Going with the Speed of Sound: Pushing Neural Surrogates into Highly-turbulent Transonic Regimes

Fabian Paischer[*1,2], **Leo Cotteleer**[*1], **Yann Dreze**[*1], **Richard Kurle**[*1], **Dylan Rubini** [1], **Maurits Bleeker** [1], **Tobias Kronlachner** [1], **Johannes Brandstetter** [1,2]
[1] Emmi AI GmbH, Linz
[2] ELLIS Unit, Institute for Machine Learning, JKU Linz

Abstract

The widespread use of neural surrogates in automotive aerodynamics, enabled by datasets such as DrivAerML and DrivAerNet++, has primarily focused on bluff-body flows with large wakes. Extending these methods to aerospace, particularly in the transonic regime, remains challenging due to the high level of non-linearity of compressible flows and 3D effects such as wingtip vortices. Existing aerospace datasets predominantly focus on 2D airfoils, neglecting these critical 3D phenomena. To address this gap, we present a new dataset of high-fidelity CFD simulations for 3D wings in the transonic regime. The dataset comprises volumetric and surface-level fields for around $30,000$ samples with unique geometry and inflow conditions. This allows computation of lift coefficients, drag coefficients, providing a foundation for data-driven aerodynamic optimization of the drag–lift Pareto front. We evaluate several state-of-the-art neural surrogates on our dataset, including Transolver and AB-UPT, focusing on their out-of-distribution (OOD) generalization over geometry and inflow variations. AB-UPT demonstrates strong performance for transonic flowfields and reproduces physically consistent drag–lift Pareto fronts even for unseen wing configurations. Our results demonstrate that AB-UPT can approximate drag–lift Pareto fronts for unseen geometries, highlighting its potential as an efficient and effective tool for rapid aerodynamic design exploration.

1 Introduction

Machine learning–based surrogates have recently emerged as powerful tools for accelerating aerodynamic design and analysis [1, 2]. In automotive aerodynamics, large-scale datasets such as DrivAerML [3] and DrivAerNet++ [4, 5] have enabled neural models to predict complex bluff-body flows with remarkable accuracy, potentially reducing reliance on expensive Computational Fluid Dynamics (CFD) simulations. In aerospace applications, the design goals and flow physics differ significantly from the automotive domain. Automobiles are bluff bodies with early flow separation, leading to high pressure drag [6]. In contrast, aircrafts are streamlined bodies designed to maintain attached flow and optimize the lift-to-drag ratio [7], which we refer to as drag–lift Pareto front.

While CFD is a fundamental design tool, its high computational cost creates a bottleneck in the design cycle. This has spurred the development of data-driven surrogate models [1, 8, 9], which learn the complex mapping between geometry and flow fields [10]. However, extending these approaches to aerospace applications remains challenging. Transonic flight regimes involve intricate 3D phenomena, such as shock–boundary layer interactions and wingtip vortices, that are not captured with existing 2D airfoil datasets [11, 12, 13, 14]. Moreover, the lack of publicly available high-fidelity 3D flow

[*] Equal contribution

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Track on Datasets and Benchmarks.

data limits the development and benchmarking of neural surrogates capable of generalizing across realistic aircraft geometries and operating conditions, as well as optimizing lift-to-drag performance.

To address these limitations, we introduce a new dataset of high-fidelity RANS simulations for 3D wings in the transonic regime. To the best of our knowledge, this is the first publicly available dataset that comprehensively captures both geometric and inflow variations for realistic 3D configurations. The dataset comprises approximately $30,000$ simulations with a unique combination of geometry and inflow parameters, providing volumetric as well as surface-level flow-field data. These data enable the computation of aerodynamic performance metrics such as lift, drag, and drag–lift polars, thereby supporting data-driven design-space exploration and optimization.

Using our new dataset, we evaluate several state-of-the-art neural surrogates, including Transolver [8] and the recently proposed AB-UPT [1], with a focus on out-of-distribution generalization to unseen geometries and flow conditions. Our results show that AB-UPT accurately predicts transonic flow fields and can reproduce physically consistent lift–drag trade-offs for unseen configurations, outperforming all competitors. Overall, our study demonstrates that AB-UPT can approximate lift–drag Pareto fronts for unseen geometries, highlighting its potential as practical tool for rapid aerodynamic design exploration.

2 Methodology

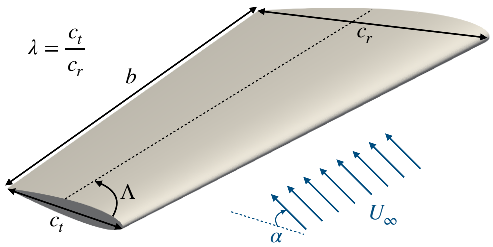
Our methodology centers on the generation of a high-fidelity dataset of 3D wings operating in the transonic regime, followed by the evaluation of state-of-the-art neural surrogate models. Existing aerospace CFD datasets predominantly focus on subsonic 2D airfoils, as summarized in Table 1, and therefore neglect critical 3D flow phenomena such as shock–boundary layer interactions and wingtip vortices. We generate a new dataset comprising around $30,000$ cases with a unique combination of geometry and inflow parameters. The data generation process is organized into three key components: (i) the design of experiments for geometry and flow conditions, (ii) the setup and execution of high-fidelity CFD simulations, and (iii) the dataset split strategy used for training and evaluating neural surrogates.

Table 1: Comparison of publicly available aerodynamic CFD datasets for machine learning.

Dataset	**Size**	**Dim.**	**Regime**	**Notes**
AirfRANS [11]	∼1 000	2D	Subsonic	ML benchmark, varied AoA/Re
UniFoil [12]	500 K	2D	Sub-/Transonic	Very large 2D dataset, wide Re/Mach range
AirFoilCFD [13]	∼18 K	2D	Subsonic	9k shapes, 2 AoA, fixed inflow
AirFoilML [14]	2 600	2D	Subsonic	NACA airfoils, fixed AoA/Re
BlendedNet [15]	∼10 K	3D	Subsonic	Blended wing-body, 9 flight conds.
Emmi-Wing (Ours)	∼ 30 K	3D	**Sub-/Transonic**	30K unique geometries + inflows, drag/lift polars

Design of experiments. We select a 2D NACA0012 airfoil profile, which is extruded into 3D based on four geometric parameters: the span (b), the taper ratio (λ), the sweep angle (Λ), and the root chord (c_r). In addition, we vary two inflow parameters, namely velocity (U_∞) and angle of attack (α). To ensure a broad and diverse range of flow conditions, a total of $29,727$ unique simulation cases were created by randomly sampling each of the six variables from a uniform distribution within a application-relevant predefined range. The specific ranges for each parameter and their visual representation are depicted in Figure 1.

CFD setup. We conduct simulations using the open-source CFD solver, OpenFOAM-v2506 [16]. The steady-state, compressible flow rhoSimpleFoam solver [17] is used for all cases using the perfect gas assumption. For turbulence modeling, we selected the Spalart-Allmaras model which offers good computational efficiency and robust convergence for transonic wing flows with predominantly attached boundary layers, despite the presence of weak-to-moderate shocks [18]. Spatial discretization utilized second-order schemes including a van Leer limiter for momentum and a bounded upwind scheme for energy/pressure terms, with first-order upwind differencing applied to turbulence quantities for enhanced stability. The wing surfaces are set with a no-slip condition, and freestream conditions are applied at the far-field boundaries. The angle of attack is imposed via the inflow boundary condition rather than by altering the wing geometry.



Parameter	Range	
c_r	[0.7, 1.2]	m
b	[1.0, 1.5]	m
λ	[0.4, 0.7]	–
Λ	[0.0, 40.0]	deg
U_∞	[150, 300]	m/s
α	[-10, 10]	deg

Figure 1: A visual representation of the parameterized 3D wing geometry alongside the four key geometric design parameters and the two inflow condition parameters with their respective sampling ranges used for the design of experiments. The geometry parameters include the span (b), taper ratio (λ), sweep angle (Λ), and chord root (c_r); inflow parameters are the inflow velocity (U_∞) and angle of attack (α).

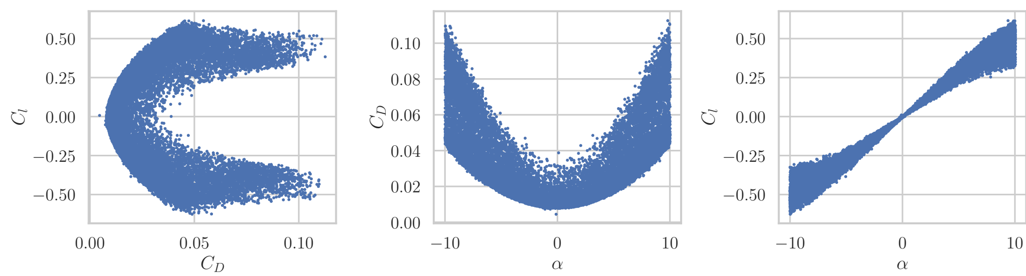


Figure 2: **Left:** Pareto frontier of drag C_D versus lift C_l coefficients for all cases in the dataset. **Middle:** C_D as a function of angles of attack present in the dataset. **Right:** C_l as a function of the angles of attack present in the dataset.

The meshing process relies on OpenFOAM's built-in snappyHexMesh to generate a high-quality, body-fitted mesh. Prismatic boundary layer meshing is used to achieve low y^+ values in the range $[50 - 200]$ in all cases.

The resulting dataset comprises $29,727$ cases, including both surface and volumetric data. The surface data consists of pressure p_s and wall shear-stress $\boldsymbol{\tau}$ on the wing. The volumetric data contains the full 3D flow field, including pressure p_v, the velocity vector $\boldsymbol{u} = [u_x, u_y, u_z]$, its magnitude $||\boldsymbol{u}||$, and the vorticity vector $\boldsymbol{\omega} = [\omega_x, \omega_y, \omega_z]$ and its magnitude $||\boldsymbol{\omega}||$ at each point of the CFD mesh. We provide visualizations of the lift-to-drag Pareto front as well as drag/lift coefficients over varying angles of attack for all cases in Figure 2. In addition we visualize a set of 3D wing geometries for the minimum and maximum of each parameter in Figure 14 in the Appendix. Finally, Figure 3a shows a representative sample for a transonic sample of our dataset exhibiting flow detachment at high angles of attack and the formation of turbulent structures at the wingtip.

Dataset split. We partition the dataset into a training, a validation, and three test sets, including two in-distribution (ID) sets and one out-of-distribution (OOD) set. The two ID sets are designed as (i) **random selection**: a random selection of $1,000$ cases within the convex hull of the training set, and (ii) **interpolation**: a parameter region within the convex hull of the parameter space that spans $1,000$ unseen cases. The former evaluates in-distribution performance, while the latter specifically targets the ability of the surrogate to interpolate between cases in an unseen ID parameter region. The $1,000$ OOD cases are selected as the outermost points of the convex hull of the parameter space. To create the OOD set and the ID interpolation set, where we used an iterative convex hull peeling method and isolate the $1,000$ outermost data points and the $1,000$ innermost points, respectively. The remaining $25,727$ cases are used for training.

Additional parameter scans. Practitioners are usually concerned with the optimization of the drag–lift Pareto front. To this end, lift and drag coefficients are usually investigated on parameter scans of various angles of attack α and varying wing geometries to obtain the wing geometry resulting in the highest lift and lowest drag forces. To asser the ability of neural surrogate models to capture these

Table 2: Evaluation of different surrogate models on the *in-distribution* and *out-of-distribution* test sets. Relative L1 and L2 errors (in %) are given for surface fields (pressure p_s, wall-shear-stress $\boldsymbol{\tau}$) and volume fields (pressure p_v, velocity $\mathbf{u}$, vorticity $\boldsymbol{\omega}$).

Test Set	**Model**	**Relative L1 Error (%)**					**Relative L2 Error (%)**				
		p_s	$\boldsymbol{\tau}$	p_v	$\boldsymbol{u}$	$\boldsymbol{\omega}$	p_s	$\boldsymbol{\tau}$	p_v	$\boldsymbol{u}$	$\boldsymbol{\omega}$
Interpol	PointNet	0.0361	0.2128	0.0344	0.1486	0.383	0.0522	0.2249	0.0496	0.1496	0.3505
	Transformer	0.001	0.0178	0.0011	0.0108	0.0819	0.0018	0.0226	0.0018	0.0133	0.1055
	Transolver	0.0008	0.0166	0.0009	0.0082	0.0656	0.0016	0.0209	0.0014	0.0108	0.0978
	AB-UPT	0.0008	0.0166	0.0008	0.0077	0.0564	0.0016	0.0211	0.0013	0.0098	0.0709
ID	PointNet	0.0707	0.4426	0.0672	0.2984	0.4962	0.101	0.4419	0.0962	0.2999	0.454
	Transformer	0.0029	0.0299	0.0029	0.0245	0.1248	0.0058	0.0438	0.0056	0.0377	0.1417
	Transolver	0.0024	0.0272	0.0024	0.0205	0.1005	0.0052	0.0407	0.0048	0.034	0.1293
	AB-UPT	0.0024	0.0273	0.0024	0.02	0.0883	0.0052	0.0408	0.0048	0.0332	0.1019
OOD	PointNet	0.0868	0.5986	0.0823	0.4006	0.597	0.1204	0.5857	0.1145	0.4015	0.5425
	Transformer	0.0044	0.0408	0.0045	0.0369	0.1727	0.0085	0.0598	0.0083	0.0561	0.1816
	Transolver	0.0036	0.0362	0.0036	0.0308	0.1337	0.0075	0.0546	0.0071	0.0505	0.1557
	AB-UPT	0.0037	0.0365	0.0036	0.0302	0.1177	0.0076	0.0548	0.0072	0.0494	0.1256

coefficients for different parameter scans we run additional cases with a unique geometry ($c_r = 0.806$, $b = 1.1963$, $\lambda = 0.562$) that does not occur in the $29,727$ cases. We run two parameter sweeps for this geometry, namely (i) $\alpha \in \{-30, -28, \ldots, 28, 30\}$, and (ii) $\Lambda \in \{0, 10, 20, 30, 40, 50, 60, 70\}$, to evaluate in-distribution and OOD generalization. The latter sweep concerns a geometry parameter (sweep angle) as it is the geometry parameter apart from AoA with the most impact on lift/drag. This results in another 248 cases solely used for evaluation.

3 Experiments

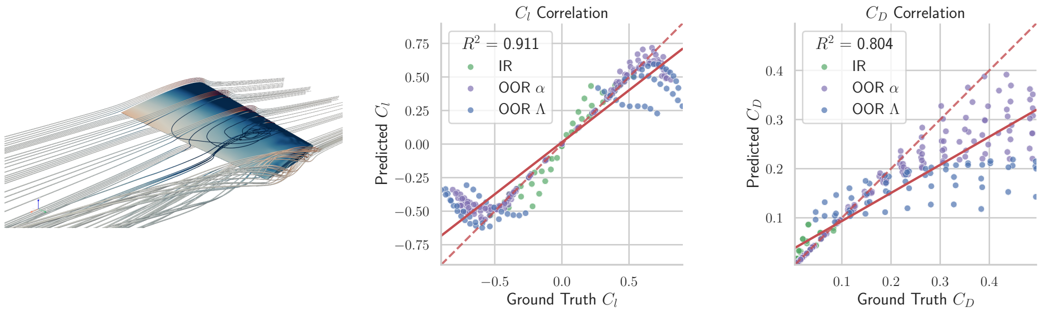
We conduct a series of experiments to evaluate the generalization capabilities of state-of-the-art neural surrogate models trained on our 3D transonic wing dataset. Our analysis focuses on two main aspects: (i) the in-distribution performance of different architectures on flowfield and aerodynamic coefficient prediction, and (ii) their out-of-distribution generalization to unseen geometries and inflow conditions.

Neural surrogates. We benchmark four neural surrogate models on our dataset: PointNet [19], Transolver [8], AB-UPT [1], and a Vision Transformer [20]. To incorporate inflow conditions and geometry design parameters, we add their embeddings to the input for PointNet, and for the transformer-based models we use a conditioning as in Diffusion Transformers [21]. All models are trained with a Mean Squared Error (MSE) loss applied to all fields.

Evaluation. We evaluate the predictive performance using the relative L_1 and L_2 errors. They are obtained by normalizing the L_i norm of the pointwise error by the norm of the ground truth values, averaged over all test cases. In addition, we report the coefficient of determination (R^2) between drag/lift coefficients obtained from the best surrogate model with the ground truth coefficients. Finally, we provide drag-lift Pareto fronts for parameter scans of the best performing surrogate model compared to ground truth to stress-test their generalization ability.

Results. We report results for volume and surface-level quantities in Table 2 for the different methods. As expected we observe a consistent increase of error the more the model is pushed towards an OOD evaluation regime. Furthermore, our results show that except for PointNet, all surrogate approaches perform similarly on surface-level quantities. However, on volume-level quantities that exhibit high variance, like vorticity, AB-UPT shows a significant improvement over competitors. Therefore, we consider AB-UPT the strongest neural surrogate approach and evaluate it on the different parameter scans.

We provide a qualitative analysis of the AB-UPT model on the OOD test set. Figure 4 in the Appendix shows pressure and friction profiles of a randomly sampled test case of the OOD set at a normalized span location $y/b = 0.75$. The corresponding 3D visualizations of the true surface fields, the predicted fields, and the prediction errors are shown in Figure 5. Remarkably, both pressure coefficient (C_p) and friction coefficient (C_f) match the ground truth closely. In Figure 8 in the



(a) 3D effects on wing. (b) C_l correlation to ground-truth. (c) C_D correlation to ground-truth.

Figure 3: Visualization of a 3D wing at $\alpha = 6.21$ and $U_\infty = 296.53$. The field lines detach at certain span lengths after the shock and vortices form at the wingtip (**3a**). Correlation of predicted C_l (**3b**) and C_D (**3c**) for parameter scans on values that are in-range (IR) and out-of-range (OOR) for α and Λ values during training. AB-UPT maintains high correlation to ground-truth, even for OOR values.

Appendix we show correlation of drag C_D and lift C_l coefficients with the ground truth for all cases in the OOD test set. The predictions of AB-UPT align closely with the ground-truth coefficients ($R^2 = 1.0$ for C_l and $R^2 = 0.998$ for C_D). Furthermore we report the error distribution across cases in the OOD test set for different combinations of drag and lift in Figure 9. Intriguingly, we find that AB-UPT is most error-prone in high C_D regimes, which are dominated by errors in τ and rather far from the tangent on the lift-drag Pareto front. This highlights that AB-UPT is a promising contender for wing design optimization.

Parameter scans. We evaluate AB-UPT on the parameter scans that comprise 248 simulations with sweeps over Λ and α with the remaining parameters fixed. We report correlation of the predicted values for C_l and C_D in Figure 3b and 3c, respectively, with highlighted cases for out-of-range (OOR) values of α and Λ. Remarkably, AB-UPT maintains high correlation with the ground-truth ($R^2 = 0.911$ for C_l, $R^2 = 0.804$ for C_D) even when observing values that are far beyond the range of values for α or Λ it has been trained on. We complement this observation with relative errors for surface quantities p_s and τ, as well as an average thereof in Figure 10. Interestingly, AB-UPT seems to perform better for positive angles of attack than for negative ones, even though the dataset was sampled uniformly.

In addition, we report the predicted drag-lift Pareto fronts for a selected set of sweep angles ($\Lambda \in \{20, 40, 50, 70\}$) and compare them to the ground-truth in Figures 11 and 12. Intriguingly, we observe minor deviations from the ground truth for values up to $\alpha \sim 20$, which is far beyond the range the model has been trained on. Furthermore, the tangent on the drag-lift Pareto front is well captured up to $\Lambda = 50$, which is again out-of-range compared to the training set. For larger $\Lambda > 50$ we observe larger divergence from the ground-truth. Finally, we report pressure coefficient C_p and friction coefficient C_f at a normalized span length of $y/b = 0.5$ for a test case with out-of-range values $\alpha = 20$ and sweep angle $\Lambda = 40$. We also show C_p and C_f for the full 3D wing in Figures 7a and 7b, respectively.

4 Conclusion

We present a first comprehensive dataset of high-fidelity RANS simulations for 3D wings operating in the transonic regime. Each simulation case comes with unique geometry and inflow parameters and provides volume and surface-level flowfields. We evaluate state-of-the-art neural surrogates on our dataset and find that AB-UPT outperforms other state-of-the-art and provides strong performance even for unseen cases. Furthermore, we demonstrate that AB-UPT accurately reproduces drag-lift Pareto fronts even for cases with parameters far beyond its training regime, establishing itself as a contender for real-time wing geometry optimization.

References

[1] Benedikt Alkin, Maurits Bleeker, Richard Kurle, Tobias Kronlachner, Reinhard Sonnleitner, Matthias Dorfer, and Johannes Brandstetter. Ab-upt: Scaling neural cfd surrogates for high-fidelity automotive aerodynamics simulations via anchored-branched universal physics transformers, 2025.

[2] Benedikt Alkin, Richard Kurle, Louis Serrano, Dennis Just, and Johannes Brandstetter. Ab-upt for automotive and aerospace applications. *arXiv preprint arXiv:2510.15808*, 2025.

[3] Neil Ashton, Charles Mockett, Marian Fuchs, Louis Fliessbach, Hendrik Hetmann, Thilo Knacke, Norbert Schonwald, Vangelis Skaperdas, Grigoris Fotiadis, Astrid Walle, Burkhard Hupertz, and Danielle Maddix. Drivaerml: High-fidelity computational fluid dynamics dataset for road-car external aerodynamics, 2025.

[4] Mohamed Elrefaie, Angela Dai, and Faez Ahmed. Drivaernet: A parametric car dataset for data-driven aerodynamic design and graph-based drag prediction. volume Volume 3A: 50th Design Automation Conference (DAC) of *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, page V03AT03A019. Curran Associates, Inc., 08 2024.

[5] Mohamed Elrefaie, Florin Morar, Angela Dai, and Faez Ahmed. Drivaernet++: A large-scale multimodal car dataset with computational fluid dynamics simulations and deep learning benchmarks. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, editors, *Advances in Neural Information Processing Systems*, volume 37, pages 499–536. Curran Associates, Inc., 2024.

[6] Pijush K. Kundu and Ira M. Cohen. *Fluid mechanics / Pijush K. Kundu, Ira M. Cohen ; with contributions by P.S. Ayyaswamy and H.H. Hu.* Academic Press, Amsterdam ;, 4th ed. edition, 2008.

[7] John D. Jr. Anderson. *Fundamentals of Aerodynamics*. McGraw-Hill Education, New York, 6th edition, 2017.

[8] Haixu Wu, Huakun Luo, Haowen Wang, Jianmin Wang, and Mingsheng Long. Transolver: A fast transformer solver for pdes on general geometries. In *International Conference on Machine Learning*, 2024.

[9] Huakun Luo, Haixu Wu, Hang Zhou, Lanxiang Xing, Yichen Di, Jianmin Wang, and Mingsheng Long. Transolver++: An accurate neural solver for pdes on million-scale geometries, 2025.

[10] Steven L. Brunton, Bernd R. Noack, and Petros Koumoutsakos. Machine learning for fluid mechanics. *Annual Review of Fluid Mechanics*, 52(Volume 52, 2020):477–508, 2020.

[11] Florent Bonnet, Ahmed Jocelyn Mazari, Paola Cinnella, and Patrick Gallinari. Airfrans: High fidelity computational fluid dynamics dataset for approximating reynolds-averaged navier-stokes solutions, 2023.

[12] Rohit Sunil Kanchi, Benjamin Melanson, Nithin Somasekharan, Shaowu Pan, and Sicheng He. Unifoil: A universal dataset of airfoils in transitional and turbulent regimes for subsonic and transonic flows, 2025.

[13] Dakota Ramos, Andrew Glaws, Ryan King, Bumseok Lee, Olga Doronina, James Baeder, Ganesh Vijayakumar, and Zachary Grey. Airfoil computational fluid dynamics - 9k shapes, 2 aoa's. Open Energy Data Initiative (OEDI), National Renewable Energy Laboratory (NREL), https://doi.org/10.25984/2222587, 2023. Accessed: 2025-10-30.

[14] Alessandro Schillaci, Maurizio Quadrio, and Giacomo Boracchi. A database of cfd-computed flow fields around airfoils for machine-learning applications, 2021. Version 1.0.

[15] Nicholas Sung, Steven Spreizer, Mohamed Elrefaie, Kaira M. Samuel, Matthew C. Jones, and Faez Ahmed. Blendednet: A blended wing body aircraft dataset and surrogate model for aerodynamic predictions. *CoRR*, abs/2509.07209, 2025.

[16] H. G. Weller, G. Tabor, H. Jasak, and C. Fureby. A tensorial approach to computational continuum mechanics using object-oriented techniques. *Computer in Physics*, 12(6):620–631, 11 1998.

[17] L. S. Caretto, A. D. Gosman, S. V. Patankar, and D. B. Spalding. Two calculation procedures for steady, three-dimensional flows with recirculation. In Henri Cabannes and Roger Temam, editors, *Proceedings of the Third International Conference on Numerical Methods in Fluid Mechanics*, pages 60–68, Berlin, Heidelberg, 1973. Springer Berlin Heidelberg.

[18] OpenFOAM Wiki. Oneram6 by michael alletto — openfoam wiki,, 2024. [Online; accessed 12-November-2025].

[19] Charles Ruizhongtai Qi, Hao Su, Kaichun Mo, and Leonidas J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 77–85. IEEE Computer Society, 2017.

[20] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *ICLR*, 2021.

[21] William S. Peebles and Saining Xie. Scalable diffusion models with transformers. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4172–4182, 2022.

[22] Xiangning Chen, Chen Liang, Da Huang, Esteban Real, Kaiyuan Wang, Hieu Pham, Xuanyi Dong, Thang Luong, Cho-Jui Hsieh, Yifeng Lu, and Quoc V. Le. Symbolic discovery of optimization algorithms. In Alice Oh, Tristan Naumann, Amir Globerson, Kate Saenko, Moritz Hardt, and Sergey Levine, editors, *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*, 2023.

[23] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 4171–4186. Association for Computational Linguistics, 2019.

A Drag/Lift coefficients

In this section we provide a definition of the drag and lift coefficients, commonly used in aerospace engineering. Aerodynamic forces quantify interesting properties of a wing such as drag and lift, which are the result of an integration of the surface pressure p_s and friction $\boldsymbol{\tau}$, i.e., force per unit area exerted by the fluid on the surface, acting tangential (parallel) to the surface. These forces are obtained from the total force acting on an object in an airflow, which is given by

$$\boldsymbol{F} = \oint_S \Big(-(p_s - p_\infty)\boldsymbol{n} + \boldsymbol{\tau} \Big) dS \;, \tag{1}$$

where p_s is the surface pressure, p_∞ the free stream pressure, $\boldsymbol{n}$ the surface normal vector, and $\boldsymbol{\tau}$ the contribution of friction. Drag and lift coefficients are defined as dimensionless numbers

$$C_\text{D} = \frac{2\,\boldsymbol{F} \cdot \boldsymbol{e}_\text{drag}}{\rho\, v^2 A_\text{ref}},\; C_\text{l} = \frac{2\,\boldsymbol{F} \cdot \boldsymbol{e}_\text{lift}}{\rho\, v^2 A_\text{ref}}, \tag{2}$$

and often used in engineering [3], where $\boldsymbol{e}_\text{drag}$ is a unit vector into the free stream direction, $\boldsymbol{e}_\text{lift}$ a unit vector into the lift direction perpendicular to the free stream direction, ρ the density, v the magnitude of the free stream velocity, and A_ref a characteristic reference area for a certain geometry. Further, the acting force $\boldsymbol{F}$ can be converted into dimensionless numbers as drag and lift forces

$$\boldsymbol{F}_\text{drag} = \boldsymbol{F} \cdot \boldsymbol{e}_\text{drag},\; \boldsymbol{F}_\text{lift} = \boldsymbol{F} \cdot \boldsymbol{e}_\text{lift} \;. \tag{3}$$

When simulating an airplane in flight, the angle of attack α is a common boundary condition which changes free stream $\boldsymbol{e}_\text{drag}$ and lift direction $\boldsymbol{e}_\text{lift}$ as given by the unit velocity vector $\boldsymbol{u}_\infty$ as

$$\boldsymbol{u}_\infty = \begin{bmatrix} \cos(\alpha) \\ 0 \\ \sin(\alpha) \end{bmatrix},\; \boldsymbol{e}_\text{drag} = \frac{\boldsymbol{u}_\infty}{\|\boldsymbol{u}_\infty\|_2},\; \boldsymbol{e}_\text{lift} = \boldsymbol{e}_\text{drag} \times \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \;, \tag{4}$$

where $\|.\|_2$ is the Euclidian norm. For $\alpha = 0$ these reduce to $\boldsymbol{e}_\text{drag} = (1,0,0)$ and $\boldsymbol{e}_\text{lift} = (0,0,1)$.

B Implementation Details

We perform a hyperparameter search over learning rate $\texttt{lr} \in \{1e-5, 3e-5, 5e-5, 1e-4\}$ for all methods and select the best performing one. We train for 10 epochs with a learning rate of $1e-5$ with linear warmup for the first 5% of training and a cosine decay thereafter. We train in `float16` precision using the Lion optimizer [22] with a weight decay of $\lambda = 0.05$ and gradient clipping of 0.25. We preprocess the different fields using z-score normalization except for vorticity where we apply normalization by the average magnitude. We elaborate on the method-specific details for the different methods as follows.

AB-UPT. For AB-UPT we use a hidden dimension of 192 and and set the number of anchor points for both volume and surface branches to $16,384$. Before processing the geometry we subsample it to $65,536$ points and perform supernode pooling to $16,384$ points. The hidden dimensionality is set to 192 and the total number of parameters for this model amounts to 7.1 million.

Transformer. We use a Transformer architecture based on coordinates encoded via continuous sine-cosine positional embeddings [23] and a hidden dimension of 192. We employ 16 Transformer blocks with three attention heads, resulting in a parameter count of roughly 7.4M parameters.

Transolver. The Transolver baseline integrates the attention mechanism from [8] into the Transformer baseline. We use 512 slices for the Transolver attention. The remainder of the architecture is the same as for the Transformer baseline, resulting in a parameter count of 7.5M parameters.

PointNet. We use a hidden dimension of 192 and an input embedding of 96. We use a global position embedding of 3072 dimensions, resulting in a model size of roughly 8.3M parameters.

C Ablation studies

Our dataset incorporates variations in geometric design parameters that are *observable* through the point cloud and *non-observable* inflow conditions (e.g., inflow Mach number). We therefore performed an ablation study on the best-performing model, AB-UPT, to assess the impact of conditioning

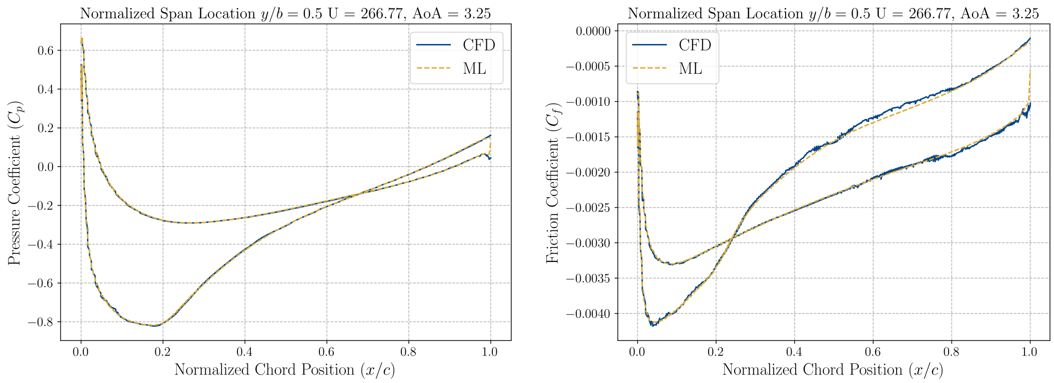


Figure 4: Comparison between CFD data (blue) and AB-UPT prediction (dashed yellow) for the pressure coefficient (C_p) and friction coefficient (C_f) on the wing surface at a normalized span location of $y/b = 0.5$ for a case of the OOD test set. 3D visualizations are shown in Figure 5.

on these parameters (relative L2 error, Table 3). The results demonstrate that the AB-UPT architecture effectively infers information about the geometry directly from the point cloud, while non-observable parameters like inflow conditions are required as explicit conditioning inputs.

Table 3: Impact of geometry and inflow conditioning on the AB-UPT surrogate model's performance, measured by relative L2 error (%) on the *in-distribution* test set. All models were trained for 10 epochs.

Conditioning		**Relative L2 error** (%)				
Geometry	Inflow	p_s	$\boldsymbol{\tau}$	p_v	$\boldsymbol{u}$	$\boldsymbol{\omega}$
✓	✓	0.53	4.16	0.49	3.36	10.5
✓	✗	0.54	4.16	0.51	3.42	10.6
✗	✓	10.8	43.5	9.58	29.5	33.7
✗	✗	10.9	43.7	9.59	29.6	34.3

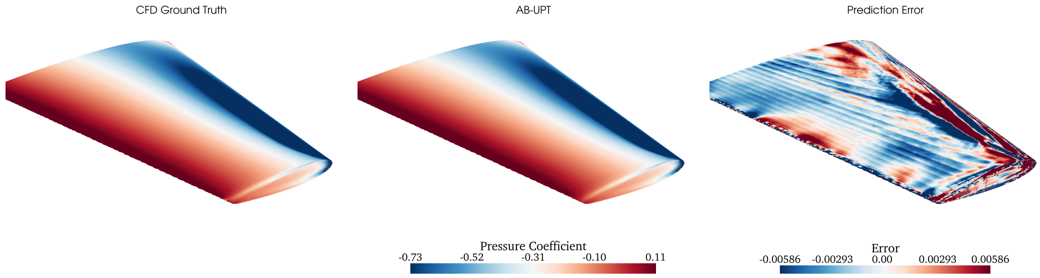
D Additional results

In Figure 4, we show the pressure and friction coefficients predicted by AB-UPT and compare it to the CFD simulation, showing that AB-UPT yields highly accurate predictions.

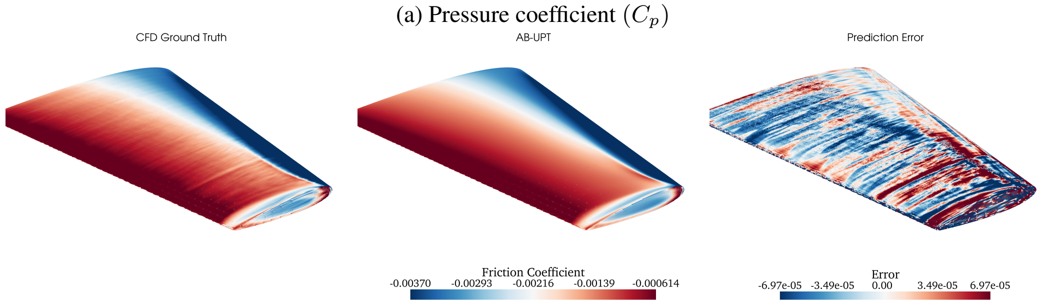
E Data visualizations

In this section, we visualize several cases obtained by our numerical simulations. An important aspect when generating a dataset is quality control to ensure that all simulations have converged to meaningful solutions. Establishing convergence criteria for transonic wing simulations is challenging. We employed classical metrics including monitoring of residuals and force/moment coefficients on the wing. However, these indicators may not always distinguish between physically unsteady and/or discontinuous flows and numerical artifacts.

To supplement these conventional approaches, we employed AB-UPT as an additional quality control tool. The model's prediction error served as an indicator of data quality—cases with anomalously high errors were flagged as potentially unconverged. This approach successfully identified failed cases where classical convergence monitors had been ambiguous. Manual inspection confirmed these cases exhibited spurious flow patterns, inadequate boundary layer mesh resolution, or early numerical divergence (Figure 13). Removing these contaminated samples ensured the final dataset comprised numerically consistent solutions.



(a) Pressure coefficient (C_p)



(b) Friction coefficient (C_f)

Figure 5: Comparison between surface field coefficients on the wing's surface of the CFD (left), AB-UPT surrogate (center) and the error between them (right). The case presented is from the the extrapolation test set with geometry and inflow conditioning parameters within the training range. Corresponding surface pressure and friction profile plots at a certain span length are shown in Figure 4.

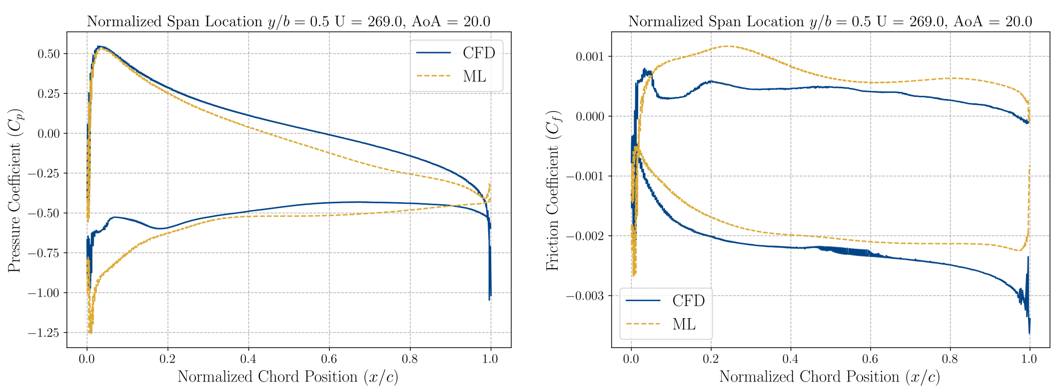
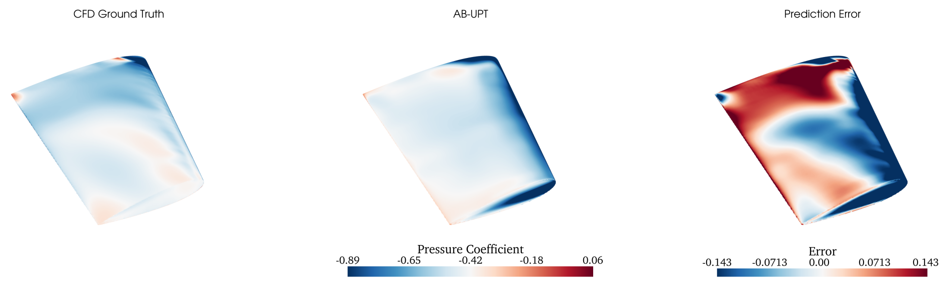
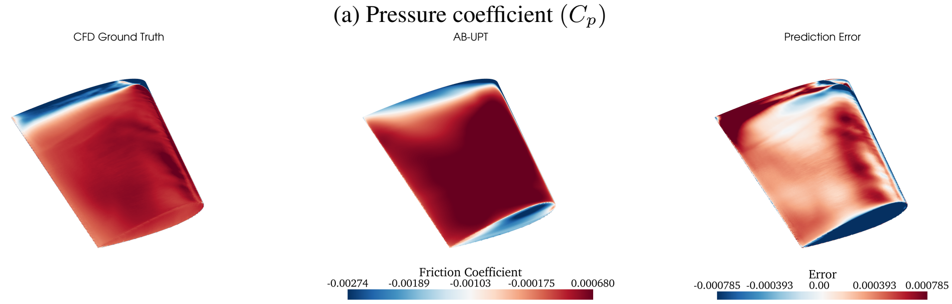


Figure 6: Comparison between CFD data (blue) and AB-UPT prediction (dashed yellow) for the pressure coefficient (C_p) and friction coefficient (C_f) on the wing surface at a normalized span location of $y/b = 0.5$ for a case of the parameter scans. 3D visualizations are shown Figure 7.

In Figure 14, we show the diversity of the dataset by visualizing the wing geometry and the corresponding surface pressure and volume velocity streamlines for various geometry parameters and inflow conditions.

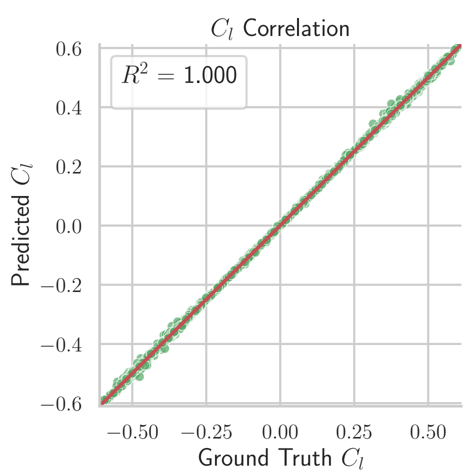


(a) Pressure coefficient (C_p)

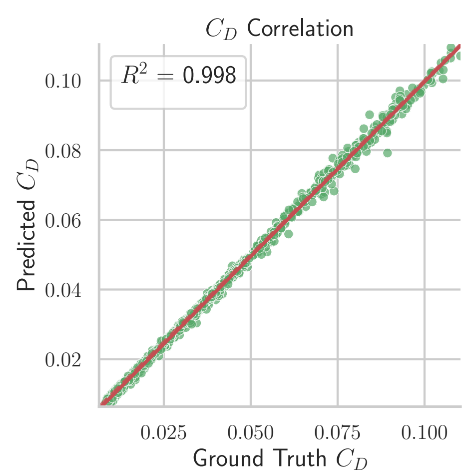


(b) Friction coefficient (C_f)

Figure 7: Comparison between surface field coefficients on the wing's surface of the CFD (left), AB-UPT surrogate (center) and the error between them (right). The case presented is from the the parameter scan with $\Lambda = 40$ and $\alpha = 20$ far outside the training range. Corresponding surface pressure and friction profile plots are shown in Figure 6.



(a) Predicted vs. ground-truth C_l for OOD test set.



(b) Predicted vs. ground-truth C_D for OOD test set.

Figure 8: Correlation of predicted C_D and C_l from AB-UPT to the ground truth for all cases in the OOD test set. AB-UPT's predictions closely match the ground truth.

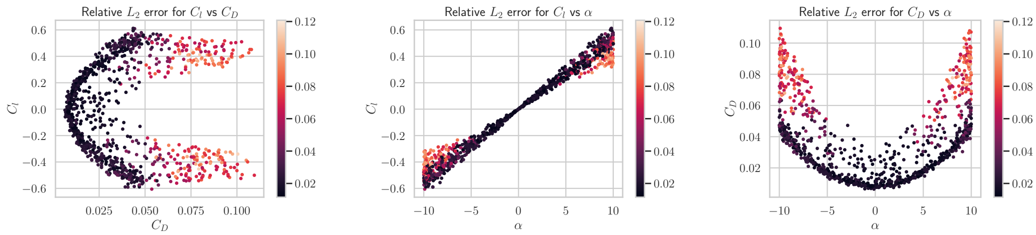


Figure 9: Distribution of relative L_2 error for C_D versus C_l (left), C_D versus α (middle), and C_l versus α (right) for the OOD test set. Most error is introduced in high C_D regimes, absent from the Pareto front.

Figure 10: Relative L_2 error for p_s (left), τ (middle), and averaged over both (right) for parameter scans over α and Λ. AB-UPT is most error-prone at out-of-range values for $\Lambda \in \{60, 70\}$ and $\alpha \in \{-26, -28, -30\}$. For other out-of-range parameters AB-UPT is surprisingly stable.

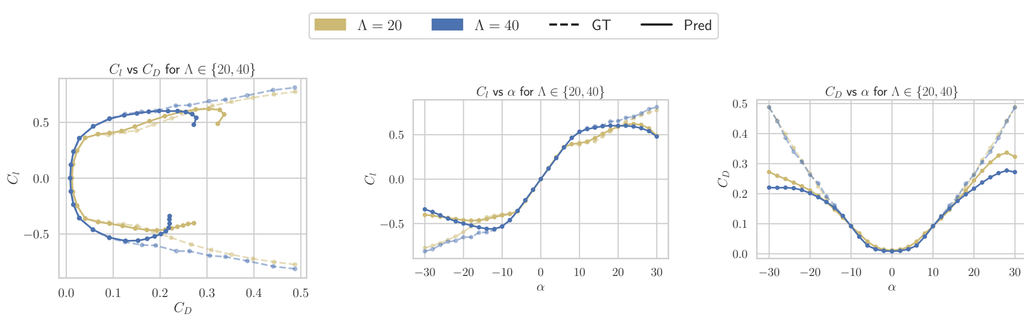


Figure 11: Ground-truth vs predicted C_l and C_D for parameter scans over α and $\Lambda \in \{20, 40\}$. AB-UPT reproduces the drag-lift Pareto front well (left), but introduces error in high regimes of $\alpha > 20$ for both C_l (middle) and C_D (right).

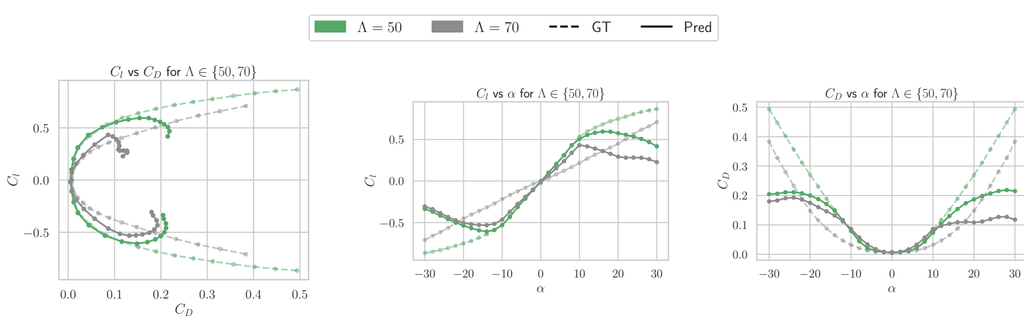


Figure 12: Ground-truth vs predicted C_l and C_D for parameter scans over α and $\Lambda \in \{50, 70\}$. AB-UPT reproduces the drag-lift Pareto front for the most part (left), but error accumulates especially for higher values of α and $\Lambda = 70$ for both C_l (middle) and C_D (right).

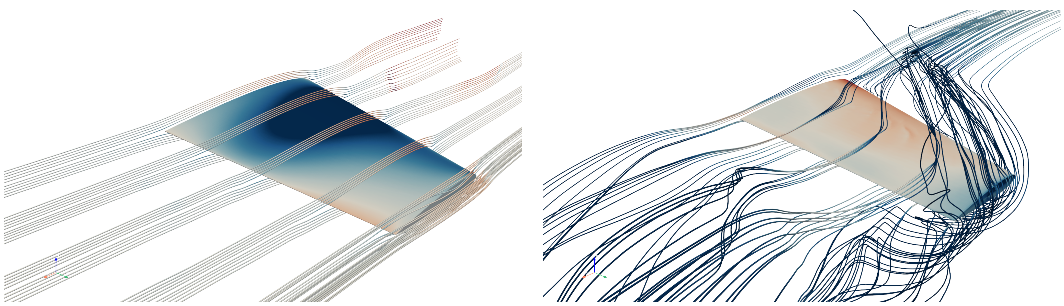
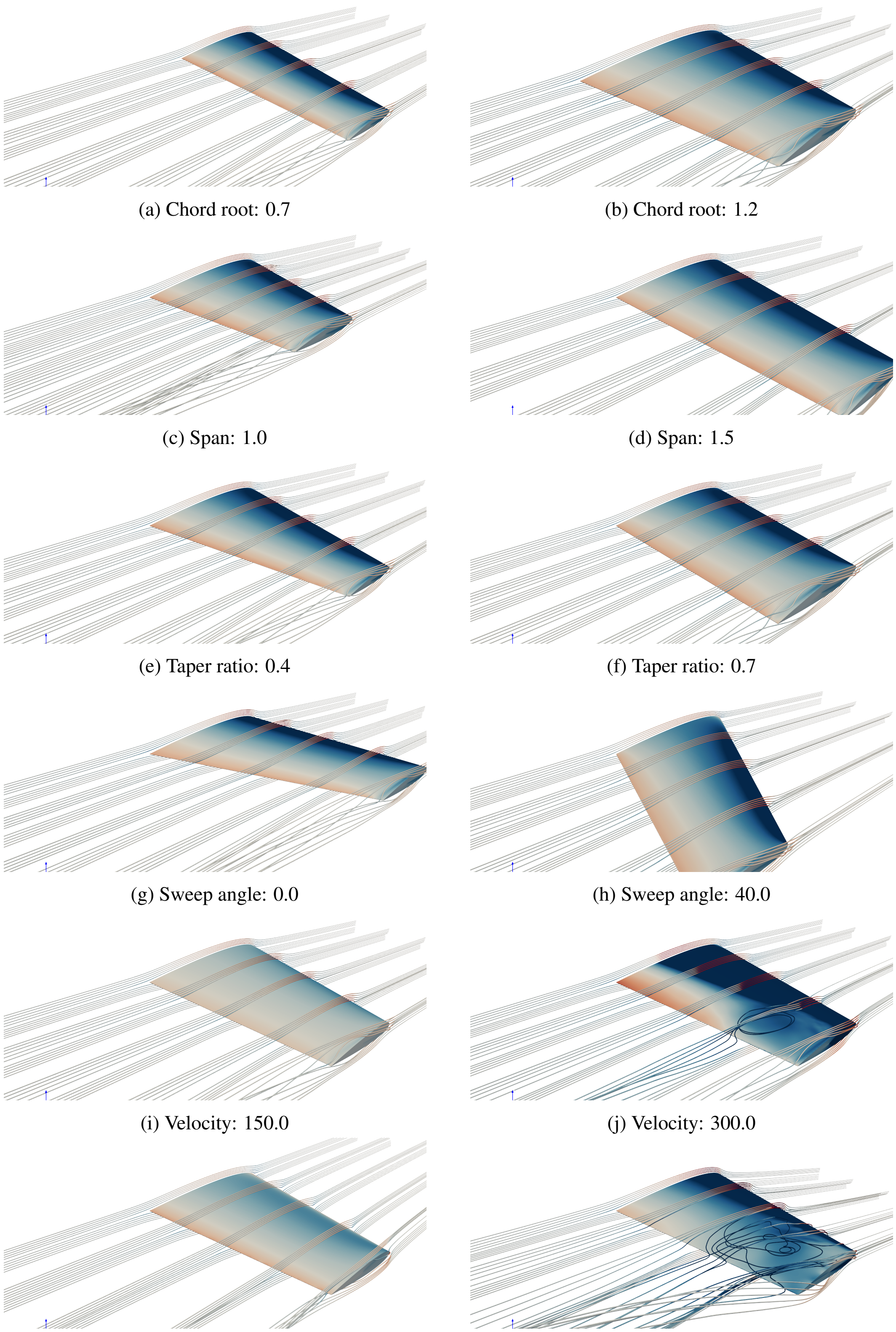


Figure 13: Illustration of failed cases caused by spanwise mesh quality variations (left) and numerically diverging case (right).



(a) Chord root: 0.7
(b) Chord root: 1.2
(c) Span: 1.0
(d) Span: 1.5
(e) Taper ratio: 0.4
(f) Taper ratio: 0.7
(g) Sweep angle: 0.0
(h) Sweep angle: 40.0
(i) Velocity: 150.0
(j) Velocity: 300.0
(k) Angle of attack: -2.0
(l) Angle of attack: 10.0

Figure 14: Wing geometry with corresponding surface pressure and volume velocity fields for various geometry design parameters and inflow conditions. In each row, the minimum and maximum values for each parameter are visualized, while all other parameters are set to their mean value. Except for k and l, the angle of attack shown is 4 degrees.