A SURROGATE-BASED OPTIMIZATION FRAMEWORK AUGMENTED BY GENERATIVE AI

A PREPRINT

Yilin Zhuang
Department of Aerospace Engineering
University of Michigan
Ann Arbor, MI 48105
ylzhuang@umich.edu

Amirpasha Hedayat
Department of Aerospace Engineering
University of Michigan
Ann Arbor, MI 48105
ahedayat@umich.edu

Christian Jacobsen
Department of Aerospace Engineering
University of Michigan
Ann Arbor, MI 48105
csjacobs@umich.edu

April 17, 2025

ABSTRACT

This research explores the deployment of a conditioned generative model for the synthesis of realistic airfoil shapes, aimed at optimizing aerodynamic performance by approximating the global minimum within the design space. By employing a surrogate model configured as a neural network, the study achieves initial and iterative refinement in the evaluation of airfoils through both offline training and online optimization processes. Principal Component Analysis serves as a dimensionality reduction method, facilitating a smooth representation of airfoil geometries.

The study demonstrates the generative model's ability to consistently generate viable airfoil configurations that meet predefined lift coefficient thresholds, illustrating its effectiveness in producing aerodynamically efficient designs.

The findings highlight the potential of constrained generative models in aerodynamic shape optimization, indicating their capacity to expedite the optimization process and encourage the discovery of innovative airfoil designs. Future work could involve incorporating geometric constraints, aiming to generate airfoils that meet structural requirements.

Keywords Optimization · Generative AI · Diffusion model · Surrogate model

1 Background & motivation

Aerodynamic shape optimization, rooted in computational fluid dynamics (CFD), stands as a fundamental element in enhancing the efficiency and performance of contemporary aircraft design. Conventional tools within this domain, such as high-cost CFD solvers, undertake substantial computational efforts to simulate and compute the target aerodynamic coefficients for each design. However, this computationally intensive process often renders these designs impractical, imposing constraints on their feasibility. Therefore, to address these challenges, a spectrum of optimization techniques has been proposed to foster more efficient aerodynamic shape optimization. Noteworthy among these approaches are surrogate-based design optimization (Jones et al. [1998], Li et al. [2019]), adjoint-based high-fidelity optimization (He et al. [2019]), data-based design optimization (Sekar et al. [2019]), and hybrid methods (Li and Cai [2020]).

Of the optimization methods mentioned, surrogate-based optimization emerges as a prominent choice, primarily attributed to its straightforward implementation (as reviewed by Yondo et al. [2018]). Surrogate models serve as effective

proxies for costly physical simulations, such as CFD in aerodynamics, making the optimization loop considerably faster. In addition, dimensionality reduction methods, such as Principal Component Analysis (PCA) (Poole et al. [2015]), could be applied to map the input data onto a substantially reduced-order subspace. Moreover, in the realm of surrogate-based optimization, dedicated attention must be directed towards refining the training samples, particularly in airfoil design where the inclusion of abnormal and unrealistic shapes could lead to misleading outcomes. The training process of a surrogate model can be divided into two phases: offline training using initial samples, and online training using infill samples. This study concentrates specifically on the online training phase, introducing an optimization framework based on generative AI to produce more realistic and higher-quality infill samples. The diffusion model aims to approximate the distribution of airfoils, however, there is no guarantee that the generated airfoils will strictly satisfy the design specification, thus, the surrogate model is used to evaluate the generated airfoils and exploiting the learned design space, thereby enhancing the efficiency of the optimization loop. This work draws inspiration from a study conducted by Li et al. [2020], where an optimization framework based on a Deep Convolutional Generative Adversarial Network (DCGAN) was introduced to generate realistic samples for the following high-fidelity CFD evaluation. Our departure from their approach lies in the utilization of a conditioned diffusion model, instructing the model to generate airfoil shapes while satisfying specific aerodynamic requirements, namely the aerodynamic coefficients under defined flow conditions.

Diffusion models Ho et al., originally stemming from statistical physics, describe the process of particle movement from high concentration regions to lower ones and this is typically referred as the forward process. The generative aspect of the diffusion model can be described as the reverse process that particles at lower concentration regions travel back to the higher ones. [Song et al., 2021] has shown that the forward and reverse processes of diffusion model can be characterized by stochastic differential equation (SDE), while the forward process can be described by Brownian motion and the reverse process solves the backward SDE. Analogous to this physical phenomenon, the diffusion model in this study learns the mapping from random noise to the distribution of the latent representation of airfoils, and steers the latent from random noise to the vicinity of latent distribution in the reverse process. Recent work by [Mazé and Ahmed, 2023] has demonstrated the promising application of diffusion models in topology optimization. In this work, we utilize the score-based generative model, a diffusion model variant that offers a unified framework integrating various sampling techniques and controllable generation Song et al. [2021], Karras et al..

For airfoil design and optimization, we adopt the classifier-free guidance (CFG) approach Ho and Salimans. This method involves implementing flow conditions, such as the angle of attack (AOA) and Reynolds number (Re), and design specifications, including lift and drag coefficients, as parallel inputs along with geometry information. These inputs compute scores during the diffusion process. Such conditions serve as 'soft' constraints in the reverse engineering process, guiding the model towards regions that satisfy or closely align with these specifications. Although it is feasible to impose 'hard' geometric constraints, like curvature and thickness, within the reverse process, our study's scope does not cover these constraints. Instead, we rely on the inherent capabilities of the diffusion model to generate feasible airfoil shapes. This approach bears some resemblance to the use of generative adversarial networks (GANs) for generating feasible initial guesses in the airfoil optimization problem, as explored in previous research Li et al. [2020]. However, our study proposes a unique design loop that effectively derives a candidate optimum design from a selected baseline airfoil. This approach depends on the airfoils generated from random starting points to explore the learned airfoil representation. The airfoil that demonstrates the best performance according to the surrogate model then serves as the foundation for high-fidelity evaluations aimed at optimizing airfoil shapes.

The contents of this paper are organized as follows: Initially, in Section 2.1, we elaborate on the intricacies of data preparation and outline the problem setup. Moving forward, Sections 2.2, 2.3, 2.4, and 2.5 introduce the methodology underlying our proposed optimization framework, providing detailed insights into the implementation of both the surrogate and generative models. The optimization loop is presented in Section 2.6, featuring an algorithm that encapsulates our innovative approach. The subsequent Section 3 meticulously presents and analyzes the results derived from this study. Finally, Section 4 contains our concluding remarks, summarizing the key findings and suggesting potential avenues for future research.

2 Problem formulation & methods

The PCA is employed to extract smooth geometry representation of airfoils in the dataset. The diffusion model is utilized to learn the distribution of the parameterized airfoils and their correlations to the flow condition, namely, Re and AOA. In Li et al. [2020], a discriminator is trained to score the generated airfoils. However, in our case, a surrogate model is trained to predict the lift and drag coefficients of the generated shapes. Starting from the flow condition of the selected baseline airfoil (RAE-5212), we gradually decrease the drag coefficient. The stopping criterion is set based on the alignment of lift and drag coefficients within the ensemble of generated airfoils. This workflow is illustrated in Fig. 1.

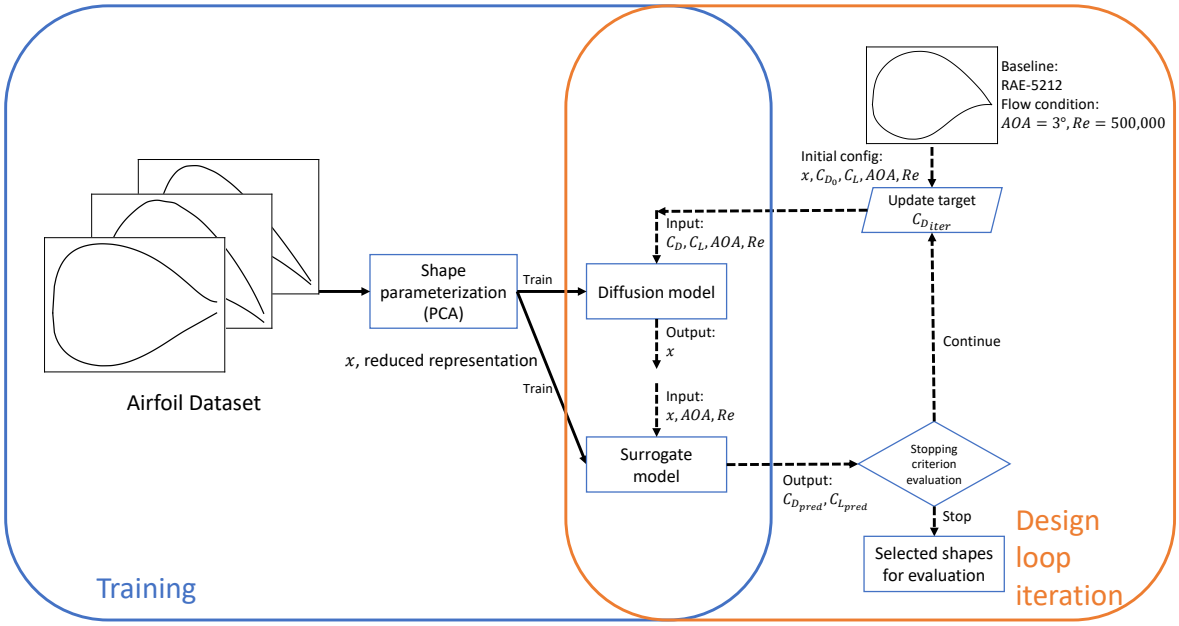

Figure 1: Workflow of the training part and the proposed close loop design optimization.

2.1 Dataset & simulation

The dataset employed in this study to obtain initial samples for training both the surrogate and generative models in the offline phase was sourced from a publicly available GitHub repository (Juangphanich [2022]). This repository contains scaled geometric data from 825 real airfoils[1], simulated using XFoil (Drela [1989]), across various Reynolds numbers (Re) (ranging from 50,000 to 1,000,000) and angles of attack (AOA) (from $-19.75°$ to $24.25°$). This compilation amounts to a total of 767,559 training samples.

XFoil is a simple program to solve the subsonic flow field and calculate the lift and drag coefficients, C_l and C_d, around a 2-dimensional isolated airfoil. Because of its simplifications, XFoil calculation is much faster than a RANS-based CFD solver, making it a fitting choice for handling large datasets, such as the one employed in this study. However, it is essential to note that XFoil may produce less reliable results in certain cases, particularly due to suboptimal N-crit selection—a hyperparameter defining the turbulence intensity of the free flow. The default N-crit value in XFoil is N-crit $= 9$, which may need adjustment based on the expected flow conditions. Given the diverse range of flows in our extensive dataset, spanning from fully laminar to fully turbulent, choosing a fixed N-crit is impractical. For simplicity, the default value of N-crit $= 9$ was retained, acknowledging that some results may deviate from reality.

The outputs of XFoil simulations, namely C_l and C_d, along with the corresponding flow conditions, Re and AOA, were put together to form the dataset for initial model training. Prior to feeding this data into the models, a preprocessing step utilizing the MinMax scaling technique was executed.

2.2 Surrogate model

To address the dimensionality challenges associated with surrogate models, the high-dimensional geometric airfoil data, comprising the coordinates information of 200 points on each airfoil's surface, is initially mapped to a lower-dimensional PCA subspace. Since the x-coordinate (along the chord line) for all airfoils spans from 0 to 1, we only consider the y-coordinate of each point. PCA is a dimensionality reduction technique based on Singular Value Decomposition (SVD), through which an orthogonal basis can be constructed. In this new space, modes are ordered according to their corresponding singular values.

[1] parsed from http://airfoiltools.com

If we denote $\mathbf{X} \in \mathbb{R}^{200}$ as the standardized airfoil data, where rows represent samples and columns correspond to the y-coordinate of each point on the airfoil, the PCA decomposition is expressed as:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{1}$$

where $\mathbf{U}$ is the matrix of left singular vectors, $\mathbf{\Sigma}$ is a diagonal matrix of singular values, and $\mathbf{V}^T$ is the transpose of the matrix of right singular vectors. The sought-after PCA modes are stacked as the columns of $\mathbf{V}^T$. To achieve dimensionality reduction while preserving enough information, we can select the truncation factor, r, to retain a desired level of explained variance:

$$r = \arg\min_r \left(\frac{\sum_{i=1}^{r} \sigma_i^2}{\sum_{i=1}^{p} \sigma_i^2} \geq \text{Threshold} \right) \tag{2}$$

and therefore, the reduced dataset, also known as the PCA coefficients, can be computed from:

$$\mathbf{X}_{\text{reduced}} = \mathbf{X}\mathbf{V}_r^T \tag{3}$$

The surrogate model can more efficiently handle this reduced-order input space compared to the original high-dimensional airfoil geometric data.

For aerodynamic optimization, Kriging methods and their variations [refs from ref1], also known as Gaussian Process Regression (GPR), stand out as the most popular surrogate models. Consequently, our initial approach involved training a basic Kriging model as our surrogate. While Kriging demonstrates effectiveness with relatively small training sets, its computational cost escalates significantly for larger datasets, as in our case, reaching a point where training becomes impractical. In such cases, neural networks present a viable alternative, particularly favored when numerous training samples are available. Hence, we employed a neural network structure featuring two hidden layers comprising 128 and 64 neurons, respectively. A few combinations of these hyperparameters were explored, revealing that the current configuration yields satisfactory results. It is worth noting that further optimization of the network structure remains a possibility to enhance predictive performance. Moreover, the input layer consisted of the airfoil PCA coefficients and the flow conditions, namely Re and AOA, and the output layer composed of two neurons, responsible for predicting the lift and drag coefficients, C_l and C_d.

The surrogate model was trained using a supervised approach, leveraging the availability of true labels in the dataset. This training process aimed to adjust the network weighting factors, w, to minimize the Root Mean Square Error (RMSE):

$$\min_w \frac{1}{N} \sum_{i=1}^{N} \|\hat{y}_i - y_i\|^2 \tag{4}$$

In this equation, N is the number of training samples, y is the true value, and $\hat{y}$ represents the network prediction. More information on the neural network implementation will be presented shortly, after the diffusion model is explained.

2.3 Score-based diffusion model

The scored matching method circumvents the difficulty of computing the normalizing constant, offering a practical approach to estimate the data distribution Hyvärinen. The sampling or the generation of score-based diffusion model is based on solving the forward and reverse SDEs and the following generalized form is proposed Karras et al.:

$$d\boldsymbol{x}_{\pm} = \underbrace{-\dot{\sigma}(t)\sigma(t)\nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x};\sigma(t))dt}_{\text{probability flow ODE}} \pm \underbrace{\beta(t)\sigma(t)^2\nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x};\sigma(t))dt + \sqrt{2\beta(t)}\sigma(t)d\omega_t}_{\text{Langevin diffusion SDE}} \tag{5}$$

where $\{\mathbf{x}(t)\}_{t=0}^{T}$ represents the design space. Here, t denotes the continuous time variable of the diffusion process within the interval $t \in [0, T]$. σ denotes the standard deviation. The term ω indicates the Brownian motion. The $\beta(t)$ denotes the noise scheduling function. The probability density of $x(t)$ is denoted as $p_t(x)$, and $\nabla_x \log p(x;\sigma)$ is the score function. When solving the probability flow (PF) ODE in reverse for generating new airfoils, the score function points to the distribution of airfoil and determines the changes in latent geometry representation, $\boldsymbol{x}$, for each discretized dt. The initial state $x(0)$ corresponds to the undisturbed data, representing the geometry of the original airfoil in the dataset, whereas $x(T)$ represent the prior distribution, which is essentially Gaussian noise. Figure 2 illustrates the forward and backward diffusion processes of a reduced airfoil representation.

Let $D(\boldsymbol{x};\sigma)$ denotes the denoiser function that is optimized by the following training objective Karras et al. to minimize the L_2 denoising error:

$$\mathbb{E}_{y\sim p_{\text{data}}}\mathbb{E}_{\boldsymbol{n}\sim\mathcal{N}(0,\sigma^2 I)}\|D(\boldsymbol{y}+\boldsymbol{n};\sigma)-\boldsymbol{y}\|_2^2, \quad \text{with } \nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x};\sigma) = \frac{D(\boldsymbol{x};\sigma)-\boldsymbol{x}}{\sigma^2} \tag{6}$$

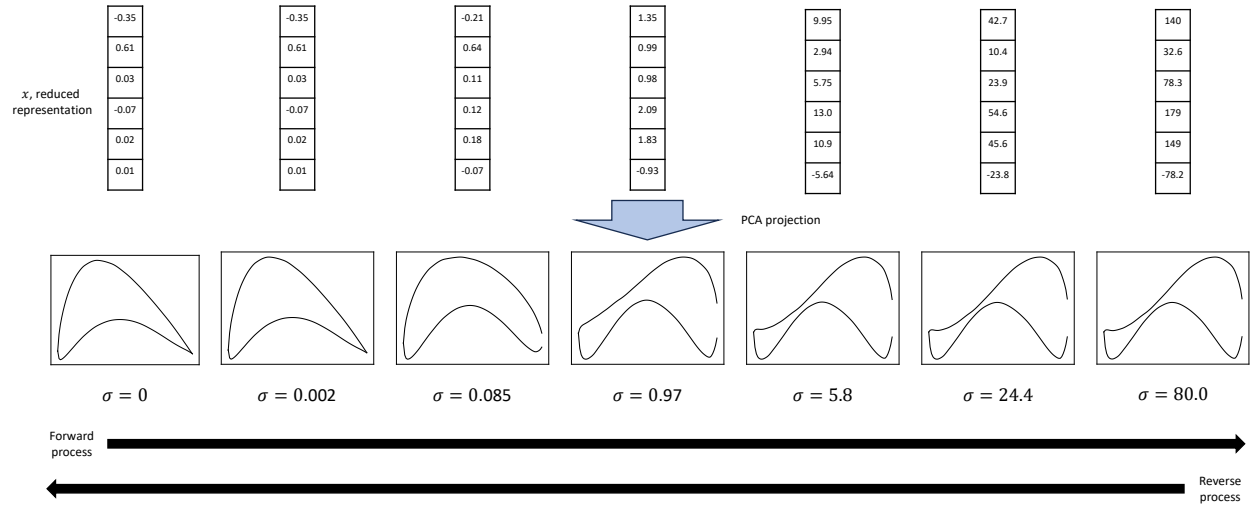

Figure 2: Illustration of forward and reverse processed of the diffusion, noised latent created by adding the product of noise and σ to the reduced representation.

where $\boldsymbol{y}$ denotes the airfoil representation draw from the dataset and n denotes the added noise. Instead of approximating the denoiser function directly with neural network, it was shown that scaling the output of denoising estimator with respect to the noise level, σ, would improve the overall performance, and the following scaling scheme is utilized in the loss function Karras et al.:

$$D_\theta(\boldsymbol{x};\sigma) = c_{skip}(\sigma)\boldsymbol{x} + c_{out}(\sigma)F_\theta(c_{in}\,(\sigma)\boldsymbol{x}; c_{noise}(\sigma)) \tag{7}$$

$$\mathbb{E}_{\sigma,\boldsymbol{y},\boldsymbol{n}}\left[\lambda(\sigma)c_{out}(\sigma)^2\|F_\theta\left(c_{in}(\sigma)\cdot(\boldsymbol{y}+\boldsymbol{n}); c_{noise}(\sigma)\right) - \frac{1}{c_{out}(\sigma)}\left(\boldsymbol{y} - c_{skip}(\sigma)\cdot(\boldsymbol{y}+\boldsymbol{n})\right)\|_2^2\right] \tag{8}$$

where $\lambda(\sigma)$ is a positive weighting function, $c_{out}(\sigma)$, $c_{noise}(\sigma)$, and $c_{in}(\sigma)$ are scaling factors. The F_θ denotes the neural network that is parameterized by θ. For solving the reverse SDE and to generate new airfoils, we solve the following deterministic PF ODE derived from substituting $\sigma(t) = t$ as the noise schedule in Eq. 5, 6

$$d\boldsymbol{x}_- = -t\nabla_{\boldsymbol{x}}\log p_t(\boldsymbol{x};\sigma(t))dt = \frac{\boldsymbol{x} - D_\theta(\boldsymbol{x};\sigma)}{t}dt \tag{9}$$

The advantage of using the PF ODE is its deterministic nature in the reverse path, given the selected unstructured prior, ensuring a more tractable generation process. To enhance the variety of the generated airfoils, increasing the sample size when sampling from the unstructured prior is recommended.

2.4 Conditional generation

Let $\boldsymbol{c}$ denotes the condition vector comprising the drag coefficient, lift coefficient, AOA, and Reynolds number. [Ho and Salimans] has shown that the score of marginal distribution given the condition c can be formulated as:

$$\nabla_{\mathbf{x}}\log p_t(\mathbf{x}(t)|\boldsymbol{c}) = \gamma\underbrace{\nabla_x \log p_t(\mathbf{x}(t)|\boldsymbol{c})}_{conditional\ score} + (1-\gamma)\underbrace{\nabla_{\mathbf{x}}\log p_t(\mathbf{x}(t))}_{unconditional\ score} \tag{10}$$

where γ is a hyperparmeter that determines the influence or 'strength' of the condition, in this study, we set $\gamma = 1$, indicating a full reliance on the conditional score. Then, for the forward SDE and the reverse PF ODE, replacing the $\nabla_x \log p_t(x(t))$ term in Eq.9 gives:

$$d\boldsymbol{x}_- = -t\nabla_{\boldsymbol{x}}\log p_t(\boldsymbol{x};\sigma(t),\boldsymbol{c})dt = \frac{\boldsymbol{x} - D_\theta(\boldsymbol{x};\sigma,\boldsymbol{c})}{t}dt \tag{11}$$

2.5 Neural network implementation

In typical implementations, diffusion models employ a U-Net-based network architecture, leveraging 2D convolutional layers to extract spatial information. However, in our project, we observed that a U-Net-based diffusion model struggles

to produce smooth airfoil shapes. Consequently, we have adopted a ResMLP-based structure, as outlined in Touvron et al. [2021] for constructing both the surrogate and the diffusion models. The surrogate model functions by mapping the PCA-reduced geometric representation of airfoils to their corresponding drag and lift coefficients. On the other hand, the diffusion model is tasked with approximating the score, which encapsulates the reduced-shape and flow conditions of the airfoil. This adaptation, utilizing the modes produced by PCA, addresses the specific challenges in smoothly representing airfoil shapes. The condition information is encoded to the forward pass of the ResMLP through the Feature-wise Linear Modulation Perez et al. [2017].

For the training process, a cosine annealing learning rate scheduler is used for training the model Loshchilov and Hutter [2017], and the batch size for the training the neural networks are set to 512. The surrogate model is trained for 1000 epochs and the diffusion model is trained for 500 epochs.

2.6 Design loop formulation

Algorithm 1 Pseudo code of one close loop iteration

Require: Trained surrogate $f_s : (\mathbf{x}, AOA, Re) \mapsto (C_d, C_l)$, initialized Gen. model $f_{G,\theta} : (\mathcal{N}(\mathbf{0}, \mathbf{I}), c) \mapsto \mathbf{x}$, Acceptable error based on RMSE of surrogate model $\boldsymbol{\epsilon_{error}}$
Model reconciliation **(Not included in this project)**: $\arg\min_\theta \, ||(C_d, C_l) - f_s\left(f_{G,\theta}(C_d, C_l)\right)||$ for $(C_d, C_l) \in$ Database
Initialize **Stopping Criterion** based on errors of models on baseline airfoil, $\mathbf{x}_0$
while True do
  Generating samples: $\mathbf{S} \leftarrow f_{G,\theta}(\mathcal{N}(\mathbf{0}, \mathbf{I}), C_{d,iter}, C_{l,iter}, AOA, Re)$
  $n_{pass} = 0$
  for $\mathbf{x} \in \mathbf{S}$ **do**
    if $||(C_{d,iter}, C_{l,iter}) - f_s\left(\mathbf{x}, AOA, Re\right)|| \leq \boldsymbol{\epsilon_{error}}$ **then**
      $n_{pass} = n_{pass} + 1$
    end if
  end for
  if $n_{pass} \leq$ **Stopping Criterion**, **then**
    Break
  end if
  $C_{d,iter} = C_{d,iter} -$ learning rate
end while
Evaluate $\mathbf{x}$ (Xfoil, N-crit=9), and update the Database

The dataset comprises two variables: chord length and thickness. Consequently, for a given flow condition, multiple local optima could exist as there are various shapes that could meet or in close proximity to the specifications of drag and lift coefficients. In this project, we initialize the optimization using the RAE-5212 as the baseline airfoil, with the drag and lift coefficients $C_D = 0.01086$, $C_L = 0.56230$, measured under flow conditions AOA$= 3^\circ$ and Re$= 500000$. The presented initialized configuration is the most similar one to the initialized baseline in Li et al. [2020].

The procedure of the one design loop iteration is outlined in Algo. 1. Assuming that the surrogate model and the generative model are aligned, the error criterion, $\boldsymbol{\epsilon_{error}}$ is assigned as half of the RMSE of the prediction made by the surrogate model on the full dataset. The termination criterion is set upon the percentage of generated samples which, under baseline flow conditions and measurements, fulfill the error criterion. In this project, the threshold for the stopping criterion is established at half of the initial percentage.

Additionally, the project explores an alternative approach involving Scipy.minimize Virtanen et al. [2020]. The surrogate model is treated as a black box in this context. Both gradient-based (BFGS) and gradient-free (Nelder-Mead) optimization methods were evaluated, given the unconstrained nature of the setup.

3 Results

3.1 Surrogate model

Firstly, we present the outcomes of implementing PCA on the geometric airfoil data, aiming to project them onto a lower-dimensional subspace. Figure 3 provides an illustration of the cumulative explained variance ratio as a function of the number of modes retained in the reduced system. By setting the threshold to %99, it becomes apparent that merely 6 PCA modes are adequate for reconstructing any airfoil shape in the dataset with almost %99 accuracy. Given

the variety of airfoil shapes within the dataset, we anticipate that these same PCA modes offer sufficient representation for airfoils beyond our dataset, including those newly generated by the generative model within the optimization loop. For a more in-depth understanding of the PCA modes, their shapes are visually presented in Fig. 4.

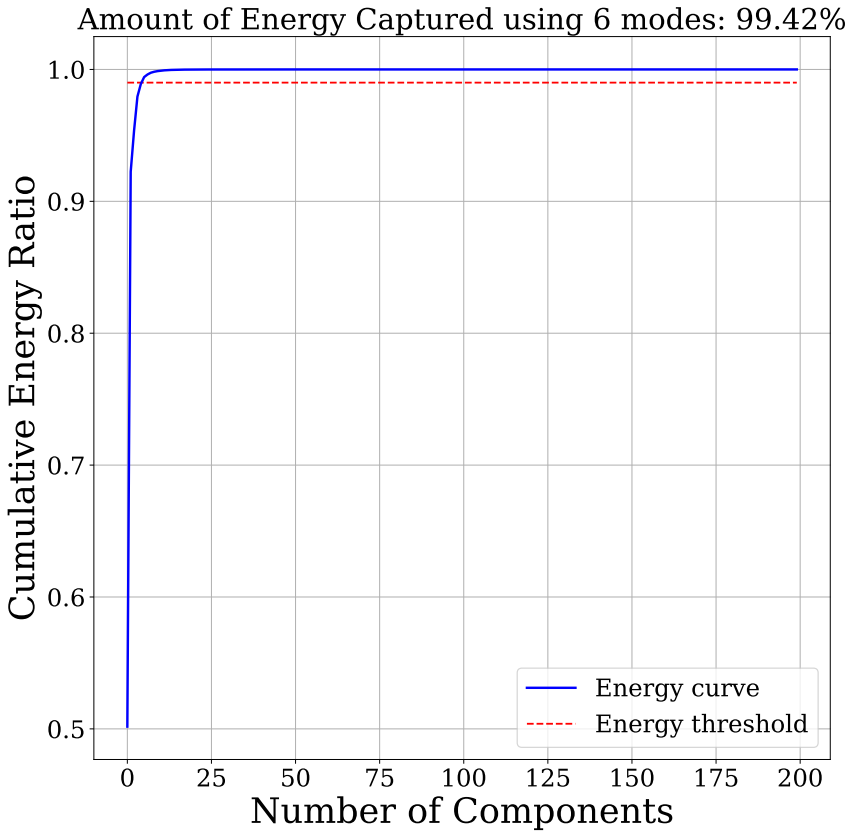

Figure 3: Cumulative explained variance (or in other terms, the energy) ratio of PCA, when applied to the airfoil dataset, as a function of the number of modes considered.

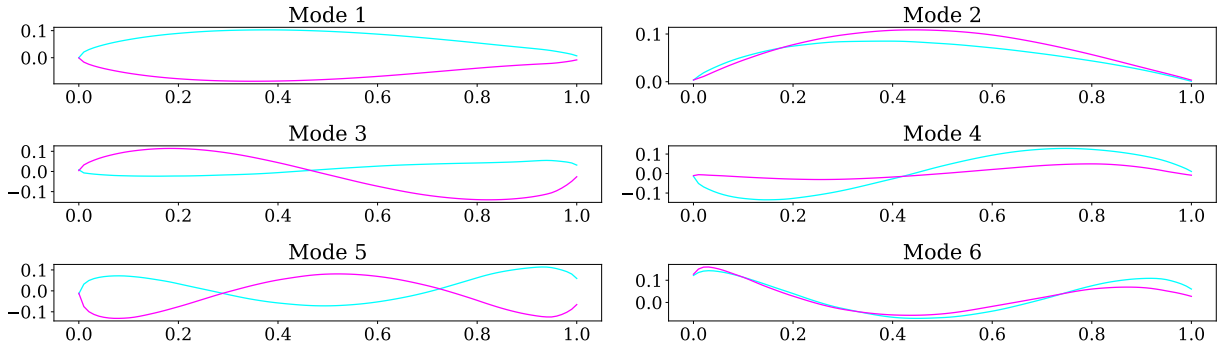

Figure 4: The first 6 PCA modes of the dataset. These capture nearly 99% of the information in the airfoil dataset and can be used as a basis for dimensionality reduction.

The implementation of PCA as a preprocessing step has yielded a significant reduction in the dimensionality of the airfoil shape parameters. Originally residing in a 200-dimensional space, the parameters have now been efficiently

mapped onto a 6-dimensional subspace. This substantial order reduction in the input information streamlines the process of airfoil shape design in a surrogate-based framework.

After the dimensionality reduction achieved through PCA, we delve into assessing the predictive performance of the surrogate model, which was trained using the 6 PCA coefficients representing airfoil parameters along with flow conditions (Re and AOA). The initial dataset underwent a random partitioning into a training set and a validation set. The model was then trained on the designated training set and subsequently tested on the validation set to evaluate its accuracy in predicting aerodynamic coefficients.

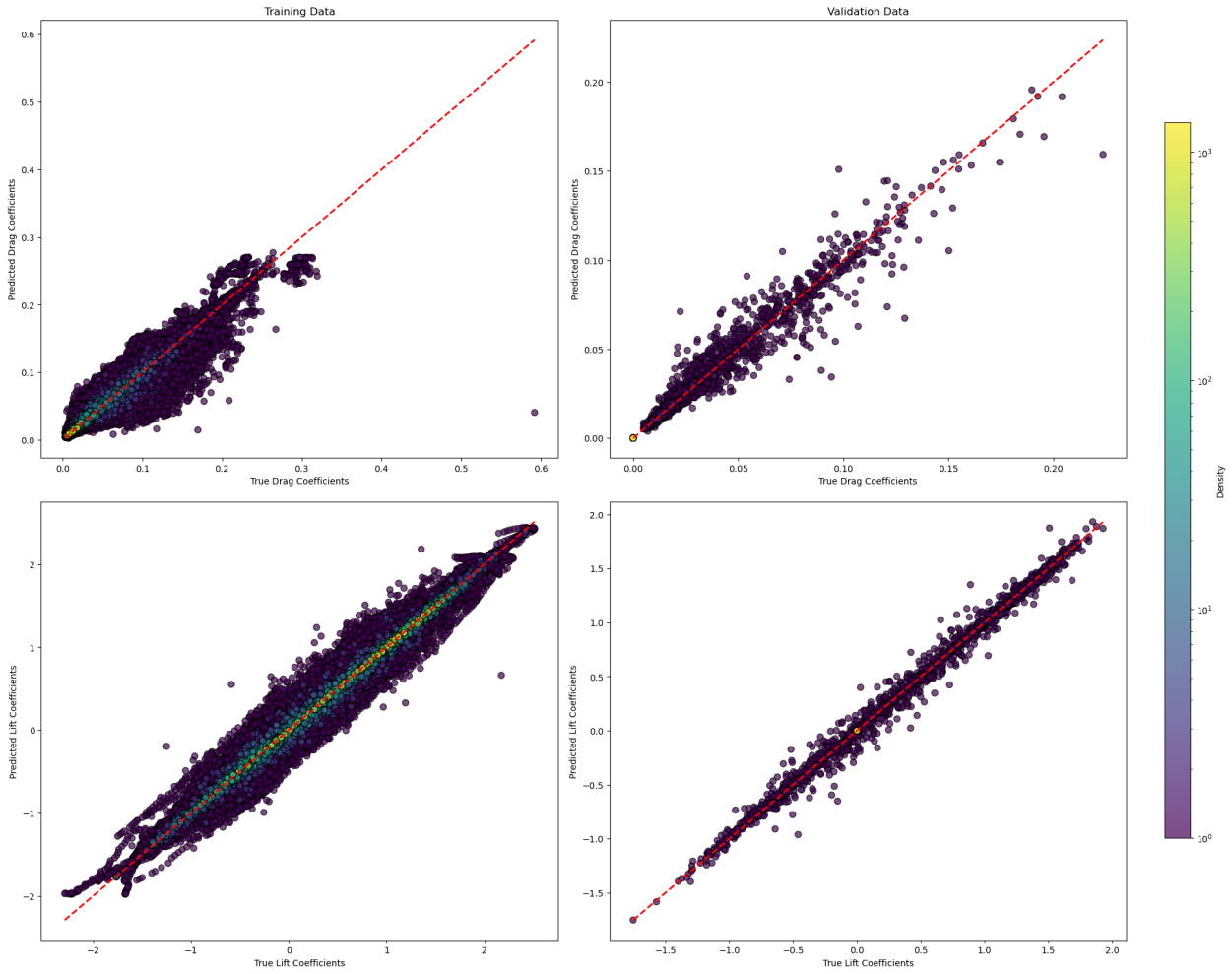

Figure 5: Performance of the surrogate model, C_D (Train) RMSE: 7.547E-3, C_D (Validation) RMSE: 8.74E-4; C_L (Train) RMSE: 6.478E-2, C_L (Validation) RMSE: 7.77E-3.

The results of this analysis are depicted in Fig. 5. Here, the x-axis represents the true labels, while the y-axis corresponds to the predicted values. In an ideal scenario, the model's predictions would align with a 45°-angled straight line (illustrated by the dashed red line). In practice, deviations from this linearity are expected due to various sources of error, stemming from the challenge of approximating a complex phenomenon like flow simulation around an airfoil using a simplified model such as a neural network. Additionally, discrepancies in some true values may arise from the inherent simplifications in XFoil calculations, which does not always produce reliable results. Despite these challenges, the surrogate model demonstrates good accuracy in predicting C_l and C_d on both the training and test sets. Furthermore, the density-based coloring of data points on the plot reveals concentration along the dashed red line, indicating the surrogate model's overall accuracy in a majority of cases. Quantitative insights into the model's performance are provided through the RMSE values provided, offering a numerical basis for comparison across different cases. Note that both C_D and C_L have equal weights in the loss function as they are scaled between 0 and 1 during training. These RMSE values underscore the model's acceptable performance across the training and test sets. Furthermore, as noted in

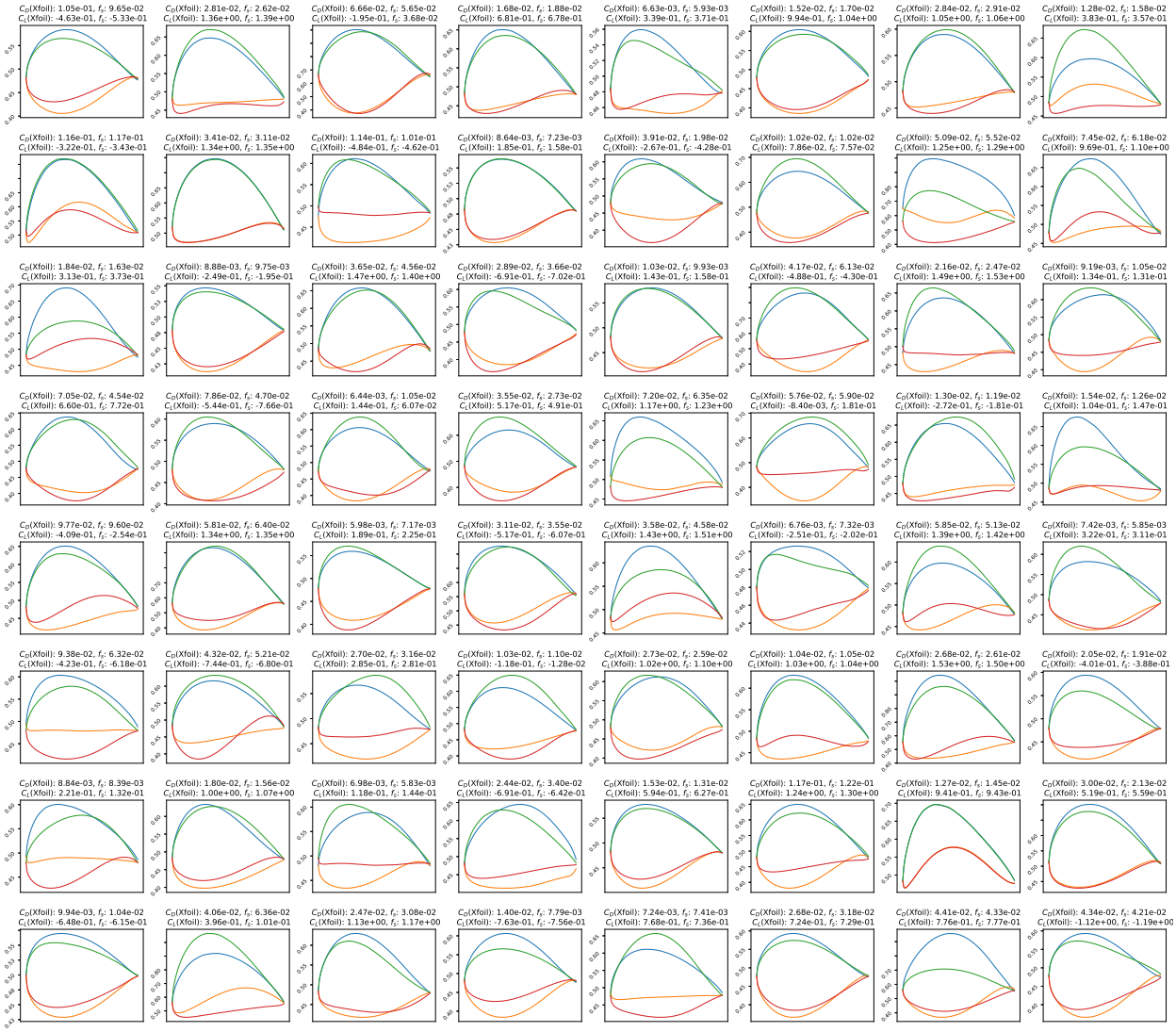

Figure 6: 64 generated samples of randomly selected condition from the full dataset. f_s denotes the surrogate model's predictions on drag and lift coefficients of the generated airfoil. The Blue and orange lines correspond to the original airfoil in the dataset while the green and red lines denote the shapes of generated airfoil.

Algo. 1, The RMSEs of the surrogate model on the drag and lift coefficients are taken into account for constructing the corresponding stopping criterion for the design loop.

The presented results exclusively depict the surrogate model's performance during the offline training phase. As the optimization loop (online phase) begins, more candidate airfoil designs are produced through the generative model, and a certain deviation between the surrogate's predictions and target values is anticipated. This variance will be regulated by the stopping criterion outlined in Algo. 1. Upon reaching the termination condition, the optimization loop concludes, and the generated infill samples are employed to refine the surrogate for subsequent optimization loops.

3.2 Generative model & model alignment

Fig. 6 shows the generated airfoils that generated from the flow conditions in the dataset, it can be seen that there could be multiple airfoil shapes that satisfy the given conditions. Further more, it can be observed that there are several cases where the generative model almost align with the airfoil used in the simulation.

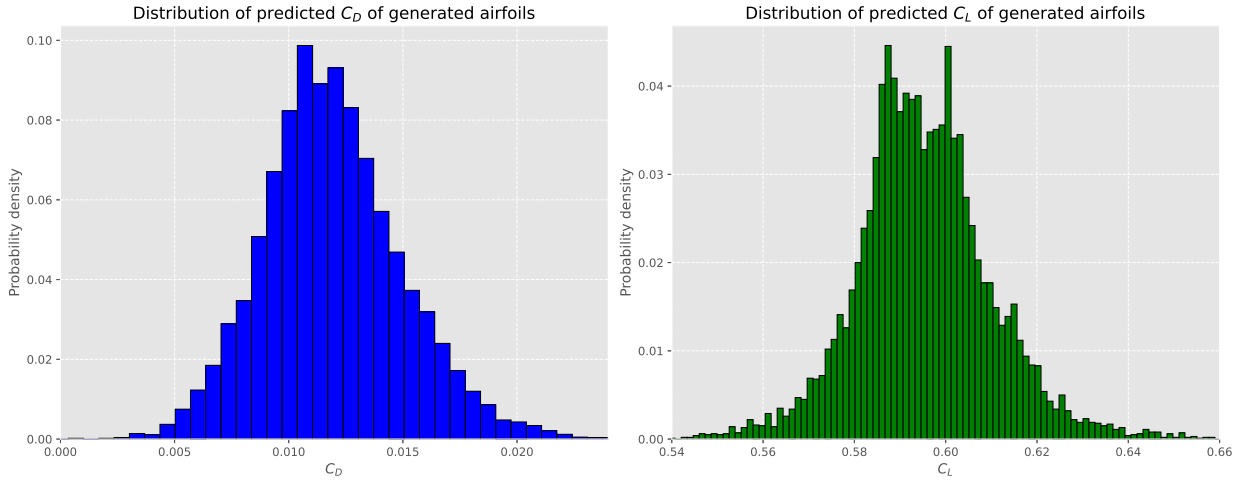

Figure 7: Probability density plot of C_D and C_L of 10000 generated samples wrt the specified baseline condition. Xfoil: C_D=1.09E-2, C_L=0.562, Surrogate: C_D=9.35E-3, C_L=0.565

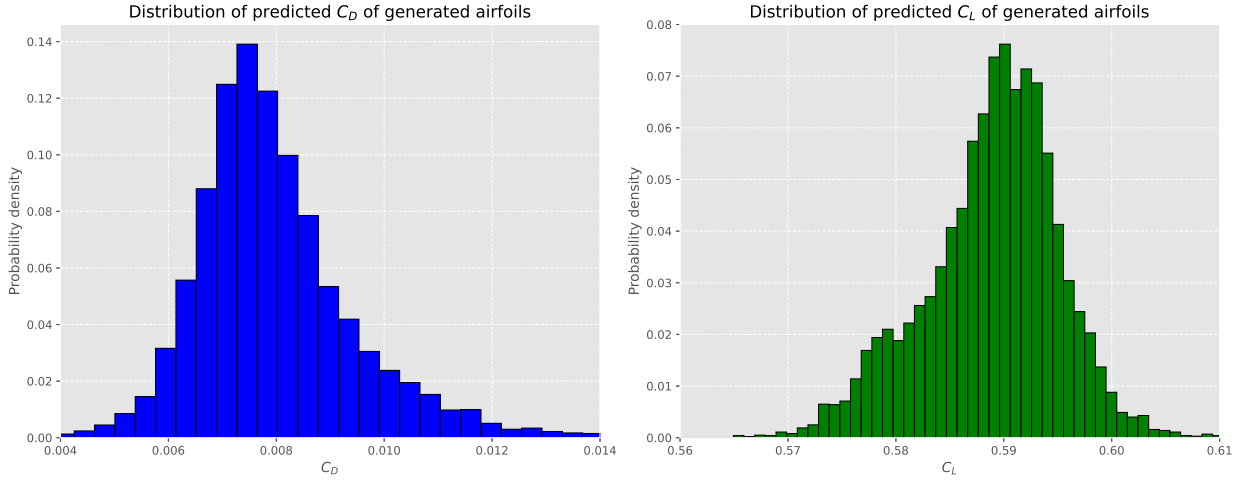

Figure 8: Probability density plot of C_D and C_L of 10000 generated samples wrt the specified baseline condition and the end of design loop iteration.

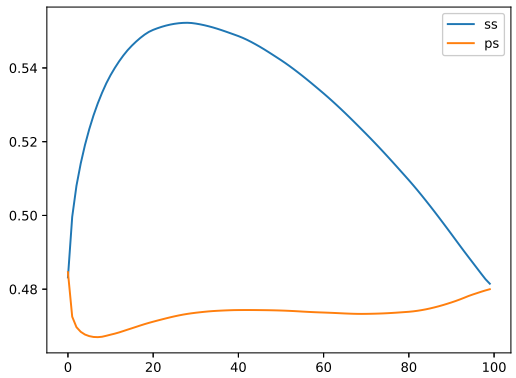

Figure 9: Better design obtained by the design loop. Surrogate prediction: C_D=4.93E-3, C_L=0.563

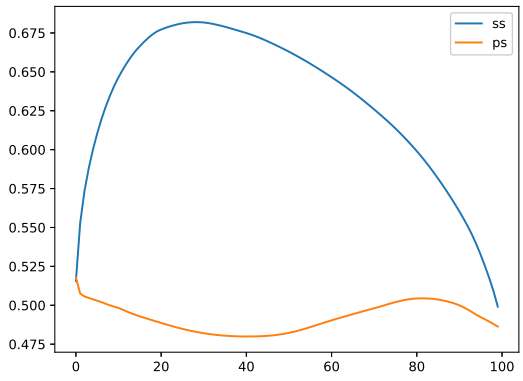

Figure 10: A datapoint/simulation that worth to be added to the dataset.Surrogate prediction: C_D=-7.87E-3, C_L=0.546

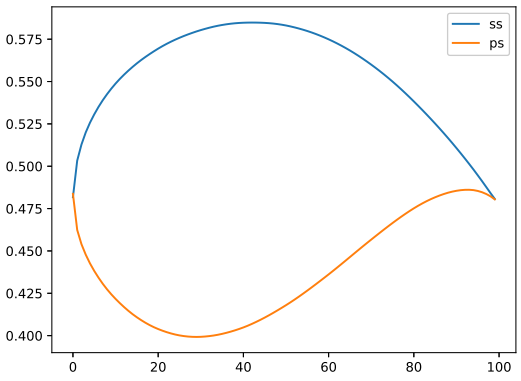

Figure 11: Design obtained by combing the surrogate model with the BFGS method. Surrogate prediction: C_D=5.39E-3, C_L=0.532

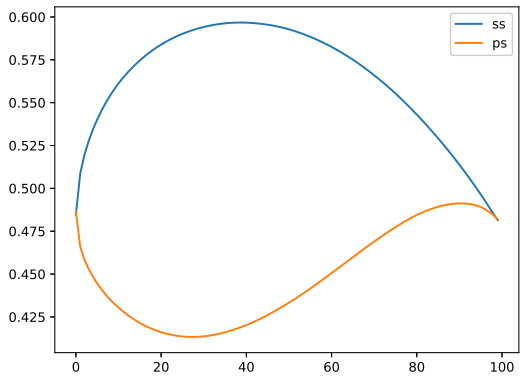

Figure 12: Design obtained by combing the surrogate model with the Nelder-Mead method. Surrogate prediction: C_D=-2.23E-4, C_L=0.710

For the baseline model, the histogram of the surrogate model's predictions on the drag and lift coefficients of 10000 the generated airfoils is shown in Fig. 7. It can be seen from the diagram that the drag coefficient aligns relatively well with that of the condition whereas the lift coefficient slightly deviates from the recorded value (0.562) in Xfoil. Figure 8 presents the histogram detailing the surrogate model's prediction for drag and lift coefficients of 10000 generated airfoils at the end of the design loop iteration.

Table 1: Combined Results of XFoil Simulations and Surrogate Predictions on the four airfoils

	Better Design Candidate		Design Worth Exploration		Scipy - BFGS		Scipy - Nelder-Mead	
	XFoil	**Surrogate**	**XFoil**	**Surrogate**	**XFoil**	**Surrogate**	**XFoil**	**Surrogate**
C_d	7.47E-3	4.93E-3	1.83E-2	-7.87E-3	8.28E-3	5.39E-3	NA	-2.23E-4
C_l	0.585	0.563	0.704	0.546	0.614	0.532	NA	0.710

Figures 9 and 10 display the two designs obtained through the design loop. The optimal design is identified from the generated airfoils that meet specific criteria for the drag and lift coefficients. Conversely, the design worth exploring

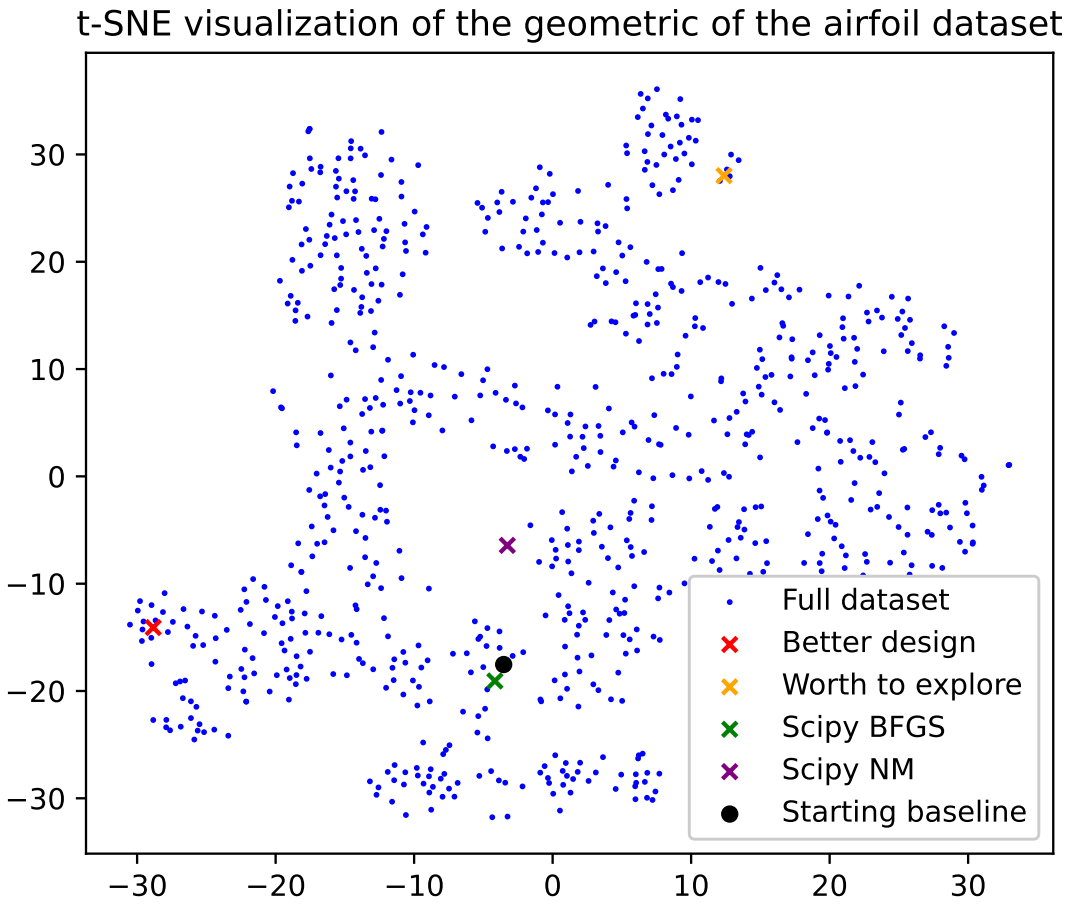

Figure 13: t-SNE plot of the geometry of the airfoil dataset, and the illustration of the better design and the design point that could potentially improve the learned representation in both the diffusion and surrogate models.

is selected based on the airfoil that exhibits the lowest drag coefficient, provided it also meets the lift constraint. Furthermore, Table 1 reveals that the surrogate model's prediction for the exploration-worthy candidate diverges from the XFoil simulation results and the difference is larger than the RMSE of the surrogate model, thus, demonstrating that this is a potential point that could be added to the dataset to help both models to learn a better representation of the airfoil data.

Figures 11 and 12 present the two designs derived by integrating the surrogate model with the BFGS and Nelder-Mead optimization methods, as implemented in Scipy. The stopping criterion for optimization with Scipy is set to be triggered when the drag coefficient is below than the lowest value measured in the dataset. Tab. 1 shows that the airfoil given by Nelder-Meah fails to converge in XFoil due to the overlapping tail, whereas the BFGS method obtain a better design in terms of drag coefficient, however, it does not satisfy the lift coefficient requirement. Such outcomes were anticipated, given the implementation of these methods as unconstrained optimization problems.

Fig 13 illustrates a t-SNE visualization of the original airfoil shapes, highlighting the diversity in designs that meet the specified flow conditions., recall that there could be multiple airfoil designs that satisfy the flow condition, thus, the obtained results could largely deviate from the shape of the baseline. The BFGS method identified a locally optimal design as indicated by the surrogate model, whereas the Nelder-Mead method, along with the diffusion model, were capable of exploring across the distribution of airfoil shapes.

4 Concluding remarks

This study investigates the application of a conditioned generative model for the production of realistic airfoil samples, strategically aimed at approximating the global minimum. The evaluation of generated airfoils in each optimization loop involves the utilization of a surrogate model, structured as a neural network. This surrogate model undergoes an offline training phase and further refinement during the online optimization process. The initial accuracy of the surrogate model is deemed satisfactory, highlighting the effectiveness of the chosen PCA subspace to address the challenges associated with the curse of dimensionality in surrogate-based optimization.

The close loop design demonstrates that it can produce a feasible solution while satisfying the threshold set on the lift coefficients; however, we found that hyperparameter tuning is required for generating physical airfoils. It should be noted that it can be solved by implementing shape constraints in the reverse process of solving the PF ODE.

It is essential to emphasize that this study currently resides in the proof-of-concept stage, with several aspects warranting further attention. Notably, optimization of hyperparameters for the neural networks constitutes a pertinent area for refinement. In future endeavors, the incorporation of physical and geometric constraints can be explored to assess their impact on the quality of airfoils generated by the diffusion model.

Nevertheless, the findings presented in this study offer promising insights into the potential of constrained generative models for aerodynamic shape optimization. These models hold the prospect of significantly reducing the optimization loop duration, and there is potential for the generation of novel and innovative airfoil designs that might be beyond the reach of traditional optimization methods.

5 Acknowledgment

We would like to thank Shugo Kaneko for their assistance in sourcing data and facilitating our communication with other scholars. We are also thankful to Christian Jacobsen for their support in the code development process. They offered valuable ideas about potentially applying shape constraints in the reverse solving of the PF ODE. Although these ideas were not implemented in this report, they present intriguing possibilities for future work. We are very grateful for their help.

References

Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, Dec 1998. ISSN 1573-2916. doi:10.1023/A:1008306431147. URL `https://doi.org/10.1023/A:1008306431147`.

Jichao Li, Jinsheng Cai, and Kun Qu. Surrogate-based aerodynamic shape optimization with the active subspace method. *Structural and Multidisciplinary Optimization*, 59(2):403–419, Feb 2019. ISSN 1615-1488. doi:10.1007/s00158-018-2073-5. URL `https://doi.org/10.1007/s00158-018-2073-5`.

Xiaolong He, Jichao Li, Charles A. Mader, Anil Yildirim, and Joaquim R.R.A. Martins. Robust aerodynamic shape optimization—from a circle to an airfoil. *Aerospace Science and Technology*, 87:48–61, 2019. ISSN 1270-9638. doi:https://doi.org/10.1016/j.ast.2019.01.051. URL `https://www.sciencedirect.com/science/article/pii/S1270963818319072`.

Vinothkumar Sekar, Mengqi Zhang, Chang Shu, and Boo Cheong Khoo. Inverse design of airfoil using a deep convolutional neural network. *AIAA Journal*, 57(3):993–1003, 2019. doi:10.2514/1.J057894. URL `https://doi.org/10.2514/1.J057894`.

Jichao Li and Jinsheng Cai. Massively multipoint aerodynamic shape design via surrogate-assisted gradient-based optimization. *AIAA Journal*, 58(5):1949–1963, 2020. doi:10.2514/1.J058491. URL `https://doi.org/10.2514/1.J058491`.

Raul Yondo, Esther Andrés, and Eusebio Valero. A review on design of experiments and surrogate models in aircraft real-time and many-query aerodynamic analyses. *Progress in Aerospace Sciences*, 96:23–61, 2018. ISSN 0376-0421. doi:https://doi.org/10.1016/j.paerosci.2017.11.003. URL `https://www.sciencedirect.com/science/article/pii/S0376042117300611`.

Daniel J. Poole, Christian B. Allen, and Thomas C. S. Rendall. Metric-based mathematical derivation of efficient airfoil design variables. *AIAA Journal*, 53(5):1349–1361, 2015. doi:10.2514/1.J053427. URL `https://doi.org/10.2514/1.J053427`.

Jichao Li, Mengqi Zhang, Joaquim R. R. A. Martins, and Chang Shu. Efficient aerodynamic shape optimization with deep-learning-based filtering. *AIAA Journal*, 58(10):4243–4259, October 2020. doi:10.2514/1.J059254.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising Diffusion Probabilistic Models. In *Advances in Neural Information Processing Systems*, volume 33, pages 6840–6851. Curran Associates, Inc. URL https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=PxTIG12RRHS.

François Mazé and Faez Ahmed. Diffusion models beat gans on topology optimization. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(8):9108–9116, Jun. 2023. doi:10.1609/aaai.v37i8.26093. URL https://ojs.aaai.org/index.php/AAAI/article/view/26093.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 26565–26577. Curran Associates, Inc. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/a98846e9d9cc01cfb87eb694d946ce6b-Paper-Conference.pdf.

Jonathan Ho and Tim Salimans. Classifier-Free Diffusion Guidance. URL http://arxiv.org/abs/2207.12598.

Paht Juangphanich. airfoil-learning. https://github.com/nasa/airfoil-learning, 2022.

Mark Drela. Xfoil: An analysis and design system for low reynolds number airfoils. In Thomas J. Mueller, editor, *Low Reynolds Number Aerodynamics*, pages 1–12, Berlin, Heidelberg, 1989. Springer Berlin Heidelberg. ISBN 978-3-642-84010-4.

Aapo Hyvärinen. Estimation of non-normalized statistical models by score matching. 6(24):695–709. URL http://jmlr.org/papers/v6/hyvarinen05a.html.

Hugo Touvron, Piotr Bojanowski, Mathilde Caron, Matthieu Cord, Alaaeldin El-Nouby, Edouard Grave, Gautier Izacard, Armand Joulin, Gabriel Synnaeve, Jakob Verbeek, and Hervé Jégou. Resmlp: Feedforward networks for image classification with data-efficient training, 2021.

Ethan Perez, Florian Strub, Harm de Vries, Vincent Dumoulin, and Aaron Courville. Film: Visual reasoning with a general conditioning layer, 2017.

Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts, 2017.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi:10.1038/s41592-019-0686-2.