb{R}^{n_{ws}\times n_x}$ and $\mathbf{\Phi}_s(\mathbf{x},\mathbf{u}) \in \mathbb{R}^{n_{ws}}$ are the ideal weights and nonlinear activation functions of the NN, respectively.

Motivated by Modares et al. (2013), the system states $\mathbf{x}$ can be expressed as:

$$\mathbf{x} = \mathbf{W}_s^{\mathrm{T}}\boldsymbol{\eta}(\mathbf{x}) + \mathbf{A}_f\boldsymbol{\mu}(\mathbf{x}) + \boldsymbol{\varepsilon}_x, \tag{49}$$

with

$$\dot{\boldsymbol{\eta}}(\mathbf{x}) = -\mathbf{A}_f\boldsymbol{\eta}(\mathbf{x}) + \mathbf{\Phi}_s(\mathbf{x},\mathbf{u}), \quad \boldsymbol{\eta}(\mathbf{x}_0) = 0, \tag{50}$$

and

$$\dot{\boldsymbol{\mu}}(\mathbf{x}) = -\mathbf{A}_f\boldsymbol{\mu}(\mathbf{x}) + \mathbf{x}, \quad \boldsymbol{\mu}(\mathbf{x}_0) = 0, \tag{51}$$

where $\boldsymbol{\eta}(\mathbf{x}) \in \mathbb{R}^{n_{ws}}$ and $\boldsymbol{\mu}(\mathbf{x}) \in \mathbb{R}^{n_x}$ are auxiliary regressors, $\mathbf{A}_f = \sigma\mathbf{I}_{n_x\times n_x}$ with $\sigma \in \mathbb{R}^{+}$, and

$$\boldsymbol{\varepsilon}_x = e^{-\mathbf{A}_f t}\mathbf{x}_0 + \int_0^t e^{-\mathbf{A}_f(t-\tau)}\varepsilon_s d\tau.$$

Denote the estimate of $\mathbf{x}$ by $\hat{\mathbf{x}}$. For fast estimation of $\hat{\mathbf{W}}_s$ towards $\mathbf{W}_s$, the EKF is considered for online tuning. In this study, multiple single-input-single-output (SISO) EKFs in a parallel configuration are employed to reduce the computational expense. On this basis, we have:

$$\begin{cases} \dot{\hat{\mathbf{W}}}_{s(i)} = \mathbf{K}_{s(i)}(\mathbf{x}_{(i)} - \hat{\mathbf{x}}_{(i)}) \\ \hat{\mathbf{x}}_{(i)} = \hat{\mathbf{W}}_{s(i)}\boldsymbol{\eta}(\mathbf{x}) + \mathbf{A}_f\boldsymbol{\mu}(\mathbf{x}) \end{cases}, \tag{52}$$

where $\mathbf{K}_s \in \mathbb{R}^{n_w s}$ is the EKF gain and subscript (i) restricts the parameters to the i^{th} decoupled EKF.

Each EKF gain vector $\mathbf{K}_{s(i)}$ can be computed using:

$$\mathbf{K}_{s(i)} = \mathbf{S}_{s(i)}\mathbf{H}_{s(i)}^{\mathrm{T}}R_s^{-1}, \tag{53}$$

$$\mathbf{H}_{s(i)}^{\mathrm{T}} = \frac{\partial \hat{\mathbf{x}}_{(i)}}{\partial \hat{\mathbf{W}}_{s(i)}} = \boldsymbol{\eta}(\mathbf{x}), \tag{54}$$

$$\dot{\mathbf{S}}_{s(i)} = \mathbf{Q}_s - \mathbf{S}_{s(i)}\mathbf{H}_{s(i)}^{\mathrm{T}}R_s^{-1}\mathbf{H}_{s(i)}\mathbf{S}_{s(i)}. \tag{55}$$

where $\mathbf{Q}_s \in \mathbb{R}^{n_w \times n_w} > 0$ and $R_s > 0$ are defined the same as $\mathbf{Q}_f$ and $\mathbf{R}_f$.

The input dynamics $\mathbf{g}(\mathbf{x})$ can then be obtained as:

$$\hat{\mathbf{g}}(\mathbf{x}) = \frac{\partial \hat{\mathbf{F}}(\mathbf{x}, \hat{\mathbf{u}})}{\partial \hat{\mathbf{u}}} = \frac{\partial \hat{\mathbf{W}}_s^{\mathrm{T}}\boldsymbol{\Phi}_s(\mathbf{x}, \hat{\mathbf{u}})}{\partial \hat{\mathbf{u}}}. \tag{56}$$

That is,

$$\mathbf{g}(\mathbf{x}) = \frac{\partial \boldsymbol{\Phi}_s(\mathbf{x}, \hat{\mathbf{u}})}{\partial \hat{\mathbf{u}}}\hat{\mathbf{W}}_s. \tag{57}$$

Remark 6: *The known analytical model is embedded into the Identifier NN in the form of the initial values of* $\mathbf{W}_s$ *obtained via pre-training the NN using the model. For the case of ASAF, as the on-board controller is normally switched on prior to the airspeed reaching the flutter boundary, pre-training the Identifier NN offline using known dynamics at this airspeed suffices. When the airspeed increases, un-modeled and mismatching dynamics can be captured, with the pre-trained NN updated in real time accordingly.*

4 Wind-Tunnel Experiments

Experiments were performed in a temperature regulated closed-loop wind tunnel at the University of Adelaide, Australia, with the setup shown in Figure 2. The wind tunnel has a 0.5×0.5 m testing duct, and can generate up to 30 m/s smooth airflow. A virtual-spring-damper system (VSDS) was used to introduce custom stiffness and damping setting. The parameters of the overall aeroelastic system used in experiments are listed in Table 1, and the system has a flutter boundary around 14.6 m/s.

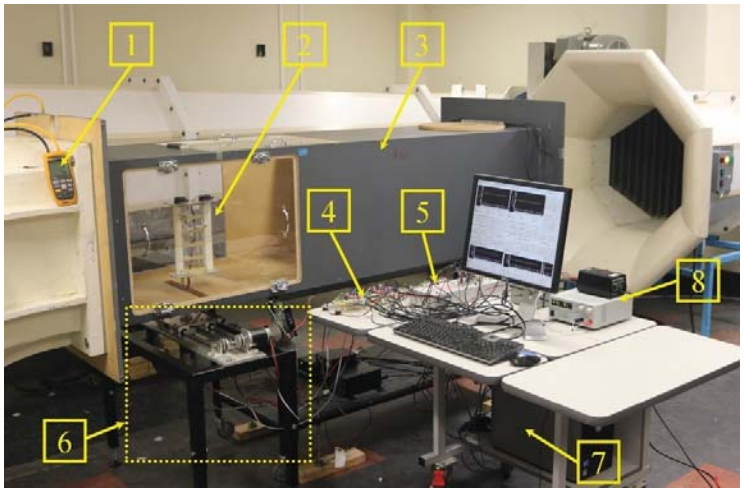

Figure 1. Wind tunnel experiment setup (1, pressure transducer connected to a Pitot tube; 2, airfoil section; 3, wind tunnel test duct; 4, custom I/O board; 5, dSPACE® DS1104 R&D controller board; 6, virtual spring-damper system; 7, controller PC; 8, power supply).

Table 1: Parameters of the Experimental Aeroelastic System

Parameters	Values	Parameters	Values
a	−0.5685	c_h	14 kg/s
b	0.0753 m	c_α	0.042 $\mathrm{kg\cdot m^2/s}$
m_α	0.851 kg	c_β	4.231×10^{-4} $\mathrm{kg\cdot m^2/s}$
m_β	0.030 kg	c_γ	4.327×10^{-4} $\mathrm{kg\cdot m^2/s}$
m_γ	0.058 kg	k_h	$50+300h^2$ N/m
s	0.26 m	k_α	$0.3+30\alpha^2$ Nm/rad
ρ	1.225 $\mathrm{kg/m^3}$	k_β	4.57×10^{-3} Nm/rad
r_α	0.0329 m	k_γ	4.70×10^{-3} Nm/rad
r_β	1.019×10^{-2} m	I_α	2.431×10^{-3} $\mathrm{kg\cdot m^2}$
r_γ	4.401×10^{-3} m	I_β	2.307×10^{-6} $\mathrm{kg\cdot m^2}$

$C_{l\alpha}$	6.573	I_γ	4.791×10^{-6} kg·m^2
$C_{l\beta}$	3.472	L_β	0.0875 m
$C_{l\gamma}$	−0.1453	L_γ	−0.01 m
$C_{m\alpha}$	0	$C_{m\gamma}$	0.0983
$C_{m\beta}$	−0.6307		

NN are prepared as Remark 3, with key information summarized in Table 2. $\mathbf{M}(U_\infty)$ and $\mathbf{G}(U_\infty)$ were designed using the parameters in Table 1 for the airspeed range from 14.6 m/s to 20 m/s with a gridding of 50 evenly spaced points. $\mathbf{P}(U_\infty)$ was calculated in real time using (47). $\bar{Q}(\mathbf{x})$ in (5) was structured as $\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x}$, with $\mathbf{Q}$ and other parameters listed in Table 3.

Table 2: Summary of the NNs used in the proposed controller.

	Input	Order	Neurons Number	Weights Initialization
Identifier NN	$[\mathbf{x}^{\mathrm{T}}, \mathbf{u}^{\mathrm{T}}]^{\mathrm{T}}$	$\mathbf{x}$: up to 4$^{\text{th}}$ $\mathbf{u}$: 1$^{\text{st}}$	135	simulation-based training for 14.6 m/s airspeed
Critic NN	$\mathbf{x}$	up to 4$^{\text{th}}$	65	zeros

Table 3: Other parameters of the proposed NN controller.

Parameters	Values
$\mathbf{A}_f$	$1\times10^3\mathbf{I}$
$\mathbf{Q}$	$\mathrm{diag}(1, 1, 10^{-4}, 10^{-4}, 0.1, 0.1, 10^{-4}, 10^{-4})$
$\mathbf{R}$	$1\times10^2\mathbf{I}$
$\mathbf{Q}_f$	$1\times10^3\mathbf{I}$

$\mathbf{R}_f$	$\mathbf{I}$
$\mathbf{Q}_{s(i)}$	$1\times10^5\mathbf{I}$
$\mathbf{R}_{s(i)}$	1

Tests were conducted at two different airspeeds, and flutter was allowed to develop to reach LCO before the controller under testing was turned on. To ensure consistent initial conditions $\mathbf{x}(t_c)$ throughout all tests under the same settings, where t_c is the time when the controller is switched on, the controller was configured to be triggered when α crossed zero immediately after 15 seconds. This means $t_c > 15$ s.

As discussed in Section 1 and throughout the paper, there is no existing policy-iteration algorithm suitable for ASAF without modification and improvement. Therefore, no suitable NN-based optimal controller counterparts can be compared in experiments. In order to evaluate the ASAF performance improvement gained by using the proposed controller, a linear-parameter-varying (LPV) controller in the form of linear-quadratic-regulator (LQR) synthesized by means of LMIs (Prime, 2010) was reconstructed for the 4-DOF model as in (1) with the parameters in Table 1 and the weighting $\mathbf{Q}$ and $\mathbf{R}$ same as those used by the proposed NN controller.

Plunge and pitch responses as well as control surfaces deflections of the airfoil section in the wind-tunnel tests under the proposed NN controller and the LPV-LQR controller at different airspeeds are plotted in Figure 3 for 14.8 m/s and Figure 5 for 18m/s. Higher airspeeds were not tested due to the torque output limit of the VSDS motors. The trajectories of NN weights are presented in Figures 4 and 6.

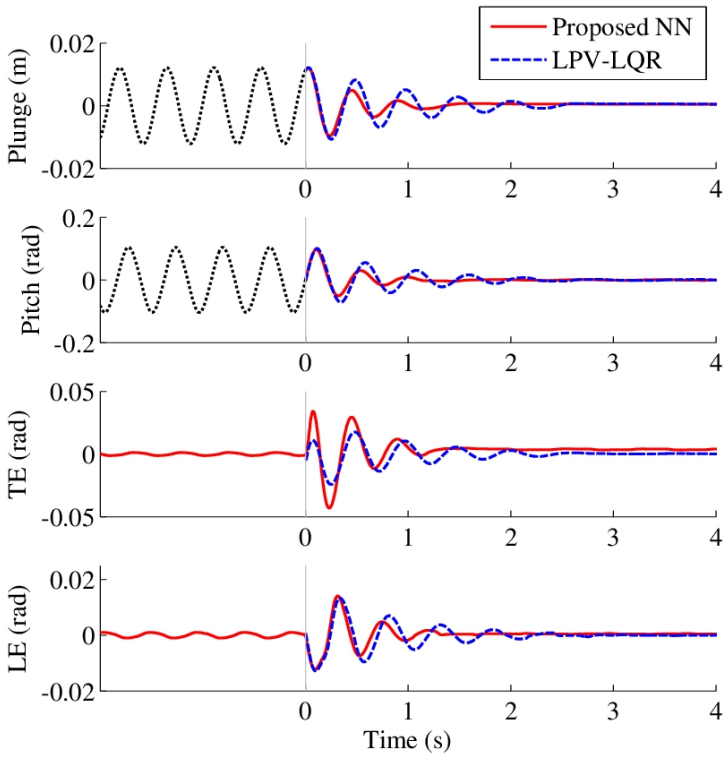

Figure 3. Suppressing developed flutter at 14.8 m/s airflow speed.

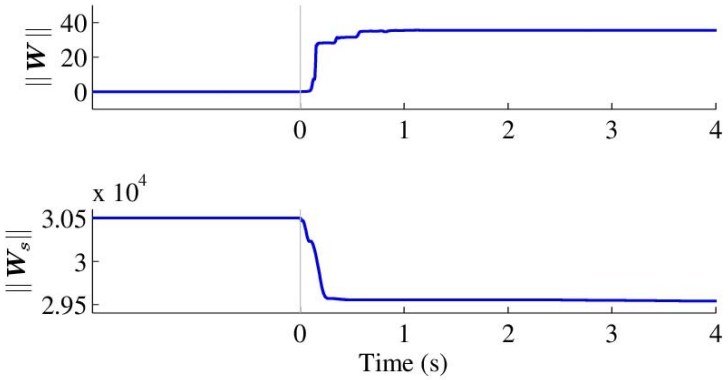

Figure 4. Convergence trajectories of the Critic and Identifier NN weights of the proposed controller at 14.8 m/s airflow speed.

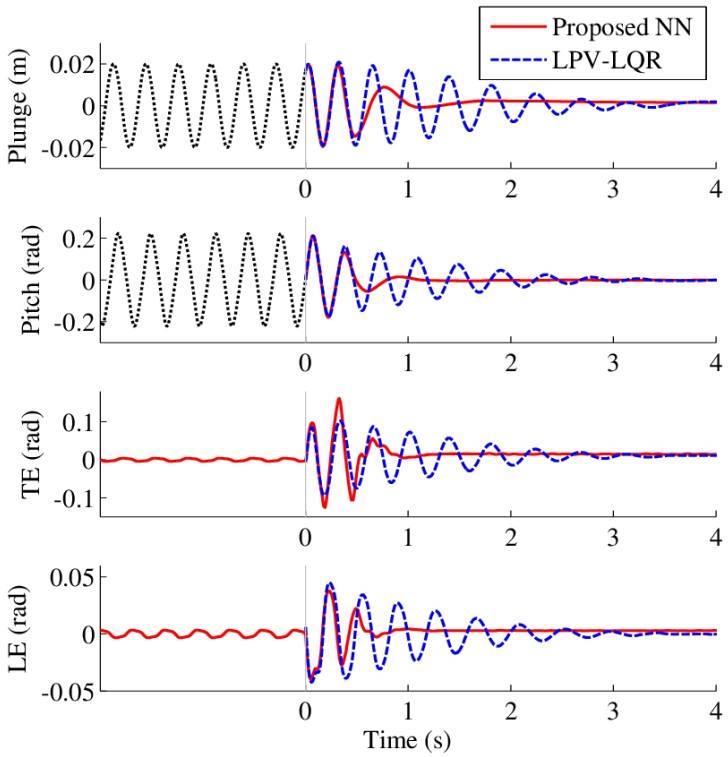

Figure 5. Suppressing developed flutter at 18 m/s airflow speed.

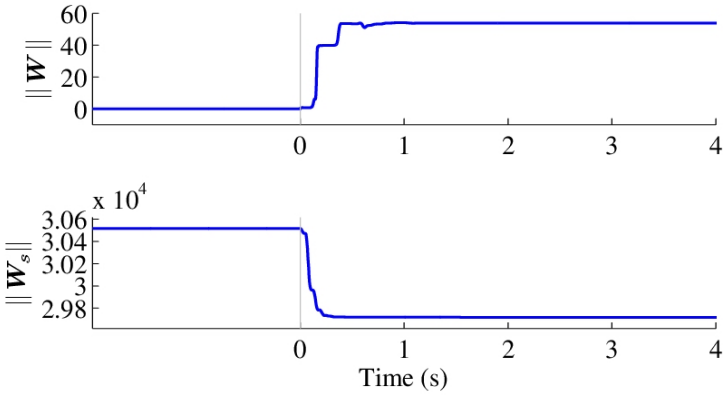

Figure 6. Convergence trajectories of the Critic and Identifier NN weights of the proposed controller at 18 m/s airflow speed.

At 14.8 m/s, the flutter was effectively suppressed within 1.5 seconds under the proposed NN controller, with only mild demands on the deflection of control surfaces. By comparing Figures 3 and 4, it can be seen the Identifier NN has higher rate of convergence than that of the Critic NN, which means the latter is able to access updated and more accurate system dynamics for control law improvement. The Critic NN also settles 1 second before the flutter is fully suppressed, indicating satisfactory parameter

convergence. This validates the selection of the activation functions sets for both the Identifier NN and the Critic NN, and also indicates that optimal control was obtained under experiment conditions. In comparison, it takes longer for the LPV-LQR controller to fully suppress the flutter. Similar phenomena can be observed for 18 m/s, as shown in Figures 5 and 6, where however, relatively larger differences between the responses under the two controllers can be observed.

To better capture the performance differences between the two controllers, performance cost is evaluated for $t = 0 \to 4$ s according to (5) using the experiment data with discrete approximation. Costs are each calculated and averaged from 4 tests under the same settings to ensure data consistency, and are listed in Table 4. It can be concluded from Table 4 that the proposed NN control suppresses the flutter better with lower cost at both airspeeds, compared with the LPV-LQR control.

Table 4: Performance costs calculated from experiment data.

Airspeed	LPV-LQR	Proposed NN Controller
14.8 m/s	5.372	4.893
18.0 m/s	0.627	0.545

5. CONCLUSIONS

The proposed novel form of VFA for synchronous policy iteration contributes to a simplified and compact controller without jeopardizing the closed-loop stability. The proposed procedure based on LMIs for designing a scheduled parameter matrix further generalizes the new VFA for globally nonlinear systems to suit ASAF applications. As validated in the wind-tunnel experiments, the proposed controller successfully improves ASAF from the optimal control perspective, with the impact of modeling uncertainties mitigated.

References

Abu-Khalaf M and Lewis FL. (2005) Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. *Automatica* 41(5): 779-791.

Bichiou Y, Nuhait AO, Abdelkefi A, et al. (2016) Unsteady aeroelastic behaviors of rigid airfoils with preset angles of attack. *Journal of Vibration and Control* 22(4): 1010-1022.

Brillante C and Mannarino A. (2016) Improvement of aeroelastic vehicles performance through recurrent neural network controllers. *Nonlinear Dynamics* 84(3): 1479-1495.

Carnahan JJ and Richards CM. (2008) A modification to filtered-X LMS control for airfoil vibration and flutter suppression. *Journal of Vibration and Control* 14(6): 831-848.
Chen G, Sun J and Li Y-M. (2012) Adaptive reduced-order-model-based control-law design for active flutter suppression. *Journal of Aircraft* 49(4): 973-980.
Chen L, He F and Sammut K. (2009) Vibration suppression of a principal parametric resonance. *Journal of Vibration and Control* 15(3): 439-463.
Cui X, Zhang H, Luo Y, et al. (2017) Finite-horizon optimal control of unknown nonlinear time-delay systems. *Neurocomputing* 238: 277-285.
Fazelzadeh SA, Azadi M and Azadi E. (2017) Suppression of nonlinear aeroelastic vibration of a wing/store under gust effects using an adaptive-robust controller. *Journal of Vibration and Control* 23(7): 1206-1217.
Fazelzadeh SA, Rasti A and Sadat-Hoseini H. (2014) Optimal flutter suppression of nonlinear typical wing section using time-domain finite elements method. *Journal of Aerospace Engineering* 27(5): 04014028 (1-10).
Feng T, Zhang H, Luo Y, et al. (2015) Stability analysis of heuristic dynamic programming algorithm for nonlinear systems. *Neurocomputing* 149, Part C(0): 1461-1468.
Finlayson BA. (1972) *The method of weighted residuals and variational principles: With application in fluid mechanics, heat and mass transfer*, New York, NY: Academic Press.
Fung YC. (1955) *An introduction to the theory of aeroelasticity,* NY: John Wiley & Sons.
Gao M-Z and Cai G-P. (2016) Finite-time fault-tolerant control for flutter involving control delay. *Journal of the Franklin Institute* 353(9): 2009-2029.
Gao M-Z, Cai G-P and Nan Y. (2016) Finite-time fault-tolerant control for flutter of wing. *Control Engineering Practice* 51: 26-47.
Gujjula S, Singh SN and Yim W. (2005) Adaptive and neural control of a wing section using leading- and trailing-edge surfaces. *Aerospace Science and Technology* 9(2): 161-171.
Jiang Y and Jiang Z-P. (2015) Global adaptive dynamic programming for continuous-time nonlinear systems. *IEEE Transactions on Automatic Control* 60(11): 2917-2929.
Keyser RD, Copot C, Hernandez A, et al. (2017) Discrete-time internal model control with disturbance and vibration rejection. *Journal of Vibration and Control* 23(1): 3-15.
Kiumarsi B, Lewis FL and Levine DS. (2015) Optimal control of nonlinear discrete time-varying systems using a new neural network approximation structure. *Neurocomputing* 156: 157-165.
Ko J, Strganac TW, Junkins JL, et al. (2002) Structured model reference adaptive control for a wing section with structural nonlinearity. *Journal of Vibration and Control* 8(5): 553-573.
Lee KW and Singh SN. (2013) L_1 adaptive control of a nonlinear aeroelastic system despite gust load. *Journal of Vibration and Control* 19(12): 1807-1821.
Lewis FL, Dawson DM and Abdallah CT. (2003) *Robot manipulator control: Theory and practice, second edition, revised and expanded,* New York, NY: Marcel Dekker, Inc.

Luo M, Gao M and Cai G. (2016) Delayed full-state feedback control of airfoil flutter using sliding mode control method. *Journal of Fluids and Structures* 61: 262-273.
Modares H, Lewis FL and Naghibi-Sistani M-B. (2013) Adaptive optimal control of unknown constrained-input systems using policy iteration and neural networks. *Neural Networks and Learning Systems, IEEE Transactions on* 24(10): 1513-1525.
Nayfeh AH, Hammad BK and Hajj MR. (2012) Discretization effects on flutter aspects and control of wing/store configurations. *Journal of Vibration and Control* 18(7): 1043-1055.
O'Neil T and Strganac TW. (1998) Aeroelastic response of a rigid wing supported by nonlinear springs. *Journal of Aircraft* 35(4): 616-622.
Pak C-G, Friedmann PP and Livne E. (1995) Digital adaptive flutter suppression and simulation using approximate transonic aerodynamics. *Journal of Vibration and Control* 1(4): 363-388.
Platanitis G and Strganac TW. (2004) Control of a nonlinear wing section using leading- and trailing-edge surfaces. *Journal of Guidance, Control, and Dynamics* 27(1): 52-58.
Prime Z, Cazzolato B, Doolan C, et al. (2010) Linear-parameter-varying control of an improved three-degree-of-freedom aeroelastic model. *Journal of Guidance, Control, and Dynamics* 33(2): 615-619.
Prime ZD. (2010) Robust scheduling control of aeroelasticity. School of Mechanical Engineering, The University of Adelaide.
Rao VM, Behal A, Marzocca P, et al. (2006) Adaptive aeroelastic vibration suppression of a supersonic airfoil with flap. *Aerospace Science and Technology* 10(4): 309-315.
Saaed TE, Nikolakopoulos G, Jonasson J-E, et al. (2017) A state-of-the-art review of structural control systems. *Journal of Vibration and Control* 21(5): 919-937.
Schmidt DK. (2016) Stability augmentation and active flutter suppression of a flexible flying-wing drone. *Journal of Guidance, Control, and Dynamics* 39(3): 409-422.
Singh SN and Brenner M. (2003) Modular adaptive control of a nonlinear aeroelastic system. *Journal of Guidance, Control, and Dynamics* 26(3): 443-451.
Singh SN and Wang L. (2002) Output feedback form and adaptive stabilization of a nonlinear aeroelastic system. *Journal of Guidance, Control, and Dynamics* 25(4): 725-732.
Song Z-G and Li F-M. (2014) Optimal locations of piezoelectric actuators and sensors for supersonic flutter control of composite laminated panels. *Journal of Vibration and Control* 20(14): 2118-2132.
Strganac TW, Ko J and Thompson DE. (2000) Identification and control of limit cycle oscillations in aeroelastic systems. *Journal of Guidance, Control, and Dynamics* 23(6): 1127-1133.
Tang D, Chen L, Tian ZF, et al. (2015) Neural-network based online policy iteration for continuous-time infinite-horizon optimal control of nonlinear systems. In: *3rd IEEE China Summit and International Conference on Signal and Information Processing*, Chengdu, China, 12-15 July 2015.

Vamvoudakis KG and Lewis FL. (2010) Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica* 46(5): 878-888.

Vasconcellos RMG, Abdelkefi A, Hajj MR, et al. (2016) Airfoil control surface discontinuous nonlinearity experimental assessment and numerical model validation. *Journal of Vibration and Control* 22(6): 1633-1644.

Viswamurthy SR and Ganguli R. (2008) Using the complete authority of multiple active trailing-edge flaps for helicopter vibration control. *Journal of Vibration and Control* 14(8): 1175-1199.

Wang Y, Zhang Q and Zhu L. (2015) Active control of hypersonic airfoil flutter via adaptive fuzzy sliding mode method. *Journal of Vibration and Control* 21(1): 134-141.

Wang Z, Behal A and Marzocca P. (2011) Model-free control design for multi-input multi-output aeroelastic system subject to external disturbance. *Journal of Guidance, Control, and Dynamics* 34(2): 446-458.

Zhang K and Behal A. (2016) Continuous robust control for aeroelastic vibration control of a 2-D airfoil under unsteady flow. *Journal of Vibration and Control* 22(12): 2841-2860.

Zhao W, Li R and Zhang H. (2017) Leader--follower optimal coordination tracking control for multi-agent systems with unknown internal states. *Neurocomputing* 249: 171-181.