THE UNIVERSITY
of ADELAIDE

Active Suppression of Aerofoil Flutter via Neural-Network-Based Adaptive Nonlinear Optimal Control

Difan Tang

School of Mechanical Engineering
The University of Adelaide
South Australia 5005
Australia

Thesis submitted in fulfillment of the requirements for the degree of Ph. D. in Engineering on November 18, 2019.

Ph. D. Thesis

November 18, 2019

School of Mechanical Engineering
The University of Adelaide
South Australia 5005
Australia

Typeset by the author using LaTeX.
Printed in Australia.

Contents

List of Publications

The following four journal papers are included as main chapters of the thesis:

- **Article-1**

  Tang D, Chen L, Tian ZF and Hu E (2019) Developing a virtual stiffness-damping system for airfoil aeroelasticity testing. *Journal of Sound and Vibration* DOI: 10.1016/j.jsv.2019.115061.

- **Article-2**

  Tang D, Chen L, Tian ZF and Hu E (2019) Modified value-function-approximation for synchronous policy iteration with single-critic configuration for nonlinear optimal control. *International Journal of Control* DOI: 10.1080/00207179.2019.1648874.

- **Article-3**

  Tang D, Chen L, Tian ZF and Hu E (2018) Adaptive nonlinear optimal control for active suppression of airfoil flutter via a novel neural-network-based controller. *Journal of Vibrations and Control* 24(22): 5261–5272.

- **Article-4**

  Tang D, Chen L, Tian ZF and Hu E (2018) A neural-network approach for improving airfoil flutter suppression under control-input constraints. *Journal of Vibrations and Control* (submitted for publication).

The following conference papers have arisen from this research but are not included as main chapters of the thesis:

- Tang D, Chen L, and Tian ZF (2015) Neural-network based online policy iteration for continuous-time infinite-horizon optimal control of nonlinear systems. In: *2015 IEEE China Summit and International Conference on Signal and Information Processing*, Chengdu, China, 12-15 July 2015, pp. 792–796. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE).

- Tang D, Chen L, Tian ZF and Hu E (2017) Nonlinear optimal control for active suppression of airfoil flutter via a novel neural-network-based controller. In: *2017 25th Mediterranean Conference on Control and Automation*, Valletta, Malta, 3-6 July 2017, pp. 253–258. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE).

- Tang D, Chen L, Tian ZF and Hu E (2019) Developing a virtual stiffness-damping system for airfoil aeroelasticity testing. In: *2019 20th IEEE International Conference on Industrial Technology*, Melbourne, Australia, 13-15 February 2019, pp. 96–101. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE).

- Tang D, Chen L, Tian ZF and Hu E (2019) Parametric analysis for robust force/torque tracking control of a virtual stiffness-damping system in airfoil aeroelasticity testing. In: *2019 American Control Conference*, Philadelphia, USA, 10-12 July 2019, pp. 3557–3562. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE).

List of Figures

List of Tables

Nomenclature

Acronyms

2-DOF	Two-degrees-of-freedom
4-DOF	Four-degrees-of-freedom
AFS	Active flutter suppression
CIC	Control-input constraint
EKF	Extended Kalman filter
ESO	Extended-state observer
HJB	Hamilton-Jacobi-Bellman
LCO	Limit-cycle oscillation
LMI	Linear matrix inequality
LPF	Low-pass filter
LPV	Linear-parameter-varying
LQG	Linear quadratic Gaussian
LQR	Linear quadratic regulator
LTI	Linear-time-invariant
MIMO	Multi-input multi-output

MVFA	Modified value-function-approximation
NATA	Nonlinear aeroelastic test apparatus
NLIP	Nonlinear-in-the-parameters
NN	Neural network
NOCOS	Nonlinear optimal control online synthesis
ONNC	Optimal neural-network controller
SISO	Single-input single-output
SPI	Synchronous policy iteration
SRI	Stability robustness index
UIE	Unknown-input estimator
UUB	Uniformly ultimately bounded
VFA	Value function approximation
VIV	Vortex induced vibration
VSDS	Virtual stiffness-damping System

Greek Symbols

α, α_{cnn}, α_{snn}	NN scalar learning rates
$\beta(\cdot)$	Function: nonlinear output dynamics
$\delta(j\omega)$	Uncertainties at frequency ω
Γ	Lyapunov matrix
γ	Scalar that determines the magnitude of Γ
$\hat{\boldsymbol{\vartheta}}$	Estimated states by MESO

κ_1, κ_2, κ_c, κ_s	Intermediate variables
λ	Control bound
ϵ_1, ϵ_2	Intermediate variable vectors
$\mathbf{\Lambda}$	Intermediate variable vector
$\boldsymbol{\lambda}$	Matrix of control bounds
μ_1, μ_2	Vectors of filter regressors
$\mathbf{\Omega}$	Compact set: state-space states
$\mathbf{\Phi}(\cdot)$	NN activation functions set
$\mathbf{\Pi}$	Intermediate variable vector
ψ	Mass matrix of 4-DOF aeroelastic system
$\sigma_{\tilde{W}}$	Vector of equilibrium/residual set of $\tilde{W}_{(i)}$ dynamics
θ	Vector of bounded functions
ς, $\varsigma_{(i)}$	Intermediate variable vectors
$\varsigma_{na(i)}$, $\varsigma_{nb(i)}$	Intermediate variable vectors
ϑ	States vector of MESO
$\Xi(\cdot)$	Functions set: control
$\zeta(\cdot)$	Functions set: un-modelled dynamics
ν	Performance index
ω	Frequency (in radian)
ω_c	Cutoff frequency (in radian)
ω_d	Frequency of disturbances (in radian)

$\phi(\cdot)$	NN activation function
ρ	Air density
τ	Time constant
$\theta(\cdot)$	Bounded function
ε, $\varepsilon_{(i)}$, ε_{Y^*}, ...	Approximation errors
ξ_1, ξ_2, ξ_3, ξ_4	Intermediate variables
η_1, η_2, η_3, ...	Intermediate variables

Number Sets

$\mathbb{B}$	Binary numbers
$\mathbb{R}$	Real numbers
$\mathbb{Z}$	Integers

Roman Symbols

$\bar{\boldsymbol{c}}$	Matrix of structural damping coefficients realised via VSDS
$\bar{\boldsymbol{k}}$	Matrix of structural stiffness coefficients realised via VSDS
$\bar{c}_a$	Structural damping coefficient of pitch DOF realised via VSDS
$\bar{c}_h$	Structural damping coefficient of plunge DOF realised via VSDS
$\bar{k}_a$	Structural stiffness coefficient of pitch DOF realised via VSDS
$\bar{k}_h$	Structural stiffness coefficient of plunge DOF realised via VSDS
$\hat{\boldsymbol{d}}_e$	Vector of estimated equivalent unknown-inputs or equivalent power-loss
$\hat{\boldsymbol{u}}$	Vector of approximated and estimated nominal optimal control inputs

$\hat{y}$	Estimated output
$\boldsymbol{c}$	Matrix of structural damping coefficients
$\boldsymbol{d}$	Vector of unknown inputs
$\boldsymbol{d}_e$	Vector of equivalent unknown-inputs
$\boldsymbol{e}_x$	Vector of states estimation errors
$\boldsymbol{e}_\vartheta$	Vector of position estimation errors
$\boldsymbol{f}(\cdot)$	Vector of system internal/drift dynamics
$\boldsymbol{g}(\cdot)$	Vector of system control-input dynamics
$\boldsymbol{h}$	Vector of performance outputs
$\boldsymbol{k}$	Matrix of structural stiffness coefficients
$\boldsymbol{m}$	Mass matrix of 2-DOF aeroelastic system
$\boldsymbol{q}$	Displacements vector
$\boldsymbol{r}$	Vector of general signals
$\boldsymbol{u}$	Vector of total control inputs
$\boldsymbol{u}^*$	Vector of nominal optimal control inputs
$\boldsymbol{u}_c$	Vector of nominal control inputs
$\boldsymbol{u}_{(i)}$	Vector of nominal control inputs at the i^{th} iteration step
$\boldsymbol{w}$	Process noise
$\boldsymbol{x}$	Vector of system states
$\boldsymbol{x}_0$	Vector of system initial states
$\boldsymbol{x}_e$	Vector of system equilibrium states

z	Intermediate variable vector
a_{mi}	The i^{th} coefficient of LPF numerator polynomial
$b_{\dot{x}}$, b_{ϕ}, b_G, ...	Scalar bounds for the 2-norm of various functions
c_a	Structural damping coefficient of pitch DOF
c_h	Structural damping coefficient of plunge DOF
c_{le}	Virtual damping coefficient associated with leading-edge servo
c_{te}	Virtual damping coefficient associated with trailing-edge servo
e_1	Cost function estimation error with regard to a static control policy
e_2	Value function estimation error (HJB error)
e_y	Output estimation error
k_a	Structural stiffness coefficient of pitch DOF
k_h	Structural stiffness coefficient of plunge DOF
k_{le}	Virtual stiffness coefficient associated with leading-edge servo
k_{te}	Virtual stiffness coefficient associated with trailing-edge servo
m_a	Mass of aerofoil section (excluding leading- and trailing-edge control surfaces if available)
m_s	Mass of force/torque transducer
m_{le}	Mass of leading-edge control surface
m_{te}	Mass of trailing-edge control surface
m_{VSDS}	Equivalent mass of VSDS plunge/pitch DOF transmission
n_d	Number of unknown inputs

n_F	Number of aerodynamic forces
n_f	Number of subsystem states
n_n	Number of neurons
n_u	Number of control inputs
n_x	Number of system states
p_1, p_2, p_3	Design parameters of MESO
q_a	Pitch angle
q_h	Plunge displacement
$q_{le\text{-}cmd}$	Leading-edge control surface control command
q_{le}	Leading-edge control surface deflection
$q_{te\text{-}cmd}$	Trailing-edge control surface control command
q_{te}	Trailing-edge control surface deflection
r_a	Distance from elastic axis to centre of mass
r_{le}	Distance from the hinge axis of leading-edge control surface (LECS) to LECS centre of mass
r_{te}	Distance from the hinge axis of trailing-edge control surface (TECS) to TECS centre of mass
s	Laplace variable
t	Time
t_0	Starting time
t_c	Time when controller is turned on
t_d	Ending time

v	Input noise
y	Measurement output
$\bar{\boldsymbol{F}}_s$	Measured aerofoil structural forces generated by VSDS
$\bar{Q}(\cdot)$	States weighting function
$\bar{R}(\cdot)$	Control-input weighting function
$\hat{\boldsymbol{F}}_s$	Estimated aerofoil structural forces generated by VSDS
$\hat{\boldsymbol{W}}$	Vector of estimated critic NN ideal weights for VFA
$\hat{\boldsymbol{W}}_s$	Vector of estimated identifier NN ideal weights
$\hat{\boldsymbol{W}}_{(i)}$	Vector of estimated critic NN ideal weights for approximating cost function at the i^{th} iteration step
$\hat{V}(\cdot)$	Approximated and estimated value function
$\hat{V}_{(i)}(\cdot)$	Approximated and estimated cost function at the i^{th} iteration step
$\mathcal{F}(\cdot)$	Functions set: general
$\mathcal{F}_A(\cdot)$	Functions set: aeroelastic system dynamics
$\mathcal{L}$, $\mathcal{L}_V$, $\mathcal{L}_W$, ...	Lyapunov functions
$\mathcal{Z}(\cdot)$	Functions set: general
$\boldsymbol{A}_c$	Transition matrix of state-space controller model
$\boldsymbol{A}_f$	Transition matrix of state-space filter model
$\boldsymbol{A}_p$	Transition matrix of nominal state-space system model
$\boldsymbol{A}_s$,	Transition matrix of identified state-space system model
$\boldsymbol{B}_c$	Control-input distribution matrix of state-space controller model

$\boldsymbol{B}_d$	Unknown-input distribution matrix
$\boldsymbol{B}_f$	Input distribution matrix of state-space filter model
$\boldsymbol{B}_p$	Control-input distribution matrix of nominal state-space system model
$\boldsymbol{B}_s$	Control-input distribution matrix of identified state-space system model
$\boldsymbol{C}_c$	Output matrix of state-space controller model
$\boldsymbol{C}_f$	Output matrix of state-space filter model
$\boldsymbol{C}_p$	Output matrix of nominal state-space system model
$\boldsymbol{C}_s$	Output matrix of identified state-space system model
$\boldsymbol{D}_s$	Pass-through matrix of identified state-space system model
$\boldsymbol{F}_c$	Centripetal forces vector
$\boldsymbol{F}_s$	Actual forces resulted from aerofoil structural stiffness and damping
$\boldsymbol{F}_{aero}$	Aerodynamic forces vector
$\boldsymbol{G}$	Intermediate variable matrix
$\boldsymbol{H}$	Observation matrix of EKF
$\boldsymbol{I}$	Identity matrix
$\boldsymbol{J}(\cdot)$	Transformed parameter-dependent variable matrix
$\boldsymbol{K}_d$	UIE gain matrix
$\boldsymbol{K}_f$	Gain of EKF for critic NN
$\boldsymbol{K}_x$, $\boldsymbol{K}_w$, $\boldsymbol{K}_r$	Gain matrices for LQG tracking control
$\boldsymbol{K}_{s(i)}$	Gain of the i^{th} decoupled EKF for identifier NN

$\boldsymbol{L}$	Gain matrix of state observer
$\boldsymbol{M}_1$, $\boldsymbol{M}_2$, $\boldsymbol{M}_3$	Intermediate variable matrices
$\boldsymbol{N}_1$, $\boldsymbol{N}_2$, $\boldsymbol{N}_3$	Parameter matrices
$\boldsymbol{P}$	Parameter matrix
$\boldsymbol{P}(\cdot)$	Scheduled parameter matrix
$\boldsymbol{Q}$	Weighting matrix for plant states
$\boldsymbol{Q}_f$	Weighting matrix for critic NN weights
$\boldsymbol{Q}_s$	Weighting matrix for identifier NN weights
$\boldsymbol{R}$	Weighting matrix for plant control inputs
$\boldsymbol{R}_r$	Vector containing diagonal elements of $\boldsymbol{R}$
$\boldsymbol{S}$	Covariance matrix of the EKF used in critic NN
$\boldsymbol{S}_{s(i)}$	Covariance matrix of the i^{th} EKF used in identifier NN
$\boldsymbol{W}$	Vector of critic NN weights
$\boldsymbol{W}^*$	Vector of critic NN ideal weights for VFA
$\boldsymbol{W}_s^*$	Vector of identifier NN ideal weights
$\boldsymbol{W}_{(i)}$	Vector of critic NN ideal weights for approximating cost function at the i^{th} iteration step
$\boldsymbol{X}$	Lyapunov matrix
$\boldsymbol{Y}$	Lyapunov matrix
$\boldsymbol{Z}$	Performance variable matrix
$\tilde{\boldsymbol{W}}$	Vector of estimation errors of critic NN ideal weights for VFA

$\tilde{\boldsymbol{W}}_{(i)}$	Vector of estimation errors of critic NN ideal weights for approximating cost function at the i^{th} iteration step
$C_{l\text{-}a}$	Lift coefficient of aerofoil section (excluding leading- and trailing-edge control surfaces, if available) about the elastic axis
$C_{l\text{-}le}$	Lift coefficient of aerofoil leading-edge control surface about the elastic axis
$C_{l\text{-}te}$	Lift coefficient of aerofoil trailing-edge control surface about the elastic axis
$C_{m\text{-}a}$	Moment coefficient of aerofoil section (excluding leading- and trailing-edge control surfaces, if available) about the elastic axis
$C_{m\text{-}le}$	Moment coefficient of aerofoil leading-edge control surface about the elastic axis
$C_{m\text{-}te}$	Moment coefficient of aerofoil trailing-edge control surface about the elastic axis
$C_{me\text{-}a}$	Effective moment coefficient of aerofoil section (excluding leading- and trailing-edge control surfaces, if available) about the elastic axis
$C_{me\text{-}le}$	Effective moment coefficient of aerofoil leading-edge control surface about the elastic axis
$C_{me\text{-}te}$	Effective moment coefficient of aerofoil trailing-edge control surface about the elastic axis
E	Error function
F_L	Aerodynamic lift force on aerofoil
F_M	Aerodynamic moment on aerofoil about the elastic axis
$F_{L\text{-}f}$	Total equivalent friction along the entire plunge-DOF power transmission of VSDS from motor output to the sliding carriage

$F_{L\text{-}in1}$, $F_{L\text{-}in2}$	Internal forces of VSDS plunge DOF transmission
$F_{L\text{-}in3}$, $F_{L\text{-}in4}$	Internal forces of VSDS plunge DOF transmission
$F_{M\text{-}in1}$, $F_{M\text{-}in2}$	Internal torques of VSDS pitch DOF transmission
$F_{M\text{-}in3}$, $F_{M\text{-}in4}$	Internal torques of VSDS pitch DOF transmission
F_{motor}	Translational force (tension) on synchronous belt B1
$G_f(s)$	Transfer function of LPF
I_a	Aerofoil section moment of inertia (excluding that of leading- and trailing-edge control surfaces, if available) about the elastic axis
I_{le}	Moment of inertia of leading-edge control surface (LECS) about LECS hinge axis
I_{te}	Moment of inertia of trailing-edge control surface (TECS) about TECS hinge axis
L_s	Aerofoil section span
$L_{3c/4}$	Distance from elastic axis to 3/4-chord position
L_{fc}	Distance from elastic axis to aerofoil aerodynamic centre
L_{hc}	Half-chord length of aerofoil section
L_{le}	Distance from elastic axis to the hinge axis of leading-edge control surface
L_{te}	Distance from elastic axis to the hinge axis of trailing-edge control surface
$P_n(s)$	Transfer function of VSDS plunge/pitch DOF
R_f	Weighting matrix for inputs of the EKF used in critic NN
$R_{s(i)}$	Weighting matrix for inputs of the EKF used in identifier NN

$S_s(j\omega)$	System sensitivity at frequency ω
T_1, T_2, T_3, ...	Intermediate variables
U_∞	Relative velocity of aerofoil to the air
$V(\cdot)$	Cost function
$V^*(\cdot)$	Value function
$V_{(i)}(\cdot)$	Cost function at the i^{th} iteration step
C_l	Lift coefficient of aerofoil section at the aerodynamic centre
C_m	Moment coefficient of aerofoil section at the aerodynamic centre

Abstract

This thesis deals with active flutter suppression (AFS) on aerofoils via adaptive nonlinear optimal control using neural networks (NNs).

Aeroelastic flutter can damage aerofoils if not properly controlled. AFS not only ensures flutter-free flight but also enables the use of aerodynamically more efficient lightweight aerofoils. However, existing optimal controllers for AFS are generally susceptible to modelling errors while other controllers less prone to uncertainties do not provide optimal control. This thesis, thus, aims to reduce the impact of the dilemma by proposing new solutions based on nonlinear optimal control online synthesis (NOCOS) according to online updated dynamics.

Existing NOCOS methods, with NNs as essential elements, require a separate initial stabilising control law for the overall system, an additional stabilising tuning loop for the actor NN, or an additional stabilising term in the critic NN tuning law, to guarantee the closed-loop stability for unstable and marginally stable systems. The resulting complexity is undesired in AFS applications due to computational concerns in real-time implementation. Moreover, the existing NOCOS methods are confined to locally nonlinear systems, while aeroelastic systems under consideration are globally nonlinear. These make all the existing NOCOS algorithms inapplicable to AFS without modification and improvement. Therefore, this thesis solves the aforementioned problems through the following step-by-step approaches.

Firstly, a four degrees-of-freedom (4-DOF) aeroelastic model is considered, where leading- and trailing-edge control surfaces of the aerofoil are used to actively suppress flutter. Accordingly, a virtual stiffness-damping system (VSDS) is developed

to simulate physical stiffness in the aeroelastic system. The VSDS, together with a scaled-down typical aerofoil section placed in a wind tunnel, serve as an experimental 4-DOF aeroelastic test-bed for synthesis and validation of proposed AFS controllers that follow.

Secondly, a Modified form of NN-based Value Function Approximation (MVFA), tuned by gradient-descent learning, is proposed for NOCOS to address the closed-loop stability in a compact controller configuration suitable for real-time implementation. Its validity and efficacy are examined by the Lyapunov stability analysis and numerical studies.

Thirdly, a systematic procedure based on linear matrix inequalities is further proposed for synthesising a scheduled parameter matrix to generalise the MVFA to to globally nonlinear cases, so that the new NN controller suits AFS applications. In addition, the extended Kalman filter (EKF) is proposed for the new NN controller for fast parameter convergence. An identifier NN is also derived to capture and update aeroelastic dynamics in real time to mitigate the impact of modelling errors. Wind-tunnel experiments were conducted for validation.

Finally, a non-quadratic functional is introduced to generalise the performance index to tackle the problem where control inputs are constrained. The feasibility of including the non-quadratic cost function under the proposed control scheme based on the MVFA is examined via the Lyapunov stability analysis and was also experimentally evaluated through the wind-tunnel testings.

The proposed NN controllers are compact in structure and shown capable of maintaining the closed-loop stability while eliminating the need for a separate initial stabilising control law for the overall system, an additional tuning loop for the actor NN, and an additional stabilising term in the critic NN tuning law. Under the new control schemes, online synthesised nonlinear control laws are optimal in the cases with and without constraints in control. Comparisons drawn with a popular linear-parameter-varying (LPV) controller in the form of the widely used linear quadratic regulator (LQR) in experiments show that the proposed NN controllers

outperform the LPV-LQR algorithm and improve AFS from the optimal control perspective. Specifically, the proposed NN controllers can effectively mitigate the impact of modelling errors, successfully solving the mentioned dilemma involved in AFS. The results also confirm that the proposed NN controllers are suitable for real-time implementation.

Declarations

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship and the Sir Ross and Sir Keith Smith Fund.

Signature Date

Acknowledgements

I wish to give my sincere gratitude to my supervisors Dr Lei Chen, Dr Zhao Feng Tian and Associate Professor Eric Hu, for their support, guidance and assistance, from general aspects of project management to specific technical details.

I would also like to thank the staff in the mechanical and electronic workshops at the School of Mechanical Engineering, the University of Adelaide, for building the virtual stiffness-damping system (VSDS) and the scaled-down aerofoil section for wind-tunnel experiments. In particular, my thanks go to Mr Philip Schmidt, Ms Lydia Zhang, Mr Derek Franklin and Mr Norio Itsumi, for providing handy helps with, and advice on, upgrading the electronic system of the VSDS.

An additional mention goes to Mr Ashfaqul Bari, Mr Lawrence Berry, Mr Jaime Liew, Mr Scott Poland, Ms Hoi Ki NG, Mr Ka Kit Chan, Mr Lee Chan Tee and Mr Raymond Sze Hwa Fung, who contributed to the project in mechanical design of the VSDS and the aerofoil.

Special appreciation is given to my friend Dr Hao Huang for enjoyable times together and helps of all kinds, in particular, tremendous supports during my hard time when I was recovering from a chest surgery. Thanks are also given to Dr Guoqiang Qi, for his valuable advice and encouragement.

Last but not least, I wish to acknowledge my great debt to my parents, Mr Canming Tang and Mrs Wen Wen, as well as my wife Wei Wei, for their consistent support and encouragement along my journey of study.

Chapter 1

Introduction

1.1 Motivation

Aerofoil flutter is a type of self-feeding oscillation due to the interaction between aerodynamic loads and non-rigid aerofoil structures, occurring at and above a certain airflow velocity depending on the aerofoil structural characteristics. Such dynamic aeroelastic instability can develop dramatically in a short instance and may cause structure fatigue and eventually catastrophic failure if not controlled properly.

To prevent aerofoil flutter from happening within an aircraft's flight envelope, traditional practice relies on stiffness distribution optimisation, mass balancing, or geometry modification performed at design stage (Haftka et al., 1975; Markowitz and Isakson, 1978; Reed et al., 1980). These are known as passive approaches for aerofoil flutter suppression, which are effective but on the other hand lead to heavier aerofoil designs (Karpel, 1981). To further postpone the onset of aerofoil flutter to a higher airspeed without excessive structure modification and weight penalty, one may apply proper control efforts (such as generating structural forces via piezoelectric actuation or changing aerodynamic loads through aerofoil control surfaces) to counteract flutter once it happens. This is categorised as an active approach, which emerged since the 1970s and has been proven effective theoretically and experimentally in numerous leading research projects (Mukhopadhyay, 2003).

More importantly, active flutter suppression (AFS) renders it possible for more efficient flexible lightweight aerofoils to be used to achieve weight reduction and aerodynamic improvement, as successfully demonstrated in the US Active Flexible Wing Program (Perry et al., 1995). These achievements through AFS are of great significance to the aviation industry because modern and future aircraft are desired to be more efficient than ever before (NSTC, 2010).

Despite some success in AFS by various means in past studies, challenges remain, especially in seeking better control algorithms to treat nonlinearities and uncertainties in time-varying aeroelastic systems. As to be detailed in the literature review in Chapter 2, the challenges are:

(1) Existing model-based optimal controllers for AFS applications are susceptible to modelling errors, while other methods less prone to uncertainties do not consider optimal control. This has become a dilemma of AFS controller synthesis hindering further improvement of AFS performance.

(2) In addition, the available control authority for AFS is subjected to constraints in practice, for instance, actuator saturation. Saturated control, being one of the nonlinearities involved in AFS, has nevertheless not been addressed in an optimal sense.

Therefore, deriving new control schemes to improve AFS performance from the optimal control perspective is of particular significance. To address the AFS controller synthesis dilemma, this thesis suggests an approach based on synthesising nonlinear optimal control in real time. Despite the availability of techniques under the theory framework of nonlinear optimal control online synthesis (NOCOS) to be detailed in Chapter 2, none of the existing approaches are applicable to AFS without modification and improvement, due to realistic problems in relation to stability, application scope, and real-time implementation. To be specific, the problems are:

(1) Aeroelastic dynamics are generally nonlinear for a constant airspeed (i.e. locally nonlinear) and also vary nonlinearly against the airspeed (i.e. globally

nonlinear). However, available NOCOS methods are all confined to locally nonlinear cases, and thus, not capable of AFS without modification and improvement.

(2) Stability during controller online self-learning is one of the major concerns among the existing NOCOS algorithms, in which sophisticated structures with various stabilising mechanisms are used to maintain the closed-loop stability. In the interests of real-time implementation, a compact NOCOS method in place of the existing ones is desired. In particular, Chapter 2 points out that existing studies all share a common form of value function approximation (VFA) as a fundamental element of synchronous policy iteration (SPI) involved in NOCOS. Could the SPI configuration be further simplified for AFS in the interests of reliable and efficient real-time implementation? Would a different form of VFA give a different stability result under the SPI framework?

Accordingly, deriving methods to tackle these NOCOS related problems for AFS applications are deemed necessary.

1.2 Aims and objectives

The primary aim of the research in this thesis is to derive new control schemes to address the AFS controllers synthesis dilemma and improve AFS performance from the optimal control perspective. Associated objectives include:

(1) Developing an experimental test-bed based on the concept of the virtual stiffness-damping system (VSDS) to facilitate testing and validation of AFS controllers in wind-tunnel experiments;

(2) Deriving a novel NOCOS scheme with a compact configuration without compromising the closed-loop stability, for locally nonlinear systems, assuming no constraints in control;

(3) Deriving a new adaptive nonlinear optimal controller for AFS, by generalising the novel NOCOS scheme to globally nonlinear scenarios. At this stage, it is assumed that there are no constraints in control. The proposed control is to be validated in wind-tunnel experiments using the developed VSDS;

(4) Deriving a new adaptive nonlinear optimal controller for AFS under control-input constraints (CICs), by generalising the proposed NOCOS scheme to take CICs into account, from the optimal performance perspective. The proposed control is to be validated in wind-tunnel experiments using the developed VSDS;

1.3 Methodologies in general

Methodologies developed to accomplish the aim and objectives are given in detail in Chapters 3, 4, 5, and 6. This section describes the methodologies in general, and also sets out the methodology-related context in support of the discussions in subsequent chapters.

1.3.1 Mechanical mechanism for active flutter suppression on aerofoils

AFS can be realised by using existing control surfaces such as leading-edge slats and trailing-edge flaps or ailerons on wings, which has been numerically and experimentally shown effective in an extensive amount of studies as given in the review by Livne (2018). Proper deployment of control surfaces can change aerodynamic forces exerted on aerofoils accordingly to counteract and alleviate aerofoil oscillatory motions. Well established experimental test-beds include but are not limited to the benchmark active control technology (BACT) (Farmer, 1982, 1984), active flexible wing (AFW) wind-tunnel prototype (Miller, 1988), nonlinear aeroelastic test apparatus (NATA) (O'Neil and Strganac, 1998), and the newly

developed Lockheed Martin X-56A demonstrator (Burnett et al., 2016). Depending on the physical configuration of these test-beds, the control surfaces utilised for flutter suppression consist of trailing-edge flaps, ailerons, upper spoilers, or leading-edge slats, functioned solely or jointly (Pendleton et al., 2000; Perry et al., 1995; Waszak, 2001). Associated merits of this class of solution include: (1) little or no modifications to existing aerofoil mechanical structures, and hence no considerable additional hardware investments; (2) improved ride quality as a result of direct aerodynamic treatment.

As the work in this thesis focuses on improving AFS performance by making scientific and technical breakthroughs from the control aspect instead of the mechanical mechanism, the control-surface actuation method, given its well established basis, is adopted as the mechanical mechanism for AFS studied in this research.

1.3.2 Aeroelastic model

Aerofoil flutter is generally studied and characterised by a typical rigid wing section of a finite span with two-degrees-of-freedom (2-DOF) motions: plunge and pitch (Fung, 1955). For studies in this thesis, the 2-DOF aeroelastic system with leading- and trailing-edge control surfaces based on the NATA test-bed is selected given its well-established theoretical and experimental basis and that it captures representative nonlinearities in real aeroelastic systems (Ko et al., 2002; O'Neil and Strganac, 1998; Platanitis and Strganac, 2004). The experimental test-bed developed in this thesis is based on this 2-DOF model and details are given in Chapter 3. When the dynamics of both control surfaces are taken into account, this 2-DOF model is improved and extended to the four-degrees-of-freedom (4-DOF) one developed in Prime (2010). This 4-DOF model not only describes the first plunge and first pitch mode oscillations of a rigid wing section under linear unsteady aerodynamic loads in subsonic flow, but also captures the delay in control-surface responses to corresponding commands. As the 4-DOF model allows better AFS results as validated in wind-tunnel experiments using the NATA test-bed (Prime

et al., 2010; Prime, 2010), it is adopted for controller synthesis in this thesis. More details of the model are presented in Chapters 5 and 6. Nonlinear transnational (plunge) and torsional (pitch) stiffness can be introduced in a polynomial form up to specified order in both the 2-DOF and 4-DOF aeroelastic models. The AFS controllers proposed in this thesis, although derived by referring to the 4-DOF aeroelastic model, can be easily extended to more complicated aeroelastic systems.

1.3.3 Controller verification and validation

The research in this thesis seeks theoretical breakthroughs in controller synthesis to accomplish the aim, which necessitates stability analysis in the form of establishing new theorems and giving corresponding proofs. In addition, numerical studies are deemed appropriate for verification of the proposed novel NOCOS scheme using representative dynamic systems that allow convincing comparison, while wind-tunnel experiments were performed to validate the proposed AFS controllers.

The VSDS used in experiments is based on the NATA test-bed, but replaces the physical plunge and pitch springs each with an electric drive. It offers advantages over NATA in that the stiffness and damping of the plunge and pitch DOFs can be adjusted conveniently without frequent change of mechanical components. However, whether the VSDS can precisely mimic the behaviour of the NATA test-bed highly depends on how well the electric drives track the desired plunge force and pitch torque calculated from the assumed physical stiffness and damping. This requires the high-performance force/torque tracking of the developed VSDS, which was addressed first (in Chapter 3) before the VSDS was used for AFS controller testing in the wind-tunnel experiments.

1.4 Thesis format and publications

This thesis is presented in publication format according to the 'Academic Program Rules' of the University of Adelaide and contains four papers published or submitted

for publication by peer-reviewed journals. These papers serve as main chapters of the thesis, with slight amendments made for concise, consistent and coherent presentation, which include:

- British English spelling is used throughout the thesis;
- Manuscript titles are changed to chapter headings of a shorter form;
- Bibliography of each paper are converted uniformly to the Harvard referencing style, and integrated at the end of the thesis;
- Figures and tables are renumbered and renamed, with repetition in different papers omitted;
- Equations are renumbered;
- Repeated mathematical assumptions and definitions from later papers are removed;
- Additional notes are given where necessary in chapters based on publications, to enhance the overall coherence.

All publications arising from the research in this thesis are included in the List of Publications on Page v, where the four journal papers are identified as, and hereafter referred to as, *Article-1*, *Article-2*, *Article-3*, and *Article-4*, respectively.

1.5 Thesis Outline

From this Chapter onward, the thesis is organised as follows.

Past and recent studies as well as latest advances in the fields of AFS and NOCOS are reviewed in Chapter 2, where a summary of identified research gaps is given at the end.

As this research elects to validate the proposed AFS controllers via wind-tunnel experiments, Chapter 3, which is based on *Article-1*, introduces the experimental

test-bed specially developed for subsequent AFS tests. It is essential that the test-bed is reliable in providing trustful results, and therefore, this chapter describes how the VSDS was developed, presents and explains special approaches taken to ensure reliable performance. The test-bed also sets out the aeroelastic model as well as associated parameters used in the AFS controller derivation and synthesis in Chapters 5 and 6 that follow. The first objective (Page 4) of the research is completed on successful development of the VSDS.

To derive the AFS controllers that accomplish the aim, the two realistic problems of the existing NOCOS algorithms (Page 3) are required to be solved first. The NOCOS algorithm structure problem under the locally nonlinear setting is addressed in Chapter 4 based on *Article-2*, by proposing a novel NOCOS scheme that is compact in configuration without compromising the closed-loop stability. This chapter provides detailed derivation, stability analysis, and numerical verification of the proposed compact NOCOS scheme for locally nonlinear systems, completing the second objective (Page 4) of the research.

On completion of the second objective, the research moves on to generalising the novel NOCOS scheme to globally nonlinear scenarios so that it suits AFS applications. This work completes the third objective (Page 4), contributes to *Article-3*, and forms Chapter 5, where details of the proposed adaptive nonlinear optimal controller based on the novel NOCOS scheme are given, together with wind-tunnel experiment results obtained using the VSDS in Chapter 3. The research aim is partially accomplished at this stage, assuming no CICs.

By further generalising the novel NOCOS scheme in Chapter 4 and the new AFS controller in Chapter 5 to treat CICs from the optimal control perspective, a new adaptive nonlinear optimal controller for AFS under CICs is derived. This work completes the fourth objective (Page 4), delivers *Article-4*, forms Chapter 6, and most importantly, accomplishes the primary aim of the research in this thesis.

Finally, Chapter 7 gives concluding remarks regarding contributions and significance of the overall project and also suggests possible directions of future work.

Chapter 2

Literature review

The literature review first looks into the basic mechanism of aerofoil flutter in Section 2.1, providing a background of the control problems discussed in sections that follow. A review on advances in various controllers for active flutter suppression (AFS) on aerofoils is then given in Section 2.2. With the potential to further improve AFS performance, methods for solving nonlinear optimal control problems are discussed in Section 2.3. Research gaps to be addressed in this project are summarised in Section 2.4.

2.1 Basic mechanism of aerofoil flutter

Aerofoil flutter is a type of self-feeding oscillation due to the interaction between aerodynamic loads and non-rigid aerofoil structures, occurring at and above a certain airflow velocity depending on structural characteristics of the aerofoil (Balakrishnan, 2012; Fung, 1955). With linear stiffness in structure, flutter develops rapidly into infinitely large amplitude that can cause immediate damage. For aeroelastic systems with nonlinear stiffness generating a 'hardening' effect, the oscillation may not infinitely grow in amplitude but maintains at a certain level, termed as 'limit-cycle oscillation (LCO)'.

Aerofoil flutter can be initiated by any off-equilibrium state, which may be either because of aerofoil self-excitation during control surfaces deployment or more com-

monly, caused by unsteady airflow (Kimberlin, 2003), as tiny variation/perturbation in airflow velocity can result in disequilibrium of an aeroelastic system (Marzocca et al., 2001). In fact, aerofoil flutter only occurs at and above a particular airspeed which is called the 'flutter boundary'. Prior to an aircraft reaching its flutter boundary, the damping in the aeroelastic system reduces both its kinetic and elastic energy and brings down any excited oscillations. On reaching or exceeding the flutter boundary, the total energy dissipated from the aeroelastic system becomes less than or equal to the total energy absorbed, resulting in the amplitude of any excited oscillations continuously growing until reaching the LCO state or structure failure.

An aerofoil in flutter can have several degrees of freedom referred to as 'flutter modes', the dominant two of which are the first plunge mode (bending) and first pitch mode (torsion), respectively (Kimberlin, 2003). The aerofoil structure and aerodynamic forces together contribute to coupled damping of each mode, and the damping ratio of the critical flutter mode may start to decrease beyond some point with increasing airspeed (Heeg, 1993; Wright and Cooper, 2008). When the damping ratio of the critical flutter mode reaches zero, flutter occurs.

2.2 Controllers for active flutter suppression on aerofoils

In light of the basic mechanism of aerofoil flutter, practically preventing such aeroelastic instability means to move the flutter boundary to a higher airspeed so that flutter does not occur within the operation envelop of an aircraft. This applies to both passive and active approaches for aerofoil flutter suppression. Among the active control approaches, there have been enormous studies on solutions involving actuation of structures embedded with piezoelectric materials and deployment of aerofoil control surfaces. Despite different AFS solutions in terms of mechanical realisation, the underlying control algorithms play a crucial role in successful implementation. Specifically, the AFS controllers are required to properly utilise the

available actuator authority to counteract oscillations and contribute to a closed-loop system with positive damping ratio of the critical flutter mode at the airspeed where an open-loop (uncontrolled) system goes unstable. This thus moves the flutter boundary of an aerofoil to a higher airspeed.

Suppressing LCOs and even bifurcations can be achieved via various control methods, such as the hybrid linear and nonlinear-velocity feedback control proposed by Chen et al. (2009) and the position feedback controller in Yabuno et al. (2012). Similarly in terms of AFS, a wide variety of control techniques are available. As exploratory attempts and for ease of implementation, non-adaptive control laws with constant gains were employed in early studies. Some are single-input single-output (SISO) classical controllers designed in frequency domain (Edwards, 1983; Schmidt and Chen, 1986; Waszak and Srinathkumar, 1992). Although being effective in early studies and useful in latest research (Schmidt, 2016), these methods are inconvenient for multiple-input multiple-output (MIMO) systems as a well-functioning design may need excessive trials, which can be costly and time-consuming. Alternatively, modern control theory in state space provides some relatively more convenient solutions to both SISO and MIMO cases. Under this category a wide range of variations in controller synthesis exist. For example, aerodynamic energy eigenvalues assignment (Nissim, 1971), suboptimal reduced-order control law synthesis (Mukhopadhyay et al., 1981), internal model control (Viswamurthy and Ganguli, 2008), linear-quadratic-Gaussian (LQG) algorithms (Newsom and Mukhopadhyay, 1985; Vipperman et al., 1998), singular-value-gradient based robustness optimisation (Newsom and Mukhopadhyay, 1985), eigenvalue placement with eigenvector shaping (Ghiringhelli et al., 1990; Liebst et al., 1986), H_∞ and μ-synthesis robust control design methods (Waszak, 2001), etc. Regardless of these various forms of optimisation and robustness recovery techniques used, very limited expansion of the flutter boundary were achieved because the controllers mentioned were designed with constant parameters. The resulting control laws are not adaptive to change of system parameters such as airspeed, and may have poor performance if there

are off-design variations in system parameters. Moreover, aeroelastic systems are generally nonlinear, thus a linear optimal controller is actually suboptimal given that it does not take nonlinearities into account. In order to go beyond these limitations, AFS controllers are desired to be nonlinear and adaptive to varying factors throughout the entire flight envelope.

A direct way of enabling adaptive control in nonlinear systems is to put the linear quadratic regulator (LQR) design online (Friedmann et al., 1997; Guillot and Friedmann, 2000; Pak et al., 1995). This is basically a discrete procedure that operates on a series of linear models estimated in real time at small enough intervals technically allowed and solves the algebraic Riccati equation at every time step. A major limitation is, identified system parameters need to be passed on to the LQR routine online, requiring some forms of transformation that are not simple in most cases, even though the auto-regressive-moving-average model is employed. In addition, since the LQR itself is solved via an iterative approach, such online implementation is computationally intensive. Reducing the iteration number to one as in Pak et al. (1995) can reduce the computation time to some extent but leads to approximation errors of the Riccati matrix and degraded performance.

Instead of running control synthesis online, methods of gain scheduling predefine a set of control laws offline and choose an appropriate one according to specific operating conditions (Leith and Leithead, 2000). In particular, advanced gain scheduling can be accomplished via systematic procedures based upon linear parameter-varying (LPV) techniques (Rugh and Shamma, 2000). Specifically for AFS applications, Prime et al. (2010) synthesise a gain-scheduled LQR controller via linear matrix inequalities (LMIs) as a generalised LPV control problem. The resulting LPV-LQR controller can self-schedule with airspeed and can effectively suppress developed LCOs at post-flutter-boundary airspeeds in wind tunnel experiments. In the work of Fazelzadeh et al. (2014) the time-domain finite element concept is employed as a special LPV formulation to address aeroelastic nonlinearities. The obtained controller is also shown effective in numerical studies. It is worth noting

that for controller synthesis, aeroelastic models are required to be known a priori, and the modelling quality directly affects the controller performance. Though a high-fidelity reduced-order model can be used to reduce the impact of modelling errors while facilitating controller synthesis (Chen et al., 2012), building a suitable LPV model that balances accuracy and simplicity remains a challenge because of the presence of uncertainties in aeroelastic systems in reality (Marcos and Balas, 2004). It is also worth noticing that aeroelastic systems possess a complicated combination of structural and aerodynamic nonlinearities (Li et al., 2012; Sheta et al., 2002), some of which vary as fast as aeroelastic dynamics, posing significant challenges for gain scheduling. When all or most of these nonlinearities are to be accounted for by the controller, LPV techniques are still capable but may face other realistic issues related to implementation due to increased complexity in computing automatically scheduled control laws.

To cope with nonlinearities in aeroelastic systems more effectively, the control law itself can be nonlinear, giving rise to a variety of more sophisticated methods. Ko et al. (1997) utilise partial and full feedback linearisation to derive locally and globally stabilising adaptive nonlinear controllers for an aeroelastic system of known dynamics with known structural nonlinearities. The associated stability and bifurcation structure of the resulting closed-loop systems are discussed in Ko et al. (1998). In the work of Ko et al. (1998), the need for explicit knowledge of the nonlinear pitch stiffness is exempted by employing high-order polynomial approximation with corresponding coefficients updated online. To gain global stability using only one control surface, the partial feedback linearisation scheme is integrated with a different form of parameterisation for the structural nonlinearity using a linear combination of transformed system states, coefficients of which are tuned online (Ko et al., 1999; Strganac et al., 2000). The case with one control surface is further extended to the configuration with two control surfaces as in Platanitis and Strganac (2004), by using a similar online tuning scheme to approximate the nonlinear torsional stiffness. Note that not all nonlinear systems can be completely linearised,

and other nonlinearities not included in the mathematical approximation can also pose stability risks (Ko et al., 1999). Moreover, although feedback linearisation yields direct cancellation of undesired nonlinearities and allows implementation of optimal control designed on the basis of the linearised model, the nonlinearities are not treated by considering optimal performance.

Compared with non-adaptive control laws, online tuning of some prescribed parameters enables synthesising control laws that are adaptive and robust. Instead of parameterised approximation, Zhang and Singh (2001) treat model uncertainties as a completely unknown function. The resulting adaptive controller is in the form of feedback linearisation but has the unknown function compensated by an estimate via a high-gain observer. This method is effective as demonstrated in experiments, but has a major drawback in that it is prone to sensor noises due to the use of the high-gain observer. Ko et al. (2002) apply model reference adaptive control (MRAC) method to update a pre-defined control law in real time according to the performance measure in the form of the error between desired and actual system outputs. Although being theoretically robust and valid for a wide range of airspeeds allowed by the actuator authority, the MRAC controller is unable to suppress LCOs at high airspeeds in experiments due to modelling errors not handled by the control law. Following the similar principle but being different in mathematical representation, direct adaptive controllers (DACs) are able to synthesise adaptive control laws online in a model-free manner with measurements of only the plunge displacement and pitch angle (Singh and Wang, 2002; Xing and Singh, 1999, 2000). Parameters of these controllers are modified directly in accordance with some prescribed performance measure by the use of backstepping techniques, with unmeasured states estimated by a reduced-order filter. For more freedom in choosing appropriate tuning laws, an alternative way can be the modular adaptive control (MAC), which redesigns the pre-defined control law in light of an online updated plant model rather than the performance measure used in the DACs (Bhoir and Singh, 2004; Rao et al., 2006; Singh and Brenner, 2003). AFS results

can be further improved with actuator dynamics taken into account (Behal et al., 2006a) or through various forms of matrix decomposition to handle coupled input dynamics when both leading- and trailing-edge control surfaces are used (Behal et al., 2006b; Lee and Singh, 2010; Reddy et al., 2007). However, the DACs and MACs have a major limitation that the approximation of nonlinearities is based on a linear-in-the-system-parameters assumption. As a result, a regression matrix, the determination of which is often complicated and time-consuming, is required to provide a suitable basis for approximating nonlinearities.

With state transformation and introduction of a Hurwitz parameter matrix, the need for a regression matrix can be waived, nonetheless at the cost of increased number of variables to be dynamically tuned (Lee and Singh, 2013). By using error dynamics from output feedback only, the methods in Carnahan and Richards (2008) and Zhang and Behal (2016) do not require a regression matrix and are numerically shown stabilising. The sliding-mode control is another nonlinear and adaptive control strain seen in AFS applications with the need for a regression matrix eliminated (Dilmi and Bouzouia, 2016; Luo et al., 2016b; Wang et al., 2015). If actuator faults involved in AFS are specifically targeted, one may apply the finite-time H_∞ adaptive fault-tolerant control (Gao and Cai, 2016; Gao et al., 2016). The controller robustness is further enhanced in the work of Fazelzadeh et al. (2017). Nevertheless, none of these methods treat control performance in an optimal sense.

Besides the nonlinear and adaptive control methods mentioned, neural networks (NNs) in AFS have also been attracting research attention. An NN can learn system dynamics online and adapt itself in real time, thus no system model is required a priori. A multi-layer NN has universal approximation properties (Hornik et al., 1989) that are particularly suitable for function approximation in nonlinear control. Even if an NN that is linear in the parameters (LIP) is used, its nonlinear activation functions set can still effectively provide a basis for function approximation. Therefore, there is no need to perform extensive modelling and preliminary analysis to find a regression matrix that is essential to other adaptive control techniques assuming

linearity in the system parameters. In an early study conducted by Spencer et al. (1999), a double-layer NN controller is used to command a trailing-edge flap on rotor blades to attenuate its vibrations caused by a wide variety of unknown and periodic disturbances. No offline training is needed. Collocated and non-collocated sensor and actuator, as well as extension to time-varying systems are considered. Successfully implementation was demonstrated in experimental tests where two different mechanisms – trailing-edge flaps and active wing tip twisting were controlled using the same NN controller (Spencer et al., 2002). An LIP adaptive controller and an NN controller were synthesised and compared in the work of Gujjula et al. (2005), where both leading- and trailing-edge control surfaces are used. Both controllers are found to have similar performance except that slightly less smooth transient responses are found for the NN controller. But this problem can be solved by simply increasing the number of neurons as explained by the author. It is worth emphasising that the LIP adaptive controller assumes known structure of uncertainties, while the NN controller approximates unstructured uncertainties. This is a significant advantage of NN controller over other LIP adaptive controllers. To address uncertainties associated with external disturbances, a multi-layer NN that is nonlinear in the parameters (NLIP) is proposed for AFS in Wang et al. (2011). A singularity-free controller design is achieved through symmetric-diagonal-upper triangular factorisation for matrix decomposition to decouple the input dynamics associated with the leading- and trailing-edge control surfaces. A proportion control component is added to stabilise the NN controller during the online training period, and a vector of nonlinear auxiliary signals is formed to improve robustness of the NN controller. The resulting controller shows considerable improvement over the LIP adaptive controller in Reddy et al. (2007) in terms of faster suppression and better robustness to disturbances. Further AFS improvement was made in the work of Brillante and Mannarino (2016), where two separate NNs dynamically tuned online are employed for system identification and control, respectively. Again, it is worth noting that optimal control is not considered in all the mentioned NN-based

controllers.

To allow aircraft normal and tactical maneuver while performing AFS, it is important to limit the amplitude of control used for AFS by setting constraints smaller than actuators saturation bounds. Although control-input constraints (CICs) have been considered in some AFS studies (Gao and Cai, 2016; Gao et al., 2016; Ko et al., 2002; Viswamurthy and Ganguli, 2008; Wang et al., 2011), none of the existing solutions address the problem in the sense of optimal control.

It can be seen from the existing studies that:

- The existing model-based optimal controllers for AFS applications are susceptible to modelling errors, while other methods less prone to uncertainties do not consider optimal control. This has become a dilemma of AFS controller synthesis hindering further improvement of AFS performance.
- In addition, the available control authority for AFS is subjected to constraints in practice, for instance, actuator saturation. Saturated control, being one of the nonlinearities involved in AFS, has nevertheless not been addressed in an optimal sense.

2.3 Optimal control for nonlinear systems

Optimal control for nonlinear systems involves solving a Hamilton-Jacobi-Bellman (HJB) equation. This differential equation is nonlinear and difficult to solve directly. As an alternative, the 'policy iteration' or 'value iteration' can be used (Howard, 1960; Sutton and Barto, 1998), which are based on a successive two-step iteration between policy evaluation and policy improvement. As distinguished by naming, these two approaches are slightly different, depending on which step initiates the process. The term 'policy' is specifically used in the field of adaptive dynamic programming (ADP) (Howard, 1960) and refers to a control law. The policy iteration process starts with a given admissible control law and proceeds to solve for the cost function associated with this initial control. The obtained cost function is then

used to compute an improved control law that possesses a lower cost. By repeating these two steps, the initial non-optimal control law evolves to an optimal one which corresponds to an optimal cost function, namely, the value function. With regard to the value iteration, a similar procedure is followed except that the overall process is initiated by a given value (cost). To implement the ADP, the value function in the HJB equation needs to be properly approximated. Possessing universal approximation properties, neural networks (NNs) are ideal candidates (Hornik et al., 1989).

For continuous-time systems, the policy iteration instead of the value iteration is commonly adopted. An early approach of NN-based policy iteration for continuous-time nonlinear systems is an offline method that takes control saturation into account (Abu-Khalaf and Lewis, 2005). The resulting control law is nonlinear and optimal for saturated actuators, outperforming its linear counterpart, the linear-quadratic regulator (LQR), which is only optimal when actuators are not saturated. Offline ADP was and remains effective and useful for handling optimal control in various challenging problems, including non-affine systems (Luo et al., 2016b; Mu et al., 2017; Wang et al., 2012), actuator saturation (Abu-Khalaf and Lewis, 2005; Heydari and Balakrishnan, 2013; Luo et al., 2015), unknown system dynamics (Li et al., 2017; Luo et al., 2015, 2016b; Mu et al., 2017, 2018; Wang and Liu, 2013; Wang et al., 2012; Wei et al., 2017; Zhao et al., 2015b), fixed final time (Heydari and Balakrishnan, 2013), finite approximation error (Wei et al., 2014), finite horizon(Mu et al., 2018), algorithm simplification (Heydari, 2014; Heydari and Balakrishnan, 2013; Wang and Liu, 2013), optimal tracking (Luo et al., 2016b), non-zero initial condition for value iteration (Wei et al., 2016), and extension to multi-agent system applications (Li et al., 2017). Nevertheless, offline methods are not suitable for AFS that requires controller online adaptation.

With some modification, some offline algorithms can be put online and does not need the knowledge of system internal dynamics (Feng et al., 2015; Jiang and Jiang, 2015; Liu et al., 2013b; Vrabie et al., 2009). In Vrabie et al. (2009) and Liu et al. (2013b),

by setting a non-zero initial condition and providing an initial stabilising control law, NN training begins when the system states start to settle back to equilibrium under the periodically evolving control. This method requires non-zero initial condition for NN training, posing a limitation in that the same updated control strategy and the same initial condition must be set for a next round of tuning if the system settles before convergence is reached. This is impractical in the AFS case. In addition, an initial stabilising controller is needed and largely affects the convergence time. A good choice of an initial control law can yield faster convergence but requires knowledge of the system. Moreover, the policy iteration is a sequential process that takes place with one step starting on completion of the other step, resulting in discrete update of control. Discontinuities in control signals should be smoothed by appropriate methods which are nonetheless not addressed. Further discussions on the impact of using a discount factor for the infinite-horizon cost are provided in Liu et al. (2013b), but the aforementioned limitations are not addressed. The necessity of having an initial stabilising control is removed in Feng et al. (2015), and the associated problem of system stability under the control law synthesised online after limited number of iterations is investigated. More advances are presented in the work of Jiang and Jiang (2015) where modelling uncertainties are taken into account and computation efficiency is improved, although an initial stabilising control is still required. The policy iteration involved in the approaches of Feng et al. (2015); Jiang and Jiang (2015); Liu et al. (2013b); Vrabie et al. (2009) is a sequential process with one step commencing upon completion of the previous step, resulting in discrete update of the control law.

Comparatively, the synchronous policy iteration (SPI) approach as in Vamvoudakis and Lewis (2010) offers more advantages. Instead of the step-by-step iteration, both the policy evaluation and policy improvement steps are performed simultaneously and continuously in real time, contributing to continuous update of the control law and hence smoother control signals. The critic-actor configuration used in Vamvoudakis and Lewis (2010) is formed by two separate NNs with dedicated

tuning for each. It guarantees the stability of the entire closed-loop system during online tuning without the necessity of providing an initial stabilising control law. A limitation is that both the internal/drift dynamics and control input dynamics of the system are assumed to be known, which can lead to sensitivity to modelling errors.

The theory framework of online SPI for synthesising optimal control for continuous-time nonlinear systems as first proposed in Vamvoudakis and Lewis (2010) have been enormously enriched by recent and latest advances in dealing with more complicated nonlinearities and system uncertainties as well as application scope extension and performance improvement (Xu et al., 2014). In Vamvoudakis and Lewis (2010), only partial knowledge of the system is required. Complete model-free design is also possible with integrated NN-based online system identification (Bhasin et al., 2013; Modares et al., 2013a). Without assuming the availability of all system states for feedback, those unmeasured states can be estimated in the meantime while optimal control laws are being approximated (Liu et al., 2013c). When disturbances are presented, one may consider the algorithms of Wang et al. (2014b). In particular, to cope with disturbances entering a system through control-input channels, a modified representation of the optimal control law with an added positive coefficient to produce a proportionally increasing gain is proposed in Wang et al. (2014b). This approach is especially suitable for applications in relation to decentralised control of large-scale plants consisting of several interconnected subsystems. In terms of unstructured uncertainties, the optimal control problem under discussion can be interpreted into another appropriate form for uncertain systems where the algorithm is required to provide robust performance with upper-bounded costs (Liu et al., 2014). To take actuator saturation into account, the use of a nonquadratic cost function is put to discussion on continuous-time nonlinear systems, with successful results (Modares and Lewis, 2014; Modares et al., 2013a, 2014, 2013b). In regard to optimal trajectory tracking control, similar but more complex principles apply, as proposed in Modares et al. (2014), with actuator saturation considered as well. If

faster online tuning is desired, one may consider the method in Bhasin et al. (2013), where a number of means are introduced to increase NN parameter convergence speed while maintaining closed-loop system stability. In the method proposed in Modares and Lewis (2014), the condition of persistence of excitation (PE) for online tuning is relaxed by means of introducing an experience-relay approach based on recorded past data. In addition, more insights with regard to the mechanism of policy iteration, its convergence, the uniqueness of the solution as well as the sufficient conditions are discussed in Heydari (2014). However, to guarantee closed-loop system stability, these online synchronous policy iteration algorithms either require an additional tuning loop with a non-standard tuning term for the actor NN, or rely on a logic algorithm to switch between different tuning modes for the critic NN. The additional tuning loop and the logic switch mechanism can both introduce more uncertainties into the system, and moreover, the logic switch mechanism can also cause discontinuities in control.

Compared with continuous-time cases, it is interesting to note that there are relatively more studies based on discrete-time systems in terms of solving for optimal control using the ADP approach (Si et al., 2004; Yu, 2009). While most recent ADP methods based on value iteration do not need initial stabilising control laws and do not require the non-standard stabilising term for tuning the actor NN as those used in continuous-time systems, these methods are limited to offline implementation (Heydari and Balakrishnan, 2013; Wang and Liu, 2013; Wei et al., 2014; Zhao et al., 2015b). Similar comments on the limitation of offline implementation can also be found in Al-Tamimi et al. (2008) and Wei and Chen (2014). The convergence of online implementation of a subclass of ADP, the heuristic dynamic programming (Werbos, 1990), is revisited and proven in Al-Tamimi et al. (2008) under the no-approximation-error assumption. Approximation errors are dealt with in Wei and Chen (2014) where a new θ-ADP technique is proposed with the potential for both online and offline implementation. Although the methods in Al-Tamimi et al. (2008) and Wei and Chen (2014) can be implemented online without stability

concerns, they does not provide continuous control law update because an inner tuning loop is required for successive approximation of the value function and the associated control at every iteration step. Significant improvements have been made in Kiumarsi and Lewis (2015) where both the critic and actor NNs are updated simultaneously and continuously with time. However, the generality of the method of Kiumarsi and Lewis (2015) is nonetheless unconfirmed due to the coupling of online tuning between the two NNs that results in increased difficulty in proving stability and parameter convergence.

Therefore, it is of significance and of particular interest if the tuning loop for the actor NN can be eliminated from the synchronous policy iteration process without adding complicated stabilising mechanisms that may bring additional uncertainties or cause control discontinuities. To simplify SPI implementation and reduce computational load, there have been efforts on event-based methods with single-critic configuration (Wang et al., 2017a,b). The event-based method reduces the needed online data, while the single-critic configuration uses the same NN for both the critic and actor and removes the necessity of separate actor NN tuning. The instability resulted from direct simplification of the actor-critic configuration is recognised in the work of Liu et al. (2013c), where initial weights of the critic NN need to be determined carefully by trial-and-error. Guaranteed stability can be achieved by adding a stabilising term to the critic tuning law (Huang et al., 2017; Liu et al., 2014; Wang et al., 2017b, 2014b).

It is worth noting that:

- All the existing nonlinear optimal control online synthesis (NOCOS) methods are 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).

The above facts render all these existing NOCOS methods inapplicable to AFS without modification and improvement.

Moreover, it is also interesting to note that the SPI schemes in the aforementioned studies all share a common form of value function approximation (VFA) with an NN directly employed. The questions are:

- Could the SPI configuration be further simplified for AFS in the interests of reliable and efficient real-time implementation?
- Would a different form of VFA give a different stability result under the SPI framework?

2.4 Research gaps

From Section 2.2 it can been seen that adaptive nonlinear control has been receiving enormous research attention under the increasing demand on higher AFS performance. However, existing methods for synthesising adaptive control strategies are subject to limitations associated with controller design, implementation, and performance. First, design of self-tuning controllers requires regression or transformation matrices that are application dependent and difficult to determine; Second, model-based controllers require knowledge of the plant and can be sensitive to modelling errors; Third, model-free controllers that do not require regression matrices are prone to instability caused by high-gain components. In comparison, some NN controllers offer better solutions to these limitations and can be a better choice for AFS, but a lack of solutions to optimal control is also identified. Thus, NOCOS featuring real-time learning of the changing aeroelastic dynamics and updating control in an optimal sense accordingly is deemed a potential direction that may give promising AFS improvement. Nonetheless, none of the available NOCOS methods are applicable to AFS due to limitations related to stability, application scope, and real-time implementation.

Concisely, three gaps have been identified and are to be addressed in this project:

(1) Optimal controllers among the existing methods are susceptible to un-modelled dynamics, while other controllers that are more adaptive to the changing environments and tolerant to un-modelled dynamics do not provide nonlinear optimal control. As a whole, these problems have become a critical dilemma in AFS control synthesis, hindering further improvement of AFS performance.

(2) Existing NOCOS methods are all confined to locally nonlinear systems and thus not applicable to AFS. Generalisation of NOCOS to globally nonlinear scenarios is deemed necessary to suit AFS applications;

(3) Existing NOCOS methods are also subject to issues related to stability and real-time implementation that are operationally undesired for AFS applications, and there is potential for further improvement. Arising questions of particular interest are: Could the SPI configuration be further simplified for AFS in the interests of reliable and efficient real-time implementation? Would a different form of VFA give a different stability result under the SPI framework?

Chapter 3

Experimental aeroelastic system

This chapter is based on *Article-1*, which introduces the experimental test-bed specially developed for active flutter suppression (AFS) tests in Chapters 5 and 6 that follow. The test-bed also sets out the aeroelastic model as well as the associated parameters used in the AFS controller derivation and synthesis. It is essential that the test-bed is reliable in providing trustful results, and therefore, this chapter describes how the VSDS was developed, presents and explains special approaches taken to ensure reliable performance. The work fulfils the first objective (Page 4) of the thesis.

Statement of Authorship

Title of Paper	Developing a virtual stiffness-damping system for airfoil aeroelasticity testing
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Tang D, Chen L, Tian ZF and Hu E (2019) Developing a virtual stiffness-damping system for airfoil aeroelasticity testing. *Journal of Sound and Vibration,* 10.1016/j.jsv.2019.115061.

Principal Author

Name of Principal Author (Candidate)	Difan Tang		
Contribution to the Paper	Proposing and deriving methods and theories, performing theoretical and numerical analysis, conducting experiments, collecting data, interpreting results, and writing manuscript.		
Overall percentage (%)			
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	11 November 2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Lei Chen		
Contribution to the Paper	Supervising development of work, evaluating methods and theories, helping with setting up experiments, providing advice in mathematical aspects, and evaluating manuscript.		
Signature		Date	11/11/19

Name of Co-Author	Zhao Feng Tian		
Contribution to the Paper	Supervising development of work, providing advice in mathematical aspects, and evaluating manuscript.		
Signature		Date	11/11/2019

Name of Co-Author	Eric Hu		
Contribution to the Paper	Supervising development of work, providing advice in mathematical aspects, and evaluating manuscript.		
Signature		Date	11/11/2019

abstract

In this research a two-degrees-of-freedom (2-DOF) virtual stiffness-damping system (VSDS) is developed to facilitate industrial and laboratory testing of aerofoil aeroelasticity instability. Other existing test-beds in this field rely on elastic elements or structures to set aerofoil elasticity in tests, which can be costly and inconvenient in cases of frequent stiffness adjustment across a wide range. A possible alternative is the VSDS that utilises electric drives to simulate structural elasticity and damping, as seen in marine and bio-mechanical engineering, which however, cannot be directly applied to aerofoil aeroelasticity testing (AAT) due to operation requirements and conditions being different. Therefore, in this study a new VSDS is developed specifically for AAT. Firstly, the concept of 1-DOF VSDS is extended to 2 DOFs, with the dynamics coupling between each DOF addressed at the stage of operation principle determination, by the proposed direct force/torque regulation with force/torque feedback. Secondly, resolution loss in position/velocity measurement is identified as a main problem associated with the non-reduction transmission required, and is solved by a modified extended-state observer (MESO) proposed for fast position/velocity estimation. Thirdly, system identification and calibration procedures involved in developing the new VSDS are reduced to minimum through a robust force/torque tracking controller design, with detailed numerical study on parametric analysis given. As validated in wind-tunnel experiments the new VSDS can closely track the desired force/torque and provide satisfactory virtual stiffness and damping in AAT.

3.1 Introduction

Aeroelasticity studies the interaction between aerodynamic loads and elastic bodies (Fung, 1955). Aerofoil flutter, as typical dynamic aeroelastic instability, can cause aerofoil structure fatigue and failure. Although advances in computing has allowed more convenient investigation on various aspects of aerofoil aeroelasticity instability as seen in many studies (Lee et al., 2005; Lum et al., 2017; Mardanpour et al., 2018; Price and Keleris, 1996; Rezaei et al., 2018; Sales et al., 2018; Tang and Dowell, 1993), experimental approaches remains essential in terms of discovering new phenomena (Vasconcellos et al., 2016b; Venkatramani et al., 2017, 2018), providing data for modelling and validation (Abdelkefi et al., 2012; Ghiringhelli et al., 1987; Khalil et al., 2016; Lum et al., 2017; Popescu et al., 2009; Price and Keleris, 1996; Sedaghat et al., 2001; Tang and Dowell, 2006; Tang et al., 2004; Tang and Dowell, 1993), as well as validating active control strategies (Burnett et al., 2016; Farmer, 1982; Lum et al., 2017; Miller, 1988; O'Neil and Strganac, 1998). For aerofoil aeroelasticity testing (AAT), numerous facilities have been built, and well established experimental test-beds include the benchmark active control technology (Farmer, 1982), active flexible wing wind-tunnel prototype (Miller, 1988), nonlinear aeroelastic test apparatus (O'Neil and Strganac, 1998), and the newly developed Lockheed Martin X-56A demonstrator (Burnett et al., 2016). Development and use of these existing test-beds can be costly and time-consuming when a wide range of stiffness needs to be investigated, requiring different materials or structures for scaled prototypes (Burnett et al., 2016; Miller, 1988) or different physical springs for typical aerofoil sections (Farmer, 1982; O'Neil and Strganac, 1998). A more efficient and low-cost alternative arose in marine engineering, using electric drives to physically simulate the effects of springs and dampers. The pioneer work was an apparatus for investigating the forces acting on marine cables due to vortex induced vibrations (VIV) (Hover et al., 1997, 1998). It utilises position control to mimic the dynamic behaviours of the subject under the effects of virtual stiffness and damping, with

the reference position trajectory computed from force measurement in real time. In the work of Hover et al. (1997), for the operation frequency range of interest, the position control loop was well tuned and brought less than 1 degree phase lag in displacement, but an additional 12 degrees phase difference between displacement and force was introduced by signal filtering. The phase loss due to signal filtering was reduced to 5 degrees in a later work (Hover et al., 1998). For the study of wake induced vibrations, a similar platform was built and used in Derakhshandeh et al. (2015), where the additional phase lag resulted from signal filtering was further reduced to less than 4 degrees for 0.9 Hz vibration frequency. Another direction towards replacing physical springs with virtual stiffness and damping does not require force measurement and position control, and thus minimises the delay in generating commanded forces, by deriving an explicit formula for direct calculation of the required motor torque output according to the desired stiffness and damping (Lee and Bernitsas, 2011; Lee et al., 2011; Sun et al., 2015). Satisfactory performance is achieved with extensive system identification procedures and calibration (Lee and Bernitsas, 2011; Lee et al., 2011; Sun et al., 2015). Some other related studies can also be seen in bio-mechanical applications, with corresponding purposes, requirements, and operation conditions being much different (Kelly, 2013). For unified naming and convenience of reference hereinafter, these approaches are collectively termed as the 'virtual stiffness-damping system' (VSDS). It is worth noting that

- The existing VSDSs are of one degree-of-freedom (DOF) whereas at least two DOFs are required to capture the dynamics of an aeroelastic system. The coupled dynamics between each DOF is one of the concerns to be addressed in developing a 2-DOF VSDS for AAT;

- Reduction transmission such as lead-screw drive and reduction gearbox are commonly used in existing VSDSs, where the tendency of back-driving by external loads is low or none. This however, is not preferred in AAT, where back-driving is desired to allow easy perturbation of each DOF by aerody-

namic loads as one of the flutter initiation/triggering conditions. Accordingly, potential problems arising from the use of non-reduction transmission on VSDS should be identified and solved.

- Power loss due to inertial loads and various frictions across the entire transmission of VSDSs requires proper treatments, which affect position controller design and tuning (Derakhshandeh et al., 2015; Hover et al., 1997, 1998) or derivation of explicit force generation formulas (Lee and Bernitsas, 2011; Lee et al., 2011; Sun et al., 2015). In particular, frictions are generally modelled and estimated via offline system identification procedures, and can be inaccurate in the presence of uncertainties and parameter change over time or under different operation conditions. In the interests of enhanced robustness and ease of maintenance with simplified system identification and calibration procedures, a new VSDS capable of online estimation of power loss is desired.

Therefore, in this study a new VSDS is developed specifically for AAT, addressing the aforementioned three main concerns, as major technical contributions. Some preliminary works can be found from Tang et al. (2019a,b). In the following: Section 3.2 briefly describes a 2-DOF aeroelastic model on which the new VSDS is based. Section 3.3 presents the proposed operation principle, mechanical design, and control methodology for the new VSDS to tackle coupled dynamics, solve problems associated with non-reduction transmission configuration, and address requirements on enhanced robustness and ease of maintenance with simplified system identification and calibration procedures. Section 3.4 gives detailed numerical analysis of the VSDS controller in terms of system sensitivity and stability robustness against different controller parameters for satisfactory implementation. Section 3.5 describes how the new VSDS was experimentally validated. Conclusions are drawn in Section 3.6.

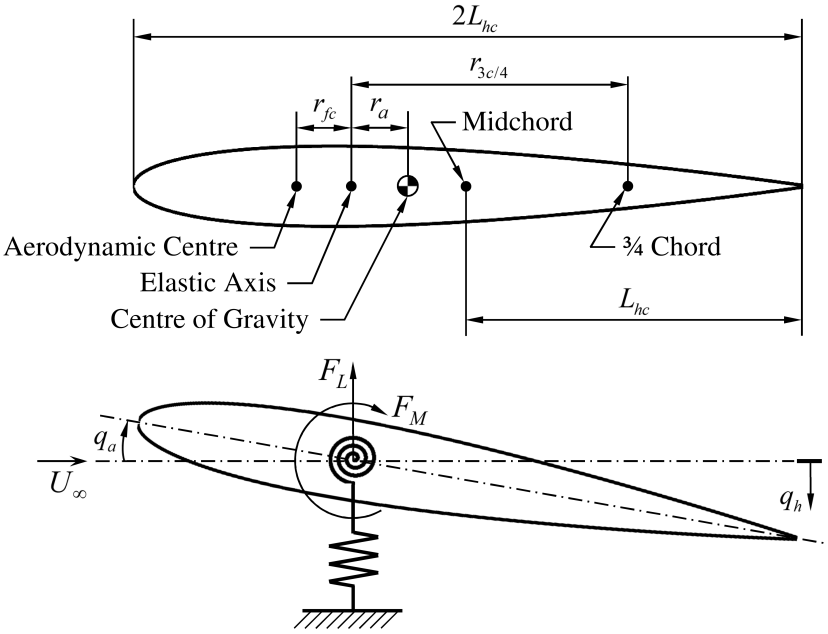


Figure 3.1: Schematic diagram of a 2-DOF aeroelastic system.

3.2 Aeroelastic model

Aeroelastic flutter, viewed from a rigid aerofoil section attached to elastic elements, can be described by oscillations in plunge and pitch DOFs. The corresponding 2-DOF aeroelastic system is illustrated in Figure 3.1 and in this study we consider a subsonic-regime model (O'Neil and Strganac, 1998) with an equation of motion taking the form of

$$\boldsymbol{m}\ddot{\boldsymbol{q}} + \boldsymbol{c}\dot{\boldsymbol{q}} + \boldsymbol{k}\boldsymbol{q} + \boldsymbol{F}_c = \boldsymbol{F}_{aero}, \tag{3.1}$$

with

$$\boldsymbol{q} = \begin{bmatrix} q_h \\ q_a \end{bmatrix}, \quad \boldsymbol{F}_{aero} = \begin{bmatrix} -F_L \\ F_M \end{bmatrix}, \quad \boldsymbol{F}_c = \begin{bmatrix} -m_a \dot{q}_a^2 r_a L_{hc} \sin(q_a) \\ 0 \end{bmatrix},$$

$$\boldsymbol{m} = \begin{bmatrix} m_a & m_a r_a L_{hc} \cos(q_a) \\ m_a r_a L_{hc} \cos(q_a) & I_a \end{bmatrix},$$

$$\boldsymbol{c} = \begin{bmatrix} c_h & 0 \\ 0 & c_a \end{bmatrix}, \quad \boldsymbol{k} = \begin{bmatrix} k_h & 0 \\ 0 & k_a \end{bmatrix},$$

$$F_L = \rho U_\infty^2 L_{hc} L_s C_{l\text{-}a} \left[q_a + \frac{\dot{q}_h}{U_\infty} + L_{3c/4} \frac{\dot{q}_a}{U_\infty} \right],$$

$$F_M = \rho U_\infty^2 L_{hc}^2 L_s C_{me\text{-}a} \left[q_a + \frac{\dot{q}_h}{U_\infty} + L_{3c/4} \frac{\dot{q}_a}{U_\infty} \right],$$

$$C_{l\text{-}a} = \frac{\partial C_l}{\partial q_a}, \quad C_{m\text{-}a} = \frac{\partial C_m}{\partial q_a}, \quad C_{me\text{-}a} = \frac{r_{fc}}{L_{hc}} C_{l\text{-}a} + 2C_{m\text{-}a},$$

where geometry and force related parameters and variables are defined in Figure 3.1, and other terms are defined as

q_h, q_a: translational/angular displacements;

m_a: aerofoil mass;

I_a: aerofoil rotational inertia about its elastic axis;

k_h, k_a: stiffness coefficients;

c_h, c_a: damping coefficients;

L_s: aerofoil span;

C_l: lift coefficient at the aerofoil aerodynamic centre;

C_m: moment coefficient at the aerofoil aerodynamic centre;

U_∞: airflow velocity;

ρ: air density.

As can be seen from Eq. (3.1), F_L and F_M are modelled by quasi-steady aerodynamics. Using a fully unsteady aerodynamic model is also possible, without affecting the VSDS methodology developed in this chapter. In addition, by special design, the proposed VSDS also suit cases where the 2-DOF aeroelastic model is extended by more DOFs to account for actuator dynamics, such as the 3-DOF aeroelastic model in Prime et al. (2010) or a 4-DOF one in Prime (2010). Explanations are given in Sections 3.3.1 and 3.3.4.

Denote the forces resulted from structural stiffness and damping by $\boldsymbol{F}_s$. Then according to Eq. (3.1), there is

$$\boldsymbol{F}_s(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \boldsymbol{c}\dot{\boldsymbol{q}} + \boldsymbol{k}\boldsymbol{q}. \tag{3.2}$$

In the absence of physical springs, $\boldsymbol{F}_s(\boldsymbol{q}, \dot{\boldsymbol{q}})$ can be generated by a VSDS instead, to be detailed next.

3.3 Virtual stiffness-damping system

3.3.1 Operation principle determination

The VSDS for AAT is required to have at least two DOFs, and the first concern to be addressed is the dynamics coupling between each DOF, as mentioned in Section 3.1.

By looking at Equation (3.1), it is clear that the plunge and pitch aeroelastic dynamics are coupled in displacements (q_h and q_a). If position control via force feedback (Derakhshandeh et al., 2015; Hover et al., 1997, 1998) is implemented to deliver equivalent virtual stiffness and damping, the reference position needs to be calculated in real-time, where the coupling from inertial and aerodynamic terms must be properly treated. A direct result is that the accuracy of simulated stiffness and damping is sensitive to modelling errors. Note that Eq. (3.1) approximates general nonlinear aeroelastic systems in a second-order sense, which means resulted modelling errors can be the direct cause of inaccuracy in virtual stiffness and damping delivered by position control.

It can be seen from Eq. (3.1) that the sum $c_h\dot{q}_h + k_h q_h$ and $c_a\dot{q}_a + k_a q_a$ (see Eq. (3.2) for the corresponding matrix form), each considered as a single variable, are mutually independent between the two DOFs. If the stiffness and damping of the aerofoil structure are physically simulated by taking direct control on corresponding force/torque generation according to the two independent sums, then subsequent mechanical design and control synthesis will not be affected by the dynamics coupling in plunge displacement and pitch angle, and thus not affected by the aerodynamic model used (more details follow in Section 3.3.4). In this case it means, the quasi-steady aerodynamic model employed in this chapter can be simply replaced by a fully unsteady aerodynamic model if needed. Moreover, the number of DOFs does not affect the effectiveness of this VSDS methodology since force/torque control is governed by Eq. (3.2) and irrelevant to the additional DOFs that model aerofoil actuator dynamics. To allow direct force/torque control according to Eq. (3.2), the computed force/torque control without the need for force/torque feedback (Lee

and Bernitsas, 2011; Lee et al., 2011; Sun et al., 2015) fits.

On the other hand however, to meet the requirements on enhanced robustness and ease of maintenance with simplified system identification and calibration procedures, it is desired to close the control loop with force/torque measurement feedback. This drives the VSDS operation principle into a new direction – direct force/torque regulation with force/torque feedback. Nevertheless, the additional phase lag between displacement/velocity and force/torque becomes another rising concern that requires special attention in subsequent control system development (to be quantitatively detailed at the implementation stage in Section 3.5).

3.3.2 Mechanical design

Following the operation principle proposed in Section 3.3.1, a new VSDS prototype is developed in this study. An overview of the computer-aided-design (CAD) model is given in Figure 3.2, where the setup is intended for airflow going along the X axis while the plunge-DOF path and pitch-DOF axis are aligned with the Y and Z axes, respectively. Structural details of the VSDS are shown in Figure 3.3. The plunge and pitch DOFs are each driven by an electric motor with an embedded encoder for position feedback. Motor shaft rotation is converted to linear displacement for the plunge DOF via two pairs of synchronous-belt transmission, and the sliding carriage (on roller bearings) can travel along a pair of parallel linear rails. It is worth noting that in real-world scenarios where flutter and active control are considered, the displacement and velocity of a certain point of interest on actual aerofoils can also be obtained from accelerometers, strain gauges, or laser measurements.

To allow back-driving with least resistance so that the plunge DOF can be freely perturbed by aerodynamic loads, speed amplification instead of reduction is introduced via pulleys P1, P2, and P3 as arranged in Figure 3.3(a). Similarly, the pitch-DOF motor is directly connected to the pitching shaft (elastic axis) of the aerofoil without any speed reduction mechanism (Figure 3.3(b)). As discussed in Section 3.1, this transmission configuration is a first trial for VSDS in the field of

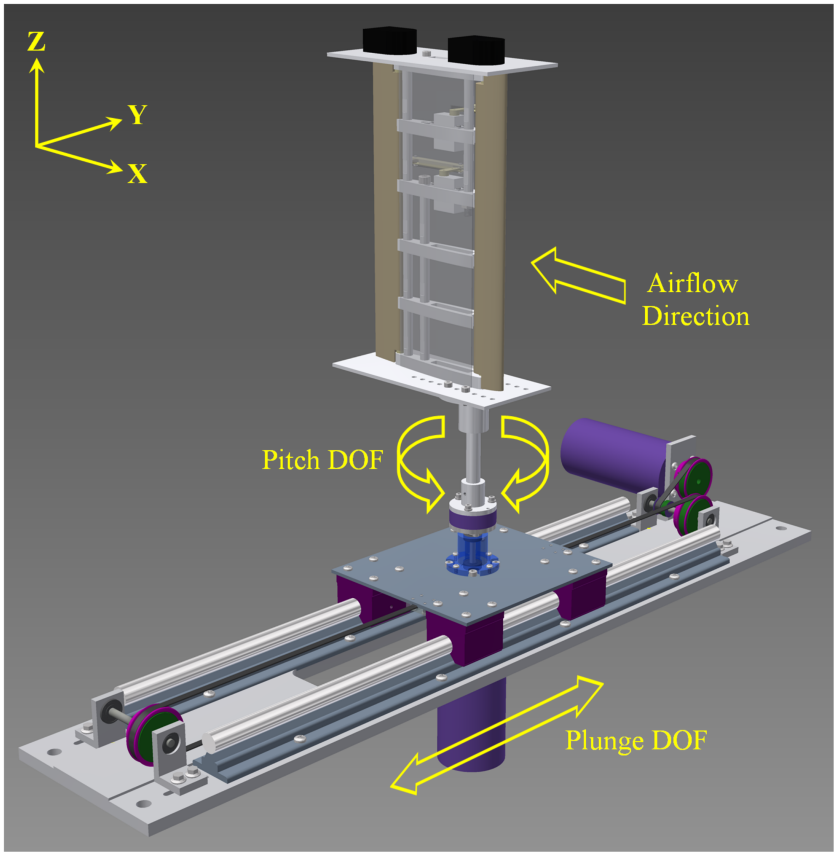
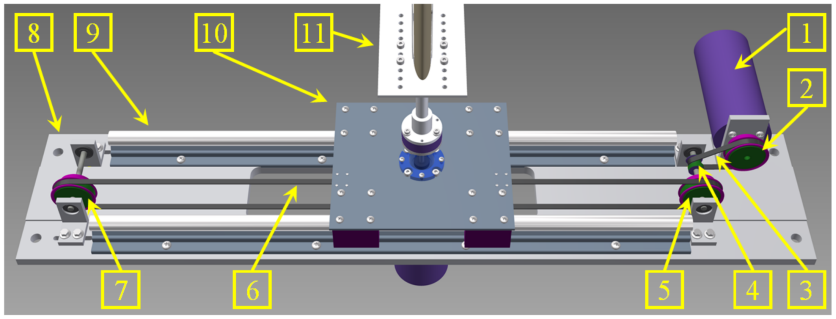


Figure 3.2: Overview of the CAD modelling of the new VSDS prototype.

AAT and hence associated potential problems need to be identified and addressed (to be covered next in Section 3.3.3).

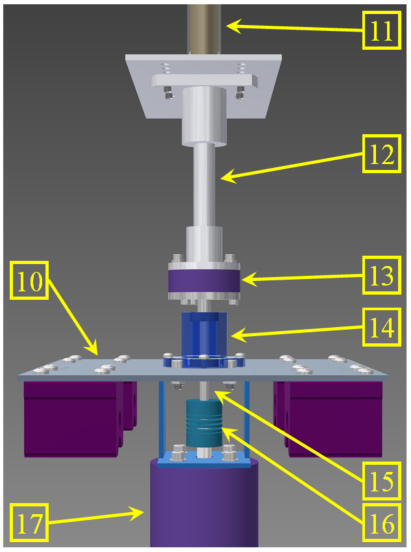
3.3.3 Velocity measurement and estimation

It has been known that velocity measurement using encoders suffers from loss of precision at low velocities alternating around and crossing zero, and the problem is more obvious for low-resolution encoders (Lorenz and Van Patten, 1991; Petrella et al., 2007; Shi et al., 2015). In the case of the proposed VSDS, the non-reduction transmission on the other hand also results in some extent of resolution loss (achievable resolution being 0.2mm for plunge DOF and 0.18 degrees for pitch DOF) despite high-resolution encoders used. This leads to imprecise measurement of low velocities. Some possible remedies to obtain improved velocity measurements are given in Lorenz and Van Patten (1991), Petrella et al. (2007), and Shi et al. (2015), which however, are not ideal solutions to the proposed VSDS due to the dynamics coupling



(a) Plunge DOF (perspective top-front view)

1, plunge-DOF motor
2, pulley P1 (44 grooves)
3, synchronous belt B1
4, pulley P2 (14 grooves)
5, pulley P3 (44 grooves)
6, synchronous belt B2
7, pulley P4 (44 grooves)
8, main platform
9, linear rail
10, sliding carriage
11, airfoil
12, pitching shaft (airfoil)
13, force/torque transducer
14, shaft housing
15, pitching shaft (VSDS)
16, coupler
17, pitch-DOF motor



(b) Pitch DOF
(perspective front view)

Figure 3.3: Structural details of the new VSDS prototype.

in plunge/pitch displacements. Motivated by the concept of the extended-state observer (ESO) (Han, 2009) in applications with strong nonlinearities (Chang et al., 2015; Erenturk, 2013; Herbst, 2016; Li et al., 2014; Yang et al., 2018; Yuan et al., 2017), we derived a modified extended-state observer (MESO) for velocity estimation. Let $\boldsymbol{\vartheta}_1 = \boldsymbol{q}$, $\boldsymbol{\vartheta}_2 = \dot{\boldsymbol{q}}$. Then Eq. (3.1) can be rewritten as

$$\begin{cases} \dot{\boldsymbol{\vartheta}}_1 = \boldsymbol{\vartheta}_2, \\ \dot{\boldsymbol{\vartheta}}_2 = \boldsymbol{m}^{-1}(\boldsymbol{F}_{aero} - \boldsymbol{F}_c - \boldsymbol{c}\dot{\boldsymbol{q}} - \boldsymbol{k}\boldsymbol{q}), \end{cases} \tag{3.3}$$

Considering the presence of modelling errors, we have

$$\begin{cases} \dot{\boldsymbol{\vartheta}}_1 = \boldsymbol{\vartheta}_2, \\ \dot{\boldsymbol{\vartheta}}_2 = \mathcal{F}(\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2), \end{cases} \tag{3.4}$$

where

$$\mathcal{F}(\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2) = \boldsymbol{m}^{-1}(\boldsymbol{F}_{aero} - \boldsymbol{F}_c - \boldsymbol{c}\dot{\boldsymbol{q}} - \boldsymbol{k}\boldsymbol{q}) + \boldsymbol{\zeta}(t, \boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2),$$

with $\boldsymbol{\zeta}(t, \boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2)$ containing un-modelled dynamics.

Let $\hat{\boldsymbol{\vartheta}}_1$, $\hat{\boldsymbol{\vartheta}}_2$, and $\hat{\boldsymbol{\vartheta}}_3$ denote the estimate of $\boldsymbol{\vartheta}_1$, $\boldsymbol{\vartheta}_2$, and $\mathcal{F}(\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2)$ respectively, and $\boldsymbol{e}_\vartheta = \hat{\boldsymbol{\vartheta}}_1 - \boldsymbol{q}$. The MESO for Eq. (3.3) is constructed as

$$\begin{cases} \dot{\hat{\boldsymbol{\vartheta}}}_1 = \hat{\boldsymbol{\vartheta}}_2 - p_1 \boldsymbol{e}_\vartheta, \\ \dot{\hat{\boldsymbol{\vartheta}}}_2 = \hat{\boldsymbol{\vartheta}}_3 - p_2 \dfrac{\boldsymbol{e}_\vartheta}{\sqrt{\boldsymbol{e}_\vartheta^{\mathrm{T}} \boldsymbol{e}_\vartheta + 1}}, \\ \dot{\hat{\boldsymbol{\vartheta}}}_3 = -p_3 \dfrac{\boldsymbol{e}_\vartheta}{\sqrt{\boldsymbol{e}_\vartheta^{\mathrm{T}} \boldsymbol{e}_\vartheta + 1}}, \end{cases} \tag{3.5}$$

where $p_1 \in \mathbb{R}^+$, $p_2 \in \mathbb{R}^+$, and $p_3 \in \mathbb{R}^+$ are design parameters of choice.

The use of MESO can introduce phase lag between position/velocity measurements and estimations. As to be covered in Section 3.5, the induced phase lag can be minimised by carefully adjusting the design parameters and do not have noticeable negative impact on VSDS performance.

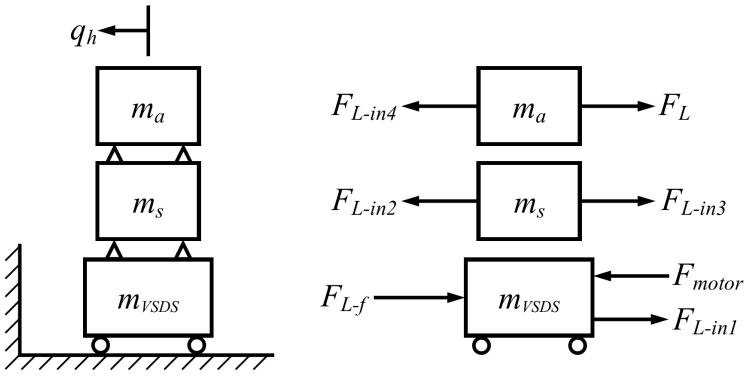
3.3.4 Force/torque measurement and control

For force/torque measurement, a 6-axis force/torque transducer (ATI® Mini40) is mounted between the pitching shafts of the VSDS and the aerofoil test section, with its Z axis aligned with the Z axis (pitching axis) of the CAD model. Under this arrangement, the aerofoil body and its pitching shaft together are considered as one whole rigid body, the total mass of which is m_a according to Eq. (3.1).

With the force/torque transducer regarded as an independent rigid body, the plunge DOF can be considered as a multi-body system illustrated in Figure 3.4(a), which has the same reference frame as Figure 3.1. Corresponding forcing diagram on isolated multiple bodies is given in Figure 3.4(b). Accordingly,

$$m_a \ddot{q}_h + F_{L\text{-}in4} = F_L, \tag{3.6}$$

$$m_s \ddot{q}_h + F_{L\text{-}in2} = F_{L\text{-}in3}, \tag{3.7}$$



(a) Reference frame (b) Forces on isolated bodies

Figure 3.4: The VSDS plunge DOF as a multi-body system.

$$m_{VSDS}\ddot{q}_h + F_{L\text{-}f} + F_{L\text{-}in1} = F_{motor}, \tag{3.8}$$

$$F_{L\text{-}in1} = F_{L\text{-}in2}, \tag{3.9}$$

$$F_{L\text{-}in3} = F_{L\text{-}in4}, \tag{3.10}$$

where F_{motor} is the translational force (tension) on the synchronous belt B1 with Pulleys P1 and P2; $F_{L\text{-}f}$ is the total equivalent friction along the entire plunge-DOF power transmission from motor output to the sliding carriage, consisting of viscous, Coulomb, and Stribeck frictions; $F_{L\text{-}in1}$, $F_{L\text{-}in2}$, $F_{L\text{-}in3}$ and $F_{L\text{-}in4}$ are internal forces defined according to Newton's 3rd law of motion; m_s is mass of the force/torque transducer, and m_{VSDS} is the equivalent mass of the plunge-DOF transmission.

Let $\bar{k}_h$, $\bar{c}_h$, $\bar{k}_a$, and $\bar{c}_a$ denote the coefficient of simulated virtual stiffness and damping of respective DOFs. Comparing Eqs. (3.1) and (3.6) gives

$$F_{L\text{-}in4} = -\bar{c}_h\dot{h} - \bar{k}_h h. \tag{3.11}$$

Similar forcing applies to the pitch DOF (related forcing diagram and equations are not repeated herein) and we have

$$F_{M\text{-}in4} = \bar{c}_a\dot{q}_a + \bar{k}_a q_a, \tag{3.12}$$

where $F_{M\text{-}in4}$ is the internal torque that contributes to

$$I_a\ddot{q}_a + F_{M\text{-}in4} = F_M. \tag{3.13}$$

Let $\bar{\boldsymbol{F}}_s \triangleq [-F_{L\text{-}in4} \quad F_{M\text{-}in4}]^{\mathrm{T}}$. Then writing Eqs. (3.11) and (3.12) in matrix form yields

$$\bar{\boldsymbol{F}}_s(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \bar{\boldsymbol{c}}\dot{\boldsymbol{q}} + \bar{\boldsymbol{k}}\boldsymbol{q}, \tag{3.14}$$

where $\bar{\boldsymbol{c}} = \mathrm{diag}(\bar{c}_h, \bar{c}_a)$ and $\bar{\boldsymbol{k}} = \mathrm{diag}(\bar{k}_h, \bar{k}_a)$.

According to the structure of the force/torque transducer together with Eqs. (3.6) and (3.13), force/torque measurements closely take the value of $\bar{\boldsymbol{F}}_s$. Note that $\bar{\boldsymbol{k}}$ and $\bar{\boldsymbol{c}}$ need not be explicitly approximated, and properly controlling the motor torque can make $\bar{\boldsymbol{F}}_s \to \boldsymbol{F}_s$, as suggested by Eq. (3.8). This means the VSDS control system is decoupled from aerodynamic terms and thus not affected by the types of aerodynamic model used. Moreover, it can also be seen that increasing the number of DOFs to account for actuator dynamics does not have any impact on the effectiveness of the proposed method. Note that precise tracking of the trajectories of $\boldsymbol{F}_s$ requires correct knowledge and proper compensation of power loss caused by frictions as well as other un-modelled dynamics and exogenous disturbances. Without force/torque feedback, extensive system identification procedures are necessary (Lee and Bernitsas, 2011; Lee et al., 2011; Sun et al., 2015). To meet requirements on enhanced robustness and ease of maintenance with simplified system identification and calibration procedures, the total power loss can be dynamically estimated online by evaluating $\bar{\boldsymbol{F}}_s$ against reference $\boldsymbol{F}_s$. For this purpose, the VSDS control system is proposed as in Figure 3.5.

Accurately estimating power loss due to inertial loads and mechanical frictions has been shown challenging (Lee and Bernitsas, 2011; Lee et al., 2011; Sun et al., 2015), and this also raises high requirements to the VSDS controller. To assist VSDS controller synthesis, dynamics of both DOFs of the VSDS prototype were obtained via black-box system identification with voltages as inputs and force/torque measurement $\bar{\boldsymbol{F}}_s$ as outputs. Chirp signals of 2 to 5 Hz were used to explore up to 5% and 60% rated capacity (sufficient for required force/torque) of pitch- and plunge-DOF motors, respectively. Estimated models are in state-space linear time-invariant form

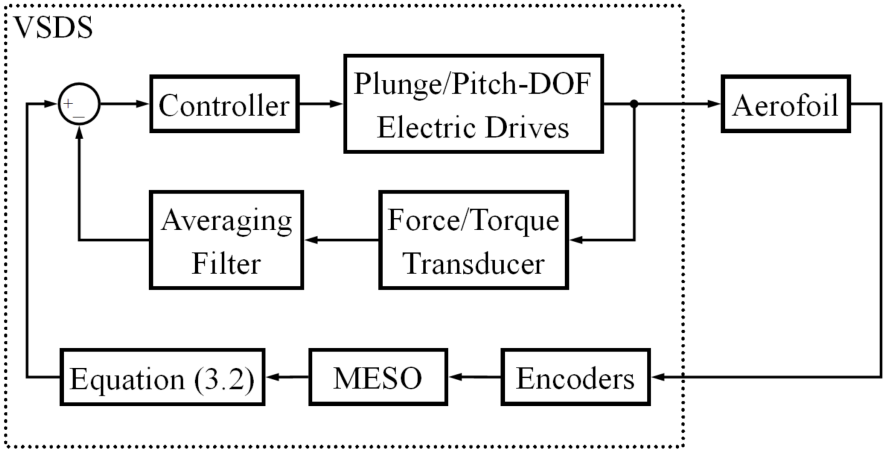


Figure 3.5: Schematic diagram of the proposed VSDS control system.

as

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}_s \boldsymbol{x} + \boldsymbol{B}_s \boldsymbol{u}, \\ \bar{\boldsymbol{F}}_s = \boldsymbol{C}_s \boldsymbol{x} + \boldsymbol{D}_s \boldsymbol{u}, \end{cases} \tag{3.15}$$

where $\boldsymbol{x} \in \mathbb{R}^{n_x}$ contains n_x system states; $\bar{\boldsymbol{F}}_s \in \mathbb{R}^{n_F}$ contains n_F measured outputs; $\boldsymbol{u} \in \mathbb{R}^{n_u}$ contains n_u control inputs; $\boldsymbol{A}_s \in \mathbb{R}^{n_x \times n_x}$, $\boldsymbol{B}_s \in \mathbb{R}^{n_x \times n_u}$, $\boldsymbol{C}_s \in \mathbb{R}^{n_F \times n_x}$, and $\boldsymbol{D}_s \in \mathbb{R}^{n_F \times n_u}$ are system matrices.

Power loss due to inertial loads and mechanical frictions are partially captured in the identified models (Table. 3.1). However, modelling accuracy is insufficient for direct force/torque regulation without force/torque feedback, as the system identification procedures involved are reduced to minimum compared with those in Lee et al. (2011), Lee and Bernitsas (2011), and Sun et al. (2015). Note that modelling errors and other types of disturbances generally affect a system in the form of unknown non-control inputs (collectively termed as 'unknown inputs'). For satisfactory control, these unknown inputs need to be properly treated, and Eq. (3.15) can be reformulated as

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}_s \boldsymbol{x} + \boldsymbol{B}_s \boldsymbol{u} + \boldsymbol{B}_d \boldsymbol{d}, \\ \bar{\boldsymbol{F}}_s = \boldsymbol{C}_s \boldsymbol{x}, \end{cases} \tag{3.16}$$

where $\boldsymbol{d} \in \mathbb{R}^{n_d}$ contains n_d unknown inputs perturbing the system, and $\boldsymbol{B}_d \in \mathbb{R}^{n_x \times n_d}$ is the unknown-input distribution matrix.

Table 3.1: Experimentally identified dynamics of the VSDS prototype.

DOF	Matrices of System Dynamics				Fit (%)
	$\boldsymbol{A}_s$	$\boldsymbol{B}_s$	$\boldsymbol{C}_s$	$\boldsymbol{D}_s$	
Plunge	$\begin{bmatrix} -8.791 & 364 \\ -273.8 & -165.9 \end{bmatrix}$	$\begin{bmatrix} 45.5 \\ 29.19 \end{bmatrix}$	$\begin{bmatrix} 25.53 \\ 3.903 \end{bmatrix}^{\mathrm{T}}$	0	93.88
Pitch	$\begin{bmatrix} 18.67 & -464 \\ 429.1 & -716.3 \end{bmatrix}$	$\begin{bmatrix} 5823 \\ 7501 \end{bmatrix}$	$\begin{bmatrix} 0.4412 \\ -0.0669 \end{bmatrix}^{\mathrm{T}}$	0	94.04

Since the models in Table. 3.1 are controllable, observable, and on the imaginary axis there is no zeros, according to She et al. (2008), an equivalent of $\boldsymbol{d}$ exists which enters the system via $\boldsymbol{B}_s$. Therefore, system Eq. (3.16) is equivalent to

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}_s \boldsymbol{x} + \boldsymbol{B}_s \left(\boldsymbol{u} + \boldsymbol{d}_e \right), \\ \bar{\boldsymbol{F}}_s = \boldsymbol{C}_s \boldsymbol{x}, \end{cases} \tag{3.17}$$

where $\boldsymbol{d}_e \in \mathbb{R}^{n_u}$ contains the equivalent unknown inputs.

As discussed in Section 3.2 and as can be seen from Figure 3.5, the VSDS needs to generate $\boldsymbol{F}_s(\boldsymbol{q}, \dot{\boldsymbol{q}})$ in the absence of physical springs. Hence, it is a force/torque tracking problem, where the reference trajectories $\boldsymbol{F}_s(\boldsymbol{q}, \dot{\boldsymbol{q}})$ change in real time with respect to $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$ in accordance with Eq. (3.2). To deal with transmission power-loss, the linear-quadratic-Gaussian (LQG) tracking control enhanced by unknown-input estimation (UIE) as in Tang et al. (2016) is employed for the VSDS prototype, given the existence of an equivalent system as in Eq. (3.17). The controller has a structure illustrated in Figure 3.6, which consists of a standard LQG tracking component and a UIE add-on. The LQG control provides nominal force/torque tracking, while the UIE estimates and compensates transmission power-loss. Detailed composition of the controller, without distinguishing between specific DOFs, is given below in a general multiple-input multiple-output form.

The total control $\boldsymbol{u}$ is

$$\boldsymbol{u} = \boldsymbol{u}_c - \hat{\boldsymbol{d}}_e, \tag{3.18}$$

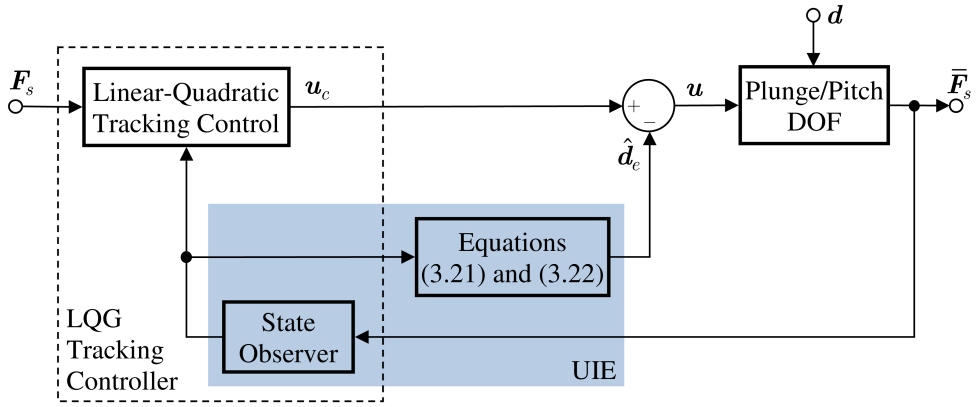


Figure 3.6: Schematic diagram of the UIE-LQG controller.

where $\boldsymbol{u}_c$ is from the LQG component for trajectory tracking, and $\hat{\boldsymbol{d}}_e$ is the estimated equivalent power loss.

The nominal control $\boldsymbol{u}_c$ takes the form of

$$\boldsymbol{u}_c = -\boldsymbol{K}_x\hat{\boldsymbol{x}} - \boldsymbol{K}_w\boldsymbol{x}_w + \boldsymbol{K}_r\boldsymbol{F}_s(\boldsymbol{q}, \dot{\boldsymbol{q}}), \tag{3.19}$$

with

$$\dot{\boldsymbol{x}}_w = \boldsymbol{F}_s(\boldsymbol{q}, \dot{\boldsymbol{q}}) - \bar{\boldsymbol{F}}_s = \boldsymbol{F}_s(\boldsymbol{q}, \dot{\boldsymbol{q}}) - \boldsymbol{C}_s\boldsymbol{x}, \tag{3.20}$$

where $\boldsymbol{K}_x \in \mathbb{R}^{n_u \times n_x}$, $\boldsymbol{K}_w \in \mathbb{R}^{n_u \times n_F}$, and $\boldsymbol{K}_f \in \mathbb{R}^{n_u \times n_F}$ are gains of the LQG tracking control which can be selected following standard LQG design procedures (Anderson and Moore, 1990).

The power-loss estimation $\hat{\boldsymbol{d}}_e$ is obtained via

$$\hat{\boldsymbol{d}}_e = \hat{\boldsymbol{d}}_{ev} + \boldsymbol{K}_d\left(\bar{\boldsymbol{F}}_s - \hat{\boldsymbol{F}}_s\right), \tag{3.21}$$

with

$$\begin{cases} \dot{\boldsymbol{x}}_f = \boldsymbol{A}_f\boldsymbol{x}_f + \boldsymbol{B}_f\hat{\boldsymbol{d}}_e, \\ \hat{\boldsymbol{d}}_{ev} = \boldsymbol{C}_f\boldsymbol{x}_f, \end{cases} \tag{3.22}$$

and

$$\begin{cases} \dot{\hat{\boldsymbol{x}}} = \boldsymbol{A}_s\hat{\boldsymbol{x}} + \boldsymbol{B}_s\boldsymbol{u}_c + \boldsymbol{L}\left(\bar{\boldsymbol{F}}_s - \hat{\boldsymbol{F}}_s\right), \\ \hat{\boldsymbol{F}}_s = \boldsymbol{C}_s\hat{\boldsymbol{x}}, \end{cases} \tag{3.23}$$

where $\boldsymbol{K}_d \in \mathbb{R}^{n_u \times n_F}$ is the UIE gain; $\hat{\boldsymbol{F}}_s$ contains estimated system outputs; $\boldsymbol{A}_f \in \mathbb{R}^{n_f \times n_f}$, $\boldsymbol{B}_f \in \mathbb{R}^{n_f \times 1}$, and $\boldsymbol{C}_f \in \mathbb{R}^{1 \times n_f}$ are matrices of a low-pass-filter-characterised subsystem $(\boldsymbol{A}_f, \boldsymbol{B}_f, \boldsymbol{C}_f)$ with n_f states; $\boldsymbol{L} \in \mathbb{R}^{n_x \times 1}$ is the state observer gain; $\hat{\boldsymbol{x}}$ contains estimated system states.

The UIE gain $\boldsymbol{K}_d$ can be calculated via linear-quadratic optimisation based on the dynamics of states estimation:

$$\begin{cases} \dot{\boldsymbol{e}}_x = (\boldsymbol{A}_s - \boldsymbol{L}\boldsymbol{C}_s)\boldsymbol{e}_x - \boldsymbol{B}_s\boldsymbol{K}_d\boldsymbol{C}_s\boldsymbol{e}_x + \boldsymbol{e}_d, \\ \boldsymbol{e}_F = \boldsymbol{C}_s\boldsymbol{e}_x, \end{cases} \tag{3.24}$$

where $\boldsymbol{e}_F = \bar{\boldsymbol{F}}_s - \hat{\boldsymbol{F}}_s$ and $\boldsymbol{e}_d = \boldsymbol{d}_e - \hat{\boldsymbol{d}}_{ev}$.

Upon $\boldsymbol{e}_d \approx 0$, we have

$$\dot{\boldsymbol{e}}_x = (\boldsymbol{A}_s - \boldsymbol{L}\boldsymbol{C}_s - \boldsymbol{B}_s\boldsymbol{K}_d\boldsymbol{C}_s)\boldsymbol{e}_x. \tag{3.25}$$

For observable systems, the pair $(\boldsymbol{A}_s - \boldsymbol{L}\boldsymbol{C}_s,\ \boldsymbol{C}_s)$ is observable. Under duality, there exists $\boldsymbol{K}_v = (\boldsymbol{B}_s\boldsymbol{K}_d)^{\mathrm{T}}$ that minimises

$$V_d = \int_0^\infty \left\{ \boldsymbol{e}_x^{\mathrm{T}}\boldsymbol{Q}_d\boldsymbol{e}_x + \boldsymbol{e}_{dv}^{\mathrm{T}}\boldsymbol{R}_d\boldsymbol{e}_{dv} \right\} dt, \tag{3.26}$$

with $\boldsymbol{Q}_d \in \mathbb{R}^{n_x \times n_x}$ and $\boldsymbol{R}_d \in \mathbb{R}^{n_F \times n_F}$ being symmetric positive-definite weighting matrices, and $\boldsymbol{e}_{dv} = -\boldsymbol{K}_d\boldsymbol{C}_s\boldsymbol{e}_x$.

Therefore,

$$\boldsymbol{K}_d = \boldsymbol{B}_s^{\dagger}\boldsymbol{K}_v^{\mathrm{T}}, \tag{3.27}$$

with $\boldsymbol{B}_s^{\dagger}$ being the Moore-Penrose pseudo inverse of $\boldsymbol{B}_s$.

As can be seen from Eqs. (3.21) to (3.23), the subsystem $(\boldsymbol{A}_f, \boldsymbol{B}_f, \boldsymbol{C}_f)$, UIE gain $\boldsymbol{K}_d$, and state observer gain $\boldsymbol{L}$ are major design parameters related to the UIE component.

3.4 Controller analysis

System sensitivity and stability robustness, being crucial to estimation and compensation of transmission power-loss, are major concerns in controller synthesis for the VSDS prototype in AAT. Although brief guidelines for selecting UIE related

parameters are available in literature, there is lack of understanding on the influence of these design variables on system sensitivity and stability robustness. As one of the technical contributions, numerical studies on controller parametric analysis are presented in this section.

As can be seen from Eqs. (3.21) to (3.23), the subsystem $(\boldsymbol{A}_f, \boldsymbol{B}_f, \boldsymbol{C}_f)$, UIE gain $\boldsymbol{K}_d$, and state observer gain $\boldsymbol{L}$ are major design parameters related to the UIE component.

Given the single-input single-output (SISO) feature of VSDS plunge/pitch-DOF dynamics (Table. 3.1), the following analysis is performed on an SISO basis.

In frequency domain, Eq. (3.17) takes the form of

$$\bar{F}_s(s) = P_n(s)\left[u(s) + d_e(s)\right], \tag{3.28}$$

where $P_n(s)$ is the nominal model of the plunge/pitch DOF.

In this SISO case, any individual tracking trajectory from the set $\boldsymbol{F}_s(\boldsymbol{q}, \dot{\boldsymbol{q}})$ is denoted by $F_s(s) \subseteq \boldsymbol{F}_s$ for convenience in notation.

With $G_w(s)$ denoting the transfer function of the integral action, Eq. (3.19) can be transformed into

$$u_c(s) = -\boldsymbol{K}_x\hat{\boldsymbol{x}}(s) - K_w x_w(s) + K_r F_s(s), \tag{3.29}$$

with $x_w(s) = G_w(s)\left[F_s(s) - \bar{F}_s(s)\right]$.

By using Eq. (3.29) in Eq. (3.23), and with $F_s(s) = 0$, there is

$$\begin{aligned}\hat{\boldsymbol{x}}(s) &= \left(s\boldsymbol{I}_{n_x} - \boldsymbol{A}_s + \boldsymbol{B}_s\boldsymbol{K}_x + \boldsymbol{L}\boldsymbol{C}_s\right)^{-1}\left[\boldsymbol{B}_s K_w G_w(s) + \boldsymbol{L}\right]\bar{F}_s(s) \\ &= \boldsymbol{G}_{yx}(s)\bar{F}_s(s),\end{aligned} \tag{3.30}$$

where $\boldsymbol{G}_{yx}(s)$ contains transfer functions from $\bar{F}_s(s)$ to $\hat{\boldsymbol{x}}(s)$.

From Eqs. (3.21) and (3.22), we have

$$\hat{d}_{ev}(s) = G_f(s)\hat{d}_e(s), \tag{3.31}$$

$$\hat{d}_e(s) = G_{yd}(s)\left[\bar{F}_s(s) - \hat{F}_s(s)\right], \tag{3.32}$$

where

$$G_f(s) = \boldsymbol{C}_f(s\boldsymbol{I}_{n_f} - \boldsymbol{A}_f)^{-1}\boldsymbol{B}_f\hat{d}_e(s),$$

$$G_{yd}(s) = \left[1 - G_f(s)\right]^{-1} K_d.$$

Based on Eqs. (3.18), (3.28) to (3.30), and (3.32), we reach

$$\bar{F}_s(s) = P_n(s)[1 + H_s(s)P_n(s)]^{-1} d_e(s), \tag{3.33}$$

where

$$H_s(s) = \boldsymbol{K}_x \boldsymbol{G}_{yx}(s) - K_w G_w(s) + G_{yd}(s)\left[1 - \boldsymbol{C}_s \boldsymbol{G}_{yx}(s)\right].$$

Thus, the system sensitivity to unknown inputs (referred to as 'sensitivity' in short hereinafter) is

$$S_s(j\omega) = \frac{1}{1 + H_s(j\omega)P_n(j\omega)}, \quad \forall \omega \in [0, +\infty). \tag{3.34}$$

To allow evaluating sensitivity against different cutoff frequency ω_c of the subsystem $(\boldsymbol{A}_f, \boldsymbol{B}_f, \boldsymbol{C}_f)$, the magnitude of sensitivity with regard to inputs of a certain natural frequency (i.e., $|S_s(j\omega_d)|$) is considered. In the analysis, $\omega_d = 5$ rad/s is considered, according to the spectrum characteristics of system identification data.

In the presence of uncertainties $\delta(j\omega), \forall \omega \in [0, +\infty)$, closed-loop stability is guaranteed if

$$|\delta(j\omega)| < \left|1 + \frac{1}{H_s(j\omega)P_n(j\omega)}\right|, \ \forall \omega \in [0, +\infty). \tag{3.35}$$

As it is common that there is little or no priori knowledge about $\delta(j\omega)$, a conservative choice is to assume $\delta(j\omega) = 1, \forall \omega \in [0, +\infty)$. On this basis, the following stability robustness index (SRI) is proposed:

$$\text{SRI} = \min \left|1 + \frac{1}{H_s(j\omega)P_n(j\omega)}\right| - 1, \ \forall \omega \in [0, +\infty). \tag{3.36}$$

It is worth noting that due to the conservative assumption of $\delta(j\omega)$, a negative SRI is not necessarily a sign of instability. Instead, SRI gives a relative measure on stability robustness. That is, a larger SRI suggests better stability robustness.

To investigate the effects of using a low-pass filter (LPF) of different orders for the subsystem $(\boldsymbol{A}_f, \boldsymbol{B}_f, \boldsymbol{C}_f)$, the following generalised formulation $G_f(s)$ is considered:

$$G_f(s) = \frac{\sum_{i=0}^{p_n} a_{mi}(\tau s)^i}{(\tau s + 1)^m}, \tag{3.37}$$

where m is the denominator order and p_n is the numerator order; $a_{mi} = \frac{m!}{(m-i)!i!}$ is the i^{th} coefficient of the numerator polynomial; τ is the time constant.

Since the closed-loop system is unstable if $p_n \geq 1$, the analysis herein only considers cases where $m \geq 1$ and $p_n = 0$. This yields LPFs with a unity passband gain.

Only the plunge-DOF dynamics of the VSDS prototype are used throughout the analysis, given similar dynamics of both DOFs. LQG parameters exclusive to UIE are kept unchanged throughout the analysis while the UIE-related parameters are varied across regions of interest for investigation. Phase lag introduced by the UIE-LQG controller is to be quantitatively given in Section 3.5.

3.4.1 Low-pass filter cutoff frequency

In this analysis, the UIE component takes parameters of $\boldsymbol{L} = [9.49,\ 1.07]^{\mathrm{T}}$, $K_d = 6.64$, and $1 \leq m \leq 4,\ \forall m \in \mathbb{Z}$, with the LPF cutoff frequency ω_c being varied. Figure 3.7 shows that sensitivity decreases with increased ω_c, when ω_c is times higher than ω_d. A first-order LPF requires the least gap between ω_d and ω_c to achieve useful sensitivity, while a fourth-order LPF introduces a mild peak in sensitivity, and ω_c needs to be 3 to 4 times higher than ω_d for reduced sensitivity. In Figure 3.8, similar stability robustness can be observed among the four filters for $\omega_c < 10^2$ rad/s, all showing relatively good stability robustness. But differences start to grow for $\omega_c > 10^2$ rad/s, where the LPF of first-order is more sensitive to the increase of ω_c, having an earlier drop of stability robustness compared with LPFs of higher orders. Figures 3.7 and 3.8 indicate that a first-order LPF can be a preferred choice from the sensitivity perspective, and is most suitable for low-frequency uncertainties in the interests of stability robustness. Figures 3.7 and 3.8 also recommend that $\omega_c = 100$ rad/s is a relatively better choice that balances sensitivity and stability robustness for the case under discussion. Therefore, $\omega_c = 100$ rad/s is used in the subsequent analysis.

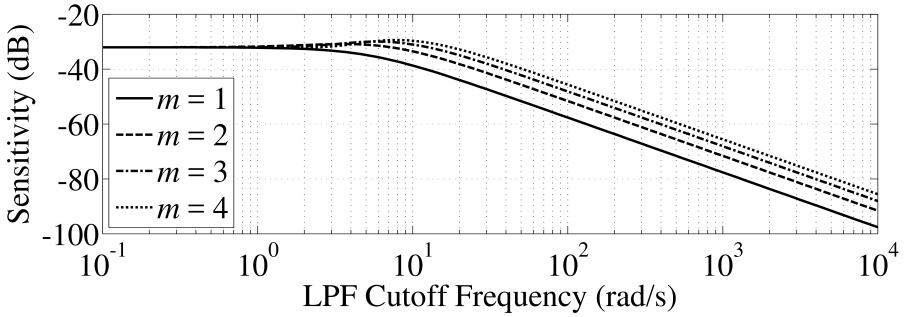


Figure 3.7: Sensitivity at $\omega_d = 5$ rad/s against LPF cutoff frequency.

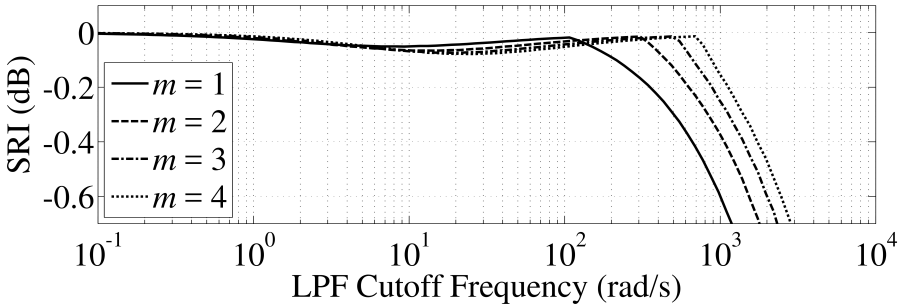


Figure 3.8: SRI against LPF cutoff frequency.

3.4.2 Unknown-input estimation gain

To study the effect of the UIE gain, K_d is considered as a variable while other parameters are set as $\boldsymbol{L} = [9.49,\ 1.07]^{\mathrm{T}}$, $\omega_c = 100$ rad/s, and $1 \leq m \leq 4$, $\forall m \in \mathbb{Z}$. It is straightforward to see from Figure 3.9 that larger K_d contributes to smaller sensitivity, delivering better rejection of unknown inputs. However, a dip can been seen on each of the SRI curves in Figure 3.10, indicating a weak stability robustness region, which should be avoided in design. K_d to the left of this potentially unstable region, being too small to make the UIE component effective, is not preferred. A larger K_d beyond the SRI dip can be considered in UIE design, as the stability robustness recovers to an acceptable level. Hence, Figures 3.9 and 3.10 both support the choice for a larger K_d. LPFs of different orders do not have significant impact on K_d selection from the stability robustness perspective, although some shift of the dip on the SRI curve can be seen between LPFs of different orders. In the interests of better estimation and compensation of transmission power-loss, a first-order LPF can be considered for the relatively smaller sensitivity achieved with the same K_d.

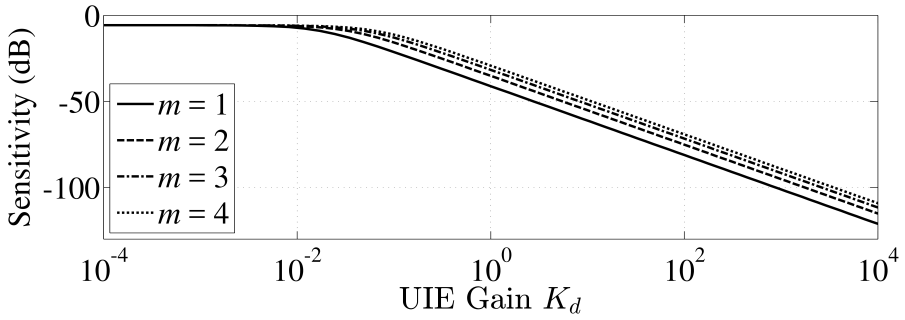


Figure 3.9: Sensitivity at $\omega_d = 5$ rad/s against UIE gain.

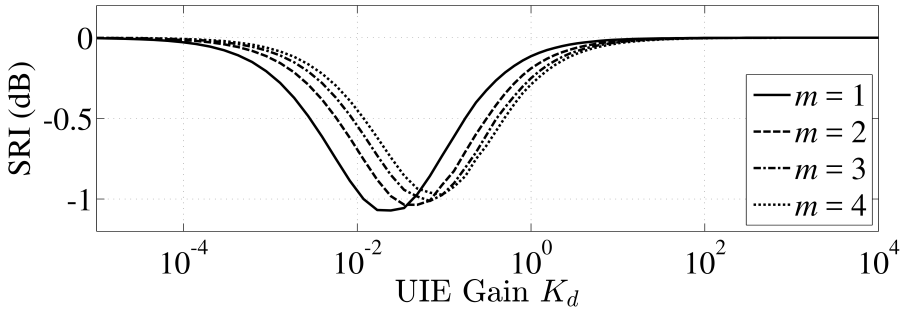


Figure 3.10: SRI against UIE gain.

3.4.3 State observer gain

With parameters $K_d = 6.64$, $\omega_c = 100$ rad/s, and $1 \leq m \leq 4$, $\forall m \in \mathbb{Z}$, the influence of $\boldsymbol{L}$ is evaluated. As can be seen in Figure 3.11, the sensitivity remains at a low level and is insensitive to variation of $\boldsymbol{L}$ in the region where $\|\boldsymbol{L}\|_2 \in (0, 10^2]$, favoured for estimation and compensation of transmission power-loss. $\boldsymbol{L}$ within this range is also acceptable in terms of stability robustness, as shown in Figure 3.12. Continuously increasing $\boldsymbol{L}$ not only weakens unknown-inputs rejection capability with raised sensitivity (Figure 3.11) but can also brings instability issues as indicated by the dip around $\|\boldsymbol{L}\|_2 \in [10^2, 10^5]$ (Figure 3.12). Although stability robustness returns to an acceptable level with much larger $\boldsymbol{L}$, $\boldsymbol{L}$ within this high-value range is nevertheless undesired in consideration of weak rejection of unknown inputs and increased computation load. With regard to the effects of the LPF order, it is easy to see from Figure 3.11 that a first-order LPF appears to be a better choice for the relatively smaller sensitivity, with stability robustness property similar to LPFs of higher orders.

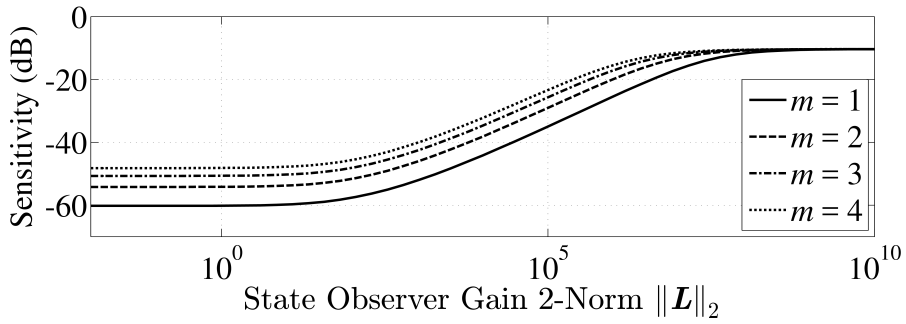


Figure 3.11: Sensitivity at $\omega_d = 5$ rad/s against state observer gain.

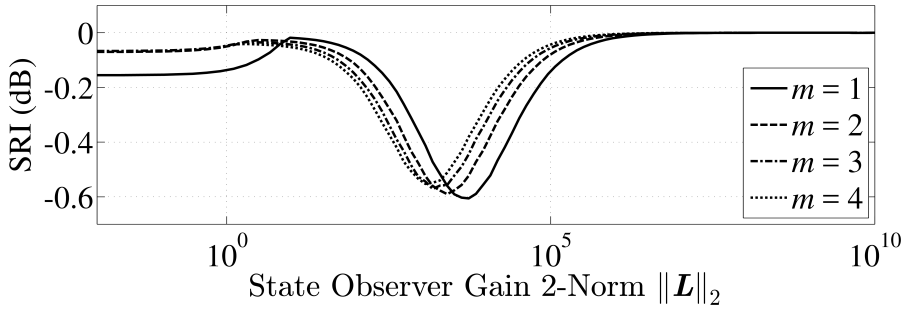


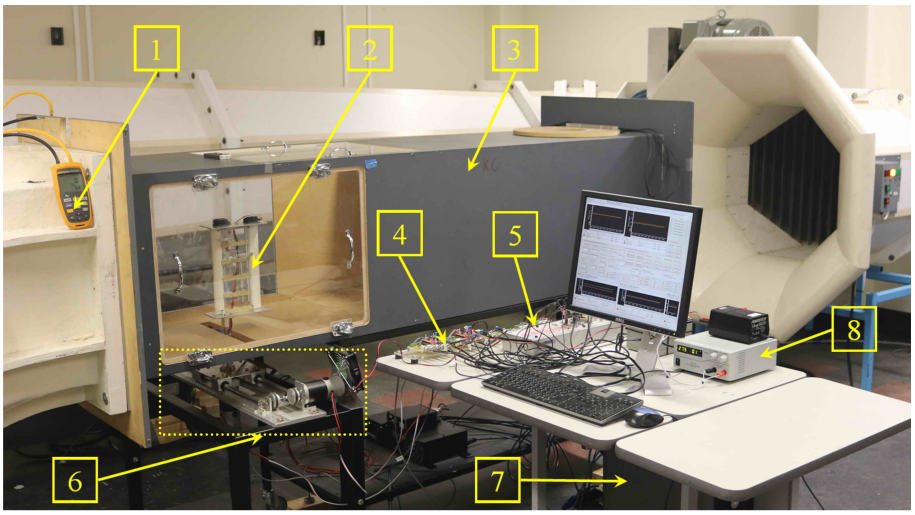
Figure 3.12: SRI against state observer gain.

3.5 Experimental validation

To validate the new VSDS prototype in providing virtual stiffness and damping to the aeroelastic system, wind-tunnel experiments were conducted, and the performance of the generated virtual structural forces $\bar{\boldsymbol{F}}_s$ tracking the reference $\boldsymbol{F}_s$ is evaluated. Realisation of virtual stiffness and damping via electric drives is considered satisfactory if the generated force/torque closely track the reference. The experimental setup for wind-tunnel testing is shown in Figure 3.13, and corresponding parameters of the 2-DOF aeroelastic system under the experimental setting are listed in Table. 5.1.

Two test scenarios are presented in this paper, with corresponding settings listed in Table. 3.3. The numerical values of the eigenfrequencies of the VSDS (with the aerofoil section mounted on) and those of the overall aeroelastic system are given in Table. 3.4. In each scenario, comparisons are drawn between standard LQG tracking control (by disconnecting the UIE component) and the UIE-LQG control.

In experiment, the VSDS controller ran at 1000 Hz while readings were taken



(1 – pressure transducer connected to a Pitot tube; 2 – airfoil; 3 – wind tunnel testing duct; 4 – I/O board; 5 – dSPACE® DS1104 R&D board; 6 – VSDS; 7 – PC; 8 – power.)

Figure 3.13: Wind-tunnel experiment setup.

Table 3.2: Parameters of wind-tunnel experiment setup.

Parameters	Values	Parameters	Values
L_{hc}	0.0753 m	c_h, c_a, k_h, k_a	See Table. 3.3
L_s	0.26 m	m_a	0.851 kg
r_a	0.0329 m	I_a	2.431×10^{-3} kg·m^2
r_{fc}	-0.0685 m	$C_{l\text{-}a}$	6.573
ρ	1.225 kg/m^3	$C_{m\text{-}a}$	0

Table 3.3: Settings of wind-tunnel test cases.

Case	Flutter Boundary	Airflow Speed	Stiffness & Damping
1	13.92 m/s	14.8 m/s	$k_h = 50 + 300h^2$ N/m $k_a = 0.3 + 30q_a^2$ Nm/rad $c_h = 14$ kg/s $c_a = 0.042$ kg·m^2/s
2	16.02 m/s	16.8 m/s	k_h, c_h, c_a: same as Test 1 $k_a = 0.77 + 30q_a^2$ Nm/rad

from the force/torque transducer and encoders at 5000 Hz. The force/torque transducer used on the VSDS features high signal-to-noise ratio with near-zero noise distortion. This allowed the use of a simple digital averaging filter with a sample

Table 3.4: Numerical values of the eigenfrequencies of the VSDS (with aerofoil mounted on) and overall aeroelastic system.

Case	VSDS (with Aerofoil)		Aeroelastic System	
	Plunge	Pitch	Plunge	Pitch
1	0 rad/s	8.06 rad/s	0 rad/s	11.91 rad/s
2	0 rad/s	14.71 rad/s	0 rad/s	16.20 rad/s

Table 3.5: VSDS controller parameters.

$\boldsymbol{K}_x$	K_w	K_r	$\boldsymbol{L}$	K_d	$G_f(s)$
$\begin{bmatrix}6.27 & 3.98\end{bmatrix}$	-3.16	0.4	$\begin{bmatrix}9.49 & 1.07\end{bmatrix}^{\mathrm{T}}$	6.64	$\frac{1}{(0.01s+1)}$
$\begin{bmatrix}4.39 & -0.74\end{bmatrix}$	-3.16	10.08	$\begin{bmatrix}7.33 & 2.19\end{bmatrix}^{\mathrm{T}}$	4.34	$\frac{1}{(0.01s+1)}$

size of 5, which gave finer force/torque readings while only induced 0.4 degree phase delay for the presented scenarios in experiments.

Given the MESO being driven by position error, proper velocity estimation can be verified by comparing measured and estimated positions. The MESO was fined tuned to give fast estimation with the phase lag minimised to less than 0.5 degree in experiments, and corresponding parameters were $p_1 = 90$, $p_2 = 20000$, and $p_3 = 80000$.

The UIE-LQG controller was synthesised following the analysis in Section 3.4, with parameters given in Table. 3.5. The total additional phase difference introduced by the VSDS control system, between actual displacement/velocity and generated force/torque, was found to be 4 degrees at maximum in tests.

Aerofoil plunge and pitch responses in the two test scenarios are shown in Figures 3.14 and 3.15. According to Eqs. (3.4) and (3.5), proper velocity estimation by MESO can be verified using position information only. To demonstrate, Figure 3.16 shows a close-up of position/velocity MESO estimation for UIE-LQG control in Case 1, where conventional velocity approximation using a filter is also presented. Subject to the limited resolution available, the filter was designed as $\frac{s}{1/50s+1}$ to provide relatively acceptable velocity approximation, which however, introduced a

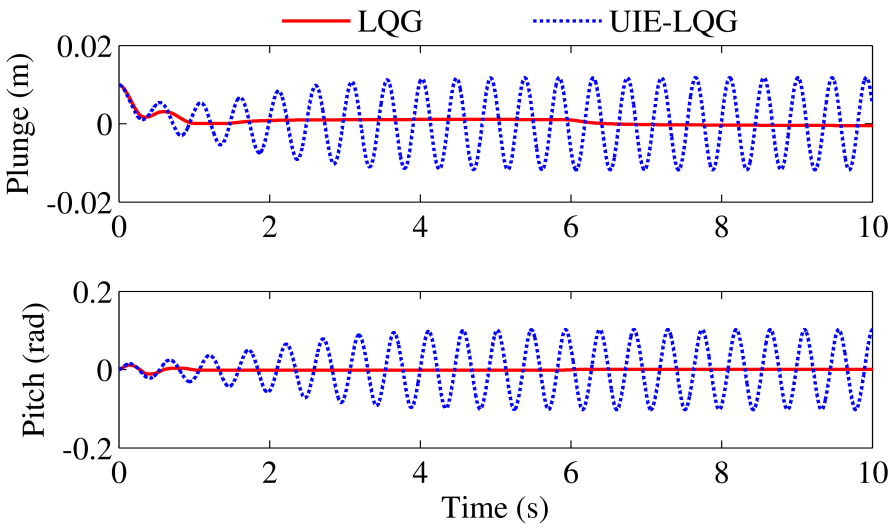


Figure 3.14: Aeroelastic responses in Case 1 tests.

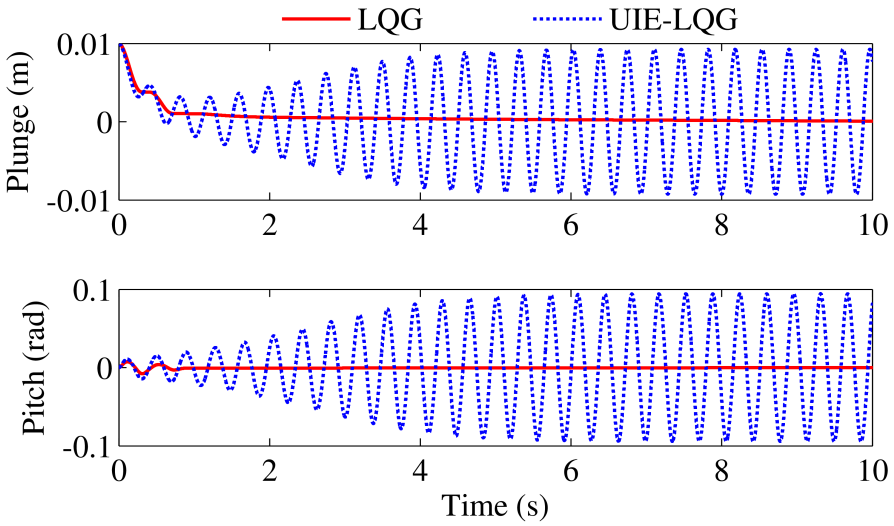


Figure 3.15: Aeroelastic responses in Case 2 tests.

significant amount of phase lag compared with MESO estimation, as can be seen in Figure 3.16. It can also be observed that position estimations using MESO closely follow encoder measurements, indicating fast and proper velocity estimation by MESO.

Note the different aeroelastic responses under the same experiment settings with different controllers, causes of which are revealed in Figures 3.17 to 3.20. Significant differences between measurements and reference tracking trajectories can be observed in tests where standard LQG control was applied (Figures 3.17 and 3.19). As a result, flutter failed to initiate under the LQG control (Figures 3.14

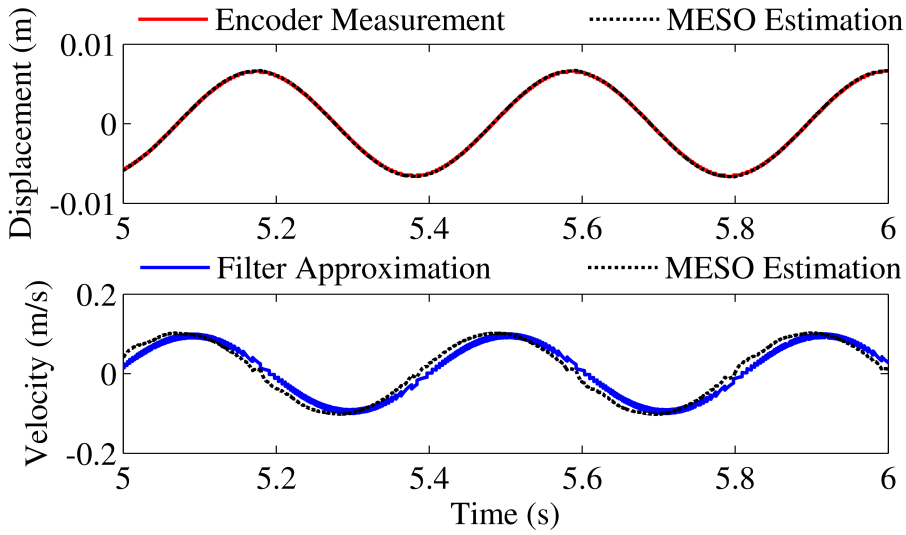


Figure 3.16: Position/velocity estimation using MESO (Case 1, UIE-LQG).

and 3.15), although the airflow speeds in tests were higher than corresponding flutter boundaries. In comparison, the measured plunge-DOF forces and pitch-DOF torques strictly follow the desired trajectories under the UIE-LQG control, with tracking deviations barely identified (Figures 3.18 and 3.20). This enabled successful initiation and development of flutter (Figures 3.14 and 3.15).

The impact of the additional phase lag from displacement/velocity measurement to force/torque generation can be examined by calculating the equivalent stiffness and damping achieved in experiments (using data of position/velocity estimations by MESO and force/torque measurements by transducer). A maximum of 0.68% error is found for all stiffness and damping coefficients in the two test scenarios. Since the phase delay caused by MESO estimation is less than 1/8 of that caused by the UIE-LQG controller, its addition to the total error of equivalent stiffness and damping is deemed negligible.

It can be seen from the experiment results that:

- Careful design of the MESO can mitigate the problem of resolution loss due to using non-reduction transmission, with the additional phase lag between actual displacement/velocity and estimated ones minimised and the associated negative impact being negligible;

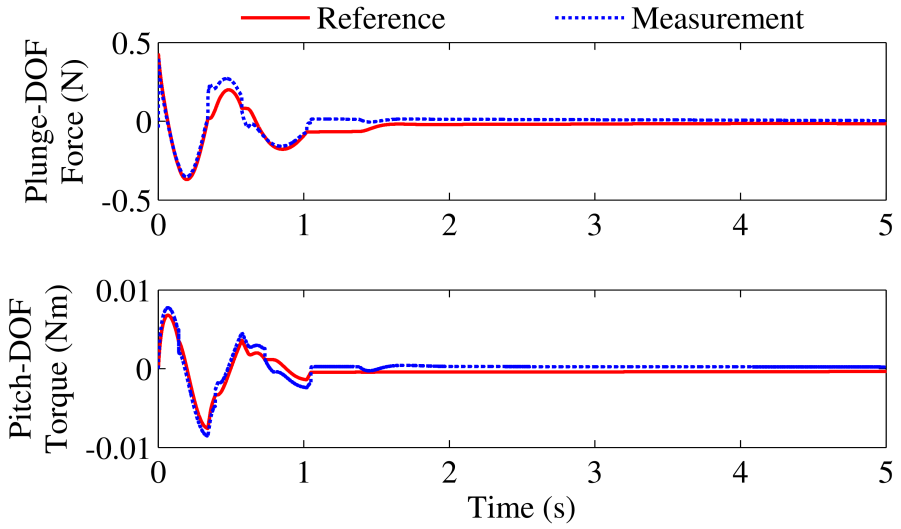


Figure 3.17: VSDS force/torque under LQG control in Case 1 tests.

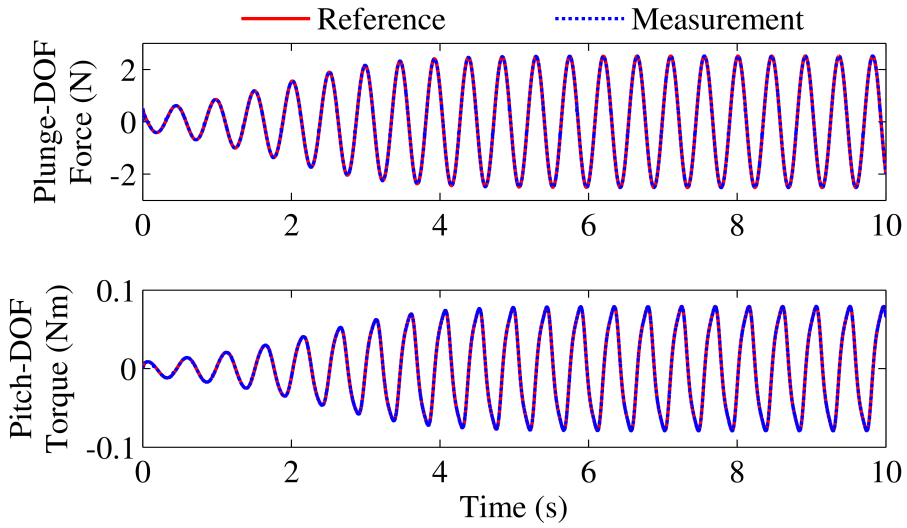


Figure 3.18: VSDS force/torque under UIE-LQG control in Case 1 tests.

- Power loss due to inertial loads, mechanical frictions, and other exogenous disturbances and un-modelled dynamics, if not properly treated, have considerable impacts on the performance of the new VSDS in AAT;

- The power loss, in the form of equivalent unknown inputs, can be effectively estimated and compensated by the UIE-LQG controller, with superior force/torque tracking achieved. The proposed approach allows simplified system identification and calibration procedures.

- The maximum of 4 degrees additional phase lag experienced in experiments,

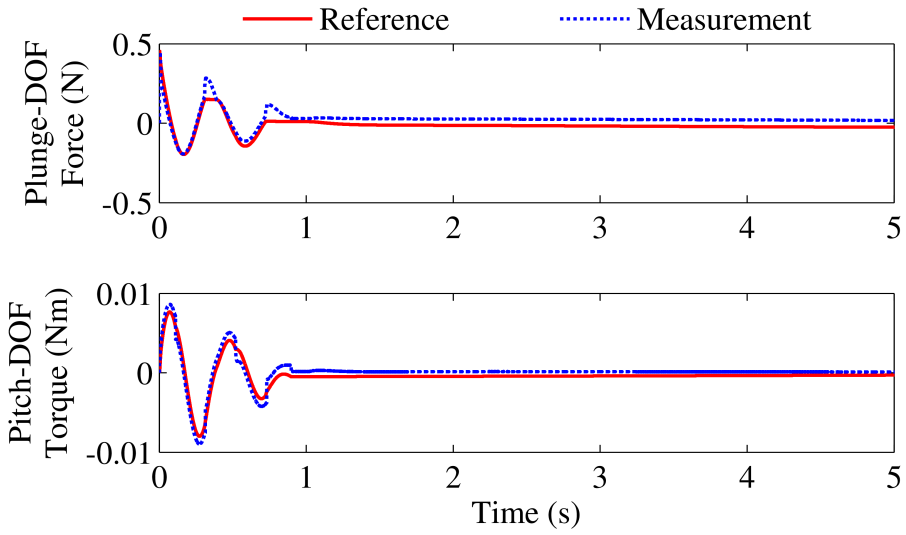


Figure 3.19: VSDS force/torque under LQG control in Case 2 tests.

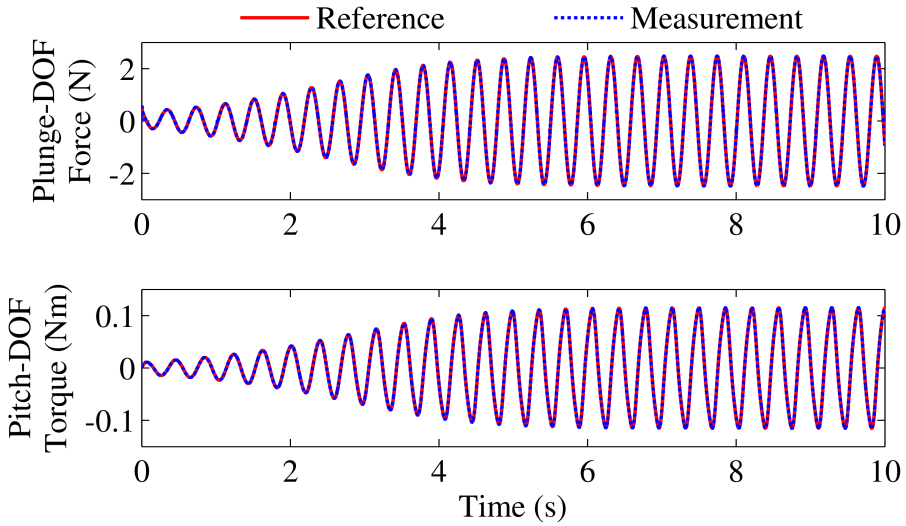


Figure 3.20: VSDS force/torque under UIE-LQG control in Case 2 tests.

introduced by the VSDS control system, showed minor impact on the overall realisation of virtual stiffness and damping for AAT.

3.6 Conclusions

Motivated by the disadvantages of conventional test-beds for AAT and limitations of existing VSDSs in other fields, a new VSDS is developed in this study specifically for AAT. The proposed new operation principle based on direct force/torque regulation with force/torque feedback effectively addresses dynamics coupling between plunge

and pitch DOFs without the need for sophisticated aeroelastic modelling. Resolution loss in velocity measurement is identified as a main problem associated with the first trial of non-reduction transmission on the new VSDS prototype and can be solved by the proposed MESO given that parameters of the MESO are carefully tuned to minimise the additional phase lag between estimated displacement/velocity and actual ones. Based on the new operation principle, the proposed VSDS control system with systematically synthesised UIE-LQG control enables superior force/torque tracking with enhanced robustness and significantly reduces system identification and calibration procedures in VSDS development, although with some phase lag introduced. The phase lag from the actual displacement/velocity to the generated force/torque due to the involvement of the closed-loop force/torque tracking controller, is found to cause minor error in the realised virtual stiffness and damping, which are shown acceptable. In general, as confirmed in wind-tunnel experiments, the developed new 2-DOF VSDS prototype can provide satisfactory realisation of virtual stiffness and damping to facilitate industrial and laboratory AAT. The developed VSDS also has the potential for other industrial applications involving oscillatory tests that require frequent change of stiffness and damping settings. Further reduction of phase lag introduced by the control system can be a possible direction of future work.

Chapter 4

Nonlinear optimal control online synthesis

To derive active flutter suppression (AFS) controllers that accomplish the thesis aim, the two practical problems of the existing nonlinear optimal control online synthesis (NOCOS) algorithms (see Section 1.1) are required to be solved first. This chapter, based on *Article-2*, addresses the NOCOS algorithm structure problem under the locally nonlinear setting, by proposing a novel NOCOS scheme that is compact in configuration without compromising closed-loop stability. Detailed derivation, stability analysis, and numerical verification of the proposed compact NOCOS scheme for locally nonlinear systems are given, completing the second objective (Page 4) of the thesis.

Statement of Authorship

Title of Paper	Modified value-function-approximation for synchronous policy iteration with single-critic configuration for nonlinear optimal control
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Tang D, Chen L, Tian ZF and Hu E (2019) Modified value-function-approximation for synchronous policy iteration with single-critic configuration for nonlinear optimal control. *International Journal of Control,* DOI: 10.1080/00207179.2019.1648874.

Principal Author

Name of Principal Author (Candidate)	Difan Tang		
Contribution to the Paper	Proposing and deriving methods and theories, performing theoretical and numerical analysis, interpreting results, and writing manuscript.		
Overall percentage (%)	80 %		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	28 October 2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Lei Chen		
Contribution to the Paper	Supervising development of work, evaluating methods and theories, providing advice in mathematical aspects and evaluating manuscript.		
Signature		Date	1/11/19

Name of Co-Author	Zhao Feng Tian		
Contribution to the Paper	Supervising development of work, providing advice in mathematical aspects, and evaluating manuscript.		
Signature		Date	1/11/19

Name of Co-Author	Eric Hu		
Contribution to the Paper	Supervising development of work, providing advice in mathematical aspects, and evaluating manuscript.		
Signature		Date	4/11/2019

abstract

This study proposes a modified value-function-approximation (MVFA) and investigates its use under a single-critic configuration based on neural networks (NNs) for synchronous policy iteration (SPI) to deliver compact implementation of optimal control online synthesis for control-affine continuous-time nonlinear systems. Existing single-critic algorithms require stabilising critic tuning laws while eliminating actor tuning. This paper thus studies alternative single-critic realisation aiming to relax the needs for stabilising mechanisms in the critic tuning law. Optimal control laws are determined from the Hamilton-Jacobi-Bellman equality by solving for the associated value function via SPI in a single-critic configuration. Different from other existing single-critic methods, an MVFA is proposed to deal with the closed-loop stability during online learning. Gradient-descent tuning is employed to adjust the critic NN parameters in the interests of not complicating the problem. Parameters convergence and the closed-loop system states stability are examined. The proposed MVFA-based approach yields an alternative single-critic SPI method with uniformly ultimately bounded NN parameter convergence and asymptotic closed-loop system states stability throughout the process of online learning without the need for stabilising mechanisms in the tuning law for critic NN. The proposed approach is verified via simulations.

4.1 Introduction

Nonlinear optimal control generally involves the determination of control laws that minimise the associated performance cost, where the Hamilton-Jacobi-Bellman (HJB) equality (Bellman, 1957) or its nonlinear variations are to be solved, or where an inverse approach without solving the HJB equation (Lopez et al., 2017) may apply. In our study, the discussion is focused on the former, where the HJB equality and its variants, being partial differential equations that are nonlinear, are difficult to be solved analytically. Practical methods to solve the HJB equation and its variants are provided through approximation methods, one class of which is the widely studied adaptive/approximate dynamic programming (ADP) (Werbos, 1974). ADP techniques are basically iterative approaches built upon the concept of reinforcement learning (Sutton and Barto, 1998), which approximates optimal control laws as well as corresponding value functions through policy evaluation and improvement, where a 'policy' is referred to as a control law. Some good reviews are provided by Wang et al. (2009), Jiang and Jiang (2013), and Wang et al. (2017a). To implement the ADP, the value function in the HJB equation needs to be properly structured, and neural networks (NNs) are ideal candidates given their universal approximation properties (Hornik et al., 1989).

Offline ADP has been an effective and useful tool for handling optimal control in various challenging problems, including nonaffine systems (Luo et al., 2016a; Mu et al., 2017; Wang et al., 2012), actuator saturation (Abu-Khalaf and Lewis, 2005; Heydari and Balakrishnan, 2013; Luo et al., 2015), unknown system dynamics (Li et al., 2017; Luo et al., 2016a, 2014, 2015; Mu et al., 2017, 2018; Wang and Liu, 2013; Wang et al., 2012; Wei et al., 2017; Zhao et al., 2015a), fixed final time (Heydari and Balakrishnan, 2013), finite approximation error (Wei et al., 2014), finite horizon (Mu et al., 2018), algorithm simplification (Heydari, 2014; Heydari and Balakrishnan, 2013; Wang and Liu, 2013), optimal tracking (Luo et al., 2016a), non-zero initial condition for value iteration (Wei et al., 2016), and extension to multi-agent system

applications (Li et al., 2017).

With increasing demands on synthesising optimal controllers in real time, online ADP has been receiving intensive research attention. Online ADP, in contrast to offline methods, features real-time synthesis of optimal control policies for dynamic systems. The iteration procedures performed on a regular- or irregular-time-interval basis, where the cost function corresponding to an admissible control being approximated undergoes evaluation before the next iteration commences, can be characterised as being *sequential*. These algorithms collect real-time data prior to batch processing for policy evaluation and policy update at each discrete iteration under either continuous-time setting (Feng et al., 2015; Jiang and Jiang, 2014, 2015; Liu et al., 2013b; Vrabie and Lewis, 2009) or in discrete-time domain (Al-Tamimi et al., 2008; Feng et al., 2015; Kiumarsi et al., 2015; Škach et al., 2018; Wei and Liu, 2014). The study by Vamvoudakis and Lewis (2010) proposes an attractive ADP algorithm, termed as *synchronous policy iteration* (SPI), where policy evaluation and policy update are implemented continuously in time and simultaneously. The SPI theory framework initiated by Vamvoudakis and Lewis (2010) has been enormously enriched by latest advances in dealing with faster convergence (Bhasin et al., 2013), actuator saturation (Huang et al., 2017; Kiumarsi and Lewis, 2015; Modares and Lewis, 2014; Modares et al., 2013a, 2014, 2013b; Yang et al., 2014), completely unknown dynamics with unknown nonlinear structures (Liu et al., 2013a; Yang et al., 2014), unknown affine nonlinear systems (Lv et al., 2019, 2016; Modares et al., 2013a; Na and Herrmann, 2014; Song et al., 2016; Wang et al., 2016; Zhong et al., 2018), partially unknown dynamics (Bhasin et al., 2013; Kiumarsi and Lewis, 2015; Modares and Lewis, 2014; Modares et al., 2014; Vamvoudakis et al., 2014), multi-agent systems (Heydari and Balakrishnan, 2014; Jiang and He, 2018; Luy, 2018), optimal tracking (Kiumarsi and Lewis, 2015; Modares and Lewis, 2014; Na and Herrmann, 2014), relaxation of persistent-excitation condition (Modares et al., 2014), exponential convergence driven directly by estimation error assuming known ideal parameters rather than being driven by the HJB error (Lv et al., 2019, 2016; Na

and Herrmann, 2014), algorithm simplification (Huang et al., 2017; Liu et al., 2013a, 2014; Luy, 2018; Lv et al., 2019, 2016; Na and Herrmann, 2014; Wang et al., 2017b, 2014a, 2017c; Zhang et al., 2013), and disturbances and uncertainties (Huang et al., 2017; Liu et al., 2014; Lv et al., 2019; Song et al., 2016; Vamvoudakis and Lewis, 2012; Wang et al., 2014a, 2016).

For stabilisation purpose, most SPI schemes implement separate NNs for the critic and actor, respectively, each dynamically tuned with a different learning law. Specifically, actor tuning laws generally contain stabilising terms derived from Lyapunov stability analysis. To simplify SPI implementation and reduce computational load, there have been efforts on single-critic approaches where the same NN is used for both components with the critic NN weights directly passed on to the actor NN (Huang et al., 2017; Liu et al., 2013a, 2014; Luy, 2018; Lv et al., 2019, 2016; Na and Herrmann, 2014; Wang et al., 2017b, 2014a, 2017c; Zhang et al., 2013). Further improvements are seen in event-based methods based on the single-critic configuration (Wang et al., 2017b,c), where the data needed for online learning are reduced. The instability resulted from direct simplification of the actor-critic configuration is recognised in Liu et al. (2013a), and critic-NN initial weights need to be determined carefully by trial-and-error. Guaranteed stability can be achieved by introducing a stabilising mechanism to the critic tuning law (Huang et al., 2017; Liu et al., 2014; Luy, 2018; Lv et al., 2019, 2016; Na and Herrmann, 2014; Wang et al., 2017b, 2014a; Zhang et al., 2013). The stabilising mechanism is generally a stabilising term derived on the basis of Lyapunov stability, either conditionally activated upon instability being detected (Huang et al., 2017; Liu et al., 2014; Luy, 2018; Wang et al., 2014a; Zhang et al., 2013), or continuously in effect throughout online learning (Lv et al., 2019, 2016; Na and Herrmann, 2014; Wang et al., 2017b). It is interesting to note that the SPI schemes in the aforementioned studies share a common form of value function approximation (VFA) with an NN of standard structure directly employed. The question is: Can a different form of VFA deliver alternative realisation of the single-critic configuration for SPI without introducing

additional stabilising mechanisms in the NN tuning law?

Therefore, as our major contributions, this study proposes a modified value-function-approximation (MVFA) and study its feasibility and efficacy as an alternative approach under the single-critic configuration. Specifically, the closed-loop stability are investigated.

In the remainder of the paper: Section 4.2 introduces the problem under discussion together with some preliminaries; Section 4.3 proposes an MVFA for alternative realisation of the single-critic configuration for SPI; Section 4.4 analyses overall closed-loop stability during online learning; Section 4.5 gives a simulation example. Section 4.6 draws conclusions.

4.2 Problem and preliminaries

4.2.1 Problem

The following control-affine nonlinear systems in continuous-time domain is considered:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})\boldsymbol{u}, \tag{4.1}$$

where $\boldsymbol{x} \in \mathbb{R}^{n_x}$ contains system states of dimension n_x, $\boldsymbol{x}(0) = \boldsymbol{x}_0$, with $\boldsymbol{x}_0$ being a vector containing the initial states; $\boldsymbol{u} \in \mathbb{R}^{n_u}$ collects control inputs of dimension n_u; $\boldsymbol{f}(\boldsymbol{x}) \in \mathbb{R}^{n_x}$ refers to internal dynamics of the system; $\boldsymbol{g}(\boldsymbol{x}) \in \mathbb{R}^{n_x \times n_u}$ denotes distribution dynamics of control inputs.

Assumption 4.1. *For the system as in Eq. (4.1), there is $\boldsymbol{f}(0) = 0$. Given a set $\boldsymbol{\Omega} \subseteq \mathbb{R}^{n_x}$ including zero, Eq. (4.1) is Lipschitz continuous with respect to $\boldsymbol{\Omega}$, and there exist admissible control $\boldsymbol{u} \in \boldsymbol{\Xi}(\boldsymbol{\Omega})$ that can stabilise system (4.1). $\boldsymbol{f}(\boldsymbol{x})$ as well as $\boldsymbol{g}(\boldsymbol{x})$ are assumed known.*

Assumption 4.2. *There exist $\|\boldsymbol{f}(\boldsymbol{x})\| \leq b_f \|\boldsymbol{x}\|$ with constant $b_f \in \mathbb{R}^+$ and $\|\boldsymbol{g}(\boldsymbol{x})\| \leq b_g$ with constant $b_g \in \mathbb{R}^+$ (Modares et al., 2014, 2013b; Vamvoudakis and Lewis, 2010).*

A proper control law $\boldsymbol{u}$ is desired to minimise

$$V(\boldsymbol{x}_0) = \int_0^\infty \left[\bar{Q}(\boldsymbol{x}(t)) + \boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}\right] dt, \tag{4.2}$$

which is also known as a cost function with a positive-definite function $\bar{Q}(\boldsymbol{x})$ and symmetric positive-definite weighting $\boldsymbol{R} \in \mathbb{R}^{n_u \times n_u}$.

Definition 4.1 (Admissible control)**.** Given a continuously differentiable control set $\boldsymbol{u}(\boldsymbol{x}) \in \boldsymbol{\Xi}(\boldsymbol{\Omega})$ with initial condition $\boldsymbol{u}(0) = 0$, if on $\boldsymbol{\Omega}$ it stabilises system (4.1) and if the cost $V(\boldsymbol{x}_0)$, $\forall \boldsymbol{x}_0 \in \boldsymbol{\Omega}$, as given in Eq. (4.2) is finite, then the control is considered as being admissible (Beard et al., 1997).

4.2.2 Continous-time HJB equation

If $V \in C^1$, differentiating Eq. (4.2) yields

$$\bar{Q}(\boldsymbol{x}) + \boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u} + (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u})^{\mathrm{T}}\nabla V = 0, \tag{4.3}$$

with $V(0) = 0$ and $\nabla V \triangleq \frac{\partial V(\boldsymbol{x})}{\partial \boldsymbol{x}} \in \mathbb{R}^{n_x}$.

The control that minimises Eq. (4.2) for the same initial conditions is deemed optimal and denoted as $\boldsymbol{u}^*$. The associated cost is $V^* = \min(V)$ for $\boldsymbol{u} \in \boldsymbol{\Xi}(\boldsymbol{\Omega})$ and generally known as the 'value function'. Specifically,

$$\boldsymbol{u}^* = -\frac{1}{2}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla V^*, \tag{4.4}$$

with which there is

$$\bar{Q} + \boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}^* + (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*)^{\mathrm{T}}\nabla V^* = 0, \tag{4.5}$$

with $V^*(0) = 0$, which then gives the following HJB equation:

$$-\frac{1}{4}\nabla V^{*\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla V^* + \nabla V^{*\mathrm{T}}\boldsymbol{f} + \bar{Q} = 0, \tag{4.6}$$

with $V^*(0) = 0$.

Remark 4.1. Note that $\boldsymbol{u}$ in Eq. (4.3) can be any admissible control, and there exists a corresponding cost V as in Eq. (4.2) that makes Eq. (4.3) hold. However, Eq. (4.5) is

a special case of Eq. (4.3) where $\boldsymbol{u}$ is associated with V through Eq. (4.4). A residual error arises to the right of Eqs. (4.5) and (4.6) if the condition of $V^* = \min(V)$ for $\boldsymbol{u} \in \Xi(\boldsymbol{\Omega})$ is unsatisfied.

4.2.3 Policy iteration

To analytically determine $V^*(\boldsymbol{x})$ from the nonlinear HJB equation has been known difficult. Instead, $V^*(\boldsymbol{x})$ can be obtained through an iterative procedure termed as 'policy iteration' (Sutton and Barto, 1998), which requires $V^*(\boldsymbol{x})$ being appropriately structured and successively approximated (Saridis and Lee, 1979), basically involving two steps in a 'actor-critic' configuration:

- The 'critic' for policy evaluation: using Eq. (4.3) to evaluate $V_{(i)}$ resulted from $\boldsymbol{u}_{(i)}$. This is to solve for $V_{(i)}$ from

$$\bar{Q}(\boldsymbol{x}) + \boldsymbol{u}_{(i)}^{\mathrm{T}} \boldsymbol{R} \boldsymbol{u}_{(i)} + (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}_{(i)})^{\mathrm{T}} \nabla V_{(i)} = 0, \tag{4.7}$$

  with $V_{(i)}(0) = 0$.

- The 'actor' for policy improvement: implementing updated control, which is

$$\boldsymbol{u}_{(i+1)} = -\frac{1}{2} \boldsymbol{R}^{-1} \boldsymbol{g}^{\mathrm{T}} \nabla V_{(i)}. \tag{4.8}$$

The iteration procedure begins with $\boldsymbol{u}_{(0)}$ which is an initial admissible control, and proceeds with the above two iterative steps until reaching convergence at V^* and $\boldsymbol{u}^*$ or proximity to V^* and $\boldsymbol{u}^*$. It is worth emphasising that for synchronous policy iteration (SPI), the procedure performs continuously in time, and the above two steps take place simultaneously (Vamvoudakis and Lewis, 2010). The subscript '(i)' in $V_{(i)}$ and $\boldsymbol{u}_{(i)}$ are unnecessary in the SPI case. However, for ease of explanation of SPI at an infinitesimal time step, these subscripts are used, only to indicate a general time step being considered rather than iteration number.

Remark 4.2. In terms of the single-critic configuration, an actor component is still necessary for a complete policy iteration procedure including SPI. The term 'single-critic' refers to the case where the separate tuning for the actor component is

eliminated in comparison to the general actor-critic structure in which both of the actor and critic components require individual tuning.

4.3 Modified single-critic configuration

4.3.1 Modified value function approximation

Analytically obtaining $V_{(i)}(\boldsymbol{x})$ from Eq. (4.7) is difficult, and hence implementing policy iteration requires proper approximation of the solution. Neural networks (NNs), with universal approximation properties (Hornik et al., 1989), can be used for this purpose. Different from other existing studies that use a common form of NN-based representation for approximating the value function, in this paper a modified value-function-approximation (MVFA) is proposed, being:

$$V^* = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} + \boldsymbol{W}^{*\mathrm{T}}\boldsymbol{\Phi} + \varepsilon, \tag{4.9}$$

where hidden-layer neurons are contained in $\boldsymbol{\Phi} \in \mathbb{R}^{n_n}$, with ideal NN weights being $\boldsymbol{W}^* \in \mathbb{R}^{n_n}$; $\boldsymbol{P} \in \mathbb{R}^{n_x \times n_x}$ is an additional parameter matrix that is diagonal and positive-definite; the error of approximation is denoted by $\varepsilon \in \mathbb{R}$.

Accordingly, there is

$$\nabla V^* = \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* + \boldsymbol{P}\boldsymbol{x} + \nabla\varepsilon, \tag{4.10}$$

with $\bar{\nabla}\boldsymbol{\Phi} = \nabla\boldsymbol{\Phi}^{\mathrm{T}} = \left[\frac{\partial\boldsymbol{\Phi}}{\partial\boldsymbol{x}}\right]^{\mathrm{T}} \in \mathbb{R}^{n_n \times n_x}$ and $\nabla\varepsilon = \frac{\partial\varepsilon}{\partial\boldsymbol{x}} \in \mathbb{R}^{n_x}$.

Remark 4.3. The discussion in Section 4.1 has revealed that VFA in existing methods takes a common NN-based representation, the convergence of which in online learning necessitates separate actor tuning or stabilising mechanisms in critic tuning laws for stabilisation. Differently in this study, the proposed MVFA features an auxiliary term in addition to the standard structure of an NN. The advantages of introducing the auxiliary term is to be discussed in the remainder of this paper.

Remark 4.4. The hidden-layer neurons in $\boldsymbol{\Phi}$ are nonlinear activation functions, which can be obtained by applying Weierstrass approximation using high-order polynomi-

als (Finlayson, 1972). The resulting activation functions are the individual terms of a polynomial of specified order with the NN inputs as variables.

Assumption 4.3. *There exist inequalities* $\left\|\bar{\nabla}\boldsymbol{\Phi}\right\| \leq b_\phi \left\|\boldsymbol{x}\right\|$ *for* $b_\phi \in \mathbb{R}^+$ *and* $\left\|\nabla\varepsilon\right\| \leq b_\varepsilon \left\|\boldsymbol{x}\right\|$ *for* $b_\varepsilon \in \mathbb{R}^+$*, where* b_ϕ *and* b_ε *are constants.*

4.3.2 Single-critic structure and tuning

On considering the 'Policy Evaluation' step only (i.e., a control law remains fixed for evaluation), the associated cost function $V_{(i)}$ takes

$$V_{(i)} = \boldsymbol{W}_{(i)}^{\mathrm{T}}\boldsymbol{\Phi} + \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} + \varepsilon_{(i)}, \tag{4.11}$$

with its gradient being

$$\nabla V_{(i)} = \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}_{(i)} + \boldsymbol{P}\boldsymbol{x} + \nabla\varepsilon_{(i)}, \tag{4.12}$$

with $\boldsymbol{W}_{(i)}$ being NN ideal weights that approximate $V_{(i)}$ with the least error $\varepsilon_{(i)}$.

Remark 4.5. Note that $V_{(i)}$ in Eq. (4.11) and V in Eq. (4.2) are equal only in terms of value, given the same initial conditions and the same control policy, but different in structure. V in Eq. (4.2) is structured to give physical interpretation of cost while $V_{(i)}$ in Eq. (4.11) is specially constructed for mathematical approximation. The term $\boldsymbol{W}_{(i)}^{\mathrm{T}}\boldsymbol{\Phi}$ in Eq. (4.11) is not equal to $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}$ in Eq. (4.2) but includes the information of $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}$, since the set $\boldsymbol{\Phi}$ contains activation functions in polynomial forms consisting of both $\boldsymbol{x}$ and $\boldsymbol{u}$.

Remark 4.6. The discussion at this stage only considers the case of approximating the cost function for a known control policy $\boldsymbol{u}_{(i)}$. That is, $\boldsymbol{u}_{(i)}$ is known and not approximated by NN. The NN used at this stage only approximates the cost function associated with the known control $\boldsymbol{u}_{(i)}$.

Using an estimate $\hat{\boldsymbol{W}}_{(i)}$ to replace $\boldsymbol{W}_{(i)}$ in Eqs. (4.11) and (4.12) gives

$$\hat{V}_{(i)} = \hat{\boldsymbol{W}}_{(i)}^{\mathrm{T}}\boldsymbol{\Phi} + \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}, \tag{4.13}$$

$$\nabla \hat{V}_{(i)} = \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}}_{(i)} + \boldsymbol{P}\boldsymbol{x}, \tag{4.14}$$

and

$$\bar{Q} + \boldsymbol{u}_{(i)}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}_{(i)} + (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}_{(i)})^{\mathrm{T}}\nabla \hat{V}_{(i)} = e_1, \tag{4.15}$$

where $\hat{V}_{(i)}(0) = 0$, and e_1 is the error arises as a result (as commented in Remark 4.1 and to be discussed in Section 4.4).

To minimise e_1 so that $\hat{\boldsymbol{W}}_{(i)} \to \boldsymbol{W}_{(i)}$, gradient-descent tuning is adopted, by considering the quadratic error function

$$E = \frac{1}{2}e_1^2. \tag{4.16}$$

This yields

$$\dot{\hat{\boldsymbol{W}}}_{(i)} = -\kappa_1 \boldsymbol{N}_1 \frac{\partial E}{\partial \hat{\boldsymbol{W}}_{(i)}} = -\frac{\alpha}{\sqrt{\boldsymbol{\varsigma}_{(i)}{}^{\mathrm{T}}\boldsymbol{\varsigma}_{(i)} + 1}}\boldsymbol{N}_1\boldsymbol{\varsigma}_{(i)}e_1, \tag{4.17}$$

where $\kappa_1 = \frac{\alpha}{\sqrt{\boldsymbol{\varsigma}_{(i)}^{\mathrm{T}}\boldsymbol{\varsigma}_{(i)}+1}}$ is added for normalisation, with $\alpha \in \mathbb{R}^+$ being a scalar learning rate and $\boldsymbol{\varsigma}_{(i)} = \bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}_{(i)})$; $\boldsymbol{N}_1$ is an auxiliary term added to adjust contribution of individual state to tuning, and $\boldsymbol{N}_1 = \mathrm{diag}(\boldsymbol{N}_2\boldsymbol{N}_3)$, with $\boldsymbol{N_2} \in \mathbb{R}^{n_n \times n_x}$ being a constant matrix related to $\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{x})$ with its element $\boldsymbol{N}_{2(jk)} \in \mathbb{B}$, $(j = 1, 2, \cdots, n_n; k = 1, 2, \cdots, n_x)$, and $\boldsymbol{N}_3 \in \mathbb{R}^{n_x \times 1}$ being a weighting vector.

Specifically, the constant matrix $\boldsymbol{N}_2$, in connection with the expression of every single element of $\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{x})$, namely, $\bar{\nabla}\boldsymbol{\Phi}_{(jk)}(\boldsymbol{x})$, is given in the following form:

$$\boldsymbol{N}_{2(jk)} = \begin{cases} 0 & \text{if } \bar{\nabla}\boldsymbol{\Phi}_{(jk)}(\boldsymbol{x}) = 0,\ \forall \boldsymbol{x} \neq 0, \\ 1 & \text{if } \bar{\nabla}\boldsymbol{\Phi}_{(jk)}(\boldsymbol{x}) \neq 0,\ \forall \boldsymbol{x} \neq 0. \end{cases}$$

Similarly, for the complete synchronous policy iteration (SPI), the ideal weights $\boldsymbol{W}^*$ are unknown and should be determined so that Eq. (4.9) approximates a target value function. With $\hat{\boldsymbol{W}}$ being the estimated weights, the approximated value function and its gradient become

$$\hat{V} = \hat{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{\Phi} + \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} \tag{4.18}$$

and

$$\nabla \hat{V} = \bar{\nabla} \boldsymbol{\Phi}^{\mathrm{T}} \hat{\boldsymbol{W}} + \boldsymbol{P}\boldsymbol{x}, \tag{4.19}$$

respectively, and the associated control is given by

$$\hat{\boldsymbol{u}} = -\frac{1}{2} \boldsymbol{R}^{-1} \boldsymbol{g}^{\mathrm{T}} \nabla \hat{V}. \tag{4.20}$$

Note the absence of the subscript '(i)' in Eqs. (4.18) and (4.19) for complete SPI, which are different from Eqs. (4.13) and (4.14) corresponding to a fixed control law at a general infinitesimal time step for 'Policy Evaluation' only.

In the SPI case involving the single-critic structure with Eq. (4.18) and direct implementation of Eq. (4.20), there is

$$\bar{Q} + \hat{\boldsymbol{u}}^{\mathrm{T}} \boldsymbol{R} \hat{\boldsymbol{u}} + (\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}})^{\mathrm{T}} \nabla \hat{V} = e_2, \tag{4.21}$$

where $\hat{V}(0) = 0$, and e_2 is the resulting approximation error as commented in Remark 4.1 (details to be given in Section 4.4).

To minimise e_2 so that $\hat{\boldsymbol{W}} \to \boldsymbol{W}^*$, Eq. (4.17) is modified as

$$\dot{\hat{\boldsymbol{W}}} = -\frac{\alpha}{\sqrt{\boldsymbol{\varsigma}^{\mathrm{T}} \boldsymbol{\varsigma} + 1}} \boldsymbol{N}_1 \boldsymbol{\varsigma} e_2 = -\kappa_2 \boldsymbol{N}_1 \boldsymbol{\varsigma} e_2, \tag{4.22}$$

where $\kappa_2 = \frac{\alpha}{\sqrt{\boldsymbol{\varsigma}^{\mathrm{T}} \boldsymbol{\varsigma} + 1}}$, and $\boldsymbol{\varsigma} = \bar{\nabla} \boldsymbol{\Phi} (\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}})$, with $\boldsymbol{N}_1$, $\boldsymbol{N}_2$, and $\boldsymbol{N}_3$ defined the same as in Eq. (4.17).

It now gives a single-critic structure consisting of critic tuning only, without additional stabilising mechanisms in the tuning law as in Eq. (4.22).

Remark 4.7. For conventional VFA as in the SPI pioneer work of Vamvoudakis and Lewis (2010) (also commonly used in other studies discussed in Section 3.1), it has been known that closing the loop by directly passing $\hat{\boldsymbol{W}}$ on to the actor NN can lead to instability issues during online learning without any stabilising mechanism. This is because in these cases some intermediate values along the evolution path of $\hat{\boldsymbol{W}}$ may not necessarily yield admissible intermediate control policies that satisfy $\frac{d\hat{V}}{dt} < 0$.

Remark 4.8. Compared with the existing single-critic approaches with conventional VFA and stabilising critic tuning laws, the proposed method with MVFA also differs in that the critic tuning law does not need to be stabilising, allowing the use of simpler tuning laws. Accordingly, in this paper the critic tuning based on traditional yet simple gradient descent is used without additional stabilising mechanisms in the tuning law. Closed-loop stability is to be investigated next in Section 4.4.

4.4 Convergence and stability analysis

Similar to most adaptive control problems that require online tuning of parameters (Ioannou and Sun, 1996), proper convergence of NN parameters in this paper also relies on the persistence of excitation (PE) condition to ensure sufficiently rich training set being obtained.

Definition 4.2 (Persistence of Excitation)**.** A bounded vector signal $\boldsymbol{r}(t)$ is considered to be persistently excited (PE) if

$$b_{PE1}\boldsymbol{I} \preceq \int_{t_0}^{t_0+t_d} \boldsymbol{r}(t)\boldsymbol{r}(t)^{\mathrm{T}} dt \preceq b_{PE2}\boldsymbol{I}; \quad \forall t_0 \geq 0,$$

where $\boldsymbol{I}$ is an identity matrix, $b_{PE1} \in \mathbb{R}^+$, $b_{PE2} \in \mathbb{R}^+$, and $t_d \in \mathbb{R}^+$ (Ioannou and Sun, 1996).

Assumption 4.4. *During online tuning, states $\boldsymbol{x}(t)$ of the system (4.1) satisfy the PE condition.*

4.4.1 Policy evaluation

As an assistive step to evaluate the position of the proposed MVFA in a single-critic synchronous policy iteration (SPI) structure, this section separates 'Policy Evaluation' from the SPI and treats it as a single process that approximates a cost function for a known admissible control policy.

In addition, the stability of a linear time-varying system as given by the lemma below, is to be used in the stability analysis that follows.

Lemma 4.1. *For a given system being linear and time-varying in the form of*

$$\dot{\boldsymbol{x}} = -\boldsymbol{r}(t)\boldsymbol{r}^T(t)\boldsymbol{x}, \tag{4.23}$$

where vector $\boldsymbol{x}$ contains system states, its origin is exponentially stable if vector $\boldsymbol{r}(t)$ satisfies the condition of PE (Ioannou and Sun, 1996).

The following theorem presents the convergence property of *Policy Evaluation* with the MVFA under the tuning given by Eq. (4.17).

Theorem 4.1. *Let Eq. (4.11) approximate the cost function as in Eq. (4.2) corresponding to a given admissible control $\boldsymbol{u}_{(i)}$. Under Assumptions 4.1, 4.3, and 4.4, and the tuning algorithm as in Eq. (4.17), the error $\tilde{\boldsymbol{W}}_{(i)} = \boldsymbol{W}_{(i)} - \hat{\boldsymbol{W}}_{(i)}$ from NN weights estimation converges to a residual set $\boldsymbol{\sigma}_{\tilde{W}}$ exponentially, and $\|\boldsymbol{\sigma}_{\tilde{W}}\| \leq b_{\tilde{W}}$ for a finite scalar $b_{\tilde{W}} \in \mathbb{R}^+$ with $b_{\tilde{W}} \to 0$ as $n_n \to \infty$.*

Proof. Comparing Eqs. (4.7) and (4.15), with Eq. (4.12) substituted for $\nabla V_{(i)}$, and with Eq. (4.14) substituted for $\nabla \hat{V}_{(i)}$, yields

$$e_1 = -\tilde{\boldsymbol{W}}_{(i)}^{\mathrm{T}}\boldsymbol{\varsigma}_{(i)} + \epsilon_1, \tag{4.24}$$

where $\boldsymbol{\varsigma}_{(i)} = \bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}_{(i)})$, and $\epsilon_1 = -\nabla\varepsilon_{(i)}^{\mathrm{T}}(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}_{(i)})$.

As can be seen from Eq. (4.24), if $\varepsilon_{(i)}(\boldsymbol{x}) = 0$ for any $\boldsymbol{x} \neq 0$, then $\epsilon_1 = 0$. For the case of $\varepsilon_{(i)}(\boldsymbol{x}) \neq 0$, it is easy to see that $\epsilon_1 \leq b_{\epsilon_1}$ for $b_{\epsilon_1} \in \mathbb{R}^+$, given Assumption 4.3 and $(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}_{(i)})$ as well as $\boldsymbol{x}$ being bounded under Assumption 4.1. Since $\varepsilon_{(i)} \to 0$ and $\nabla\varepsilon_{(i)} \to 0$ given proper activation functions with sufficiently large n_n (Finlayson, 1972), it is straightforward to show that $\epsilon_1 \to 0$ when $n_n \to \infty$.

By using Eqs. (4.17) and (4.24), we have the time derivative of $\tilde{\boldsymbol{W}}_{(i)}$

$$\dot{\tilde{\boldsymbol{W}}}_{(i)} = -\alpha\boldsymbol{N}_1\boldsymbol{\varsigma}_{na(i)}\boldsymbol{\varsigma}_{na(i)}^{\mathrm{T}}\tilde{\boldsymbol{W}}_{(i)} + \alpha\boldsymbol{N}_1\boldsymbol{\varsigma}_{nb(i)}\epsilon_1, \tag{4.25}$$

where $\boldsymbol{\varsigma}_{na(i)} = \frac{\boldsymbol{\varsigma}_{(i)}}{(\boldsymbol{\varsigma}_{(i)}^{\mathrm{T}}\boldsymbol{\varsigma}_{(i)}+1)^{\frac{1}{4}}}$ and $\boldsymbol{\varsigma}_{nb(i)} = \frac{\boldsymbol{\varsigma}_{(i)}}{\sqrt{\boldsymbol{\varsigma}_{(i)}^{\mathrm{T}}\boldsymbol{\varsigma}_{(i)}+1}}$.

Let $\boldsymbol{u}_\epsilon = \alpha\boldsymbol{N}_1\boldsymbol{\varsigma}_{nb(i)}\epsilon_1$. If $\epsilon_1 = 0$, then $\boldsymbol{u}_\epsilon = 0$, and Eq. (4.25) reduces to

$$\dot{\tilde{\boldsymbol{W}}}_{(i)} = -\alpha\boldsymbol{N}_1\boldsymbol{\varsigma}_{na(i)}\boldsymbol{\varsigma}_{na(i)}^{\mathrm{T}}\tilde{\boldsymbol{W}}_{(i)}. \tag{4.26}$$

Denote the equilibrium of system (4.25) by $\sigma_{\tilde{W}}$. Under Assumption 4.4, $\varsigma_{na(i)}$ is PE. Recall that α and N_1 in Eq. (4.26) are constant parameters. Therefore, under Lemma 4.1, the origin (i.e. $\sigma_{\tilde{W}} = 0$) of the system (4.26) is exponentially stable. That is, $\tilde{W}_{(i)}$ converges to zero exponentially.

In the case of $u_\epsilon \neq 0$, it is straightforward to show that Eq. (4.25) has non-zero equilibrium (i.e. $\sigma_{\tilde{W}} \neq 0$), and that $\tilde{W}_{(i)}$ converges to $\sigma_{\tilde{W}}$ exponentially. Since $\|\varsigma_{nb(i)}\| < 1$ and $\epsilon_1 \leq b_{\epsilon_1}$, we have $\|u_\epsilon\| \leq b_{u_\epsilon}$ for $b_{u_\epsilon} \in \mathbb{R}^+$ that can be arbitrarily small given sufficient number of suitable activation functions being provided. Therefore, there exists a bound $b_{\tilde{W}} \in \mathbb{R}^+$ such that $\|\sigma_{\tilde{W}}\| \leq b_{\tilde{W}}$, and $b_{\tilde{W}} \to 0$ with the number of activation functions $n_n \to \infty$. □

Remark 4.9. As can be seen from the proof of Theorem 4.1, the MVFA has no direct influence on critic NN weights convergence when considering the 'Policy Evaluation' step only. Exponential stability is primarily due to the admissible control being evaluated. However, the overall system stability in the case of complete SPI needs to be further analysed (to be dealt with in Section 4.4.2), where the control policy is replaced by a dynamically varying approximation.

4.4.2 Synchronous policy iteration

As discussed in Remark 4.7, instability may result when directly implementing the approximated control policy as in Eq. (4.20) for complete SPI. In this subsection, the closed-loop stability under the proposed alternative single-critic scheme with the MVFA is investigated.

Definition 4.3 (Uniformly Ultimately Bounded)**.** The states $x(t)$ of a dynamic system with initial states $x_0 \triangleq x(t_0)$ is regarded as uniformly ultimately bounded (UUB) about equilibrium $x_e \in \mathbb{R}^{n_x}$ if there exist a compact set $\Omega \in \mathbb{R}^{n_x}$, a finite constant $b_e \in \mathbb{R}^+$ and a time $t_d(b_e, x_0) \in \mathbb{R}^+$ such that $\|x(t) - x_e\| \leq b_e$ for any $x_0 \in \Omega$ and $t \geq t_0 + t_d$ (Lewis et al., 1999).

Theorem 4.2. *Consider a system as in Eq. (4.1). Let Eq. (4.9) approximate its value function, with the control policy given by Eq. (4.20). Under Assumptions 4.1 to 4.4 and the online tuning law as in Eq. (4.22), the critic NN weights estimation error* $\tilde{\boldsymbol{W}} = \boldsymbol{W}^* - \hat{\boldsymbol{W}}$ *remain UUB during online tuning while the states* $\boldsymbol{x}$ *of the system are asymptotically stable, if the parameter matrix* $\boldsymbol{P}$ *in Eq. (4.9) is selected to satisfy* $\|\boldsymbol{P}\| > b_{mP}$*, for a finite scalar* $b_{mP} \in \mathbb{R}^+$.

Proof. Consider

$$\begin{aligned} \mathcal{L} &= \hat{V} + \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}(\kappa_2 \boldsymbol{N}_1)^{-1}\tilde{\boldsymbol{W}} \\ &= \mathcal{L}_V + \mathcal{L}_W, \end{aligned} \tag{4.27}$$

where $\mathcal{L}_V = \hat{V}$ and $\mathcal{L}_W = \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}(\kappa_2 \boldsymbol{N}_1)^{-1}\tilde{\boldsymbol{W}}$.

With Eqs. (4.1), (4.19), and (4.20), there is

$$\begin{aligned} \dot{\mathcal{L}}_V &= (\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}})^{\mathrm{T}}\nabla\hat{V} \\ &= (\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}})^{\mathrm{T}}\left[\boldsymbol{f} - \frac{1}{2}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}})\right] \\ &= \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{f} + \hat{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} - \frac{1}{2}\hat{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}} \\ &\quad - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}} - \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}. \end{aligned} \tag{4.28}$$

Let $\boldsymbol{G} = \boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}$. With $\hat{\boldsymbol{W}} = \boldsymbol{W}^* - \tilde{\boldsymbol{W}}$, Eq. (4.28) becomes

$$\begin{aligned} \dot{\mathcal{L}}_V &= \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{f} + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} \\ &\quad - \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* \\ &\quad + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} \\ &\quad + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \frac{1}{2}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*. \end{aligned} \tag{4.29}$$

In regard to the second term in Eq. (4.27), considering Eq. (4.22), we have

$$\begin{aligned} \dot{\mathcal{L}}_W &= \tilde{\boldsymbol{W}}^{\mathrm{T}}(\kappa_2 \boldsymbol{N}_1)^{-1}\dot{\tilde{\boldsymbol{W}}} = -\tilde{\boldsymbol{W}}^{\mathrm{T}}(\kappa_2 \boldsymbol{N}_1)^{-1}\dot{\hat{\boldsymbol{W}}} \\ &= \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}})e_2. \end{aligned} \tag{4.30}$$

By comparing Eqs. (4.4) and (4.20), there is

$$\hat{\boldsymbol{u}} = \boldsymbol{u}^* + \frac{1}{2}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}(\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} + \nabla\varepsilon). \tag{4.31}$$

Let $\boldsymbol{z} = \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}$. Rewriting Eq. (4.30) using Eq. (4.31) gives

$$\begin{aligned}\dot{\mathcal{L}}_W &= \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon e_2 + \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*)e_2 \\ &\quad + \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}e_2 \\ &= \boldsymbol{z}^{\mathrm{T}}(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*)e_2 + \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon e_2 + \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}e_2.\end{aligned} \tag{4.32}$$

Subtracting Eq. (4.21) from Eq. (4.5) yields

$$\begin{aligned}e_2 &= (\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}})^{\mathrm{T}}(\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}}) + \hat{\boldsymbol{u}}^{\mathrm{T}}\boldsymbol{R}\hat{\boldsymbol{u}} - \boldsymbol{u}^{*\mathrm{T}}\boldsymbol{R}\boldsymbol{u}^* \\ &\quad - (\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* + \nabla\varepsilon)^{\mathrm{T}}(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*).\end{aligned} \tag{4.33}$$

By using Eqs. (4.4), (4.10), (4.19), and (4.20), the individual terms in Eq. (4.33) have expressions of

$$\begin{aligned}&(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}})^{\mathrm{T}}(\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}}) \\ &= (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*)^{\mathrm{T}}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* - \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}) \\ &\quad + \frac{1}{2}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*)^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}(\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} + \nabla\varepsilon) \\ &\quad - \frac{1}{2}(\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}})^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}(\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} + \nabla\varepsilon),\end{aligned} \tag{4.34}$$

$$\begin{aligned}&\hat{\boldsymbol{u}}^{\mathrm{T}}\boldsymbol{R}\hat{\boldsymbol{u}} \\ &= \frac{1}{4}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*)^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*) \\ &\quad - \frac{1}{2}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*)^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} \\ &\quad + \frac{1}{4}(\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}})^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}},\end{aligned} \tag{4.35}$$

$$\begin{aligned}&\boldsymbol{u}^{*\mathrm{T}}\boldsymbol{R}\boldsymbol{u}^* \\ &= \frac{1}{4}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*)^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*) \\ &\quad + \frac{1}{2}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*)^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon \\ &\quad + \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon.\end{aligned} \tag{4.36}$$

Substituting Eqs. (4.34), (4.35), and (4.36) back into Eq. (4.33) gives

$$\begin{aligned} e_2 = &- (\nabla \boldsymbol{\Phi}^{\mathrm{T}} \tilde{\boldsymbol{W}})^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) - \nabla \varepsilon^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) \\ &- \frac{1}{2} (\nabla \boldsymbol{\Phi}^{\mathrm{T}} \tilde{\boldsymbol{W}})^{\mathrm{T}} \boldsymbol{g} \boldsymbol{R}^{-1} \boldsymbol{g}^{\mathrm{T}} \nabla \varepsilon - \frac{1}{4} \nabla \varepsilon^{\mathrm{T}} \boldsymbol{g} \boldsymbol{R}^{-1} \boldsymbol{g}^{\mathrm{T}} \nabla \varepsilon \\ &- \frac{1}{4} (\nabla \boldsymbol{\Phi}^{\mathrm{T}} \tilde{\boldsymbol{W}})^{\mathrm{T}} \boldsymbol{g} \boldsymbol{R}^{-1} \boldsymbol{g}^{\mathrm{T}} \nabla \boldsymbol{\Phi}^{\mathrm{T}} \tilde{\boldsymbol{W}}. \end{aligned} \tag{4.37}$$

It can be seen from Eq. (4.37) that for a given set of NN hidden-layer neurons of a finite number n_n, the minimum of e_2, denoted by ϵ_2, is reached when $\tilde{\boldsymbol{W}} = 0$:

$$\epsilon_2 = -\nabla \varepsilon^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) - \frac{1}{4} \nabla \varepsilon^{\mathrm{T}} \boldsymbol{g} \boldsymbol{R}^{-1} \boldsymbol{g}^{\mathrm{T}} \nabla \varepsilon. \tag{4.38}$$

Under Assumptions 4.1 and 4.3, $\nabla \varepsilon$ and $(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*)$ are bounded. Thus, there exist a finite constant $b_{\epsilon_2} \in \mathbb{R}^+$ such that $\epsilon_2 \leq b_{\epsilon_2}$. Since $\varepsilon \to 0$ and $\nabla \varepsilon \to 0$ as the number of suitable activation functions n_n increases infinitely (Finlayson, 1972), it is straightforward to show that $\epsilon_2 \to 0,\ \forall \boldsymbol{x} \neq 0$, if $n_n \to \infty$. As a special case, $\epsilon_2 = 0$ if $\nabla \varepsilon = 0,\ \forall \boldsymbol{x} \neq 0$.

Substituting Eq. (4.33) for e_2 in Eq. (4.32) yields

$$\begin{aligned} \dot{\mathcal{L}}_W = &- \boldsymbol{z}^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*)(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*)^{\mathrm{T}} \boldsymbol{z} - \frac{3}{8} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{z} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \nabla \varepsilon \\ &- \boldsymbol{z}^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*)(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*)^{\mathrm{T}} \nabla \varepsilon - \frac{1}{8} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{z} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{z} \\ &- \frac{3}{4} \boldsymbol{z}^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{z} - \frac{1}{2} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{z} \nabla \varepsilon^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) \\ &- \boldsymbol{z}^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \nabla \varepsilon - \frac{1}{8} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \nabla \varepsilon \nabla \varepsilon^{\mathrm{T}} \boldsymbol{G} \nabla \varepsilon \\ &- \frac{1}{4} \boldsymbol{z}^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) \nabla \varepsilon^{\mathrm{T}} \boldsymbol{G} \nabla \varepsilon - \frac{1}{8} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{z} \nabla \varepsilon^{\mathrm{T}} \boldsymbol{G} \nabla \varepsilon \\ &- \frac{1}{2} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \nabla \varepsilon \nabla \varepsilon^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) - \frac{1}{4} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \nabla \varepsilon \nabla \varepsilon^{\mathrm{T}} \boldsymbol{G} \boldsymbol{z}. \end{aligned} \tag{4.39}$$

Note that the first term in Eq. (4.39) can be expanded as

$$\begin{aligned} &- \boldsymbol{z}^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*)(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*)^{\mathrm{T}} \boldsymbol{z} \\ = &- \boldsymbol{z}^{\mathrm{T}} \boldsymbol{f} \boldsymbol{f}^{\mathrm{T}} \boldsymbol{z} - \frac{1}{4} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \bar{\nabla} \boldsymbol{\Phi}^{\mathrm{T}} \boldsymbol{W}^* \boldsymbol{W}^{*\mathrm{T}} \bar{\nabla} \boldsymbol{\Phi} \boldsymbol{G} \boldsymbol{z} \\ &- \frac{1}{4} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{P} \boldsymbol{x} \boldsymbol{x}^{\mathrm{T}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{z} - \frac{1}{4} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \nabla \varepsilon \nabla \varepsilon^{\mathrm{T}} \boldsymbol{G} \boldsymbol{z} \\ &+ \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{P} \boldsymbol{x} \boldsymbol{f}^{\mathrm{T}} \boldsymbol{z} - \frac{1}{2} \boldsymbol{z}^{\mathrm{T}} \boldsymbol{G} \bar{\nabla} \boldsymbol{\Phi}^{\mathrm{T}} \boldsymbol{W}^* \nabla \varepsilon^{\mathrm{T}} \boldsymbol{G} \boldsymbol{z} \end{aligned}$$

$$+ z^{\mathrm{T}} G \bar{\nabla} \boldsymbol{\Phi}^{\mathrm{T}} W^* f^{\mathrm{T}} z - \frac{1}{2} z^{\mathrm{T}} G P x \nabla \varepsilon^{\mathrm{T}} G z$$
$$+ z^{\mathrm{T}} G \nabla \varepsilon f^{\mathrm{T}} z - \frac{1}{2} z^{\mathrm{T}} G P x W^{*\mathrm{T}} \bar{\nabla} \boldsymbol{\Phi} G z. \tag{4.40}$$

Combining Eqs. (4.29) and (4.39) gives:

$$\dot{\mathcal{L}} = T_1 + T_2 + T_3 + T_4 + T_5, \tag{4.41}$$

where

$$T_1 = -\frac{1}{2} x^{\mathrm{T}} P^{\mathrm{T}} G P x + x^{\mathrm{T}} P^{\mathrm{T}} f + W^{*\mathrm{T}} \bar{\nabla} \boldsymbol{\Phi} f$$
$$- x^{\mathrm{T}} P^{\mathrm{T}} G \bar{\nabla} \boldsymbol{\Phi}^{\mathrm{T}} W^*, \tag{4.42}$$

$$T_2 = -\frac{1}{4} z^{\mathrm{T}} G P x x^{\mathrm{T}} P^{\mathrm{T}} G z - \frac{1}{2} z^{\mathrm{T}} G P x W^{*\mathrm{T}} \bar{\nabla} \boldsymbol{\Phi} G z$$
$$- \frac{1}{2} z^{\mathrm{T}} G P x \nabla \varepsilon^{\mathrm{T}} G z - \frac{1}{2} z^{\mathrm{T}} G \bar{\nabla} \boldsymbol{\Phi}^{\mathrm{T}} W^* \nabla \varepsilon^{\mathrm{T}} G z$$
$$+ z^{\mathrm{T}} G P x f^{\mathrm{T}} z + z^{\mathrm{T}} G \bar{\nabla} \boldsymbol{\Phi}^{\mathrm{T}} W^* f^{\mathrm{T}} z + z^{\mathrm{T}} G \nabla \varepsilon f^{\mathrm{T}} z$$
$$- z^{\mathrm{T}} (f + g u^*) z G \nabla \varepsilon - \frac{1}{2} z^{\mathrm{T}} G z \nabla \varepsilon^{\mathrm{T}} (f + g u^*)$$
$$- z^{\mathrm{T}} f + x^{\mathrm{T}} P^{\mathrm{T}} G z + W^{*\mathrm{T}} \bar{\nabla} \boldsymbol{\Phi} G z, \tag{4.43}$$

$$T_3 = -\frac{1}{2} z^{\mathrm{T}} G \nabla \varepsilon \nabla \varepsilon^{\mathrm{T}} G z - \frac{1}{4} z^{\mathrm{T}} (f + g u^*) \nabla \varepsilon^{\mathrm{T}} G \nabla \varepsilon$$
$$- \frac{1}{8} z^{\mathrm{T}} G z \nabla \varepsilon^{\mathrm{T}} G \nabla \varepsilon - \frac{1}{2} z^{\mathrm{T}} G \nabla \varepsilon \nabla \varepsilon^{\mathrm{T}} (f + g u^*)$$
$$- z^{\mathrm{T}} f f^{\mathrm{T}} z - z^{\mathrm{T}} (f + g u^*) (f + g u^*)^{\mathrm{T}} \nabla \varepsilon$$
$$- \frac{1}{4} z^{\mathrm{T}} G \bar{\nabla} \boldsymbol{\Phi}^{\mathrm{T}} W^* W^{*\mathrm{T}} \bar{\nabla} \boldsymbol{\Phi} G z$$
$$- \frac{1}{8} z^{\mathrm{T}} G \nabla \varepsilon \nabla \varepsilon^{\mathrm{T}} G \nabla \varepsilon, \tag{4.44}$$

$$T_4 = -\frac{1}{8} z^{\mathrm{T}} G z z^{\mathrm{T}} G z - \frac{3}{4} z^{\mathrm{T}} (f + g u^*) z^{\mathrm{T}} G z$$
$$- \frac{3}{8} z^{\mathrm{T}} G z z^{\mathrm{T}} G \nabla \varepsilon, \tag{4.45}$$

$$T_5 = -\frac{1}{2} W^{*\mathrm{T}} \bar{\nabla} \boldsymbol{\Phi} G \bar{\nabla} \boldsymbol{\Phi}^{\mathrm{T}} W^* - \frac{1}{2} z^{\mathrm{T}} G z. \tag{4.46}$$

Now introduce bounds to Eq. (4.42).

As $G = g R^{-1} g^{\mathrm{T}}$, the rank of G is

$$\operatorname{rank}(G) = \operatorname{rank}(g) < n_x. \tag{4.47}$$

It follows that there exist kernel

$$\ker(\boldsymbol{GP}) = \{\boldsymbol{r} \in \mathbb{R}^{n_x} \mid \boldsymbol{GPr} = 0\}. \tag{4.48}$$

For nonlinear systems as in Eq. (4.1), since $\boldsymbol{x}$ and $\boldsymbol{z}$ are explicitly governed by Eq. (4.1) instead of being random, the existence of $\boldsymbol{x} = \ker(\boldsymbol{GP})$ and corresponding effects to the system is rendered negligible. Accordingly, we focus on $\boldsymbol{x} \neq \ker(\boldsymbol{GP})$ in this paper. In this case, $\boldsymbol{G}$ is positive-definite and symmetric, and under Assumptions 4.2 and 4.3, there is

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{GPx} \geq b_{m1}\|\boldsymbol{P}\|^2\|\boldsymbol{x}\|^2, \tag{4.49}$$

where constant $b_{m1} \in \mathbb{R}^+$. Also, there is $\|\boldsymbol{G}\| \leq b_G$ for constant $b_G \in \mathbb{R}^+$.

Together with Assumption 4.3, the following inequality holds:

$$\begin{aligned} T_1 &\leq \left(-\frac{1}{2}b_{m1}\|\boldsymbol{P}\|^2 + b_f\|\boldsymbol{P}\| + b_G b_\Phi \|\boldsymbol{W}^*\|\,\|\boldsymbol{P}\| + b_\Phi b_f\|\boldsymbol{W}^*\|\right)\|\boldsymbol{x}\|^2 \\ &= -\frac{1}{2}b_{m1}\|\boldsymbol{x}\|^2\left(\|\boldsymbol{P}\|^2 - \eta_1\|\boldsymbol{P}\| - \eta_2\right), \end{aligned} \tag{4.50}$$

where

$$\begin{aligned} \eta_1 &= \frac{2\left(b_f + b_G b_\Phi\|\boldsymbol{W}^*\|\right)}{b_{m1}}, \\ \eta_2 &= \frac{2b_\Phi b_f\|\boldsymbol{W}^*\|}{b_{m1}}. \end{aligned}$$

Under Assumption 4.4, if $\|\boldsymbol{P}\|^2 - \eta_1\|\boldsymbol{P}\| - \eta_2 \geq 0$, then $T_1 \leq 0$. This requires

$$\|\boldsymbol{P}\| \geq \frac{\eta_1 + \sqrt{\eta_1^2 + 4\eta_2}}{2} \triangleq b_{p1}. \tag{4.51}$$

In T_2, similarly to the case of T_1, we consider the circumstances of $\boldsymbol{x} \neq \ker(\boldsymbol{GP})$ and $\boldsymbol{z} \neq \ker(\boldsymbol{G})$. Then there is a finite constant $b_{m2} \in \mathbb{R}^+$ such that

$$\boldsymbol{z}^{\mathrm{T}}\boldsymbol{GPx}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{Gz} \geq b_{m2}\|\boldsymbol{P}\|^2\|\boldsymbol{x}\|^2\|\boldsymbol{z}\|^2. \tag{4.52}$$

Given Assumption 4.3, we have $\|\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*\| \leq b_{\dot{x}}\|\boldsymbol{x}\|$, for a finite constant $b_{\dot{x}} \in \mathbb{R}^+$. Hence,

$$\begin{aligned}
T_2 \leq & \left(-\frac{1}{4}b_{m2}\|\boldsymbol{P}\|^2+\frac{1}{2}b_G^2 b_\varepsilon\|\boldsymbol{P}\|+b_G b_f\|\boldsymbol{P}\|+\frac{1}{2}b_G^2 b_\Phi\|\boldsymbol{W}^*\|\|\boldsymbol{P}\|\right. \\
& \left.+\frac{1}{2}b_G^2 b_\Phi b_\varepsilon\|\boldsymbol{W}^*\|+b_G b_\varepsilon b_f+b_G b_\Phi b_f\|\boldsymbol{W}^*\|+\frac{3}{2}b_{\dot{x}}b_G b_\varepsilon\right)\|\boldsymbol{x}\|^2\|\boldsymbol{z}\|^2 \\
& +\left(b_G\|\boldsymbol{P}\|+b_f+b_G b_\Phi\|\boldsymbol{W}^*\|\right)\|\boldsymbol{x}\|\|\boldsymbol{z}\| \\
= & -\frac{1}{4}b_{m2}\left(\|\boldsymbol{P}\|^2-\xi_1\|\boldsymbol{P}\|-\xi_2\right)\|\boldsymbol{x}\|^2\|\boldsymbol{z}\|^2 \\
& +\left(b_G\|\boldsymbol{P}\|+b_f+b_G b_\Phi\|\boldsymbol{W}^*\|\right)\|\boldsymbol{x}\|\|\boldsymbol{z}\|, \qquad (4.53)
\end{aligned}$$

where

$$\begin{aligned}
\xi_1 &= \frac{2b_G^2 b_\Phi\|\boldsymbol{W}^*\|+2b_G^2 b_\varepsilon+4b_G b_f}{b_{m2}}, \\
\xi_2 &= \frac{(2b_G^2 b_\Phi b_\varepsilon+4b_G b_\Phi b_f)\|\boldsymbol{W}^*\|+4b_G b_\varepsilon b_f+6b_{\dot{x}}b_G b_\varepsilon}{b_{m2}}.
\end{aligned}$$

Let

$$\begin{aligned}
\eta_3 &= \|\boldsymbol{P}\|^2-\xi_1\|\boldsymbol{P}\|-\xi_2, \\
\eta_4 &= b_G\|\boldsymbol{P}\|+b_f+b_G b_\Phi\|\boldsymbol{W}^*\|.
\end{aligned}$$

Then Eq. (4.53) can be rewritten as

$$T_2 \leq -\frac{1}{4}b_{m2}\eta_3\left(\|\boldsymbol{x}\|\|\boldsymbol{z}\|-\frac{4\eta_4}{b_{m2}\eta_3}\right)\|\boldsymbol{x}\|\|\boldsymbol{z}\|. \qquad (4.54)$$

It is clear that $T_2 \leq 0$ if $\eta_3\left(\|\boldsymbol{x}\|\|\boldsymbol{z}\|-\frac{4\eta_4}{b_{m2}\eta_3}\right) \geq 0$. That is, if

$$\|\boldsymbol{P}\| > \frac{\xi_1+\sqrt{\xi_1^2+4\xi_2}}{2} \triangleq b_{p2}, \qquad (4.55)$$

then $\eta_3 > 0$, and in this case, $T_2 \leq 0$ whenever

$$\|\boldsymbol{x}\|\|\boldsymbol{z}\| \geq \frac{4\eta_4}{b_{m2}\eta_3}. \qquad (4.56)$$

Regarding T_3, for cases of $\boldsymbol{z} \neq \ker(\boldsymbol{G})$ and $\nabla\varepsilon \neq \ker(\boldsymbol{G})$, there exist constants $b_{m3},\ b_{m4},\ b_{m5},\ b_{m6} \in \mathbb{R}^+$ such that the following inequalities hold:

$$\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z} \geq b_{m3}\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|^2, \tag{4.57}$$

$$\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon \geq b_{m4}\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|^2, \tag{4.58}$$

$$\boldsymbol{z}^{\mathrm{T}}\boldsymbol{f}\boldsymbol{f}^{\mathrm{T}}\boldsymbol{z} \geq b_{m5}\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|^2, \tag{4.59}$$

$$\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\boldsymbol{z} \geq b_{m6}\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|^2. \tag{4.60}$$

Therefore, we have from Eq. (4.44) that

$$\begin{aligned} T_3 \leq & \left(-\frac{1}{2}b_{m3}\left\|\tilde{\boldsymbol{W}}\right\| - \frac{1}{8}b_{m4}\left\|\tilde{\boldsymbol{W}}\right\| - b_{m5}\left\|\tilde{\boldsymbol{W}}\right\| - \frac{1}{4}b_{m6}\left\|\tilde{\boldsymbol{W}}\right\|\right. \\ & \left. + \frac{1}{4}b_{\dot{x}}b_{\varepsilon}^2b_Gb_\Phi + \frac{1}{2}b_{\dot{x}}b_{\varepsilon}^2b_Gb_\Phi + b_{\dot{x}}^2b_\varepsilon b_\Phi + \frac{1}{8}b_{\varepsilon}^3b_G^2b_\Phi\right)\left\|\tilde{\boldsymbol{W}}\right\|\|\boldsymbol{x}\|^4 \\ = & -\eta_5\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|\left(\left\|\tilde{\boldsymbol{W}}\right\| - \frac{\eta_6}{\eta_5}\right), \end{aligned} \tag{4.61}$$

where

$$\begin{aligned} \eta_5 &= \frac{1}{2}b_{m3} + \frac{1}{8}b_{m4} + b_{m5} + \frac{1}{4}b_{m6}, \\ \eta_6 &= \frac{3}{4}b_{\dot{x}}b_{\varepsilon}^2b_Gb_\Phi + b_{\dot{x}}^2b_\varepsilon b_\Phi + \frac{1}{8}b_{\varepsilon}^3b_G^2b_\Phi. \end{aligned}$$

As a result, under Assumption 4.4, $T_3 \leq 0$ when

$$\left\|\tilde{\boldsymbol{W}}\right\| \geq \frac{\eta_6}{\eta_5}. \tag{4.62}$$

With regard to T_4, for $\boldsymbol{z} \neq \ker(\boldsymbol{G})$, there is

$$\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z} \geq b_{m7}\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|^4 \tag{4.63}$$

for a constant $b_{m7} \in \mathbb{R}^+$. Then from Eq. (4.45),

$$\begin{aligned} T_4 \leq & -\frac{1}{8}b_{m7}\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|^4 + \frac{3}{4}b_{\dot{x}}b_Gb_\Phi^3\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|^3 \\ & + \frac{3}{8}b_\varepsilon b_G^2b_\Phi^3\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|^3 \\ = & -\frac{1}{8}b_{m7}\left(\left\|\tilde{\boldsymbol{W}}\right\| - \frac{6b_{\dot{x}}b_Gb_\Phi^3 + 3b_\varepsilon b_G^2b_\Phi^3}{b_{m7}}\right)\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|^3. \end{aligned} \tag{4.64}$$

Therefore, under Assumption 4.4, $T_4 \leq 0$ requires

$$\left\|\tilde{\boldsymbol{W}}\right\| \geq \frac{6b_{\dot{x}}b_G b_\Phi^3 + 3b_\varepsilon b_G^2 b_\Phi^3}{b_{m7}}. \tag{4.65}$$

It is easy to see from Eq. (4.46) that $T_5 \leq 0$. Thus, it can be concluded from Eqs. (4.51), (4.55), (4.56), (4.62), and (4.65) that Eq. (4.41) is negative if

$$\|\boldsymbol{P}\| > \max(b_{p1}, b_{p2}), \tag{4.66}$$

$$\|\boldsymbol{x}\| \, \|\boldsymbol{z}\| > \frac{4\eta_4}{b_{m2}\eta_3}, \tag{4.67}$$

$$\left\|\tilde{\boldsymbol{W}}\right\| > \max\left(\frac{\eta_6}{\eta_5}, \frac{6b_{\dot{x}}b_G b_\Phi^3 + 3b_\varepsilon b_G^2 b_\Phi^3}{b_{m7}}\right). \tag{4.68}$$

Since $\boldsymbol{z} = \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}$, and $\boldsymbol{x}$ is PE under Assumption 4.4, Eq. (4.67) also establishes a bound for $\|\tilde{\boldsymbol{W}}\|$. Thus, Eqs. (4.67) and (4.68) together, show that $\tilde{\boldsymbol{W}}$ is UUB during online learning. To be specific, there exists a bound $b_{\tilde{W}} \in \mathbb{R}^+$ such that $\tilde{\boldsymbol{W}} \leq b_{\tilde{W}}$.

In the following, the closed-loop stability of system states are examined. The corresponding Lyapunov function candidate is $\mathcal{L}_V$ in Eq. (4.27), the time derivative of which, according to Eq. (4.29), can be rewritten as

$$\dot{\mathcal{L}}_V = \dot{\mathcal{L}}_{V1} + \dot{\mathcal{L}}_{V2}, \tag{4.69}$$

where

$$\begin{aligned}\dot{\mathcal{L}}_{V1} = {} & \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{f} + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} \\ & - \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* \\ & + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}},\end{aligned} \tag{4.70}$$

$$\dot{\mathcal{L}}_{V2} = -\frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \frac{1}{2}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*. \tag{4.71}$$

With $\|\tilde{\boldsymbol{W}}\| \leq b_{\tilde{W}}$, together with Assumptions 4.2 and 4.3, the following inequality yields from Eq. (4.70):

$$\begin{aligned}\dot{\mathcal{L}}_{V1} \leq & (-\frac{1}{2}b_{m1}\|\boldsymbol{P}\|^2 + b_f \|\boldsymbol{P}\| + b_G b_\phi \|\boldsymbol{W}^*\| \, \|\boldsymbol{P}\| + b_G b_\phi b_{\tilde{W}} \|\boldsymbol{P}\| \\ & + b_f b_\phi \|\boldsymbol{W}^*\| + b_G b_\phi^2 b_{\tilde{W}} \|\boldsymbol{W}^*\| + b_f b_\phi b_{\tilde{W}})\|\boldsymbol{x}\|^2 \\ = & -\frac{1}{2}b_{m1}\|\boldsymbol{x}\|^2 \left(\|\boldsymbol{P}\|^2 - \xi_3 \|\boldsymbol{P}\| - \xi_4\right),\end{aligned} \tag{4.72}$$

where

$$\xi_3 = \frac{2b_f + 2b_G b_\phi \|\mathbf{W}^*\| + 2b_G b_\phi b_{\tilde{W}}}{b_{m1}},$$
$$\xi_4 = \frac{2b_f b_\phi \|\mathbf{W}^*\| + 2b_G b_\phi^2 b_{\tilde{W}} \|\mathbf{W}^*\| + 2b_f b_\phi b_{\tilde{W}}}{b_{m1}}.$$

Under Assumption 4.4, we have $\dot{\mathcal{L}}_{V1} \leq 0$ as long as $\|\mathbf{P}\|^2 - \xi_3 \|\mathbf{P}\| - \xi_4 \geq 0$, requiring

$$\|\boldsymbol{P}\| \geq \frac{\xi_3 + \sqrt{\xi_3^2 + 4\xi_4}}{2} \triangleq b_{p3}. \tag{4.73}$$

Since $\dot{\mathcal{L}}_{V2} \leq 0$, it follows from Eqs. (4.66) and (4.73) that $\dot{\mathcal{L}}_V \leq 0$ if

$$\|\boldsymbol{P}\| > \max(b_{p1}, b_{p2}, b_{p3}) \triangleq b_{mP}. \tag{4.74}$$

At this stage, upon satisfaction of Eqs. (4.67), (4.68), and (4.74), the UUB stability holds for the NN weights estimation errors $\tilde{\boldsymbol{W}}$ and the system states $\boldsymbol{x}$.

It further follows that $\ddot{\mathcal{L}}_V = \frac{d\dot{\mathcal{L}}_V}{dt}$ is a function of $\tilde{\boldsymbol{W}}$ and $\boldsymbol{x}$, and $\ddot{\mathcal{L}}_V$ is also bounded since $\|\tilde{\boldsymbol{W}}\|$ and $\|\boldsymbol{x}\|$ are bounded. As a result, asymptotic stability also holds true for the system states $\boldsymbol{x}$ throughout online training. □

Remark 4.10. As can be seen from Theorem 4.2, the proposed MVFA establishes a direct link to the closed-loop stability through the auxiliary quadratic term with a design parameter matrix $\boldsymbol{P}$. With the MVFA, no special stabilising tuning laws are required for the NNs in critic and actor, and during online learning the SPI under the resulted single-critic configuration remains stable with simple gradient descent tuning.

Remark 4.11. Upon initialisation of $\boldsymbol{P}$ and $\hat{\boldsymbol{W}}$, Eq. (4.20) delivers an initial admissible control at the time when the algorithm starts. Specifically, $\boldsymbol{P}$ can be chosen following Theorem 4.2 with some trial-and-error, while the initialisation of $\hat{\boldsymbol{W}}$ is trivial which can be small random numbers or simply zeros.

4.5 Numerical studies

This section presents a simulation example of finding the optimal control law for a nonlinear model with a known value function to verify the proposed method.

The following nonlinear system is considered (Vamvoudakis and Lewis, 2010), with

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} -x_1 + x_2 \\ -0.5x_1 - 0.5x_2 \left\{1 - [\cos(2x_1) + 2]^2\right\} \end{bmatrix},$$

and

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{bmatrix} 0 \\ \cos(2x_1) + 2 \end{bmatrix}.$$

For $\boldsymbol{Q} = \boldsymbol{I}_{2\times 2}$ and $R = 1$, the corresponding V^* and $\boldsymbol{u}^*$ are known to be

$$V^* = \frac{1}{2}x_1^2 + x_2^2, \tag{4.75}$$

and

$$u^* = -\left[\cos(2x_1) + 2\right] x_2, \tag{4.76}$$

respectively, as given in Vamvoudakis and Lewis (2010).

The critic NN has activation functions of

$$\boldsymbol{\Phi} = [x_1^2,\ x_1x_2,\ x_2^2]^{\mathrm{T}},$$

with NN weights being

$$\hat{\boldsymbol{W}} = [\hat{W}_1,\ \hat{W}_2,\ \hat{W}_3]^{\mathrm{T}}.$$

In simulation, $\boldsymbol{P} = 10\boldsymbol{I}_{2\times 2}$, $\alpha = 10$, and $\boldsymbol{N}_3 = [5,\ 1]^{\mathrm{T}}$. System states $\boldsymbol{x}$ and NN weights $\hat{\boldsymbol{W}}$ are initialised to zeros. An exogenous signal

$$\begin{aligned} u_e(t) = 2[&\cos(0.8t) + \sin(t)^2\cos(t) + \sin(2t)^2\cos(0.1t) \\ &+ \sin(-1.2t)^2\cos(0.5t) + \sin(t)^5] \end{aligned}$$

is used to perturb the system for exploration. Note that the total control that enters the process during exploration is the sum of $\hat{u}$ and u_e, which also perturbs the system states $\boldsymbol{x}$ in the meantime. For efficient and effective training with Eqs. (4.21)

and (4.22) involved, exploration is implemented in the following manner: The excitation of $u_e(t)$ lasts 0.05 s for every 0.1 s time interval, while the HJB error e_2 in Eq. (4.21) is periodically fed back for calculation during the intervals when $u_e(t)$ is temporarily off (i.e., $e_2 = 0$ if $u_e(t) \neq 0$). $u_e(t)$ is completely turned off at 40 s. The trajectories of system states $\boldsymbol{x}$, approximated optimal control $\hat{u}$ and the excitation signal u_e during online learning are plotted in Figures 4.1 and 4.2, respectively. Close-up of the excitation signal u_e for the first 2 seconds is shown in Figure 4.3 for clearer illustration of the special excitation implemented. Weights convergence history of the critic NN is given in Figure 4.4.

Note in Figure 4.2 that the approximated optimal control $\hat{u}$ generally mirrors the contour of the excitation signal u_e with slight difference in amplitude. It shows $\hat{u}$ effectively counteracts u_e and maintains closed-loop system states stability during online training.

Figure 4.4 shows that all NN weights settle within 10 s. At the end of training,

$$\hat{\boldsymbol{W}} = [-4.4999,\ -0.0003,\ -3.9996]^{\mathrm{T}}.$$

This yields

$$\begin{aligned}\hat{V}(x) &= 0.5001{x_1}^2 - 0.0003x_1x_2 + 1.0004{x_2}^2 \\ &\approx \frac{1}{2}x_1^2 + x_2^2,\end{aligned}$$

and

$$\begin{aligned}\hat{u}(x) &= -\left[\cos(2x_1) + 2\right](-0.0002x_1 + 1.0004x_2) \\ &\approx -\left[\cos(2x_1) + 2\right]x_2,\end{aligned}$$

which are close approximation to Eqs. (4.75) and (4.76), showing that the convergence of NN weights is reached with good accuracy.

In situations when the PE condition may not be satisfied (for example, the closed-loop response is subject to none-zero initial conditions only), $\hat{\boldsymbol{W}}$ may not reach its ideal set $\boldsymbol{W}^*$ as a result. In the following simulations, the convergence of NN parameters and the closed-loop stability is investigated under unsatisfied PE condition.

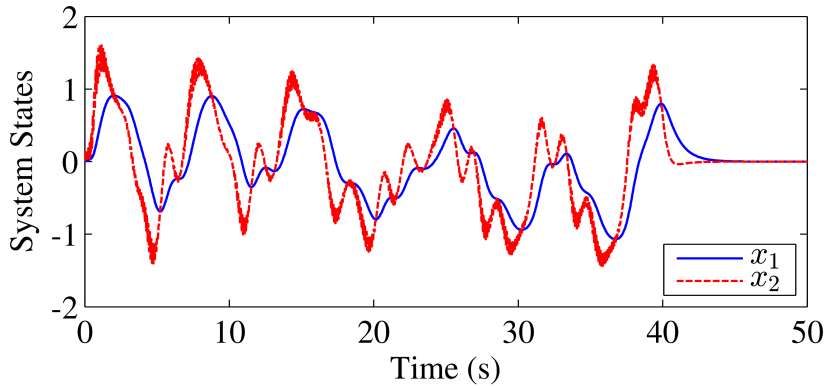


Figure 4.1: Trajectories of system states during online tuning.

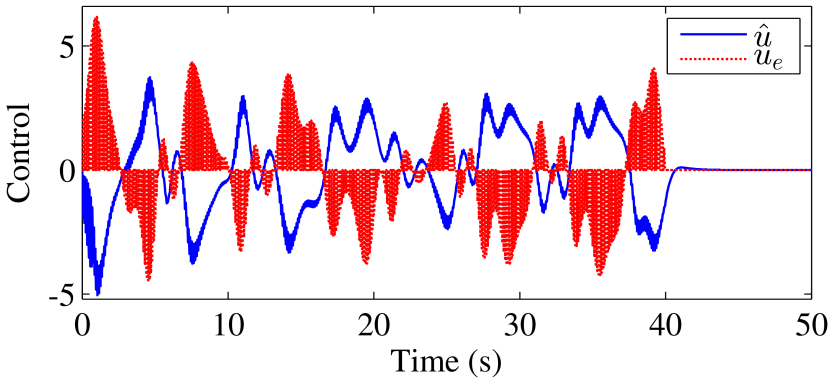


Figure 4.2: Trajectories of control signals during online tuning.

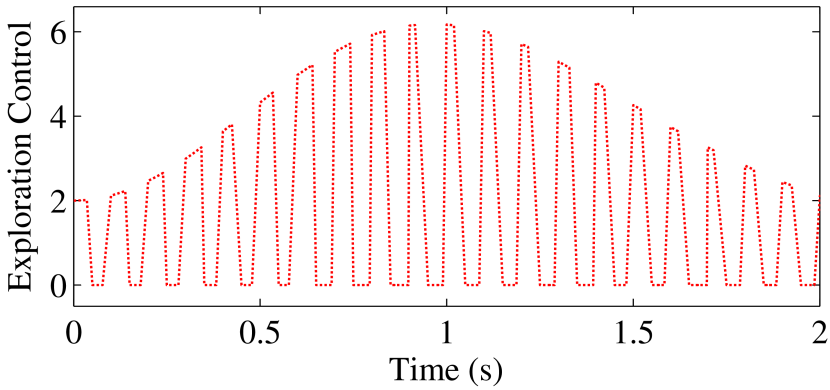


Figure 4.3: Close-up of excitation signal u_e for the first 2 seconds.

Accordingly, $\boldsymbol{x}(0) = [0.5 \quad 0]^{\mathrm{T}}$ is applied as an initial condition, no probing noise is added, and controller parameters remain the same. The corresponding closed-loop states responses are plotted in Figure 4.5, and the corresponding control action is given in Figure 4.6, together with responses under the ideal optimal control supplied for comparison. As can be seen from the figures, states trajectories and control signal under the proposed control scheme are similar to those of the ideal optimal

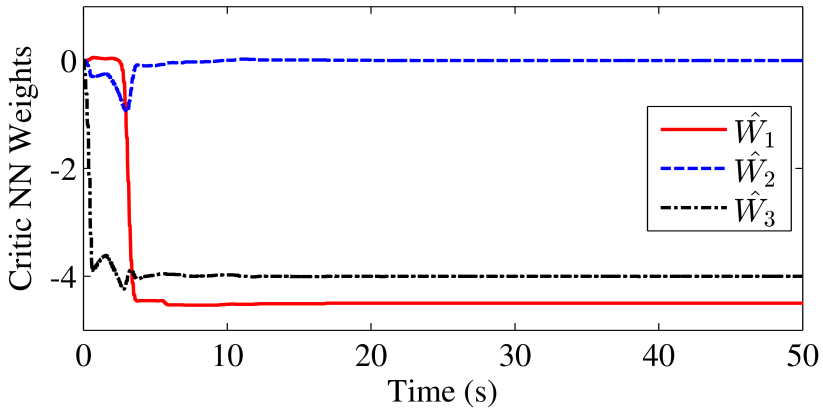


Figure 4.4: NN parameters convergence during online tuning.

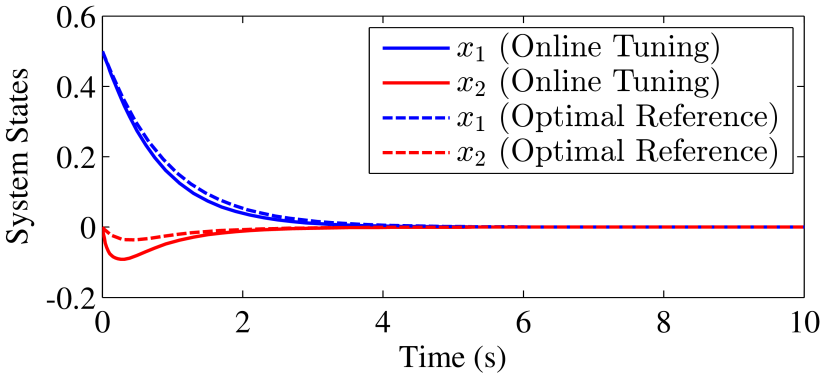


Figure 4.5: State trajectories of the closed-loop response to the non-zero initial condition under the proposed online tuning scheme (PE unsatisfied in this case) and the known ideal optimal control.

control. The difference in response is due to the approximation error resulted from lack of PE. The NN parameters convergence history is plotted in Figure 4.7, where the settling value of $\hat{W}_2$ and $\hat{W}_3$ is still far from the ideal one. However, stable closed-loop responses are observed under the proposed algorithm regardless of the differences, as shown by Figures 4.5 and 4.6. The cost of the closed-loop response to the none-zero initial condition under the proposed algorithm (i.e., $\hat{V}(\boldsymbol{x}(0))$) together with that under the known ideal optimal control (i.e., $V^*(\boldsymbol{x}(0))$) are evaluated in Figure 4.8. By recalculating the cost using the continuously updated NN weights, the approximated value function $\hat{V}$ is shown to be converging to the optimal one, in the presence of some approximation error.

Remark 4.12. As shown by the example, the compact controller under the single-critic configuration with the proposed MVFA is able to maintain the closed-loop stability

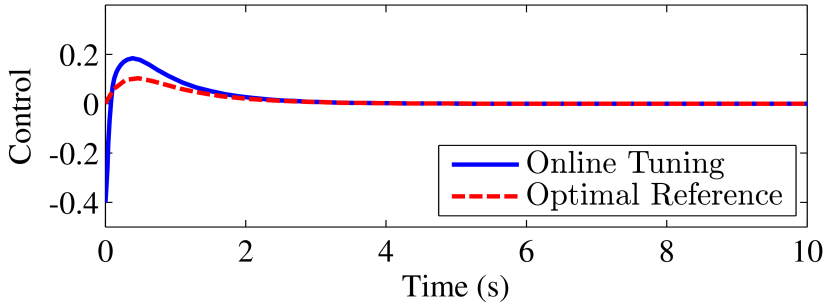


Figure 4.6: Control input in response to the non-zero initial condition under the proposed online tuning scheme (PE unsatisfied in this case) and the known ideal optimal control.

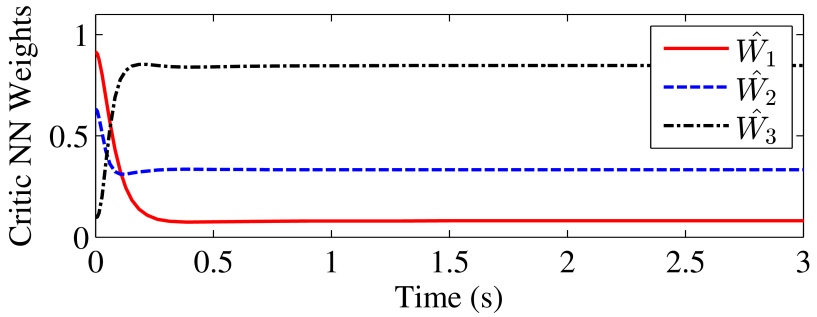


Figure 4.7: NN parameters convergence history during the closed-loop response to the non-zero initial condition (PE unsatisfied case).

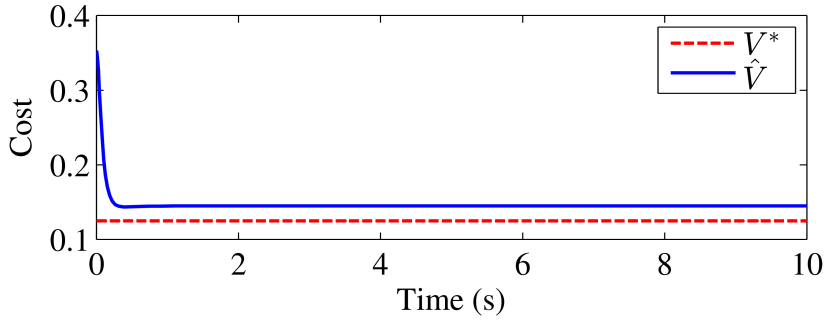


Figure 4.8: The minimal cost $V^*(\boldsymbol{x}(0))$ of the closed-loop response to the non-zero initial condition and the evolution of the approximated $\hat{V}(\boldsymbol{x}(0))$ (PE unsatisfied case).

during online learning with the traditional yet simple gradient descent tuning law without stabilising mechanisms in either critic or actor tuning. The proposed MVFA does not complicate the problem either, as can be seen from the parameter selection and initialisation setting.

4.6 Conclusions

It is shown in stability analysis that using the proposed MVFA to provide alternative realisation of the single-critic configuration for SPI is feasible and effective. The proposed method eliminates the need for stabilising mechanisms in either the critic or actor NN tuning, without jeopardising the closed-loop stability, and without complicating the problem, as confirmed in theoretical proof and demonstrated in numerical studies. In general, the proposed MVFA used in a single-critic configuration for SPI, together with the study on parameters convergence and the closed-loop stability, serve as a new development to the online SPI theory framework.

It is worth noting that the proposed MVFA scheme in this paper is model-based. Many successful model-free applications (Abouheaf et al., 2018; Luo et al., 2018; Radac et al., 2018) have motivated future works on advanced model-free MVFA based schemes that: (1) feature better adaptability and robustness in circumstances with complex, unknown, uncertain or time-varying system dynamics; (2) deliver simplified online implementation enabled by the MVFA approach.

Chapter 5

Flutter suppression by input-unconstrained optimal control

The novel nonlinear optimal control online synthesis (NOCOS) scheme proposed in Chapter 4 needs to be generalised to globally nonlinear scenarios so that it suits active flutter suppression (AFS) applications, as discussed in Chapter 2. This chapter, based on *Article-3*, presents details of the proposed adaptive nonlinear optimal controller based on the novel NOCOS scheme, together with wind-tunnel experiment results obtained using the VSDS described in Chapter 3. By assuming no control-input constraints (CICs), the third objective (Page 4) of the thesis is fulfilled by the work in this chapter, which, at this stage, partially accomplishes the thesis aim.

Statement of Authorship

Title of Paper	Adaptive nonlinear optimal control for active suppression of airfoil flutter via a novel neural-network-based controller
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Tang D, Chen L, Tian ZF and Hu E (2018) Adaptive nonlinear optimal control for active suppression of airfoil flutter via a novel neural-network-based controller. *Journal of Vibrations and Control*, DOI: 10.1177/1077546317750504.

Principal Author

Name of Principal Author (Candidate)	Difan Tang		
Contribution to the Paper	Proposing and deriving methods and theories, performing theoretical analysis, conducting experiments, collecting data, interpreting results, and writing manuscript.		
Overall percentage (%)	80 %		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	5 April 2018

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Lei Chen		
Contribution to the Paper	Supervising development of work, evaluating methods and theories, helping with setting up experiments, providing advice in mathematical aspects, and evaluating manuscript.		
Signature		Date	5/04/18

Name of Co-Author	Zhao Feng Tian		
Contribution to the Paper	Supervising development of work, providing advice in mathematical aspects, and evaluating manuscript.		
Signature		Date	05/04/18

Name of Co-Author	Eric Hu		
Contribution to the Paper	Supervising development of work, providing advice in mathematical aspects, and evaluating manuscript.		
Signature		Date	12/4/18

abstract

This paper proposes a novel adaptive nonlinear controller based on neural networks (NNs) for active flutter suppression (AFS) on aerofoils from the optimal control perspective. A four-degrees-of-freedom aeroelastic system that has nonlinear translational and torsional stiffness and employs leading- and trailing-edge control surfaces as control inputs is considered. Optimal control for the nonlinear aeroelastic system at a constant airspeed is synthesised by solving the Hamilton-Jacobi-Bellman equation through synchronous policy iteration with a Modified form of NN-based Value Function Approximation (MVFA). An extended Kalman filter is proposed to tune the MVFA. A systematic procedure involving linear matrix inequalities is further proposed for designing a scheduled parameter matrix to generalise the MVFA to globally nonlinear cases where the aeroelastic dynamics vary nonlinearly with the airspeed. An identifier NN is also derived to capture un-modelled dynamics in real time. Parameter convergence and the closed-loop stability are examined through the Lyapunov stability analysis. Comparisons drawn with a linear-parameter-varying optimal controller in wind-tunnel experiments confirm the effectiveness and validity of the proposed control scheme.

5.1 Introduction

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

Suppressing LCOs in broad engineering practice can be done using various control methods (Chen et al., 2009; Keyser et al., 2017; Saaed et al., 2017). Similarly in terms of AFS, a wide variety of control strategies are available. In early studies, there are non-adaptive classical and modern control (Edwards, 1983; Mukhopadhyay et al., 1981; Newsom and Mukhopadhyay, 1985; Schmidt and Chen, 1986; Waszak, 2001) as well as adaptive online linear quadratic regulator (LQR) that updates the control in real time to suit the changing dynamics (Friedmann et al., 1997; Guillot and Friedmann, 2000; Pak et al., 1995). In recent studies, conventional frequency-domain analysis remains a useful tool for control synthesis (Schmidt, 2016), while advanced methods in adaptive, nonlinear, and robust control have received more attention due to the time-varying nature and nonlinear characteristics of an aeroelastic system (Bichiou et al., 2016; Biskri et al., 2008; Chen and Liu, 2010; Nayfeh et al., 2012; Rebolho et al., 2014; Vasconcellos et al., 2016a) and the increasing demand on a wider operation range beyond the flutter boundary. These advanced methods include but are not limited to: optimal control synthesised via time-domain finite elements method (Fazelzadeh et al., 2014), self-tuning regulator (Viswamurthy and Ganguli, 2008), linear-parameter-varying techniques (Chen et al., 2012; Prime, 2010), feedback linearisation (Ko et al., 1997, 1998, 1999; Platanitis and Strganac, 2004; Strganac et al., 2000), model reference adaptive control (Ko et al., 2002), back-stepping-based adaptive output feedback control (Singh and Wang, 2002; Xing and Singh, 1999, 2000), robust output feedback control (Zhang and Behal, 2016), modular

adaptive control (Bhoir and Singh, 2004; Rao et al., 2006; Singh and Brenner, 2003), modified filtered-X least-mean-square control (Carnahan and Richards, 2008), $\mathcal{L}_1$ adaptive control (Lee and Singh, 2013), sliding-mode control (Dilmi and Bouzouia, 2016; Luo et al., 2016b; Wang et al., 2015), finite-time H_∞ adaptive fault-tolerant control (Gao and Cai, 2016; Gao et al., 2016) and neural networks (NNs) based adaptive control (Brillante and Mannarino, 2016; Gujjula et al., 2005; Spencer et al., 1999, 2002; Wang et al., 2011), etc.

However, optimal controllers among the mentioned methods are sensitive to modelling errors, which means suboptimal or unsatisfactory performance may result in the presence of uncertainties or faults. Though some other controllers are designed to be more adaptive to the changing environments and tolerant to un-modelled dynamics, these methods do not provide nonlinear optimal control. These two problems, together being a dilemma in AFS controller synthesis, have nevertheless not been addressed. Therefore, the study in this paper proposes an approach that synthesises nonlinear optimal control in real time for AFS according to online updated aeroelastic dynamics, aiming to reduce the impact of the aforementioned two problems.

Optimal control for nonlinear systems involves iteratively solving a nonlinear Hamilton-Jacobi-Bellman (HJB) equation for the associated value function via NN-based approximation. Most existing methods for online synthesis of nonlinear optimal control (NOCOS) employ a common form of value function approximation (VFA), subject to limitations related to stability and algorithm structure, as commented in Tang et al. (2015). Accordingly, a solution to these limitations arising from the use of the traditional VFA is addressed in the work of Tang et al. (2015), where a modified value function approximation (MVFA) is proposed. Nonetheless, all the existing NOCOS methods are 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

NOCOS methods inapplicable to AFS without modification and improvement.

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

- Under the locally nonlinear setting, the MVFA is proposed for the AFS controller in the interests of compact algorithm structure suitable for real-time implementation. An extended Kalman filter is proposed to tune the MVFA online. NN parameters convergence and the closed-loop stability are examined through the Lyapunov stability analysis.

- A systematic procedure based on linear matrix inequalities is further purposed for the design of a scheduled parameter matrix for the MVFA to generalise the proposed method to globally nonlinear cases, so that the proposed NN controller suits AFS applications.

- The proposed method successfully solves the aforementioned controller synthesis dilemma involved in AFS applications, with the capability of learning in real time to improve AFS performance from the nonlinear optimal control perspective. 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.

The remainder of the paper is arranged as follows. The nonlinear aeroelastic model under consideration is introduced in Section 5.2. A new adaptive control scheme featuring NN-based NOCOS with integrated system identification is proposed in Section 5.3. Experiment results are presented and discussed in Section 5.4. Conclusions are drawn in Section 5.5.

5.2 Aeroelastic system

To provide an appropriate aeroelasticity platform for the investigation and discussion of AFS under the new controller proposed in this study, a typical rigid aerofoil

section featuring two-dimensional vibration modes (i.e. the first plunge and first pitch mode oscillations) is considered for its well-established theory basis and experimental validation (Ko et al., 2002; O'Neil and Strganac, 1998; Platanitis and Strganac, 2004; Prime et al., 2010; Prime, 2010; Strganac et al., 2000). 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 et al. (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 aerofoil 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 aerofoil 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 modelling 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 5.1.

The aeroelastic system has equation of motion

$$\boldsymbol{\psi}\ddot{\boldsymbol{q}} + \boldsymbol{\Lambda} = \boldsymbol{F}, \tag{5.1}$$

for which

$$\ddot{\boldsymbol{q}} = \begin{bmatrix} \ddot{q}_h & \ddot{q}_a & \ddot{q}_{te} & \ddot{q}_{le} \end{bmatrix}^{\mathrm{T}},$$

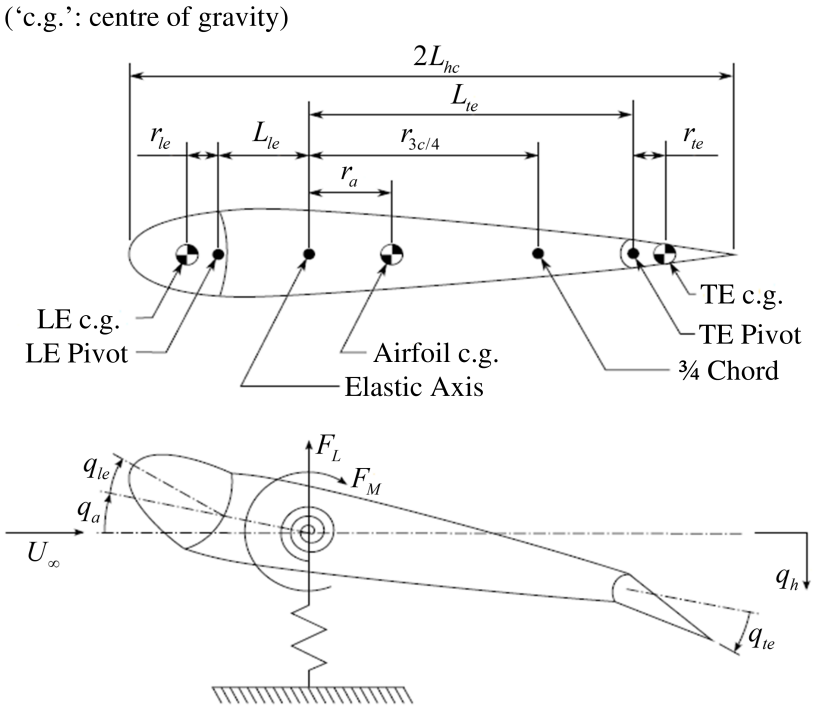


Figure 5.1: Schematic figure of the 4-DOF aeroelastic system ('LE' – leading edge; 'TE' – trailing edge; 'c.g.' – centre of gravity).

$$\boldsymbol{\psi} = \begin{bmatrix} \psi_{11} & \psi_{12} & \psi_{13} & \psi_{14} \\ \psi_{21} & \psi_{22} & \psi_{23} & \psi_{24} \\ \psi_{31} & \psi_{32} & I_{te} & 0 \\ \psi_{41} & \psi_{42} & 0 & I_{le} \end{bmatrix},$$

$$\boldsymbol{\Lambda} = \begin{bmatrix} \Lambda_1 & \Lambda_2 & \Lambda_3 & \Lambda_4 \end{bmatrix}^{\mathrm{T}},$$

$$\boldsymbol{F} = \begin{bmatrix} -F_L & F_M & k_{te} q_{te\text{-}cmd} & k_{le} q_{le\text{-}cmd} \end{bmatrix}^{\mathrm{T}},$$

with

$$\psi_{11} = m_a + m_{te} + m_{le},$$

$$\psi_{12} = \psi_{21} = (m_a r_a + m_{te} L_{te} + m_{le} L_{le}) \cos(q_a) + m_{te} r_{te} \cos(q_a + q_{te}) + m_{le} r_{le} \cos(q_a + q_{le}),$$

$$\psi_{13} = \psi_{31} = m_{te} r_{te} \cos(q_a + q_{te}),$$

$$\psi_{14} = \psi_{41} = m_{le} r_{le} \cos(q_a + q_{le}),$$

$$\psi_{22} = I_a + I_{te} + I_{le} + m_{te}L_{te}^2 + m_{le}L_{le}^2 + 2L_{te}m_{te}r_{te}\cos(q_{te}) + 2L_{le}m_{le}r_{le}\cos(q_{le}),$$

$$\psi_{23} = \psi_{32} = I_{te} + L_{te}m_{te}r_{te}\cos(q_{te}),$$

$$\psi_{24} = \psi_{42} = I_{le} + L_{le}m_{le}r_{le}\cos(q_{le}),$$

$$\Lambda_1 = k_h q_h + c_h \dot{q}_h - (\dot{q}_a + \dot{q}_{te})^2 m_{te}r_{te}\sin(q_a + q_{te}) - (\dot{q}_a + \dot{q}_{le})^2 m_{le}r_{le}\sin(q_a + q_{le}) - \dot{q}_a^2 \sin(q_a)(m_a r_a + m_{te}L_{te} + m_{le}L_{le}),$$

$$\Lambda_2 = k_a q_a + c_a \dot{q}_a - \dot{q}_{te}(\dot{q}_{te} + 2\dot{q}_a)m_{te}r_{te}L_{te}\sin(q_{te}) - \dot{q}_{le}(\dot{q}_{le} + 2\dot{q}_a)m_{le}r_{le}L_{le}\sin(q_{le}),$$

$$\Lambda_3 = k_{te}q_{te} + c_{te}\dot{q}_{te} + \dot{q}_a^2 m_{te}r_{te}L_{te}\sin(q_{te}),$$

$$\Lambda_4 = k_{le}q_{le} + c_{le}\dot{q}_{le} + \dot{q}_a^2 m_{le}r_{le}L_{le}\sin(q_{le}),$$

$$F_L = \rho U_\infty^2 L_{hc}L_s C_{l\text{-}a}(q_a + \frac{\dot{q}_h}{U_\infty} + r_{3c/4}\frac{\dot{q}_a}{U_\infty}) + \rho U_\infty^2 L_{hc}L_s C_{l\text{-}te}q_{te} + \rho U_\infty^2 L_{hc}L_s C_{l\text{-}le}q_{le},$$

$$F_M = \rho U_\infty^2 L_{hc}^2 L_s C_{me\text{-}a}(q_a + \frac{\dot{q}_h}{U_\infty} + r_{3c/4}\frac{\dot{q}_a}{U_\infty}) + \rho U_\infty^2 L_{hc}^2 L_s C_{me\text{-}te}q_{te} + \rho U_\infty^2 L_{hc}^2 L_s C_{me\text{-}le}q_{le},$$

$$C_{l\text{-}a} = \frac{\partial C_l}{\partial q_a}, \quad C_{l\text{-}te} = \frac{\partial C_l}{\partial q_{te}}, \quad C_{l\text{-}le} = \frac{\partial C_l}{\partial q_{le}},$$

$$C_{m\text{-}a} = \frac{\partial C_m}{\partial q_a}, \quad C_{m\text{-}te} = \frac{\partial C_m}{\partial q_{te}}, \quad C_{m\text{-}le} = \frac{\partial C_m}{\partial q_{le}},$$

$$C_{me\text{-}a} = \frac{r_{fc}}{L_{hc}}C_{l\text{-}a} + 2C_{m\text{-}a},$$

$$C_{me\text{-}te} = \frac{r_{fc}}{L_{hc}}C_{l\text{-}te} + 2C_{m\text{-}te},$$

$$C_{me\text{-}le} = \frac{r_{fc}}{L_{hc}}C_{l\text{-}le} + 2C_{m\text{-}le},$$

where geometry and force related parameters and variables are defined in Figure 5.1, and other terms are defined as

q_h, q_a, q_{te}, q_{le}:	translational/angular displacements;
$q_{te\text{-}cmd}$, $q_{le\text{-}cmd}$:	TE and LE commands;
m_a:	aerofoil mass (excluding TE and LE);
m_{te}, m_{le}:	mass of TE and LE;
I_a:	aerofoil rotational inertia (excluding TE and LE) about its elastic axis;
I_{te}, I_{le}:	TE and LE rotational inertia about respective pivot;
k_h, k_a, k_{te}, k_{le}:	stiffness coefficients;
c_h, c_a, c_{te}, c_{le}:	damping coefficients;
ρ:	air density;
L_s:	aerofoil span;
C_l:	aerofoil lift coefficient;
C_m:	aerofoil moment coefficient at 1/4-chord;
U_∞:	airflow velocity.

The system as in Eq. (5.1) can be transformed into

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(U_\infty, \boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})\boldsymbol{u}, \tag{5.2}$$

with $\boldsymbol{f}(\boldsymbol{x}, U_\infty) \in \mathbb{R}^{n_x}$ being drift dynamics, $\boldsymbol{g}(\boldsymbol{x}) \in \mathbb{R}^{n_x \times n_u}$ denoting control distribution, $\boldsymbol{x} \in \mathbb{R}^{n_x}$ being a states vector of dimension n_x, and $\boldsymbol{u} \in \mathbb{R}^{n_u}$ containing n_u control inputs.

Specifically,

$$\boldsymbol{x} = \begin{bmatrix} q_h & q_a & q_{te} & q_{le} & \dot{q}_h & \dot{q}_a & \dot{q}_{te} & \dot{q}_{le} \end{bmatrix}^{\mathrm{T}},$$

$$\boldsymbol{u} = \begin{bmatrix} q_{te\text{-}cmd} & q_{le\text{-}cmd} \end{bmatrix}^{\mathrm{T}},$$

$$\boldsymbol{f}(\boldsymbol{x}, U_\infty) = \begin{bmatrix} \dot{\boldsymbol{q}} \\ \boldsymbol{\psi}^{-1}\bar{\boldsymbol{F}} \end{bmatrix}, \quad \boldsymbol{g}(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{\psi}^{-1}\bar{\boldsymbol{g}} \end{bmatrix},$$

with

$$\bar{\boldsymbol{F}} = \begin{bmatrix} F_L - \Lambda_1 \\ F_M - \Lambda_2 \\ -\Lambda_3 \\ -\Lambda_4 \end{bmatrix}, \quad \bar{\boldsymbol{g}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ k_{te} & 0 \\ 0 & k_{le} \end{bmatrix}.$$

5.3 Proposed controller

Aeroelastic dynamics in general, as shown in Eqs. (5.1) and (5.2), are nonlinear for a constant freestream airflow velocity (i.e. locally nonlinear) and also change nonlinearly with respect to the airspeed (i.e. globally nonlinear). As discussed in Section 5.1, existing methods capable of synthesising optimal control laws for nonlinear systems are all limited to locally nonlinear cases and in the meantime subject to other practical limitations. Therefore, there are no existing methods suitable for direct implementation in AFS without modification and improvements. In this section, preliminaries are given in Subsection 5.3.1; NOCOS with the MVFA tuned by an EKF for locally nonlinear systems is proposed in Subsection 5.3.2, with convergence and stability properties discussed; The discussion then moves forward with a systematic procedure proposed in Subsection 5.3.3 to generalise the new control method to globally nonlinear cases, specifically, for AFS; As the proposed controller requires some dynamics of the aeroelastic system to be known, an online system identification scheme is introduced in Subsection 5.3.4 to tackle this.

5.3.1 Continuous-time HJB equation and policy iteration

For a fixed velocity U_∞, Eq. (5.2) can be reduced to:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t))|_{U_\infty = U} + \boldsymbol{g}(\boldsymbol{x}(t))\boldsymbol{u}(t); \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{5.3}$$

which can be written in a compact form:

$$\dot{\boldsymbol{x}}(t) = \mathcal{F}_A(\boldsymbol{x}(t), \boldsymbol{u}(t))|_{U_\infty = U}; \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{5.4}$$

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

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

The properties of (5.2) are given by Assumptions 4.1 and 4.2 in Chapter 4.

The control problem is to determine a control policy/law $\boldsymbol{u}(t)$ to minimise the following performance index (cost function):

$$V(\boldsymbol{x}_0) = \int_0^\infty [\bar{Q}(\boldsymbol{x}(\tau)) + \bar{R}(\boldsymbol{u}(\tau))]d\tau, \tag{5.5}$$

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

Differentiating Eq. (5.5) yields its infinitesimal version that is a nonlinear Lyapunov equation (Abu-Khalaf and Lewis, 2005), written as:

$$\nabla \boldsymbol{V}^{\mathrm{T}}(\boldsymbol{x})\mathcal{F}_A(\boldsymbol{x}, \boldsymbol{u}) + \bar{Q}(\boldsymbol{x}) + \bar{R}(\boldsymbol{u}) = 0; \quad V(0) = 0. \tag{5.6}$$

Let $V^*(\boldsymbol{x})$ denote the optimal (minimal) cost function, named as the 'value function', and let $\nabla \boldsymbol{V}^*(\boldsymbol{x}) \triangleq \partial V^*(\boldsymbol{x})/\partial \boldsymbol{x}$ denote its derivative with respect to $\boldsymbol{x}$, The corresponding optimal control policy is then given by:

$$\boldsymbol{u}^*(\boldsymbol{x}) = -\frac{1}{2}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{x})\nabla \boldsymbol{V}^*(\boldsymbol{x}), \tag{5.7}$$

which satisfies the Hamilton-Jacobi-Bellman (HJB) equation based on Eq. (5.6):

$$-\frac{1}{4}[\nabla \boldsymbol{V}^*(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x})\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{x})\nabla \boldsymbol{V}^*(\boldsymbol{x}) + \bar{Q}(\boldsymbol{x}) + [\nabla \boldsymbol{V}^*(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{f}(\boldsymbol{x}) = 0; \quad V^*(0) = 0. \tag{5.8}$$

That is, by solving Eq. (5.8) for $V^*(\boldsymbol{x})$, the optimal control policy can then be obtained as in Eq. (5.7), given that the system internal dynamics $\boldsymbol{f}(\boldsymbol{x})$ and control input dynamics $\boldsymbol{g}(\boldsymbol{x})$ are known.

5.3.2 Neural-network-based value function approximation

Note that the HJB equation is nonlinear and analytically solving for $V^*(\boldsymbol{x})$ is difficult. Instead, the policy-iteration approach applies. To allow implementation of the policy iteration, an appropriately structured representation of $V^*(\boldsymbol{x})$ is necessary, which

can be a neural-network (NN) approximation. For this purpose and for AFS, the MVFA proposed in Tang et al. (2015) is adopted, being:

$$V^*(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} + \boldsymbol{W}^{*\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{x}) + \varepsilon(\boldsymbol{x}), \tag{5.9}$$

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

In regard to Eq. (5.9), the derivative of $V^*(\boldsymbol{x})$ with respect to $\boldsymbol{x}$ is then given by:

$$\nabla \boldsymbol{V}^*(\boldsymbol{x}) \triangleq \frac{\partial V^*(\boldsymbol{x})}{\partial \boldsymbol{x}} = \boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{W}^* + \nabla\varepsilon(\boldsymbol{x}), \tag{5.10}$$

where $\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{x}) \triangleq \left[\frac{\partial \boldsymbol{\Phi}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right]^{\mathrm{T}}$ is the gradient of $\boldsymbol{\Phi}(\boldsymbol{x})$, and $\nabla\varepsilon(\boldsymbol{x}) \triangleq \frac{\partial \varepsilon(\boldsymbol{x})}{\partial \boldsymbol{x}}$ is the gradient of $\varepsilon(\boldsymbol{x})$.

The boundedness properties of $\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{x})$ and $\nabla\varepsilon(\boldsymbol{x})$ are given in Assumption 4.3 in Chapter 4.

Due to the approximation error $\varepsilon(\boldsymbol{x})$, the associated control law under the proposed approximation scheme is a near-optimal control as:

$$\boldsymbol{u}(\boldsymbol{x}) = -\frac{1}{2}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{x})\left[\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{W}^*\right]. \tag{5.11}$$

It is worth noting that Eq. (5.11) is an optimal control law when $\varepsilon(\boldsymbol{x}) = 0$ in ideal conditions where the provided activation functions are the ideal basis set. To avoid confusion and for ease of discussion, Eq. (5.11) is referred to as optimal control hereinafter.

This optimal control results in:

$$\varepsilon_{_H} = \left[\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{x})\right][\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})\boldsymbol{u}(\boldsymbol{x})] + \bar{Q}(\boldsymbol{x}) + \bar{R}(\boldsymbol{u}), \tag{5.12}$$

where $\varepsilon_{_H}$ is the residual caused by the approximation error $\varepsilon(\boldsymbol{x})$ in Eq. (5.9).

Recall that Eq. (5.9) contains a double-layer NN (i.e. $\boldsymbol{W}^{*\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{x})$), which is nonlinear in the hidden layer $\boldsymbol{\Phi}(\boldsymbol{x})$ but linear in the output layer weights $\boldsymbol{W}^*$. Let $\hat{\boldsymbol{W}}$ be the estimate of the ideal weights. To implement policy iteration, $\hat{\boldsymbol{W}}$ need to be

tuned dynamically so that $\hat{\boldsymbol{W}} \rightarrow \boldsymbol{W}^*$ and thus Eq. (5.9) approximates a target value function. In this case,

$$\hat{V}(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} + \hat{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{x}), \tag{5.13}$$

and

$$\hat{\boldsymbol{u}}(\boldsymbol{x}) = -\frac{1}{2}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{x})\left[\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}(\boldsymbol{x})\hat{\boldsymbol{W}}\right]. \tag{5.14}$$

The resulting nonlinear Lyapunov equation then becomes:

$$\left[\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} + \hat{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{x})\right][\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})\hat{\boldsymbol{u}}(\boldsymbol{x})] + \bar{Q}(\boldsymbol{x}) + \bar{R}(\hat{\boldsymbol{u}}) = \varepsilon_{_H} + e_{_H}, \tag{5.15}$$

where $e_{_H}$ is the error of weights estimation during a tuning process.

For fast convergence of $\hat{\boldsymbol{W}}$ to the ideal $\boldsymbol{W}^*$ so that $e_{_H}$ is minimised, an EKF is proposed in this paper. Since $\hat{\boldsymbol{W}}$ is the parameter vector to be estimated, Eq. (5.15) can be rearranged in the following form:

$$\begin{cases} \dot{\hat{\boldsymbol{W}}} = \boldsymbol{0} + \boldsymbol{w}, \\ y = \beta(\boldsymbol{x}, \hat{\boldsymbol{W}}) - \varepsilon_{_H} - e_{_H} + v, \end{cases} \tag{5.16}$$

with

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

and

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

where $\boldsymbol{0}$ is a null matrix (since $\hat{\boldsymbol{W}}$ is a constant vector), $\boldsymbol{w}$ and v are white-noise inputs with covariance $\boldsymbol{Q}_f \succ 0$ and $R_f > 0$, respectively.

Remark 5.1. In the system described by Eq. (5.16), $\hat{\boldsymbol{W}}$ are system states, and there are no drift dynamics for $\hat{\boldsymbol{W}}$. However, nonlinearities are present in the output dynamics, which are associated with the gradient of the NN activation functions as well as the time derivative of $\boldsymbol{x}$. Thus a nonlinear observer is needed to estimate the system states $\hat{\boldsymbol{W}}$.

Remark 5.2. White-noise inputs $\boldsymbol{w}$ and v in fact do not physically exist in the system of Eq. (5.16). Therefore, the corresponding covariance $\boldsymbol{Q}_f$ and R_f have no physical

implication. The expression of $\boldsymbol{w}$ and v is given in Eq. (5.16) purely in support of the use of $\boldsymbol{Q}_f$ and R_f for EKF implementation (Simon, 2002).

Note that y is known from measurements. The unknown ideal $\boldsymbol{W}^*$ is to be estimated according to y and the known dynamics $\beta(\boldsymbol{x}, \hat{\boldsymbol{W}})$. Introducing an EKF into the system described by Eq. (5.16) yields:

$$\begin{cases} \dot{\hat{\boldsymbol{W}}} = \kappa_c \boldsymbol{K}_f (y - \hat{y}), \\ \hat{y} = \beta(\boldsymbol{x}, \hat{\boldsymbol{W}}), \end{cases} \tag{5.17}$$

where $\hat{y}$ denotes the estimated output, $\boldsymbol{K}_f \in \mathbb{R}^{n_n \times 1}$ is the EKF gain, and $\kappa_c = \frac{\alpha_{cnn}}{\|\boldsymbol{K}_f\|+1}$ is a normalisation term with a scalar learning rate $\alpha_{cnn} \in \mathbb{R}^+$.

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

$$\boldsymbol{K}_f = \boldsymbol{S}\boldsymbol{H}^{\mathrm{T}} R_f^{-1}, \tag{5.18}$$

with

$$\boldsymbol{H}^{\mathrm{T}} = \frac{\partial \beta(\boldsymbol{x}, \hat{\boldsymbol{W}})}{\partial \hat{\boldsymbol{W}}} = \bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{x})[\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})\hat{\boldsymbol{u}}(\boldsymbol{x})], \tag{5.19}$$

and

$$\dot{\boldsymbol{S}} = \boldsymbol{Q}_f - \boldsymbol{S}\boldsymbol{H}^{\mathrm{T}} R_f^{-1} \boldsymbol{H}\boldsymbol{S}, \tag{5.20}$$

where $\boldsymbol{H} \in \mathbb{R}^{n_n \times 1}$ is defined as in Eq. (5.19), and $\boldsymbol{S} \in \mathbb{R}^{n_n \times n_n}$ is a symmetrical positive-definite matrix with initial state $\boldsymbol{S}(0) = \boldsymbol{S}(0)^{\mathrm{T}} \succeq 0$.

Similar to most adaptive control problems that require online tuning of parameters, persistence of excitation (PE) is needed for proper convergence of NN parameters (Ioannou and Sun, 1996). For online tuning, the PE condition as given in Assumption 4.4 is assumed to hold.

Remark 5.3. Despite the wide use of EKF in literature, it is one of the contributions in this paper that an EKF other than traditional gradient-based methods is employed to tune the weights of the NN involved in NOCOS with the MVFA for continuous-time nonlinear systems. Parameter convergence of the NN in the MVFA under the EKF tuning scheme is shown in the following theorem.

Theorem 5.1. *Under Assumptions 4.1 to 4.4 and the EFK estimation scheme provided by Eqs. (5.17) to (5.20), optimal or near-optimal control laws for the nonlinear system as in Eq. (5.3) are given by Eq. (5.14) in an online-learning manner, with the adaptive variable* $\hat{\boldsymbol{W}}$ *converging to the ideal value* $\boldsymbol{W}^*$ *within an error bound* $\|\tilde{\boldsymbol{W}}\| \leq b_{\tilde{W}}$.

Proof. Given the ideal NN parameter $\boldsymbol{W}^*$, the estimation error $\tilde{\boldsymbol{W}}$ is thus $\tilde{\boldsymbol{W}} = \boldsymbol{W}^* - \hat{\boldsymbol{W}}$. The Lyapunov function candidate $\mathcal{L}_W = \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}(\kappa_c \boldsymbol{S})^{-1}\tilde{\boldsymbol{W}}$ is considered.

According to Eqs. (5.16), (5.17), and (5.18),

$$\begin{aligned}\dot{\mathcal{L}}_W &= -\tilde{\boldsymbol{W}}^{\mathrm{T}}(\kappa_c \boldsymbol{S})^{-1}\dot{\hat{\boldsymbol{W}}} \\ &= -\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}R_f^{-1}(y - \hat{y}) \\ &= \tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}R_f^{-1}[\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}}) + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}}(\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}}) + \bar{Q} + \hat{\boldsymbol{u}}^{\mathrm{T}}\boldsymbol{R}\hat{\boldsymbol{u}}]. \end{aligned} \tag{5.21}$$

With Eq. (5.12), we have from Eq. (5.21):

$$\begin{aligned}\dot{\mathcal{L}}_W = \tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}R_f^{-1}\Big[&(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}})^{\mathrm{T}}(\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}}) \\ &- (\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*)^{\mathrm{T}}(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}) \\ &+ \hat{\boldsymbol{u}}^{\mathrm{T}}\boldsymbol{R}\hat{\boldsymbol{u}} - \boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u} + \varepsilon_{_H}\Big]. \end{aligned} \tag{5.22}$$

Substituting Eq. (5.10) for $\nabla \boldsymbol{V}^*$ in Eq. (5.8) with some manipulation yields:

$$\begin{aligned}&\left[\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\right][\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}] + \bar{Q} + \bar{R} \\ &= -\nabla\varepsilon^{\mathrm{T}}\boldsymbol{f} + \frac{1}{2}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} + \frac{1}{2}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* + \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon. \end{aligned} \tag{5.23}$$

Comparing Eqs. (5.12) and (5.23) thus gives:

$$\begin{aligned}\varepsilon_{_H} &= -\nabla\varepsilon^{\mathrm{T}}\boldsymbol{f} + \frac{1}{2}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} + \frac{1}{2}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* + \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon \\ &= -\nabla\varepsilon^{\mathrm{T}}(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) - \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon. \end{aligned} \tag{5.24}$$

With Eq. (5.24) substituted for $\varepsilon_{_H}$ in Eq. (5.22), and $\boldsymbol{G} \triangleq \boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}$, we have:

$$\begin{aligned}\dot{\mathcal{L}}_W = \tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}R_f^{-1}[&\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}(\hat{\boldsymbol{u}} - \boldsymbol{u}) - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} \\ &+ \hat{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\hat{\boldsymbol{u}} - \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{u} + \hat{\boldsymbol{u}}^{\mathrm{T}}\boldsymbol{R}\hat{\boldsymbol{u}} \\ &- \boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u} - \nabla\varepsilon^{\mathrm{T}}(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) - \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon] \end{aligned}$$

$$
\begin{aligned}
&= \tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}R_f^{-1}\left[\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f}\right.\\
&\qquad -\frac{1}{4}\hat{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}} - \nabla\varepsilon^{\mathrm{T}}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*)\\
&\qquad \left.+\frac{1}{4}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* - \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon\right]\\
&= -\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}R_f^{-1}\left[\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} - \frac{1}{4}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\right.\\
&\qquad +\frac{1}{4}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* - \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\\
&\qquad \left.-\nabla\varepsilon^{\mathrm{T}}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*) - \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon\right]\\
&= -\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}R_f^{-1}\left\{\frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\left[\boldsymbol{f} - \frac{1}{2}\boldsymbol{G}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}})\right]\right.\\
&\qquad -\frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\left[\boldsymbol{f} - \frac{1}{2}\boldsymbol{G}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*)\right]\\
&\qquad \left.-\nabla\varepsilon^{\mathrm{T}}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*) - \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon\right\}\\
&= -\tilde{\boldsymbol{W}}^{\mathrm{T}}\left[\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{f}+\boldsymbol{g}\hat{\boldsymbol{u}})R_f^{-1}\right]\left[\frac{1}{2}(\boldsymbol{f}+\boldsymbol{g}\hat{\boldsymbol{u}})^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\right.\\
&\qquad \left.+\frac{1}{2}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u})^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \nabla\varepsilon^{\mathrm{T}}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*) - \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon\right]. \qquad (5.25)
\end{aligned}
$$

Let $\tilde{\boldsymbol{u}} = \boldsymbol{u}^* - \hat{\boldsymbol{u}}$. Since $\boldsymbol{g}\boldsymbol{u} = \boldsymbol{g}\boldsymbol{u}^* + \frac{1}{2}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon$, we have:

$$
\begin{aligned}
\dot{\mathcal{L}}_W =& -\tilde{\boldsymbol{W}}^{\mathrm{T}}\left[\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{f}+\boldsymbol{g}\hat{\boldsymbol{u}})R_f^{-1}\right]\left[\frac{1}{2}(\boldsymbol{f}+\boldsymbol{g}\hat{\boldsymbol{u}})^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\right.\\
&\qquad +\frac{1}{2}\left(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*+\frac{1}{2}\boldsymbol{G}\nabla\varepsilon\right)^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\\
&\qquad \left.-\nabla\varepsilon^{\mathrm{T}}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*) - \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon\right]\\
=& -\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*)R_f^{-1}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*)^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\\
&-\frac{1}{4}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*)R_f^{-1}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\\
&-\frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*)R_f^{-1}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\\
&-\frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}R_f^{-1}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*)^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\\
&-\frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\nabla\varepsilon R_f^{-1}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*)^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\\
&-\frac{1}{8}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}R_f^{-1}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\\
&-\frac{1}{4}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}R_f^{-1}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}
\end{aligned}
$$

$$-\frac{1}{8}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\nabla\varepsilon R_f^{-1}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}$$
$$-\frac{1}{4}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\nabla\varepsilon R_f^{-1}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}. \tag{5.26}$$

Now we introduce bounds. Let

$$\boldsymbol{M}_1 = \bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*)R_f^{-1}(\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*)^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}},$$

$$\boldsymbol{M}_2 = \bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}R_f^{-1}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}},$$

and

$$\boldsymbol{M}_3 = \bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\nabla\varepsilon R_f^{-1}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}.$$

Because of $\operatorname{rank}(\boldsymbol{G}) = \operatorname{rank}(\boldsymbol{g}) < n_x$, there exist kernels

$$\ker(\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}) = \left\{\boldsymbol{r}\in\mathbb{R}^{n_n} \;\middle|\; \bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{r}=0\right\},$$
$$\ker(\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}) = \left\{\boldsymbol{r}\in\mathbb{R}^{n_n} \;\middle|\; \boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{r}=0\right\},$$
$$\ker(\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}) = \left\{\boldsymbol{r}\in\mathbb{R}^{n_x} \;\middle|\; \bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\boldsymbol{r}=0\right\}.$$

Since $\hat{\boldsymbol{W}}$ is explicitly governed by Eq. (5.17), we consider cases where

$$\tilde{\boldsymbol{W}} \neq \ker(\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}),$$
$$\tilde{\boldsymbol{W}} \neq \ker(\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}),$$
$$\nabla\varepsilon \neq \ker(\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}).$$

Let $\|R_f^{-1}\| = b_{Rf}$. Under Assumptions 4.1 to 4.4, and with Eq. (5.11), there exist constants $b_{\dot{x}}$, b_{GU}, b_{m1}, b_{m2}, $b_{m3} \in \mathbb{R}^+$ so that the following inequalities hold:

$$\|\boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}^*\| \leq b_{\dot{x}}\|\boldsymbol{x}\|, \tag{5.27}$$
$$\|\boldsymbol{G}\| \leq b_{GU}, \tag{5.28}$$
$$b_{m1}\Pi\|\tilde{\boldsymbol{W}}\|^2 \leq \left\|\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{M}_1\tilde{\boldsymbol{W}}\right\|, \tag{5.29}$$
$$b_{m2}\Pi\|\tilde{\boldsymbol{W}}\|^4 \leq \left\|\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{M}_2\tilde{\boldsymbol{W}}\right\|, \tag{5.30}$$
$$b_{m3}\Pi\|\tilde{\boldsymbol{W}}\|^2 \leq \left\|\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{M}_3\tilde{\boldsymbol{W}}\right\|, \tag{5.31}$$

where

$$\Pi = \|\boldsymbol{x}\|^4.$$

As a result, Eq. (5.26) can be upper bounded as:

$$\begin{aligned}\dot{\mathcal{L}}_W \leq & -\frac{1}{8}b_{m2}\Pi\left\|\tilde{\boldsymbol{W}}\right\|^4+\frac{3}{4}b_{\phi}^3 b_{\dot{x}}b_{GU}b_{Rf}\Pi\left\|\tilde{\boldsymbol{W}}\right\|^3 \\ & +\frac{3}{4}b_{\phi}^3 b_{GU}^2 b_{\varepsilon}b_{Rf}\Pi\left\|\tilde{\boldsymbol{W}}\right\|^3+b_{\phi}^2 b_{\dot{x}}b_{GU}b_{\varepsilon}b_{Rf}\Pi\left\|\tilde{\boldsymbol{W}}\right\|^2 \\ & -\frac{1}{4}b_{m3}\Pi\left\|\tilde{\boldsymbol{W}}\right\|^2-b_{m1}\Pi\left\|\tilde{\boldsymbol{W}}\right\|^2 \\ = & -\frac{1}{8}b_{m2}\Pi\left\|\tilde{\boldsymbol{W}}\right\|^2\left[\left(\left\|\tilde{\boldsymbol{W}}\right\|-\frac{4\eta_1}{b_{m2}}\right)^2-\frac{16\eta_1^2+8\eta_2 b_{m2}}{b_{m2}^2}\right], \end{aligned} \tag{5.32}$$

where

$$\eta_1=\frac{3}{4}b_{\phi}^3\left(b_{\dot{x}}b_{GU}+b_{GU}^2 b_{\varepsilon}\right)b_{Rf},$$

and

$$\eta_2=b_{\phi}^2 b_{\dot{x}}b_{GU}b_{\varepsilon}b_{Rf}-\frac{1}{4}b_{m3}-b_{m1}.$$

According to Eqs. (5.29) to (5.31),

$$\begin{aligned}\eta_2 &\geq b_{\phi}^2 b_{\dot{x}}b_{GU}b_{\varepsilon}b_{Rf}-\frac{1}{4}b_{\phi}^2 b_{GU}^2 b_{\varepsilon}^2 b_{Rf}-b_{\phi}^2 b_{\tilde{x}}^2 b_{Rf} \\ &=-b_{\phi}^2 b_{Rf}\left(b_{\dot{x}}-\frac{1}{2}b_{GU}b_{\varepsilon}\right)^2. \end{aligned} \tag{5.33}$$

If $-b_{\phi}^2 b_{Rf}\left(b_{\dot{x}}-\frac{1}{2}b_{GU}b_{\varepsilon}\right)^2\leq\eta_2<0$, then:

$$\begin{aligned}& 16\eta_1^2+8\eta_2 b_{m2} \\ =&\frac{1}{4}\left(64\eta_1^2+32\eta_2 b_{m2}\right) \\ \geq&\frac{1}{4}\left[64\left(\frac{3}{4}b_{\phi}^3\left(b_{\dot{x}}b_{GU}+b_{GU}^2 b_{\varepsilon}\right)b_{Rf}\right)^2-32b_{\phi}^6 b_{GU}^2 b_{Rf}^2\left(b_{\dot{x}}-\frac{1}{2}b_{GU}b_{\varepsilon}\right)^2\right] \\ =&\frac{1}{4}b_{\phi}^6 b_{GU}^2 b_{Rf}^2\left(4b_{\tilde{x}}^2+68b_{\dot{x}}b_{GU}b_{\varepsilon}+b_{GU}^2 b_{\varepsilon}^2\right)>0. \end{aligned}$$

If $\eta_2\geq 0$, then:

$$16\eta_1^2+8\eta_2 b_{m2}\leq 16\eta_1^2+8\eta_2 b_{\phi}^4 b_{GU}^2 b_{Rf}.$$

As a result, the inequality in Eq. (5.33) gives:

$$0<16\eta_1^2+8\eta_2 b_{m2}\leq 16\eta_1^2+8\eta_2 b_{\phi}^4 b_{GU}^2 b_{Rf}.$$

Under Assumption 4.4, Π is PE. As a result, Eq. (5.32) shows that $\dot{\mathcal{L}}_W$ is negative definite if

$$\|\tilde{\boldsymbol{W}}\| > \frac{4\eta_1 + \sqrt{16\eta_1^2 + 8\eta_2 b_{m2}}}{b_{m2}} \triangleq b_{\tilde{W}}.$$

That is, $\left\|\boldsymbol{W}^* - \hat{\boldsymbol{W}}\right\|$ is uniformly ultimately bounded (UUB) within $b_{\tilde{W}}$, offering an optimal or near-optimal control solution as given by Eq. (5.14). This completes the proof. □

Remark 5.4. It is shown that the convergence of $\|\tilde{\boldsymbol{W}}\|$ is UUB under the MVFA and the EKF tuning scheme. More importantly, any interim value of $\hat{\boldsymbol{W}}$ along its evolution path can be directly used by the actor NN (as in Eq. (5.14)) to provide stabilising control in the meantime. The closed-loop stability of the overall system under the new approach is discussed next.

According to Assumption 4.1 and Eq. (5.15), the system (5.3) has only one equilibrium where $\boldsymbol{x} = 0$ and $\boldsymbol{u} = 0$, whereas other equilibrium points require sustained control inputs to maintain. Therefore, both the system states and control inputs are to be evaluated in terms of the Lyapunov stability, where the control inputs are also required to decay to zero with time. A suitable Lyapunov function candidate can be the approximated value function $\hat{V}^{\boldsymbol{u}_{(i)}}(\boldsymbol{x})$ of the system associated with an admissible control $\boldsymbol{u}_{(i)}$ at the i^{th} infinitesimal time step. This is because in a stabilised closed-loop system the value function $V^{\boldsymbol{u}_{(i)}}(\boldsymbol{x})$ is positive from definition as in Eq. (5.5) and its time derivative should be negative in accordance with Eq. (5.6). On the other hand, if some intermediate values of $\hat{\boldsymbol{W}}$ during convergence to $\boldsymbol{W}^*$ can lead to system instability, the resulted value function time derivative $d\hat{V}/dt$ will turn positive. Therefore, if $d\hat{V}/dt$ is shown to be always negative regardless how $\hat{\boldsymbol{W}}$ converges to $\boldsymbol{W}^*$, then the closed-loop system is said to be stable under admissible control during online tuning.

Theorem 5.2. *Given Assumptions 4.1 to 4.3 and the EKF estimation scheme provided by Eqs. (5.17) to (5.20), there exists a scalar $b_{mP} \in \mathbb{R}^+$ and a matrix $\boldsymbol{P}$ for Eq. (5.9) such that*

$\|\boldsymbol{P}\| > b_{mP}$ and the nonlinear system as in Eq. (5.4) remains asymptotically stable during online tuning with the control law given by Eq. (5.14).

Proof. The Lyapunov function candidate is selected to be $\mathcal{L}_V = \hat{V}$ as discussed. Its time derivative is:

$$
\begin{aligned}
\dot{\mathcal{L}}_V =& \left(\frac{\partial \hat{V}}{\partial \boldsymbol{x}}\right)^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}}) \\
=& (\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}})^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}}) \\
=& (\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* - \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}})^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^* - \boldsymbol{g}\tilde{\boldsymbol{u}}) \\
=& \left(\frac{\partial V^*}{\partial \boldsymbol{x}}\right)^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) - \nabla\varepsilon^{\mathrm{T}} (\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) \\
&+ \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}(\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} + \nabla\varepsilon) \\
&+ \frac{1}{2}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}(\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} + \nabla\varepsilon) - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} \\
&+ \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* + \nabla\varepsilon) \\
&- \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}(\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} + \nabla\varepsilon) \\
=& -\bar{Q} - \boldsymbol{u}^{*\mathrm{T}}\boldsymbol{R}\boldsymbol{u}^* - \nabla\varepsilon^{\mathrm{T}}\boldsymbol{f} - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} + \nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x} \\
&+ \frac{1}{2}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon + \nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* + \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x} \\
&+ \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* - \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} \\
=& -\bar{Q} - \frac{1}{4}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x} - \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* - \nabla\varepsilon^{\mathrm{T}}\boldsymbol{f} \\
&+ \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon - \frac{1}{4}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* + \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{G}\nabla\varepsilon \\
&+ \frac{1}{2}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\nabla\varepsilon - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} + \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x} \\
&+ \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* - \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}.
\end{aligned}
\tag{5.34}
$$

As $\bar{Q}(\boldsymbol{x}) > 0$, we have $b_Q\|\boldsymbol{x}\|^2 \leq \bar{Q}(x)$ for constants $b_Q \in \mathbb{R}^+$. Given that $\boldsymbol{x}$ is explicitly governed by system (5.2), the case of $\boldsymbol{x} \neq \ker(\boldsymbol{G})$ is considered. It is straightforward to see that there exist constants $b_{GL} \in \mathbb{R}^+$ and $b_{m4} \in \mathbb{R}^+$ such that

$$
b_{GL} \leq \|\boldsymbol{G}\|,
$$

and

$$b_{m4}\|\boldsymbol{x}\|^2 \leq \left\| \boldsymbol{W}^{*\mathrm{T}} \bar{\nabla} \boldsymbol{\Phi} \boldsymbol{G} \bar{\nabla} \boldsymbol{\Phi}^{\mathrm{T}} \boldsymbol{W}^* \right\|.$$

Following the results of Theorem 5.1, it is known that $0 \leq \|\tilde{\boldsymbol{W}}\| \leq b_{\tilde{W}}$. Then Eq. (5.34) can be upper bounded as:

$$\begin{aligned}
\dot{\mathcal{L}}_V \leq & -b_Q\|\boldsymbol{x}\|^2 - \frac{1}{4}b_{GL}\|\boldsymbol{P}\|^2\|\boldsymbol{x}\|^2 + \frac{1}{2}b_{GU}b_\varepsilon\|\boldsymbol{P}\|\,\|\boldsymbol{x}\|^2 \\
& + \frac{1}{2}b_{GU}b_\phi\|\boldsymbol{W}^*\|\,\|\boldsymbol{P}\|\,\|\boldsymbol{x}\|^2 - \frac{1}{4}b_{m4}\|\boldsymbol{x}\|^2 + b_\varepsilon b_f\|\boldsymbol{x}\|^2 \\
& + \frac{1}{2}b_\phi b_{GU}b_\varepsilon\|\boldsymbol{W}^*\|\|\boldsymbol{x}\|^2 + \frac{1}{4}b_\varepsilon^2 b_{GU}\|\boldsymbol{x}\|^2 + b_\phi b_f b_{\tilde{W}}\|\boldsymbol{x}\|^2 \\
& + b_\phi b_{GU}b_{\tilde{W}}\|\boldsymbol{P}\|\,\|\boldsymbol{x}\|^2 + b_\phi^2 b_{GU}b_{\tilde{W}}\|\boldsymbol{W}^*\|\|\boldsymbol{x}\|^2 \\
= & -\|\boldsymbol{x}\|^2\left(\frac{1}{4}b_{GL}\|\boldsymbol{P}\|^2 - \eta_3\|\boldsymbol{P}\| - \eta_4\right).
\end{aligned} \tag{5.35}$$

where

$$\eta_3 = \frac{1}{2}b_{GU}b_\phi\|\boldsymbol{W}^*\| + \frac{1}{2}b_{GU}b_\varepsilon + b_\phi b_{GU}b_{\tilde{W}},$$

and

$$\begin{aligned}
\eta_4 = & \frac{1}{2}b_\phi b_{GU}b_\varepsilon\|\boldsymbol{W}^*\| + \frac{1}{4}b_\varepsilon^2 b_{GU} + b_\varepsilon b_f + b_\phi b_f b_{\tilde{W}} \\
& + b_\phi^2 b_{GU}b_{\tilde{W}}\|\boldsymbol{W}^*\| - b_Q - \frac{1}{4}b_{m4}.
\end{aligned} \tag{5.36}$$

Equation (5.35) shows that $\dot{\mathcal{L}}_V$ is negative, and thus $\|\boldsymbol{x}\|$ is bounded, as long as

$$\|\boldsymbol{P}\| \geq \frac{2\eta_3 + 2\sqrt{\eta_3^2 + b_{GL}\eta_4}}{b_{GL}} \triangleq b_{mP}. \tag{5.37}$$

It can be easily derived that the second-order derivative of $\mathcal{L}_V$ with respect to time is a function of $\boldsymbol{x}$ and $\tilde{\boldsymbol{W}}$. Since $\|\boldsymbol{x}\|$ and $\|\tilde{\boldsymbol{W}}\|$ are both bounded, $\ddot{\mathcal{L}}_V$ is also bounded. Therefore, it can be concluded that the system states $\boldsymbol{x}$ are asymptotically stable. This completes the proof. □

Remark 5.5. This new control scheme under EKF tuning maintains the stability of the closed-loop system during online adaptation without the necessity of providing an initial stabilising control, adding a stabilising logic-switch mechanism to the critic NN, or adding an additional stabilising tuning loop to the actor NN. Moreover, the

proposed algorithm provides proven asymptotic stability rather than the relatively weaker UUB stability to nonlinear systems as in Eq. (5.3) during online tuning. Theorems 5.1 and 5.2 verify the EKF tuning for the MVFA and build an important contribution of this paper.

5.3.3 Generalisation of modified value-function-approximation

Note that the discussion in Subsections 5.3.1 and 5.3.2 are limited to locally nonlinear systems, with the parameter matrix $\boldsymbol{P}$ being constant, which as a result, does not suit a wider flight envelop with varying travelling 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 $\boldsymbol{P}$ to cope with U_∞ dependent dynamics as in Eq. (5.2), generalising the new NN-based VFA to globally nonlinear cases.

Linearising Eq. (5.2) about $\boldsymbol{x} = 0$ gives:

$$\dot{\boldsymbol{x}} = \boldsymbol{A}_p(U_\infty)\boldsymbol{x} + \boldsymbol{B}_p\boldsymbol{u} + \boldsymbol{w}_p, \tag{5.38}$$

where

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

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

With performance output $\boldsymbol{h}$ considered, there is:

$$\begin{bmatrix} \dot{\boldsymbol{x}} \\ \boldsymbol{h} \\ \boldsymbol{y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_p(U_\infty) & \boldsymbol{I} & \boldsymbol{B}_p \\ \boldsymbol{C}_h & 0 & \boldsymbol{E}_h \\ \boldsymbol{C}_p & 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{w} \\ \boldsymbol{u} \end{bmatrix}, \tag{5.39}$$

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

Let $\boldsymbol{P}(U_\infty)$ be a scheduled matrix which varies with the freestream airspeed U_∞. According to Theorem 5.2, a stable closed-loop system under the dynamically tuned control law as in Eq. (5.14) requires $\|\boldsymbol{P}\| > b_{mP}$ where the value of the scalar $b_{mP} \in \mathbb{R}^+$ depends on the system dynamics. In the case of AFS, b_{mP} is not constant

but varies with U_∞. That is, $b_{mP} = b^*_{mP}(U_\infty)|_{U_\infty=U}$ for any valid airspeed U, where $b^*_{mP}(U_\infty)$ is a generalised function. To find $\boldsymbol{P}(U_\infty)$ that satisfies the condition of $\|\boldsymbol{P}(U_\infty)\| > b^*_{mP}(U_\infty)$, a Lyapunov matrix $\boldsymbol{X}(U_\infty) = \boldsymbol{X}^{\mathrm{T}}(U_\infty) \succ 0$ and an auxiliary parameter-dependent performance variable $\boldsymbol{Z}(U_\infty)$ are introduced to form the following linear matrix inequalities (LMIs):

$$\begin{bmatrix} \dot{\boldsymbol{X}} + \boldsymbol{A}_c^{\mathrm{T}}\boldsymbol{X} + \boldsymbol{X}\boldsymbol{A}_c & \boldsymbol{X} \\ \boldsymbol{X} & -\nu\boldsymbol{I} \end{bmatrix} \prec 0, \tag{5.40}$$

$$\begin{bmatrix} \boldsymbol{X} & \boldsymbol{C}_{ch}^{\mathrm{T}} \\ \boldsymbol{C}_{ch} & \boldsymbol{Z} \end{bmatrix} \succ 0, \tag{5.41}$$

and

$$\mathrm{Tr}(\boldsymbol{Z}) < \nu, \tag{5.42}$$

where

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

$$\boldsymbol{C}_{ch} \triangleq \boldsymbol{C}_h - \frac{1}{2}\boldsymbol{E}_h\boldsymbol{R}^{-1}\boldsymbol{B}_p^{\mathrm{T}}\boldsymbol{P}(U_\infty),$$

and ν is a performance index.

Let $\bar{\boldsymbol{B}}_c \triangleq \frac{1}{2}\boldsymbol{B}_p\boldsymbol{R}^{-1}\boldsymbol{B}_p^{\mathrm{T}}$, $\boldsymbol{Y}(U_\infty) = \boldsymbol{P}(U_\infty)\boldsymbol{X}^{-1}(U_\infty)$, and $\boldsymbol{J}(U_\infty) = \boldsymbol{X}^{-1}(U_\infty)$. Then Eqs. (5.40) and (5.41) can be transformed into:

$$-\dot{\boldsymbol{J}} + \boldsymbol{A}_p\boldsymbol{J} + \boldsymbol{J}\boldsymbol{A}_p^{\mathrm{T}} + \bar{\boldsymbol{B}}_c\boldsymbol{Y} + \boldsymbol{Y}^{\mathrm{T}}\bar{\boldsymbol{B}}_c^{\mathrm{T}} \prec 0, \tag{5.43}$$

$$\begin{bmatrix} \boldsymbol{J} & (\boldsymbol{C}_h\boldsymbol{J} + \boldsymbol{E}_h\boldsymbol{Y})^{\mathrm{T}} \\ \boldsymbol{C}_h\boldsymbol{J} + \boldsymbol{E}_h\boldsymbol{Y} & \boldsymbol{Z} \end{bmatrix} \succ 0. \tag{5.44}$$

In light of Eq. (5.1), $\boldsymbol{A}(U_\infty)$ can be structured as:

$$\boldsymbol{A}_p(U_\infty) = \boldsymbol{A}_{p1} + \boldsymbol{A}_{p2}U_\infty + \boldsymbol{A}_{p3}U_\infty^2. \tag{5.45}$$

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

$$\boldsymbol{J}(U_\infty) = \boldsymbol{J}_1 + \boldsymbol{J}_2U_\infty + \boldsymbol{J}_3U_\infty^2, \tag{5.46}$$

and

$$\boldsymbol{Y}(U_\infty) = \boldsymbol{Y}_1 + \boldsymbol{Y}_2 U_\infty + \boldsymbol{Y}_3 U_\infty^2. \tag{5.47}$$

Solving for $\boldsymbol{J}(U_\infty)$ and $\boldsymbol{Y}(U_\infty)$ through Eqs. (5.43), (5.44), and (5.42) gives $\|\boldsymbol{P}(U_\infty)\| > b^*_{mP}(U_\infty)$ in the form of:

$$\boldsymbol{P}(U_\infty) = \boldsymbol{Y}(U_\infty)\boldsymbol{J}^{-1}(U_\infty). \tag{5.48}$$

with $b^*_{mP}(U_\infty)$ implicitly embedded in the LMIs derived.

Remark 5.6. The proposed procedure for designing the parameter matrix $\boldsymbol{P}$ yields $\|\boldsymbol{P}(U_\infty)\| > b^*_{mP}(U_\infty)$ across U_∞ of interest, satisfying the condition of $\|\boldsymbol{P}\| > b_{mP}$ in Theorem 5.2, and thus generalises Theorem 5.2 to a more general scenario, the globally nonlinear case, with nonlinear dynamics described in Eq. (5.2). This forms the second contribution of this paper.

5.3.4 Online system identification

Note that the information of $\boldsymbol{f}(\boldsymbol{x}, U_\infty)$ and $\boldsymbol{g}(\boldsymbol{x})$ is required for real-time synthesis of nonlinear optimal control laws. Although the knowledge of $\boldsymbol{f}(\boldsymbol{x}, U_\infty)$ and $\boldsymbol{g}(\boldsymbol{x})$ is analytically available, the presence of un-modelled dynamics or uncertainties can degrade controller performance as discussed in Section 5.1. To mitigate this problem, an NN-based identifier is proposed in the following form:

$$\dot{\boldsymbol{x}} = \boldsymbol{W}_s^{*\mathrm{T}} \boldsymbol{\Phi}_s(\boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{\varepsilon}_s, \tag{5.49}$$

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

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

$$\boldsymbol{x} = \boldsymbol{W}_s^{*\mathrm{T}} \boldsymbol{\mu}_1(\boldsymbol{x}) + \boldsymbol{\Gamma} \boldsymbol{\mu}_2(\boldsymbol{x}) + \boldsymbol{\varepsilon}_x, \tag{5.50}$$

with

$$\dot{\boldsymbol{\mu}}_1(\boldsymbol{x}) = -\boldsymbol{\Gamma} \boldsymbol{\mu}_1(\boldsymbol{x}) + \boldsymbol{\Phi}_s(\boldsymbol{x}, \boldsymbol{u}), \quad \boldsymbol{\mu}_1(\boldsymbol{x}_0) = 0, \tag{5.51}$$

and

$$\dot{\boldsymbol{\mu}}_2(\boldsymbol{x}) = -\boldsymbol{\Gamma}\boldsymbol{\mu}_2(\boldsymbol{x}) + \boldsymbol{x}, \quad \boldsymbol{\mu}_2(\boldsymbol{x}_0) = 0, \tag{5.52}$$

where $\boldsymbol{\mu}_1(\boldsymbol{x}) \in \mathbb{R}^{n_{ws}}$ and $\boldsymbol{\mu}_2(\boldsymbol{x}) \in \mathbb{R}^{n_x}$ are auxiliary regressors, $\boldsymbol{\Gamma} = \gamma \boldsymbol{I}_{n_x \times n_x}$ with $\gamma \in \mathbb{R}^+$, and

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

Denote the estimate of $\boldsymbol{x}$ by $\hat{\boldsymbol{x}}$. For fast estimation of $\hat{\boldsymbol{W}}_s$ towards $\boldsymbol{W}_s^*$, the EKF is considered for online tuning. In this study, multiple EKFs in a parallel configuration instead of a single EKF are employed, based on the fact that the columns of NN weights in the $\boldsymbol{W}_s^*$ matrix are exclusively associated with respective single state in the $\boldsymbol{x}$ vector and thus independent from each other (i.e. uncoupled). By doing so, the computational expense is significantly less than standard implementation using only one EKF, according to Simon (2002). On this basis, we have:

$$\begin{cases} \dot{\hat{\boldsymbol{W}}}_{s(i)} = \kappa_s \boldsymbol{K}_{s(i)}(\boldsymbol{x}_{(i)} - \hat{\boldsymbol{x}}_{(i)}), \\ \hat{\boldsymbol{x}}_{(i)} = \hat{\boldsymbol{W}}_{s(i)}\boldsymbol{\mu}_1(\boldsymbol{x}) + \boldsymbol{\Gamma}\boldsymbol{\mu}_2(\boldsymbol{x}), \end{cases} \tag{5.53}$$

where $\boldsymbol{K}_s \in \mathbb{R}^{n_w s}$ is the EKF gain and the subscript (i) restricts the parameters to the i^{th} decoupled EKF; $\kappa_s = \frac{\alpha_{snn}}{\left\|\boldsymbol{K}_{s(i)}\right\|+1}$ is a normalisation term with a constant learning rate $\alpha_{snn} \in \mathbb{R}^+$.

Each EKF gain vector $\boldsymbol{K}_{s(i)}$ can be computed following Eqs. (5.18) to (5.20), with some variation to Eq. (5.19). That is:

$$\boldsymbol{K}_{s(i)} = \boldsymbol{S}_{s(i)}\boldsymbol{H}_{s(i)}^{\mathrm{T}} R_s^{-1}, \tag{5.54}$$

$$\boldsymbol{H}_{s(i)}^{\mathrm{T}} = \frac{\partial \hat{\boldsymbol{x}}_{(i)}}{\partial \hat{\boldsymbol{W}}_{s(i)}} = \boldsymbol{\mu}_1(\boldsymbol{x}), \tag{5.55}$$

$$\dot{\boldsymbol{S}}_{s(i)} = \boldsymbol{Q}_s - \boldsymbol{S}_{s(i)}{\boldsymbol{H}_{s(i)}}^{\mathrm{T}} R_s^{-1}\boldsymbol{H}_{s(i)}\boldsymbol{S}_{s(i)}. \tag{5.56}$$

where $\boldsymbol{Q}_s \succ 0$ and $R_s > 0$ are defined the same as $\boldsymbol{Q}_f$ and $\boldsymbol{R}_f$.

As can be seen from Eq. (5.55), $\boldsymbol{H}_{s(i)}$ is the same among all individual EKF in the parallel configuration. It is also to be noted that $\boldsymbol{H}_{s(i)}$ is not constant but state dependent, indicating the nonlinearities involved in system identification, which justifies the use of an EKF instead of linear observers.

At this stage, the internal dynamics $f(x, U_\infty)$ can be given by the identifier NN in an indirect manner, since the estimated derivative of system states used by Eq. (5.17) can be obtained through Eq. (5.49), with W_s^* and u replaced by $\hat{W}_s$ and $\hat{u}$, respectively.

The input dynamics $g(x)$ can then be obtained as:

$$\hat{g}(x) = \frac{\partial \hat{\mathcal{F}}_A(x, \hat{u})}{\partial \hat{u}} = \frac{\partial \left[\hat{W}_s^{\mathrm{T}} \Phi_s(x, \hat{u})\right]}{\partial \hat{u}}. \tag{5.57}$$

That is,

$$\hat{g}(x) = \frac{\partial \Phi_s(x, \hat{u})}{\partial \hat{u}} \hat{W}_s. \tag{5.58}$$

Convergence of the identifier NN weights under the state filtering scheme and the decoupled EKF is given in the following theorem.

Theorem 5.3. *Under Assumption 4.4 and the EKF estimation scheme provided by Eqs. (5.53) to (5.56), the nonlinear system as in Eq. (5.3) can be reconstructed by Eqs. (5.49) to (5.52), with* $\|\tilde{W}_s\| = \|W_{s(i)}^* - \hat{W}_{s(i)}\| \leq b_{\tilde{W}s}$, *where* $b_{\tilde{W}s} \in \mathbb{R}^+$.

Proof. The following Lyapunov candidate is considered:

$$\mathcal{L}_s = \frac{1}{2} \tilde{W}_{s(i)}^{\mathrm{T}} (\kappa_s S_{s(i)})^{-1} \tilde{W}_{s(i)}. \tag{5.59}$$

With an EKF as in Eq. (5.53), the time derivative of Eq. (5.59) is:

$$\begin{aligned} \dot{\mathcal{L}}_s &= -\tilde{W}_{s(i)}^{\mathrm{T}} (\kappa_s S_{s(i)})^{-1} \dot{\hat{W}}_{s(i)} \\ &= -\tilde{W}_{s(i)}^{\mathrm{T}} (S_{s(i)})^{-1} K_{s(i)} (x_{(i)} - \hat{x}_{(i)}). \end{aligned} \tag{5.60}$$

Using Eqs. (5.50), (5.53), and (5.54), we have:

$$\dot{\mathcal{L}}_s = -\tilde{W}_{s(i)}^{\mathrm{T}} H_{s(i)}^{\mathrm{T}} R_s^{-1} \left(\tilde{W}_{s(i)}^{\mathrm{T}} \mu_1 + \varepsilon_{s(i)} \right).$$

Substituting Eq. (5.55) for $H_{s(i)}^{\mathrm{T}}$ yields:

$$\dot{\mathcal{L}}_s = -\tilde{W}_{s(i)}^{\mathrm{T}} \mu_1 R_s^{-1} \mu_1{}^{\mathrm{T}} \tilde{W}_{s(i)} - \tilde{W}_{s(i)}^{\mathrm{T}} \mu_1 R_s^{-1} \varepsilon_{s(i)}.$$

Under Assumption 4.4, μ_1 is PE. With Eq. (5.51), there exist constants $b_{m\mu 1} \in \mathbb{R}^+$ so that

$$b_{m\mu 1} \|x\|^2 \left\| \tilde{W}_{s(i)} \right\|^2 \leq \left\| \tilde{W}_{s(i)}^{\mathrm{T}} M_{s(i)} \tilde{W}_{s(i)} \right\|,$$

where $\boldsymbol{M}_{s(i)} = \boldsymbol{\mu}_1 R_s^{-1} \boldsymbol{\mu}_1^{\mathrm{T}}$.

Therefore, with $\left\|\boldsymbol{R}_s^{-1}\right\| = b_{Rs}$, we have

$$\begin{aligned}\dot{\mathcal{L}}_s &\leq -b_{m\mu 1}\|\boldsymbol{x}\|^2\left\|\tilde{\boldsymbol{W}}_{s(i)}\right\|^2 + b_\varepsilon b_\eta^2 b_{Rs}\|\boldsymbol{x}\|^2\left\|\tilde{\boldsymbol{W}}_{s(i)}\right\| \\ &= -\|\boldsymbol{x}\|^2\left(b_{m\mu 1}\left\|\tilde{\boldsymbol{W}}_{s(i)}\right\|^2 - b_\varepsilon b_\eta^2 b_{Rs}\left\|\tilde{\boldsymbol{W}}_{s(i)}\right\|\right). \end{aligned} \tag{5.61}$$

Under Assumption 4.4, $\|\boldsymbol{x}\|$ is PE. As a result, $\dot{\mathcal{L}}_s$ is negative definite if

$$\left\|\tilde{\boldsymbol{W}}_{s(i)}\right\| > \frac{b_\varepsilon b_\eta^2 b_{Rs}}{b_{m\mu 1}},$$

which shows $\tilde{\boldsymbol{W}}_s$ is UUB within bound $b_{\tilde{W}s} \triangleq \frac{b_\varepsilon b_\eta^2 b_{Rs}}{b_{m\mu 1}}$. This completes the proof. □

Remark 5.7. As stated at the beginning of this subsection, the use of the identifier NN is mainly for updating the knowledge of system dynamics accessed by the real-time optimal-control synthesis scheme. The known dynamics (i.e. known analytical model) are embedded into the identifier NN in the form of the initial values of $\boldsymbol{W}_s^*$ obtained via pre-training the NN using known dynamics. Therefore, any mismatch between the actual dynamics and the analytical model can be captured and used to update $\boldsymbol{W}_s^*$. Specifically for the case of AFS, 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 given the learning ability of the identifier NN. When the airspeed increases, un-modelled and mismatching dynamics can be captured, with the pre-trained NN updated in real time accordingly.

5.4 Wind-tunnel experiments

Experiments were performed in a temperature regulated closed-loop wind tunnel at the University of Adelaide, Australia, and the setup is shown in Figure 3.13. The wind tunnel has a 0.5×0.5 m testing duct, and can generate up to 30 m/s smooth airflow with 0.5% turbulence intensity. The leading- and trailing-edge control surfaces of the aerofoil section are each driven by a servo motor, with the

Table 5.1: Parameters of the experimental aeroelastic system

Parameters	Values	Parameters	Values
ρ	1.225 kg/m^3	c_h	14.0 kg/s
m_a	0.851 kg	c_a	0.042 kg·m^2/s
m_{te}	0.030 kg	c_{te}	4.231×10^{-4} kg·m^2/s
m_{le}	0.058 kg	c_{le}	4.327×10^{-4} kg·m^2/s
r_{fc}	-0.0685 m	k_h	$50 + 300h^2$ N/m
$r_{3c/4}$	0.081 m	k_a	$0.3 + 30q_a^2$ Nm/rad
r_a	0.033 m	k_{te}	4.570×10^{-3} Nm/rad
r_{te}	1.019×10^{-2} m	k_{le}	4.704×10^{-3} Nm/rad
r_{le}	4.401×10^{-3} m	I_a	2.431×10^{-3} kg·m^2
$C_{l\text{-}a}$	6.573	I_{te}	2.307×10^{-6} kg·m^2
$C_{l\text{-}te}$	3.472	I_{le}	4.791×10^{-6} kg·m^2
$C_{l\text{-}le}$	-0.145	L_{te}	0.088 m
$C_{m\text{-}a}$	0	L_{le}	-0.010 m
$C_{m\text{-}te}$	-0.631	L_s	0.260 m
$C_{m\text{-}le}$	0.098	L_{hc}	0.075 m

corresponding deployment angle fed back via an optical encoder. Instead of using physical springs for plunge and pitch stiffness, a virtual stiffness-damping system (VSDS) was developed in this study, where two electric motors were used to mimic the compound structural forces acting on the aerofoil using force/torque feedback from a 6-axis force/torque transducer. The VSDS allows custom setting of the structural stiffness and damping within the output capacity of the motors used. This enables convenient adjustment of the structural stiffness and damping so that the flutter phenomenon of the aerofoil section can be observed around an desired airspeed. The parameters of the overall aeroelastic system used in experiments are listed in Table 5.1. With the stiffness and damping setting of the VSDS in Table 5.1, the system had a flutter boundary around 14.6m/s. Flutter was initiated by means of giving the VSDS a pulse signal along the plunge DOF.

In consideration of the 3^{rd} order nonlinear plunge and pitch stiffness in polynomial form, a power series of activation functions containing the powers of 8 system

Table 5.2: Other parameters of the proposed NN controller.

Parameters	Values
$\boldsymbol{Q}$	diag (1, 1, 10^{-4}, 10^{-4}, 0.1, 0.1, 10^{-4}, 10^{-4})
$\boldsymbol{R}$	$100\boldsymbol{I}$
$\boldsymbol{Q}_f$	$1000\boldsymbol{I}$
$\boldsymbol{R}_f$	$\boldsymbol{I}$
$\boldsymbol{Q}_{s(i)}$	$1 \times 10^5 \boldsymbol{I}$
$R_{s(i)}$	1

states up to 4^{th} order and 2 control inputs limited to 1^{st} order were used for the identifier NN in accordance with the high-order Weierstrass approximation theorem (Finlayson, 1972). This renders 135 significant activation functions for $\boldsymbol{\Phi}_s(\boldsymbol{x})$. Initial weights were determined via simulation-based training for 14.6 m/s airspeed. Similarly, $\boldsymbol{\Phi}(\boldsymbol{x})$ of the critic NN contains the powers of 8 system states only (no control inputs) up to 4^{th} order. This gives 65 significant activation functions. Weights $\hat{\boldsymbol{W}}$ were initialised to zeros. $\boldsymbol{J}(U_\infty)$ and $\boldsymbol{Y}(U_\infty)$ were designed using the parameters in Table 5.1 for the airspeed range from 14.6 m/s to 20 m/s with a gridding of 50 evenly spaced points. $\boldsymbol{P}(U_\infty)$ was calculated in real time using Eq. (5.48). $\bar{Q}(\boldsymbol{x})$ in Eq. (5.5) was structured as $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}$, with $\boldsymbol{Q}$ and other parameters listed in Table 5.2.

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 $\boldsymbol{x}(t_c)$ 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 5.1 and throughout the paper, there is no existing policy-iteration algorithms suitable for AFS without modification and improvement. Therefore, no suitable NN-based optimal controller counterparts can be compared in experiments. In order to evaluate the AFS performance improvement gained by using the proposed controller, a linear-parameter-varying (LPV) controller in the form of linear-quadratic-regulator (LQR) synthesised by means of LMIs (Prime,

2010) was reconstructed for the 4-DOF model as in Eq. (5.1) with the parameters in Table 5.1 and the weighting $\boldsymbol{Q}$ and $\boldsymbol{R}$ same as those used by the proposed NN controller.

Plunge and pitch responses as well as control surfaces deflections of the aerofoil section in the wind-tunnel tests under the proposed NN controller and the LPV-LQR controller at different airspeeds are plotted in Figure 5.2 for 14.8 m/s and Figure 5.4 for 18 m/s. Higher airspeeds were not tested due to the torque output limit of the VSDS motors. For elegance of presentation and ease of reading, t_c is offset to zero in each plot, and LCOs before controllers are activated are presented in dotted curves. Since both controllers are off before $t = t_c$, only the full trajectories of control surfaces deflections under the proposed NN controller are shown for illustration purpose. In the plots, leading- and trailing-edge control surfaces are expressed in short as 'LE' and 'TE', respectively. The trajectories of NN weights are presented in Figures 5.3 and 5.5.

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 5.2 and 5.3, 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 the PE condition being met, which leads to 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 near-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 5.4 and 5.5, where however, relatively larger differences between the the responses under the two controllers can be observed.

To better capture the performance differences between the two controllers, performance cost is evaluated for $t = 0 \rightarrow 4$ s according to Eq. (5.5) using the experiment

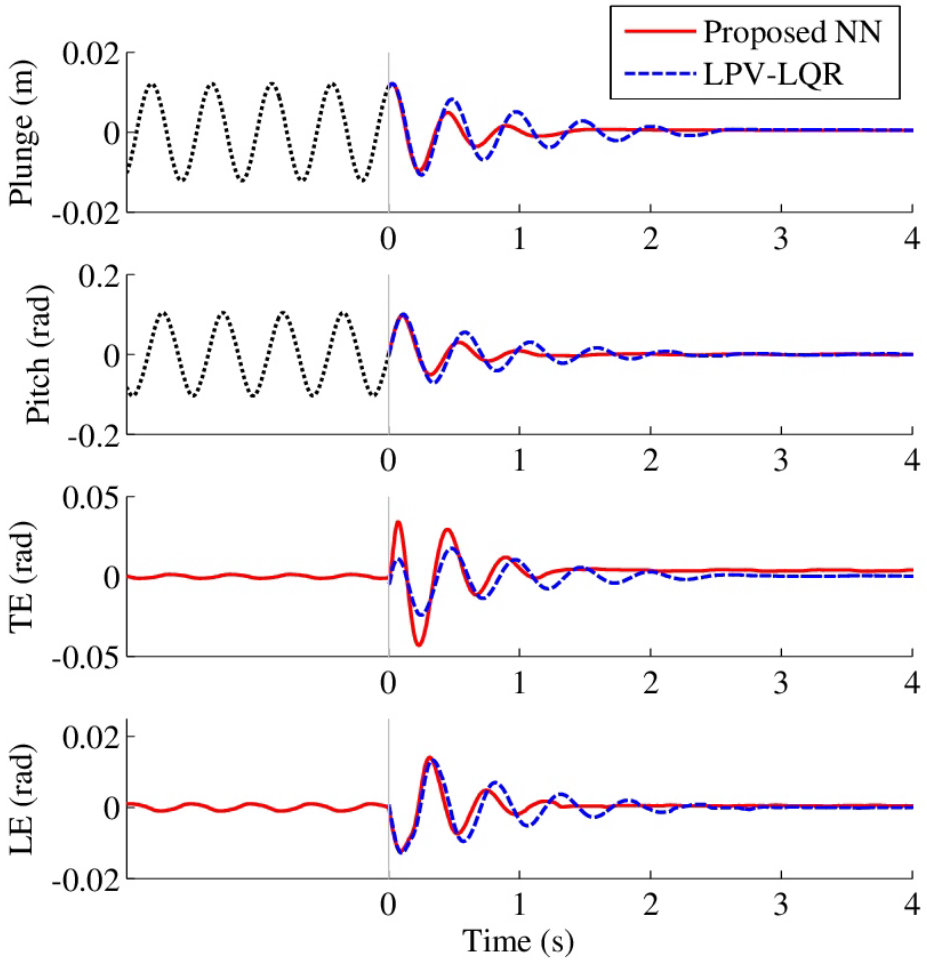


Figure 5.2: Suppressing developed flutter at 14.8 m/s airflow speed using the proposed NN controller and an LPV-LQR controller.

Table 5.3: Performance costs calculated from experiment data.

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

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 5.3. It can be concluded from Table 5.3 that the proposed NN control suppresses the flutter better with lower cost at both airspeeds, compared with the LPV-LQR control.

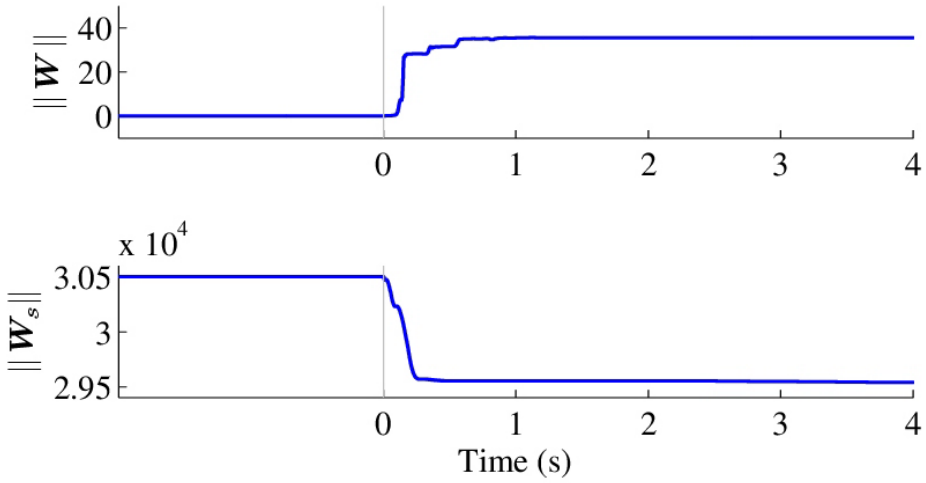


Figure 5.3: Convergence trajectories of the critic and identifier NN weights of the proposed controller at 14.8 m/s airflow speed.

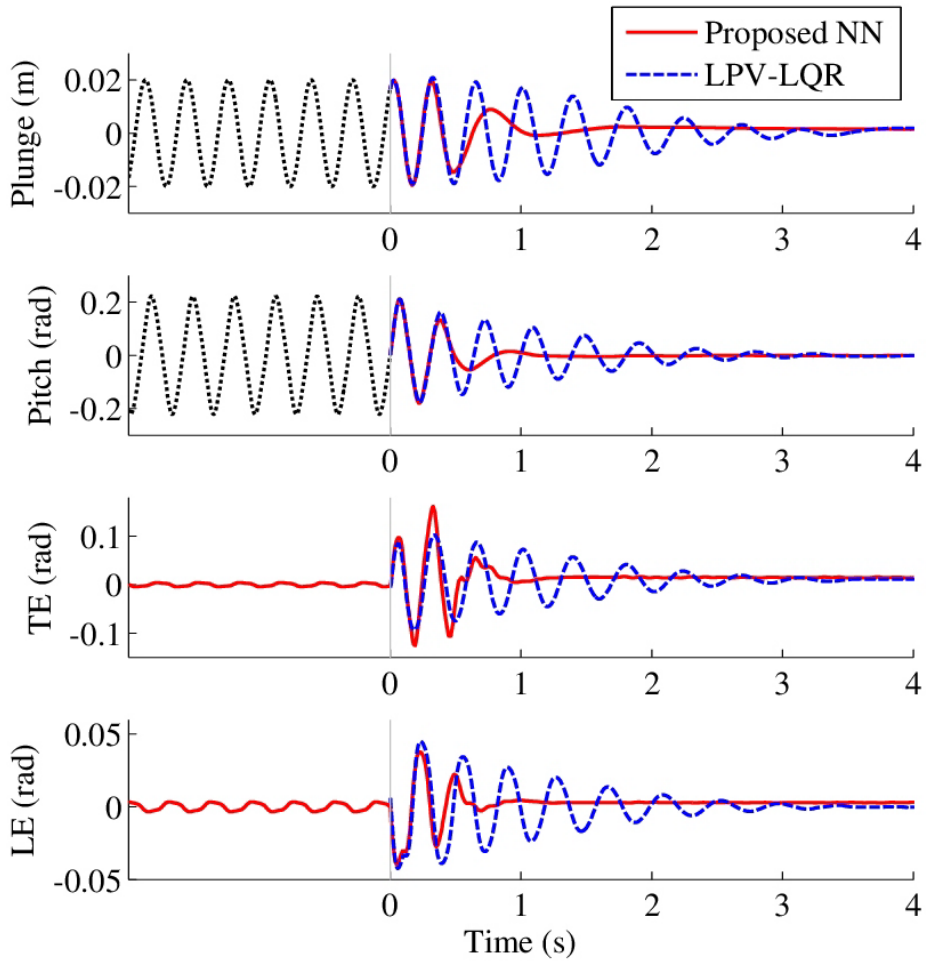


Figure 5.4: Suppressing developed flutter at 18 m/s airflow speed using the proposed NN controller and an LPV-LQR controller.

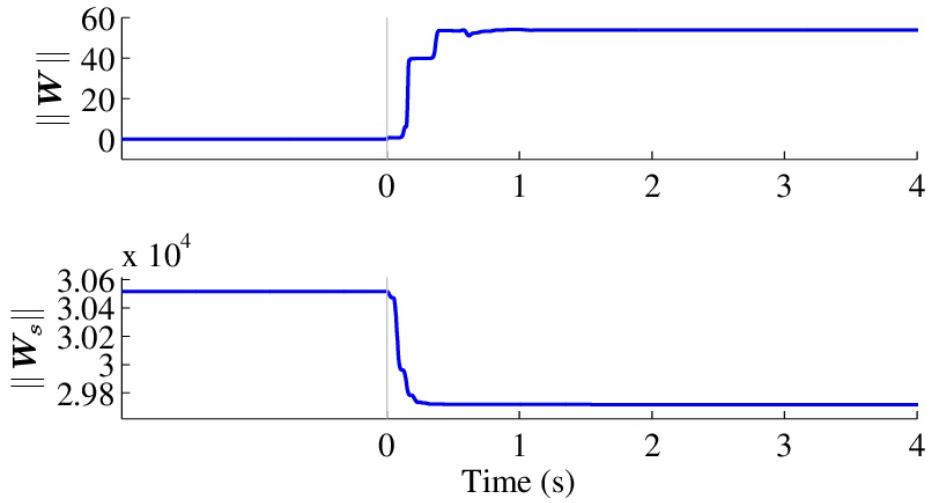


Figure 5.5: Convergence trajectories of the critic and identifier NN weights of the proposed controller at 18 m/s airflow speed.

5.5 Conclusions

In this paper a novel controller is proposed for AFS on aerofoils, featuring NOCOS with an MVFA tuned by an EKF. Convergence and stability analysis shows an important advantage of the new method that asymptotic convergence of the closed-loop system states is guaranteed during online learning. This is a stronger stability than the uniformly ultimately bounded stability of other existing NOCOS algorithms. The proposed systematic procedure based on linear matrix inequalities for the design of a scheduled parameter matrix further generalises the MVFA from locally nonlinear cases to globally nonlinear scenarios to suit AFS and other potential applications with strong nonlinearities that also vary nonlinearly with non-state independent variables. With an NN-based online system identification scheme, un-known dynamics can be estimated in real time for online control improvement. As validated in the wind-tunnel experiments, the proposed controller satisfactorily mitigates the impact of modelling uncertainties and improves AFS from the optimal control perspective, with the AFS controller synthesis dilemma successfully addressed. Experiments also confirm that the proposed controller is suitable for real-time implementation.

Chapter 6

Flutter suppression by input-constrained optimal control

This chapter is based on *Article-4*, focusing on nonlinear optimal control online synthesis (NOCOS) under control-input constraints (CICs). By further generalising the novel NOCOS scheme proposed in Chapter 4 and the new active flutter suppression (AFS) controller proposed in Chapter 5 to treat CICs from the optimal control perspective, a new adaptive nonlinear optimal controller for AFS under CICs is derived. This work completes the fourth objective (Page 4) of the thesis, and most importantly, accomplishes the primary aim of the research in this thesis.

Statement of Authorship

Title of Paper	Neural network based adaptive nonlinear optimal control for active suppression of airfoil flutter with constrained inputs
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Tang D, Chen L, Tian ZF and Hu E (2018) A neural-network approach for improving airfoil flutter suppression under control-input constraints. *Journal of Vibrations and Control.*

Principal Author

Name of Principal Author (Candidate)	Difan Tang		
Contribution to the Paper	Proposing and deriving methods and theories, performing theoretical analysis, conducting experiments, collecting data, interpreting results, and writing manuscript.		
Overall percentage (%)	80 %		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	11 December 2018

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Lei Chen		
Contribution to the Paper	Supervising development of work, evaluating methods and theories, helping with setting up experiments, providing advice in mathematical aspects, and evaluating manuscript.		
Signature		Date	18/12/18

Name of Co-Author	Zhao Feng Tian		
Contribution to the Paper	Supervising development of work, providing advice in mathematical aspects, and evaluating manuscript.		
Signature		Date	13/12/18

Name of Co-Author	Eric Hu		
Contribution to the Paper	Supervising development of work, providing advice in mathematical aspects, and evaluating manuscript.		
Signature		Date	12/12/2018

abstract

This paper deals with improving aerofoil flutter suppression (AFS) under control-input constraints (CICs) from the optimal control perspective by proposing a novel optimal neural-network control (ONNC). The proposed CIC-ONNC approach uses a modified value function approximation (MVFA) dynamically tuned by an extended Kalman filter to solve the Hamilton-Jacobi-Bellman equality online for continuously improved optimal control to address optimality in globally nonlinear systems. CICs are integrated into the controller synthesis by introducing a generalised nonquadratic cost function for control inputs. The feasibility of using a performance index involving the nonquadratic control-input cost with the MVFA is examined through the Lyapunov stability analysis. Wind-tunnel experiments were conducted for controller validation, where an optimal controller synthesised offline via linear-parameter-varying technique was used as a benchmark and compared. It is shown, both theoretically and experimentally, that the proposed CIC-ONNC method can effectively improve AFS under CICs.

6.1 Introduction

Aerofoil flutter is destructive vibration that occurs at and beyond a particular airspeed (flutter boundary) and can damage the aerofoil. For active flutter suppression (AFS), there have been enormous studies on solutions involving piezoelectric-material-actuated structures and aerofoil control-surface deployment. Despite different AFS solutions in terms of mechanical realisation, the underlying control algorithms play a crucial role in successful implementation.

Conventional frequency-domain analysis and basic state-space methods are useful in control synthesis (Schmidt, 2016) and actuators location optimisation (Song and Li, 2014). However, aeroelastic systems, being generally nonlinear and time-varying, pose significant challenges in characterisation and modelling, and make controller synthesis faced with numerous difficulties under increasing demands on performance improvement over extended flight envelopes. Accordingly, recent AFS studies mostly focus on advanced methods dealing with adaptive, nonlinear, and robust control. Pak et al. (1995) developed an approach to synthesise the linear-quadratic regulator online at the cost of optimality loss due to limiting the number of iterations in order to reduce computation loads; Strganac et al. (2000) tackled the nonlinearities using feedback linearisation, where optimality was retained for the linearised dynamics. Platanitis and Strganac (2004) extended the feedback-linearisation method to the case where control surfaces at both leading and trailing edges were involved. Ko et al. (2002) employed adaptive control built upon a reference model of known structure. Singh and Wang (2002) examined the possibility of utilising only plunge or pitch displacement measurement for AFS without exact knowledge of the aeroelastic system using back-stepping technique. Differently, Singh and Brenner (2003) used a passive observer and treated parameter estimation errors as disturbances, with the closed-loop stability guaranteed by an input-to-state stabilising control law. Similar modular approach can also be seen in the work of Rao et al. (2006). Viswamurthy and Ganguli (2008) focused on

helicopter blades with multiple flaps and derived deferentially weighted control. Carnahan and Richards (2008) removed the filter from the feed-forward path of the least-mean-square algorithm and proposed a modified version without the need to adjust its convergence coefficient. Prime et al. (2010) synthesised suboptimal control scheduled across a wide range of airspeeds by employing the linear-parameter-varying (LPV) technique. In the work of Chen et al. (2012), the LPV control was investigated on a proposed high-fidelity LPV model of reduced order. Lee and Singh (2013) tackled the presence of modelling uncertainties as well as various exogenous aerodynamic disturbances by developing an $\mathcal{L}_1$ adaptive control law. Fazelzadeh et al. (2014) dealt with nonlinearities and optimality using time-domain finite-element approach performed offline. Wang et al. (2015) introduced sliding-mode control (SMC) for AFS in hypersonic scenarios. Under the SMC framework the circumstances involving control time-delay were studied in Luo et al. (2016b). In addition to control delay, Gao and Cai (2016) took actuator faults into account. Beside actuator faults, Gao et al. (2016) also considered sensor faults in control design. By using pitch angle measurement only, Zhang and Behal (2016) derived and investigated a robust output-feedback scheme for AFS under aerodynamic perturbations. Fazelzadeh et al. (2017) proposed a hybrid control structure featuring both adaptive and robust control. With the advances in computing, neural networks (NNs) have been increasingly used. Gujjula et al. (2005) raised two hypotheses of structural nonlinearities and used adaptive control to treat linearly parameterised case and NN control to tackle non-parameterisable circumstance. In the work of Wang et al. (2011), a different NN was used to deal with scenarios under various exogenous disturbances. Brillante and Mannarino (2016) employed two recurrent NNs for system identification and control respectively. Tang et al. (2018) made further improvement to AFS by deriving an algorithm powered by two NNs to synthesise nonlinear optimal control in real time.

To allow aircraft normal and tactical maneuver while performing AFS, it is important to limit the amplitude of control used for AFS by setting constraints

smaller than actuators saturation bounds. Although control-input constraints (CICs) have been considered in some AFS studies (Gao and Cai, 2016; Gao et al., 2016; Ko et al., 2002; Viswamurthy and Ganguli, 2008; Wang et al., 2011), none of the existing solutions address the problem in the sense of optimal control. Despite numerous methods of nonlinear optimal control online synthesis (NOCOS) for systems with CICs being available, these methods are inapplicable to AFS due to problems related to stability, application scope, and real-time implementation. Moreover, these methods are limited to locally nonlinear systems whereas aeroelastic systems are globally nonlinear as the dynamics also change nonlinearly with the freestream airflow speed.

Therefore, in this study we focus on optimal AFS under CICs, and deliver the following two contributions:

- As a major technical contribution, a new algorithm is proposed for AFS, which performs NOCOS for globally nonlinear systems with CICs taken into account. The resulting controller is optimal under CICs and adaptive to airspeeds, and to our knowledge, a pioneer solution to adaptive nonlinear optimal control for AFS under CICs. Controller validation was conducted in wind-tunnel experiments.

- Unlike existing NOCOS approaches considering CICs, a modified value function approximation (MVFA) is used in this paper specifically for AFS. For the first time, the feasibility of using the MVFA with a performance index involving a nonquadratic control-input cost function to address CICs is examined through the Lyapunov stability analysis. This delivers an important scientific contribution to the NOCOS theory framework.

6.2 Aeroelastic system

In this study, AFS is mechanically realised by proper deployment of the aerofoil slat at leading-edge (LE) and the flap at trailing-edge (TE). The aeroelastic system

involves a typical rigid aerofoil section with plunge and pitch oscillations. For controller design, the analytical model takes the form studied in Prime (2010), which describes aeroelastic behaviours at subsonic regime using four degrees-of-freedom (DOF), capturing the coupled dynamics between plunge, pitch, LE, and TE DOFs including servo dynamics (see Figure 5.1). The corresponding analytical model is given in Section 5.2 of Chapter 5.

6.3 Proposed controller

6.3.1 Problem formulation

If the airspeed U_∞ is constant, Eq. (5.2) reduces to

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})\boldsymbol{u}, \tag{6.1}$$

which is locally nonlinear due to the absence of the variable U_∞ as compared with Eq. (5.2) .

The properties of (6.1) are given by Assumptions 4.1 and 4.2 in Chapter 4.

A proper control law $\boldsymbol{u}$ is desired to minimise

$$V(\boldsymbol{x}_0) = \int_0^\infty \bar{Q}(\boldsymbol{x}(\tau))d\tau + \int_0^\infty \bar{R}(\boldsymbol{u}(\tau))d\tau, \tag{6.2}$$

where the penalty to states is set by $\bar{Q}(\boldsymbol{x}(\tau))$ and the control is weighted by $\bar{R}(\boldsymbol{u}(\tau))$, both of which are functions being positive-definite and monotonically increasing.

The control that minimises Eq. (6.2) for the same initial conditions is deemed optimal and denoted as $\boldsymbol{u}^*(\boldsymbol{x})$. The associated cost is generally known as the 'value function', denoted by $V^*(\boldsymbol{x})$.

6.3.2 Input-constrained optimal control

In conventional optimal control problems, $\bar{Q}(\boldsymbol{x})$ and $\bar{R}(\boldsymbol{u})$ take quadratic forms, which however, do not apply to cases under CICs. To take CICs into account, a

positive-definite integrand function is introduced for $\bar{R}(\boldsymbol{u})$ in the following form (Lyshevski, 1998):

$$\bar{R}(\boldsymbol{u}) = 2\int \left(\boldsymbol{\Theta}^{-1}(\boldsymbol{u})\right)^{\mathrm{T}} \boldsymbol{R} d\boldsymbol{u}, \tag{6.3}$$

where $\boldsymbol{u} = \boldsymbol{\Theta}(\boldsymbol{v}) = [\theta(v_1),\ \cdots,\ \theta(v_{n_u})]^{\mathrm{T}}$ with $\theta(\cdot)$ being a bounded function that is p^{th}-order continuous ($p \geq 1$) and $\boldsymbol{v} \in \mathbb{R}^{n_u}$; weighting $\boldsymbol{R} \in \mathbb{R}^{n_u \times n_u}$ and $\boldsymbol{R} = \boldsymbol{R}^{\mathrm{T}} \succ 0$.

For $\boldsymbol{\Theta}(\boldsymbol{v})$ in Eq. (6.3), it is practically feasible to choose

$$\boldsymbol{\Theta}(\boldsymbol{v}) = \boldsymbol{\lambda}\tanh(\boldsymbol{\lambda}^{-1}\boldsymbol{v}),$$

and accordingly,

$$\begin{aligned}\bar{R}(\boldsymbol{u}) &= 2\int \left[\boldsymbol{\lambda}\tanh^{-1}(\boldsymbol{\lambda}^{-1}\boldsymbol{u})\right]^{\mathrm{T}} \boldsymbol{R} d\boldsymbol{u} \\ &= 2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\lambda}\boldsymbol{R}\tanh^{-1}\left(\boldsymbol{\lambda}^{-1}\boldsymbol{u}\right) + \boldsymbol{R}_r\boldsymbol{\lambda}^2 \ln\left(\overrightarrow{\boldsymbol{1}} - \left(\boldsymbol{\lambda}^{-1}\boldsymbol{u}\right)^2\right),\end{aligned} \tag{6.4}$$

where the functions $\tanh(\cdot)$, $\tanh^{-1}(\cdot)$ and $\ln(\cdot)$ perform element-wise operation, the vector $\boldsymbol{R}_r = [R_1, R_2, \cdots, R_{n_u}]$ contains diagonal elements of $\boldsymbol{R}$, diagonal matrix $\boldsymbol{\lambda} = \mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_{n_u}]$ contains bounds for respective control ($-\lambda_i \leq u_i \leq \lambda_i$, $\forall \lambda_i \in R^+$, $i = 1, 2, \cdots n_u$), and $\overrightarrow{\boldsymbol{1}} = [1, 1, \cdots, 1]^{\mathrm{T}}$.

Remark 6.1. It can be seen from Eq. (6.4) that for any unbounded control $\boldsymbol{v}$, if $|\boldsymbol{v}| \leq \boldsymbol{\lambda}$ then $\bar{R}(\boldsymbol{u}) \approx \boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u} \approx \boldsymbol{v}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{v}$, and if $|\boldsymbol{v}| > \boldsymbol{\lambda}$ then $\bar{R}(\boldsymbol{u}) \ll \boldsymbol{v}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{v}$.

With the introduction of Eq. (6.4), differentiating Eq. (6.2) yields

$$\begin{aligned}&\nabla V^{\mathrm{T}}(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}) + \bar{Q} + \boldsymbol{R}_r\boldsymbol{\lambda}^2 \ln\left(\overrightarrow{\boldsymbol{1}} - \left(\boldsymbol{\lambda}^{-1}\boldsymbol{u}\right)^2\right) \\ &\quad + 2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\lambda}\boldsymbol{R}\tanh^{-1}\left(\boldsymbol{\lambda}^{-1}\boldsymbol{u}\right) = 0; \quad V(0) = 0.\end{aligned} \tag{6.5}$$

On the basis of Eq. (6.5), the following Hamilton-Jacobi-Bellman (HJB) equality holds under CICs:

$$\begin{aligned}&\bar{Q} + \boldsymbol{R}_r\boldsymbol{\lambda}^2 \ln\left(\overrightarrow{\boldsymbol{1}} - \tanh^2\left(\frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla V^*\right)\right) \\ &\quad + \nabla V^{*\mathrm{T}}\left[\boldsymbol{f} - \boldsymbol{g}\boldsymbol{\lambda}\tanh\left(\frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla V^*\right)\right] \\ &\quad + \nabla V^{*\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\tanh\left(\frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla V^*\right) = 0,\end{aligned} \tag{6.6}$$

the solution of which is V^*, with

$$\boldsymbol{u}^* = -\boldsymbol{\lambda} \tanh\left(\frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla V^*\right), \tag{6.7}$$

where $V^*(0) = 0$ and $\nabla V^* \triangleq \frac{\partial V^*}{\partial \boldsymbol{x}}$.

6.3.3 Value function approximation for cases with constant U_∞

To analytically determine $V^*(\boldsymbol{x})$ is difficult, which instead, can be obtained through an iterative procedure termed as 'policy iteration' (Jiang and Jiang, 2015), which requires $V^*(\boldsymbol{x})$ being appropriately structured and approximated. With universal approximation properties, neural networks (NNs) can be used (Hornik et al., 1989). The discussion in Section 6.1 reveals that a majority of the available methods for NOCOS, being subjected to issues related to stability, application scope, and real-time implementation, are inapplicable to AFS. Accordingly, a modified value function approximation (MVFA) as in Tang et al. (2018) is introduced specifically for AFS as

$$V^*(\boldsymbol{x}) = \boldsymbol{W}^{*\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{x}) + \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} + \varepsilon(\boldsymbol{x}), \tag{6.8}$$

where hidden-layer neurons are contained in $\boldsymbol{\Phi}(\boldsymbol{x}) \in \mathbb{R}^{n_n}$, with ideal NN weights being $\boldsymbol{W}^* \in \mathbb{R}^{n_n}$; $\boldsymbol{P} \in \mathbb{R}^{n_x \times n_x}$ is an additional parameter matrix that is diagonal and positive-definite; the error of approximation is denoted by $\varepsilon(\boldsymbol{x}) \in \mathbb{R}$.

Remark 6.2. For the first time, in this paper the MVFA is used with a performance cost involving the generalised nonquadratic functional as in Eq. (6.3) to solve for optimal control under CICs. As one of the major contributions, this input-constrained case with MVFA is examined through the Lyapunov stability analysis, as to be presented in Theorems 6.1 and 6.2.

Remark 6.3. To select proper activation functions for the hidden-layer neurons of the NN in Eq. (6.8), one may apply Weierstrass approximation using high-order polynomials (Finlayson, 1972). The resulting activation functions are the individual terms of a polynomial of specified order with the NN inputs as variables.

Accordingly, there is

$$\nabla V^* = \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* + \boldsymbol{P}\boldsymbol{x} + \nabla\varepsilon, \tag{6.9}$$

with $\bar{\nabla}\boldsymbol{\Phi} = \nabla\boldsymbol{\Phi}^{\mathrm{T}} = \left[\frac{\partial\boldsymbol{\Phi}}{\partial\boldsymbol{x}}\right]^{\mathrm{T}}$ and $\nabla\varepsilon = \frac{\partial\varepsilon}{\partial\boldsymbol{x}}$.

The boundedness properties of $\bar{\nabla}\boldsymbol{\Phi}(\boldsymbol{x})$ and $\nabla\varepsilon(\boldsymbol{x})$ are given in Assumption 4.3 in Chapter 4.

Note that $\boldsymbol{W}^*$ is unknown and is replaced by an estimate $\hat{\boldsymbol{W}}$ during real-time implementation. Under the proposed approximation scheme, the corresponding estimations $\hat{V}(\boldsymbol{x})$, $\nabla\hat{V}(\boldsymbol{x})$, and $\hat{\boldsymbol{u}}(\boldsymbol{x})$, expressed in short omitting the variable $\boldsymbol{x}$, are

$$\hat{V} = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} + \hat{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{\Phi}, \tag{6.10}$$

$$\nabla\hat{V} = \boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}}, \tag{6.11}$$

$$\hat{\boldsymbol{u}} = -\boldsymbol{\lambda}\tanh\left(\frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\hat{V}\right). \tag{6.12}$$

As a result, Eq. (6.6) becomes

$$\begin{aligned}\bar{Q} &+ \boldsymbol{R}_r\boldsymbol{\lambda}^2\ln\left(\overrightarrow{\mathbf{1}} - \tanh^2(\frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\hat{V})\right)\\ &+ \nabla\hat{V}^{\mathrm{T}}\left[\boldsymbol{f} - \boldsymbol{g}\boldsymbol{\lambda}\tanh\left(\frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\hat{V}\right)\right]\\ &+ \nabla\hat{V}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\tanh\left(\frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\hat{V}\right) = e,\end{aligned} \tag{6.13}$$

where $\hat{V}(0) = 0$ and $e \in \mathbb{R}$ is the error resulted from approximation.

It is clear that e is to be minimised so that $\hat{\boldsymbol{W}} \to \boldsymbol{W}^*$. The extended Kalman filter (EKF) can be applied and Eq. (6.13) can be rewritten as a dynamic system

$$\begin{cases}\dot{\hat{\boldsymbol{W}}} = 0 + \boldsymbol{w}_v,\\ \quad y = \beta(\boldsymbol{x}, \hat{\boldsymbol{W}}) - e + v_v,\end{cases} \tag{6.14}$$

with

$$y = -\bar{Q} - 2\hat{\boldsymbol{u}}^{\mathrm{T}}\boldsymbol{\lambda}\boldsymbol{R}\tanh^{-1}\left(\boldsymbol{\lambda}^{-1}\hat{\boldsymbol{u}}\right) - \boldsymbol{R}_r\boldsymbol{\lambda}^2\ln\left(\overrightarrow{\mathbf{1}} - \left(\boldsymbol{\lambda}^{-1}\hat{\boldsymbol{u}}\right)^2\right),$$

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

where $\boldsymbol{w}_v$ contains virtual process noises of covariance $\boldsymbol{Q}_f \in \mathbb{R}^{n_n \times n_n} \succ 0$ and v_v is virtual measurement noise of covariance $R_f \in \mathbb{R}^+$.

Now applying an EKF to Eq. (6.14) yields

$$\begin{cases} \hat{y} = \beta(\boldsymbol{x}, \hat{\boldsymbol{W}}), \\ \dot{\hat{\boldsymbol{W}}} = \kappa_c \boldsymbol{K}_f e_y, \\ \kappa_c = \dfrac{\alpha_{cnn}}{\left\|\boldsymbol{K}_f\right\| + 1}, \\ e_y = y - \hat{y}, \end{cases} \tag{6.15}$$

where the EKF has a gain of $\boldsymbol{K}_f \in \mathbb{R}^{n_n \times 1}$, y has an estimate of $\hat{y}$, and $\alpha_{cnn} \in \mathbb{R}^+$ is an auxiliary constant normalisation gain.

Specifically,

$$\boldsymbol{K}_f = \boldsymbol{S}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}_f^{-1}, \tag{6.16}$$

$$\boldsymbol{H}^{\mathrm{T}} = \frac{\partial \beta(\boldsymbol{x}, \hat{\boldsymbol{W}})}{\partial \hat{\boldsymbol{W}}} = \bar{\nabla}\boldsymbol{\Phi}\left(\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}}\right), \tag{6.17}$$

$$\dot{\boldsymbol{S}} = \boldsymbol{Q}_f - \boldsymbol{S}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}_f^{-1}\boldsymbol{H}\boldsymbol{S}, \tag{6.18}$$

where $\boldsymbol{S} \in \mathbb{R}^{n_n \times n_n}$ is symmetric and positive-definite.

For online tuning, the persistence of excitation (PE) condition as in Assumption 4.4 is assumed to hold.

Theorem 6.1. *For nonlinear systems as in Eq. (6.1), let the value function be approximated by Eq. (6.8), the control be in the form of Eq. (6.12), and the NN weights $\hat{\boldsymbol{W}}$ be dynamically adjusted using the EKF as in Eqs. (6.15) to (6.18) with estimation error of $\tilde{\boldsymbol{W}} = \boldsymbol{W}^* - \hat{\boldsymbol{W}}$. On the basis of Assumptions 4.1 to 4.4, $\tilde{\boldsymbol{W}}$ remain uniformly ultimately bounded (UUB) during online learning, if the parameter matrix $\boldsymbol{P}$ in Eq. (6.8) is selected to be $\|\boldsymbol{P}\| > b_{mP}$ for a finite constant $b_{mP} \in \mathbb{R}^+$.*

Proof. Consider

$$\begin{aligned} \mathcal{L} &= \hat{V} + \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}(\kappa_c \boldsymbol{S})^{-1}\tilde{\boldsymbol{W}} \\ &= \mathcal{L}_V + \mathcal{L}_W, \end{aligned} \tag{6.19}$$

where $\mathcal{L}_V = \hat{V}$ and $\mathcal{L}_W = \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}(\kappa_c \boldsymbol{S})^{-1}\tilde{\boldsymbol{W}}$.

With Eqs. (6.10) to (6.12), the time derivative of $\mathcal{L}_V$ is

$$\dot{\mathcal{L}}_V = \left(\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\hat{\boldsymbol{W}} + \boldsymbol{P}\boldsymbol{x}\right)^{\mathrm{T}}\left[f - g\boldsymbol{\lambda}\tanh(\hat{\boldsymbol{\Upsilon}})\right], \tag{6.20}$$

where

$$\hat{\boldsymbol{\Upsilon}} = \frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}g^{\mathrm{T}}\nabla\hat{V}. \tag{6.21}$$

By using Maclaurin series, there is

$$\tanh(\hat{\boldsymbol{\Upsilon}}) = \hat{\boldsymbol{\Upsilon}} + \mathcal{O}(\hat{\boldsymbol{\Upsilon}}^3), \quad \text{for } \left|\hat{\boldsymbol{\Upsilon}}\right| < \frac{\pi}{2}. \tag{6.22}$$

Based on Eqs. (6.21), (6.22), and Assumptions 4.2 and 4.3, we have

$$\tanh(\hat{\boldsymbol{\Upsilon}}) \approx \hat{\boldsymbol{\Upsilon}} + \boldsymbol{\varepsilon}_{m\hat{\Upsilon}}, \quad \forall\hat{\boldsymbol{\Upsilon}}, \tag{6.23}$$

and there is $\left\|\boldsymbol{\varepsilon}_{m\hat{\boldsymbol{\Upsilon}}}\right\| \leq b_{\varepsilon m\hat{\boldsymbol{\Upsilon}}}\left\|\boldsymbol{x}\right\|$ where constant $b_{\varepsilon m\hat{\boldsymbol{\Upsilon}}} \in \mathbb{R}^+$.

With Eqs. (6.11), (6.21) and (6.23), Eq. (6.20) becomes

$$\begin{aligned}
\dot{\mathcal{L}}_V \approx{}& \left(\boldsymbol{P}\boldsymbol{x} + \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* - \bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}\right)^{\mathrm{T}}\Bigg(f - \frac{1}{2}g\boldsymbol{R}^{-1}g^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} \\
&- \frac{1}{2}g\boldsymbol{R}^{-1}g^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* + \frac{1}{2}g\boldsymbol{R}^{-1}g^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - g\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\boldsymbol{\Upsilon}}}\Bigg) \\
={}& -\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}g\boldsymbol{R}^{-1}g^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} - \frac{1}{2}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}g\boldsymbol{R}^{-1}g^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* \\
&- \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}g\boldsymbol{R}^{-1}g^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}g\boldsymbol{R}^{-1}g^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* \\
&+ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}f + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}g\boldsymbol{R}^{-1}g^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}g\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\boldsymbol{\Upsilon}}} \\
&+ \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}f + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}g\boldsymbol{R}^{-1}g^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} \\
&- \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}g\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\boldsymbol{\Upsilon}}} - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}f + \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}g\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\boldsymbol{\Upsilon}}}.
\end{aligned} \tag{6.24}$$

In regard to $\mathcal{L}_W$ in Eq. (6.19), by using Eqs. (6.14) to (6.16), the time derivative of Eq. (6.19) is

$$\begin{aligned}
\dot{\mathcal{L}}_W &= -\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{S}^{-1}\boldsymbol{K}_f(y - \hat{y}) \\
&= -\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}R_f^{-1}(y - \hat{y}).
\end{aligned} \tag{6.25}$$

First we look at the term $y - \hat{y}$ of Eq. (6.25). From Eqs. (6.14) and (6.15), there is

$$\begin{aligned} y - \hat{y} = & -\bar{Q} - 2\hat{\boldsymbol{u}}^{\mathrm{T}}\boldsymbol{\lambda}\boldsymbol{R}\tanh^{-1}\left(\boldsymbol{\lambda}^{-1}\hat{\boldsymbol{u}}\right) \\ & - \boldsymbol{R}_r\boldsymbol{\lambda}^2\ln\left(\overrightarrow{\mathbf{1}} - \left(\boldsymbol{\lambda}^{-1}\hat{\boldsymbol{u}}\right)^2\right) \\ & - \left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} + \hat{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\right)(\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}}). \end{aligned} \tag{6.26}$$

By using Eq. (6.5) with $\boldsymbol{u}^*$ substituted for $\boldsymbol{u}$ as an equivalent expression to Eq. (6.6), together with Eqs. (6.7), (6.9), (6.11) and (6.12), Eq. (6.26) becomes

$$y - \hat{y} = e_{\bar{R}} + e_{\dot{V}}, \tag{6.27}$$

where

$$\begin{aligned} e_{\bar{R}} = & -2\boldsymbol{u}^{*\mathrm{T}}\boldsymbol{\lambda}\boldsymbol{R}\mathbf{Y}^* + \boldsymbol{R}_r\boldsymbol{\lambda}^2\ln\left(\overrightarrow{\mathbf{1}} - \tanh^2(\mathbf{Y}^*)\right) \\ & + 2\hat{\boldsymbol{u}}^{\mathrm{T}}\boldsymbol{\lambda}\boldsymbol{R}\hat{\mathbf{Y}} - \boldsymbol{R}_r\boldsymbol{\lambda}^2\ln\left(\overrightarrow{\mathbf{1}} - \tanh^2(\hat{\mathbf{Y}})\right), \end{aligned} \tag{6.28}$$

$$\begin{aligned} e_{\dot{V}} = & \left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi} + \nabla\varepsilon^{\mathrm{T}}\right)(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*) \\ & - \left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} + \hat{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\right)(\boldsymbol{f} + \boldsymbol{g}\hat{\boldsymbol{u}}), \end{aligned} \tag{6.29}$$

with

$$\mathbf{Y}^* = \frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla V^*. \tag{6.30}$$

Using Eq. (6.30) and substituting $\boldsymbol{u}^*$ with Eqs. (6.7) and (6.9) gives

$$\begin{aligned} 2\boldsymbol{u}^{*\mathrm{T}}\boldsymbol{\lambda}\boldsymbol{R}\mathbf{Y}^* &= 2\left[-\boldsymbol{\lambda}\tanh(\frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla V^*)\right]^{\mathrm{T}}\boldsymbol{\lambda}\boldsymbol{R}\mathbf{Y}^* \\ &= -2\left(\frac{1}{2}\boldsymbol{\lambda}^{-1}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla V^*\right)^{\mathrm{T}}\boldsymbol{\lambda}^2\boldsymbol{R}\tanh(\mathbf{Y}^*) \\ &= -\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi} + \nabla\varepsilon^{\mathrm{T}}\right)\boldsymbol{g}\boldsymbol{\lambda}\tanh(\mathbf{Y}^*). \end{aligned} \tag{6.31}$$

Similarly, with Eqs. (6.11), (6.12) and (6.21), there is

$$\begin{aligned} 2\hat{\boldsymbol{u}}^{\mathrm{T}}\boldsymbol{\lambda}\boldsymbol{R}\hat{\mathbf{Y}} &= -\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} + \hat{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\right)\boldsymbol{g}\boldsymbol{\lambda}\tanh(\hat{\mathbf{Y}}) \\ &= -\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi} - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\right)\boldsymbol{g}\boldsymbol{\lambda}\tanh(\hat{\mathbf{Y}}). \end{aligned} \tag{6.32}$$

In the following, we deal with the term $\boldsymbol{R}_r\boldsymbol{\lambda}^2 \ln\left(\overrightarrow{\mathbf{1}} - \tanh^2(\mathbf{Y}^*)\right)$ in Eq. (6.28). Since $\left|\overrightarrow{\mathbf{1}} - \tanh^2(\mathbf{Y}^*)\right| < 1$, by using Mercator series, there is

$$\begin{aligned}\ln\left(\overrightarrow{\mathbf{1}} - \tanh^2(\mathbf{Y}^*)\right) &= \ln\left(\overrightarrow{\mathbf{1}} + \tanh(\mathbf{Y}^*)\right) + \ln\left(\overrightarrow{\mathbf{1}} - \tanh(\mathbf{Y}^*)\right) \\ &= -\tanh^2(\mathbf{Y}^*) + \mathcal{O}\left(\tanh^4(\mathbf{Y}^*)\right) \\ &= -\tanh^2(\mathbf{Y}^*) + \boldsymbol{\varepsilon}_{\mathrm{Y}^*}, \end{aligned} \tag{6.33}$$

where $\tanh^2(\cdot)$ and $\tanh^4(\cdot)$ perform element-wise operations, $\boldsymbol{\varepsilon}_{\mathrm{Y}^*} = \mathcal{O}\left(\tanh^4(\mathbf{Y}^*)\right)$.

Then,

$$\boldsymbol{R}_r\boldsymbol{\lambda}^2 \ln\left(\overrightarrow{\mathbf{1}} - \tanh^2(\mathbf{Y}^*)\right) = -\tanh^{\mathrm{T}}(\mathbf{Y}^*)\boldsymbol{\lambda}^2\boldsymbol{R}\tanh(\mathbf{Y}^*) + \boldsymbol{R}_r\boldsymbol{\lambda}^2\boldsymbol{\varepsilon}_{\mathrm{Y}^*}. \tag{6.34}$$

Similarly, the term $\boldsymbol{R}_r\boldsymbol{\lambda}^2 \ln\left(\overrightarrow{\mathbf{1}} - \tanh^2(\hat{\mathbf{Y}})\right)$ in Eq. (6.28) can be expressed as

$$\boldsymbol{R}_r\boldsymbol{\lambda}^2 \ln\left(\overrightarrow{\mathbf{1}} - \tanh^2(\hat{\mathbf{Y}})\right) = -\tanh^{\mathrm{T}}(\hat{\mathbf{Y}})\boldsymbol{\lambda}^2\boldsymbol{R}\tanh(\hat{\mathbf{Y}}) + \boldsymbol{R}_r\boldsymbol{\lambda}^2\boldsymbol{\varepsilon}_{\hat{\mathrm{Y}}}, \tag{6.35}$$

where $\boldsymbol{\varepsilon}_{\hat{\mathrm{Y}}} = \mathcal{O}\left(\tanh^4(\hat{\mathbf{Y}})\right)$.

Substituting Eqs. (6.31), (6.32), (6.34) and (6.35) for respective terms in Eq. (6.28) yields

$$\begin{aligned} e_{\bar{R}} = {} & \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\left[\tanh(\mathbf{Y}^*) - \tanh(\hat{\mathbf{Y}})\right] \\ & + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\left[\tanh(\mathbf{Y}^*) - \tanh(\hat{\mathbf{Y}})\right] \\ & + \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\tanh(\hat{\mathbf{Y}}) + \nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\tanh(\mathbf{Y}^*) + \boldsymbol{R}_r\boldsymbol{\lambda}^2\boldsymbol{\varepsilon}_{e\mathrm{Y}} \\ & - \tanh^{\mathrm{T}}(\mathbf{Y}^*)\boldsymbol{\lambda}^2\boldsymbol{R}\tanh(\mathbf{Y}^*) + \tanh^{\mathrm{T}}(\hat{\mathbf{Y}})\boldsymbol{\lambda}^2\boldsymbol{R}\tanh(\hat{\mathbf{Y}}). \end{aligned} \tag{6.36}$$

where $\boldsymbol{\varepsilon}_{e\mathrm{Y}} = \boldsymbol{\varepsilon}_{\mathrm{Y}^*} - \boldsymbol{\varepsilon}_{\hat{\mathrm{Y}}}$, and it is straightforward to see that $\|\boldsymbol{\varepsilon}_{e\mathrm{Y}}\| \leq b_{\varepsilon e\mathrm{Y}}$ with constant $b_{\varepsilon e\mathrm{Y}} \in \mathbb{R}^+$ and $b_{e\varepsilon\mathrm{Y}} \ll 1$.

Similar to Eq. (6.23), according to Eq. (6.30) with Maclaurin series expansion under Assumptions 4.2 and 4.3, we have

$$\tanh(\mathbf{Y}^*) \approx \mathbf{Y}^* + \boldsymbol{\varepsilon}_{m\mathrm{Y}^*}, \quad \forall \mathbf{Y}^*, \tag{6.37}$$

and $\|\boldsymbol{\varepsilon}_{m\mathbf{Y}^*}\| \leq b_{\varepsilon m\mathbf{Y}^*}\|\boldsymbol{x}\|$ with $b_{\varepsilon m\mathbf{Y}^*} \in \mathbb{R}^+$ being a constant.

By using the approximations as in Eqs. (6.23) and (6.37) with Eqs. (6.21) and (6.30), we obtain

$$\begin{aligned} e_{\bar{R}} = & \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} + \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon \\ & + \frac{1}{2}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} + \frac{1}{2}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon \\ & - \frac{1}{4}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} + \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon \\ & - \boldsymbol{\varepsilon}_{m\mathbf{Y}^*}^{\mathrm{T}}\boldsymbol{\lambda}^2\boldsymbol{R}\boldsymbol{\varepsilon}_{m\mathbf{Y}^*} + \boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}}^{\mathrm{T}}\boldsymbol{\lambda}^2\boldsymbol{R}\boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}} + \boldsymbol{R}_r\boldsymbol{\lambda}^2\boldsymbol{\varepsilon}_{e\mathrm{Y}}. \end{aligned} \tag{6.38}$$

In the following we deal with $e_{\dot{V}}$ as in Eq. (6.29). Substituting Eqs. (6.7), (6.9), (6.11), and (6.12) into Eq. (6.29) yields

$$\begin{aligned} e_{\dot{V}} = & \left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi} + \nabla\varepsilon^{\mathrm{T}}\right)\left[f - \boldsymbol{g}\boldsymbol{\lambda}\tanh(\mathbf{Y}^*)\right] \\ & - \left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi} - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\right)\left[f - \boldsymbol{g}\boldsymbol{\lambda}\tanh(\hat{\mathbf{Y}})\right] \\ = & - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\left[\tanh(\mathbf{Y}^*) - \tanh(\hat{\mathbf{Y}})\right] \\ & - \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\left[\tanh(\mathbf{Y}^*) - \tanh(\hat{\mathbf{Y}})\right] \\ & + \nabla\varepsilon^{\mathrm{T}}f - \nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\tanh(\mathbf{Y}^*) + \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}f \\ & - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\tanh(\hat{\mathbf{Y}}). \end{aligned} \tag{6.39}$$

Again, with approximations in Eqs. (6.23) and (6.37), Eq. (6.39) becomes

$$\begin{aligned} e_{\dot{V}} = & - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_m \\ & - \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon - \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_m \\ & + \nabla\varepsilon^{\mathrm{T}}f - \frac{1}{2}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon - \nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\mathbf{Y}^*} \\ & - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} \\ & + \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}f + \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}}, \end{aligned} \tag{6.40}$$

where $\boldsymbol{\varepsilon}_m = \boldsymbol{\varepsilon}_{m\mathbf{Y}^*} - \boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}}$, and $\|\boldsymbol{\varepsilon}_m\| \leq b_{m\mathrm{Y}^*}\|\boldsymbol{x}\|$, according to Eqs. (6.23) and (6.37).

With Eqs. (6.38) and (6.40) substituted for $e_{\bar{R}}$ and $e_{\dot{V}}$, respectively, Eq. (6.27) becomes

$$y - \hat{y} = e_{\bar{R}} + e_{\dot{V}}$$

$$
\begin{aligned}
&= -\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon \\
&\quad - \frac{1}{2}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \frac{1}{2}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon \\
&\quad + \frac{1}{4}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \frac{1}{4}\nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\nabla\varepsilon + \boldsymbol{R}_r\boldsymbol{\lambda}^2\boldsymbol{\varepsilon}_{e\mathrm{Y}} \\
&\quad - \boldsymbol{\varepsilon}_{m\mathbf{Y}^*}^{\mathrm{T}}\boldsymbol{\lambda}^2\boldsymbol{R}\boldsymbol{\varepsilon}_{m\mathbf{Y}^*} + \boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}}^{\mathrm{T}}\boldsymbol{\lambda}^2\boldsymbol{R}\boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_m - \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_m \\
&\quad + \nabla\varepsilon^{\mathrm{T}}\boldsymbol{f} - \nabla\varepsilon^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\mathbf{Y}^*} + \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}}. \qquad (6.41)
\end{aligned}
$$

Now we rewrite the term $\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}$ of Eq. (6.25) using Eqs. (6.17), (6.11), (6.12), (6.21), and (6.23). This gives

$$
\begin{aligned}
\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}} &= \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\left[\boldsymbol{f} - \boldsymbol{g}\boldsymbol{\lambda}\tanh(\hat{\mathbf{Y}})\right] \\
&= \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} - \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} - \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^* \\
&\quad + \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}}. \qquad (6.42)
\end{aligned}
$$

After assembling Eqs. (6.41) and (6.42) for Eq. (6.25), together with Eq. (6.24), we have the time derivative of Eq. (6.19) as

$$
\dot{\mathcal{L}} = \dot{\mathcal{L}}_V + \dot{\mathcal{L}}_W = T_1 + T_2 + T_3 + T_4 + T_5 + T_6, \qquad (6.43)
$$

with

$$
\begin{aligned}
T_1 = &-\frac{1}{16}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}R_f^{-1}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z} - \frac{3}{8}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\varepsilon R_f^{-1}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z} + \frac{3}{4}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}R_f^{-1}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}} \\
&- \frac{3}{4}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}R_f^{-1}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\mathbf{Y}^*} - \frac{3}{4}\boldsymbol{z}^{\mathrm{T}}\left(\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*\right)R_f^{-1}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}, \qquad (6.44)
\end{aligned}
$$

$$
T_2 = T_{2a} + T_{2b} + T_{2c} + T_{2d} + T_{2e}, \qquad (6.45)
$$

$$
\begin{aligned}
T_{2a} = &-\frac{1}{8}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x}R_f^{-1}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z} - \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x}R_f^{-1}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\boldsymbol{z} \\
&+ 2\boldsymbol{z}^{\mathrm{T}}\boldsymbol{f}R_f^{-1}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}} - \boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^*R_f^{-1}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}} \\
&- \boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x}R_f^{-1}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}} - \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}R_f^{-1}(\boldsymbol{f} - \boldsymbol{g}\boldsymbol{u}^*)^{\mathrm{T}}\nabla\varepsilon \\
&+ \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}R_f^{-1}\boldsymbol{\varepsilon}_{m\mathbf{Y}^*}^{\mathrm{T}}\boldsymbol{\lambda}^2\boldsymbol{R}\boldsymbol{\varepsilon}_{m\mathbf{Y}^*} - \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}R_f^{-1}\boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}}^{\mathrm{T}}\boldsymbol{\lambda}^2\boldsymbol{R}\boldsymbol{\varepsilon}_{m\hat{\mathbf{Y}}} \\
&+ \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}R_f^{-1}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_m + \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z}R_f^{-1}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_m, \qquad (6.46)
\end{aligned}
$$

$$
T_{2b} = -\boldsymbol{z}^{\mathrm{T}}\boldsymbol{f}R_f^{-1}(\boldsymbol{f} - \boldsymbol{g}\boldsymbol{u}^*)^{\mathrm{T}}\nabla\varepsilon + \boldsymbol{z}^{\mathrm{T}}\boldsymbol{f}R_f^{-1}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z} + \boldsymbol{z}^{\mathrm{T}}\boldsymbol{f}R_f^{-1}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\boldsymbol{z}
$$

$$
\begin{aligned}
&- \frac{1}{4} z^{\mathrm{T}} f R_f^{-1} \nabla \varepsilon^{\mathrm{T}} G \nabla \varepsilon + z^{\mathrm{T}} f R_f^{-1} \varepsilon_{m\mathbf{Y}^*}^{\mathrm{T}} \lambda^2 R \varepsilon_{m\mathbf{Y}^*} - z^{\mathrm{T}} f R_f^{-1} \varepsilon_{m\hat{\mathbf{Y}}}^{\mathrm{T}} \lambda^2 R \varepsilon_{m\hat{\mathbf{Y}}} \\
&+ z^{\mathrm{T}} f R_f^{-1} x^{\mathrm{T}} P^{\mathrm{T}} g \lambda \varepsilon_m + z^{\mathrm{T}} f R_f^{-1} W^{*\mathrm{T}} \bar{\nabla} \Phi g \lambda \varepsilon_m, \qquad (6.47)
\end{aligned}
$$

$$
\begin{aligned}
T_{2c} = &\frac{1}{2} z^{\mathrm{T}} G P x R_f^{-1} (f - g u^*)^{\mathrm{T}} \nabla \varepsilon + \frac{1}{8} z^{\mathrm{T}} G P x R_f^{-1} \nabla \varepsilon^{\mathrm{T}} G \nabla \varepsilon \\
&- \frac{1}{2} z^{\mathrm{T}} G P x R_f^{-1} \varepsilon_{m\mathbf{Y}^*}^{\mathrm{T}} \lambda^2 R \varepsilon_{m\mathbf{Y}^*} + \frac{1}{2} z^{\mathrm{T}} G P x R_f^{-1} \varepsilon_{m\hat{\mathbf{Y}}}^{\mathrm{T}} \lambda^2 R \varepsilon_{m\hat{\mathbf{Y}}} \\
&- \frac{1}{2} z^{\mathrm{T}} G P x R_f^{-1} x^{\mathrm{T}} P^{\mathrm{T}} g \lambda \varepsilon_m - \frac{1}{2} z^{\mathrm{T}} G P x R_f^{-1} W^{*\mathrm{T}} \bar{\nabla} \Phi g \lambda \varepsilon_m, \qquad (6.48)
\end{aligned}
$$

$$
\begin{aligned}
T_{2d} = &\frac{1}{2} z^{\mathrm{T}} G \bar{\nabla} \Phi^{\mathrm{T}} W^* R_f^{-1} (f - g u^*)^{\mathrm{T}} \nabla \varepsilon + \frac{1}{8} z^{\mathrm{T}} G \bar{\nabla} \Phi^{\mathrm{T}} W^* R_f^{-1} \nabla \varepsilon^{\mathrm{T}} G \nabla \varepsilon \\
&- \frac{1}{2} z^{\mathrm{T}} G \bar{\nabla} \Phi^{\mathrm{T}} W^* R_f^{-1} \varepsilon_{m\mathbf{Y}^*}^{\mathrm{T}} \lambda^2 R \varepsilon_{m\mathbf{Y}^*} + \frac{1}{2} z^{\mathrm{T}} G \bar{\nabla} \Phi^{\mathrm{T}} W^* R_f^{-1} \varepsilon_{m\hat{\mathbf{Y}}}^{\mathrm{T}} \lambda^2 R \varepsilon_{m\hat{\mathbf{Y}}} \\
&- \frac{1}{2} z^{\mathrm{T}} G \bar{\nabla} \Phi^{\mathrm{T}} W^* R_f^{-1} x^{\mathrm{T}} P^{\mathrm{T}} g \lambda \varepsilon_m - \frac{1}{2} z^{\mathrm{T}} G \bar{\nabla} \Phi^{\mathrm{T}} W^* R_f^{-1} W^{*\mathrm{T}} \bar{\nabla} \Phi g \lambda \varepsilon_m, \qquad (6.49)
\end{aligned}
$$

$$
\begin{aligned}
T_{2e} = &z^{\mathrm{T}} g \lambda \varepsilon_{m\hat{\mathbf{Y}}} R_f^{-1} (f - g u^*)^{\mathrm{T}} \nabla \varepsilon + \frac{1}{4} z^{\mathrm{T}} g \lambda \varepsilon_{m\hat{\mathbf{Y}}} R_f^{-1} \nabla \varepsilon^{\mathrm{T}} G \nabla \varepsilon \\
&- z^{\mathrm{T}} g \lambda \varepsilon_{m\hat{\mathbf{Y}}} R_f^{-1} \varepsilon_{m\mathbf{Y}^*}^{\mathrm{T}} \lambda^2 R \varepsilon_{m\mathbf{Y}^*} + z^{\mathrm{T}} g \lambda \varepsilon_{m\hat{\mathbf{Y}}} R_f^{-1} \varepsilon_{m\hat{\mathbf{Y}}}^{\mathrm{T}} \lambda^2 R \varepsilon_{m\hat{\mathbf{Y}}} \\
&- z^{\mathrm{T}} g \lambda \varepsilon_{m\hat{\mathbf{Y}}} R_f^{-1} x^{\mathrm{T}} P^{\mathrm{T}} g \lambda \varepsilon_m - z^{\mathrm{T}} g \lambda \varepsilon_{m\hat{\mathbf{Y}}} R_f^{-1} W^{*\mathrm{T}} \bar{\nabla} \Phi g \lambda \varepsilon_m, \qquad (6.50)
\end{aligned}
$$

$$
\begin{aligned}
T_3 = &- \frac{1}{2} x^{\mathrm{T}} P^{\mathrm{T}} G P x + x^{\mathrm{T}} P^{\mathrm{T}} f - x^{\mathrm{T}} P^{\mathrm{T}} G \bar{\nabla} \Phi^{\mathrm{T}} W^* - x^{\mathrm{T}} P^{\mathrm{T}} g \lambda \varepsilon_{m\hat{\mathbf{Y}}} \\
&+ W^{*\mathrm{T}} \bar{\nabla} \Phi f - W^{*\mathrm{T}} \bar{\nabla} \Phi g \lambda \varepsilon_{m\hat{\mathbf{Y}}}, \qquad (6.51)
\end{aligned}
$$

$$
\begin{aligned}
T_4 = &- \frac{1}{8} z^{\mathrm{T}} G P x R_f^{-1} x^{\mathrm{T}} P^{\mathrm{T}} G z + x^{\mathrm{T}} P^{\mathrm{T}} G z - z^{\mathrm{T}} f + z^{\mathrm{T}} g \lambda \varepsilon_{m\hat{\mathbf{Y}}} \\
&+ W^{*\mathrm{T}} \bar{\nabla} \Phi G z - z^{\mathrm{T}} f R_f^{-1} R_r \lambda^2 \varepsilon_{e\mathrm{Y}} + \frac{1}{2} z^{\mathrm{T}} G P x R_f^{-1} R_r \lambda^2 \varepsilon_{e\mathrm{Y}} \\
&+ \frac{1}{2} z^{\mathrm{T}} G \bar{\nabla} \Phi^{\mathrm{T}} W^* R_f^{-1} R_r \lambda^2 \varepsilon_{e\mathrm{Y}} + z^{\mathrm{T}} g \lambda \varepsilon_{m\hat{\mathbf{Y}}} R_f^{-1} R_r \lambda^2 \varepsilon_{e\mathrm{Y}}, \qquad (6.52)
\end{aligned}
$$

$$
T_5 = - \frac{1}{16} z^{\mathrm{T}} G z R_f^{-1} z^{\mathrm{T}} G z - \frac{1}{2} z^{\mathrm{T}} G z R_f^{-1} R_r \lambda^2 \varepsilon_{e\mathrm{Y}}, \qquad (6.53)
$$

$$
\begin{aligned}
T_6 = &- z^{\mathrm{T}} f R_f^{-1} z^{\mathrm{T}} f - \frac{1}{4} z^{\mathrm{T}} G \bar{\nabla} \Phi^{\mathrm{T}} W^* R_f^{-1} W^{*\mathrm{T}} \bar{\nabla} \Phi G z \\
&- z^{\mathrm{T}} g \lambda \varepsilon_{m\hat{\mathbf{Y}}} R_f^{-1} z^{\mathrm{T}} g \lambda \varepsilon_{m\hat{\mathbf{Y}}} - \frac{1}{8} z^{\mathrm{T}} G z R_f^{-1} \nabla \varepsilon^{\mathrm{T}} G \nabla \varepsilon \\
&- \frac{1}{2} z^{\mathrm{T}} G z - \frac{1}{2} W^{*\mathrm{T}} \bar{\nabla} \Phi G \bar{\nabla} \Phi^{\mathrm{T}} W^*, \qquad (6.54)
\end{aligned}
$$

where $G = g R^{-1} g^{\mathrm{T}}$ and $z = \bar{\nabla} \Phi^{\mathrm{T}} \tilde{W}$.

Now we introduce bounds. As $\boldsymbol{G} = \boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}$, the rank of $\boldsymbol{G}$ is

$$\operatorname{rank}(\boldsymbol{G}) = \operatorname{rank}(\boldsymbol{g}) < n_x.$$

It follows that there exist kernels

$$\ker(\boldsymbol{G}) = \{\boldsymbol{r} \in \mathbb{R}^{n_x} \mid \boldsymbol{G}\boldsymbol{r} = 0\},$$
$$\ker(\boldsymbol{G}\boldsymbol{P}) = \{\boldsymbol{r} \in \mathbb{R}^{n_x} \mid \boldsymbol{G}\boldsymbol{P}\boldsymbol{r} = 0\}.$$

For nonlinear systems as in Eq. (6.1), since vectors $\boldsymbol{x}$ and $\boldsymbol{z}$ are explicitly governed by Eq. (6.1) instead of being random, the existence of $\boldsymbol{x} = \ker(\boldsymbol{G}\boldsymbol{P})$ and $\boldsymbol{z} = \ker(\boldsymbol{G})$ as well as corresponding effects to the system are rendered negligible. Accordingly, we focus on cases where $\boldsymbol{x} \neq \ker(\boldsymbol{G}\boldsymbol{P})$ and $\boldsymbol{z} \neq \ker(\boldsymbol{G})$.

Given that $\boldsymbol{G}$ is positive-definite and symmetric, under Assumptions 4.2 and 4.3, there are

$$\mathbf{z}^{\mathrm{T}}\mathbf{G}\mathbf{z}R_f^{-1}\mathbf{z}^{\mathrm{T}}\mathbf{G}\mathbf{z} \geq b_{m1}\left\|\tilde{\mathbf{W}}\right\|^4\|\mathbf{x}\|^4,$$
$$\boldsymbol{z}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x}R_f^{-1}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{z} \geq b_{m2}\|\boldsymbol{P}\|^2\left\|\tilde{\boldsymbol{W}}\right\|^2\|\boldsymbol{x}\|^4,$$
$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x} \geq b_{m3}\|\boldsymbol{P}\|^2\|\boldsymbol{x}\|^2,$$

where constants $b_{m1} \in \mathbb{R}^+$, $b_{m2} \in \mathbb{R}^+$ and $b_{m3} \in \mathbb{R}^+$. Also, there is $\|\boldsymbol{G}\| \leq b_G$ for constant $b_G \in \mathbb{R}^+$.

Under Assumptions 4.1 and 4.2, we also have $\|\boldsymbol{f} + \boldsymbol{g}\boldsymbol{u}^*\| \leq b_{\dot{x}}\|\boldsymbol{x}\|$ for constant $b_{\dot{x}} \in \mathbb{R}^+$. Let $b_{Rf} = \left\|R_f^{-1}\right\|$ and $b_\lambda = \|\boldsymbol{\lambda}\|$. Then for T_1,

$$\begin{aligned} T_1 \leq & -\frac{1}{16}\Big(b_{m1}\left\|\tilde{\boldsymbol{W}}\right\| - 6b_G^2 b_\varepsilon b_{Rf} - 12b_G b_{Rf} b_g b_\lambda b_{\varepsilon m\hat{\mathrm{Y}}} \\ & -12b_G b_{Rf} b_g b_\lambda b_{\varepsilon m\mathrm{Y}^*} - 12b_{\dot{x}} b_{Rf} b_G\Big)\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|^3 \\ = & -\frac{1}{16}\left(b_{m1}\left\|\tilde{\boldsymbol{W}}\right\| - \eta_1\right)\|\boldsymbol{x}\|^4\left\|\tilde{\boldsymbol{W}}\right\|^3, \end{aligned} \tag{6.55}$$

where

$$\eta_1 = 6b_G^2 b_\varepsilon b_{Rf} + 12b_G b_{Rf} b_g b_\lambda b_{\varepsilon m\hat{\mathrm{Y}}} + 12b_G b_{Rf} b_g b_\lambda b_{\varepsilon m\mathrm{Y}^*} + 12b_{\dot{x}} b_{Rf} b_G.$$

On satisfying Assumption 4.4, it is clear that $T_1 \leq 0$ if

$$b_{m1}\left\|\tilde{\boldsymbol{W}}\right\| - \eta_1 \geq 0,$$

which requires

$$\left\|\tilde{\boldsymbol{W}}\right\| \geq \frac{\eta_1}{b_{m1}}. \tag{6.56}$$

Note that T_2 has several terms as shown in Eq. (6.45). Let $b_R = \|\boldsymbol{R}\|$. Then for T_{2a} we have

$$\begin{aligned}
T_{2a} \leq & -\frac{1}{8}\Big(b_{m2}\|\boldsymbol{P}\|^2 - 4b_G^2 b_{Rf} b_\phi \|\boldsymbol{P}\| \left\|\boldsymbol{W}^*\right\| \\
& - 8b_G b_{Rf} b_g b_\lambda b_{\varepsilon m\hat{\mathbf{Y}}} \|\boldsymbol{P}\| - 4b_G b_{Rf} b_g b_\lambda b_{\varepsilon m\mathbf{Y}^*} \|\boldsymbol{P}\| \\
& - 8b_g b_\lambda b_{\varepsilon m\hat{\mathbf{Y}}} \left\|\boldsymbol{W}^*\right\| b_G b_{Rf} b_\phi - 4b_G b_{Rf} b_\phi b_g b_\lambda b_{\varepsilon m\mathbf{Y}^*} \left\|\boldsymbol{W}^*\right\| \\
& - 16b_f b_{Rf} b_g b_\lambda b_{\varepsilon m\hat{\mathbf{Y}}} - 4b_G b_{Rf} b_{\dot{x}} b_\varepsilon - 4b_G b_{Rf} b_{\varepsilon m\mathbf{Y}^*}^2 b_\lambda^2 b_R \\
& - 4b_G b_{Rf} b_{\varepsilon m\hat{\mathbf{Y}}}^2 b_\lambda^2 b_R\Big) \left\|\tilde{\boldsymbol{W}}\right\|^2 \|\boldsymbol{x}\|^4 \\
= & -\frac{1}{8}\left(b_{m2}\|\boldsymbol{P}\|^2 - \eta_2 \|\boldsymbol{P}\| - \eta_3\right) \left\|\tilde{\boldsymbol{W}}\right\|^2 \|\boldsymbol{x}\|^4,
\end{aligned} \tag{6.57}$$

where

$$\begin{aligned}
\eta_2 &= 4b_G^2 b_{Rf} b_\phi \left\|\boldsymbol{W}^*\right\| + 8b_G b_{Rf} b_g b_\lambda b_{\varepsilon m\hat{\mathbf{Y}}} + 4b_G b_{Rf} b_g b_\lambda b_{\varepsilon m\mathbf{Y}^*}, \\
\eta_3 &= 8b_g b_\lambda b_{\varepsilon m\hat{\mathbf{Y}}} b_G b_{Rf} b_\phi \left\|\boldsymbol{W}^*\right\| + 4b_G b_{Rf} b_\phi b_g b_\lambda b_{\varepsilon m\mathbf{Y}^*} \left\|\boldsymbol{W}^*\right\| \\
&\quad + 16b_f b_{Rf} b_g b_\lambda b_{\varepsilon m\hat{\mathbf{Y}}} + 4b_G b_{Rf} b_{\dot{x}} b_\varepsilon + 4b_G b_{Rf} b_{\varepsilon m\mathbf{Y}^*}^2 b_\lambda^2 b_R \\
&\quad + 4b_G b_{Rf} b_{\varepsilon m\hat{\mathbf{Y}}}^2 b_\lambda^2 b_R.
\end{aligned}$$

For T_{2b} there is

$$\begin{aligned}
T_{2b} \leq & \big(b_f b_{Rf} b_G \|\boldsymbol{P}\| + b_f b_{Rf} b_g b_\lambda b_{\varepsilon m\mathbf{Y}^*} \|\boldsymbol{P}\| + b_f b_{Rf} b_{\dot{x}} b_\varepsilon \\
& + \frac{1}{4} b_f b_{Rf} b_\varepsilon^2 b_G + b_f b_{Rf} b_{\varepsilon m\mathbf{Y}^*}^2 b_\lambda^2 b_R + b_f b_{Rf} b_{\varepsilon m\hat{\mathbf{Y}}}^2 b_\lambda^2 b_R \\
& + b_f b_{Rf} b_\phi b_G \left\|\boldsymbol{W}^*\right\| + b_f b_{Rf} b_\phi b_g b_\lambda b_{\varepsilon m\mathbf{Y}^*} \left\|\boldsymbol{W}^*\right\|\big) \left\|\tilde{\boldsymbol{W}}\right\| \|\boldsymbol{x}\|^4 \\
= & \left(\eta_4 \|\boldsymbol{P}\| + \eta_5\right) \left\|\tilde{\boldsymbol{W}}\right\| \|\boldsymbol{x}\|^4,
\end{aligned} \tag{6.58}$$

where

$$\begin{aligned}
\eta_4 &= b_f b_{Rf} b_G + b_f b_{Rf} b_g b_\lambda b_{\varepsilon m\mathbf{Y}^*}, \\
\eta_5 &= b_f b_{Rf} b_{\dot{x}} b_\varepsilon + \frac{1}{4} b_f b_{Rf} b_\varepsilon^2 b_G + b_f b_{Rf} b_{\varepsilon m\mathbf{Y}^*}^2 b_\lambda^2 b_R \\
&\quad + b_f b_{Rf} b_{\varepsilon m\hat{\mathbf{Y}}}^2 b_\lambda^2 b_R + b_f b_{Rf} b_\phi b_G \left\|\boldsymbol{W}^*\right\| \\
&\quad + b_f b_{Rf} b_\phi b_g b_\lambda b_{\varepsilon m\mathbf{Y}^*} \left\|\boldsymbol{W}^*\right\|.
\end{aligned}$$

For T_{2c},

$$\begin{aligned} T_{2c} \leq & \left(\frac{1}{2} b_G b_{Rf} b_g b_\lambda b_{\varepsilon m \mathbf{Y}^*} \|\boldsymbol{P}\|^2 + \frac{1}{2} b_G b_{Rf} b_{\dot{x}} b_\varepsilon \|\boldsymbol{P}\| + \frac{1}{8} b_G^2 b_{Rf} b_\varepsilon^2 \|\boldsymbol{P}\| \right. \\ & + \frac{1}{2} b_G b_{Rf} b_{\varepsilon m \mathbf{Y}^*}^2 b_\lambda^2 b_R \|\boldsymbol{P}\| + \frac{1}{2} b_G b_{Rf} b_{\varepsilon m \hat{\mathbf{Y}}}^2 b_\lambda^2 b_R \|\boldsymbol{P}\| \\ & \left. + \frac{1}{2} b_G b_{Rf} b_\phi b_g b_\lambda b_{\varepsilon m \mathbf{Y}^*} \left\| \boldsymbol{W}^* \right\| \|\boldsymbol{P}\| \right) \left\| \tilde{\boldsymbol{W}} \right\| \|\boldsymbol{x}\|^4 \\ = & \left(\eta_6 \|\boldsymbol{P}\|^2 + \eta_7 \|\boldsymbol{P}\| \right) \left\| \tilde{\boldsymbol{W}} \right\| \|\boldsymbol{x}\|^4, \end{aligned} \tag{6.59}$$

where

$$\begin{aligned} \eta_6 &= \frac{1}{2} b_G b_{Rf} b_g b_\lambda b_{\varepsilon m \mathbf{Y}^*}, \\ \eta_7 &= \frac{1}{2} b_G b_{Rf} b_{\dot{x}} b_\varepsilon + \frac{1}{8} b_G^2 b_{Rf} b_\varepsilon^2 + \frac{1}{2} b_G b_{Rf} b_{\varepsilon m \mathbf{Y}^*}^2 b_\lambda^2 b_R \\ &\quad + \frac{1}{2} b_G b_{Rf} b_{\varepsilon m \hat{\mathbf{Y}}}^2 b_\lambda^2 b_R + \frac{1}{2} b_G b_{Rf} b_\phi b_g b_\lambda b_{\varepsilon m \mathbf{Y}^*} \left\| \boldsymbol{W}^* \right\|. \end{aligned}$$

For T_{2d},

$$\begin{aligned} T_{2d} \leq & \left(\frac{1}{2} b_G b_\phi b_{Rf} b_g b_\lambda b_{\varepsilon m \mathrm{Y}^*} \|\boldsymbol{W}^*\| \|\boldsymbol{P}\| + \frac{1}{2} b_G b_\phi b_{Rf} b_{\dot{x}} b_\varepsilon \|\boldsymbol{W}^*\| \right. \\ & + \frac{1}{8} b_G^2 b_\phi b_{Rf} b_\varepsilon^2 \|\boldsymbol{W}^*\| + \frac{1}{2} b_G b_\phi b_{Rf} b_{\varepsilon m \mathrm{Y}^*}^2 b_\lambda^2 b_R \|\boldsymbol{W}^*\| \\ & \left. + \frac{1}{2} b_G b_\phi b_{Rf} b_{\varepsilon m \hat{\mathrm{Y}}}^2 b_\lambda^2 b_R \|\boldsymbol{W}^*\| + \frac{1}{2} b_G b_\phi^2 b_{Rf} b_g b_\lambda b_{\varepsilon m \mathrm{Y}^*} \|\boldsymbol{W}^*\|^2 \right) \left\| \tilde{\boldsymbol{W}} \right\| \|\boldsymbol{x}\|^4 \\ = & \left(\eta_8 \|\boldsymbol{P}\| + \eta_9 \right) \left\| \tilde{\boldsymbol{W}} \right\| \|\boldsymbol{x}\|^4, \end{aligned} \tag{6.60}$$

where

$$\begin{aligned} \eta_8 &= \frac{1}{2} b_G b_\phi b_{Rf} b_g b_\lambda b_{\varepsilon m \mathrm{Y}^*} \|\boldsymbol{W}^*\|, \\ \eta_9 &= \frac{1}{2} b_G b_\phi b_{Rf} b_{\dot{x}} b_\varepsilon \|\boldsymbol{W}^*\| + \frac{1}{8} b_G^2 b_\phi b_{Rf} b_\varepsilon^2 \|\boldsymbol{W}^*\| + \frac{1}{2} b_G b_\phi b_{Rf} b_{\varepsilon m \mathrm{Y}^*}^2 b_\lambda^2 b_R \|\boldsymbol{W}^*\| \\ &\quad + \frac{1}{2} b_G b_\phi b_{Rf} b_{\varepsilon m \hat{\mathrm{Y}}}^2 b_\lambda^2 b_R \|\boldsymbol{W}^*\| + \frac{1}{2} b_G b_\phi^2 b_{Rf} b_g b_\lambda b_{\varepsilon m \mathrm{Y}^*} \|\boldsymbol{W}^*\|^2. \end{aligned}$$

For T_{2e},

$$\begin{aligned} T_{2e} \leq & \left(b_g^2 b_\lambda^2 b_{\varepsilon m \hat{\mathrm{Y}}} b_{Rf} b_{\varepsilon m \mathrm{Y}^*} \|\boldsymbol{P}\| + b_g b_\lambda b_{\varepsilon m \hat{\mathrm{Y}}} b_{Rf} b_{\dot{x}} b_\varepsilon \right. \\ & + \frac{1}{4} b_g b_\lambda b_{\varepsilon m \hat{\mathrm{Y}}} b_{Rf} b_G b_\varepsilon^2 + b_g b_{\varepsilon m \hat{\mathrm{Y}}} b_{Rf} b_{\varepsilon m \mathrm{Y}^*}^2 b_\lambda^3 b_R + b_g b_{Rf} b_{\varepsilon m \hat{\mathrm{Y}}}^3 b_\lambda^3 b_R \\ & \left. + b_g^2 b_\lambda^2 b_{\varepsilon m \hat{\mathrm{Y}}} b_{Rf} b_\phi b_{\varepsilon m \mathrm{Y}^*} \|\boldsymbol{W}^*\| \right) \left\| \tilde{\boldsymbol{W}} \right\| \|\boldsymbol{x}\|^4 \\ = & \left(\eta_{10} \|\boldsymbol{P}\| + \eta_{11} \right) \left\| \tilde{\boldsymbol{W}} \right\| \|\boldsymbol{x}\|^4, \end{aligned} \tag{6.61}$$

where

$$\begin{aligned}
\eta_{10} &= b_g^2 b_\lambda^2 b_{\varepsilon m\hat{\Upsilon}} b_{Rf} b_{\varepsilon m\Upsilon^*}, \\
\eta_{11} &= b_g b_\lambda b_{\varepsilon m\hat{\Upsilon}} b_{Rf} b_{\dot{x}} b_\varepsilon + \frac{1}{4} b_g b_\lambda b_{\varepsilon m\hat{\Upsilon}} b_{Rf} b_G b_\varepsilon^2 + b_g b_{\varepsilon m\hat{\Upsilon}} b_{Rf} b_{\varepsilon m\Upsilon^*}^2 b_\lambda^3 b_R \\
&\quad + b_g b_{Rf} b_{\varepsilon m\hat{\Upsilon}}^3 b_\lambda^3 b_R + b_g^2 b_\lambda^2 b_{\varepsilon m\hat{\Upsilon}} b_{Rf} b_\phi b_{\varepsilon m\Upsilon^*} \left\| \boldsymbol{W}^* \right\|.
\end{aligned}$$

Assembling T_{2a}, T_{2b}, T_{2c}, T_{2d}, and T_{2e} together gives

$$\begin{aligned}
T_2 \leq & -\frac{1}{8}\left(b_{m2}\|\boldsymbol{P}\|^2 - \eta_2\|\boldsymbol{P}\| - \eta_3\right)\left\|\tilde{\boldsymbol{W}}\right\|^2\|\boldsymbol{x}\|^4 \\
& + \left[\eta_6\|\boldsymbol{P}\|^2 + (\eta_4+\eta_7+\eta_8+\eta_{10})\|\boldsymbol{P}\| + (\eta_5+\eta_9+\eta_{11})\right]\left\|\tilde{\boldsymbol{W}}\right\|\|\boldsymbol{x}\|^4 \\
= & -\frac{1}{8}\Big[\left(b_{m2}\|\boldsymbol{P}\|^2 - \eta_2\|\boldsymbol{P}\| - \eta_3\right)\left\|\tilde{\boldsymbol{W}}\right\| - 8\eta_6\|\boldsymbol{P}\|^2 \\
& -8(\eta_4+\eta_7+\eta_8+\eta_{10})\|\boldsymbol{P}\| - 8(\eta_5+\eta_9+\eta_{11})\Big]\left\|\tilde{\boldsymbol{W}}\right\|\|\boldsymbol{x}\|^4 \\
= & -\frac{1}{8}\left(\xi_1\left\|\tilde{\boldsymbol{W}}\right\| - \xi_2\right)\left\|\tilde{\boldsymbol{W}}\right\|\|\boldsymbol{x}\|^4,
\end{aligned} \tag{6.62}$$

where

$$\begin{aligned}
\xi_1 &= b_{m2}\|\boldsymbol{P}\|^2 - \eta_2\|\boldsymbol{P}\| - \eta_3, \\
\xi_2 &= 8\eta_6\|\boldsymbol{P}\|^2 + 8(\eta_4+\eta_7+\eta_8+\eta_{10})\|\boldsymbol{P}\| + 8(\eta_5+\eta_9+\eta_{11}).
\end{aligned} \tag{6.63}$$

Under Assumption 4.4, it is clear that $T_2 \leq 0$ if $\left(\xi_1\left\|\tilde{\boldsymbol{W}}\right\| - \xi_2\right) \geq 0$, that is,

$$\left\|\tilde{\boldsymbol{W}}\right\| \geq \frac{\xi_2}{\xi_1}, \text{ for } \xi_1 > 0 \text{ and } \xi_2 \geq 0. \tag{6.64}$$

In order to ensure $\xi_1 > 0$, it can be derived from Eq. (6.63) that the condition to be satisfied is

$$\|\boldsymbol{P}\| > \frac{\eta_2 + \sqrt{\eta_2^2 + 4b_{m2}\eta_3}}{2b_{m2}} \triangleq b_{p1}. \tag{6.65}$$

Now we look at T_3, which is bounded as

$$\begin{aligned}
T_3 \leq & \left(-\frac{1}{2}b_{m3}\|\boldsymbol{P}\|^2 + b_f\|\boldsymbol{P}\| + b_g b_\lambda b_{\varepsilon m\hat{\Upsilon}}\|\boldsymbol{P}\| + b_G b_\phi\|\boldsymbol{W}^*\|\|\boldsymbol{P}\|\right. \\
& \left. + b_\phi b_f\left\|\boldsymbol{W}^*\right\| + b_\phi b_g b_\lambda b_{\varepsilon m\hat{\Upsilon}}\left\|\boldsymbol{W}^*\right\|\right)\|\boldsymbol{x}\|^2 \\
= & -\left(\frac{1}{2}b_{m3}\|\boldsymbol{P}\|^2 - \eta_{12}\|\boldsymbol{P}\| - \eta_{13}\right)\|\boldsymbol{x}\|^2,
\end{aligned} \tag{6.66}$$

where

$$\eta_{12} = b_f + b_g b_\lambda b_{\varepsilon m\hat{\Upsilon}} + b_G b_\phi \left\| \boldsymbol{W}^* \right\|,$$
$$\eta_{13} = b_\phi b_f \left\| \boldsymbol{W}^* \right\| + b_\phi b_g b_\lambda b_{\varepsilon m\hat{\Upsilon}} \left\| \boldsymbol{W}^* \right\|.$$

According to Eq. (6.66) and under Assumption 4.4, to guarantee $T_3 \leq 0$ requires

$$\frac{1}{2} b_{m3} \|\boldsymbol{P}\|^2 - \eta_{12} \|\boldsymbol{P}\| - \eta_{13} \geq 0,$$

which can be satisfied by

$$\|\boldsymbol{P}\| \geq \frac{\eta_{12} + \sqrt{\eta_{12}^2 + 2 b_{m3} \eta_{13}}}{b_{m3}} \triangleq b_{p2}. \tag{6.67}$$

In regard to T_4, let $b_{Rr} = \|\boldsymbol{R}_r\|$, then

$$\begin{aligned} T_4 \leq & -\frac{1}{8} \Big(b_{m2} \|\boldsymbol{P}\|^2 \left\|\tilde{\boldsymbol{W}}\right\| \|\boldsymbol{x}\|^2 - 8 b_G \|\boldsymbol{P}\| - 4 b_G b_{Rf} b_{Rr} b_\lambda^2 b_{\varepsilon e\Upsilon} \|\boldsymbol{P}\| \\ & - 8 b_\phi b_G \left\| \boldsymbol{W}^* \right\| - 4 b_G b_\phi b_{Rf} b_{Rr} b_\lambda^2 b_{\varepsilon e\Upsilon} \|\boldsymbol{W}^*\| - 8 b_f - 8 b_g b_\lambda b_{\varepsilon m\hat{\Upsilon}} \\ & - 8 b_f b_{Rf} b_{Rr} b_\lambda^2 b_{\varepsilon e\Upsilon} - 8 b_g b_{\varepsilon m\hat{\Upsilon}} b_{Rf} b_{Rr} b_\lambda^3 b_{\varepsilon e\Upsilon} \Big) \left\|\tilde{\boldsymbol{W}}\right\| \|\boldsymbol{x}\|^2 \\ = & -\frac{1}{8} \left[b_{m2} \|\boldsymbol{P}\|^2 \left\|\tilde{\boldsymbol{W}}\right\| \|\boldsymbol{x}\|^2 - \eta_{14} \|\boldsymbol{P}\| - \eta_{15} \right] \left\|\tilde{\boldsymbol{W}}\right\| \|\boldsymbol{x}\|^2, \end{aligned} \tag{6.68}$$

where

$$\begin{aligned} \eta_{14} &= 8 b_G + 4 b_G b_{Rf} b_{Rr} b_\lambda^2 b_{\varepsilon e\Upsilon}, \\ \eta_{15} &= 8 b_\phi b_G \left\| \boldsymbol{W}^* \right\| + 4 b_G b_\phi b_{Rf} b_{Rr} b_\lambda^2 b_{\varepsilon e\Upsilon} \|\boldsymbol{W}^*\| + 8 b_f + 8 b_g b_\lambda b_{\varepsilon m\hat{\Upsilon}} \\ &\quad + 8 b_f b_{Rf} b_{Rr} b_\lambda^2 b_{\varepsilon e\Upsilon} + 8 b_g b_{\varepsilon m\hat{\Upsilon}} b_{Rf} b_{Rr} b_\lambda^3 b_{\varepsilon e\Upsilon}. \end{aligned}$$

As can been seen from Eq. (6.68), with PE condition applied, $T_4 \leq 0$ requires

$$b_{m2} \|\boldsymbol{P}\|^2 \left\|\tilde{\boldsymbol{W}}\right\| \|\boldsymbol{x}\|^2 - \eta_{14} \|\boldsymbol{P}\| - \eta_{15} \geq 0.$$

That is,

$$\left\|\tilde{\boldsymbol{W}}\right\| \|\boldsymbol{x}\|^2 \geq \frac{\eta_{14} \|\boldsymbol{P}\| + \eta_{15}}{b_{m2} \|\boldsymbol{P}\|^2}. \tag{6.69}$$

Regarding T_5,

$$\begin{aligned} T_5 &\leq \left(-\frac{1}{16} b_{m1} \left\|\tilde{\boldsymbol{W}}\right\|^2 \|\boldsymbol{x}\|^2 + \frac{1}{2} b_G b_{Rf} b_{Rr} b_\lambda^2 b_{\varepsilon e\Upsilon} \right) \left\|\tilde{\boldsymbol{W}}\right\|^2 \|\boldsymbol{x}\|^2 \\ &= -\frac{1}{16} b_{m1} \left(\left\|\tilde{\boldsymbol{W}}\right\|^2 \|\boldsymbol{x}\|^2 - \frac{8 b_G b_{Rf} b_{Rr} b_\lambda^2 b_{\varepsilon e\Upsilon}}{b_{m1}} \right) \left\|\tilde{\boldsymbol{W}}\right\|^2 \|\boldsymbol{x}\|^2. \end{aligned} \tag{6.70}$$

It is straightforward to see from Eq. (6.70) that $T_5 \leq 0$ under the PE condition as long as

$$\left\|\tilde{\boldsymbol{W}}\right\| \|\boldsymbol{x}\| \geq \sqrt{\frac{8 b_G b_{Rf} b_{Rr} b_\lambda^2 b_{\varepsilon e Y}}{b_{m1}}}. \tag{6.71}$$

As for T_6, it is clear from Eq. (6.54) that $T_6 \leq 0$. Therefore, from Eq. (6.43), we have $\dot{\mathcal{L}} < 0$ upon conditions in Eqs. (6.56), (6.64), (6.65), (6.67), (6.69) and (6.71) being satisfied.

Recall from Eqs. (6.69) and (6.71) that a bound exists for $\left\|\tilde{\boldsymbol{W}}\right\|$, given $\boldsymbol{x}$ is PE under Assumption 4.4. Denote this bound as $b_w \in \mathbb{R}^+$. Together with Eqs. (6.56), (6.64), (6.65), and (6.67), we can reach a conclusion that $\tilde{\boldsymbol{W}}$ is UUB by a bound $b_{\tilde{W}} \triangleq \max\left(\frac{\eta_1}{b_{m1}}, \frac{\xi_2}{\xi_1}, b_w\right)$ during online tuning if the parameter matrix $\boldsymbol{P}$ is selected to be

$$\|\boldsymbol{P}\| > \max(b_{p1}, b_{p2}) \triangleq b_{mP}. \tag{6.72}$$

The proof is now completed. □

In the next, we further examine the closed-loop stability with respect to system states $\boldsymbol{x}$ regardless of whether Assumption 4.4 is applied or not.

Theorem 6.2. *For nonlinear systems as in Eq. (6.1), let the value function be approximated by Eq. (6.8), the control be in the form of Eq. (6.12), and the NN weights $\hat{\boldsymbol{W}}$ be dynamically adjusted using the EKF as in Eqs. (6.15) to (6.18) with estimation error of $\tilde{\boldsymbol{W}} = \boldsymbol{W}^* - \hat{\boldsymbol{W}}$. On the basis of Assumptions 4.1 to 4.3, the closed-loop system is asymptotically stable in terms of the states $\boldsymbol{x}$ about the origin, if the parameter matrix $\boldsymbol{P}$ in Eq. (6.8) is selected to be $\|\boldsymbol{P}\| > b_{mP}$ for a finite constant $b_{mP} \in \mathbb{R}^+$.*

Proof. Consider

$$\mathcal{L}_V = \hat{V} \tag{6.73}$$

as the Lyapunov candidate, the time derivative of which has been given by Eq. (6.24).

With $\boldsymbol{G} = \boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^{\mathrm{T}}$, Eq. (6.24) can be rewritten as

$$\dot{\mathcal{L}}_V = \dot{\mathcal{L}}_{V1} + \dot{\mathcal{L}}_{V2}, \tag{6.74}$$

where

$$\begin{aligned}\dot{\mathcal{L}}_{V1} = &-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}\boldsymbol{x} + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{f} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^{*} \\ &+ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\Upsilon}} + \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} \\ &+ \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}} - \boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\Upsilon}} \\ &- \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{f} + \tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{g}\boldsymbol{\lambda}\boldsymbol{\varepsilon}_{m\hat{\Upsilon}},\end{aligned} \tag{6.75}$$

$$\dot{\mathcal{L}}_{V2} = -\frac{1}{2}\boldsymbol{W}^{*\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{W}^{*} - \frac{1}{2}\tilde{\boldsymbol{W}}^{\mathrm{T}}\bar{\nabla}\boldsymbol{\Phi}\boldsymbol{G}\bar{\nabla}\boldsymbol{\Phi}^{\mathrm{T}}\tilde{\boldsymbol{W}}. \tag{6.76}$$

If Assumption 4.4 holds, then $\tilde{\boldsymbol{W}}$ is UUB during online tuning, as shown in Theorem 6.1. In the case of the PE condition being not satisfied, $\hat{\boldsymbol{W}}$ may not reach $\boldsymbol{W}^{*}$ but remains bounded as a result. Let $b_{\tilde{W}}$ denote the bound of $\tilde{\boldsymbol{W}}$ in both cases. Then Eq. (6.75) is bounded as

$$\begin{aligned}\dot{\mathcal{L}}_{V1} \leq &\left(-\frac{1}{2}b_{m3}b_{G}\|\boldsymbol{P}\|^{2} + b_{f}\|\boldsymbol{P}\| + b_{G}b_{\phi}\|\boldsymbol{W}^{*}\|\|\boldsymbol{P}\| + b_{G}b_{\phi}b_{\tilde{W}}\|\boldsymbol{P}\|\right. \\ &+ b_{g}b_{\lambda}b_{\varepsilon m\hat{\Upsilon}}\|\boldsymbol{P}\| + b_{\phi}b_{f}\left\|\boldsymbol{W}^{*}\right\| + b_{\phi}^{2}b_{G}b_{\tilde{W}}\|\boldsymbol{W}^{*}\| \\ &\left.+b_{\phi}b_{g}b_{\lambda}b_{\varepsilon m\hat{\Upsilon}}\left\|\boldsymbol{W}^{*}\right\| + b_{\tilde{W}}b_{\phi}b_{f} + b_{\tilde{W}}b_{\phi}b_{g}b_{\lambda}b_{\varepsilon m\hat{\Upsilon}}\right)\|\boldsymbol{x}\|^{2} \\ = &-\frac{1}{2}\left(b_{m3}b_{G}\|\boldsymbol{P}\|^{2} - \eta_{16}\|\boldsymbol{P}\| - \eta_{17}\right)\|\boldsymbol{x}\|^{2},\end{aligned} \tag{6.77}$$

where

$$\begin{aligned}\eta_{16} &= b_{G}b_{\phi}\|\boldsymbol{W}^{*}\| + b_{f} + b_{G}b_{\phi}b_{\tilde{W}} + b_{g}b_{\lambda}b_{\varepsilon m\hat{\Upsilon}}, \\ \eta_{17} &= b_{\phi}b_{f}\left\|\boldsymbol{W}^{*}\right\| + b_{\phi}^{2}b_{G}b_{\tilde{W}}\|\boldsymbol{W}^{*}\| + b_{\phi}b_{g}b_{\lambda}b_{\varepsilon m\hat{\Upsilon}}\left\|\boldsymbol{W}^{*}\right\| \\ &\quad + b_{\tilde{W}}b_{\phi}b_{f} + b_{\tilde{W}}b_{\phi}b_{g}b_{\lambda}b_{\varepsilon m\hat{\Upsilon}}.\end{aligned}$$

According to Eq. (6.77), when $\boldsymbol{x} \neq 0$, to make $\dot{\mathcal{L}}_{V1} \leq 0$, it requires

$$b_{m3}b_{G}\|\boldsymbol{P}\|^{2} - \eta_{16}\|\boldsymbol{P}\| - \eta_{17} \geq 0,$$

that is,

$$\|\boldsymbol{P}\| \geq \frac{\eta_{16} + \sqrt{\eta_{16}^{2} + 4b_{m3}b_{G}\eta_{17}}}{2b_{m3}b_{G}} \triangleq b_{p3}. \tag{6.78}$$

Since $\dot{\mathcal{L}}_{V2} \leq 0$ according to Eq. (6.76), by using Eq. (6.74) together with the result of Eq. (6.72) in Theorem 6.1 we have $\dot{\mathcal{L}}_V \leq 0$, if

$$\|\boldsymbol{P}\| > \max\left(b_{p1},\ b_{p2},\ b_{p3}\right) \triangleq b_{mP}. \tag{6.79}$$

Upon satisfying Eq. (6.79), it follows that the second-order time derivative of Eq. (6.73) is bounded since both $\boldsymbol{x}$ and $\tilde{\boldsymbol{W}}$ are bounded. Therefore, the closed-loop system is asymptotically stable in terms of the states $\boldsymbol{x}$ about the origin. The proof is now completed. □

Remark 6.4. Theorems 6.1 and 6.2 show that it is viable to introduce the nonquadratic control-input cost function to NOCOS involving the MVFA. The new control scheme considering CICs guarantees the closed-loop stability during NN online learning. Theorems 6.1 and 6.2 contribute new knowledge to the NOCOS theory framework and serve as major scientific contributions.

6.3.4 Value function approximation for cases with non-constant airspeed U_∞

The NOCOS under CICs proposed in Subsections 6.3.1 to 6.3.3 assumes constant U_∞. For cases with non-constant U_∞ that renders system (5.2) globally nonlinear, the parameter matrix $\boldsymbol{P}$ in Eq. (6.10) needs to be changed accordingly to satisfy the condition in Theorems 6.1 and 6.2, and can be determined and implemented through linear-parameter-varying technique using linear matrix inequalities (LMIs) as follows.

By linearising Eq. (5.2) at equilibrium and denoting $\boldsymbol{h}$ as the performance output, we have

$$\begin{bmatrix} \dot{\boldsymbol{x}} \\ \boldsymbol{y} \\ \boldsymbol{h} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_p(U_\infty) & \boldsymbol{I} & \boldsymbol{B}_p \\ \boldsymbol{C}_p & 0 & 0 \\ \boldsymbol{C}_h & 0 & \boldsymbol{E}_h \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{u} \\ \boldsymbol{w}_p \end{bmatrix}, \tag{6.80}$$

with

$$\boldsymbol{A}_p(U_\infty) = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=0},\quad \boldsymbol{B}_p = \left.\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=0},$$

where

$$\boldsymbol{C}_p = \boldsymbol{I}, \quad \boldsymbol{C}_h = \begin{bmatrix} \sqrt{\boldsymbol{Q}} \\ 0 \end{bmatrix}, \quad \boldsymbol{E}_h = \begin{bmatrix} 0 \\ \sqrt{\boldsymbol{R}} \end{bmatrix},$$

and vector $\boldsymbol{w}_p$ contains process noises.

In the case of AFS, the value of b_{mP} in Theorems 6.1 and 6.2 varies with U_∞, which thus, is denoted by $\varpi(U_\infty)$ herein instead. Given the $\|\boldsymbol{P}\| > b_{mP}$ condition in Theorem 6.1 and 6.2, a parameter-varying $\boldsymbol{P}(U_\infty)$ is now introduced, and accordingly, it is required that $\|\boldsymbol{P}(U_\infty)\| > \varpi(U_\infty)$ for a stable closed loop. To properly determine $\boldsymbol{P}(U_\infty)$, the following LMIs are constructed by introducing a performance variable $\boldsymbol{Z}(U_\infty)$ and a positive-definite symmetric matrix $\bar{\boldsymbol{X}}(U_\infty)$:

$$\begin{bmatrix} \dot{\bar{\boldsymbol{X}}} + \boldsymbol{A}_c^{\mathrm{T}}\bar{\boldsymbol{X}} + \bar{\boldsymbol{X}}\boldsymbol{A}_c & \bar{\boldsymbol{X}} \\ \bar{\boldsymbol{X}} & -\nu\boldsymbol{I} \end{bmatrix} \prec 0, \tag{6.81}$$

$$\begin{bmatrix} \bar{\boldsymbol{X}} & \boldsymbol{C}_{ch}^{\mathrm{T}} \\ \boldsymbol{C}_{ch} & \boldsymbol{Z} \end{bmatrix} \succ 0, \tag{6.82}$$

$$\mathrm{Tr}(\boldsymbol{Z}(U_\infty)) < \nu, \tag{6.83}$$

where ν is a performance index, and

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

$$\boldsymbol{C}_{ch} \triangleq \boldsymbol{C}_h - \frac{1}{2}\boldsymbol{E}_h\boldsymbol{R}^{-1}\boldsymbol{B}_p^{\mathrm{T}}\boldsymbol{P}(U_\infty).$$

Let $\bar{\boldsymbol{B}}_c \triangleq \frac{1}{2}\boldsymbol{B}_p\boldsymbol{R}^{-1}\boldsymbol{B}_p^{\mathrm{T}}$, $\boldsymbol{Y} = \boldsymbol{P}\bar{\boldsymbol{X}}^{-1}$, and $\boldsymbol{J} = \bar{\boldsymbol{X}}^{-1}$. Then from Eqs. (6.81) and (6.82), we have

$$-\dot{\boldsymbol{J}} + \boldsymbol{A}_p\boldsymbol{J} + \boldsymbol{J}\boldsymbol{A}_p^{\mathrm{T}} + \bar{\boldsymbol{B}}_c\boldsymbol{Y} + \boldsymbol{Y}^{\mathrm{T}}\bar{\boldsymbol{B}}_c^{\mathrm{T}} \prec 0, \tag{6.84}$$

$$\begin{bmatrix} \boldsymbol{J} & (\boldsymbol{C}_h\boldsymbol{J} + \boldsymbol{E}_h\boldsymbol{Y})^{\mathrm{T}} \\ \boldsymbol{C}_h\boldsymbol{J} + \boldsymbol{E}_h\boldsymbol{Y} & \boldsymbol{Z} \end{bmatrix} \succ 0. \tag{6.85}$$

Based on Eq. (5.1), $\boldsymbol{A}_p(U_\infty)$ can be in the form of

$$\boldsymbol{A}_p(U_\infty) = \boldsymbol{A}_{p1} + \boldsymbol{A}_{p2}U_\infty + \boldsymbol{A}_{p3}U_\infty^2. \tag{6.86}$$

Accordingly, $\boldsymbol{J}(U_\infty)$ and $\boldsymbol{\Upsilon}(U_\infty)$ are given by

$$\boldsymbol{J}(U_\infty) = \boldsymbol{J}_1 + \boldsymbol{J}_2 U_\infty + \boldsymbol{J}_3 U_\infty^2, \tag{6.87}$$

$$\boldsymbol{\Upsilon}(U_\infty) = \boldsymbol{\Upsilon}_1 + \boldsymbol{\Upsilon}_2 U_\infty + \boldsymbol{\Upsilon}_3 U_\infty^2, \tag{6.88}$$

and can be obtained by solving Eqs. (6.83), (6.84), and (6.85). Finally, there is

$$\boldsymbol{P}(U_\infty) = \boldsymbol{\Upsilon}(U_\infty)\boldsymbol{J}^{-1}(U_\infty), \tag{6.89}$$

which satisfies $\|\boldsymbol{P}(U_\infty)\| > \varpi(U_\infty)$, with $\varpi(U_\infty)$ implicitly expressed in the LMIs.

Remark 6.5. Different from other existing NOCOS methods which only apply to locally nonlinear systems (i.e., not suitable for the AFS cases with non-constant U_∞), the MVFA with the scheduled parameter matrix $\boldsymbol{P}$ enables the proposed NOCOS scheme under CICs to cover globally nonlinear scenarios involving non-constant U_∞.

6.3.5 Online system identification

Despite aeroelastic dynamics being analytically known, online system identification is desired to reduce the impact of modelling errors. A suitable solution can be realised via an NN as:

$$\dot{\boldsymbol{x}} = \boldsymbol{W}_s^{\mathrm{T}} \boldsymbol{\Phi}_s(\boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{\varepsilon}_s, \tag{6.90}$$

where $\boldsymbol{W}_s \in \mathbb{R}^{n_x \times n_{ws}}$ collects the ideal weights, $\boldsymbol{\Phi}_s(\boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}^{n_{ws}}$ contains nonlinear activation functions, and $\boldsymbol{\varepsilon}_s \in \mathbb{R}^{n_x}$ is a vector of approximation errors.

According to Modares et al. (2013a), the solution to Eq. (6.90) has a filtered representation of

$$\boldsymbol{x} = \boldsymbol{W}_s^{\mathrm{T}} \boldsymbol{\mu}_1 + \boldsymbol{\Gamma} \boldsymbol{\mu}_2 + \boldsymbol{\varepsilon}_x, \tag{6.91}$$

with

$$\begin{bmatrix} \dot{\boldsymbol{\mu}}_1 \\ \dot{\boldsymbol{\mu}}_2 \end{bmatrix} = \begin{bmatrix} -\boldsymbol{\Gamma} & \boldsymbol{I} & 0 & 0 \\ 0 & 0 & -\boldsymbol{\Gamma} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\Phi}_s(\boldsymbol{x}, \boldsymbol{u}) \\ \boldsymbol{\mu}_2 \\ \boldsymbol{x} \end{bmatrix},$$

where $\boldsymbol{\mu}_1 \in \mathbb{R}^{n_{ws}}$ with zero initial condition and $\boldsymbol{\mu}_2 \in \mathbb{R}^{n_x}$ also with zero initial condition are both auxiliary regressors; $\boldsymbol{\Gamma} \in \mathbb{R}^{n_x \times n_x}$ is a constant matrix and $\boldsymbol{\Gamma} = \gamma \boldsymbol{I}_{n_x \times n_x}$ with $\gamma \in \mathbb{R}^+$ being a constant; $\boldsymbol{\varepsilon}_x = e^{-\boldsymbol{\Gamma} t}\boldsymbol{x}_0 + \int_0^t e^{-\boldsymbol{\Gamma}(t-\tau)}\boldsymbol{\varepsilon}_s d\tau$ is the approximation error.

Note that $\boldsymbol{W}_s$ needs to be found so that $\boldsymbol{x}$ can be correctly obtained as in Eq. (6.91). The corresponding estimation $\hat{\boldsymbol{W}}_s$ and $\hat{\boldsymbol{x}}$ can be obtained by employing the EKF. In the interests of light computational load during real-time implementation, the EKF can be constructed in a decoupled form, consisting of multiple EKFs of one input and one output, each approximating a single state in $\boldsymbol{x}$. Accordingly,

$$\begin{cases} \hat{x}^{(j)} = \hat{\boldsymbol{W}}_s^{(j)}\boldsymbol{\mu}_1(\boldsymbol{x}) + \boldsymbol{\Gamma}\boldsymbol{\mu}_2(\boldsymbol{x}), \\ \dot{\hat{\boldsymbol{W}}}_s^{(j)} = \kappa_s \boldsymbol{K}_s^{(j)} e_x^{(j)}, \\ e_x^{(j)} = x^{(j)} - \hat{x}^{(j)}, \\ \kappa_s = \dfrac{\alpha_{snn}}{\left\| \boldsymbol{K}_s^{(j)} \right\| + 1}, \end{cases} \tag{6.92}$$

where $x^{(j)} \in \mathbb{R}$ is the j^{th} measured state; $\hat{x}^{(j)} \in \mathbb{R}$ is the j^{th} estimated state; $\alpha_{snn} \in \mathbb{R}^+$ is a constant; $\boldsymbol{K}_s^{(j)} \in \mathbb{R}^{n_{ws}}$ is the gain of the j^{th} decoupled EKF; and $\hat{\boldsymbol{W}}_s^{(j)} \in \mathbb{R}^{1 \times n_{ws}}$ is a row vector corresponding to the j^{th} row of $\hat{\boldsymbol{W}}_s$.

The gain $\boldsymbol{K}_s^{(j)}$ of each decoupled EKF can be obtained from

$$\boldsymbol{K}_s^{(j)} = \boldsymbol{S}_s^{(j)} \boldsymbol{\mu}_1 R_s^{-1}, \tag{6.93}$$

$$\dot{\boldsymbol{S}}_s^{(j)} = \boldsymbol{Q}_s - \boldsymbol{S}_s^{(j)} \boldsymbol{\mu}_1 R_s^{-1} \boldsymbol{\mu}_1^{\mathrm{T}} \boldsymbol{S}_s^{(j)}. \tag{6.94}$$

where $\boldsymbol{Q}_s \in \mathbb{R}^{n_{ws} \times n_{ws}}$ and $R_s \in \mathbb{R}^+$ are respective weightings.

Based on Eqs. (6.90), (6.91), (6.92) to (6.94), we now have an estimate of $\dot{\hat{\boldsymbol{x}}}$ as

$$\dot{\hat{\boldsymbol{x}}} = \hat{\boldsymbol{W}}_s^{\mathrm{T}} \boldsymbol{\Phi}_s(\boldsymbol{x}, \boldsymbol{u}). \tag{6.95}$$

By using Eq. (6.95), there is

$$\frac{\partial \dot{\hat{\boldsymbol{x}}}}{\partial \boldsymbol{u}} = \frac{\partial \boldsymbol{\Phi}_s(\boldsymbol{x}, \boldsymbol{u})}{\partial \boldsymbol{u}} \hat{\boldsymbol{W}}_s. \tag{6.96}$$

Let $\hat{\boldsymbol{g}}(\boldsymbol{x})$ denote the estimation of $\boldsymbol{g}(\boldsymbol{x})$. Since $\hat{\boldsymbol{g}}(\boldsymbol{x}) = \frac{\partial \dot{\hat{\boldsymbol{x}}}}{\partial \boldsymbol{u}}$, from Eq. (6.96) we have

$$\hat{\boldsymbol{g}}(\boldsymbol{x}) = \frac{\partial \boldsymbol{\Phi}_s(\boldsymbol{x}, \boldsymbol{u})}{\partial \boldsymbol{u}} \hat{\boldsymbol{W}}_s. \tag{6.97}$$

Equations (6.95) and (6.97) now provide necessary information required by the proposed control scheme.

Remark 6.6. $\boldsymbol{W}_s$ can be initialised by offline training of the identifier NN performed prior to online implementation. One may use a known analytical model with U_∞ set to the flutter boundary (denoted by U_b herein) for the offline training given the fact that the AFS controller only needs to be activated when the aircraft is approaching its flutter onset airspeed. Modelling errors (for $U_\infty \geq U_b$) can then be captured in real time and used to update the pre-trained NN model online.

6.4 Wind-tunnel experiments

For validation of the proposed optimal NN control (ONNC) scheme that considers CICs (in short, CIC-ONNC), wind-tunnel experiments were conducted (see setup in Figure 3.13). A scaled-down NACA 0012 aerofoil section was placed in a testing duct measuring 0.5×0.5 m in cross-section, and was vertically mounted on a virtual stiffness-damping system (VSDS). The VSDS uses electric drives to generate plunge force and pitch torque to mimic the effects of physical translational and torsional springs. This allows convenient adjustment of the structural stiffness and damping of the aeroelastic system to suit different test scenarios. Table 6.1 lists the parameters of the wind-tunnel experiment setup which together deliver an aeroelastic system subjected to flutter when the airflow reaches or increases beyond 14.1 m/s.

To implement the proposed CIC-ONNC, the activation functions of the NNs were generated as in Remark 6.3, and Table 6.2 gives an overview of these NNs. For synthesis of $\boldsymbol{J}(U_\infty)$ and $\boldsymbol{Y}(U_\infty)$, 50 scheduling points are equally spaced over $U_\infty \subseteq [14.0,\ 20.0]$ m/s. Then $\boldsymbol{P}(U_\infty)$ was obtained online according to Eq. (6.89). The conventional quadratic structure of $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}$ was adopted for $\bar{Q}(\boldsymbol{x})$ in Eq. (6.2). All parameters related to the proposed CIC-ONNC are given in Table 6.3.

As discussed in Section 6.1, the existing methods for AFS do not consider CICs from the optimal control perspective and no existing NOCOS algorithms that

Table 6.1: Wind-tunnel experiment parameters.

Parameters	Values	Parameters	Values
ρ	1.225 kg/m^3	c_h	14.0 kg/s
m_a	0.851 kg	c_a	0.042 kg·m^2/s
m_{te}	0.030 kg	c_{te}	4.231×10^{-4} kg·m^2/s
m_{le}	0.058 kg	c_{le}	4.327×10^{-4} kg·m^2/s
r_{fc}	-0.0685 m	k_h	$50 + 300q_h^2$ N/m
$r_{3c/4}$	0.081 m	k_a	$0.3 + 5q_a^2$ Nm/rad
r_a	0.033 m	k_{te}	4.570×10^{-3} Nm/rad
r_{te}	1.019×10^{-2} m	k_{le}	4.704×10^{-3} Nm/rad
r_{le}	4.401×10^{-3} m	I_a	2.431×10^{-3} kg·m^2
$C_{l\text{-}a}$	6.573	I_{te}	2.307×10^{-6} kg·m^2
$C_{l\text{-}te}$	3.472	I_{le}	4.791×10^{-6} kg·m^2
$C_{l\text{-}le}$	-0.145	L_{te}	0.088 m
$C_{m\text{-}a}$	0	L_{le}	-0.010 m
$C_{m\text{-}te}$	-0.631	L_s	0.260 m
$C_{m\text{-}le}$	0.098	L_{hc}	0.075 m

Table 6.2: Key information of the NN components of the proposed CIC-ONNC.

NN	Input	Order	Neurons	Weights Initialisation
Identifier	$\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{u} \end{bmatrix}$	$\boldsymbol{x}$: up to 4$^{\text{th}}$ $\boldsymbol{u}$: 1$^{\text{st}}$	$n_n = 135$	offline training based on simulations for 14.1 m/s airspeed
Critic	$\boldsymbol{x}$	up to 4$^{\text{th}}$	$n_n = 65$	zeros

deals with input-constrained cases apply to AFS. For these reasons, the robust controller for AFS studied in Prime (2010) was deemed appropriate for use in comparison, which has a primary structure as the linear-quadratic-regulator (LQR) optimal control scheduled via linear-parameter-varying (LPV) technique using LMIs. The LPV-LQR controller was designed using the same aeroelastic model, system parameters, and cost function weightings ($\boldsymbol{Q}$ and $\boldsymbol{R}$) as the proposed CIC-ONNC.

AFS at two different airspeeds were investigated, and the worst-case scenario was considered where the AFS controllers were expected to suppress flutter that had already developed and reached the state of limit-cycle oscillation. Automatic

Table 6.3: Parameters of the proposed CIC-ONNC.

Parameters	Values
$\boldsymbol{\lambda}$	$\pm 2\boldsymbol{I}$ degrees ($\pm 0.035\boldsymbol{I}$ rad) for $U_\infty = 14.8$ m/s tests $\pm 8\boldsymbol{I}$ degrees ($\pm 0.140\boldsymbol{I}$ rad) for $U_\infty = 18$ m/s tests $\pm 9\boldsymbol{I}$ degrees ($\pm 0.157\boldsymbol{I}$ rad) for $U_\infty = 18$ m/s tests
α_{cnn}	10^4
α_{snn}	10^4
$\boldsymbol{\Gamma}$	$10^3\boldsymbol{I}$
$\boldsymbol{Q}$	diag (1, 1, 10^{-4}, 10^{-4}, 10^{-1}, 10^{-1}, 10^{-4}, 10^{-4})
$\boldsymbol{R}$	$10^2\boldsymbol{I}$
$\boldsymbol{Q}_f$	$10^3\boldsymbol{I}$
R_f	1
$\boldsymbol{Q}_{s(i)}$	$10^5\boldsymbol{I}$
$R_{s(i)}$	1

triggers were put in place to switch the AFS controllers on when the pitch angle q_a crossed zero immediately after 15 s. Let t_c denote this triggering time. Then we have initial conditions $\boldsymbol{x}(t_c)$ under the experimental setting. This allowed all tests under different controllers to be evaluated under the same settings and consistent initial conditions.

For airspeed of $U_\infty = 14.8$ m/s, a bound of ± 2 degrees ($\boldsymbol{\lambda} = 0.035\boldsymbol{I}$ rad) was set for the TE and LE control surfaces by imposing saturation constraints to the TE and LE servos. Aeroelastic responses of the aerofoil section, together with TE and LE deflections, under the proposed CIC-ONNC and the benchmark LPV-LQR, are plotted in Figure 6.1. Corresponding trajectories of the NNs weights ($\hat{\boldsymbol{W}}_s$ and $\hat{\boldsymbol{W}}$) of the proposed CIC-ONNC in the form of their 2-norm are presented in Figure 6.2. No significant differences can be observed for plunge and pitch responses from Figure 6.1 before 0.5 s, whereas deviations start to increase afterwards (Figure 6.1) as the NNs weights converge (Figure 6.2). Figure 6.2 also shows that $\hat{\boldsymbol{W}}_s$ converges faster than $\hat{\boldsymbol{W}}$, and thus allowing proper convergence of $\hat{\boldsymbol{W}}$ in around 0.7 s, before the flutter is fully suppressed at approximately 1.5 s (under the proposed CIC-

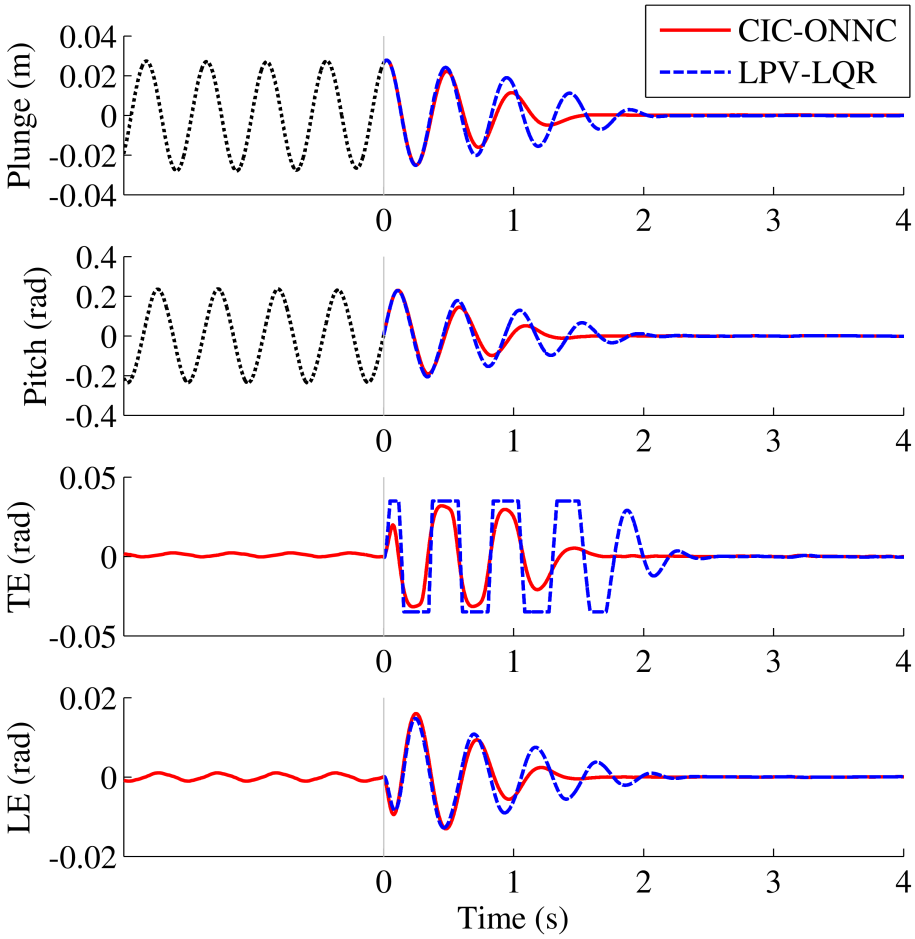


Figure 6.1: Suppression of developed flutter at $U_\infty = 14.8$ m/s airspeed with $\boldsymbol{\lambda} = 0.035\boldsymbol{I}$ rad.

ONNC). The convergence of both $\hat{\boldsymbol{W}}_s$ and $\hat{\boldsymbol{W}}$ well before the flutter settles suggests optimal control being obtained under CICs. By comparing TE and LE control surfaces deflections, it is straightforward to see that TE severely saturates the pre-set bounds under the LPV-LQR scheme due to constraints in control being not taken into account in controller synthesis. As a result, the LPV-LQR scheme takes longer to suppress the developed flutter.

In tests with $U_\infty = 18.0$ m/s airspeed, the ± 2 degrees bound was found too small to enable effective flutter suppression. Therefore, $\boldsymbol{\lambda}$ for both controllers were increased incrementally by a magnitude of 1 degree until successful flutter suppression was observed. It was found that the LPV-LQR scheme failed to suppress flutter until the saturation bound for TE and LE was raised to an absolute value of 9 degrees ($\boldsymbol{\lambda} = 0.157\boldsymbol{I}$ rad), while the proposed CIC-ONNC successfully eliminated flutter

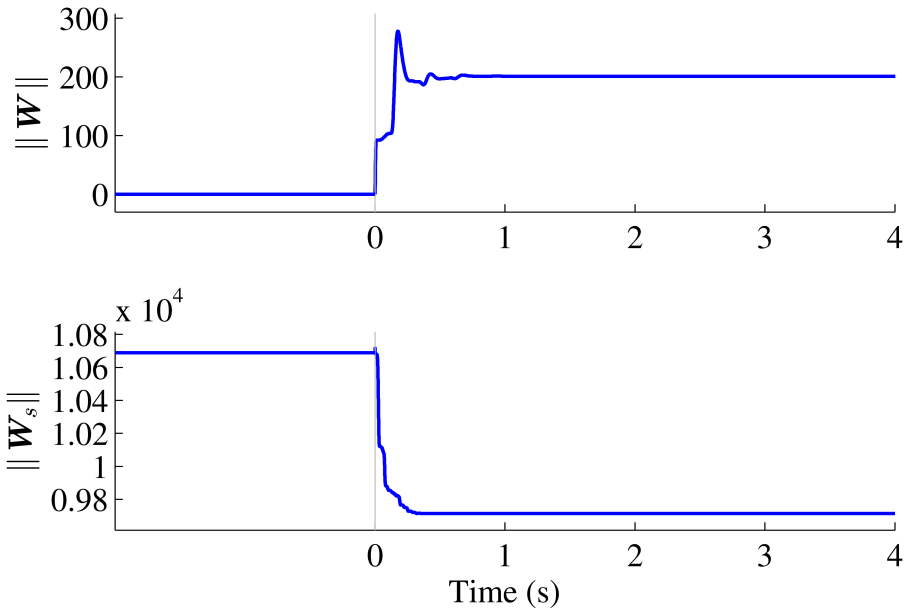


Figure 6.2: NN weights convergence of the proposed CIC-ONNC at $U_\infty = 14.8$ m/s airspeed with $\boldsymbol{\lambda} = 0.035\boldsymbol{I}$ rad.

with a ± 8 degrees bound ($\boldsymbol{\lambda} = 0.140\boldsymbol{I}$ rad), as can be seen from Figures 6.3 and 6.5. With the ± 9 degrees bound, the proposed CIC-ONNC also yielded superior AFS results, as demonstrated in Figure 6.5, where in contrast, the LPV-LQR scheme heavily saturates both control surfaces and needs significantly longer time to suppress flutter. Regarding the NNs weights during online learning, the 2-norm trajectories in Figures 6.4 and 6.6 show similar patterns to those in the $U_\infty = 14.8$ m/s case, consistently indicating satisfactory convergence and optimal control being reached.

The optimality of the control law synthesised in real time using the proposed CIC-ONNC can be further confirmed by comparing the performance costs of both controllers, calculated from the experiments data. The span of the data included for cost computation in each different test scenario is the same as the time range shown in Figures 6.1, 6.3, and 6.5. For the cost of control inputs, both the conventional quadratic formulation $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}$ and the nonquadratic one as in Eq. (6.4) in their discretised form are considered. The values in Table 6.4 are the averaged results from four repeated experiments, which show that the proposed CIC-ONNC has consistent low costs while those of the LPV-LQR scheme are much higher. It is also worth noting that the differences between the performance costs obtained by

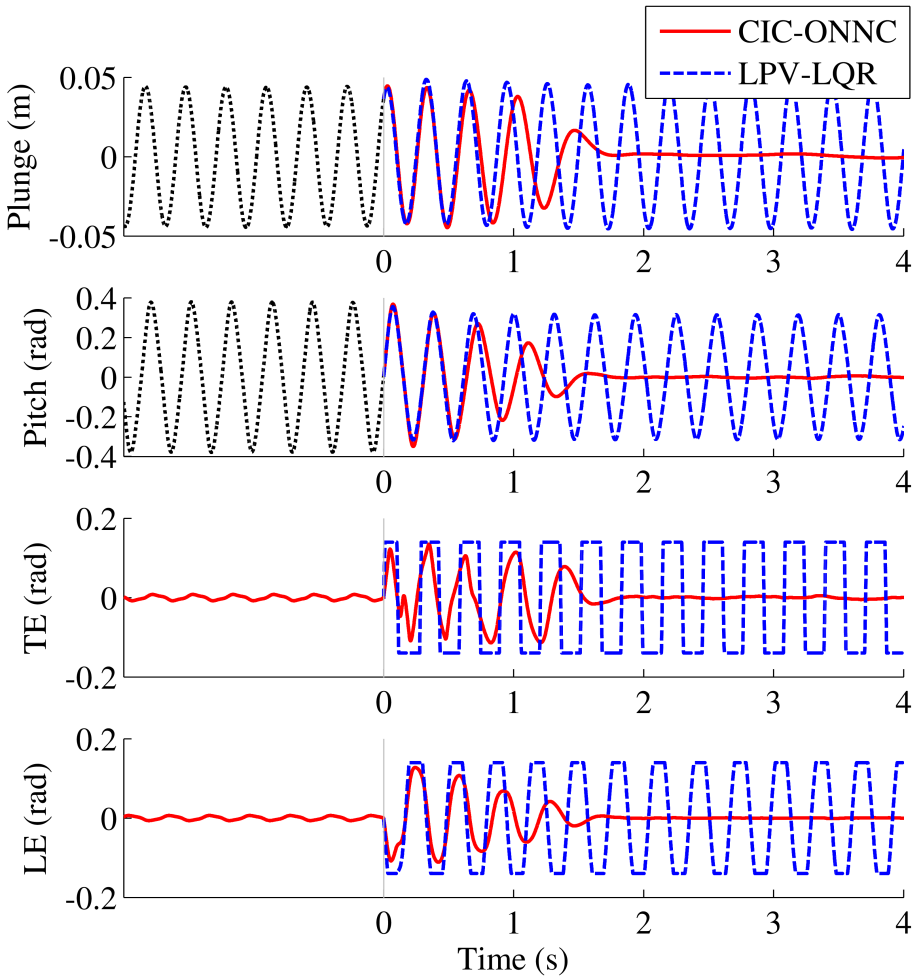


Figure 6.3: Suppression of developed flutter at $U_\infty = 18.0$ m/s airspeed with $\boldsymbol{\lambda} = 0.140\boldsymbol{I}$ rad.

Table 6.4: Performance costs calculated from experiments data.

Tests	$V(\boldsymbol{x}(t_c))$ from Eq. (6.2) with $\bar{Q}(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}$			
	$\bar{R}(\boldsymbol{u}) = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{u}$		$\bar{R}(\boldsymbol{u})$ as Eq. (6.4)	
	LPV-LQR	CIC-ONNC	LPV-LQR	CIC-ONNC
14.8 m/s with $\boldsymbol{\lambda} = 0.035\boldsymbol{I}$	1.615	0.380	0.542	0.411
18.0 m/s with $\boldsymbol{\lambda} = 0.140\boldsymbol{I}$	513.393	7.634	70.446	8.780
18.0 m/s with $\boldsymbol{\lambda} = 0.157\boldsymbol{I}$	129.213	7.633	22.242	8.786

different cost functions are small under the proposed CIC-ONNC compared with those of the LPV-LQR scheme. This finding is in agreement with Remark 6.1, as the proposed CIC-ONNC did not saturate the TE and LE whereas heavy saturation of control was observed under the LPV-LQR scheme (see Figures 6.1, 6.3, and 6.5).

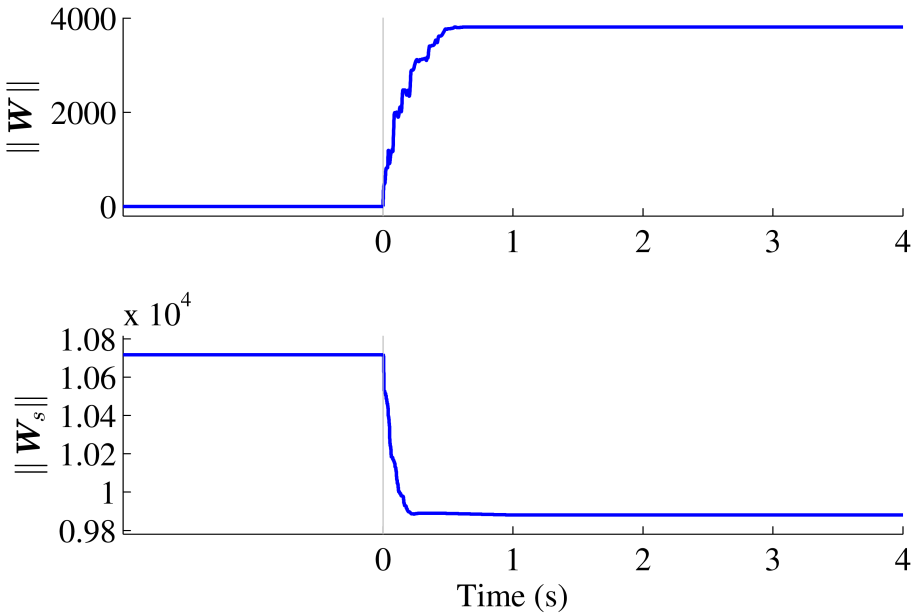


Figure 6.4: NN weights convergence of the proposed CIC-ONNC at $U_\infty = 18.0$ m/s airspeed with $\boldsymbol{\lambda} = 0.140\boldsymbol{I}$ rad.

6.5 Conclusions

A novel control scheme, CIC-ONNC based on NOCOS, is proposed and analysed in this study to improve AFS performance under CICs. The Lyapunov stability analysis shows that it is viable to introduce the generalised nonquadratic control-input cost function to the NOCOS involving the MVFA, contributing new knowledge to the NOCOS theory framework. As demonstrated and confirmed in experimental studies in a wind tunnel, the proposed CIC-ONNC using the MVFA and EKF with a performance cost containing the generalised nonquadratic control-input cost function is practically feasible and effective in optimally dealing with CICs involved in AFS, delivering satisfactory AFS results.

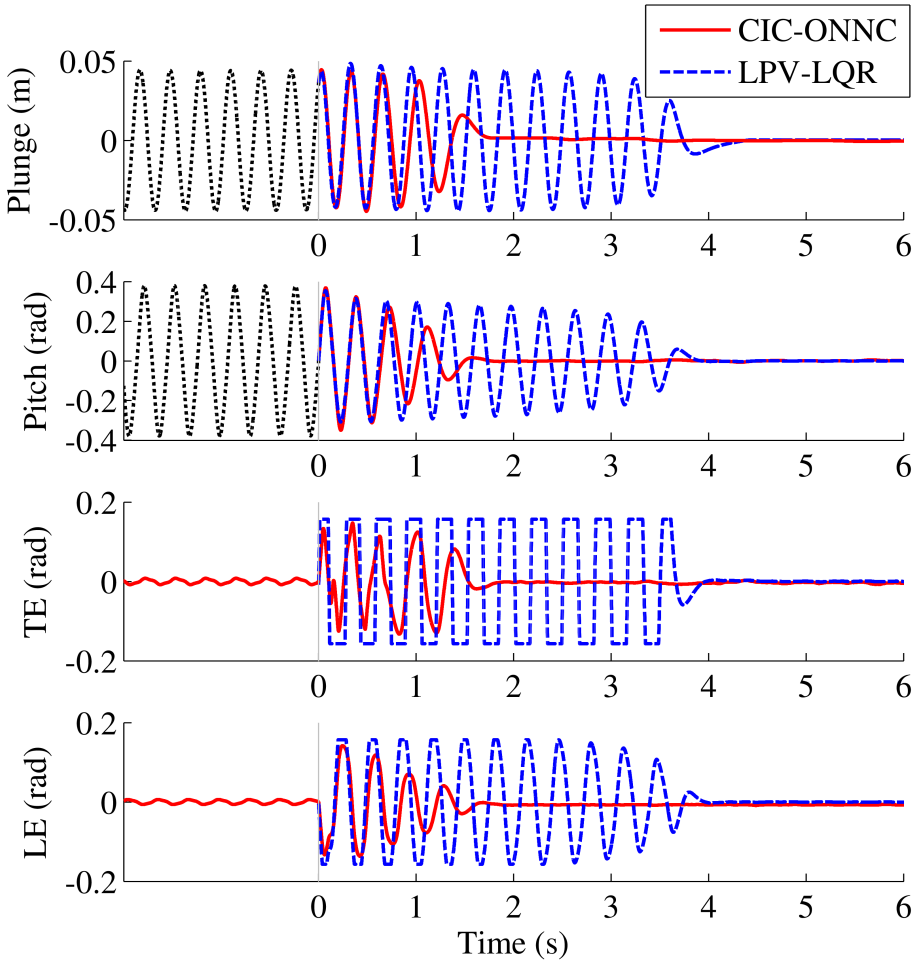


Figure 6.5: Suppression of developed flutter at $U_\infty = 18.0$ m/s airspeed with $\boldsymbol{\lambda} = 0.157\boldsymbol{I}$ rad.

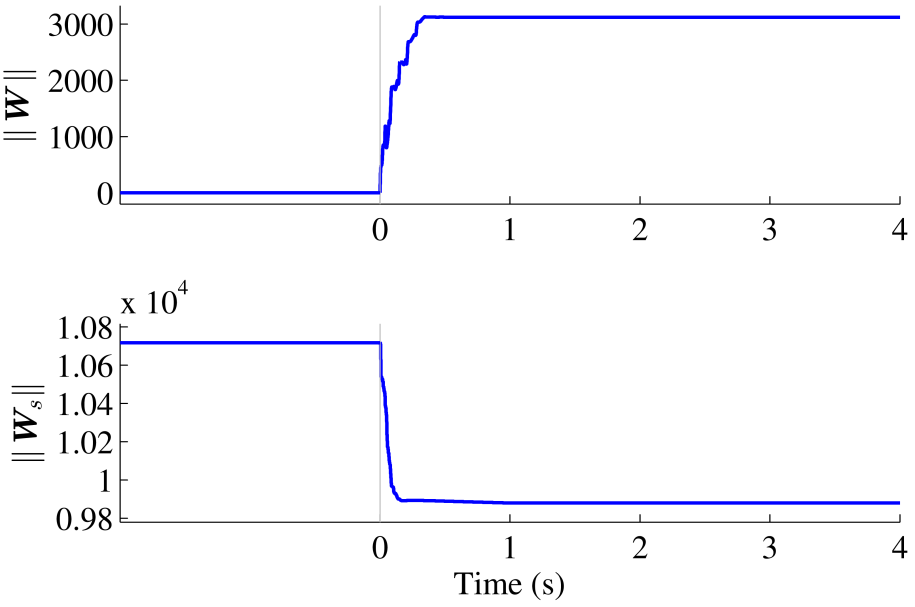


Figure 6.6: NN weights convergence of the proposed CIC-ONNC at $U_\infty = 18.0$ m/s airspeed with $\boldsymbol{\lambda} = 0.157\boldsymbol{I}$ rad.

Chapter 7

Conclusions

The contributions and significance of the overall work on active suppression of aerofoil flutter (AFS) and its essential elements are summarised in Section 7.1. Future directions of the work are recommended in Section 7.2.

7.1 Contributions and significance

In general, the proposed solutions in this thesis based on online synthesis of nonlinear optimal control (NOCOS) using neural networks (NNs) are novel and shown effective in solving the AFS controller synthesis dilemma and improving AFS from the optimal control perspective. The proposed methods were also experimentally confirmed feasible for real-time implementation.

Specifically, the methods derived in this thesis at each stage towards the ultimate goal of solving the dilemma involved in AFS controller synthesis have respective technical and scientific significance as follows:

- The virtual stiffness-damping system (VSDS) prototype developed in Chapter 3 for wind-tunnel experiments has a technical breakthrough in enabling improved and robust simulation of physical stiffness and damping in oscillatory systems. It was experimentally proven capable of mimicking traditional aeroelasticity test-beds and hence serving as a reliable platform for experi-

mental validation of the proposed AFS controllers that follow. It also has the potential for other industrial applications involving oscillatory testing requiring frequent change of stiffness and damping settings.

- The proposed NOCOS algorithm with the Modified form of NN-based Value Function Approximation (MVFA) presented in Chapter 4 is theoretically and analytically shown effective in addressing closed-loop stability in a compact configuration suitable for real-time implementation. It not only forms an essential basis for subsequent studies towards solving the AFS control synthesis dilemma but also contributes new knowledge to the NOCOS theory framework.

- In Chapter 5, the systematic procedure proposed for synthesising the scheduled parameter matrix $\boldsymbol{P}$ plays a crucial role in generalising the MVFA from locally nonlinear systems to globally nonlinear scenarios. The resulting generalisation is a vital step to make the proposed NOCOS scheme suitable for AFS applications. The extended Kalman filter (EKF) proposed for critic NN tuning with the MVFA, theoretically proven stable and experimentally confirmed feasible, allows satisfactory parameter convergence for AFS during online learning. The impact of modelling errors or uncertainties can be significantly reduced by implementing online system identification using decoupled identifier NNs, as shown effective in wind-tunnel tests. These three elements – MVFA generalisation, EKF-based MVFA tuning, and online system identification, together form the proposed NN-based control scheme for AFS and provides a complete solution to solving the AFS controller synthesis dilemma, assuming no actuator constraints.

- Optimal control for AFS in the presence of control-input constraints is tackled in Chapter 6, where the introduction of a non-quadratic functional for generalising the traditional quadratic cost function to a non-quadratic form for the NN controller based on the MVFA is theoretically and experimentally proven

viable. The proposed NN controller with the generalised cost function thus successfully addresses AFS scenarios where the available control authority is bounded by some constraints, in an optimal manner. Experimental results are promising and confirm the validity and efficacy of the proposed NN controller under control-input constraints. This is a significant enhancement to the NN control scheme proposed in Chapter 5 and provides a better solution to solving the AFS controller synthesis dilemma.

7.2 Future work

Three possible directions are recommended for future work on AFS.

The first deals with the online tuning for both the critic and identifier NNs. The proposed NN controllers for AFS, with extended Kalman filters for online tuning, were proven practical and effective in wind-tunnel experiments. For faster parameter convergence and further performance improvement, a direct and possible way is to explore more efficient online learning of the NNs involved.

The second suggestion on possible future work is about the types of NNs used for the controller. The critic and identifier NNs of the proposed controllers are linear in the parameters (LIP) with explicit nonlinear activation functions. LIP NNs on one hand can give firm approximation results for specified accuracy but on the other hand require explicitly nominating a suitable basis of activation functions. Using the high-order Weierstrass approximation theorem is a convenient way to systematically select a set of proper activation functions, but the the number of activation functions can increase significantly as the order of the power series or the dimension of the system rises. Some activation functions as part of the power series given by the Weierstrass approximation theorem may have negligible influence on approximation accuracy and can be eliminated, but the refinement requires either a prior knowledge or simulation-aided procedures. Therefore, to further ease the controller synthesis, multi-layer NNs that are nonlinear in the parameters (NLIP)

are of interest for further investigation. By using NLIP NNs, there is no need to explicitly determine a basis of activation functions but only a common form such as sigmoid, hyperbolic tangent, radial-basis or other logistic-curve-type functions. Nonetheless, nonlinearity in NN weights poses new challenges in deriving efficient tuning algorithms.

The third recommendation is to consider exogenous disturbances in the optimal control scheme. For the case of AFS, typical exogenous disturbances are turbulence. During steady cruise regime, the turbulence intensity of atmosphere is low and has little impact on AFS. As shown in wind-tunnel experiments, where the turbulence intensity was 0.5%, no considerable perturbation was observed, and all results under the proposed NN controller are promising. However, should turbulence be inevitably encountered, the NN controllers proposed in this thesis may only be suboptimal, due to exogenous disturbances being not considered in the online optimal control synthesis. For optimal AFS in the presence of turbulence, possible solutions may involve solving the Hamilton-Jacobi-Isaacs equation instead of the Hamilton-Jacobi-Bellman equation in the critic NN. In addition, the disturbance-input dynamics may need to be included in the online system identification scheme as well.

Appendix A

Controller composition

This appendix supplies detailed neural-network composition of the proposed nonlinear optimal controller used in wind-tunnel experiments as presented in Chapters 5 and 6.

A.1 Identifier neural network activation functions

The activation functions contained in $\boldsymbol{\Phi}_s(\boldsymbol{x})$ are:

$x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7,\ x_8,\ u_1,\ u_2,\ x_1^3,\ x_1^2x_2,\ x_1^2x_4,\ x_1^2x_5,\ x_1^2x_6,\ u_1x_1^2,\ u_2x_1^2,$
$x_1x_2^2,\ x_1x_2x_4,\ u_2x_1x_2,\ x_1x_3x_4,\ x_1x_3x_5,\ x_1x_4x_5,\ x_1x_5^2,\ u_1x_1x_5,\ x_2^3,\ x_2^2x_3,\ x_2^2x_5,\ x_1^4,\ x_1^3x_2,$
$x_1^3x_3,\ x_1^3x_4,\ x_1^3x_5,\ x_1^3x_6,\ x_1^3x_7,\ x_1^3x_8,\ u_1x_1^3,\ u_2x_1^3,\ x_1^2x_2^2,\ x_1^2x_2x_3,\ x_1^2x_2x_4,\ x_1^2x_2x_5,\ x_1^2x_2x_6,$
$x_1^2x_2x_7,\ u_1x_1^2x_2,\ u_2x_1^2x_2,\ x_1^2x_3^2,\ x_1^2x_3x_4,\ x_1^2x_3x_5,\ x_1^2x_3x_6,\ x_1^2x_3x_7,\ x_1^2x_3x_8,\ u_1x_1^2x_3,\ u_2x_1^2x_3,$
$x_1^2x_4^2,\ x_1^2x_4x_5,\ x_1^2x_4x_6,\ x_1^2x_4x_7,\ x_1^2x_4x_8,\ u_1x_1^2x_4,\ u_2x_1^2x_4,\ x_1^2x_5^2,\ x_1^2x_5x_6,\ x_1^2x_5x_7,\ x_1^2x_5x_8,$
$u_1x_1^2x_5,\ u_2x_1^2x_5,\ x_1^2x_6^2,\ x_1^2x_6x_8,\ u_2x_1^2x_6,\ u_1x_1^2x_7,\ u_2x_1^2x_7,\ u_1x_1^2x_8,\ u_2x_1^2x_8,\ x_1x_2^3,\ x_1x_2^2x_3,$
$x_1x_2^2x_4,\ x_1x_2^2x_5,\ x_1x_2^2x_6,\ x_1x_2^2x_7,\ x_1x_2^2x_8,\ u_1x_1x_2^2,\ x_1x_2x_3^2,\ x_1x_2x_3x_4,\ x_1x_2x_3x_5,\ u_1x_1x_2x_3,$
$u_2x_1x_2x_3,\ x_1x_2x_4^2,\ x_1x_2x_4x_5,\ u_1x_1x_2x_4,\ u_2x_1x_2x_4,\ x_1x_2x_5^2,\ x_1x_2x_5x_6,\ u_1x_1x_2x_5,\ u_2x_1x_2x_5,$
$u_1x_1x_2x_6,\ x_1x_3^2x_5,\ u_1x_1x_3^2,\ u_2x_1x_3^2,\ x_1x_3x_4^2,\ x_1x_3x_4x_5,\ u_1x_1x_3x_4,\ u_2x_1x_3x_4,\ x_1x_3x_5^2,\ u_1x_1x_3x_5,$
$u_2x_1x_3x_5,\ u_1x_1x_3x_6,\ u_2x_1x_3x_6,\ x_1x_4^3,\ x_1x_4^2x_5,\ u_1x_1x_4^2,\ x_1x_4x_5^2,\ x_1x_4x_5x_6,\ u_1x_1x_4x_5,\ u_2x_1x_4x_5,$
$u_2x_1x_4x_6,\ x_1x_5^3,\ x_1x_5^2x_6,\ u_1x_1x_5^2,\ u_2x_1x_5^2,\ u_2x_1x_5x_6,\ x_2^3x_4,\ x_2^3x_5,\ u_1x_2^3,\ u_1x_2^2x_3,\ u_2x_2^2x_3,$
$x_2^2x_5^2,\ u_1x_2^2x_5,\ u_2x_2^2x_5,\ x_2x_3^2x_4,\ u_1x_2x_3x_5,\ x_2x_4x_5^2,\ u_2x_2x_5^2,\ x_3x_5^3,\ u_2x_4x_5^2.$

A.2 Critic neural network activation functions

The activation functions contained in $\boldsymbol{\Phi}(\boldsymbol{x})$ are:

x_1^4, $x_1^3x_6$, $x_1^2x_2x_4$, $x_1^2x_3^2$, $x_1^2x_3x_8$, $x_1^2x_4x_8$, $x_1^2x_6^2$, $x_1^2x_8^2$, $x_1x_2^2x_6$, $x_1x_2x_3x_5$, $x_1x_2x_4x_5$, $x_1x_2x_5x_6$, $x_1x_2x_6x_8$, $x_1x_3^2x_4$, $x_1x_3x_4^2$, $x_1x_3x_5^2$, $x_1x_3x_6x_7$, $x_1x_4^3$, $x_1x_4x_5^2$, $x_1x_4x_6x_7$, $x_1x_5^3$, $x_1x_5x_6x_7$, $x_1x_6^3$, $x_1x_6x_8^2$, x_2^4, $x_2^3x_7$, $x_2^2x_3x_6$, $x_2^2x_4x_6$, $x_2^2x_5x_7$, $x_2^2x_7^2$, $x_2x_3^2x_5$, $x_2x_3x_4x_5$, $x_2x_3x_5x_6$, $x_2x_3x_6x_8$, $x_2x_4^2x_5$, $x_2x_4x_5x_6$, $x_2x_4x_6x_8$, $x_2x_5^2x_6$, $x_2x_5x_6x_8$, $x_2x_6^2x_7$, $x_3^2x_4^2$, $x_3^2x_5^2$, $x_3^2x_6x_7$, $x_3x_4^3$, $x_3x_4x_5^2$, $x_3x_4x_6x_7$, $x_3x_5^3$, $x_3x_5x_6x_7$, $x_3x_6^3$, $x_3x_6x_8^2$, x_4^4, $x_4^2x_5^2$, $x_4^2x_6x_7$, $x_4x_5^3$, $x_4x_5x_6x_7$, $x_4x_6^3$, $x_4x_6x_8^2$, x_5^4, $x_5^2x_6x_7$, $x_5x_6^3$, $x_5x_6x_8^2$, x_6^4, $x_6^2x_8^2$, x_7^4, x_8^4.

References

Abdelkefi A, Vasconcellos R, Marques F and Hajj M (2012) Modeling and identification of freeplay nonlinearity. *Journal of Sound and Vibration* 331(8): 1898–1907.

Abouheaf M, Gueaieb W and Sharaf A (2018) Model-free adaptive learning control scheme for wind turbines with doubly fed induction generators. *IET Renewable Power Generation* 12(14): 1675–1686.

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.

Al-Tamimi A, Lewis FL and Abu-Khalaf M (2008) Discrete-time nonlinear HJB solution using approximate dynamic programming: Convergence proof. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 38(4): 943–949.

Anderson BDO and Moore JB (1990) *Optimal Control: Linear Quadratic Methods*. Englewood Cliffs, NJ: Prentice-Hall.

Balakrishnan AV (2012) *Aeroelasticity: The Continuum Theory*. New York, NY: Springer.

Beard RW, Saridis GN and Wen JT (1997) Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation. *Automatica* 33(12): 2159–2177.

Behal A, Marzocca P, Rao VM and Gnann A (2006a) Nonlinear adaptive control of an aeroelastic two-dimensional lifting surface. *Journal of Guidance, Control, and Dynamics* 29(2): 382–390.

Behal A, Rao VM, Marzocca P and Kamaludeen M (2006b) Adaptive control for a nonlinear wing section with multiple flaps. *Journal of Guidance, Control, and Dynamics* 29(3): 744–749.

Bellman R (1957) *Dynamic Programming*. Princeton, NJ: Princeton University Press.

Bhasin S, Kamalapurkar R, Johnson M, Vamvoudakis KG, Lewis FL and Dixon WE (2013) A novel actor-critic-identifier architecture for approximate optimal control of uncertain nonlinear systems. *Automatica* 49(1): 82–92.

Bhoir N and Singh SN (2004) Output feedback modular adaptive control of a nonlinear prototypical wing section. *Nonlinear Dynamics* 37(4): 357–373.

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

Biskri D, Botez R, Stathopoulos N, Thérien S, Dickinson M and Rathe A (2008) Aeroservoelasticity analysis method based on an error analytical form applied on a business aircraft. *Journal of Vibration and Control* 14(8): 1217–1230.

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

Burnett EL, Beranek JA, Holm-Hansen BT, Atkinson CJ and Flick PM (2016) Design and flight test of active flutter suppression on the X-56A multi-utility technology test-bed aircraft. *The Aeronautical Journal* 120(1228): 893–909.

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.

Chang X, Li Y, Zhang W, Wang N and Xue W (2015) Active disturbance rejection control for a flywheel energy storage system. *IEEE Transactions on Industrial Electronics* 62(2): 991–1001.

Chen G, Sun J and Li YM (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.

Chen Y and Liu J (2010) Homotopy analysis method for limit cycle oscillations of an airfoil with cubic nonlinearities. *Journal of Vibration and Control* 16(2): 163–179.

Derakhshandeh JF, Arjomandi M, Cazzolato BS and Dally B (2015) Harnessing hydro-kinetic energy from wake-induced vibration using virtual mass spring damper system. *Ocean Engineering* 108: 115–128.

Dilmi S and Bouzouia B (2016) Improving performance for nonlinear aeroelastic systems via sliding mode controller. *Arabian Journal for Science and Engineering* 41(9): 3739–3748.

Edwards JW (1983) Flight test results of an active flutter suppression system. *Journal of Aircraft* 20(3): 267–274.

Erenturk K (2013) Fractional-order $PI^{\lambda}D^{\mu}$ and active disturbance rejection control of nonlinear two-mass drive system. *IEEE Transactions on Industrial Electronics* 60(9): 3806–3813.

Farmer MG (1982) A two-degree-of-freedom flutter mount system with low damping for testing rigid wings at different angles of attack. Technical Report NASA TM 83302, National Aeronautics and Space Administration.

Farmer MG (1984) Model mount system for testing flutter, U.S. Patent No. 4,475,385. Washington, DC: U.S. Patent and Trademark Office.

Fazelzadeh S, 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), DOI: 10.1061/(ASCE)AS.1943-5525.0000343.

Feng T, Zhang H, Luo Y and Zhang J (2015) Stability analysis of heuristic dynamic programming algorithm for nonlinear systems. *Neurocomputing* 149, Part C: 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.

Friedmann PP, Guillot D and Presente E (1997) Adaptive control of aeroelastic instabilities in transonic flow and its scaling. *Journal of Guidance, Control, and Dynamics* 20(6): 1190–1199.

Fung YC (1955) *An Introduction to the Theory of Aeroelasticity*. NY: John Wiley & Sons.

Gao MZ and Cai GP (2016) Finite-time fault-tolerant control for flutter involving control delay. *Journal of the Franklin Institute* 353(9): 2009–2029.

Gao MZ, Cai GP and Nan Y (2016) Finite-time fault-tolerant control for flutter of wing. *Control Engineering Practice* 51: 26–47.

Ghiringhelli GL, Lanz M and Mantegazza P (1987) A comparison of methods used for the identification of flutter from experimental data. *Journal of Sound and Vibration* 119(1): 39–51.

Ghiringhelli GL, Lanz M and Mantegazza P (1990) Active flutter suppression for a wing model. *Journal of Aircraft* 27(4): 334–341.

Guillot DM and Friedmann PP (2000) Fundamental aeroservoelastic study combining unsteady computational fluid mechanics with adaptive control. *Journal of Guidance, Control, and Dynamics* 23(6): 1117–1126.

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.

Haftka RT, Starnes Jr JH, Barton FW and Dixon SC (1975) Comparison of two types of structural optimization procedures for flutter requirements. *AIAA Journal* 13(10): 1333–1339.

Han J (2009) From PID to active disturbance rejection control. *IEEE Transactions on Industrial Electronics* 56(3): 900–906.

Heeg J (1993) Analytical and experimental investigation of flutter suppression by piezoelectric actuation. Technical Report NASA TP 3241, National Aeronautics and Space Administration.

Herbst G (2016) Practical active disturbance rejection control: Bumpless transfer, rate limitation, and incremental algorithm. *IEEE Transactions on Industrial Electronics* 63(3): 1754–1762.

Heydari A (2014) Revisiting approximate dynamic programming and its convergence. *IEEE Transactions on Cybernetics* 44(12): 2733–2743.

Heydari A and Balakrishnan SN (2013) Finite-horizon control-constrained nonlinear optimal control using single network adaptive critics. *IEEE Transactions on Neural Networks and Learning Systems* 24(1): 145–157.

Heydari A and Balakrishnan SN (2014) An adaptive critic-based scheme for consensus control of nonlinear multi-agent systems. *International Journal of Control* 87(12): 2463–2474.

Hornik K, Stinchcombe M and White H (1989) Multilayer feedforward networks are universal approximators. *Neural Networks* 2(5): 359–366.

Hover FS, Miller SN and Triantafyllou MS (1997) Vortex-induced vibration of marine cables: Experiments using force feedback. *Journal of Fluids and Structures* 11(3): 307–326.

Hover FS, Techet AH and Triantafyllou MS (1998) Forces on oscillating uniform and tapered cylinders in crossflow. *Journal of Fluid Mechanics* 363: 97–114.

Howard RA (1960) *Dynamic Programming and Markov Processes*. Cambridge, MA: MIT Press.

Huang Y, Wang D and Liu D (2017) Bounded robust control design for uncertain nonlinear systems using single-network adaptive dynamic programming. *Neurocomputing* 266: 128–140.

Ioannou PA and Sun J (1996) *Robust Adaptive Control*. Upper Saddle River, NJ: PTR Prentice-Hall.

Jiang H and He H (2018) Data-driven distributed output consensus control for partially observable multiagent systems. *IEEE Transactions on Cybernetics* 49(3): 848–858.

Jiang Y and Jiang ZP (2014) Robust adaptive dynamic programming and feedback stabilization of nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems* 25(5): 882–893.

Jiang Y and Jiang ZP (2015) Global adaptive dynamic programming for continuous-time nonlinear systems. *IEEE Transactions on Automatic Control* 60(11): 2917–2929.

Jiang ZP and Jiang Y (2013) Robust adaptive dynamic programming for linear and nonlinear systems: An overview. *European Journal of Control* 19(5): 417–425.

Karpel M (1981) Design for active and passive flutter suppression and gust alleviation. Technical Report NASA CR-3482, National Aeronautics and Space Administration.

Kelly A (2013) *Mobile Robotics: Mathematics, Models, and Methods*. New York, NY: Cambridge University Press.

Keyser RD, Copot C, Hernandez A and Ionescu C (2017) Discrete-time internal model control with disturbance and vibration rejection. *Journal of Vibration and Control* 23(1): 3–15.

Khalil M, Poirel D and Sarkar A (2016) Bayesian analysis of the flutter margin method in aeroelasticity. *Journal of Sound and Vibration* 384: 56–74.

Kimberlin RD (2003) *Flight Testing of Fixed-Wing Aircraft*. Reston, VA: American Institute of Aeronautics and Astronautics.

Kiumarsi B and Lewis FL (2015) Actor-critic-based optimal tracking for partially unknown nonlinear discrete-time systems. *IEEE Transactions on Neural Networks and Learning Systems* 26(1): 140–151.

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, Kurdila AJ and Strganac TW (1997) Nonlinear control of a prototypical wing section with torsional nonlinearity. *Journal of Guidance, Control, and Dynamics* 20(6): 1181–1189.

Ko J, Strganac TW, Junkins JL, Akella MR and Kurdila AJ (2002) Structured model reference adaptive control for a wing section with structural nonlinearity. *Journal of Vibration and Control* 8(5): 553–573.

Ko J, Strganac TW and Kurdila AJ (1998) Stability and control of a structurally nonlinear aeroelastic system. *Journal of Guidance, Control, and Dynamics* 21(5): 718–725.

Ko J, Strganac TW and Kurdila AJ (1999) Adaptive feedback linearization for the control of a typical wing section with structural nonlinearity. *Nonlinear Dynamics* 18(3): 289–301.

Lee BHK, Liu L and Chung KW (2005) Airfoil motion in subsonic flow with strong cubic nonlinear restoring forces. *Journal of Sound and Vibration* 281(3-5): 699–717.

Lee J and Bernitsas M (2011) High-damping, high-reynolds VIV tests for energy harnessing using the VIVACE converter. *Ocean Engineering* 38(16): 1697–1712.

Lee JH, Xiros N and Bernitsas MM (2011) Virtual damper-spring system for VIV experiments and hydrokinetic energy conversion. *Ocean Engineering* 38(5-6): 732–747.

Lee KW and Singh SN (2010) Multi-input noncertainty-equivalent adaptive control of an aeroelastic system. *Journal of Guidance, Control, and Dynamics* 33(5): 1451–1460.

Lee KW and Singh SN (2013) $\mathcal{L}_1$ adaptive control of a nonlinear aeroelastic system despite gust load. *Journal of Vibration and Control* 19(12): 1807–1821.

Leith DJ and Leithead WE (2000) Survey of gain-scheduling analysis and design. *International Journal of Control* 73(11): 1001–1025.

Lewis FL, Jagannathan S and Yesildirek A (1999) *Neural Network Control of Robot Manipulators and Nonlinear Systems*. London: Taylor & Francis.

Li D, Guo S, He Y and Xiang J (2012) Nonlinear aeroelastic analysis of a morphing flap. *International Journal of Bifurcation and Chaos* 22(5), DOI: 10.1142/S021812741250099X.

Li J, Modares H, Chai T, Lewis FL and Xie L (2017) Off-policy reinforcement learning for synchronization in multiagent graphical games. *IEEE Transactions on Neural Networks and Learning Systems* 28(10): 2434–2445.

Li S, Li J and Mo Y (2014) Piezoelectric multimode vibration control for stiffened plate using ADRC-based acceleration compensation. *IEEE Transactions on Industrial Electronics* 61(12): 6892–6902.

Liebst BS, Garrard WL and Adams WM (1986) Design of an active flutter suppression system. *Journal of Guidance, Control, and Dynamics* 9(1): 64–71.

Liu D, Huang Y, Wang D and Wei Q (2013a) Neural-network-observer-based optimal control for unknown nonlinear systems using adaptive dynamic programming. *International Journal of Control* 86(9): 1554–1566.

Liu D, Wang D, Wang FY, Li H and Yang X (2014) Neural-network-based online HJB solution for optimal robust guaranteed cost control of continuous-time uncertain nonlinear systems. *IEEE Transactions on Cybernetics* 44(12): 2834–2847.

Liu D, Yang X and Li H (2013b) Adaptive optimal control for a class of continuous-time affine nonlinear systems with unknown internal dynamics. *Neural Computing and Applications* 23(7-8): 1843–1850.

Liu RJ, Wu M, Liu GP, She J and Thomas C (2013c) Active disturbance rejection control based on an improved equivalent-input-disturbance approach. *IEEE/ASME Transactions on Mechatronics* 18(4): 1410–1413.

Livne E (2018) Aircraft active flutter suppression: State of the art and technology maturation needs. *Journal of Aircraft* 55(1): 410–450.

Lopez VG, Sanchez EN, Alanis AY and Rios JD (2017) Real-time neural inverse optimal control for a linear induction motor. *International Journal of Control* 90(4): 800–812.

Lorenz RD and Van Patten KW (1991) High-resolution velocity estimation for all-digital, AC servo drives. *IEEE Transactions on Industry Applications* 27(4): 701–705.

Lum KY, Xu CL, Lu Z, Lai KL and Cui Y (2017) Design and experiment of data-driven modeling and flutter control of a prototype wing. *Journal of Sound and Vibration* 398: 103–122.

Luo B, Liu D, Huang T and Wang D (2016a) Model-free optimal tracking control via critic-only Q-learning. *IEEE Transactions on Neural Networks and Learning Systems* 27(10): 2134–2144.

Luo B, Wu HN and Huang T (2018) Optimal output regulation for model-free Quanser helicopter with multistep Q-Learning. *IEEE Transactions on Industrial Electronics* 65(6): 4953–4961.

Luo B, Wu HN, Huang T and Liu D (2014) Data-based approximate policy iteration for affine nonlinear continuous-time optimal control design. *Automatica* 50(12): 3281–3290.

Luo B, Wu HN, Huang T and Liu D (2015) Reinforcement learning solution for HJB equation arising in constrained optimal control problem. *Neural Networks* 71: 150–158.

Luo M, Gao M and Cai G (2016b) Delayed full-state feedback control of airfoil flutter using sliding mode control method. *Journal of Fluids and Structures* 61: 262–273.

Luy NT (2018) Distributed cooperative H_∞ optimal tracking control of MIMO nonlinear multi-agent systems in strict-feedback form via adaptive dynamic programming. *International Journal of Control* 91(4): 952–968.

Lv Y, Na J and Ren X (2019) Online H_∞ control for completely unknown nonlinear systems via an identifier–critic-based ADP structure. *International Journal of Control* 92(1): 100–111.

Lv Y, Na J, Yang Q, Wu X and Guo Y (2016) Online adaptive optimal control for continuous-time nonlinear systems with completely unknown dynamics. *International Journal of Control* 89(1): 99–112.

Lyshevski SE (1998) Optimal control of nonlinear continuous-time systems: design of bounded controllers via generalized nonquadratic functionals. In: *Proceedings of the 1998 American Control Conference*, volume 1. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE), pp. 205–209.

Marcos A and Balas GJ (2004) Development of linear-parameter-varying models for aircraft. *Journal of Guidance, Control, and Dynamics* 27(2): 218–228.

Mardanpour P, Izadpanahi E, Rastkar S, Calastawad S and Levy C (2018) Effect of shooting and blast-induced gust on nonlinear aeroelastic stability and behavior of high-aspect ratio wing. *Journal of Sound and Vibration* 433: 299–313.

Markowitz J and Isakson G (1978) FASTOP-3: A strength, deflection and flutter optimization program for metallic and composite structures. Technical Report AFFDL-TR-78-50, Air Force Flight Dynamics Laboratory.

Marzocca P, Librescu L and Chiocchia G (2001) Aeroelastic response of 2-D lifting surfaces to gust and arbitrary explosive loading signatures. *International Journal of Impact Engineering* 25(1): 41–65.

Miller GD (1988) Active flexible wing (AFW) technology. Technical Report AFWAL-TR-87-3096, Air Force Wright Aeronautical Laboratories.

Modares H and Lewis FL (2014) Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. *Automatica* 50(7): 1780–1792.

Modares H, Lewis FL and Naghibi-Sistani MB (2013a) Adaptive optimal control of unknown constrained-input systems using policy iteration and neural networks. *IEEE Transactions on Neural Networks and Learning Systems* 24(10): 1513–1525.

Modares H, Lewis FL and Naghibi-Sistani MB (2014) Integral reinforcement learning and experience replay for adaptive optimal control of partially-unknown constrained-input continuous-time systems. *Automatica* 50(1): 193–202.

Modares H, Naghibi Sistani MB and Lewis FL (2013b) A policy iteration approach to online optimal control of continuous-time constrained-input systems. *ISA Transactions* 52(5): 611–621.

Mu C, Wang D and He H (2017) Novel iterative neural dynamic programming for data-based approximate optimal control design. *Automatica* 81: 240–252.

Mu C, Wang D and He H (2018) Data-driven finite-horizon approximate optimal control for discrete-time nonlinear systems using iterative HDP approach. *IEEE Transactions on Cybernetics* 48(10): 2948–2961.

Mukhopadhyay V (2003) Historical perspective on analysis and control of aeroelastic responses. *Journal of Guidance, Control, and Dynamics* 26(5): 673–684.

Mukhopadhyay V, Newsom JR and Abel I (1981) A method for obtaining reduced-order control laws for high-order systems using optimization techniques. Technical Report NASA TP 1876, National Aeronautics and Space Administration.

Na J and Herrmann G (2014) Online adaptive approximate optimal tracking control with simplified dual approximation structure for continuous-time unknown nonlinear systems. *IEEE/CAA Journal of Automatica Sinica* 1(4): 412–422.

Nayfeh A, Hammad B and Hajj M (2012) Discretization effects on flutter aspects and control of wing/store configurations. *Journal of Vibration and Control* 18(7): 1043–1055.

Newsom JR and Mukhopadhyay V (1985) A multiloop robust controller design study using singular value gradients. *Journal of Guidance, Control, and Dynamics* 8(4): 514–519.

Nissim E (1971) Flutter suppression using active control based on the concept of aerodynamic energy. Technical Report NASA TN D-6199, National Aeronautics and Space Administration.

NSTC (2010) National aeronautics researchand development plan. Government document, National Science and Technology Council, US, Washington, DC. URL `http://www.whitehouse.gov/sites/default/files/microsites/ostp/aero-rdplan-2010.pdf`.

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 CG, 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.

Pendleton EW, Bessette D, Field PB, Miller GD and Griffin KE (2000) Active aeroelastic wing flight research program: technical program and model analytical development. *Journal of Aircraft* 37(4): 554–561.

Perry IB, Cole SR and Miller GD (1995) Summary of an active flexible wing program. *Journal of Aircraft* 32(1): 10–15.

Petrella R, Tursini M, Peretti L and Zigliotto M (2007) Speed measurement algorithms for low-resolution incremental encoder equipped drives: a comparative analysis. In: *2007 International Aegean Conference on Electrical Machines and Power Electronics*. pp. 780–787.

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.

Popescu C, Wong Y and Lee B (2009) An expert system for predicting nonlinear aeroelastic behavior of an airfoil. *Journal of Sound and Vibration* 319(3): 1312–1329.

Price SJ and Keleris JP (1996) Non-linear dynamics of an airfoil forced to oscillate in dynamic stall. *Journal of Sound and Vibration* 194(2): 265–283.

Prime Z, Cazzolato B, Doolan C and Strganac T (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*. Ph.D. thesis, School of Mechanical Engineering, The University of Adelaide, Adelaide, Australia.

Radac MB, Precup RE and Roman RC (2018) Data-driven model reference control of MIMO vertical tank systems with model-free VRFT and Q-Learning. *ISA Transactions* 73: 227–238.

Rao VM, Behal A, Marzocca P and Rubillo CM (2006) Adaptive aeroelastic vibration suppression of a supersonic airfoil with flap. *Aerospace Science and Technology* 10(4): 309–315.

Rebolho DC, Belo EM and Marques FD (2014) Aeroelastic parameter identification in wind tunnel testing via the extended eigensystem realization algorithm. *Journal of Vibration and Control* 20(11): 1607–1621.

Reddy KK, Chen J, Behal A and Marzocca P (2007) Multi-input/multi-output adaptive output feedback control design for aeroelastic vibration suppression. *Journal of Guidance, Control, and Dynamics* 30(4): 1040–1048.

Reed WH III, Cazier FW Jr and Foughner JT Jr (1980) Passive control of wing/store flutter. Technical Report NASA TM-81865, National Aeronautics and Space Administration.

Rezaei M, Zohoor H and Haddadpour H (2018) Aeroelastic modeling and dynamic analysis of a wind turbine rotor by considering geometric nonlinearities. *Journal of Sound and Vibration* 432: 653–679.

Rugh WJ and Shamma JS (2000) Research on gain scheduling. *Automatica* 36(10): 1401–1425.

Saaed TE, Nikolakopoulos G, Jonasson JE and Hedlund H (2017) A state-of-the-art review of structural control systems. *Journal of Vibration and Control* 21(5): 919–937.

Sales T, Marques FD, Pereira DA and Rade DA (2018) Dynamic assessment of nonlinear typical section aeroviscoelastic systems using fractional derivative-based viscoelastic model. *Journal of Sound and Vibration* 423: 230–245.

Saridis GN and Lee CSG (1979) An approximation theory of optimal control for trainable manipulators. *IEEE Transactions on Systems, Man, and Cybernetics* 9(3): 152–159.

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.

Schmidt DK and Chen TK (1986) Frequency domain synthesis of a robust flutter suppression control law. *Journal of Guidance, Control, and Dynamics* 9(3): 346–351.

Sedaghat A, Cooper JE, Leung AYT and Wright JR (2001) Estimation of the Hopf bifurcation point for aeroelastic systems. *Journal of Sound and Vibration* 248(1): 31–42.

She JH, Fang M, Ohyama Y, Hashimoto H and Wu M (2008) Improving disturbance-rejection performance based on an equivalent-input-disturbance approach. *IEEE Transactions on Industrial Electronics* 55(1): 380–389.

Sheta EF, Harrand VJ, Thompson DE and Strganac TW (2002) Computational and experimental investigation of limit cycle oscillations of nonlinear aeroelastic systems. *Journal of Aircraft* 39(1): 133–141.

Shi T, Wang Z and Xia C (2015) Speed measurement error suppression for PMSM control system using self-adaption Kalman observer. *IEEE Transactions on Industrial Electronics* 62(5): 2753–2763.

Si J, Barto A, Powell W and Wunsch D (2004) *Handbook of Learning and Approximate Dynamic Programming*. Piscataway, NJ: Wiley-IEEE Press.

Simon D (2002) Training radial basis neural networks with the extended kalman filter. *Neurocomputing* 48(1-4): 455–475.

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 R, Lewis FL, Wei Q and Zhang H (2016) Off-policy actor-critic structure for optimal control of unknown systems with disturbances. *IEEE Transactions on Cybernetics* 46(5): 1041–1050.

Song ZG and Li FM (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.

Spencer MG, Sanner RM and Chopra I (1999) Adaptive neurocontrol of simulated rotor vibrations using trailing edge flaps. *Journal of Intelligent Material Systems and Structures* 10(11): 855–871.

Spencer MG, Sanner RM and Chopra I (2002) Closed-loop neurocontroller tests on piezoactuated smart rotor blades in hover. *AIAA Journal* 40(8): 1596–1602.

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.

Sun H, Kim ES, Bernitsas MP and Bernitsas MM (2015) Virtual spring-damping system for flow-induced motion experiments. *Journal of Offshore Mechanics and Arctic Engineering* 137(6).

Sutton RS and Barto AG (1998) *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Tang D, Chen L, Hu E and Tian ZF (2016) A novel actuator controller: Delivering a practical solution to realization of active-truss-based morphing wings. *IEEE Transactions on Industrial Electronics* 63(10): 6226–6237.

Tang D, Chen L and Tian ZF (2015) Neural-network based online policy iteration for continuous-time infinite-horizon optimal control of nonlinear systems. In: *2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)*. Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE), pp. 792–796.

Tang D, Chen L, Tian ZF and Hu E (2018) Adaptive nonlinear optimal control for active suppression of airfoil flutter via a novel neural-network-based controller. *Journal of Vibration and Control* 24(22): 5261–5272.

Tang D, Chen L, Tian ZF and Hu E (2019a) Developing a virtual stiffness-damping system for airfoil aeroelasticity testing. In: *2019 IEEE International Conference on Industrial Technology (ICIT)*. pp. 96–101.

Tang D, Chen L, Tian ZF and Hu E (2019b) Parametric analysis for robust force/torque tracking control of a virtual stiffness-damping system in aeroelastic instability testing. In: *2019 American Control Conference (ACC)*. pp. 3557–3562.

Tang D and Dowell EH (2006) Experimental and theoretical study of gust response for a wing-store model with freeplay. *Journal of Sound and Vibration* 295(3-5): 659–684.

Tang D, Gavin HP and Dowell EH (2004) Study of airfoil gust response alleviation using an electro-magnetic dry friction damper. Part 2: Experiment. *Journal of Sound and Vibration* 269(3-5): 875–897.

Tang DM and Dowell EH (1993) Comparison of theory and experiment for nonlinear flutter and stall response of a helicopter blade. *Journal of Sound and Vibration* 165(2): 251–276.

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.

Vamvoudakis KG and Lewis FL (2012) Online solution of nonlinear two-player zero-sum games using synchronous policy iteration. *International Journal of Robust and Nonlinear Control* 22(13): 1460–1483.

Vamvoudakis KG, Vrabie D and Lewis FL (2014) Online adaptive algorithm for optimal control with integral reinforcement learning. *International Journal of Robust and Nonlinear Control* 24(17): 2686–2710.

Vasconcellos RM, Abdelkefi A, Hajj MR, Almeida DP and Marques FD (2016a) Airfoil control surface discontinuous nonlinearity experimental assessment and numerical model validation. *Journal of Vibration and Control* 22(6): 1633–1644.

Vasconcellos RMG, Pereira DA and Marques FD (2016b) Characterization of nonlinear behavior of an airfoil under stall-induced pitching oscillations. *Journal of Sound and Vibration* 372: 283–298.

Venkatramani J, Nair V, Sujith R, Gupta S and Sarkar S (2017) Multi-fractality in aeroelastic response as a precursor to flutter. *Journal of Sound and Vibration* 386: 390–406.

Venkatramani J, Sarkar S and Gupta S (2018) Investigations on precursor measures for aeroelastic flutter. *Journal of Sound and Vibration* 419: 318–336.

Vipperman JS, Clark RL, Conner M and Dowell EH (1998) Experimental active control of a typical section using a trailing-edge flap. *Journal of Aircraft* 35(2): 224–229.

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.

Vrabie D and Lewis F (2009) Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems. *Neural Networks* 22(3): 237–246.

Vrabie D, Pastravanu O, Abu-Khalaf M and Lewis FL (2009) Adaptive optimal control for continuous-time linear systems based on policy iteration. *Automatica* 45(2): 477–484.

Škach J, Kiumarsi B, Lewis FL and Straka O (2018) Actor-critic off-policy learning for optimal control of multiple-model discrete-time systems. *IEEE Transactions on Cybernetics* 48(1): 29–40.

Wang D, He H and Liu D (2017a) Adaptive critic nonlinear robust control: A survey. *IEEE Transactions on Cybernetics* 47(10): 3429–3451.

Wang D, He H and Liu D (2017b) Improving the critic learning for event-based nonlinear H_∞ control design. *IEEE Transactions on Cybernetics* 47(10): 3417–3428.

Wang D and Liu D (2013) Neuro-optimal control for a class of unknown nonlinear dynamic systems using SN-DHP technique. *Neurocomputing* 121: 218–225.

Wang D, Liu D, Li H and Ma H (2014a) Neural-network-based robust optimal control design for a class of uncertain nonlinear systems via adaptive dynamic programming. *Information Sciences* 282: 167–179.

Wang D, Liu D, Wei Q, Zhao D and Jin N (2012) Optimal control of unknown nonaffine nonlinear discrete-time systems based on adaptive dynamic programming. *Automatica* 48(8): 1825–1832.

Wang D, Liu D, Zhang Q and Zhao D (2016) Data-based adaptive critic designs for nonlinear robust optimal control with uncertain dynamics. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 46(11): 1544–1555.

Wang D, Mu C, Yang X and Liu D (2017c) Event-based constrained robust control of affine systems incorporating an adaptive critic mechanism. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 47(7): 1602–1612.

Wang FY, Zhang H and Liu D (2009) Adaptive dynamic programming: An introduction. *IEEE Computational Intelligence Magazine* 4(2): 39–47.

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 YX, Yu DH and Kim YB (2014b) Robust time-delay control for the DC–DC boost converter. *IEEE Transactions on Industrial Electronics* 61(9): 4829–4837.

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.

Waszak MR (2001) Robust multivariable flutter suppression for benchmark active control technology wind-tunnel model. *Journal of Guidance, Control, and Dynamics* 24(1): 147–153.

Waszak MR and Srinathkumar S (1992) Flutter suppression for the active flexible wing - control system design and experimental validation. Technical Report AIAA Paper 92-2097, American Institute of Aeronautics and Astronautics. Inc.

Wei Q, Lewis FL, Sun Q, Yan P and Song R (2017) Discrete-time deterministic Q-learning: A novel convergence analysis. *IEEE Transactions on Cybernetics* 47(5): 1224–1237.

Wei Q and Liu D (2014) A novel iterative θ-adaptive dynamic programming for discrete-time nonlinear systems. *IEEE Transactions on Automation Science and Engineering* 11(4): 1176–1190.

Wei Q, Liu D and Lin H (2016) Value iteration adaptive dynamic programming for optimal control of discrete-time nonlinear systems. *IEEE Transactions on Cybernetics* 46(3): 840–853.

Wei Q, Wang FY, Liu D and Yang X (2014) Finite-approximation-error-based discrete-time iterative adaptive dynamic programming. *IEEE Transactions on Cybernetics* 44(12): 2820–2833.

Wei X and Chen N (2014) Composite hierarchical anti-disturbance control for nonlinear systems with DOBC and fuzzy control. *International Journal of Robust and Nonlinear Control* 24(2): 362–373.

Werbos PJ (1974) *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Ph.D. thesis, Harvard University.

Werbos PJ (1990) *Neural networks for control*, book section A menu of designs for reinforcement learning over time. Cambridge, MA: MIT Press, pp. 67–95.

Wright JR and Cooper JE (2008) *Introduction to Aircraft Aeroelasticity and Loads*. Chichester, GBR: Wiley.

Xing W and Singh SN (1999) Adaptive control of a nonlinear prototypical wing section with reduced order observer. *Dynamics and Control* 9(4): 297–317.

Xing W and Singh SN (2000) Adaptive output feedback control of a nonlinear aeroelastic structure. *Journal of Guidance, Control, and Dynamics* 23(6): 1109–1116.

Xu X, Wang C and Lewis FL (2014) Some recent advances in learning and adaptation for uncertain feedback control systems. *International Journal of Adaptive Control and Signal Processing* 28(3-5): 201–204.

Yabuno H, Hasegawa M and Ohkuma M (2012) Bifurcation control for a parametrically excited cantilever beam by linear feedback. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 226(8): 1987–1999.

Yang H, Sun J, Xia Y and Zhao L (2018) Position control for magnetic rodless cylinders with strong static friction. *IEEE Transactions on Industrial Electronics* 65(7): 5806–5815.

Yang X, Liu D and Wang D (2014) Reinforcement learning for adaptive optimal control of unknown continuous-time nonlinear systems with input constraints. *International Journal of Control* 87(3): 553–566.

Yu W (2009) *Recent Advances in Intelligent Control Systems*. London: Springer.

Yuan Y, Zhang P, Wang Z, Guo L and Yang H (2017) Active disturbance rejection control for the ranger neutral buoyancy vehicle: A delta operator approach. *IEEE Transactions on Industrial Electronics* 64(12): 9410–9420.

Zhang H, Cui L and Luo Y (2013) Near-optimal control for nonzero-sum differential games of continuous-time nonlinear systems using single-network ADP. *IEEE Transactions on Cybernetics* 43(1): 206–216.

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.

Zhang R and Singh SN (2001) Adaptive output feedback control of an aeroelastic system with unstructured uncertainties. *Journal of Guidance, Control, and Dynamics* 24(3): 502–509.

Zhao D, Xia Z and Wang D (2015a) Model-free optimal control for affine nonlinear systems with convergence analysis. *IEEE Transactions on Automation Science and Engineering* 12(4): 1461–1468.

Zhao L, Yang Y, Xia Y and Liu Z (2015b) Active disturbance rejection position control for a magnetic rodless pneumatic cylinder. *IEEE Transactions on Industrial Electronics* 62(9): 5838–5846.

Zhong X, He H, Wang D and Ni Z (2018) Model-free adaptive control for unknown nonlinear zero-sum differential game. *IEEE Transactions on Cybernetics* 48(5): 1633–1646.