TU Delft

Delft University of Technology

Unified Aeroelastic Flutter and Loads Control via Data-Enabled Policy Optimization

Wang, Xuerui; Zhao, Feiran; Jurisson, Andres; Dorfler, Florian; Smith, Roy S.

DOI
10.1109/TAES.2025.3566351

Publication date
2025

Document Version
Final published version

Published in
IEEE Transactions on Aerospace and Electronic Systems

Citation (APA)
Wang, X., Zhao, F., Jurisson, A., Dorfler, F., & Smith, R. S. (2025). Unified Aeroelastic Flutter and Loads Control via Data-Enabled Policy Optimization. *IEEE Transactions on Aerospace and Electronic Systems*, *61*(5), 11437 - 11449. https://doi.org/10.1109/TAES.2025.3566351

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Green Open Access added to TU Delft Institutional Repository
as part of the Taverne amendment.

More information about this copyright law amendment
can be found at https://www.openaccess.nl.

Otherwise as indicated in the copyright section:
the publisher is the copyright holder of this work and the
author uses the Dutch legislation to make this work public.

Unified Aeroelastic Flutter and Loads Control via Data-Enabled Policy Optimization

XUERUI WANG
Delft University of Technology, Delft, The Netherlands

FEIRAN ZHAO
ETH Zürich, Zürich, Switzerland

ANDRES JÜRISSON
Delft University of Technology, Delft, The Netherlands
Netherlands Aerospace Centre, Amsterdam, The Netherlands

FLORIAN DÖRFLER, Senior Member, IEEE
ROY S. SMITH, Life Fellow, IEEE
ETH Zürich, Zürich, Switzerland

Ultraefficient, high-aspect-ratio wings offer a promising solution for reducing emissions in next-generation aircraft. However, these designs are sensitive to atmospheric disturbances and prone to instability. While active control strategies can mitigate structural loads and stabilize the system, their development is challenging due to the uncertain and time-varying nature of aeroelastic systems. This article addresses these challenges with a direct, adaptive, data-driven approach. The proposed data-enabled policy optimization algorithm leverages sample covariance to directly learn and adapt control strategies from a single batch of persistently exciting, closed-loop input–output data. A forgetting factor mechanism enhances adaptability to time-varying dynamics during operation. The algorithm is explicit and recursive, requiring only a single step of projected gradient descent per sample, improving computational efficiency and enabling real-time application. Numerical simulations demonstrate that the proposed algorithm effectively suppresses unstable flutter, alleviates structural loads, adapts to dynamic time variations, and minimizes control effort—all without requiring prior knowledge of system dynamics or disturbances.

Received 23 December 2024; revised 18 March 2025; accepted 28 April 2025. Date of publication 1 May 2025; date of current version 13 October 2025.

DOI. No. 10.1109/TAES.2025.3566351

Refereeing of this contribution was handled by C.-H. Lee.

This work was supported by Ultra Performance Wing (UP Wing) under Project 101101974 through the Clean Aviation Joint Undertaking and its members. The work of Xuerui Wang was supported by the IDEA League Fellowship.

Authors' addresses: Xuerui Wang is with the Faculty of Aerospace Engineering, Delft University of Technology, 2629HS Delft, The Netherlands, E-mail: (X.Wang-6@tudelft.nl); Feiran Zhao, Florian Dörfler, and Roy S. Smith are with the Department of Information Technology and Electrical Engineering, ETH Zürich, 8092 Zürich, Switzerland, E-mail: (zhaofe@control.ee.ethz.ch; dorfler@ethz.ch; rsmith@control.ee.ethz.ch); Andres Jürisson is with the Faculty of Aerospace Engineering, Delft University of Technology, 2629HS Delft, The Netherlands and also with Netherlands Aerospace Centre, 1059 CM Amsterdam, The Netherlands, E-mail: (Andres.Jurisson@nlr.nl). *(Corresponding author: Xuerui Wang.)*

0018-9251 © 2025 EU

I. INTRODUCTION

Disruptive new aircraft technologies are urgently required by the European Commission to achieve climate neutrality by 2050. To meet this ambitious target, the next generation of short- to medium-range aircraft must reduce net greenhouse gas emissions by at least 30% [1]. This segment constitutes the largest contributor to emissions in commercial air transportation [1]. One promising strategy involves the development of aircraft equipped with high-aspect-ratio wings constructed from ultralightweight materials. This design offers significant advantages, such as improved aerodynamic efficiency and reduced weight. However, it intensifies the interplay between aerodynamic forces and structural elastic dynamics, a phenomenon known as aeroelasticity. Consequently, these wings become more susceptible to atmospheric disturbances and instability phenomena such as flutter [1], [2].

Active control techniques show significant potential for stabilizing aeroelastic systems and reducing structural loads caused by atmospheric disturbances [3], [4]. An appropriately designed control algorithm can utilize distributed onboard sensor data to actuate trailing-edge control surfaces along the wings. This allows local aerodynamic pressures to be manipulated, thereby alleviating loads and suppressing flutter while minimizing control effort to conserve energy.

Achieving these objectives necessitates a comprehensive understanding of the system's dynamics, typically obtained through mathematical modeling. However, the dynamics of an aeroelastic aircraft operating in the transonic regime are inherently complex: they are uncertain, nonlinear, time-varying, and infinite-dimensional [1], [5]. The infinite dimensionality arises from the continuous spectrum of structural vibration dynamics and aerodynamic vortex effects.

Modeling these dynamics usually begins with first-principles-based approaches to establish a model structure, followed by parameter estimation using high-fidelity computational fluid dynamics (CFD) and computational structural dynamics (CSD) simulations [5], [6]. Despite their accuracy, these simulations rely on underlying assumptions and require real-world data for validation and correction, typically sourced from scaled wind tunnel experiments and flight tests. This process is resource-intensive, and the integration of data from diverse sources—often collected under varying conditions—requires significant engineering expertise and iterative refinement.

Despite these efforts, the resulting models are typically nonlinear and high-dimensional. To facilitate real-time

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

control design, these models often undergo significant simplification through order reduction and linearization [7], [8]. However, such simplifications can compromise the stability and robustness of the designed controller. While robust control methods can manage model inaccuracies to some extent, they can be overly conservative [9]. Moreover, conventional aircraft control systems are typically designed using offline models. This poses challenges, as aircraft operate in uncertain environments where factors, such as icing, manufacturing tolerances, buckling, structural maintenance, and unforeseen damage can significantly alter their dynamics [10], [11]. Such variations challenge the effectiveness of traditional offline model-based control approaches.

Up to this point, it has been revealed that for complex dynamic systems, the modeling process can be cumbersome, while the resulting models often exhibit uncertainty and require substantial simplification to enable feasible controller design, thereby creating significant gaps between the model and the actual system. These limitations can be addressed through data-driven control methods, which are broadly categorized into indirect and direct approaches. The indirect approach involves system identification, where a model is derived from offline/online data and subsequently used for model-based controller design. However, this two-step process is computationally intensive, particularly for large-scale systems requiring real-time updates to accommodate dynamic property changes [12]. Conversely, the direct approach bypasses system identification, designing control policies directly from raw data. By leveraging real-time, data-rich information, this approach eliminates reliance on assumptions and simplifications [12]. It inherently captures the influences of uncertainties and disturbances on the system, seamlessly incorporating them into the control design [13]. Furthermore, it avoids the concomitant issues of unmodeled dynamics and robustness concerns inherent in model-based control, making it a more efficient solution for complex systems [14].

Direct data-driven control approaches for aeroelastic or general aerospace systems do exist [15], [16], [17], [18]. However, most of these methods—if not all—face two significant limitations: their reliance on episodic data and their computational inefficiency. The reliance on episodic data means that the control policy is derived from a single episode of offline data or multiple alternating episodes of data collection batches [19], [20]. In other words, the control policy is updated only after an entire episode is complete. This limitation stems from the need for statistically independent data and the methods used for regret analysis [21]. As a result, the dimensionality of the problem scales with the length of the data, and these methods are unable to update the policy in real time using online closed-loop data. This inability contrasts with the online adaptive approach, which allows for continuous policy updates during real-time operations.

Ideally, the control of aeroelastic systems necessitates a direct, data-driven, adaptive, and computationally efficient method with convergence guarantees.

One promising candidate is data-enabled policy optimization (DeePO) recently proposed in [21] and [22]. DeePO belongs to the broader category of policy optimization, which parameterizes the control policy and recursively updates it to minimize a cost function using gradient-based methods [23], [24]. A representative method is zeroth-order policy optimization, a cornerstone of model-free reinforcement learning [25], [26], [27], [28]. While zeroth-order policy optimization is direct and data-driven, it is unsuitable for online adaptive control. This limitation is due to the fact that the cost function required for gradient estimation can only be calculated after observing multiple complete trajectories of sufficient length. Furthermore, achieving an optimal policy using zeroth-order methods requires a prohibitively large number of trajectories [21].

Different from zeroth-order policy optimization, DeePO computes gradients directly from a single trajectory of finite length. This is accomplished by parameterizing the control policy using the sample covariance of online closed-loop data [21], [29]. By performing a single gradient descent per sample, DeePO efficiently updates the control policy from online closed-loop data. Furthermore, the algorithm can be explicitly and recursively implemented via rank-one updates, making it computationally efficient. The direct, data-driven, and online adaptive features of DeePO make it highly suitable for controlling uncertain and time-varying systems subject to unknown stochastic external disturbances. The DeePO method has been extended to linear quadratic tracking problem [30] and applied to power converter systems [31] and an autonomous bicycle with real-world experiments [32].

In this article, we further develop the DeePO method in both theory and application to address the challenge of unified flutter suppression and gust load alleviation in an uncertain and time-varying aeroelastic system.

Our theoretical contributions are twofold. First, we propose a forgetting factor mechanism for the DeePO algorithm to resolve issues of potential performance degradation and numerical stability problems encountered in state-of-the-art methods when applied to fast time-varying dynamic systems. Second, we incorporate the system's direct transmission matrix, which has often been neglected in prior research, into the algorithm's development, resulting in more general expressions and broader applicability.

From a practical relevance standpoint, this work represents the first achievement of unified aeroelastic flutter suppression and gust load alleviation control without relying on any prior models or disturbance knowledge. Moreover, it introduces the first approach capable of addressing the time-varying optimal control problem in an unstable and underactuated aeroelastic system without employing a scheduling parameter methodology. Furthermore, this work pioneers the optimal control of an output-feedback aeroelastic system without requiring explicit state estimation/observation, or system identification. In addition, the effectiveness and efficiency of the proposed algorithm are demonstrated through high-fidelity data-supported simulations.

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

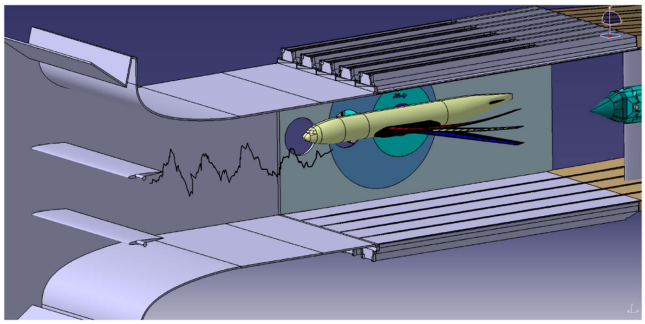

Fig. 1. Illustration of an aeroelastic aircraft subjects to unknown turbulence.

The rest of this article is organized as follows: the aeroelastic control problem is formulated in Section II; the DeePO algorithm is developed in Sections III and IV; the simulation results are presented and discussed in Section V. Finally, Section VI concludes this article.

Notations: In this article, I_k indicates a k-by-k identify matrix; $0_{p\times q}$ indicates a p-by-q zero matrix; $\mathcal{O}^\dagger$ is the left pseudoinverse of $\mathcal{O}$; $\|\cdot\|$ denotes the 2-norm of a matrix or a vector; $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix; Φ_t is the sample covariance of input-state data at t; Ψ_t is the sample covariance of input–output data at t; $\rho_r(\cdot)$ indicates the spectral radius of a matrix; ρ denotes the scaling parameter of a linear parameter-varying system, which is only used for simulation purposes and is unknown by the controller.

II. AEROELASTIC FLUTTER AND LOADS CONTROL

A. System Dynamics

This research considers the next generation of ultrahigh-performance transonic commercial aircraft equipped with high aspect-ratio, slender wings, designed to operate in transonic flow conditions. The slender configuration, along with a lightweight structure, offers significant benefits for enhancing aircraft efficiency, leading to reduced fuel consumption and emissions. However, as a tradeoff, the wing becomes more flexible, with an anticipated tip deflection of 15%–20% of the semispan. Moreover, this increased flexibility makes the wing more sensitive to atmospheric disturbances and more prone to become unstable. Fig. 1 presents an illustration of a flexible aircraft operating in disturbed flow. It also depicts a realistic wind tunnel experiment setup, representing a scaled and mirrored version of real-world conditions. The dynamics of an aeroelastic wing under the excitation of atmospheric disturbances can be modeled as

$$\begin{aligned}\dot{x}(t) &= A^c(t)x(t) + B^c(t)u(t) + B_d^c(t)\alpha_d(t) + w_n(t)\\ y(t) &= C^c(t)x(t) + D^c(t)u(t) + D_d^c(t)\alpha_d(t) + v_n(t)\end{aligned} \quad (1)$$

where $x(t) \in \mathbb{R}^n$ is the aeroelastic state vector. The vector $u(t) \in \mathbb{R}^m$ represents the control input. For the wing considered in this research, $u(t)$ corresponds to the trailing-edge control surfaces, which can actively deflect to manipulate local aerodynamic loads. The vector $\alpha_d(t) \in \mathbb{R}^q$ represents the unknown atmospheric disturbance. $y(t) \in \mathbb{R}$ is the output vector. $w_n(t) \in \mathbb{R}^n$ and $v_n(t) \in \mathbb{R}$ are process and measurement noise, respectively, that are assumed to have zero mean. The superscript $(\cdot)^c$ indicates system dynamic matrices in continuous time.

B. Control Objective and Challenges

From a physical perspective, the objective of the control design is to develop a strategy that optimally drives the trailing-edge control surfaces to stabilize the system and mitigate load variations caused by unknown atmospheric disturbances. This objective can be formulated mathematically as follows:

$$\min_{u(t)} \int_0^\infty \left(\|y(t) - y_{\mathrm{ref}}(t)\|^2 + \|u(t)\|^2\right) \mathrm{d}t \quad (2)$$

where y_{ref} is the reference load vector in trim conditions without disturbances.

The following challenges are identified for this control task.

1) Uncertain/unknown dynamics: The system dynamics matrices $A^c, B^c, C^c, D^c, B_d^c, D_d^c$ are uncertain. In particular, accurately modeling the dynamic influence matrices of atmospheric disturbances, B_d^c and D_d^c, is especially challenging. Assumptions must often be made about the dimensionality and uniform distribution of disturbances, which may differ in real-world conditions.
2) Time-varying dynamics: The system dynamics matrices $A^c, B^c, C^c, D^c, B_d^c, D_d^c$ are also time-varying, with their values changing under different operational conditions, such as airspeed and altitude. Even temperature impacts system dynamics through its influence on the Reynolds number.
3) Hard-to-identify dynamics: The influence matrices of atmospheric disturbances, B_d^c and D_d^c, are not only difficult to model accurately but also challenging to identify. System identification typically requires knowledge of both inputs and outputs. However, the dynamics of atmospheric disturbances are typically stochastic, time-varying, and difficult to be measured accurately.
4) High dimensionality: In theory, $x(t)$ possesses infinite dimensionality due to the infinite spectrum of structural modes and aerodynamic vortices. However, in practice, $x(t)$ is often truncated to a finite dimension up to a certain aeroelastic frequency limit. Despite this truncation, $x(t)$ remains high-dimensional, imposing challenges on computational efficiency and numerical stability.
5) Unmeasurable states: $x(t)$, the coupled state vector of structural modes and aerodynamic vortices, is not directly measurable. Consequently, the control algorithm must rely exclusively on the input–output information.
6) Real-time implementation: The control algorithm should be sufficiently computationally efficient

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

to enable real-time implementation on realistic hardware with limited processing power.

C. Problem Formulation

In light of these challenges, this article proposes a DeePO algorithm to directly and adaptively stabilize and dampen the uncertain time-varying dynamics without requiring any prior knowledge of the dynamic model or disturbances.

Since B_d^c and D_d^c are challenging to model or identify, and the information of atmospheric disturbances are generally unavailable for control design in real-time, the entirety of the terms $B_d^c(t)\alpha_d(t)$ and $D_d^c(t)\alpha_d(t)$ are considered to be completely unknown. Consequently, they are incorporated into the process and measurement noise in (1) as

$$\begin{aligned} w(t) &:= B_d^c(t)\alpha_d(t) + w_n(t) \\ v(t) &:= D_d^c(t)\alpha_d(t) + v_n(t). \end{aligned} \tag{3}$$

Consequently, for the purpose of control design, the continuous-time state-space system in (1) can be discretized via the bilinear transform as

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t \\ y_t &= Cx_t + Du_t + v_t \end{aligned} \tag{4}$$

where A, B, C, D are the corresponding dynamic matrices in the discrete time. They are treated as completely unknown in this research.

Considering the discrete-time system in (4), when y_{ref} is set to zero in (2), we formulate the weighted quadratic cost as

$$\limsup_{T\to\infty} \mathbb{E}_{\{w_t\}} \left[\frac{1}{T} \sum_{t=0}^{T-1} \left(u_t^\top R u_t + x_t^\top Q x_t \right) \right] \tag{5}$$

where the weighting matrices Q and R are positive definite. There exists a control policy $u_t = Kx_t$ which stabilizes (4). When the system dynamic matrices A, B are known, K can be directly solved based on the Riccati equation. However, when A and B are unknown, uncertain, and time-varying, although online system identification can be performed, following the identification, the Riccati equation has to be solved accordingly at every time step to find out the optimal gain matrix, which is computational intensive [21]. Instead, we parameterize the cost as

$$C(K) := \operatorname{Tr}\left(\left(Q + K^\top R K \right) \Sigma_K \right) \tag{6}$$

in which Σ_K represents the closed-loop state covariance matrix, determined as the positive definite solution of the Lyapunov equation

$$\Sigma_K = I_n + (A + BK)\Sigma_K (A + BK)^\top. \tag{7}$$

Equations (6) and (7) are referred to as a policy parameterization of the quadratic regulation problem in (5). In the following section, the algorithm for solving this data-driven problem will be developed.

III. DATA-ENABLED POLICY OPTIMIZATION

In this section, we introduce and further develop the DeePO algorithm, which enables adaptive and recursive learning of policy directly from online closed-loop data. First, the direct data-driven policy parameterization, based on sample covariance, is formulated in Section III-A. Then, the DeePO method for direct adaptive control is derived in Section III-B.

A. Policy Parameterization Based on Sample Covariance

Assume the full state information of the system is known, or is provided by a state observer. Then, based on (4), the following discrete-time data-driven system dynamics holds:

$$X_{1,t} = AX_{0,t} + BU_{0,t} + W_{0,t} \tag{8}$$

where $X_{0,t} := [x_0, \ldots, x_{t-1}] \in \mathbb{R}^{n\times t}$, $U_{0,t} := [u_0, \ldots, u_{t-1}] \in \mathbb{R}^{m\times t}$, $W_{0,t} := [w_0, \ldots, w_{t-1}] \in \mathbb{R}^{n\times t}$ are the t-long time series of states, inputs, and process noises, respectively. $X_{1,t} := [x_1, \ldots, x_t] \in \mathbb{R}^{n\times t}$ is the successor state vector. Under the common assumption of persistent excitation (PE), the data block matrix $D_{0,t} := [U_{0,t}^\top, X_{0,t}^\top]^\top$ has full row rank. Denote Φ_t as the sample covariance of input-state data: $\Phi_t := \frac{1}{t} D_{0,t} D_{0,t}^\top$. Then, under the PE assumption, Φ_t is positive definite, while there exists a unique parameterized policy $V \in \mathbb{R}^{(n+m)\times n}$ for any given feedback gain matrix K satisfying

$$\begin{bmatrix} K \\ I_n \end{bmatrix} = \frac{1}{t} D_{0,t} D_{0,t}^\top V =: \Phi_t V. \tag{9}$$

It is noteworthy that the dimension of V does not depend on the data length. The control problem can be represented using raw data and V, with the closed-loop system dynamic matrix given as

$$A + BK = [B, A] \begin{bmatrix} K \\ I_n \end{bmatrix} \overset{(9)}{=} [B, A]\Phi_t V \overset{(8)}{=} (\overline{X}_{1,t} - \overline{W}_{0,t})V \tag{10}$$

where $\overline{X}_{0,t} = X_{0,t} D_{0,t}^\top / t$, $\overline{U}_{0,t} = U_{0,t} D_{0,t}^\top / t$, $\overline{W}_{0,t} = W_{0,t} D_{0,t}^\top / t$ are sample matrices.

Using the certainty-equivalence principle [33], the unmeasurable uncertainty $\overline{W}_{0,t}$ will be disregarded, leading to the direct data-driven problem formulation

$$\begin{aligned} &\underset{V}{\text{minimize}}\ J_t(V) := \operatorname{Tr}\left((Q + V^\top \overline{U}_{0,t}^\top R \overline{U}_{0,t} V) \Sigma_t(V) \right) \\ &\text{subject to } \overline{X}_{0,t} V = I_n \end{aligned} \tag{11}$$

where $\Sigma_t(V) = I_n + \overline{X}_{1,t} V \Sigma_t(V) V^\top \overline{X}_{1,t}^\top$. The closed-loop gain is recovered as $K = \overline{U}_{0,t} V$.

REMARK 1 The direct data-driven covariance parameterization of the problem can be proven to have an equivalent solution to the indirect certainty-equivalence linear quadratic regulator problem [21, Lemma 1]. However, the latter is computationally demanding, as it requires solving a least-square system identification and an algebraic Riccati

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

equation at every time step. This will be avoided by the DeePO algorithm presented in the following sections.

B. DeePO for Direct Adaptive Control

The objective of the algorithm design is to determine the optimal gain matrix directly from online closed-loop data, without relying on prior model knowledge or disturbance information. The algorithm is designed to be both direct, meaning it bypasses the need for a system model, and adaptive, implying that it adjusts the control policy dynamically to real-world scenarios based on fresh online data.

The algorithm is presented in Algorithm 1. The initial policy can be set to zero if the system is open-loop stable. Otherwise, it can be determined, for example, using semidefinite programming (SDP) [34] based on offline data. The length of the offline data t_0 should be sufficiently long to ensure that the offline data matrix attains full row rank. Since $[U_{0,t}^\top, X_{0,t}^\top]^\top$ has a size of $(m+n)\times t$, the offline data length t_0 should be at least $m+n$. At every time step, the control policy is defined as $u_t = K_t x_t + e_t$, where a probing noise e_t is added to satisfy the PE condition. Given a gain matrix, the policy is parameterized as outlined in Section III-A and is then updated using one-step projected gradient descent, as shown in (12). The projection $\Pi_{\overline{X}_{0,t+1}}$ is defined as $\Pi_{\overline{X}_{0,t+1}} := I_{n+m} - \overline{X}_{0,t+1}^\dagger \overline{X}_{0,t+1}$, which is used to enforce the subspace constraint $\overline{X}_{0,t} V = I_n$. The gradient $\nabla J_{t+1}(V_{t+1})$ is calculated with P_{t+1}, which satisfies the Lyapunov equation $P_{t+1} = Q + V_{t+1}^\top \overline{U}_{0,t+1}^\top R \overline{U}_{0,t+1} V_{t+1} + V_{t+1}^\top \overline{X}_{1,t+1}^\top P_{t+1} \overline{X}_{1,t+1} V_{t+1}$. At the final step, the control gain matrix is recovered from V_{t+1}.

Algorithm 1 exhibits several key features. First, it guarantees global convergence under the assumptions of persistence of excitation, a bounded signal-to-noise ratio, and an initial parameterized policy V^0 within a feasible set [21, Th. 1]. Second, Algorithm 1 can be proven to provide convergence guarantees that are independent of noise statistics [21, Th. 2]. These guarantees establish performance bounds that are valid for a finite, and often limited, number of data samples—an essential characteristic for practical algorithm design and deployment. Third, Algorithm 1 does not require storing historical data and can be efficiently and recursively implemented using rank-one updates. This efficiency applies to all data matrices, covariance parameterization, and the parameterized policy, which are critical for practical implementation and deployment. In particular, explicit matrix inversion of Φ_{t+1} is not needed at every time step. Instead, it is efficiently computed via rank-one update based on the Sherman–Morrison formula [35] as

$$\Phi_{t+1}^{-1} = \frac{t+1}{t}\left(\Phi_t^{-1} - \frac{\Phi_t^{-1}\psi_t\psi_t^\top\Phi_t^{-1}}{t+\psi_t^\top\Phi_t^{-1}\psi_t}\right)$$

where $\psi_t := [u_t^\top, x_t^\top]^\top$.

Since only a single-step gradient descent per time step is required, DeePO is computationally more efficient than the indirect certainty-equivalence linear quadratic regulator, which necessitates solving a Riccati equation at each step. A quantitative comparison in [21] has demonstrated that, for the same number of computational steps, DeePO requires less wall time, and this advantage becomes more pronounced as the system order increases. Moreover, to achieve the same performance gap, the computational time is similar for both methods when the performance gap is large. However, as the performance gap decreases, the wall time of DeePO becomes significantly smaller, further verifying its computational efficiency.

Algorithm 1: DeePO for Direct Adaptive Control.

Require: Offline data $(X_{0,t_0}, U_{0,t_0}, X_{1,t_0})$, an initial policy K_{t_0}, and a constant stepsize η.

1: **for** $t = t_0, t_0+1, \ldots$ **do**

2: Apply $u_t = K_t x_t + e_t$ and observe x_{t+1}.

3: Update data matrices Φ_{t+1} and $\overline{X}_{1,t+1}$.

4: **Policy parameterization:** given K_t, solve V_{t+1} via

$$V_{t+1} = \Phi_{t+1}^{-1}\begin{bmatrix} K_t \\ I_n \end{bmatrix}.$$

5: **Update of the parameterized policy:** perform one-step projected gradient descent

$$V'_{t+1} = V_{t+1} - \eta \Pi_{\overline{X}_{0,t+1}} \nabla J_{t+1}(V_{t+1}), \quad (12)$$

in which the gradient term $\nabla J_{t+1}(V_{t+1})$ equals $2(\overline{U}_{0,t+1}^\top R \overline{U}_{0,t+1} + \overline{X}_{1,t+1}^\top P_{t+1} \overline{X}_{1,t+1}) V_{t+1} \Sigma_{t+1}(V_{t+1})$, and P_{t+1} satisfies the Lyapunov equation $P_{t+1} = Q + V_{t+1}^\top \overline{U}_{0,t+1}^\top R \overline{U}_{0,t+1} V_{t+1} + V_{t+1}^\top \overline{X}_{1,t+1}^\top P_{t+1} \overline{X}_{1,t+1} V_{t+1}$.

6: **Gain update:** update the control gain by

$$K_{t+1} = \overline{U}_{0,t+1} V'_{t+1}.$$

7: **end for**

IV. DEEPO WITH FORGETTING DATA FOR OUTPUT-FEEDBACK SYSTEMS

The DeePO algorithm developed in the previous section depends on the availability of state information. We eliminate this dependency entirely in Section IV-A. The design objective is to achieve direct adaptive control using only input–output data, without requiring any model or state information. Moreover, we propose a forgetting factor mechanism to account for time-varying dynamics in Section IV-B.

A. DeePO for Single-Output Feedback Systems

Denote the input and output trajectories and their stacked from time $t-h$ to $t-1$ as follows:

$$u_{t,h} = \begin{bmatrix} u_{t-1} \\ \vdots \\ u_{t-h} \end{bmatrix}, y_{t,h} = \begin{bmatrix} y_{t-1} \\ \vdots \\ y_{t-h} \end{bmatrix}, z_t = \begin{bmatrix} u_{t,h} \\ y_{t,h} \end{bmatrix}.$$

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

For the system in (4), define the observability and controllability matrices as $\mathcal{O} = [(CA^{h-1})^\top, \ldots, (CA)^\top, C^\top]^\top$ and $\mathcal{C} = [B, AB, \ldots, A^{h-1}B]$, respectively. Here, the length h is larger equal to the system lag l, which is also known as the observability index. The lag or observability index l is the smallest number of rows of the observability matrix $\mathcal{O}$ required to span its entire row space. For a system with p output, the following inequality holds: $l \le n \le pl$ [36]. In this article, we consider a special case of $p = 1$, this leads to $l = n$. Correspondingly, the online data length is chosen as $h = l = n$. Define $\mathcal{C}_w = [I, A, \ldots, A^{h-1}]$ and define $\mathcal{T}$, $\mathcal{T}_w$ as Toeplitz matrices

$$\mathcal{T} = \begin{bmatrix} D & CB & CAB & \cdots & CA^{h-2}B \\ 0 & D & CB & \cdots & CA^{h-3}B \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & & D & CB \\ 0 & 0 & 0 & 0 & D \end{bmatrix}$$

$$\mathcal{T}_w = \begin{bmatrix} 0 & C & CA & \cdots & CA^{h-2} \\ 0 & 0 & C & \cdots & CA^{h-3} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & & 0 & C \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

THEOREM IV.1 *(Nonminimal controllable realization)* If the original system in (4) is fully observable and controllable, then for $h = l = n$ and $p = 1$, the realization in (13), derived from input–output data, constitutes a nonminimal but controllable realization of (4)

$$z_{t+1} = \begin{bmatrix} 0_{m\times(m+1)n} \\ \hdashline I_{m(n-1)} \vdots 0_{m(n-1)\times m} \vdots 0_{m(n-1)\times n} \\ \hdashline S \\ \hdashline 0_{(n-1)\times mn} \vdots I_{n-1} \vdots 0_{(n-1)\times 1} \end{bmatrix} z_t + \begin{bmatrix} I_m \\ \hdashline 0_{m(n-1)\times m} \\ \hdashline D \\ \hdashline 0_{(n-1)\times m} \end{bmatrix} u_t + \begin{bmatrix} 0_{m\times 1} \\ \hdashline 0_{m(n-1)\times 1} \\ \hdashline I_1 \\ \hdashline 0_{(n-1)\times 1} \end{bmatrix} d_t$$

$$y_t = Sz_t + Du_t + d_t, \tag{13}$$

where $S = [C(\mathcal{C} - A^n\mathcal{O}^\dagger\mathcal{T}), CA^n\mathcal{O}^\dagger]$ and $d_t = C(\mathcal{C}_w - A^n\mathcal{O}^\dagger\mathcal{T}_w)w_{t,n} - CA^n\mathcal{O}^\dagger v_{t,n} + v_t$.

PROOF Using historical system trajectory, the system in (4) can be realized as

$$\begin{aligned} x_t &= A^h x_{t-h} + \mathcal{C}u_{t,h} + \mathcal{C}_w w_{t,h} \\ y_{t,h} &= \mathcal{O}x_{t-h} + \mathcal{T}u_{t,h} + \mathcal{T}_w w_{t,h} + v_{t,h}. \end{aligned} \tag{14}$$

When $h = l = n$, the observability matrix $\mathcal{O}$ has full column rank. Therefore, denoting $\mathcal{O}^\dagger = (\mathcal{O}^\top\mathcal{O})^{-1}\mathcal{O}^\top$, then (4) and (14) lead to

$$\begin{aligned} x_t &= (\mathcal{C} - A^n\mathcal{O}^\dagger\mathcal{T})u_{t,n} + A^n\mathcal{O}^\dagger y_{t,n} \\ &\quad + (\mathcal{C}_w - A^n\mathcal{O}^\dagger\mathcal{T}_w)w_{t,n} - A^n\mathcal{O}^\dagger v_{t,n} \\ y_t &= Cx_t + Du_t + v_t = Sz_t + Du_t + d_t \end{aligned} \tag{15}$$

which can be realized as in (13).

When $p = 1$, it follows from the reachability lemma [37, Lemma 3.4.7] that the nonminimal realization (13) is controllable. □

REMARK 2 In [21], the direct transmission matrix D is set to zero. However, this term can be important in transonic unsteady aerodynamic conditions [38]. Thus, we consider nonzero D, leading to more general expressions of $\mathcal{T}$ and (13) and (15).

For the physical aeroelastic wing dynamics considered in this article, the direct transmission matrix D arises due to the unsteady aerodynamic effects of control inputs. In the practical control design process, if the actuator bandwidth is not sufficiently higher than the system dynamics, the actuator dynamics can be augmented into the original system dynamics. If the actuator behaves as a typical mechanical or electromechanical system (e.g., dc motors, hydraulic actuators, or pneumatic actuators), this augmentation typically removes the direct feed-through term D. However, if the actuator is a high-bandwidth electronic system, such as an operational amplifier, a piezoelectric actuator, or a system with very low inertia, the augmented system may still exhibit direct feed-through.

The vector z_t consists of historical input and output data, which is available. Define $Z_{0,t} = [z_0, z_1, \ldots, z_{t-1}] \in \mathbb{R}^{h(m+1)\times t}$ and $\overline{Z}_{0,t} = Z_{0,t}[U_{0,t}^\top, X_{0,t}^\top]/t$. Denote Ψ_t as the sample covariance of input–output data

$$\Psi_t := \frac{1}{t}\begin{bmatrix} U_{0,t} \\ Z_{0,t} \end{bmatrix}\begin{bmatrix} U_{0,t} \\ Z_{0,t} \end{bmatrix}^\top.$$

Then, under the PE condition, the DeePO algorithm for output-feedback systems is presented in Algorithm 2. Since in the output feedback case, the data matrix $[U_{0,t}^\top, Z_{0,t}^\top]^\top$ has a size of $(h(m+p)+m)\times t$, the offline data length t_0 should be at least $h(m+p)+m$ to ensure the offline data matrix has full row rank.

B. DeePO With Forgetting Factor for Time-Varying Dynamics

To account for time-varying system dynamics, a forgetting factor $\lambda \in (0, 1)$ is introduced in this article. The covariance of exponentially weighted data is defined as

$$\widetilde{\Lambda}_t := \frac{1}{t}D_{0,t}S_\lambda D_{0,t}^\top$$

where $S_\lambda := \mathrm{diag}\{\lambda^{t-1}, \lambda^{t-2}, \ldots, 1\} \succ 0$. Since $\lambda \in (0, 1)$, the design of S_λ imposes no penalty on the most recent data sample u_{t-1} and x_{t-1}, while the oldest data u_0 and x_0 are penalized the most with a weight of λ^{t-1}.

As discussed in Section III-B, under PE condition, $D_{0,t}$ has full row rank. Therefore, $\widetilde{\Lambda}_t$ also becomes positive definite, and the covariance parameterization in (9) can be modified to

$$\begin{bmatrix} K \\ I_n \end{bmatrix} = \widetilde{\Lambda}_t V. \tag{16}$$

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

Algorithm 2: DeePO for Output-Feedback Systems.

Require: Offline data $(Z_{0,t_0}, U_{0,t_0}, Z_{1,t_0})$, an initial policy K_{t_0}, and a stepsize η.

1: **for** $t = t_0, t_0+1, \ldots$ **do**

2: Apply $u_t = K_t z_t + e_t$ and observe z_{t+1}.

3: Update data matrices Ψ_{t+1} and $\overline{Z}_{1,t+1}$.

4: **Policy parameterization:** given K_t, solve V_{t+1} via

$$V_{t+1} = \Psi_{t+1}^{-1} \begin{bmatrix} K_t \\ I_n \end{bmatrix}.$$

5: **Update of the parameterized policy:** perform one-step projected gradient descent

$$V'_{t+1} = V_{t+1} - \eta \Pi_{\overline{Z}_{0,t+1}} \nabla J_{t+1}(V_{t+1}).$$

in which the gradient term $\nabla J_{t+1}(V_{t+1})$ equals $2(\overline{U}_{0,t+1}^\top R \overline{U}_{0,t+1} + \overline{Z}_{1,t+1}^\top P_{t+1} \overline{Z}_{1,t+1}) V_{t+1} \Sigma_{t+1}(V_{t+1})$, and P_{t+1} satisfies the Lyapunov equation $P_{t+1} = Q + V_{t+1}^\top \overline{U}_{0,t+1}^\top R \overline{U}_{0,t+1} V_{t+1} + V_{t+1}^\top \overline{Z}_{1,t+1}^\top P_{t+1} \overline{Z}_{1,t+1} V_{t+1}$.

6: **Gain update:** update the control gain by

$$K_{t+1} = \overline{U}_{0,t+1} V'_{t+1}.$$

7: **end for**

With the new parameterization in (16), it follows from (11) that:

$$\begin{aligned} &\underset{V}{\text{minimize}}\ J_t(V) := \text{Tr}\left((Q + V^\top \widetilde{U}_{0,t}^\top R \widetilde{U}_{0,t} V) \Sigma_t(V)\right) \\ &\text{subject to } \widetilde{X}_{0,t} V = I_n \end{aligned} \tag{17}$$

where $\Sigma_t(V) = I_n + \widetilde{X}_{1,t} V \Sigma_t(V) V^\top \widetilde{X}_{1,t}^\top$. The exponential weighted data matrices are $\widetilde{X}_{0,t} = X_{0,t} S_\lambda D_{0,t}^\top / t$, $\widetilde{U}_{0,t} = U_{0,t} S_\lambda D_{0,t}^\top / t$, $\widetilde{W}_{0,t} = W_{0,t} S_\lambda D_{0,t}^\top / t$, $\widetilde{X}_{1,t} = X_{1,t} S_\lambda D_{0,t}^\top / t$. From here, Algorithms 1 and 2 can be applied directly. When the system dynamics are linear time-invariant (LTI), similar to the proofs of [21, Ths. 1 and 2], Algorithms 1 and 2 can be proven to guarantee global stability and convergence. Providing theoretical guarantees for linear parameter-varying (LPV) systems remain an open research challenge in the field. The primary difficulty arises from the coupling between the rate of system dynamics variation and the rate of policy updates. Qualitatively, the policy update rate must be sufficiently high to ensure timely adaptation to changes in system dynamics.

The computational efficiency advantage over the indirect certainty-equivalence control for adaptive learning of the linear quadratic regulator remains valid in the time-varying and output feedback cases. This is because the proposed algorithm still requires only a single-step gradient descent per time step and avoids solving a Riccati equation at every step. Due to the data-driven nature of the proposed algorithm, the rate of system dynamic variations influences system stability and convergence rate but does not significantly affect computational load. In other words, the actual elapsed time required to execute a given number of steps is primarily determined by the system size rather than the time-varying nature of the dynamics. Whether the algorithm employs output feedback or state feedback is also not a crucial factor for computational loads, as Algorithm 2 and Theorem 4.1 in this article effectively transform the output feedback system into an equivalent system where full state information is available.

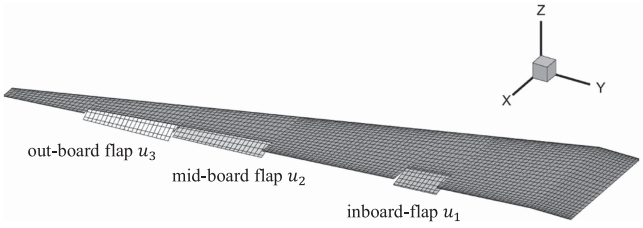

Fig. 2. Aerodynamic mesh for unsteady aerodynamic modeling and the locations of control surfaces.

V. RESULTS AND DISCUSSIONS

A. Aeroelastic Wing Modeling

The model used to validate the proposed control algorithms is a transonic aeroelastic wing model developed within the Clean Aviation project "Ultraperformance Wing (UPwing)", cofunded by the European Union [1]. The objective of this project is to innovate ultraperformance wing concepts for short-to medium-range aircraft, a class with significant impacts on transport aircraft emissions.

To achieve this objective, a wind-tunnel-scaled model of the transonic aeroelastic wing was developed to validate these innovative technologies in the German-Dutch High-Speed Tunnel (DNW-HST). The wing has a trapezoidal shape, with a mean aerodynamic chord of 0.25 m, a half-wing span of 1.34 m, and a surface area of 0.35 m^2. It is designed to operate at a Mach number of 0.78 in an air density of 0.9 kg/m^3. The geometry is derived from the full-scale reference aircraft DLR-F25 [38].

An aerodynamic model of the wing is obtained using the panel method. The corresponding aerodynamic mesh, along with the three trailing-edge control surfaces for active control, is shown in Fig. 2. Compressible effects at high Mach numbers are accounted for using Prandtl–Glauert corrections. Unsteady aerodynamic effects are modeled through a rational function approximation method, introducing six aerodynamic lag states per structural mode. Under the design condition of a lift coefficient of $C_L = 0.5$, a total static load of approximately 4100 N is imposed on the model.

Regarding the structural model of the wing, finite element analysis is conducted. Fig. 3 shows the structural mesh and the locations of linear accelerometers. The wing structure comprises glass-fiber-reinforced plastic skins with a foam core. A lumped mass at the wingtip explores potential reduction of frequency in the first structural mode.

The aerodynamic and structural models are coupled through a set of dedicated coupling nodes. The first 12

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

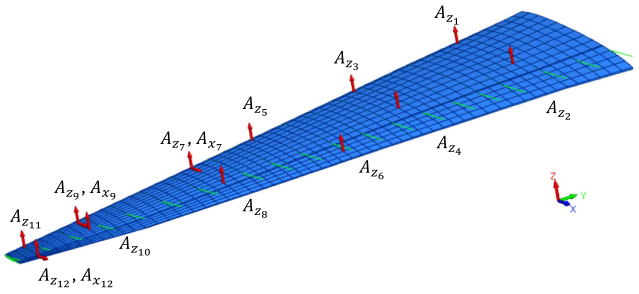

Fig. 3. Finite element structural mesh and the locations of the accelerometers.

structural modes are considered, covering frequencies up to 360 Hz. Two continuous state-space aeroelastic models of the transonic wing are used in this research: one at an airspeed of 240 m/s and the other at 257 m/s. Each model incorporates 96 aeroelastic states, three trailing-edge control inputs (see Fig. 2), and 21 outputs (see Fig. 3). The outputs include three-axis forces and moments at the wing root, 12 linear acceleration measurements along the z-axis (vertical), and three linear acceleration measurements along the x-axis (horizontal). More details of the wing can be found in our previous publication [38].

B. Model Preprocessing

In theory, the model of a transonic aeroelastic wing possesses infinite dimensionality, resulting from the infinite spectrum of structural and aerodynamic modes. However, physical actuators have limited bandwidth, beyond which the system properties become difficult to modify through closed-loop control. Consequently, model order reduction is a necessary first step in model preprocessing. Notably, the DeePO algorithm can process input–output data directly without requiring prior model knowledge. Nonetheless, model order reduction remains beneficial for enhancing computational efficiency, real-time applicability, and numerical stability.

In this research, the balanced truncation method is employed to reduce the model from 96 states to 10 states. The reduction order is determined by analyzing the singular values of a Hankel matrix derived from the system's input–output behavior and by considering the actuator bandwidth. Figs. 4 and 5 compare the eigenvalues and open-loop responses of the full-order and reduced-order systems. The eigenvalues correspond well in the low-frequency range. It is also evident that some higher-frequency, lower-damping aeroelastic modes contribute more significantly to the input–output mapping than the lower-frequency modes along the real axis.

Regarding response comparisons, when $V = 240$ m/s, the root mean square (rms) differences between the full-order and reduced-order models are within 1.18 N (0.37 %) and 1.15 N·m (0.81 %) for wing root shear force and root bending moment, respectively. At $V = 257$ m/s, where the system exhibits inherent instability (flutter), the full and

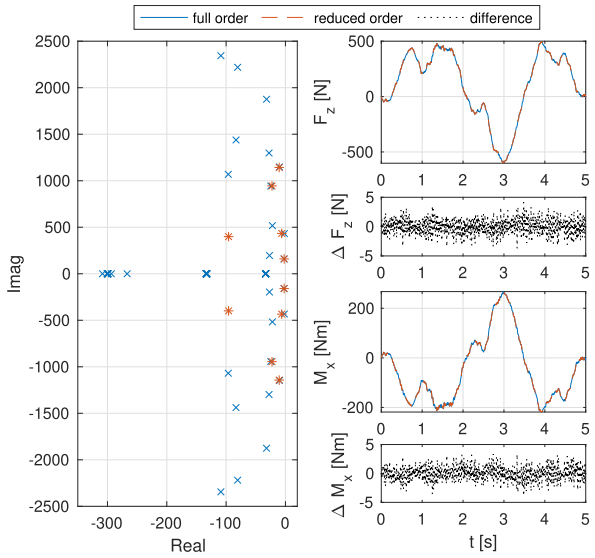

Fig. 4. Comparisons of eigenvalues [full-order (x), reduced-order (*)] and comparisons of responses (wing root shear force F_z and bending moment M_x) when $V = 240$ m/s.

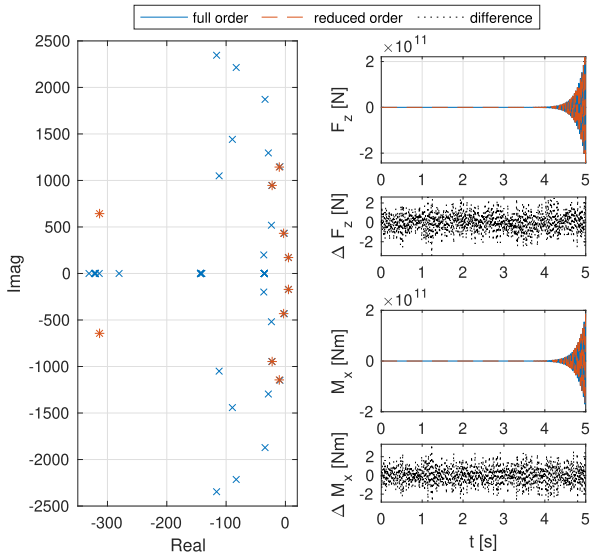

Fig. 5. Comparisons of eigenvalues [full-order (x), reduced-order (*)], and comparisons of responses (wing root shear force F_z and bending moment M_x) when $V = 257$ m/s.

reduced models remain consistent, with rms differences below 0.92 N and 0.92 N·m, respectively.

The second step in model preprocessing is discretization. The reduced-order continuous-time state-space systems are discretized using the bilinear transform. The sampling frequency is determined based on the Nyquist criterion. The highest mode frequencies for the stable and unstable systems are 181.91 and 182.08 Hz, respectively, leading to a choice of 0.002 as the sampling step. This value is also practical for real-world hardware implementations.

The final step is scaling, which is crucial for physical systems, especially when some input and output channels have different units (e.g., N, N·m, and m/s^2 in this case). This results in state, input, and output matrices with varying

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

orders of magnitude. To address this, a similarity transformation is applied to improve the system's condition number, thereby enhancing numerical stability.

These three steps of model preprocessing produce two LTI systems, each with ten states, three inputs, and 21 outputs. Both systems are controllable and observable. The stable system at $V = 240$ m/s will be used in Sections V-C and V-D for disturbance rejection tasks. These two systems at different velocities are then transformed into an LPV system to be used in Section V-E for simultaneous flutter suppression (instability stabilization) and load alleviation.

C. State Feedback

The data-driven disturbance rejection capability of the DeePO algorithm is evaluated in this section. Both offline and online data are generated using a 10-state model at $V = 240$ m/s. The algorithm is not aware of any model information, expect for the dimension of the data. It is assumed that state information is provided by a state observer, which is a common assumption in the aeroelasticity control community. Output feedback, where state information is not directly measured, will be addressed in the following subsections. Since the system is open-loop stable, the initial gain matrix is set to zero, meaning the algorithm does not rely on an initial policy learned from offline data. The offline data length is set to $t_0 = 20$, which satisfies $t_0 \geq m + n$. To satisfy the PE condition, white process and measurement noise are added. Their amplitudes are tuned based on real-world sensor properties and to ensure a proper signal-to-noise ratio.

The choice of the step size for the projected gradient descent is primarily influenced by system dimensions (the number of states, inputs, and outputs), controllability, the rate of system parameter variations, and noise levels. Qualitatively, the step size should be reduced as the system dimensions increase, the degree of controllability decreases (i.e., when the condition number of the controllability Gramian is larger), the rate of system parameter variations increases, and the noise level increases. A larger step size accelerates convergence but may induce oscillations or divergence. A smaller step size improves stability but may lead to slow convergence. In the full state feedback case, a constant step size is chosen as $\eta = 0.1$ to balance numerical stability and convergence rate.

The relative values of the Q and R matrices represent a tradeoff between control effort and performance. A higher Q emphasizes state regulation, leading to improved disturbance rejection but potentially increasing control effort. Conversely, a higher R prioritizes minimizing control input, which reduces actuation effort but may result in a less responsive system. In this state-feedback case, Q is selected as an identity matrix with a magnitude of 100, while R is an identity matrix with a magnitude of 10^{-4}. For a selected pair of Q and R, the effectiveness of the algorithm can be evaluated using the optimality gap, which quantifies the deviation of a given control strategy from the best possible

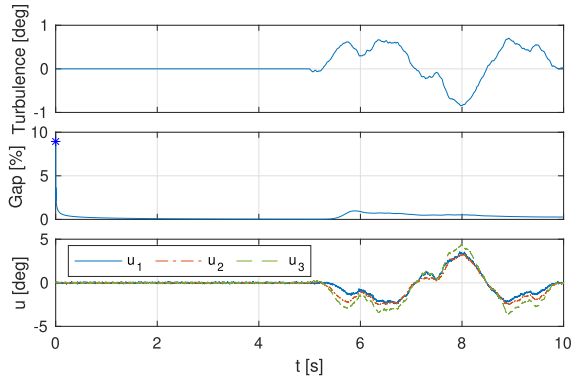

Fig. 6. Evolution of an unknown Dryden turbulence, optimality gap, and control inputs. The blue asterisk indicates the initial value of the optimality gap.

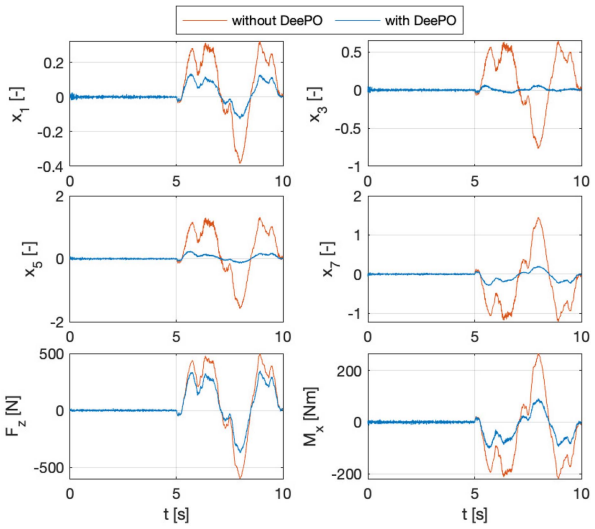

Fig. 7. Comparison of open-loop and closed-loop responses of states and outputs under the excitation of unknown external turbulence.

outcome in terms of the specified objective or cost function. The optimality gap is expressed as $\frac{C(K)-J^*}{J^*}$, where $C(K)$ is given in (6) while J^* represents the best possible cost achieved with perfect knowledge of A and B. The time history of the optimality gap can also reflect the convergence rate of the algorithm.

The performance of the DeePO algorithm is illustrated in Figs. 6 and 7. As shown, the optimality gap initially drops sharply from 8.95 to 0.15 within the first second and then gradually converges to 0.02. The first subplot of Fig. 6 shows the angle of attack (in degrees) induced by an unknown Dryden turbulence disturbance. The turbulence introduced from $t = 5$ s leads to a temporarily increasing optimality gap, after which it adapts and resumes a decreasing trend. All three control surface deflections remain within the bounds of ± 10 degrees, actively counteracting the effects of turbulence on the aeroelastic wing. Fig. 7 displays the system's state and output responses, clearly showing that the state responses are dampened by the algorithm. The peak wing root bending

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

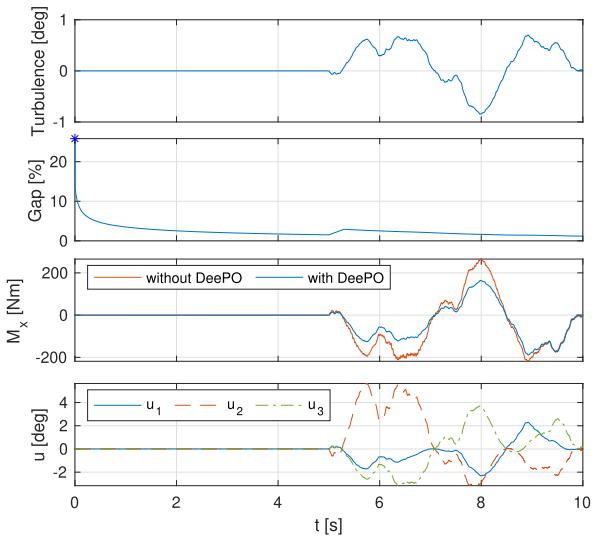

Fig. 8. Performance of DeePO using only output feedback. (a) Time history of an unknown Dryden turbulence. (b) Evolution of the optimality gap. (c) Responses of the root bending moment. (d) Evolution of the inputs.

moment M_x is reduced by 163.29 N·m (61.51 %), and its rms value decreases by 59.88 N·m (60.01 %), confirming the load alleviation effectiveness of the DeePO algorithm.

D. Output Feedback

The results in Section V-C rely on state information provided by an observer. In this section, this dependency is completely removed. The control objective is to alleviate structural loads induced by unknown external turbulence without any knowledge of the model or state information. Input–output data is generated using the model at $V = 240$ m/s. As discussed in Section V-A, this model has 21 sensory outputs. While it is possible to use all outputs in controller design, this is not necessary; a subset of outputs can be selected as long as the system remains observable. To challenge the capabilities of the DeePO algorithm, only one single output, specifically the wing root bending moment M_x, is selected without loss of generality. This output is also the most important metric for evaluating load alleviation performance in the aeroelastic control community.

Regarding the algorithm parameters, the offline data length is set to $t_0 = 250$, which satisfies $t_0 \geq h(m+p)+m$. Since after the transformation using Theorem 4.1, the system size increased from n to $h(m+p)$, which is 40 in this case, the step size for the projected gradient descent is decreased to $\eta = 1 \times 10^{-5}$. Moreover, Q is tuned as a diagonal matrix, where the diagonal elements corresponding to historical inputs are set to one, and the diagonal elements corresponding to outputs, are set to 10^4. R is selected as an identity matrix. The online data length h is selected to be equal to the system lag l, which equals the state dimension n when the number of output $p = 1$ ($l \leq n \leq pl$ [36]).

The control performance is shown in Fig. 8. Under the same Dryden turbulence perturbation, the optimality gap of the output-feedback DeePO algorithm also exhibits an initial sharp drop, specifically from 26.1 to 3.5 within the first second. It then shows a more gradual decreasing trend, with a slight increase when turbulence is introduced at the 5-second mark. Compared to the full-state feedback case, the convergence rate is slightly lower due to a reduced step size. Within the same 10-second experimental period, the optimality gap of the output-feedback algorithm converges to 1.3, achieving a load reduction of 28.68% and 29.14% in peak and rms values, respectively. Further performance improvements can be expected with additional time. All the control inputs remain within the bound of ± 10 deg while actively alleviate gust loads and redistributes the loads along the wing span.

E. Output Feedback for an LPV System

The control task is further enhanced in this section. In addition to the challenges posed by unmeasurable states and unknown external disturbances, the simulation system is also upgraded to an LPV system. A scaling parameter, ρ, is introduced to bridge two LTI systems discussed in Section V-A. When $\rho = 0$, the LPV system is equivalent to the LTI system at $V = 240$ m/s, which is stable. When $\rho = 1$, the LPV system corresponds to the unstable LTI system at $V = 257$ m/s, which would experience unstable flutter without active control. Linear interpolation is performed in between. It is noteworthy that the LPV model is used solely for simulation, i.e., to generate data. The DeePO algorithm has no knowledge of or access to any information about the model or the scaling parameter ρ.

To account for the time-varying dynamics, a forgetting factor λ is introduced. The value of $\lambda = 0.995$ is tuned, compromising convergence speed, responsiveness, and noise robustness. The offline data length remains at $t_0 = 250$, which satisfies $t_0 \geq h(m+p)+m$. Because of the time-varying system dynamics, the constant step size for the projected gradient descent is further reduced to $\eta = 1 \times 10^{-8}$. This step size could be significantly relaxed if a better initial policy were available (e.g., using the SDP method based on offline data [34]). Nevertheless, to challenge the DeePO algorithm, the initial policy is kept at zero in this research.

As shown in the first subplot of Fig. 9, during the simulation, ρ increases linearly from 0 to 1 over nine seconds and remains at 1 for the final second. The same Dryden turbulence is introduced at 5 seconds and continues throughout the remainder of the simulation. As ρ evolves, the open-loop system transitions from stable to unstable (flutter). In contrast, the DeePO algorithm is able to stabilize the system while alleviating gust loads without any knowledge of the model or state information. The optimality gap initially decreases promptly, then gradually rises due to the increasing ρ. At 5 seconds, the introduction of turbulence causes a brief increase in the optimality gap. After 9 seconds, as ρ stabilizes at 1, the optimality gap resumes its decline. All control inputs remain within the bounds of ± 10 degrees, effectively performing multiple tasks: flutter suppression, gust load alleviation, and adaptation to time-varying dynamics.

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

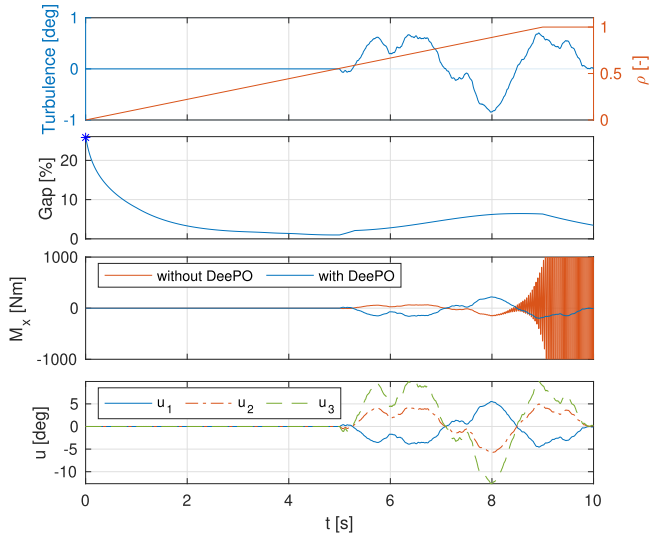

Fig. 9. Performance of DeePO using only output feedback for an LPV system. (a) Time history of an unknown Dryden turbulence and the LPV scaling parameter ρ. (b) Evolution of the optimality gap. (c) Responses of the wing root bending moment. (d) Evolution of the control inputs.

Before concluding the article, we acknowledge practical aspects such as time delay, saturation, and finite precision in digital computation, which can introduce challenges to the proposed algorithm. These challenges can be mitigated by incorporating techniques, such as Smith predictors for time delays, anti-windup compensation methods (e.g., conditional integration or observer-based approaches) for saturation, and dithering techniques for reducing quantization effects. Although these methods can be implemented, their integration and detailed analysis are beyond the scope of this article.

VI. CONCLUSION

In this article, we addressed the open challenge of unified flutter suppression and gust load alleviation for an aeroelastic wing operating in uncertain environments subject to atmospheric disturbances via DeePO. Theoretical analysis demonstrated that this approach can directly learn and update a control policy purely from input–output data, without requiring any prior knowledge of model or disturbance. Moreover, explicit system identification and state observation were avoided, significantly reducing computational costs and simplifying implementation. Furthermore, the proposed forgetting factoring mechanism enabled DeePO to actively adapt to time-varying system dynamics.

The effectiveness and efficiency of the proposed algorithm were validated via simulations. The results confirmed that the algorithm successfully stabilizes a time-varying fluttering aeroelastic system while alleviating gust loads. These objectives were achieved without relying on model or disturbance information, explicit system identification, state estimation, or gain/parameter scheduling. The effectiveness, efficiency, and global convergence properties of the algorithm suggested its potential for real-world implementation.

One limitation of this study is that the weights in the cost formulation are manually tuned by the designer. Since the optimality achieved is based on this predefined cost, optimizing the closed-loop performance, including optimizing the cost function through methods such as backpropagation, represents a promising direction for future research.

ACKNOWLEDGMENT

Disclaimer: Cofunded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or Clean Aviation Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them.

Co-funded by the European Union

CLEAN AVIATION

REFERENCES

[1] A. Eberle, B. Stefes, and D. Reckzeh, "Clean aviation ultra-performance wing (UP-Wing)," in *Proc. AIAA SCITECH Forum*, 2024, Art. no. 2109.

[2] P. Qi, X. Zhao, Y. Wang, R. Palacios, and A. Wynn, "Aeroelastic and trajectory control of high altitude long endurance aircraft," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 6, pp. 2992–3003, Dec. 2018.

[3] X. Wang, T. Mkhoyan, and R. D. Breuker, "Nonlinear incremental control for flexible aircraft trajectory tracking and load alleviation," *J. Guid. Control Dyn.*, vol. 45, no. 1, pp. 39–57, 2022.

[4] P. Qi, X. Zhao, and R. Palacios, "Autonomous landing control of highly flexible aircraft based on LiDAR preview in the presence of wind turbulence," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 5, pp. 2543–2555, Oct. 2019.

[5] C. Gao and W. Zhang, "Transonic aeroelasticity: A new perspective from the fluid mode," *Prog. Aerosp. Sci.*, vol. 113, 2020, Art. no. 100596.

[6] H. Dang, Z. Yang, and Y. Li, "Accelerated loosely-coupled CFD/CSD method for nonlinear static aeroelasticity analysis," *Aerosp. Sci. Technol.*, vol. 14, no. 4, pp. 250–258, 2010.

[7] T. Kim, "Parametric model reduction for aeroelastic systems: Invariant aeroelastic modes," *J. Fluids Struct.*, vol. 65, pp. 196–216, 2016.

[8] C. P. Moreno, P. J. Seiler, and G. J. Balas, "Model reduction for aeroservoelastic systems," *J. Aircr.*, vol. 51, no. 1, pp. 280–290, 2014.

[9] P. Dorato, "A historical review of robust control," *IEEE Control Syst. Mag.*, vol. 7, no. 2, pp. 44–47, Apr. 1987.

[10] T. Cebeci and F. Kafyeke, "Aircraft icing," *Annu. Rev. Fluid Mechan.*, vol. 35, no. 1, pp. 11–21, 2003.

[11] J. Chang, R. De Breuker, and X. Wang, "Adaptive nonlinear incremental flight control for systems with unknown control effectiveness," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 1, pp. 228–240, Feb. 2023.

[12] M. Tanaskovic, L. Fagiano, C. Novara, and M. Morari, "Data-driven control of nonlinear systems: An on-line direct approach," *Automatica*, vol. 75, pp. 1–10, 2017.

[13] F. Dörfler, "Data-driven control: Part two of two: Hot take: Why not go with models?," *IEEE Control Syst. Mag.*, vol. 43, no. 6, pp. 27–31, Dec. 2023.

[14] Z.-S. Hou and Z. Wang, "From model-based control to data-driven control: Survey, classification and perspective," *Inf. Sci.*, vol. 235, pp. 3–35, 2013.

[15] T. He and W. Su, "A parametric dynamic mode decomposition for reduced-order modeling of highly flexible aircraft," in *Proc.*

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

ASME Aerosp. Structures Struct. Dynamics, Mater. Conf., 2024, Art. no. V001T02A010.
[16] N. Fonzi, S. L. Brunton, and U. Fasel, "Data-driven nonlinear aeroelastic models of morphing wings for control," *Proc. Roy. Soc. A*, vol. 476, no. 2239, 2020, Art. no. 20200079.
[17] J. G. Pérez, A. Ghadami, L. Sanches, G. Michon, and B. I. Epureanu, "Data-driven optimization for flutter suppression by using an aeroelastic nonlinear energy sink," *J. Fluids Structures*, vol. 114, 2022, Art. no. 103715.
[18] T. He and W. Su, "Data-driven reduced-order aeroelastic modeling of highly flexible aircraft by parametric dynamic mode decomposition," 2023, *arXiv:2307.13960*.
[19] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *Found. Comput. Math.*, vol. 20, no. 4, pp. 633–679, 2020.
[20] H. Mania, S. Tu, and B. Recht, "Certainty equivalence is efficient for linear quadratic control," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 10154–10164.
[21] F. Zhao, F. Dörfler, A. Chiuso, and K. You, "Data-enabled policy optimization for direct adaptive learning of the LQR," 2024, *arXiv:2401.14871*.
[22] F. Zhao, F. Dörfler, and K. You, "Data-enabled policy optimization for the linear quadratic regulator," in *Proc. 62nd IEEE Conf. Decis. Control*, 2023, pp. 6160–6165.
[23] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017, *arXiv:1707.06347*.
[24] P. Makila and H. Toivonen, "Computational methods for parametric LQ problems—A survey," *IEEE Trans. Autom. Control*, vol. AC-32, no. 8, pp. 658–671, Aug. 1987.
[25] B. Hu, K. Zhang, N. Li, M. Mesbahi, M. Fazel, and T. Başar, "Toward a theoretical foundation of policy optimization for learning control policies," *Annu. Rev. Control Robot. Auton. Syst.*, vol. 6, no. 1, pp. 123–158, 2023.
[26] F. Zhao, K. You, and T. Başar, "Global convergence of policy gradient primal-dual methods for risk-constrained LQRs," *IEEE Trans. Autom. Control*, vol. 68, no. 5, pp. 2934–2949, May 2023.
[27] H. Mohammadi, A. Zare, M. Soltanolkotabi, and M. R. Jovanović, "Convergence and sample complexity of gradient methods for the model-free linear-quadratic regulator problem," *IEEE Trans. Autom. Control*, vol. 67, no. 5, pp. 2435–2450, May 2022.
[28] F. Zhao, X. Fu, and K. You, "Convergence and sample complexity of policy gradient methods for stabilizing linear systems," *IEEE Trans. Autom. Control*, vol. 70, no. 3, pp. 1455–1466, Mar. 2025.
[29] C. De Persis and P. Tesi, "Low-complexity learning of linear quadratic regulators from noisy data," *Automatica*, vol. 128, 2021, Art. no. 109548.
[30] S. Kang, F. Zhao, and K. You, "Linear convergence of data-enabled policy optimization for linear quadratic tracking," 2024, *arXiv:2410.05596*.
[31] F. Zhao, R. Leng, L. Huang, H. Xin, K. You, and F. Dörfler, "Direct adaptive control of grid-connected power converters via output-feedback data-enabled policy optimization," 2024, *arXiv:2411.03909*.
[32] N. Persson, F. Zhao, M. Kaheni, F. Dörfler, and A. V. Papadopoulos, "An adaptive data-enabled policy optimization approach for autonomous bicycle control," 2025, *arXiv:2502.13676*.
[33] F. Dörfler, P. Tesi, and C. D. Persis, "On the certainty-equivalence approach to direct data-driven LQR design," *IEEE Trans. Autom. Control*, vol. 68, no. 12, pp. 7989–7996, Dec. 2023.
[34] C. Helmberg, "Semidefinite programming," *Eur. J. Oper. Res.*, vol. 137, no. 3, pp. 461–482, 2002.
[35] J. Sherman and W. J. Morrison, "Adjustment of an inverse matrix corresponding to a change in one element of a given matrix," *Ann. Math. Statist.*, vol. 21, no. 1, pp. 124–127, 1950.
[36] M. Alsalti, V. G. Lopez, and M. A. Müller, "Notes on data-driven output-feedback control of linear MIMO systems," *IEEE Trans. Automa. Control*, 2025.
[37] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*. North Chelmsford, MA, USA: Courier Corporation, 2014.
[38] H. S. Timmermans et al., "Design of a high aspect ratio wind tunnel model for transonic gust load alleviation," in *Proc. Int. Forum Aeroelasticity Struct. Dyn.* 2024, pp. 1–14.

Xuerui Wang received the Ph.D. degree in aerospace engineering from the Delft University of Technology (TU Delft), Delft, the Netherlands, in 2019.

From May 2019 to May 2020, she was a Postdoctoral Researcher with the Smart and Aeroelastic Structure Laboratory, TU Delft. Since May 2020, she has been an Assistant Professor with the Faculty of Aerospace Engineering, TU Delft. She has joint affiliations with the Department of Control and Operations and the Department of Aerospace Structures and Materials. Her research interests include nonlinear control, aeroelasticity, aerial robotics, and active morphing structures.

Feiran Zhao received the B.S. degree from the Harbin Institute of Technology, Harbin, China, in 2018, and the Ph.D. degree from Tsinghua University, Beijing, China, in 2024, both in control science and engineering.

Since 2024, he has been a Postdoctoral Researcher with ETH Zürich, Zürich, Switzerland. His research interests include policy optimization, data-driven control, adaptive control and their applications.

Andres Jürisson received the M.Sc. degree in systems and control from Mechanical, Maritime and Materials Engineering Faculty, Delft University of Technology, (TU Delft), Delft, The Netherlands, in 2018. He is currently working toward the Ph.D. degree in system identification of flexible aircraft with the Netherlands Aerospace Centre (NLR) in collaboration with the Aerospace Structures and Materials Department, TU Delft.

His research interests include aeroelastic aircraft identification, active load alleviation, and aeroelasticity.

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.

Florian Dörfler (Senior Member, IEEE) received the diploma degree in engineering cybernetics from the University of Stuttgart, Stuttgart, Germany, in 2008 and the Ph.D. degree in mechanical engineering from the University of California at Santa Barbara, Santa Barbara, CA, USA, in 2013.

From 2013 to 2014, he was an Assistant Professor with the University of California Los Angeles, Los Angeles, CA, USA. He has been serving as the Associate Head with the Department of Information Technology and Electrical Engineering, ETH Zürich, Zürich, Switzerland, from 2021 to 2022. He is currently a Full Professor with the Automatic Control Laboratory, ETH Zürich. His research interests include automatic control, system theory, and optimization. His particular foci are on network systems, data-driven settings, and applications to power systems.

Dr. Dörfler was the recipient of the distinguished young research awards by IFAC (Manfred Thoma Medal 2020) and EUCA (European Control Award 2020). His students were winners or finalists for Best Student Paper awards at the European Control Conference (2013, 2019), the American Control Conference (2016, 2024), the Conference on Decision and Control (2020), the PES General Meeting (2020), the PES PowerTech Conference (2017), the International Conference on Intelligent Transportation Systems (2021), and the IEEE CSS Swiss Chapter Young Author Best Journal Paper Award (2022, 2024). He is furthermore a recipient of the 2010 ACC Student Best Paper Award, the 2011 O. Hugo Schuck Best Paper Award, the 2012–2014 Automatica Best Paper Award, the 2016 IEEE Circuits and Systems Guillemin-Cauer Best Paper Award, the 2022 IEEE Transactions on Power Electronics Prize Paper Award, and the 2015 UCSB ME Best Ph.D. award. He is currently serving on the council of the European Control Association and as a Senior Editor of Automatica.

Roy S. Smith (Life Fellow, IEEE) received the undergraduate degree from the University of Canterbury, Christchurch, New Zealand, and the Ph.D. degree from the California Institute of Technology, Pasadena, CA, USA, in 1980 and 1990, respectively, all in electrical engineering.

He is currently a Retired Adjunct Professor of electrical engineering with ETH Zürich, Zürich, Switzerland. Prior to joining ETH in 2011, he was a Faculty with the University of California, Santa Barbara, from 1990 to 2010. His research interests include the modeling, identification, and control of uncertain systems with applications in chemical processes, flexible structure vibration, spacecraft and vehicle formations, aerodynamic control of kites, automotive engines, Mars aeromaneuvering entry design, building, and energy hub control.

Authorized licensed use limited to: TU Delft Library. Downloaded on October 14,2025 at 11:26:44 UTC from IEEE Xplore. Restrictions apply.