Learning the nonlinear manifold of extreme aerodynamics

Kai Fukami and Kunihiko Taira*
Department of Mechanical and Aerospace Engineering,
University of California, Los Angeles
Corresponding author: ktaira@seas.ucla.edu

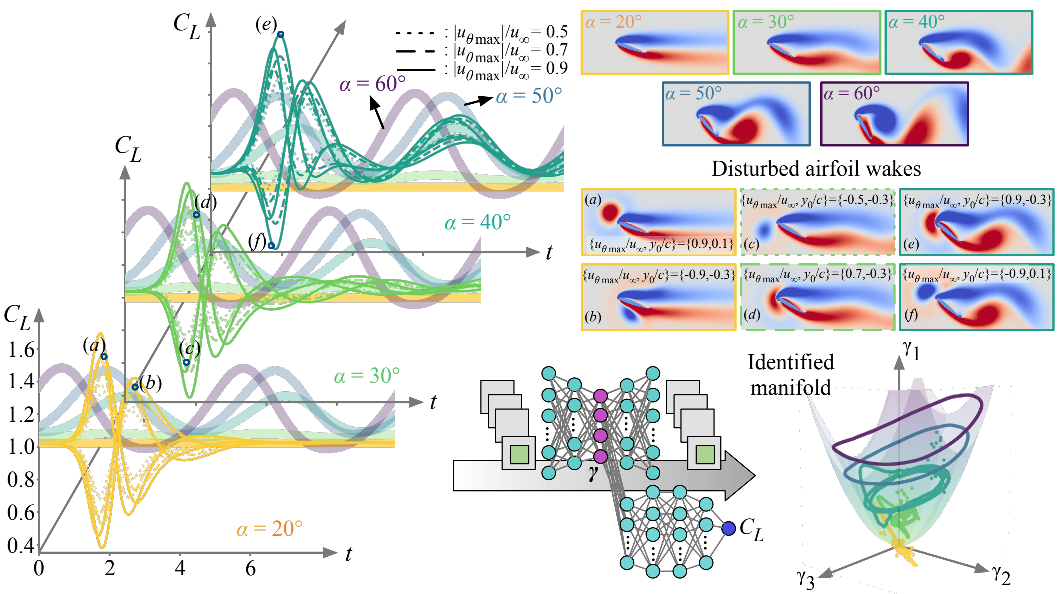

Figure 1: Physics-embedded autoencoder-based manifold identification of extreme aerodynamics.

Abstract

With the increased occurrence of extreme events and miniaturization of aircraft, it has become an urgent task to understand aerodynamics in highly turbulent flight environments. We propose a physics-embedded autoencoder to discover a low-dimensional compact manifold representation of extreme aerodynamics. The present method is demonstrated with the highly nonlinear dynamics of vortex gust-airfoil wake interaction around a NACA0012 airfoil over a range of configurations. The present model extracts key features of the high-dimensional airfoil wake dynamics on a physically interpretable and compact manifold, covering a massive number of wake scenarios across a huge parameter space that determines the characteristics of complex gusty flow conditions. Our data-driven approach offers a new avenue for expressing the seemingly high-dimensional fluid flow systems by identifying the low-dimensional data coordinates that can also be leveraged for data compression and flow control.

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

1 Introduction

Global warming fueled by climate change is contributing to extreme weather conditions, which makes flight difficult for drones, helicopters, and airplanes. It is now critically important to understand and control such highly unsteady and chaotic conditions to fly safely and efficiently. Traditionally, these flight conditions were rarely studied. However, enormously large combinations of parameters govern the extreme weather conditions associated with flow disturbances and gusts. Moreover, these parameters nonlinearly affect the overall aerodynamics [1]. These factors make analyzing unsteady flows incredibly challenging using existing theoretical techniques, calling for data-driven approaches.

This study aims to capture the key dynamics of complex airfoil wake-vortex gust interaction. Transitioning to examine fluid flow data in high-dimensional space-time coordinates to low-dimensional phase-based space enables us to achieve this objective. We leverage a nonlinear autoencoder to identify a compact manifold of high-dimensional airfoil wakes. We demonstrate that a nonlinear autoencoder becomes a powerful tool to obtain a physically-interpretable manifold for rich dynamics of transient airfoil wakes by incorporating the lift response into the low-dimensional latent space design. The present manifold offers a diverse possibility of extensions including reduced-order modeling, data compression, and flow control in complex aerodynamics.

2 Methods

2.1 Example: a family of wakes around a NACA0012 airfoil

This study finds a universal manifold that describes a broad range of complex wake behaviors in a low-dimensional space. We consider a variety of wake patterns generated around a NACA0012 airfoil at a chord-based Reynolds number $Re_c = 100$ and a Mach number $M_\infty = 0.1$, as presented in figure 1. In addition to the steady and periodic shedding cases at the angles of attack of $\alpha \in [20^\circ, 30^\circ, 40^\circ, 50^\circ, 60^\circ]$, the present data set from the numerical simulation also includes complex transient wakes associated with strong vortex disturbance interacting with the airfoil at $\alpha \in [20^\circ, 30^\circ, 40^\circ]$. These complex wakes also provide a diversity of lift responses.

For the cases with vortical gust, single disturbance modeled by a compressible Taylor vortex is initially added to the flows, enabling us to emulate the practical aerodynamics scenario when a gust traverses through an airfoil during flight. The vortex disturbance is parameterized with its radius R and max vortex velocity $u_{\theta\max}$. We set the radius R/c to be 0.25. At the initial condition, the vortex is placed at $(x_0, y_0)/c$ while x_0 being set at -2.

For the cases without any disturbances, the wake at $\alpha = 20^\circ$ is steady with no vortex shedding, while the wakes at $\alpha \geq 30^\circ$ exhibit unsteady periodic shedding. In contrast, the cases with the disturbance show the different types of nonlinear and intrinsic responses depending on the max vortex velocity $u_{\theta\max}$ and the vertical position y_0 of the vortex [2]. We consider disturbed flows with a parameter combination of $u_{\theta\max}/u_\infty \in [-0.9, -0.7, -0.5, 0.5, 0.7, 0.9]$ and $y_0/c \in [-0.3, 0.1]$ at $\alpha \in [20^\circ, 30^\circ, 40^\circ]$. Although the flows considered in the present study are laminar, these wakes involve strong nonlinearities associated with the interaction among a disturbance, an airfoil, and the wake. As a result, a variety of wake patterns are generated, as presented in figure 1.

The present wake data set includes an extremely large spatiotemporal degree of freedom corresponding to $N_x \times N_y \times N_t \times N_c$, where N_x and N_y are respectively the number of the grid points in the x and y directions, N_t is the number of the snapshots for each case, and N_c is the number of the cases considered in the present data sets. In the present study, we consider $N_t = 1200$ snapshots of vorticity field ω on $N_x \times N_y = 240 \times 120$ grids for each of $N_c = 50$ cases, resulting in $\mathcal{O}(10^9)$. We demonstrate that the present autoencoder successfully expresses such enormously high-dimensional and rich nonlinear wake dynamics in a low-dimensional latent space while conventional linear modal reduction techniques cannot distinguish these features in their traditionally identified low-dimensional coordinates.

2.2 Physics-embedded autoencoder

Principal component analysis (PCA) has often been used to study wake patterns [3]. However, the strong complexity of the present transient wakes associated with disturbance causes difficulty in

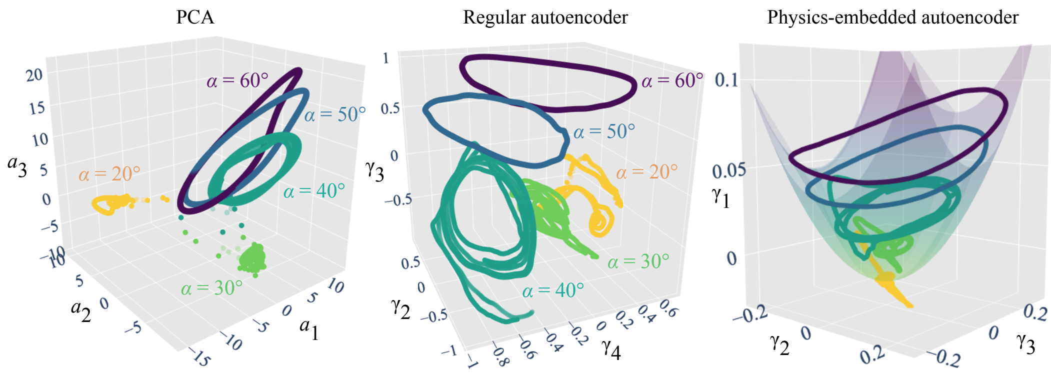

Figure 2: Comparison of low-dimensional latent spaces with various compression techniques.

acquiring a low-dimensional manifold with such conventional linear techniques (we will also demonstrate it later). Hence, we consider nonlinear dimensionality reduction [4, 5] to low-dimensionalize the present wake family. This study uses a physics-embedded autoencoder [6, 7] composed of convolutional neural network [8] and multi-layer perceptron [9] to optimally find a low-order and nonlinear representation of the high-dimensional wakes $\boldsymbol{q}$. An autoencoder outputs the same data as that given as the input through the bottleneck structure using nonlinear activation functions. The lowest-dimensional vector $\boldsymbol{\gamma}$ (pink circles in figure 1) can be regarded as a low-dimensional representation of the original data $\boldsymbol{q}$ if the model successfully outputs the same data as the input. Following our preliminary test, we set the number of latent variables to 4 and use the hyperbolic tangent function (tanh) as the activation function.

While efficient data compression can be achieved with a nonlinear autoencoder, we also aim to find coordinates that explicitly express the disturbed dynamics about the baseline states (the wakes without disturbance). To find such an interpretable manifold, we enforce the low-dimensional latent vector $\boldsymbol{\gamma}$ to have a coherent relationship with the lift coefficient C_L, as illustrated in figure 1. The lift coefficient is an observable characteristic from sensors, having a physical relationship with angles of attack. This physics embedding not only promotes the interpretability of the latent vectors but also prevents a scenario in which the latent variables cannot distinguish the wakes with different angles of attack. Our optimization technique also allows for a general representation of the manifold shape oriented towards visualization or the design of flow control.

The optimization process of the present physics-embedded autoencoder is expressed as

$$\boldsymbol{w}^* = \mathrm{argmin}_{\boldsymbol{w}} \Big[||\boldsymbol{q} - \hat{\boldsymbol{q}}||_2 + \beta_1 ||C_L - \hat{C_L}||_2 + \beta_2 ||f(\boldsymbol{\gamma})||_2 \Big], \quad (1)$$

where $\boldsymbol{w}$ denotes the weights inside the autoencoder model, and β_1 and β_2 decide the weighting (balance) in the loss function among the reconstruction loss and the lift-based regularization, and the manifold identification. This study uses a paraboloid-based manifold, such that $f(\boldsymbol{\gamma}) = (\gamma_1 + \delta_{\gamma_1}) - (\gamma_2^2 + \gamma_3^2)$, inspired by the traditional reduced-order modeling for transient wake dynamics [3]. The offset parameter δ_{γ_1} is needed to successfully represent the disturbing influence at $\alpha = 20°$, this case without the disturbance is steady. We set $\{\beta_1, \beta_2, \delta_{\gamma_1}\} = \{0.05, 1, 0.025\}$. The Adam optimizer [10] is used for updating the weights in training. The training of the autoencoder is performed with the NVIDIA Tesla V100 GPU.

3 Results and discussion

Let us first compare a three-dimensional space composed of latent variables derived from PCA, a regular autoencoder (without the lift and manifold shape loss functions), and the present physics-embedded autoencoder, in figure 2. PCA models the flow fields without disturbances at $\alpha = 20°$ and $30°$ into the different portions, suggesting that these two cases can be distinguished with the linear technique. However, the low-dimensional variables for the cases at $\alpha = 40°$, $50°$, and , $60°$ overlap around the region of $a_1 > 0$. In addition, it is also difficult to identify the influence of disturbance in the low-dimensional space, although we observe some non-interpretable fluctuations in the PCA

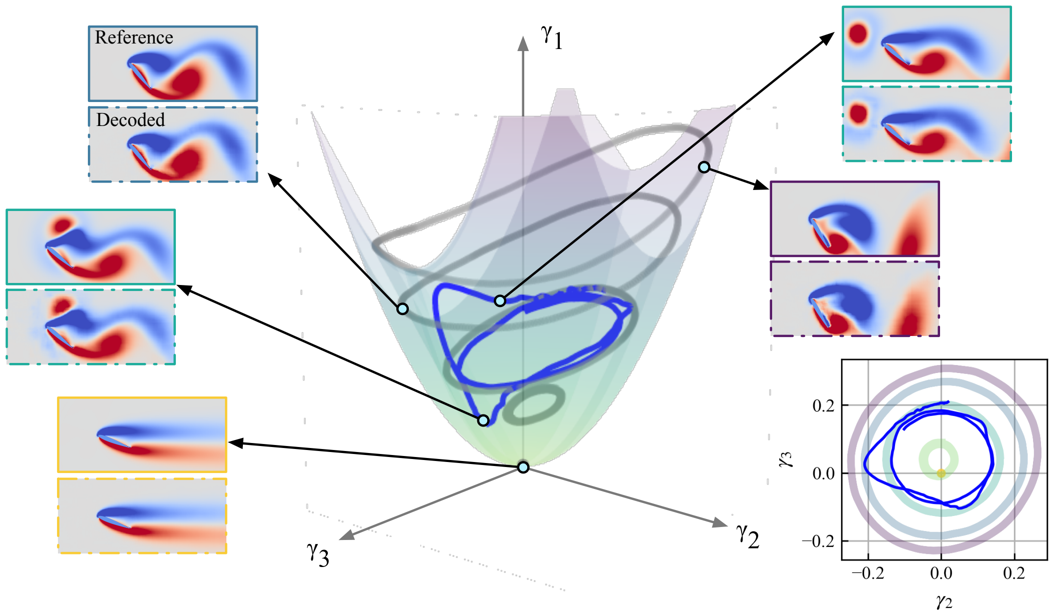

Figure 3: Wake fields projected on the present manifold. Decoded flow fields are also shown.

space that may correspond to the disturbed wake snapshots. Hence, PCA cannot find a universal mode to successfully classify these cases, especially in the presence of disturbed transient flows with strong nonlinearities.

In contrast, a regular autoencoder with tanh nonlinear activations clearly distinguishes the cases in the present wake data sets. While the cases at $\alpha = 40°$, $50°$, and, $60°$ are mapped into the different portions in the latent space, the influence of the disturbance is expressed in the radius direction of each periodic orbit. The reconstruction L_2 error of the regular autoencoder for all cases considered herein is 0.0858, while that of the PCA is 0.456, clearly indicating that the latent variables from the nonlinear autoencoder are more informative compared to that from the linear PCA. However, since there is no constraint for the data distribution in the latent space, the low-dimensional variables reside in the latent space in an unorganized manner. For instance, there are no clear axes to express the difference in angles of attack. This is because the latent distribution with this no-constraint setup is just needed to be distinguishable among the cases, implying that the autoencoder attempts to efficiently and widely use area in the latent space rather than possessing its interpretability. Towards a deeper understanding of the high-dimensional and complex wake dynamics and controlling them based on the manifold, it is desirable to obtain a more compact manifold, which clearly represent the disturbed dynamics against the baseline limit cycles, through the nonlinear autoencoder.

From this aspect, the present physics-embedded autoencoder provides a compact and understandable low-dimensional representation of high-dimensional wake dynamics. The axis of γ_1 corresponds to the angle of attack. Moreover, the influence of the disturbance is represented on the manifold surface, providing a similarity between disturbed wake fields and the undisturbed baseline flow fields, as depicted in figure 3. This indicates that the present autoencoder-based manifold captures effective angle of attacks of the disturbed wake cases from the vorticity snapshots and the lift coefficients. The reconstruction L_2 error of the present physics-embedded autoencoder is 0.0901, which is in a qualitative agreement of the decoded field with the reference simulation data as presented in figure 3. In other words, the present manifold offers a low-dimensional expression (only 3 variables) of the original key dynamics ($O(10^9)$) while reducing its complexity.

4 Concluding remarks

We proposed a physics-embedded autoencoder to optimally derive a compact representation of high-dimensional and complex extreme aerodynamic wakes. The current method successfully offered a simple and informative representation of the original wake dynamics associated with vortex gust-airfoil wake interaction. As demonstrated, the physics-embedded autoencoder can be used for compressing high-dimensional fluid flow data sets into a few latent variables while keeping their essential and key features of transient dynamics. Although not shown, we have also confirmed that the present autoencoder is robust against not only noisy field measurements but also untrained parameter conditions. We expect that the present parabolic-type manifold can be leveraged for flow control in extreme aerodynamic conditions by combining with energy-based control [11] with the modification of angle of attack to mitigate the impact of the gusty disturbance on the airfoil.

Acknowledgements

This work was supported by the US Department of Defense Vannevar Bush Faculty Fellowship (N00014-22-1-2798) and the US Air Force Office of Scientific Research (FA9550-21-1-0178).

References

[1] A. R. Jones, O. Cetiner, and M. J. Smith. Physics and modeling of large flow disturbances: Discrete gust encounters for modern air vehicles. *Annu. Rev. Fluid Mech.*, 54:469–493, 2022.

[2] Y. Zhong, K. Fukami, B. An, and K. Taira. Machine-learning-based reconstruction of transient vortex-airfoil wake interaction. AIAA paper, 2022-3244, 2022.

[3] B. R. Noack, K. Afanasiev, M. Morzynski, G. Tadmor, and F. Thiele. A hierarchy of low-dimensional models for the transient and post-transient cylinder wake. *J. Fluid Mech.*, 497:335–363, 2003.

[4] N. D. Lawrence. A unifying probabilistic perspective for spectral dimensionality reduction: Insights and new models. *J. Mach. Learn. Res.*, 13:1609–1638, 2012.

[5] J. A. Lee and M. Verleysen. *Nonlinear dimensionality reduction*, volume 1. Springer, 2007.

[6] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006.

[7] T. Murata, K. Fukami, and K. Fukagata. Nonlinear mode decomposition with convolutional neural networks for fluid dynamics. *J. Fluid Mech.*, 882:A13, 2020.

[8] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proc. IEEE*, 86(11):2278–2324, 1998.

[9] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagation errors. *Nature*, 322:533–536, 1986.

[10] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. arXiv:1412.6980, 2014.

[11] S. L. Brunton and B. R. Noack. Closed-loop turbulence control: Progress and challenges. *Appl. Mech. Rev.*, 67(5), 2015.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes]
   (c) Did you discuss any potential negative societal impacts of your work? [No] No negative social impacts.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [No] The data that support the findings of this study are available from the authors upon reasonable request.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [N/A]
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]