From CAD to Eigenshapes for Surrogate-based Optimization

David Gaudrie, Rodolphe Le Riche, Victor Picheny, Benoit Enaux, Vincent Herbert

The World Congress of Structural and Multidisciplinary Optimization, May 20-24, 2019, Beijing, China

From CAD to Eigenshapes for Surrogate-based Optimization

David Gaudrie[1,2*], **Rodolphe Le Riche**[2], **Victor Picheny**[3], **Benoît Enaux**[1], **Vincent Herbert**[1]

[1] *Groupe PSA, Vélizy-Villacoublay, France*
[*] *Corresponding author: david.gaudrie@mpsa.com*
[2] *CNRS LIMOS, École des Mines de Saint-Étienne, France*
[3] *Prowler.io, Cambridge, United Kingdom*

Abstract Parametric shape optimization aims at minimizing an objective function $f(\mathbf{x})$ where $\mathbf{x}$ are CAD parameters of a shape. This task is difficult when $f(\cdot)$ is the output of an expensive-to-evaluate numerical simulator and the number of CAD parameters is large.

Most often, the set of all considered CAD shapes reside in a manifold of lower effective dimension in which it is preferable to build the surrogate model and perform the optimization. In this work, we uncover the manifold through a high-dimensional shape mapping and build a new coordinate system that we call the eigenshape space. The surrogate model is learned in the space of eigenshapes: a regularized likelihood maximization provides the most relevant dimensions for the output. The final surrogate model is detailed (anisotropic) with respect to the most sensitive eigenshapes and rough (isotropic) in the remaining dimensions. Last, the optimization is carried out with a focus on the critical variables, the remaining ones being coarsely optimized through an embedding strategy. At low budgets, the methodology leads to a more accurate model and a faster optimization than the classical approach of directly working with the CAD parameters.

Keywords: Dimension Reduction, Principal Component Analysis, Parametric Shape Optimization, Gaussian Processes, Bayesian Optimization

1 Introduction

The most frequent approach to shape optimization is to describe the shape by a vector of d Computer Aided Design (CAD) parameters, $\mathbf{x} \in X \subset \mathbb{R}^d$ and to search for the parameters that minimize the objective function, $\mathbf{x}^* = \arg\min_{\mathbf{x}\in X} f(\mathbf{x})$. In the CAD modeling process, the set of all possible shapes has been reduced to the space of parameterized shapes, $\mathbf{\Omega} \coloneqq \{\Omega_{\mathbf{x}}, \mathbf{x} \in X\}$.

It is common for d to be large, $d \gtrsim 50$. Optimization in such a high-dimensional design space is difficult, especially when $f(\cdot)$ is the output of a high fidelity numerical simulator that can only be run a restricted number of times. In computational fluid dynamics for example, simulations easily take 12 to 24 hours and evaluation budgets range between 100 and 200 calls. Surrogate-based approaches [1] have proven their effectiveness to tackle optimization problems in a few calls to $f(\cdot)$. They rely on a surrogate model (or metamodel, e.g., Gaussian Processes [2]) built upon n past observations of $y_i = f(\mathbf{x}^{(i)})$. The metamodel is utilized by an acquisition function such as the Expected Improvement [1] to decide what will be the next parametric design $\mathbf{x}^{(n+1)}$ evaluated. However, such techniques suffer from the curse of dimensionality when d is large. The budget is also typically too narrow to perform sensitivity analysis and select variables prior to optimizing. A further issue is that the CAD parameters $\mathbf{x}$ commonly have heterogeneous impacts on the shapes $\Omega_{\mathbf{x}}$: many of them are intended to refine the shape locally whereas others have a global influence so that shapes of practical interest involve interactions between all the parameters.

Most often, the set of all CAD generated shapes, $\mathbf{\Omega}$, can be approximated in a δ-dimensional manifold, $\delta \ll d$. In [3,4] this manifold is accessed through an auxiliary description of the shape, $\phi(\Omega)$, ϕ being either its characteristic function or the signed distance to its contour. The authors aim at minimizing an objective function using diffuse approximation and gradient-based techniques, while staying on the manifold of admissible shapes. Another way to handle shapes is the Point Distribution Model, in which the contour of the shape is discretized [5,6]. In [7], a kriging surrogate is built in the space of the first Partial Least Squares axes. The consideration of modes of discretized shapes is further pursued in [8] where bounds on the modes coefficients are enforced and the gradient of the metamodel used for optimization purposes.

Following the same route, in Section 2, we retrieve a shape manifold with dimension $\delta \ll d$ by principal component analysis of shapes described in an ad hoc manner. Section 3 is devoted to the construction of a kriging surrogate model in reduced dimension (but the remaining dimensions are not completely omitted). In Section 4, we employ the metamodel to perform global optimization [1] via the maximization of the Expected Improvement. A reduction of the space dimension is achieved through a random embedding technique [9] and a pre-image problem is solved to keep the correspondence between the eigenshapes and the CAD parameters.

2 From CAD description to shape eigenbasis

CAD parameters are usually set up by engineers to automate shape generation. These parameters may be Bézier or Spline control points which locally readjust the shape. Other CAD parameters, such as the overall width or the length of a component, have a more global impact on the shape. While these parameters are intuitive to a designer, they are not chosen to achieve any specific mathematical property and in particular do not let themselves interpret to reduce dimensionality.

In order to define a better behaved description of the shapes that will help in reducing dimensionality, we exploit the fact that the time to generate a shape $\Omega_{\mathbf{x}}$ is negligible in comparison with the evaluation time of $f(\mathbf{x})$.

In the spirit of kernel methods [10,11], we analyze the designs $\mathbf{x}$ in a high-dimensional feature space $\Phi \subset \mathbb{R}^D$, $D \gg d$ (potentially infinite dimensional) that is defined via a mapping $\phi(\mathbf{x})$, $\phi : X \to \Phi$. With an appropriate ϕ, it is possible to distinguish a lower dimensional manifold embedded in Φ. As we deal with shapes, natural candidates for ϕ are shape representations. In the literature, shapes have been described by the characteristic function [3] (a grid of 0's or 1's whether we are inside or outside the shape, cast as a vector), the signed distance to its contour $\partial\Omega_{\mathbf{x}}$ [4] (again on a spatial grid which is transformed into a vector), or the discretization of $\partial\Omega_{\mathbf{x}}$ [5] (the coordinates of nodes on the contour in a given order and transformed into a vector). We map a large number (N) of uniformly sampled designs $\mathbf{x}^{(i)} \in X$ to Φ and build the matrix $\mathbf{\Phi} \in \mathbb{R}^{N \times D}$ which contains the centered $\phi(\mathbf{x}^{(i)}) \in \mathbb{R}^D$ in rows. Next, we perform a Principal Component Analysis (PCA) on $\mathbf{\Phi}$: the eigenvectors of $\mathbf{\Phi}^\top\mathbf{\Phi}$, written $\mathbf{v}^j$, form an ordered orthonormal basis of Φ with decreasing importance as measured by the PCA's eigenvalues λ_j, $j = 1, \ldots, D$. The $\mathbf{v}^j$'s are orthonormal directions in Φ that explain the most the dispersion of the high-dimensional representations of the shapes, $\phi(\mathbf{x}^{(i)})$. Any design $\mathbf{x}$ can now be expressed in the eigenbasis[1] $\mathbf{V} \coloneqq (\mathbf{v}^1, \ldots, \mathbf{v}^D)$ since $\phi(\mathbf{x}) = \overline{\boldsymbol{\phi}} + \sum_{j=1}^{D} \alpha_j \mathbf{v}^j$, where $(\alpha_1, \ldots, \alpha_D)^\top =: \boldsymbol{\alpha} = \mathbf{V}^\top(\phi(\mathbf{x}) - \overline{\boldsymbol{\phi}})$ are the coordinates in the $\mathbf{V}$ basis (principal components). α_j is a deviation from the mean shape, $\overline{\boldsymbol{\phi}} \in \Phi$, in the direction of the *eigenshape* $\mathbf{v}^j$.

In experiments that are partially reported here due to space limitations, we have generated shapes of known low intrinsic dimension: in the example of Fig. 1, the shapes are a set of circular holes of varying centers and radii, therefore described by 1, 2 or 3 parameters. PCAs were then carried out on the $\mathbf{\Phi}$'s associated to the three mappings (characteristic function, signed contour distance and contour discretization). We have observed the drop in eigenvalues λ_j and plotted the $\boldsymbol{\alpha}$'s manifolds.

The discretization of $\partial\Omega_{\mathbf{x}}$ was the only mapping ϕ for which the intrinsic shape dimension was equal to the number of non-zero eigenvalues, i.e., the shape intrinsic dimension was recovered. With this ϕ, the manifold of $\boldsymbol{\alpha}^{(i)}$'s was better unfolded, which is advantageous for metamodeling and the eigenshapes $\mathbf{v}^j$ were physically meaningful. For these reasons in the following, we will only consider the $\boldsymbol{\alpha}$'s obtained using the contour discretization as ϕ mapping.

The three first eigenshapes $\mathbf{v}^1$, $\mathbf{v}^2$ and $\mathbf{v}^3$ as well as the mean shape $\overline{\boldsymbol{\phi}}$ of the NACA 22 airfoil can be seen in Figure 2. The NACA 22 benchmark is a data set made of airfoils described by $d = 22$ CAD parameters with the associated lift and drag coefficients as outputs [12]. Notice that even though the NACA 22 shapes contain bumps, the three first eigenvectors do not include them. Bumps, which correspond to local refinements of the shapes, appear from the $4th$ eigenshape on.

3 GP model for reduced eigenspace

Building a surrogate-model in the space of principal components has already been investigated to construct reduced order models [13]. In most applications, the dimension reduction is carried out in the output space, which has large dimension when it corresponds to values on a finite element mesh. The response is approximated by a linear combination of a small number of modes, and the metamodel is a function of the modes coefficients.

Here, we aim at reducing the dimension of the input space by building a surrogate of the design principal components, $\boldsymbol{\alpha}$. A first idea to reduce the dimension of the problem is to conserve the δ first eigenvectors $\mathbf{v}^j$ according to some reconstruction quality criterion measured by the eigenvalues. Given a threshold T (e.g., 0.99), only the first δ modes such that $\frac{\sum_{j=1}^{\delta} \lambda_j}{\sum_{j=1}^{D} \lambda_j} > T$ are retained for the surrogate-based optimization process because they contribute for $100 \times T\%$ of the variance in Φ.

But such an approach relies only on considerations about the shape geometry. The output y is not taken into account for the dimension reduction even though some $\mathbf{v}^j$, $j \in \{1, \ldots, \delta\}$ may influence y or not. Two

[1] In a slight abuse of notation, we use $\mathbf{V}$ for both the basis and the matrix made of the $\mathbf{v}$'s.

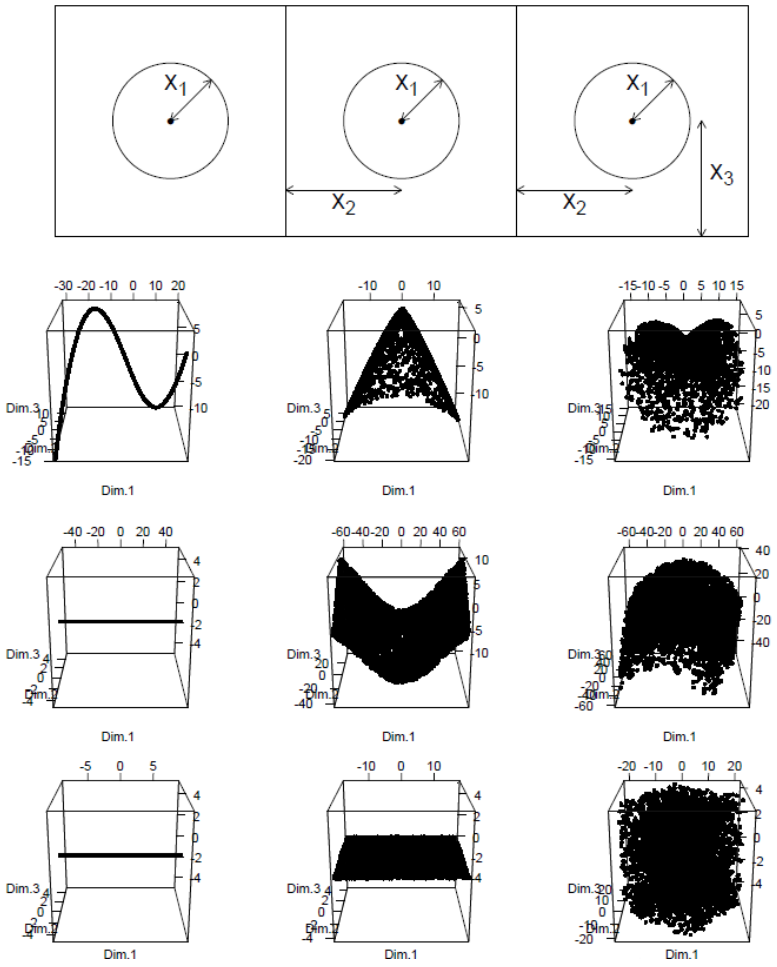

Figure 1: Three first eigencomponents of the $\boldsymbol{\alpha}^{(i)}$'s for three parametric test cases (columns) with low effective dimension equal to 1 (left), 2 (center) and 3 (right). The rows correspond to different ϕ's which are the characteristic function (top), the signed distance to the contour (middle) and the discretization of the contour (bottom).

shapes which differ in the α_j components with $j \leq \delta$ may behave similarly in terms of output y, so that further dimension reduction is possible. Vice versa, eigencomponents that have a small geometrical effect and were neglected may be reintroduced because they matter for y.

A second idea is thus to select the eigencomponents that impact y the most. This is done by maximizing the penalized log-likelihood [14] of a Gaussian process (GP) in the high (D) dimensional space of $\boldsymbol{\alpha}$'s,

$$\max_{\vartheta} pl_\lambda(\boldsymbol{\alpha}^{(1:n)}, \mathbf{y}_{1:n}; \vartheta) \quad \text{where} \quad pl_\lambda(\boldsymbol{\alpha}^{(1:n)}, \mathbf{y}_{1:n}; \vartheta) := l(\boldsymbol{\alpha}^{(1:n)}, \mathbf{y}_{1:n}; \vartheta) - \lambda \|\boldsymbol{\theta}^{-1}\|_1 \tag{1}$$

The ϑ are the GP's hyper-parameters which include the length-scales of the GP, θ_j. l stands for the log-likelihood of the GP and $\boldsymbol{\theta}^{-1} := (1/\theta_1, \ldots, 1/\theta_D)^\top$ is the vector containing the inverse length-scales of the GP. It is indeed known [15] that if $\theta_j \to +\infty$, the direction $\mathbf{v}^j$ has no influence on y. The L^1 penalty term applied to the θ_j's performs variable selection: this Lasso-like procedure promotes zeros in the vector of inverse length-scales, hence sets many θ_j's to $+\infty$. In the end, the directions with small enough θ_j, denoted $\boldsymbol{\alpha}^a \in \mathbb{R}^\delta$, are declared to be active with regard to y, and have to be emphasized during metamodeling and optimization: if $\theta_j \leq 10 \times \min_{i=1,\ldots,D} \theta_i$, then α_j is a component of the active $\boldsymbol{\alpha}^a$. Even if the maximization of pl_λ is carried out in a D-dimensional space, the problem is tractable since the gradients of pl_λ are analytically known, and because the L^1 penalty convexifies the problem. Most local optima to this problem solely differ in θ_j's that are already too large to be relevant and efficient local optimizers consistently yield the same set of active variables $\boldsymbol{\alpha}^a$. In our implementation, we have modified the likelihood maximization of the kergp package [16] to include the penalization term. After a dimensional anaysis of pl_λ, we have chosen to take $\lambda = \frac{n}{D}$ to balance both terms.

Other techniques such as cross-validation or the use of different λ's for obtaining a pre-defined number of active components can also be considered. On the NACA 22 benchmark with few observations of $f(\cdot)$ ($n = 15$ here), Figure 2 shows that only few active components among the $D = 600$ are selected by the penalized maximum likelihood procedure. The three first principal axes, $\mathbf{v}^1$, $\mathbf{v}^2$ and $\mathbf{v}^3$ are retained when considering the drag (top). Indeed, these are the eigenshapes that globally impact the shape the most and change its drag. When the output y is the lift (bottom), only the second principal axis is selected. This eigenshape modifies the camber of the shape, which is known to highly impact the lift. The other eigenvectors are detected to be less critical for y's variations. When n grows, more eigenshapes get selected because they also slightly impact the output. For instance when $n = 50$, some eigenshapes that contain bumps (the 4th, the 5th, the 8th, etc.) are selected for modeling the lift. They also contribute to changing the camber of the airfoil, hence its lift.

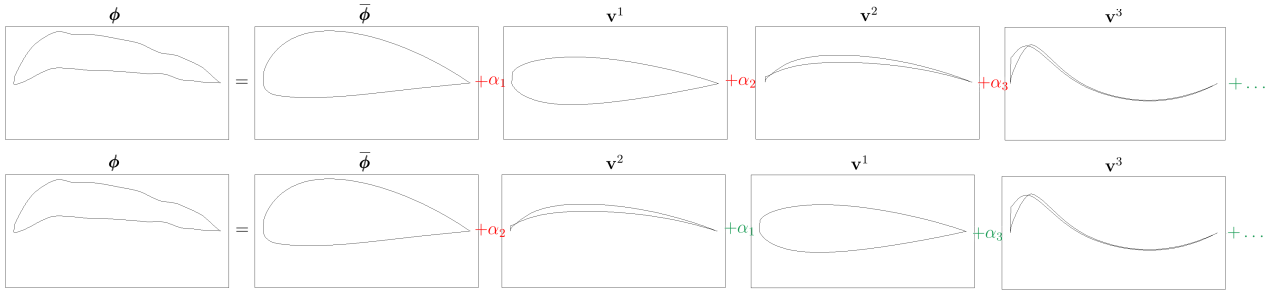

Figure 2: Variable selection on the NACA 22 benchmark by penalized maximum likelihood. For the drag (top), the three first eigenshapes that act on the shape, hence on its drag, are selected (red coefficients). For the lift, only the second eigencomponent ($\mathbf{v}^2$) is selected (bottom). Indeed $\mathbf{v}^2$ modifies the camber of the airfoil, hence it plays a major role on the lift. The other eigenbasis vectors (green coefficients) are estimated to be less influent on y.

Completely omitting the "non-active" dimensions, $\boldsymbol{\alpha}^{\overline{a}} \in \mathbb{R}^{D-\delta}$, and building the surrogate model $Y(\cdot)$ in the sole $\boldsymbol{\alpha}^a$ space is not satisfactory however, as it is equivalent to erasing some geometric patterns of the shapes which may contribute to small variations of y. For this reason, an additive GP with zonal anisotropy [17] between the active eigenshapes and the residual ones is considered: $Y(\boldsymbol{\alpha}) = Y^a(\boldsymbol{\alpha}^a) + Y^{\overline{a}}(\boldsymbol{\alpha}^{\overline{a}})$. $Y^a(\boldsymbol{\alpha}^a)$ is the anisotropic main-effect GP which works in the reduced space of active variables. It requires the estimation of $\delta + 1$ hyper-parameters (the length-scales θ_j and a GP variance) and aims at capturing most of y's variation, related to $\boldsymbol{\alpha}^a$'s effect. $Y^{\overline{a}}(\boldsymbol{\alpha}^{\overline{a}})$ is a GP over the large space of inactive components. It is a GP which solely aims at taking residual effects into account. Thus, a modeling assumption to keep it tractable is to consider it isotropic, i.e., it only has 2 hyper-parameters, a unique length-scale and a variance. In the end, even though $Y(\boldsymbol{\alpha})$ operates with $\boldsymbol{\alpha}$'s $\in \mathbb{R}^D$ and there are fewer observations than dimensions, $n \ll D$, it remains tractable since only a total of $\delta + 3$ hyper-parameters have to be learned, which remains affordable even when the number of observations is small. This additive model can be interpreted as a GP in the $\boldsymbol{\alpha}^a$ space, with an inhomogeneous noise fitted by the $Y^{\overline{a}}$ GP.

The effect of this refined surrogate model is a better predictive capacity, demonstrated by its increased performance on the NACA 22 test set shown in Figure 3. The additive GP over the space of active and inactive variables (left "Selection" boxes) achieves better performance than the GP which only uses the $\delta = 3$ first principal components (central "3 first" boxes). It also outperforms the classical approach of building the metamodel in the space of CAD parameters (right "CAD params" boxes). An interesting phenomenon is visible in the two rightmost plots of Figure 3 where lift is the output: even though the three first eigenvectors $\mathbf{V}_{1:3} = (\mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^3)$ contribute to 98.5% of the shape discretizations variability, metamodeling in the space they span leads to a deteriorated performance, the prediction being even worse than in the 22-dimensional CAD parameters space. This is explained by the fact that none of these eigenshapes contains bumps, which appear from the 4th one. Since the bumps contribute to the airfoils camber hence to the lift, critical information is lost and performance is degraded. This highlights the benefits of selecting the variables that impact y for the main effect GP, and of not completely disregarding the remaining variables.

4 Optimization in reduced dimension

We now turn to the problem of finding the shape that minimizes an expensive objective function $f(\cdot)$. To this aim, we employ the previous additive GP, which works in the space of eigencomponents $\boldsymbol{\alpha}$, in an Efficient

Global Optimization procedure [1]: at each iteration, a new shape is determined given the previous observations $\{(\boldsymbol{\alpha}^{(1)}, y_1), \ldots, (\boldsymbol{\alpha}^{(n)}, y_n)\}$ by maximizing the Expected Improvement (EI) as calculated with the GP $Y(\boldsymbol{\alpha})$: $\boldsymbol{\alpha}^{(n+1)} = \arg\max_{\boldsymbol{\alpha}\in\mathbb{R}^D} EI(Y(\boldsymbol{\alpha}))$. Although no call to the expensive $f(\cdot)$ is needed at this step, optimizing the EI is tricky as it is a multi-modal and high (D) dimensional function.

We can however take advantage of the dimension reduction beyond the construction of Y: $\boldsymbol{\alpha}^a \in \mathbb{R}^\delta$ are the variables that affect y the most and should be prioritized for the optimization of $f(\cdot)$. Maximizing the EI solely with respect to $Y^a(\boldsymbol{\alpha}^a)$ is nonetheless not an option, as the full GP $Y(\cdot)$ requires the knowledge of $\boldsymbol{\alpha} = [\boldsymbol{\alpha}^a, \boldsymbol{\alpha}^{\overline{a}}]$. Moreover, $\boldsymbol{\alpha}^{\overline{a}}$ act as local refinements to the shape that contribute a little to y, and should be optimized also.

For these reasons, instead of maximizing EI($[\boldsymbol{\alpha}^a, \boldsymbol{\alpha}^{\overline{a}}]$) , we maximize the EI with respect to $\boldsymbol{\alpha}^a$ and use a random embedding [9] to coarsely optimize the components $\boldsymbol{\alpha}^{\overline{a}}$: we maximize EI($[\boldsymbol{\alpha}^a, \alpha'\mathbf{A}]$), where $\alpha' \in \mathbb{R}$ is the coordinate along a random line in the $\boldsymbol{\alpha}^{\overline{a}}$ space, $\mathbf{A}$. Since $\boldsymbol{\alpha}^{\overline{a}}$ have been classified as inactive, it is not necessary to make a large effort for their optimization. The EI maximization is hence carried out in a much more tractable $\delta + 1$ -dimensional space and still has analytical gradients. From its optimum $\boldsymbol{\alpha}^* = (\boldsymbol{\alpha}^{a*}, \alpha'^*) \in \mathbb{R}^{\delta+1}$ arises a D-dimensional vector, $\boldsymbol{\alpha}^{(n+1)} = (\boldsymbol{\alpha}^{a*}, \alpha'^* A_1, \ldots, \alpha'^* A_{D-\delta})$ to be evaluated by the true function.

A last point needs however to be fixed: the expensive simulator neither works with shapes nor with $\boldsymbol{\alpha}$'s but with CAD parameters $\mathbf{x}$. We have to solve the *pre-image* problem, that is to say to find the CAD parameter vector $\mathbf{x}$ whose representation in the $\mathbf{V}$ basis equals $\boldsymbol{\alpha}^{(n+1)}$. Because there are more $\boldsymbol{\alpha}$'s than $\mathbf{x}$'s, $D \gg d$, a strict equality may not be feasible and the pre-image problem is relaxed into: $\mathbf{x}^{(n+1)} = \arg\min_{\mathbf{x}\in X} \|\mathbf{V}^\top(\phi(\mathbf{x}) - \overline{\boldsymbol{\phi}}) - \boldsymbol{\alpha}^{(n+1)}\|^2$. To complete an iteration, $\mathbf{x}^{(n+1)}$, the parametric shape that resembles $\boldsymbol{\alpha}^{(n+1)}$ the most, is evaluated by the computer code, which returns $y_{n+1} = f(\mathbf{x}^{(n+1)})$. Notice that the surrogate model is updated with y_{n+1} and $\boldsymbol{\alpha}^{(n+1)}$ and not with the $\mathbf{x}^{(n+1)}$.

Figure 3 shows an optimization run for the minimization of the NACA 22's drag. The main advantage of our approach (bottom left) is that it enables an early search for low drag airfoils. The standard approach (bottom center) needs much more function evaluations for building the initial surrogate model (black dots) because the inputs live in a space of higher dimension. Furthermore, the approach introduced in this paper would gain in importance in problems with much more than $d = 22$ CAD parameters, where it would almost be impossible to build a large enough initial design of experiments.

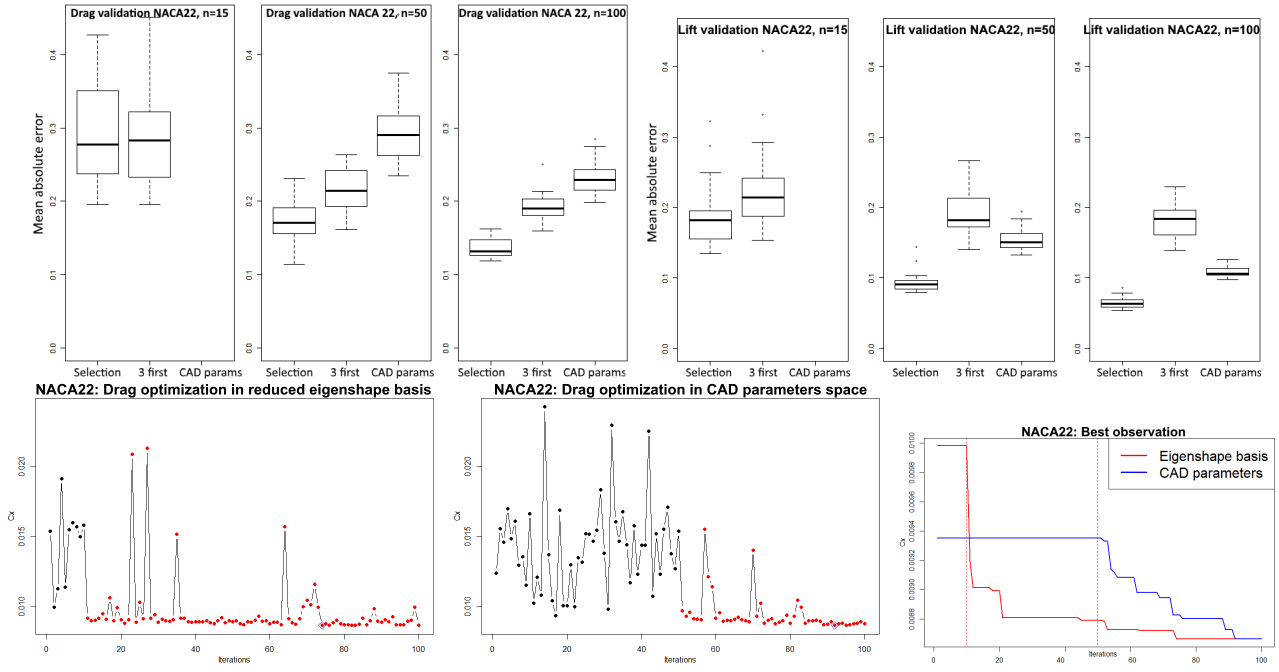

Figure 3: Top row: mean absolute prediction errors in drag (left) and lift (right) of the additive GP with zonal anisotropy and mode selection based on penalized likelihood ("Selection"), of an anisotropic GP using the first 3 eigenshapes ("3 first") and of an anisotropic GP built in the space of CAD parameters ("CAD params"). Bottom row: Drag optimization of the NACA 22 airfoil in the reduced eigenbasis (left) or carried out in the CAD parameters space (center). Low-drag airfoils are found while the classical method still evaluates the airfoils of the initial design of experiments (right).

5 Conclusions

In this paper we have proposed a new methodology to apply Bayesian optimization techniques to parametric shapes. Instead of working directly with the CAD parameters, which are too numerous for an efficient optimization and may not be the best representation of the underlying shape, we unveil the lower dimensional manifold of shapes through an auxiliary mapping. The dimensions of this manifold that contribute the most to the variation of the output are prioritized for building a surrogate model, which is utilized for Bayesian optimization. A significant dimension reduction and speed-up is achieved on an aerodynamic CAD benchmark.

Acknowledgments

This research was performed within the framework of a CIFRE grant (convention #2016/0690) established between the ANRT and the Groupe PSA for the doctoral work of David Gaudrie.

References

1. Jones D R, Schonlau M and Welch W J. *Efficient global optimization of expensive black-box functions.* Journal of Global Optimization, 1998, 13(4), 455-492.
2. Rasmussen C E and Williams C K. *Gaussian processes for machine learning.* MIT Press, 2006, Vol 2, No 3.
3. Raghavan B, Breitkopf P, Tourbier Y and Villon P. *Towards a space reduction approach for efficient structural shape optimization.* Structural and Multidisciplinary Optimization, 2013, 48(5), 987-1000.
4. Raghavan B, Le Quilliec G, Breitkopf P, Rassineux A, Roelandt J M and Villon P. *Numerical assessment of springback for the deep drawing process by level set interpolation using shape manifolds.* International Journal of Material Forming, 2014, 7(4), 487-501.
5. Stegmann M B and Gomez D D. *A brief introduction to statistical shape analysis.* Informatics and Mathematical Modelling, Technical University of Denmark, DTU, 2012, 15(11).
6. Wang Q. *Kernel principal component analysis and its applications in face recognition and active shape models.* arXiv preprint arXiv:1207.3538. 2012
7. Bouhlel M A, Bartoli N, Otsmane A and Morlier J. *Improving kriging surrogates of high-dimensional design models by Partial Least Squares dimension reduction.* Struct. and Mult. Optimization, 53(5), 935-952, 2016.
8. Li J, Bouhlel M A and Martins J. *A data-based approach for fast airfoil analysis and optimization.* AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference (p. 1383), 2018.
9. Wang Z, Zoghi M, Hutter F, Matheson D and De Freitas N. *Bayesian optimization in high dimensions via random embeddings.* 23rd International Joint Conference on Artificial Intelligence, 2013.
10. Schölkopf B, Smola A and Müller K R. *Kernel principal component analysis.* International conference on artificial neural networks, 1997, 583-588.
11. Vapnik V, *The nature of statistical learning theory.* 2013.
12. Gaudrie D, Le Riche R, Picheny V, Enaux B and Herbert V. *Budgeted Multi-Objective Optimization with a Focus on the Central Part of the Pareto Front - Extended Version.* arXiv preprint arXiv:1809.10482. 2018.
13. Berkooz G, Holmes P and Lumley J L. *The proper orthogonal decomposition in the analysis of turbulent flows.* Annual review of fluid mechanics, 1993, 25(1), 539-575.
14. Yi G, Shi J Q and Choi T. *Penalized Gaussian Process Regression and Classication for High-Dimensional Nonlinear Data.* Biometrics, 2011, 67(4), 1285-1294.
15. Ben Salem M, Bachoc F, Roustant O, Gamboa F and Tomaso L. *Sequential dimension reduction for learning features of expensive black-box functions.* 2018.
16. Deville Y, Ginsbourger D, Durrande N, Roustant O. Package 'kergp'. 2018.
17. Allard D, Senoussi R and Porcu E. *Anisotropy models for spatial data.* Mathematical Geosciences, 2016, 48(3), 305-328.