Combining Bayesian optimization with adjoint-based gradients for efficient control of flow instabilities

By J. S. Müller†, J. M. Reumschüssel†, T. L. Kaiser†, S. J. Knechtel† AND K. Oberleithner†

This work presents an efficient framework for shape optimization to control flow instabilities and coherent structures in laminar and turbulent flows by combining a Bayesian optimization approach with adjoint-based gradient information (BOA). Linear stability and resolvent analyses are used to yield physically meaningful cost functions. In addition, adjoint methods are employed to compute the sensitivity of the cost functions to shape parameter changes. These physics-based function and gradient information are embedded into a global Bayesian optimization framework, which also handles uncertainties that may arise from the provided data or linear model. The BOA framework is tested on the task of designing a hydrofoil to control the wake instability and it is compared against a gradient-free Bayesian optimizer and a purely gradient-based method.

1. Introduction

In order to control unsteadiness of coherent structures in turbulent flows, one can pursue various strategies to optimize flow devices. One approach is the use of surrogate models, which approximate an objective function, in this case the strength of flow unsteadiness, and identify potential minima. These models are informed by measurements of the objective function value within a predefined parameter space via simulations or experiments. A sample-efficient method for achieving this is Bayesian optimization (BO), which makes probabilistic predictions about the objective function from the available test data and uses these predictions to guide the selection of new samples. Although BO is particularly sample efficient compared with other global optimization strategies, design tasks that can be handled in this physics-agnostic manner are usually strongly constrained by the number of free parameters. Otherwise, a prohibitively large number of tests would be required for convergence.

A second approach to controlling flow unsteadiness is through gradient-based optimization methods such as the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm. These techniques modify an initial design based on gradient information to obtain an improved configuration. In the case of addressing flow unsteadiness, an efficient way to obtain physics-based gradient information is through linear methods such as linear stability analysis (LSA) and resolvent analysis (RA), which are powerful tools to model flow dynamics in both laminar and turbulent flows (Sipp *et al.* 2010; Taira *et al.* 2017). The corresponding adjoint sets of equations allow for the calculation of sensitivity in predicted dynamics to changes in shape parameters with very little computational effort. Gradient-based techniques usually appeal through fast convergence, but they tend to converge to local minima of the optimization landscape. In order to identify the global

† Laboratory for Flow Instabilities and Dynamics, Technische Universität Berlin, Germany

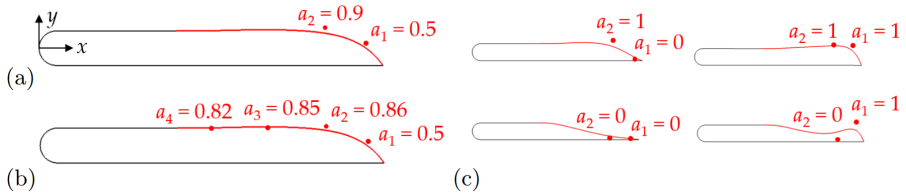

FIGURE 1. Parameterization of the hydrofoil with (a) two parameters in Blake-foil-type configuration, (b) four parameters in Blake-foil-type configuration, and (c) two parameters in extremal configurations. The red curve indicates the variable shape that is to be optimized, and the red dots indicate the B-spline control knots of the variable shape.

optimum, multiple runs can be initiated from disparate initial designs. However, this typically results in a considerable increase in the number of required tests.

In this work, we propose to augment BO with gradient information obtained through linearized and adjoint methods to overcome the limitations of each individual method. The framework integrates physically meaningful cost functions from LSA and RA quantities, growth rate and resolvent gain, respectively, along with the computational efficiency of adjoint-based gradient information that barely increases computation time. Simultaneously, it maintains the global optimization capabilities and uncertainty handling inherent in BO. Our proposed approach, combining BO with adjoint-based gradient information (BOA), thus represents a synthesis of these methodologies.

2. Toy problem for optimization

As a demonstrator case, BOA is tested and compared against other classical optimization methods on a laminar flow past a generic Blake-type hydrofoil with a variable trailing edge shape, similar to the study of Marsden *et al.* (2004). In our setup, the hydrofoil features a round leading edge, a chord length of 0.1 m, a chord-to-height ratio of $c/h = 12$, no camber and a fairly blunt trailing edge in its original baseline design, as shown in Figure 1(a,b). We consider an inflow velocity of $U = 1.5$ m/s with a Reynolds number of $Re = 3300$ and an angle of attack of $\alpha = 0°$. In the following, all quantities are nondimensionalized with the chord length c and the inflow velocity U, if not specified otherwise.

In the present configuration, the hydrofoil wake exhibits a vortex shedding mode as a result of a global instability in the flow field. For shape optimization, the downstream half of the suction side up to the trailing edge is parameterized with a B-spline consisting of d control knots, which can be varied in y direction while the rest of the shape is maintained. Each candidate design is thus specified by a parameter vector $\boldsymbol{a} = [a_1, a_2, \ldots, a_d]^T$, which defines the d y positions of the control knots. This allows us to vary the dimension of the problem d by changing the number of control knots. Figure 1(a,b) shows the hydrofoil with two and four parameters. The numbers given in the figure for each parameter correspond to the baseline Blake-foil-type configuration as considered in Marsden *et al.* (2004). The parameters of $\boldsymbol{a}$ are normalized so that they are bounded between 0 and 1 for each parameter. Figure 1(c) shows the extremal configurations that result from the predefined parameter bounds of the two-parameter case.

3. Linearized methods and shape sensitivity

In the following, we briefly introduce the mathematical fundamentals for the LSA and RA and their corresponding shape sensitivities. We start by considering the 2D incompressible Navier–Stokes equations

$$\mathcal{B}\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t} + \mathcal{N}(\boldsymbol{q}) = \boldsymbol{0}\ , \quad \mathcal{B} = \begin{pmatrix} \mathcal{I} & 0 \\ \boldsymbol{0} & 0 \end{pmatrix}, \quad \mathcal{N} = \begin{pmatrix} (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p - \frac{1}{Re}\nabla^2\boldsymbol{u} \\ \nabla\cdot\boldsymbol{u} \end{pmatrix}\ , \tag{3.1}$$

with the state vector $\boldsymbol{q} = (\boldsymbol{u}, p)^T$ consisting of the velocity field $\boldsymbol{u} = (u, v)^T$ and the pressure field p, and $\mathcal{I}$ being an identity operator. Linearization of $\mathcal{N}$ around $\boldsymbol{q}_0 = (\boldsymbol{u}_0, p_0)^T$ yields the operator

$$\mathcal{A}(\boldsymbol{q}_0)\ \boldsymbol{q}' := \left.\frac{\partial\mathcal{N}}{\partial\boldsymbol{q}}\right|_{q=q_0} \boldsymbol{q}' = \begin{pmatrix} (\boldsymbol{u}'\cdot\nabla)\boldsymbol{u}_0 + (\boldsymbol{u}_0\cdot\nabla)\boldsymbol{u}' + \nabla p' - \frac{1}{Re}\nabla^2\boldsymbol{u}' \\ \nabla\cdot\boldsymbol{u}' \end{pmatrix}\ , \tag{3.2}$$

where $\boldsymbol{q}_0$ is either a base flow or a mean flow. The base flow represents a steady flow state and is obtained by directly solving Eq. (3.1) with $\mathrm{d}\boldsymbol{q}/\mathrm{d}t = 0$. The mean flow is a statistical quantity and represents the time average of an unsteady flow. It is determined by time-averaging flow snapshots from simulations, such as large-eddy simulations, or experiments. Moreover, the mean flow can be obtained by performing Reynolds-averaged Navier–Stokes simulations. Base and mean flows differ due to the fluctuating field modifying the base flow.

For LSA, a normal-mode ansatz is assumed for $\boldsymbol{q}'$, leading to the linear direct and adjoint eigenproblem (Sipp *et al.* 2010)

$$(\boldsymbol{A}(\boldsymbol{q}_0) + \lambda\boldsymbol{B})\hat{\boldsymbol{q}} = \boldsymbol{0}, \qquad (\hat{\boldsymbol{q}}^\dagger)^T(\boldsymbol{A}(\boldsymbol{q}_0) + \lambda\boldsymbol{B}) = \boldsymbol{0}\ , \tag{3.3}$$

with the direct eigenmode $\hat{\boldsymbol{q}}$ (representing the vortex shedding mode in the hydrofoil case) and the adjoint eigenmode $\hat{\boldsymbol{q}}^\dagger$. λ is the corresponding eigenvalue, of which the real part is the growth rate σ (which will serve as a cost function for the LSA-based optimization) and the imaginary part is the angular frequency ω. Here, $\boldsymbol{A}$ and $\boldsymbol{B}$ correspond to the discretized linear operators $\mathcal{A}$ and $\mathcal{B}$, respectively.

The sensitivity of the eigenvalue λ with respect to geometry-related parameters $\boldsymbol{a}$, in our case the control knots of the B-spline, is calculated as described in Knechtel *et al.* (2024). First, the sensitivities of $\boldsymbol{q}_0$ with respect to each parameter a_i are computed by solving the linear equation system

$$\boldsymbol{A}(\boldsymbol{q}_0)\ \frac{\mathrm{d}\boldsymbol{q}_0}{\mathrm{d}a_i} = -\left.\frac{\partial\boldsymbol{N}(\boldsymbol{q})}{\partial a_i}\right|_{\boldsymbol{q}=\boldsymbol{q}_0}\ . \tag{3.4}$$

With this, the calculation of the adjoint base flow is omitted. Then, with the operator sensitivity

$$\frac{\mathrm{d}\boldsymbol{A}(\boldsymbol{q}_0)}{\mathrm{d}a_i} = \frac{\partial\boldsymbol{A}(\boldsymbol{q}_0)}{\partial a_i} + \boldsymbol{A}_{BL}(\frac{\mathrm{d}\boldsymbol{q}_0}{\mathrm{d}a_i})\ , \qquad \boldsymbol{A}_{BL}(\frac{\mathrm{d}\boldsymbol{q}_0}{\mathrm{d}a_i})\ \hat{\boldsymbol{q}} := \begin{pmatrix} (\hat{\boldsymbol{u}}\cdot\nabla)\frac{\mathrm{d}\boldsymbol{u}_0}{\mathrm{d}a_i} + (\frac{\mathrm{d}\boldsymbol{u}_0}{\mathrm{d}a_i}\cdot\nabla)\hat{\boldsymbol{u}} \\ 0 \end{pmatrix} \tag{3.5}$$

the sensitivity of the eigenvalue can be written as (Luchini & Bottaro 2014)

$$\frac{\mathrm{d}\lambda}{\mathrm{d}a_i} = \frac{(\hat{\boldsymbol{q}}^\dagger)^T\left(\frac{\mathrm{d}\boldsymbol{A}(\boldsymbol{q}_0)}{\mathrm{d}a_i} + \lambda\frac{\partial\boldsymbol{B}}{\partial a_i}\right)\hat{\boldsymbol{q}}}{(\hat{\boldsymbol{q}}^\dagger)^T\boldsymbol{B}\hat{\boldsymbol{q}}}\ . \tag{3.6}$$

For RA, a linear input-output system is considered for each frequency ω, represented

by the resolvent operator $\mathcal{R}$, which acts on the forcing $\hat{\boldsymbol{f}}$ and generates the response $\hat{\boldsymbol{u}}$

$$\hat{\boldsymbol{u}} = \mathcal{R}\hat{\boldsymbol{f}}. \tag{3.7}$$

The forcing and response modes are normalized with their respective norms ($||\hat{\boldsymbol{u}}||^2 := \hat{\boldsymbol{u}}^H \boldsymbol{Q_u} \hat{\boldsymbol{u}}$ and $||\hat{\boldsymbol{f}}||^2 := \hat{\boldsymbol{f}}^H \boldsymbol{Q_f} \hat{\boldsymbol{f}}$), and μ^2 is the gain of the input-output system (which will serve as a cost function for the RA-based optimization). The discretized resolvent operator $\boldsymbol{R}$ is defined as

$$\boldsymbol{R} \coloneqq \boldsymbol{P}(\boldsymbol{A} + j\omega\boldsymbol{B})^{-1}\, \boldsymbol{W}\boldsymbol{P}^T, \qquad \boldsymbol{P} \coloneqq \begin{pmatrix} \boldsymbol{I} & 0 \end{pmatrix}, \tag{3.8}$$

where j is the imaginary unit, $\boldsymbol{A}$ and $\boldsymbol{B}$ are defined as above, $\boldsymbol{P}$ is a rectangular matrix containing the (discrete) identity matrix $\boldsymbol{I}$, and $\boldsymbol{W}$ represents the discretization weights of the whole system. We are interested in the optimal gains

$$\mu^2 = \sup_{\hat{\boldsymbol{f}}} \frac{||\hat{\boldsymbol{u}}||^2}{||\hat{\boldsymbol{f}}||^2} = \sup_{\hat{\boldsymbol{f}}} \frac{\hat{\boldsymbol{f}}^H \widetilde{\boldsymbol{R}} \hat{\boldsymbol{f}}}{||\hat{\boldsymbol{f}}||^2}, \qquad \widetilde{\boldsymbol{R}} \coloneqq \boldsymbol{R}^H \boldsymbol{Q_u} \boldsymbol{R} \tag{3.9}$$

and thus can formulate an eigenproblem (Sipp *et al.* 2010). Since $\widetilde{\boldsymbol{R}}$ and $\boldsymbol{Q_f}$ are Hermitian matrices, the direct and the adjoint modes are the same,

$$(\widetilde{\boldsymbol{R}} - \mu^2 \boldsymbol{Q_f})\hat{\boldsymbol{f}} \;=\; \boldsymbol{0} \;=\; \hat{\boldsymbol{f}}^H(\widetilde{\boldsymbol{R}} - \mu^2 \boldsymbol{Q_f}). \tag{3.10}$$

As the flow can be unstable in some parts of the parameter space, we perform a discounted RA that effectively shifts the complex plane (Rolandi *et al.* 2024).

Analogous to the sensitivity from Eq. (3.6), here the sensitivity can be written as

$$\frac{\mathrm{d}\mu^2}{\mathrm{d}a_i} = \frac{\hat{\boldsymbol{f}}^H \left(\frac{\mathrm{d}\widetilde{\boldsymbol{R}}}{\mathrm{d}a_i} - \mu^2 \frac{\partial \boldsymbol{Q_f}}{\partial a_i} \right) \hat{\boldsymbol{f}}}{\hat{\boldsymbol{f}}^H \boldsymbol{Q_f} \hat{\boldsymbol{f}}}. \tag{3.11}$$

The operator sensitivity is computed similar to Poulain *et al.* (2024) with

$$\hat{\boldsymbol{f}}^H \frac{\mathrm{d}\widetilde{\boldsymbol{R}}}{\mathrm{d}a_i} \hat{\boldsymbol{f}} = \mu^2 \hat{\boldsymbol{u}}^H \frac{\partial \boldsymbol{Q_u}}{\partial a_i} \hat{\boldsymbol{u}} + 2\mu \Re \left\{ \hat{\boldsymbol{f}}^H \left(\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}a_i} \right)^H \boldsymbol{Q_u} \hat{\boldsymbol{u}} \right\} \tag{3.12}$$

and

$$\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}a_i} = \boldsymbol{P}(\boldsymbol{A} + j\omega\boldsymbol{B})^{-1} \frac{\partial \boldsymbol{W}}{\partial a_i} \boldsymbol{P}^T - \boldsymbol{P}(\boldsymbol{A} + j\omega\boldsymbol{B})^{-1} \left(\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}a_i} + \omega \frac{\partial \boldsymbol{B}}{\partial a_i} \right) (\boldsymbol{A} + j\omega\boldsymbol{B})^{-1} \boldsymbol{W}\boldsymbol{P}^T. \tag{3.13}$$

The base flow computations and the linearized methods are conducted using the finite-element-based FELiCS software (Kaiser *et al.* 2023). The analytically derived linear operators are built explicitly as finite-element-weighted matrices, which are then used to solve the linear equation systems described above. The partial derivatives of vectors and matrices, which appear in the sensitivities of Eqs. (3.6) and (3.11), are calculated with a finite difference of first order. The unsteady numerical simulations to obtain the mean flow fields are conducted with the finite-volume code OpenFOAM.

4. Bayesian optimization with derivatives

Consider an unknown cost function $f : \mathbb{R}^d \to \mathbb{R}$ (here the LSA growth rate σ or the resolvent gain μ^2 as a function of shape parameters $\boldsymbol{a} \in \mathbb{R}^d$), which we aim to approximate to find the minimum. In Bayesian optimization with derivatives (d-BO) this is carried

out in two steps. First, a surrogate model of f is built from available data consisting of n inputs, function value observations and gradient observations $\mathcal{D} = \{\mathfrak{a}_i, \mathfrak{f}_i, \nabla\mathfrak{f}_i\}_{i=1}^n$. This is done using Gaussian process regression (GPR), which yields a probabilistic prediction of function values across the optimization domain Ω. Second, the prediction is used to identify regions in Ω in which optima are likely to be found. To do so, a sampling policy in the form of an acquisition function is used to guide sampling within the domain. The sampled configurations are subsequently added to the training data set and the surrogate model is recalibrated. By iteratively repeating these steps, optima on Ω are identified. Both GPR and the acquisition-function-based sampling are described in the following.

GPR begins with a prior belief about the latent function $f(\boldsymbol{a})$ that describes the data. This prior is represented by a Gaussian process (GP), which is the continuous equivalent to a multivariate Gaussian distribution. A GP is an infinite collection of random variables for inputs $\boldsymbol{a}$ (infinite due to the continuous nature of Ω). The GP assumption about f is denoted $f(\boldsymbol{a}) \sim \mathcal{GP}(m(\boldsymbol{a}), k(\boldsymbol{a}, \boldsymbol{a}^*))$ and is characterized by a mean function $m(\boldsymbol{a})$ that determines the expected function value of f at $\boldsymbol{a}$ and a covariance function $k(\boldsymbol{a}, \boldsymbol{a}^*)$ that characterizes the probability of deviations from the mean,

$$m(\boldsymbol{a}) = \mathbb{E}[f(\boldsymbol{a})], \quad k(\boldsymbol{a}, \boldsymbol{a}^*) = \mathbb{E}[(f(\boldsymbol{a}) - m(\boldsymbol{a}))\,(f(\boldsymbol{a}^*) - m(\boldsymbol{a}^*))]. \tag{4.1}$$

Here, $\mathbb{E}$ marks the expectation operator and $\boldsymbol{a} = [a_1, a_2, \ldots, a_d]^T, \boldsymbol{a}^* = [a_1^*, a_2^*, \ldots, a_d^*]^T$ are variables in Ω. Throughout this work, a constant prior mean function is used. Furthermore, a kernel function of squared exponential type is chosen, for which the covariance function reads

$$k(\boldsymbol{a}, \boldsymbol{a}^*) = \exp\left(-\sum_{i=1}^{d} \frac{(a_i - a_i^*)^2}{2\theta_i^2}\right), \tag{4.2}$$

where θ_1 through θ_d are hyperparameters that are tuned to available data by maximization of the log marginal likelihood (Rasmussen & Williams 2006).

The prior assumption about f implies a covariance between function values and gradients, such that the function and its gradient follow a multi-output GP of the following form (Rasmussen & Williams 2006)

$$\begin{bmatrix} f \\ \nabla f \end{bmatrix} \sim \mathcal{GP}(\boldsymbol{m}_\nabla, \boldsymbol{K}_\nabla), \quad \boldsymbol{m}_\nabla = \begin{bmatrix} m(\boldsymbol{a}) \\ \nabla m(\boldsymbol{a}) \end{bmatrix}, \boldsymbol{K}_\nabla = \begin{bmatrix} k(\boldsymbol{a}, \boldsymbol{a}^*) & \boldsymbol{j}(\boldsymbol{a}, \boldsymbol{a}^*)^T \\ \boldsymbol{j}(\boldsymbol{a}^*, \boldsymbol{a}) & \boldsymbol{H}(\boldsymbol{a}, \boldsymbol{a}^*) \end{bmatrix}. \tag{4.3}$$

Here, $\boldsymbol{j}(\boldsymbol{a}, \boldsymbol{a}^*)^T = [\partial k(\boldsymbol{a}, \boldsymbol{a}^*)/\partial a_1^*, \partial k(\boldsymbol{a}, \boldsymbol{a}^*)/\partial a_2^*, \ldots]$ is a vector of size $d \times 1$ and $\boldsymbol{H}(\boldsymbol{a}, \boldsymbol{a}^*)$ is a symmetric $d \times d$ matrix with entries $\boldsymbol{H}_{i,j} = \partial^2 k/\partial a_i^* \partial a_j$.

When function observations are available, the prior assumption can be updated to a posterior distribution, which is also a GP. Bayes's rule on conditional probability is used to derive the posterior, which describes the conditional probability of observing function values and gradients, given observations $\mathcal{D} = \{\mathfrak{a}_i, \mathfrak{f}_i, \nabla\mathfrak{f}_i\}_{i=1}^n$. For closed-form equations, see, e.g., Wu *et al.* (2017). The posterior mean represents the most likely function and gradient value at $\boldsymbol{a}$, given the data, and serves as a prediction from GPR. The posterior covariance function in turn quantifies prediction uncertainty. It results from the prior covariance and an update term that depends on the proximity to the training samples.

Based on this posterior distribution, samples to be tested are selected from Ω sequentially. Therefore, an acquisition function $\alpha(\boldsymbol{a})$ is used as a sampling criterion. Within this work the expected improvement (EI) acquisition function is used (Mockus 1974), which

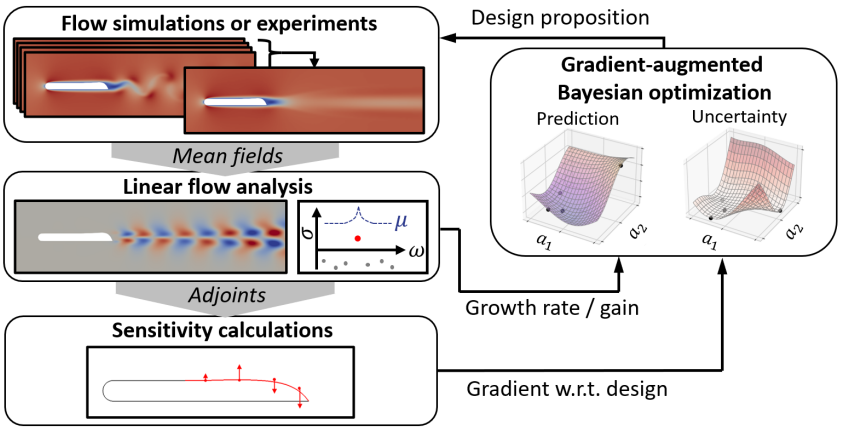

FIGURE 2. Schematic of the BOA optimization workflow.

aims to minimize the objective value throughout the observed samples,

$$\mathfrak{a}_{n+1} = \underset{\boldsymbol{a}\in\Omega}{\operatorname{argmax}}\ \alpha_{EI}(\boldsymbol{a}), \qquad \alpha_{EI}(\boldsymbol{a}) = \mathbb{E}(\max(\mathfrak{f}^* - f(\boldsymbol{a}), 0)), \tag{4.4}$$

where $\mathfrak{f}^*$ is the best function value tested so far.

The proposed methodology for d-BO is set up using the Python package BoTorch (Balandat *et al.* 2020). For the implemented GP with joint function and gradient output, the provided modules ConstantMeanGrad and RBFKernelGrad for mean and kernel functions and their respective gradients are combined into a multi-output GP model class. Input parameters are scaled to a unit hypercube, and the output data set is normalized to zero mean and unit variance in each iteration, in order to facilitate model fitting.

5. Optimization procedure

Figure 2 shows the workflow of the BOA optimizer. The cost function to be minimized is selected to be either the largest eigenvalue growth rate or the leading resolvent gain integrated over a frequency band, reading

$$\underset{\boldsymbol{a}}{\operatorname{argmin}} f(\boldsymbol{a}), \qquad f = \max_{\sigma\in\{\Re(\lambda_i)\}} \sigma(\boldsymbol{a}) \quad \text{or} \quad \int_\omega \mu_1^2(\boldsymbol{a})\ \mathrm{d}\omega. \tag{5.1}$$

To evaluate this function for a candidate design $\mathfrak{a}_{n+1}$, selected in iteration $n+1$, a corresponding base flow or mean flow is first determined. Next, the LSA or RA is conducted (Eqs. (3.3) and (3.9)) to obtain the objective value, maximum growth rate or integral leading gain, respectively. In addition to evaluating the function $f(\boldsymbol{a})$ at each candidate point, we leverage additional system knowledge by computing the gradient with regard to the design parameters (Eqs. (3.6) and (3.11)) and computing the sensitivity of the eigenvalue or the integral gain to shape parameters, $\nabla f(\boldsymbol{a})$.

Within the BOA method, the obtained data point consisting of design $\mathfrak{a}_{n+1}$, growth rate/RA gain $\mathfrak{f}_{n+1} = f(\boldsymbol{a} = \mathfrak{a}_{n+1})$ and gradient value $\nabla\mathfrak{f}_{n+1} = \nabla f(\boldsymbol{a} = \mathfrak{a}_{n+1})$ is then added to the training data set, $\mathcal{D} \leftarrow \mathcal{D}\cup\{\mathfrak{a}_{n+1}, \mathfrak{f}_{n+1}, \nabla\mathfrak{f}_{n+1}\}$, on the basis of which a new GP model is trained. The model prediction is used in an acquisition function to select the subsequent design to be tested, Eq. (4.4). The procedure is repeated until a query budget is exhausted, after which the best tested sample is then returned.

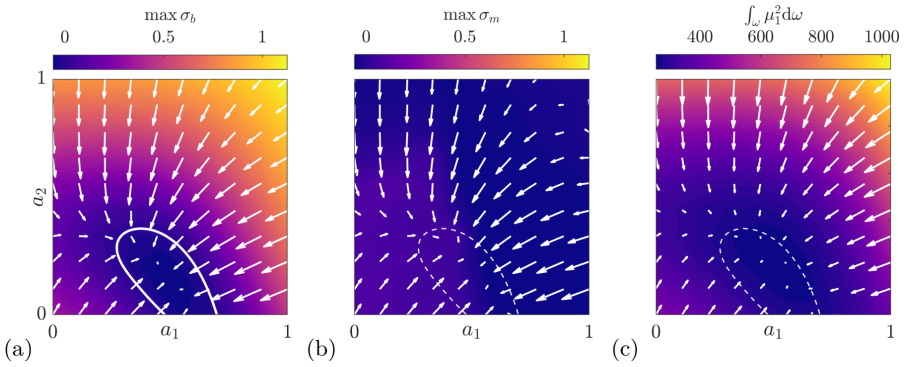

FIGURE 3. Maximum LSA growth rate in the parameter space using (a) base flow field, $\sigma_b(a_1, a_2)$, and (b) mean flow field, $\sigma_m(a_1, a_2)$; (c) RA integral leading gain $\int_\omega \mu_1^2(a_1, a_2)\mathrm{d}\omega$ in the parameter space using base flow field. Negative gradient with regard to shape is superposed as white arrows. Region enclosed with white contour line indicates the stable basin of the base flow LSA.

6. Results

6.1. *Growth rate and sensitivity maps obtained by base and mean flow LSA*

As a reference solution for the two-parameter case, the growth rate of the wake mode determined from LSA and its sensitivity are computed for the entire parameter space by brute-force evaluation on a grid. The results from the analyses of the base and mean flow fields are shown in Figure 3(a,b), respectively. The plots show growth rate maps, superposed with arrows that indicate the growth rate sensitivity, pointing in the direction of the negative gradient.

From the base flow growth rate, we observe a stable basin characterized by a single minimum. It corresponds to a moderate thickness of the hydrofoil trailing edge and the presence of a concave bump on the suction side of the airfoil (see inset in Figure 4(a)). This bump is particularly notable, as it parallels similar geometric features predicted in shape-optimized cylinder flows (Brewster & Juniper 2020). Conversely, the most unstable configuration aligns with the most blunt and thickest trailing edge (see Figure 1(c)).

The growth rate determined from the mean-field LSA (Figure 3(b)) is approximately zero throughout the parameter space. This is a classic result for mean-field-based LSA, where the nonlinearly saturated wake mode is found as a global mode at zero growth rate (Sipp & Lebedev 2007). The sensitivity gradients derived from mean flow LSA reflect the change in the linear growth mechanism but do not account for changes in nonlinear saturation. Consequently, they do not align with the actual growth rate distribution. Interestingly, they closely resemble the sensitivity gradients obtained from the base flow instead. Thus, the sensitivity gradients remain valuable in the mean flow LSA, while the growth rate itself is not a useful cost function for optimization. To still proceed with the optimization, we approximate the base flow growth rate σ_b from the nonlinear vortex shedding amplitude using the following relation

$$\sigma_b \approx \eta |A|^2, \tag{6.1}$$

which is determined from the Stuart–Landau equation (Sipp & Lebedev 2007) for a statistically stationary saturated state. Here, A represents the global saturation amplitude determined from a single probe in the nonlinear simulations. The nonlinear damping coefficient η is calibrated to a constant value for the entire parameter map using the true base flow growth rate determined near the bifurcation point. The probing in the nonlin-

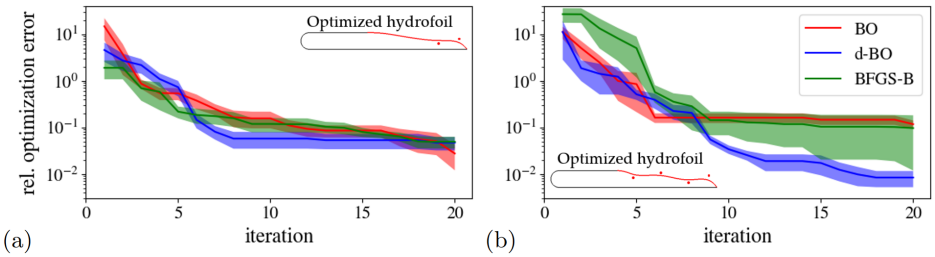

FIGURE 4. Optimization curves using mean flow fields for (a) two parameters and (b) four parameters, comparing global optimization without gradient information (BO), global optimization with gradient information (d-BO) and local optimization with gradient information (BFGS-B), averaged over five runs. The respective optimized hydrofoils are shown as insets.

ear simulations at only one position to obtain A introduces uncertainty in the growth rate predictions. Further uncertainty is associated with Eq. (6.1) of the Stuart–Landau model, which is valid only in close proximity to the bifurcation point. This, however, becomes uncritical with a progressing number of optimization steps, as the distance to the bifurcation point reduces. All uncertainties are believed to be effectively absorbed by the BO and d-BO frameworks with their inherent ability to incorporate uncertainties in the surrogate modeling, ensuring a robust optimization process.

6.2. *Shape optimization using base and mean flow LSA*

In the following, shape optimization of the hydrofoil parameterized with both two and four parameters is conducted. Three different methods are compared: Bayesian optimization without gradient information (BO), Bayesian optimization with gradient information (d-BO), and local gradient-based optimization, for which a bounded version of BFGS (BFGS-B) is used. Their performance is evaluated by comparing the optimization curves in Figure 4. It shows the relative error between the best configuration identified by each method up to a certain iteration and the global optimum within the parameter space. The latter is identified by considering all brute-force and optimization runs. Since initial sampling is random for all three methods, five independent runs are performed for each method to ensure statistically stationary results. The plots show the average and plus/minus half a standard deviation as lines and shaded areas, respectively. For brevity, only the mean flow results are presented here since very similar observations are made for the base flow case.

For the two-parameter case shown in Figure 4(a), optimization methods that incorporate gradient information (BFGS-B and d-BO) outperform the global optimization approach that lacks gradient information (BO). The local optimization method (BFGS-B) exhibits a performance comparable to that of the global optimization approach with gradient information (d-BO).

However, when the dimension of the parameter space is increased to four, the advantages of the d-BO method become more apparent (Figure 4(b)). The d-BO method clearly outperforms both BO and BFGS-B, demonstrating faster convergence to the true minimum. In contrast, the BFGS-B method appears to become trapped in local minima, failing to identify the global optimum in most runs, while the d-BO method avoids this pitfall due to its global approach of exploration and exploitation. For the present case, these results clearly illustrate the superiority of d-BO over BFGS-B, attributed to its global approach, and also underscore the advantage of d-BO over BO, stemming from the incorporation of additional gradient information.

6.3. *Parameter space of gain and sensitivity obtained by base flow RA*

In order to generate a reference solution for the resolvent gain and its sensitivity in the two-parameter case, a brute-force evaluation is again employed across the entire parameter space. Figure 3(c) shows the integral leading gain and its sensitivity using base flow fields. As intended, the negative RA sensitivity points in the direction of steepest descent of the RA gain. Compared to the LSA results in Figure 3(a), the regions with the lowest gain correspond to the most stable configurations. This observation is consistent with the nature of the wake dynamics, which is driven by an isolated eigenmode with a relatively weak nonorthogonality to other modes. As a result, the resolvent gain is inversely proportional to the distance of the eigenmode to the stability limit (Symon *et al.* 2018).

The strong correlation between the distributions of gain and eigenvalue growth rate associated with the wake mode reinforces the RA's applicability to oscillator-type flows. Furthermore, the agreement between the RA gain distribution and its sensitivity permits its use for shape optimization in the BOA framework to flows with multiple modes and flows with broadband coherent structures, in particular under fully turbulent conditions.

7. Conclusion and outlook

In this work, we demonstrate the integration of linear flow analysis methods (in the form of linear stability analysis (LSA) and resolvent analysis (RA)) and their respective adjoint-based gradients with derivative-augmented Bayesian optimization to enable efficient shape optimization of flow dynamics. LSA and RA provide physically meaningful objective functions that are cost-effective to evaluate. The incorporation of gradient information significantly enhances the otherwise gradient-agnostic Bayesian framework. Particularly in large parameter spaces the speedup becomes increasingly pronounced.

Shape optimization was performed with respect to the growth rate of the vortex shedding mode occurring in the wake of a Blake-type hydrofoil, employing both base and mean flow LSA. Furthermore, the sensitivity of the resolvent gain was derived and computed, showing trends consistent with the LSA results. This finding opens up the possibility for future studies addressing fully turbulent flows at higher Reynolds numbers to target the amplification of oscillations specifically (via the resolvent gain) instead of the stability (via the growth rate). It will also enable us to conduct shape optimization in amplifier-type flows that are globally stable. Optimizing flow dynamics based on the mean field allows the use of cost-efficient Reynolds-averaged Navier–Stokes simulations. In future work, these may be complemented with large-eddy simulations using multi-fidelity surrogate methods or they could be improved through data assimilation. Open research questions also arise from the implementation of additional meaningful constraints and the expansion to more complex parameter spaces with multiple local minima, potentially requiring the evaluation of (many) more designs to fully capture the optimization landscape. Overall, the integration of LSA and RA, adjoint-based methods, and Bayesian optimization offers a powerful toolkit for the efficient and precise optimization of aerodynamic shapes in unsteady flow regimes.

Acknowledgments

The funding from the Stanford CTR stipend, the Technische Universität Berlin and the Deutsche Forschungsgemeinschaft within project numbers 429772199 and 523881008

is gratefully acknowledged. The authors express their gratitude to Stanford University for generously providing the high-performance computing resources. Furthermore, the authors thank Salvador Gomez, Matthew Juniper, Roney Thompson, Benjamin Herrmann and Josef Winter for the fruitful discussions during the summer program, and Peter Schmid for the careful review of the manuscript.

REFERENCES

Balandat, M., Karrer, B., Jiang, D. R., Daulton, S., Letham, B., Wilson, A. G. & Bakshy, E. 2020 BoTorch: A framework for efficient Monte-Carlo Bayesian optimization. In *Adv. Neur. In.* **80**, 21524–21538.

Brewster, J. & Juniper, M. P. 2020 Shape sensitivity of eigenvalues in hydrodynamic stability, with physical interpretation for the flow around a cylinder. *Eur. J. Mech. B-Fluid* **80**, 80–91.

Kaiser, T. L., Demange, S., Müller, J. S., Knechtel, S. & Oberleithner, K. 2023 Felics: a versatile linearized solver addressing dynamics in multi-physics flows. In *AIAA Aviation 2023 Forum*, p. 3434.

Knechtel, S., Kaiser, T., Orchini, A. & Oberleithner, K. 2024 Arbitrary-order sensitivities of the incompressible base flow and its eigenproblem. *J. Fluid Mech.* **985**, A32.

Luchini, P. & Bottaro, A. 2014 An introduction to adjoint problems. *Annu. Rev. Fluid Mech.* **46**, 493–517.

Marsden, A. L., Wang, M., Dennis, J. E. & Moin, P. 2004 Optimal aeroacoustic shape design using the surrogate management framework. *Optim. Eng.* **5**, 235–262.

Mockus, J. 1974 On Bayesian methods for seeking the extremum. In *Proceedings of the IFIP Technical Conference*, pp. 400–404.

Poulain, A., Content, C., Rigas, G., Garnier, E. & Sipp, D. 2024 Adjoint-based linear sensitivity of a supersonic boundary layer to steady wall blowing-suction/heating-cooling. *J. Fluid Mech.* **978**, A16.

Rasmussen, C. E. & Williams, C. K. I. 2006 *Gaussian Processes for Machine Learning*. MIT Press.

Rolandi, L. V., Ribeiro, J. H. M., Yeh, C.-A. & Taira, K. 2024 An invitation to resolvent analysis. *Theor. Comp. Fluid Dyn.* **38**, 603–639.

Sipp, D. & Lebedev, A. 2007 Global stability of base and mean flows: a general approach and its applications to cylinder and open cavity flows. *J. Fluid Mech.* **593**, 333–358.

Sipp, D., Marquet, O., Meliga, P. & Barbagallo, A. 2010 Dynamics and control of global instabilities in open-flows: a linearized approach. *Appl. Mech. Rev.* **63**, 030801.

Symon, S., Rosenberg, K., Dawson, S. T. & McKeon, B. J. 2018 Non-normality and classification of amplification mechanisms in stability and resolvent analysis. *Phys. Rev. Fluids* **3**, 053902.

Taira, K., Brunton, S. L., Dawson, S. T., Rowley, C. W., Colonius, T., McKeon, B. J., Schmidt, O. T., Gordeyev, S., Theofilis, V. & Ukeiley, L. S. 2017 Modal analysis of fluid flows: an overview. *AIAA J.* **55**, 4013–4041.

Wu, J., Poloczek, M., Wilson, A. G. & Frazier, P. I. 2017 Bayesian optimization with gradients. *Adv. Neur. In.* **30**, 5268–5279.