Disciplinary Proper Orthogonal Decomposition and Interpolation for the resolution of parametrized Multidisciplinary Analysis

Gaspard Berthelin, Sylvain Dubreuil, Michel Salaün, Nathalie Bartoli, Christian Gogu

Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

Disciplinary Proper Orthogonal Decomposition and Interpolation for the resolution of parametrized Multidisciplinary Analysis

G. Berthelin[1,2] | S. Dubreuil[1] | M. Salaün[2] | N. Bartoli[1] | C. Gogu[2]

[1]DTIS, ONERA, Université de Toulouse, Toulouse, France

[2], Université de Toulouse, CNRS, UPS, INSA, ISAE, Mines Albi, Institut Clément Ader (ICA), Toulouse, France

Correspondence
Gaspard Berthelin. Email: gaspard.berthelin@onera.fr

Abstract

This article proposes a new method for the resolution of parametrized multidisciplinary analysis (MDA) using disciplinary surrogate models within a multi-query context. The main idea is to replace the costly disciplinary solvers of the MDA by Proper Orthogonal Decomposition and Interpolation models. The main challenge we address is the high dimensional coupling variables whose ranges are unknown. To overcome this issue, a training strategy is developed by uncoupling the disciplinary solvers from the MDA context. This new surrogate MDA called Disciplinary Proper Orthogonal Decomposition and Interpolation (DPOD+I) is iteratively enriched to solve the analysis with an estimation of the error made by the surrogate model. The disciplinary surrogate model which has the most influence on this error is determined by a sensitivity analysis and thus enriched. This approach allows to uncouple the disciplinary solvers during the training and enrichment phases. The approach is applied to various aeroelastic problems of an aircraft wing and allows to reduce by a factor 5 the mean number of disciplinary solver calls needed for the resolution of the MDA.

KEYWORDS:
MDA solver, surrogate models, Gaussian Process, POD, POD+I, multi-query context

1 | INTRODUCTION

This article addresses the numerical solving of parametrized multidisciplinary analysis (MDA) in a multi-query context such as multidisciplinary design optimization (MDO) or reliability analysis among others. In both cases, numerous calls to a quantity of interest depending on the solution of a complex parametrized coupled problem is necessary. This coupled system, defined by the interaction between several disciplines through some coupling variables is called MDA. The basic idea of the MDA is to take into account these interactions to accurately describe complex physical phenomena. A common example of MDA is the determination of the static aeroelastic equilibrium of an aircraft wing. The aeroelastic equilibrium is affected by parameters or design variables such as the angle of attack, the dimensions and properties of the structure. This fluid structure interaction problem couples an aerodynamic model and a structural mechanics one. Indeed the elastic deformation of a wing is derived from the forces that are applied to the structure and the aerodynamic loads are defined by the wing shape and therefore by the elastic deformation of the structure. Thus, the solution of this problem for a given parameter needs the resolution of a coupled problem involving the two disciplines. For the last decades, the intensive use of MDA has been struggling with the use of high fidelity

[0]**Abbreviations:** MDA, multidisciplinary analysis; POD, proper orthogonal decomposition; POD+I, proper orthogonal decomposition and interpolation; MDO, multidisciplinary design optimization; GP, Gaussian process; pROM, parametric model order reduction methods; DoE, design of experiments; PCE, polynomial chaos expansion; SVD, singular value decomposition; LSH, Latin hypercube sampling; DPOD+I, disciplinary proper orthogonal decomposition and interpolation; MC, Monte Carlo; VLM, vortex lattice method; CV, coefficient of variation

models for which the cost of the disciplinary solvers can be prohibitive when used for more than a single analysis. Here, a focus is made on the resolution of MDA problems where the coupling variables are fields discretized over a mesh (e.g. displacement and pressure fields) and the disciplinary solvers are black boxes (e.g. industrial softwares), limiting the use of intrusive methods. In this study, the focus is on partitioned techniques for independent disciplinary solvers [1].

In MDO context, a survey of several formulations for the resolution of the MDO have been proposed by [2] where the coupling and design variables are managed in various ways. Those formulations can be split into two categories: the one-level strategies are those where the design variables and the coupling ones are handled simultaneously (all-at-one (AAO), Simultaneous Analysis and Design (SAND), etc.). Those strategies might be difficult to use as handling large dimensional coupling variables like fields discretized over a mesh is challenging. Then, a particular interest is given to the feasible strategy [3] where the optimization and the resolution of the MDA are considered separately. The resolution of the MDA is solved by a dedicated method that can handle a large number of coupling variables and that can be used for various multi-query problems (MDO, reliability analysis,...). However, the repeated resolutions of the MDA involve a high number of calls to the disciplinary solvers leading to some intractable computational cost issues when high fidelity models are involved. To reduce this computational cost, a surrogate model of the quantity of interest given by the MDA can be built providing an inexpensive approximation of this quantity (see for examples [4,5,6]). Those methods are not dedicated to the MDA and do not exploit the partitioned procedure used to solve the MDA. Even if this kind of strategy succeeds in reducing the number of MDA calls, each enrichment of the surrogate model implies to solve the whole MDA. Consequently, in order to further reduce the cost of the MDA computation, few studies propose to take advantage of the partitioned methods used to solve the MDA.

One way to benefit from this partitioned approach is to replace the costly disciplinary solvers by surrogate models. For example, authors of [7,8,9,10] use surrogate models of the disciplinary solvers and couple them to solve an approximate version of the MDA. These studies rely on an independent construction of each surrogate model and do not consider the resulting approximation error made by coupling the disciplinary surrogate models. To improve the disciplinary surrogate model construction in an MDO context, it is proposed in [11] to use Gaussian Processes (GPs) to approximate each disciplinary solver. Then, a dedicated strategy is applied to solve the MDO by taking into consideration the randomness brought by the GPs during the MDA resolution. Nevertheless, it should be noted that learning a GP surrogate model from the disciplinary solvers requires a low number of coupling variables (low dimensional input space) which is a strong limitation of the previous approach. Indeed, the approximation of models with a high number of inputs and outputs is still a challenging task.

With respect to this issue, in the context of single disciplinary approximation, the parametric model order reduction methods (pMOR) have been widely developed in the last decades. Among all the existing pMOR approaches, a particular attention is given to projection methods. The main idea is to project the physical equations of the disciplinary solver onto a low dimensional subspace using a dedicated projection basis. Several methods exist to compute the projection basis such as snapshot Proper Orthogonal Decomposition (POD) [12], greedy procedures [13], goal-oriented methods [14] among others. Although these methods allow to drastically reduce the number of degrees of freedom associated to the physical solver, they require the assembly and projection of large scale operators when no affine decomposition of these operators with respect to the parameters of interest is available. To overcome this issue, several techniques using hyper-reduced-order models [15] based on discrete empirical interpolation [16] were proposed. They are based on computing few components of the large scale operator to build approximations but remain intrusive as the sparse sampling is not always achievable. Moreover, all of them are based on an offline-online strategy. In fact, a computationally expensive offline stage is needed to build the reduced model using many calls to the disciplinary solver. As no information regarding the interesting regions of the design space are available, the reduced models have to be accurate on the whole input space. The computational burden of the offline stages can be prohibitive and can dominate the computational cost of the entire process. To tackle this issue, Proper Orthogonal Decomposition and Interpolation (POD+I) [17,18,19] aims at approximating the POD coefficients using surrogate models. Once the POD basis is assembled, a regression-based approach is used to establish a mapping from design variables to projection coefficients onto the POD basis. This ensures a complete decoupling of offline and online stages as the online solutions only require direct outputs from the reduced-order regression model. Therefore, those methods are non-intrusive but suffer from the lack of information concerning the error made by the regression model. One of the possible ways to address this issue is the use of GPs [20] that give an estimation on the error made on the interpolation. It should be noted that POD+I relies on a Design of Experiments (DoE) containing samples of the disciplinary solver solutions. This DoE is challenging to build in the context of coupled simulations. Indeed, in that particular case, the definition sets of the coupling variables are unknown which creates a difficulty for the construction of the DoE used to train the disciplinary surrogate models. In fact, POD+I has mainly been used to give an approximation of the solution of the whole MDA such as in [19,21,22].

In this paper an approximation of each disciplinary solver is proposed using a POD+I type technique to adapt the strategy developed in [11] in a coupled context simulations with high-dimensional coupling variables. A similar approach has been proposed in [23] where only one of the disciplinary solvers was approximated. This approach allows to drastically reduce the number of solver calls for one discipline in an optimization context. However, both disciplinary solvers are considered costly in our study. The training of the POD+I in [23] is made using incomplete MDA solutions; in this study, another training strategy will be proposed in addition to an estimation of the error made on the interpolation. Our approach is based on a reduction of the coupling variables using POD. This reduction of the coupling variables defines a new reduced MDA that can be learned using surrogate models to get an inexpensive approximation of each discipline. The challenge is the training of the surrogate models in a multidisciplinary environment where coupling variables depend on each other output through the feedback loops. Those coupling variables live in a large, unknown space leading to a dedicated sampling strategy. More precisely, the training strategy of both the POD basis and the DoE for the surrogate models is the first contribution of this paper. The second contribution is an original adaptive enrichment of the surrogate models allowing to give precise estimation of the MDA. The enrichment strategy relies on a sensitivity analysis. Thus, a probabilistic estimation of the error is developed based on the surrogate model information. The tolerance on this online error can be adjusted arbitrarily by the user.

Presentation of these two contributions is organized as followed. In Section 2 the construction of the POD basis and GPs in the MDA context are presented. This section highlights the issue of the construction of the disciplinary DoE and proposes a sequential sampling for the initial disciplinary DoE. In Section 3 an enrichment strategy is described to solve the MDA using the surrogate model. Finally, the performance of the proposed method is assessed on various static aeroelastic examples in Section 4.

2 | SURROGATE MODEL OF THE MDA

In this study, the approach is presented on an aeroelastic example with two disciplines. Note that the same methodology can be adapted with more than two disciplines.

2.1 | Problem formulation

In the following, the superscript ^{1} will refer to the discipline 1 and the superscript ^{2} to the discipline 2. A parametrized MDA problem, of two disciplines involving feedback loops, could be written as a system of two equations,

$$\begin{cases} y^1 = f^1(z, y^2) \\ y^2 = f^2(z, y^1) \end{cases} \tag{1}$$

where the design variables z belong to a design space $\mathcal{Z} \subset \mathbb{R}^q$. The coupling variables of the system are $y^1 \in C^1 \subset \mathbb{R}^{d^1}$ and $y^2 \in C^2 \subset \mathbb{R}^{d^2}$, and f^1 and f^2 are the solvers of the two disciplines. The solutions of this system, which only depend on the design variables are denoted by $y^1_*(z)$ and $y^2_*(z)$. This system is illustrated in Figure 1. In an aerostructure example, y^1 are the displacements of the wing nodes and y^2 are the aerodynamic pressures that are applied to the structure. As we assume that those solvers are black boxes, one can solve the system using an iterative algorithm. Here a non linear Gauss Seidel solver is used and described in the Appendix. The two user defined parameters of this MDA solver are the initial guess y^1_0 and a convergence tolerance ϵ_{MDA}. This first guess y^1_0 defines the initial point of the iterative solver and is usually obtained by the solution of the MDA for the design variable z_0 corresponding to the center of the design space $\mathcal{Z}$ if it's available or from physical considerations (zero displacement for the structure for example). The solver ends when the relative error between two successive iterates is lower than ϵ_{MDA}. In the following, existence of a unique solution for Eq. (1) is assumed by physical considerations and the solution obtained using the MDA solver is denoted by exact MDA.

The proposed methodology to solve this system is to build a surrogate model of each disciplinary solver. Unfortunately, surrogate models like Gaussian Processes (GPs), radial basis functions, Polynomial Chaos Expansion (PCE) cannot be used in this context as coupling outputs are high dimensional vectors ($d^1 \gg 1$ and $d^2 \gg 1$). To overcome this issue, a reduction of the coupling variables is proposed. The objective is to replace both disciplines by surrogate models using a POD+I strategy. This method will be split into two steps: the reduction of the coupling variables using POD and the interpolation with the use of surrogate models to approximate POD coefficients. In this approach, both disciplinary solvers are replaced by surrogate models unlike the approach developed in [21,22] where only one of the disciplinary solver was approximated. To train the surrogate models, physical simplifications are developed to avoid expensive exact MDA [10].

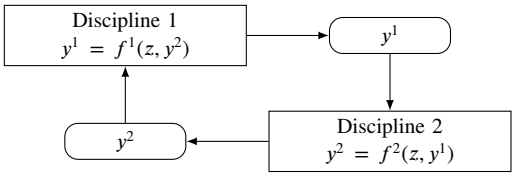

FIGURE 1 Illustration of the MDA problem involving 2 disciplines.

2.2 | Reducing the dimensionality of the coupling variables

The coupling variables are the outputs of numerical solvers representing physical phenomena and a reduction of their dimensionality is needed. Thankfully, it has been shown that for many physical problems, those outputs can be well described by few dominant modes allowing to reduce drastically the dimensionality. To define this low dimensional subspace, called projection space, several methods have been studied[24,25]. In this paper, the POD method is considered as it has been widely used and has shown its efficiency on computational fluid dynamics and structural problems, for both linear and non linear solvers.

The POD is described here for the first discipline and the same methodology holds for the second. The POD aims at representing the output of each solver on a linear subspace:

$$y^1(z, y^2) \approx \tilde{y}^1(z, y^2) = c^1 + \sum_{j=1}^{n^1} \alpha_j^1(z, y^2)\phi_j^1 \tag{2}$$

where $\tilde{y}^1$ is an approximation of the vector y^1, ϕ_j^1 is the j^{th} basis vector for the solver 1, α_j^1 is the corresponding coefficient, and c^1 is a constant vector associated with the projection basis. The number of terms n^1 is chosen for the approximation and is described in the following. In practice, the POD basis is built using the method of snapshots. First, several computations of disciplinary solver are performed to get a representative set of the variety of possible solutions. Let $((y^1)^k)_{k=1}^m$ be a set of m snapshots ($m \ll d^1$) computed for different inputs. The mean of the snapshots reads:

$$c^1 = \sum_{k=1}^{m} \frac{(y^1)^k}{m} \tag{3}$$

The centered solutions are stored in a $d^1 \times m$ matrix where d^1 is the dimension of y^1.

$$\psi = \left[(y^1)^0 - c^1 | \dots | (y^1)^{m-1} - c^1\right] \tag{4}$$

Then, a Singular Value Decomposition (SVD)[26] on ψ is achieved leading to

$$\psi = USV^T \tag{5}$$

where U is a $(d^1 \times d^1)$ orthonormal matrix containing the left singular vectors, S is a $(d^1 \times m)$ diagonal matrix with diagonal entries containing the singular values σ_i and V is a $(m \times m)$ orthonormal matrix containing the right singular vectors. The singular values are sorted decreasingly: $\sigma_1 \geq \dots \geq \sigma_m \geq 0$. For a given number of basis vector s, the optimal basis, i.e that minimizes the mean of the quadratic projection error, is given by the s first columns of U. The number n^1 ($n^1 \ll d^1$) of basis vector retained is

$$n^1 = \min_s s \text{ subject to } \frac{\sum_{i=1}^{s} \sigma_i}{\sum_{i=1}^{m} \sigma_i} > \eta \tag{6}$$

where η represents the percentage of captured energy by the POD basis from the original model. Let ϕ_j^1 with $j \in [|1, \dots, n^1|]$ be the columns of ϕ^1 defining the linear basis in Eq. (2). With this basis, the projection application Φ^1 is now defined by:

$$\begin{aligned} \Phi^1 : C^1 &\to \tilde{C}^1 \subset \mathbb{R}^{n^1} \\ u^1 &\mapsto \Phi^1(u^1) = \{\alpha_1^1, \dots, \alpha_{n^1}^1\} = (\phi^1)^T(u_1 - c_1) \end{aligned} \tag{7}$$

where $\tilde{C}^1$ is the image of C^1 by Φ^1. The reverse of Φ^1, noted Φ^{-1}, allowing to reconstruct the full vector from the coefficients is defined by:

$$\begin{aligned} \Phi^{-1} : \quad \tilde{C}^1 \quad &\to C^1 \\ \alpha_1^1, \dots, \alpha_{n^1}^1 &\mapsto \Phi^{-1}(\alpha_1^1, \dots, \alpha_{n^1}^1) = c^1 + \textstyle\sum_{j=1}^{n^1} \alpha_j^1 \phi_j^1 \end{aligned} \tag{8}$$

The same methodology applied to the second discipline allows to define a projection function Φ^2 and its inverse Φ^{-2}.

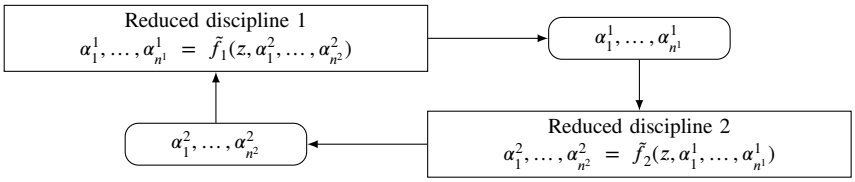

FIGURE 2 Illustration of the reduced MDA using POD coefficients.

Finally, two functions allowing the transition from a large dimensional space to a reduced one have been built and will be used to efficiently represent the solutions from each solver by mean of a few coefficients (usually less than 20). The new MDA system using the coefficients of the POD as coupling variables reads

$$\begin{cases} \alpha_1^1(z, \alpha_1^2, \dots, \alpha_{n^2}^2), \dots, \alpha_{n^1}^1(z, \alpha_1^2, \dots, \alpha_{n^2}^2) = \tilde{f}^1(z, \alpha_1^2, \dots, \alpha_{n^2}^2) \\ \alpha_1^2(z, \alpha_1^1, \dots, \alpha_{n^1}^1), \dots, \alpha_{n^2}^2(z, \alpha_1^1, \dots, \alpha_{n^1}^1) = \tilde{f}^2(z, \alpha_1^1, \dots, \alpha_{n^1}^1) \end{cases} \tag{9}$$

where $\tilde{f}^1(z, \alpha_1^2, \dots, \alpha_{n^2}^2) = \Phi^1(f^1(z, \Phi^{-2}(\alpha_1^2, \dots, \alpha_{n^2}^2)))$ and $\tilde{f}^2(z, \alpha_1^1, \dots, \alpha_{n^1}^1) = \Phi^2(f^2(z, \Phi^{-1}(\alpha_1^1, \dots, \alpha_{n^1}^1)))$. In the following, $\tilde{f}^1$ and $\tilde{f}^2$ are called reduced disciplinary solvers. This new reduced MDA is illustrated in Figure 2. Once the coupling variables reduced, the next objective is to build Gaussian Process (GP) approximations of the POD coefficients to surrogate the reduced MDA.

Remark:

One of the main difficulties here is the dependency of the coupling variables on each other. Indeed, this phenomena has to be taken into account for the construction of the snapshots. A partial solution is proposed in Section 2.4 by mean of a specific sampling strategy.

2.3 | Surrogate model of the reduced MDA

2.3.1 | Interpolation

Let us consider a continuous function we seek to approximate by a surrogate model:

$$\begin{aligned} f : \Omega \in \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto f(x) \end{aligned} \tag{10}$$

The idea is to approximate the function using GPs conditioned on a few observations of the function output [27]. The conditioned GP is built over a prior GP defined by:

- a prior mean $\mu_{prior}(x) = F(x)\beta$ computed using a regression on a set basis of functions defined by $F(x)$ and by their associated coefficients β.
- a prior variance σ^2_{prior} defined by the covariance function (sometimes called kernel):

$$k(v, w) = \text{cov}(\mathcal{G}(v), \mathcal{G}(w)) = \sigma_0 \Psi(v, w, \theta) \tag{11}$$

  where Ψ is a user defined correlation function.

An estimation of hyperparameters σ_0, β and θ is necessary to ensure an accurate response for any unknown point of the domain. For a fixed kernel, several techniques exist to obtain the optimal values of these hyperparameters, for example Maximum Likelihood Estimation or cross-validation.

Then, the GP is conditioned relatively to the observations of the function f. For any input $x \in \mathbb{R}^n$, the conditioned GP is defined by its mean μ and its variance σ:

$$\mathcal{G}(x) \sim \mathcal{N}(\mu(x), \sigma^2(x)) \tag{12}$$

For additional details on the construction of the GP, readers are referred to [27].

2.3.2 | Surrogate models of the reduced disciplinary solvers

The purpose of this section is to apply a POD+I like methodology for MDA analysis. It should be noted that the POD+I strategy cannot be applied to replace the disciplinary solver directly. Indeed, the inputs of the surrogate models are composed of design

variables z and coupling variables y^1 or y^2, so their dimensions are quite high. To handle this issue, a surrogate model of the reduced MDA illustrated in Figure 2 is proposed. This approach is called Disciplinary Proper Orthogonal Decomposition and Interpolation (DPOD+I) in the following.

To distinguish deterministic variables from random ones, all random variables will be denoted by a hat symbol. Thanks to the reduced MDA (system of Eq. (9)), the dimension of the coupling variables has been drastically reduced. After the POD projection, the projected disciplines $\tilde{f}^1$ and $\tilde{f}^2$ can be approximated by disciplinary surrogate models. The idea here is to replace each coefficient on the POD basis by a GP, denoted by the hat symbol:

$$\begin{array}{l} \hat{\alpha}_i^1(z, \alpha_1^2, \dots, \alpha_{n^2}^2) \text{ for } \alpha_i^1(z, \alpha_1^2, \dots, \alpha_{n^2}^2) \ \forall i = 1, \dots, n^1 \\ \hat{\alpha}_j^2(z, \alpha_1^1, \dots, \alpha_{n^1}^1) \text{ for } \alpha_j^2(z, \alpha_1^1, \dots, \alpha_{n^1}^1) \ \forall j = 1, \dots, n^2 \end{array} \tag{13}$$

Consequently the coefficients of the reduced MDA are replaced by

$$\begin{array}{l} \hat{\alpha}_i^1(z, \alpha_1^2, \dots, \alpha_{n^2}^2) = \mu_i^1(z, \alpha_1^2, \dots, \alpha_{n^2}^2) + \hat{\epsilon}_i^1(z, \alpha_1^2, \dots, \alpha_{n^2}^2) \ \forall i = 1, \dots, n^1 \\ \hat{\alpha}_j^2(z, \alpha_1^1, \dots, \alpha_{n^1}^1) = \mu_j^2(z, \alpha_1^1, \dots, \alpha_{n^1}^1) + \hat{\epsilon}_j^2(z, \alpha_1^1, \dots, \alpha_{n^1}^1) \ \forall j = 1, \dots, n^2 \end{array} \tag{14}$$

where μ_i^1 and μ_j^2 are the means of the GP; $\hat{\epsilon}_i^1$ and $\hat{\epsilon}_j^2$ are random variables following a zero mean GP whose covariance function is the one of the corresponding GP.

It should be noted that the training of the disciplinary surrogate models implies the construction of a DoE and thus to sample the disciplinary solver with respect to the design space and the coupling variables spaces. If sampling in the design space is trivial, efficient sampling in the large dimensional coupling variables spaces is more difficult due to the unknown range of the corresponding variables. An original sampling method is thus proposed in Section 2.4.

Once the disciplinary surrogate models have been trained, the solution of the surrogate MDA is no longer deterministic as the exact coefficients onto the POD basis have been replaced by GPs. To handle the random variables introduced in the random MDA, a methodology is proposed in Section 3.1.

2.4 | Training of surrogate models

In this section, the process allowing to initiate the snapshots to build the POD basis and the DoE to train the GPs is described. One naive way to do is to use a sample of exact MDA (solving Eq. (1) for different design variables z). A sample of p design variables z_i, $i = 1, \dots, p$ over the space $\mathcal{Z}$ can be generated using an appropriate DoE method, Latin Hypercube Sampling (LHS)[28] for example. Then, the exact MDA associated to each design sample is computed using the MDA solver (or unconverged ones as proposed by[23]). For each design z_i, a certain number of iterations is needed. All those solver solutions are used as snapshots to build Φ^1 and Φ^2. Knowing Φ^1 and Φ^2, a DoE of $\tilde{f}^1$ and $\tilde{f}^2$ is obtained by projecting all the coupling variables inputs and the outputs of each solver call. However, this DoE is not efficient for the exploration of the coupling variable spaces. Indeed, during one MDA resolution, the design variables are constant (equal to z_i) and the coupling variables y^1 and y^2 converge to the exact MDA. The distribution of the DoE obtained with this basic strategy is illustrated on Figure 3. This figure shows the first two POD coefficients and the first design variable of an aerostructure example which will be presented in the following. In this figure, each color corresponds to the resolution of the MDA for a given design variable z_i. It had been observed that the training of the surrogate models with this DoE is not efficient as the space of the coupling variables is not explored by the successive iterations of the fixed point algorithm used to solve the MDA. One way to improve this method could be to not use exact MDA but rather some incomplete MDA analysis[23] to avoid the coupling variables y^1 and y^2 to remain around the coupled solutions.

Here, a specific methodology exploiting a decorrelation between design and coupling variables is proposed to train the GPs. The main purpose is to get a well distributed DoE allowing to build a global model of the parametric MDA. This DoE is challenging because no sampling method can be used to generate a well distributed sample for the coupling variables. Indeed, the spaces $\mathcal{C}^1$ and $\mathcal{C}^2$ are unknown or hard to define: $\mathcal{C}^1$ is the image of $\mathcal{Z} \times \mathcal{C}^2$ under f^1 and $\mathcal{C}^2$ is the image of $\mathcal{Z} \times \mathcal{C}^1$ under f^2 leading to high correlations. Thus, a method inspired by the MDA solver is used; instead of keeping the same design variables through the iterates, the design variables are randomly changed at each iteration. As the MDA, this method involves a first sample of initial guesses to begin the iterations. Then, the iterations and the new stopping criterion are described.

2.4.1 | First sample of the coupling variables

First of all, no solver computation can be achieved without an initial guess of the other solver solution. The objective of this task is to generate a sample of r initial guesses for each discipline. Once this sample is built, exact computation of the other

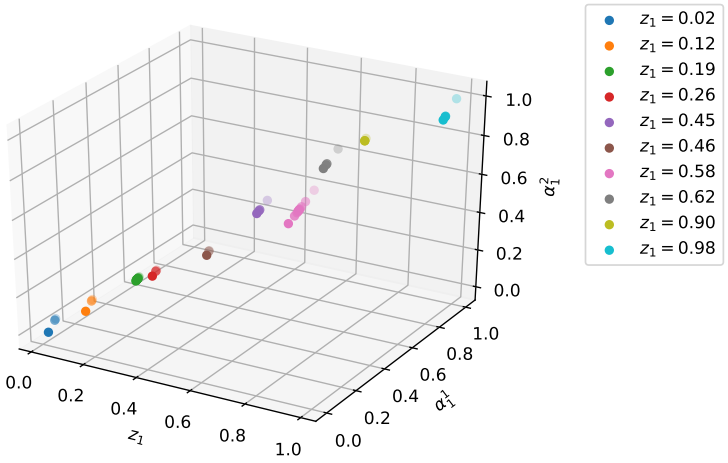

FIGURE 3 Distribution of the inputs of each disciplinary call during the resolution of 10 MDA associated to 10 different design variables. Here, only the first coefficient on each POD basis and the first design variable are displayed. Each MDA resolution is differentiated by different colors, the resolution of the 10 MDA involved 54 calls to each disciplinary solver.

discipline can be achieved using the given sample as inputs. To do so, one way to generate samples from $\mathcal{C}^1$ is to build a 1D manifold spanned by the initial guess of the MDA solver y_0^1. We seek to generate a manifold $\mathcal{F} = \{\lambda y_0^1, \lambda \in [\lambda^-; \lambda^+]\}$ where every element of $\mathcal{F}$ is likely to be an element of $\mathcal{C}^1$. To obtain the bounds λ^+ and λ^- a physical analysis is required. For example, in an aerostructure context, the maximum displacement of the wing lies between zero and a value given by some physical considerations over the wing structure. Next, the idea is to use a sample from $\mathcal{F}$ to generate a sample for the second discipline.

The algorithm corresponding to this method is described in Algorithm 1. The following steps of the associated algorithm are described in the following:

1. Generate r samples from $[\lambda^-; \lambda^+]$ by mean of a sampling method (LHS).
2. The initial guesses of the first discipline are given by the product between the r samples from $[\lambda^-; \lambda^+]$ and y_0^1.
3. Generate r design variables z_i from the design space $\mathcal{Z}$ by LHS for example.
4. The initial guesses of the second discipline are given by outputs of the exact disciplinary solver. The inputs of the disciplinary solver are given by the sample of design variables z_i and the initial guesses of the first discipline $(y^1)_i^0$.

Remarks:

- As for the MDA solver initial guess, high accuracy for the generated sample is not required because the influence of the initial guess will decrease during the iterative training process. The same holds for the estimation of λ^+ and λ^- whose influence will decrease through the iteration of the training process.
- The solver solutions computed in Algorithm 1 are not used to train the GPs.

2.4.2 | Training Algorithm

The objectives are now to get an accurate POD basis and data to train the GPs. The idea is to generate batchs of disciplinary solutions using the first samples obtained during Algorithm 1. To evaluate the accuracy of the POD basis, the relative projection error is computed for the newly computed batch. This allows to evaluate the accuracy of the POD basis on solutions that are independent from the one used to train the basis. The POD basis is accurate enough when the mean relative error does not exceed a tolerance ϵ_{PE} chosen by the user.

The steps of the method, called Algorithm 2, are described in the following:

1. Initiate the projection error and the iteration counter n

Algorithm 1: Compute r samples of coupling variables.

input : An initial guess y_0^1, a number $r > 0$, bounds λ^- and λ^+
$\{\lambda_i, i = 1, \dots, r\} \leftarrow r$ samples from $[\lambda^-; \lambda^+]$ obtained by LHS;
for $i = 1, \dots, r$ **do**
  $(y^1)_i^0 \leftarrow \lambda_i y_0^1$;
end
$\{z_i, i = 1, \dots, r\} \leftarrow$ containing r samples from $\mathcal{Z}$ obtained by LHS;
for $i = 1, \dots, r$ **do**
  $(y^2)_i^0 \leftarrow f_2(z_i, (y^1)_i^0)$;
end
output: $\{(y^1)_i^0, i = 1, \dots, r\}$, $\{(y^2)_i^0, i = 1, \dots, r\}$

2. Compute samples of both disciplines until the mean projection error is lower than ϵ_{PE}. To do so, a sample of design variables $\{(z^1)_i^n,\ i = 1, \dots, r\}$ is generated using LHS. Then, each design variable is associated to a sample of $\{(y^2)_i^{n-1},\ i = 1, \dots, r\}$ to get the inputs of the first discipline solver. The exact solution to the corresponding input is computed using the disciplinary solver. The relative error between the new computed solution and its POD approximation is computed. The mean projection error for the first discipline is obtained by the mean of the relative errors obtained.

3. Another sample of design variables $\{(z^2)_i^n,\ i = 1, \dots, r\}$ is generated using LHS. Each design variable is associated to the sample of the newly computed solution of the first discipline $\{(y^1)_i^n,\ i = 1, \dots, r\}$. The exact solution of the corresponding input is computed using the disciplinary solver. The relative error between the newly computed solution and the POD approximation is computed. The mean projection error for the second discipline is thus obtained by the mean of the relative errors obtained.

4. The POD basis are generated by SVD with the information given by all the computed solver solutions.

5. If the criterion on the relative error is satisfied, the algorithm ends and a robust POD basis is obtained defining Φ^1 and Φ^2. Then, each solver call of the disciplinary solver is used to build a DoE for the reduced disciplinary solver. $\{(z^1)_i^j, \Phi^2((y^2)_i^{j-1})\} \rightarrow \Phi^1((y^1)_i^j),\ i = 1, \dots, r,\ j = 1, \dots, n$ are observations of $\alpha_1^1, \dots, \alpha_{n^1}^1$ and $\{(z^2)_i^j, \Phi^1((y^1)_i^j)\} \rightarrow \Phi^2((y^2)_i^j),\ i = 1, \dots, r,\ j = 1, \dots, n$ are observations of $\alpha_1^2, \dots, \alpha_{n^2}^2$.

Finally we succeed in building a POD basis for each discipline and a DoE to train the GPs. The obtained DoE is illustrated in Figure 4 where the inputs for each disciplinary solver call are displayed with respect to the first design variable and the first POD coefficient for each discipline. One can note that, compared to Figure 3, the obtained DoE is spread in the input spaces thus the distribution of points is much better suited for a surrogate construction.

Remark:

As no MDA is solved for the training, the accuracy of the POD basis is not proven for the whole coupling variables spaces and even more for the space where the exact MDA lives. A local correction of the POD basis will be proposed in the *online* phase to solve this issue.

3 | RESOLUTION OF THE MDA USING SURROGATE MODELS

The objective of this section is to use a MDA solver, called DPOD+I, to solve the MDA analysis with an enrichment strategy based on the work presented in [11]. In Section 2.4.2, an initial model is trained to obtain an accurate POD basis for each discipline. However, the GPs constructed on this initial basis are not accurate enough. Thus an enrichment of the GPs is proposed. To do so, a local model is initialized as a copy of the initial model obtained during the training phasis in Section 2.4.2 and locally enriched. Then, the data gathered during the local enrichment are filtered and used to enrich the initial model. The purpose of the filtering is to avoid poorly distributed inputs for the GPs that could lead to ill-conditioned GPs as explained in Section 2.4.

The first step is the enrichment of the local model to get an accurate approximation of the exact MDA for a given value $z = z_*$. To illustrate this process through figures, a simplification is made by considering only one vector in each POD basis. Thus each

Algorithm 2: Compute a POD basis and a training set for the GPs.

input : $\{(y^2)_i^0,\ i = 1, \dots, r\}$, a tolerance over the projection error ϵ_{PE}
projection error $\leftarrow 1 + \epsilon_{PE}$;
$n \leftarrow 1$;
while *projection error* $> \epsilon_{PE}$ **do**
  sample $\{(z^1)_i^n,\ i = 1, \dots, r\}$ obtained by LHS in $\mathcal{Z}$;
  for $i = 1, \dots, r$ **do**
    $(y^1)_i^n = f^1((z^1)_i^n, (y^2)_i^{n-1}))$;
    $(e^1)_i^n \leftarrow \frac{\|(y^1)_i^n) - \Phi^{-1}(\Phi^1((y^1)_i^n)))\|_2}{\|(y^1)_i^n)\|_2}$;
  end
  $e^1 \leftarrow \frac{1}{r}\sum_{i=1}^r (e^1)_i^n$;
  sample $\{(z^2)_i^n,\ i = 1, \dots, r\}$ obtained by LHS in $\mathcal{Z}$;
  for $i = 1, \dots, r$ **do**
    $(y^2)_i^n \leftarrow f^2((z^2)_i^n, (y^1)_i^n)$;
    $(e^2)_i^n \leftarrow \frac{\|(y^2)_i^n - \Phi^{-2}(\Phi^2((y^2)_i^n))\|_2}{\|(y^2)_i^n\|_2}$;
  end
  $e^2 \leftarrow \frac{1}{r}\sum_{i=1}^r (e^2)_i^n$;
  $\Phi^1 \leftarrow$ POD basis built from $\{(y^1)_i^j,\ i = 1, \dots, r,\ j = 1, \dots, n\}$;
  $\Phi^2 \leftarrow$ POD basis built from $\{(y^2)_i^j, i = 1, \dots, r,\ j = 1, \dots, n\}$;
  projection error $\leftarrow \max(e^1, e^2)$;
  $n \leftarrow n + 1$;
end
output: Φ^1, Φ^2, $\{(y^1)_i^j,\ i = 1, \dots, r,\ j = 1, \dots, n\}$, $\{(y^2)_i^j, i = 1, \dots, r,\ \ j = 0, \dots, n\}$, $\{(z^1)_i^j,\ i = 1, \dots, r,\ j = 1, \dots, n\}$, $\{(z^2)_i^j,\ i = 1, \dots, r,\ j = 1, \dots, n\}$

reduced disciplinary solver only depends on one coupling variable. The reduced MDA problem is shown in Figure 5 where the green dotted curve represents the response of the first discipline according to the coupling variable α_1^2 for a fixed parameter z_* and the purple dotted curve represents the response of the second discipline according to the coupling variable α_1^1. The solution of the MDA is the intersection between the two curves displayed with a red star.

From the training steps, a surrogate model of the MDA is initiated:

$$\begin{cases} \hat{\alpha}_i^1 = \mu_i^1(z_*, \hat{\alpha}_1^2, \dots, \hat{\alpha}_{n^2}^2) + \hat{\epsilon}_i^1(z_*, \hat{\alpha}_1^2, \dots, \hat{\alpha}_{n^2}^2)\ \forall i = 1, \dots, n^1 \\ \hat{\alpha}_j^2 = \mu_j^2(z_*, \hat{\alpha}_1^1, \dots, \hat{\alpha}_{n^1}^1) + \hat{\epsilon}_j^2(z_*, \hat{\alpha}_1^1, \dots, \hat{\alpha}_{n^1}^1)\ \forall j = 1, \dots, n^2 \end{cases} \tag{15}$$

where μ_i^1 and μ_j^2 are the mean values of the GPs. The terms $\hat{\epsilon}_i^1$ and $\hat{\epsilon}_j^2$ are random processes following a zero mean GP whose covariance function σ_i^1 or σ_j^2 is the one of the corresponding GP.

From the mean values of the disciplinary GPs, an approximation of the MDA can be computed:

$$\begin{cases} \bar{\alpha}_i^1 = \mu_i^1(z_*, \bar{\alpha}_1^2, \dots, \bar{\alpha}_{n^2}^2)\ \forall i = 1, \dots, n^1 \\ \bar{\alpha}_j^2 = \mu_j^2(z_*, \bar{\alpha}_1^1, \dots, \bar{\alpha}_{n^1}^1)\ \forall j = 1, \dots, n^2 \end{cases} \tag{16}$$

This mean solution is displayed with an orange star in Figure 6. The surrogate approximations of the reduced disciplinary solvers $\tilde{f}^1$ and $\tilde{f}^2$ are displayed with continuous lines while the exact reduced disciplinary solvers are displayed with dotted lines. This information does not manage to evaluate the uncertainty of the surrogate models, as this uncertainty is provided by the variances of the GPs. Then, even if the initial problem is deterministic, taking into account the variances of the GP brings randomness to the MDA. Thus, the first step will be to generate a sample of random solutions to evaluate the accuracy of the surrogate models. Then, an enrichment strategy will be proposed to enhance the accuracy of the surrogate models until its accuracy is equivalent to the one of the MDA solver.

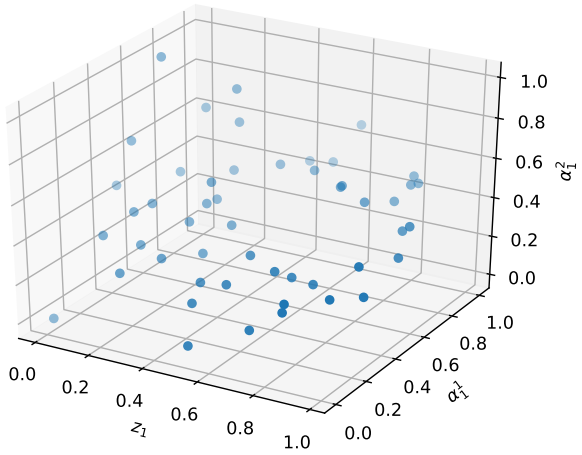

FIGURE 4 Distribution of the first coefficients α_1^1 and α_1^2 on the POD basis for both disciplines and one design variable given by the training algorithm samples involving 50 disciplinary solver solution.

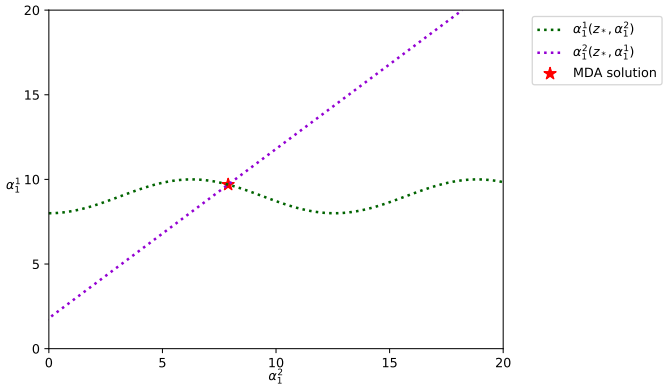

FIGURE 5 Illustration of the reduced MDA problem on a 1D POD basis. Reduced disciplinary solvers and MDA solution for a fixed parameter z_*.

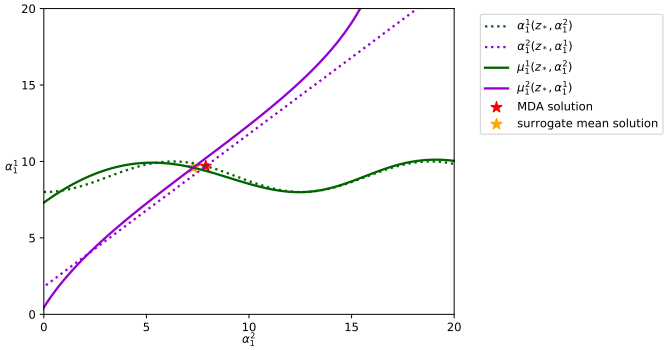

FIGURE 6 Illustration of the reduced MDA problem on a 1D POD basis. Reduced disciplinary solvers and their associated surrogate models with the solution of both the reduced and mean MDA.

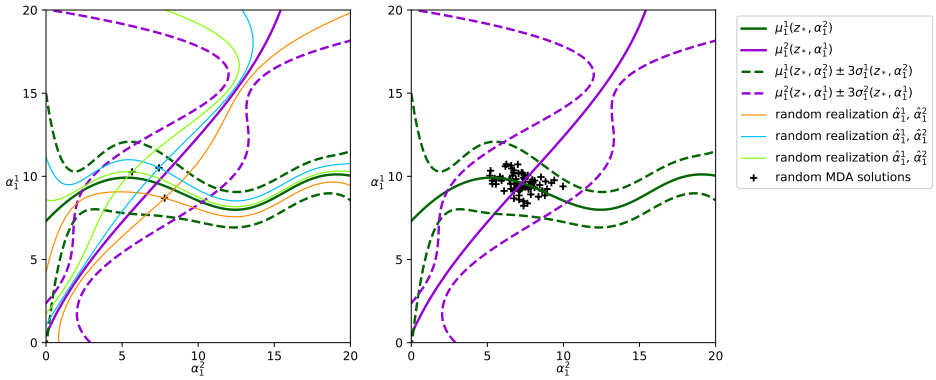

FIGURE 7 Illustration of the reduced MDA problem on a 1D POD basis. On the left, 3 random solutions obtained with simplified trajectories. On the right, a Monte Carlo simulation of 75 random solutions.

3.1 | How to generate random solutions considering the variance of the GPs?

One way to get random solutions of this system is to consider the following trajectories:

$$\begin{array}{ll} \hat{t}_i^1(z_*, \hat{\alpha}_1^2, \dots, \hat{\alpha}_{n^2}^2) & \text{trajectory of } \ \hat{\epsilon}_i^1(z_*, \hat{\alpha}_1^2, \dots, \hat{\alpha}_{n^2}^2) \ \forall i = 1, \dots, n^1 \\ \hat{t}_j^2(z_*, \hat{\alpha}_1^1, \dots, \hat{\alpha}_{n^1}^1) & \text{trajectory of } \ \hat{\epsilon}_j^2(z_*, \hat{\alpha}_1^1, \dots, \hat{\alpha}_{n^1}^1) \ \forall j = 1, \dots, n^2 \end{array}$$

Then, the following system is solved

$$\left\{ \begin{array}{l} \hat{\alpha}_i^1 = \mu_i^1(z_*, \hat{\alpha}_1^2, \dots, \hat{\alpha}_{n^2}^2) + \hat{t}_i^1(z_*, \hat{\alpha}_1^2, \dots, \hat{\alpha}_{n^2}^2) \ \forall i = 1, \dots, n^1 \\ \hat{\alpha}_j^2 = \mu_j^2(z_*, \hat{\alpha}_1^1, \dots, \hat{\alpha}_{n^1}^1) + \hat{t}_j^2(z_*, \hat{\alpha}_1^1, \dots, \hat{\alpha}_{n^1}^1) \ \forall j = 1, \dots, n^2 \end{array} \right. \tag{17}$$

This allows to generate random solutions that can be reached by the GP trajectories. Unfortunately, trajectories of GPs are difficult to simulate when the number of inputs of the GP (design variables and POD coefficients) is high leading to a representation with a large number of random variables which is not suitable for the following. One solution proposed in [11] is to use perfectly dependant GPs. The idea is to model the unknow functions as a conditionned GP with mean and variance obtained as described in Section 2.3.1 but with a correlation function constant and equal to one. This model is thus expressed as

$$\hat{\alpha}_i^1 = \mu_i^1(z_*, \hat{\alpha}_1^2, \dots, \hat{\alpha}_{n^2}^2) + \xi_i^1 \sigma_i^1(z_*, \hat{\alpha}_1^2, \dots, \hat{\alpha}_{n^2}^2), \ \forall i = 1, \dots, n^1$$

In the following, Figure 7 illustrates this idea.

Thus, we propose to solve the system:

$$\left\{ \begin{array}{l} \hat{\alpha}_i^1 = \mu_i^1(z_*, \hat{\alpha}_1^2, \dots, \hat{\alpha}_{n^2}^2) + \xi_i^1 \sigma_i^1(z_*, \hat{\alpha}_1^2, \dots, \hat{\alpha}_{n^2}^2) \ \forall i = 1, \dots, n^1 \\ \hat{\alpha}_j^2 = \mu_j^2(z_*, \hat{\alpha}_1^1, \dots, \hat{\alpha}_{n^1}^1) + \xi_j^2 \sigma_j^2(z_*, \hat{\alpha}_1^1, \dots, \hat{\alpha}_{n^1}^1) \ \forall j = 1, \dots, n^2 \end{array} \right. \tag{18}$$

Here, the solution depends on the draw ξ_i^1 and ξ_j^2 of the standard Gaussian variables $\hat{\xi}_i^1$ and $\hat{\xi}_j^2$. Thus, several random solutions of the surrogate model are given by different draws of $\hat{\xi}_i^1$ and $\hat{\xi}_j^2$. To characterize the variability of these random solutions a direct Monte Carlo (MC) is used. On the left of Figure 7, 3 random solutions are generated using these simplified trajectories. On the right, the same methodology is used to generate 75 MC simulations of the random solutions. In the next section, a method is proposed to deal with the random solutions given by the surrogate models.

Remarks:

- This simplification in the GP modelization is not detrimental in the evaluation of the interpolation uncertainty as this GP model shares the same variance as the original one, only the covariance is different. It has been shown in [11] that this choice of modelization is numerically efficient to quantify the uncertainty due to GP interpolation.

- Getting a random solution of such surrogate model is inexpensive. Indeed, for a given draw of $\hat{\xi}_i^1$ and $\hat{\xi}_j^2$, the solution of Eq. (18) only requires calls to the analytical means and variances of the GPs. This resolution is achieved using a non linear Jacobi solver.

3.2 | How to evaluate the accuracy of the surrogate model?

$$\begin{aligned}\hat{q}^1(\hat{\alpha}_1^1,\dots,\hat{\alpha}_{n^1}^1) &= \frac{\|\Phi^{-1}(\hat{\alpha}_1^1,\dots,\hat{\alpha}_{n^1}^1)-\Phi^{-1}(\bar{\alpha}_1^1,\dots,\bar{\alpha}_{n^1}^1)\|_2}{\|\Phi^{-1}(\bar{\alpha}_1^1,\dots,\bar{\alpha}_{n^1}^1)\|_2}\\ \hat{q}^2(\hat{\alpha}_1^2,\dots,\hat{\alpha}_{n^2}^2) &= \frac{\|\Phi^{-2}(\hat{\alpha}_1^2,\dots,\hat{\alpha}_{n^2}^2)-\Phi^{-2}(\bar{\alpha}_1^2,\dots,\bar{\alpha}_{n^2}^2)\|_2}{\|\Phi^{-2}(\bar{\alpha}_1^2,\dots,\bar{\alpha}_{n^2}^2)\|_2}\end{aligned} \tag{19}$$

The surrogate model is considered accurate enough if the τ-quantiles q_τ^1 and q_τ^2 of $\hat{q}^1$ and $\hat{q}^2$ are lower than ϵ_q where ϵ_q is the tolerance of the resolution (that can be equal to ϵ_{MDA}) and τ can be typically chosen as 0.9. This criterion ensures that a τ proportion of the random MDA is close enough to the mean solution. To estimate the quantiles, n_{MC} MC simulations are used. This quantity is similar to the convergence criterion of the MDA solver where the relative error between the coupling variables over two successive iterations is evaluated.

3.3 | How to locally enrich the surrogate model?

If the criterion on the quantile is not fulfilled, a local enrichment of the surrogate model is proposed. One could enrich both disciplines but this solution is not optimal as only one of the surrogate disciplinary solvers could be inaccurate. Then, enriching only one discipline is of interest to reduce the number of calls to the exact disciplinary solvers. The less accurate discipline to be enriched is chosen according to the approximation of the following Sobol sensitivity indices[29]:

$$\begin{aligned}S^1 &= \frac{\mathrm{Var}(\mathbb{E}(\hat{q}|\hat{\xi}_1^1,\dots,\hat{\xi}_{n^1}^1))}{\mathrm{Var}(\mathbb{E}(\hat{q}))}\\ S^2 &= \frac{\mathrm{Var}(\mathbb{E}(\hat{q}|\hat{\xi}_1^2,\dots,\hat{\xi}_{n^2}^2))}{\mathrm{Var}(\mathbb{E}(\hat{q}))}\end{aligned} \tag{20}$$

with

$$\hat{q}(\hat{\xi}_1^1,\dots,\hat{\xi}_{n^1}^1,\hat{\xi}_1^2,\dots,\hat{\xi}_{n^2}^2) = \hat{q}^1(\hat{\alpha}_1^1,\dots,\hat{\alpha}_{n^1}^1) + \hat{q}^2(\hat{\alpha}_1^2,\dots,\hat{\alpha}_{n^2}^2) \tag{21}$$

where, $\hat{\alpha}_1^1,\dots,\hat{\alpha}_{n^1}^1$ and $\hat{\alpha}_1^2,\dots,\hat{\alpha}_{n^2}^2$ are solutions of the surrogate MDA from Eq. (18) with the corresponding Gaussian variables $\hat{\xi}_1^1,\dots,\hat{\xi}_{n^1}^1$ and $\hat{\xi}_1^2,\dots,\hat{\xi}_{n^2}^2$. Every sample of $\hat{q}$ needs a solution of the surrogate model. The Sobol indices are estimated using a Polynomial Chaos Expansion approximation.

Then, the discipline associated to the highest value of S^1 or S^2 is enriched. If the projection error on the newly computed solution is considered as too important, the POD basis is enriched. This step ensures that even if the POD basis was not accurate enough at the end of Algorithm 1, the POD basis can be enriched to maintain a small error between the reduced MDA and the exact one. This enrichment strategy is applied repeatedly until the criterion over the quantiles is fulfilled. In other words, the objective is to decrease the variability of the random solutions by some successive enrichments. This process is described in Algorithm 3.

1. The mean solution defined by Eq. (16) is computed. This is used as a reference for the computation of the variability of the random solutions and for the enrichment.

2. First, a sampling of standard Gaussian $(\xi_i^1)_k,\ i = 1,\dots,n^1$ and $(\xi_j^2)_k,\ j = 1,\dots,n^2$ is generated using a MC sampling method. Those samples define the simplified trajectories needed to solve the random MDA.

3. The surrogate MDA defined in Eq. (18) is solved for each sample $(\xi_i^1)_k,\ i = 1,\dots,n^1$ and $(\xi_j^2)_k,\ j = 1,\dots,n^2$. For each random solution, the distance to the mean solution is computed by Eq. (19).

4. If the τ-quantile of $\hat{q}$ defined in Eq. (21) is higher than ϵ_q, an enrichment is needed.

5. To decide which of the two surrogate models to enrich, the estimation of S^1 and S^2 defined in Eq. (20) is made using a PCE approximation. The higher Sobol sensitivity indice determines which discipline has to be enriched. A call to the corresponding discipline is carried out using the information given by the mean solution and the surrogate discipline is enriched accordingly.

6. If the relative projection error is greater than ϵ_{PE}, the POD basis is enriched using an on-the-fly method[30,31]. For example, if the first discipline is enriched, the column $\bar{y}^1 - \Phi^1(\Phi^{-1}(\bar{y}^1))$ is added to the POD basis ϕ^1 after orthonormalisation. The applications Φ^1 and Φ^{-1} are modified to suit the new basis.

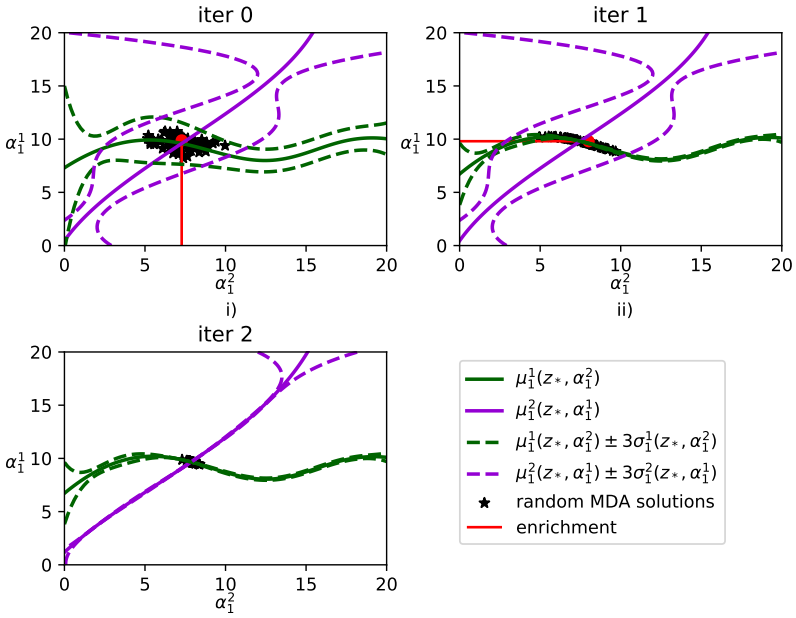

FIGURE 8 Illustration of the reduced MDA problem on a 1D POD basis. 3 iterations of the solving algorithm: i) initial GPs leading to first disciplinary surrogate model enrichment. ii) enriched GPs leading to second disciplinary surrogate model enrichment. iii) final GPs where no enrichment is required.

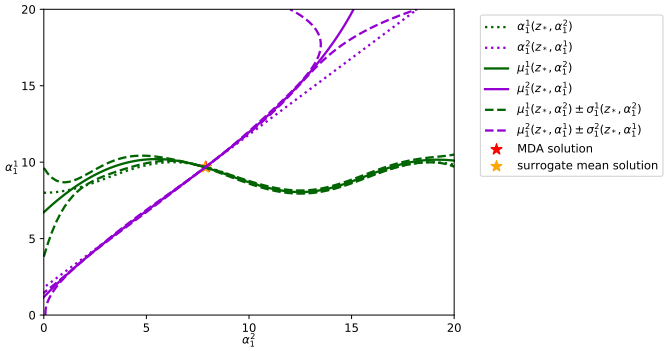

FIGURE 9 Illustration of the reduced MDA problem on a 1D POD basis. Reduced disciplinary solvers and their surrogate models after enrichment with the solution of the reduced and mean MDA where the surrogate models are only accurate close to the intersection.

If the variability is small enough, the surrogate solution is given by projecting back the mean solution to obtain $\tilde{y}^1(z_*)$ and $\tilde{y}^2(z_*)$. These steps of the algorithm are illustrated in Figure 8. The GP means and variances are displayed and MC simulations of the random solutions obtained are showed. In Figure 8 i), when no enrichment has been done, the random solutions are spread over a large area. The sensitivity analysis leads to the enrichment of the first disciplinary model reducing the variance of the corresponding GP in Figure 8 ii). Then, the second disciplinary model is enriched. In Figure 8 iii), the random solutions are concentrated near the mean solution of the surrogate model. Thus, the surrogate model is considered to be close enough to the exact model as illustrated in Figure 9 showing the exact and surrogate MDA solution.

Remarks:

- The enrichment of the POD basis requires to save all the previous solver solutions.

- Sometimes, the successive enrichment of the GPs around the reduced MDA solution leads to poorly trained GPs due to ill-conditioning. Then, the convergence of the random MDA solutions is not possible as no solution exists. In this case, the resolution is aborted and the exact MDA solution is used instead.

- It should be noted that the GPs do not have to be accurate in the whole coupling variable space but only where the MDA solution lies. This remark is illustrated in Figure 9, where one can see that the GPs are not very accurate near the boundaries of the domain, without this affecting the quality of the MDA solution.

Algorithm 3: Compute MDA solution using surrogate models.

input : z_*, n_{MC}
while τ-*quantile of* $\hat{q} > \epsilon_q$ **do**
  $\bar{\alpha}_i^1,\ i = 1,\dots,n^1, \bar{\alpha}_j^2,\ j = 1,\dots,n^2 \leftarrow$ solution of Eq. (16);
  $(\xi_i^1)_k,\ i = 1,\dots,n^1,\ k = 1,\dots,n_{MC} \leftarrow n^1 \times n_{MC}$ standard Gaussian samples.;
  $(\xi_j^2)_k,\ j = 1,\dots,n^2,\ k = 1,\dots,n_{MC} \leftarrow n^2 \times n_{MC}$ standard Gaussian samples.;
  for $k = 1,\dots,n_{MC}$ **do**
    $(\alpha_i^1)_k,\ i = 1,\dots,n^1, (\alpha_j^2)_k,\ j = 1,\dots,n^2 \leftarrow$ solution of Eq. (18) associated to $(\xi_i^1)_k,\ i = 1,\dots,n^1$ and $(\xi_j^2)_k,\ j = 1,\dots,n^2$;
    $q_k^1 \leftarrow \frac{\|\Phi^{-1}((\alpha_1^1)_k,\dots,(\alpha_{n^1}^1)_k)-\Phi^{-1}(\bar{\alpha}_1^1,\dots,\bar{\alpha}_{n^1}^1)\|_2}{\|\Phi^{-1}(\bar{\alpha}_1^1,\dots,\bar{\alpha}_{n^1}^1)\|_2}$;
    $q_k^2 \leftarrow \frac{\|\Phi^{-2}((\alpha_1^2)_k,\dots,(\alpha_{n^2}^2)_k)-\Phi^{-2}(\bar{\alpha}_1^2,\dots,\bar{\alpha}_{n^2}^2)\|_2}{\|\Phi^{-2}(\bar{\alpha}_1^2,\dots,\bar{\alpha}_{n^2}^2)\|_2}$;
    $q_k \leftarrow q_k^1 + q_k^2$;
  end
  $q_l^1 \leftarrow \tau$-quantile of $(q_k^1)_{k=1}^{n_{MC}}$;
  $q_l^2 \leftarrow \tau$-quantile of $(q_k^2)_{k=1}^{n_{MC}}$;
  if $q_l^1 > \epsilon_q$ *or* $q_l^2 > \epsilon_q$ **then**
    $\tilde{S}^1 \leftarrow$ estimation of S^1 by PCE;
    $\tilde{S}^2 \leftarrow$ estimation of S^2 by PCE;
    if $\tilde{S}^1 > \tilde{S}^2$ **then**
      Enrich the first disciplinary surrogate model;
      $\bar{y}^1 \leftarrow f^1(z^*, \Phi^2(\bar{\alpha}_1^2,\dots,\bar{\alpha}_{n^2}^2))$;
      if $\frac{\|\bar{y}^1-\Phi^1(\Phi^{-1}(\bar{y}^1))\|_2}{\|\bar{y}^1\|_2} > \epsilon_{PE}$ **then**
        $\phi^1 \leftarrow \{\phi^1, \frac{\bar{y}^1-\Phi^1(\Phi^{-1}(\bar{y}^1))}{\|\bar{y}^1-\Phi^1(\Phi^{-1}(\bar{y}^1))\|_2}\}$;
      end
    else
      Enrich the second disciplinary surrogate model;
      $\bar{y}^2 \leftarrow f^2(z^*, \Phi^1(\bar{\alpha}_1^1,\dots,\bar{\alpha}_{n^1}^1))$;
      if $\frac{\|\bar{y}^2-\Phi^2(\Phi^{-2}(\bar{y}^2))\|_2}{\|\bar{y}^2\|_2} > \epsilon_{PE}$ **then**
        $\phi^2 \leftarrow \{\phi^2, \frac{\bar{y}^2-\Phi^2(\Phi^{-2}(\bar{y}^2))}{\|\bar{y}^2-\Phi^2(\Phi^{-2}(\bar{y}^2))\|_2}\}$;
      end
    end
  end
end
output: $\tilde{y}^1(z_*) = \Phi^1(\bar{\alpha}_1^1,\dots,\bar{\alpha}_{n^1}^1)$, $\tilde{y}^2(z_*) = \Phi^2(\bar{\alpha}_1^2,\dots,\bar{\alpha}_{n^2}^2)$

3.4 | How to globally enrich the surrogate model?

In this section, the main objective is to add information from the local enrichment to the global surrogate model. The local model allows to get accurate MDA solution for a given design variable through local enrichment. The global model is the starting model of the local one before the enrichment steps. To lower the local enrichment cost, an enrichment of the initial model is proposed to enhance its accuracy on the design space after each local resolution.

An intuitive idea is to add all the information gathered during the local enrichment but this results in a poorly trained GPs and leads to stability issues (as explained in Section 2.4 concerning the initial DoE creation).

To avoid this issue, we filter the information and only add the last call to each disciplinary solver. Here, a choice has been made to not enrich the POD basis of the initial model. The POD basis is thus given by the initialization process and is only enriched during the local enrichment in Algorithm 3. This solution allows to handle local phenomenon through local POD basis enrichment without getting a large global POD basis affected by all the local singularities.

3.5 | POD+I

The classic POD+I aims at approximating directly the exact MDA solutions $y^1_*(z)$ and $y^2_*(z)$ as proposed in [19]. Then, to train the POD+I only exact MDA data shall be used. First, a DoE is needed. To do so, n_{DoE} different design variable points are obtained through a LHS: $\{z_i^{DoE},\ i = 1, \dots, n_{DoE}\}$. The MDA solutions $\{y^1_*(z_i^{DoE}),\ i = 1, \dots, n_{DoE}\}$ and $\{y^2_*(z_i^{DoE}),\ i = 1, \dots, n_{DoE}\}$ associated to each design variable is obtained using the MDA solver. Then, a POD basis is created for the first and the second disciplines. A DoE for the surrogate models of each POD coefficient is obtained by projecting each MDA solution on the POD basis. The surrogate models used are GPs:

$$\begin{cases} (\hat{\alpha}^1_*)_i(z) = \mu^1_i(z) + \hat{\epsilon}^1_i(z)\ \forall i = 1, \dots, n^1 \\ (\hat{\alpha}^2_*)_j(z) = \mu^2_j(z) + \hat{\epsilon}^2_j(z)\ \forall j = 1, \dots, n^2 \end{cases} \tag{22}$$

where $\hat{\epsilon}^1_i$ and $\hat{\epsilon}^2_j$ are zero mean GP whose covariance function is the one of corresponding kriging (i.e. σ^1_i and σ^2_j respectively). From this offline training, it is possible to compare the performance of the POD+I by analyzing the online approximation given by the mean values of the GPs.

The idea is now to develop a new POD+I solver that could be used as comparison for the DPOD+I based on an enrichment strategy derived from the enrichment strategy obtained in Section 3.3. This method is called enriched POD+I in the following. The enrichment algorithm described in Algorithm 4 follows the following steps:

- The mean values of the GPs give an approximation of the solution. It should be noted that the solution is easier to get as the POD+I is approximating the solution of the coupled system rather than each disciplinary solver.
- MC simulations are generated using the variance values of the GPs.
- A similar analysis is made on the variability of the random solutions around the mean solution.

4 | APPLICATION TO A STATIC AEROELASTIC PROBLEM

In the following, an application involving aerodynamic and structure disciplines is presented to illustrate the proposed approach. First, the physical problem is described. Then, the algorithms are tested and compared.

4.1 | Description of the MDA

The static aeroelastic problem is a common test case in multidisciplinary analysis. This study case involves two disciplines: the displacement of a wing is derived from the forces applied to the structure and the aerodynamic forces are defined by the wing shape and therefore by the displacement of the structure. This leads to the following non linear system:

$$\begin{cases} u = \mathcal{M}_s(z, f) \\ f = \mathcal{M}_a(z, u) \end{cases} \tag{23}$$

Algorithm 4: Compute MDA solution using POD+I.

input : z_*, n_{MC}
$(\bar{\alpha}^1_*)_i \leftarrow \mu^1_i(z),\ i = 1, \dots, n^1$;
$(\bar{\alpha}^2_*)_j \leftarrow \mu^2_j(z),\ j = 1, \dots, n^2$;
$(\xi^1_i)_k,\ i = 1, \dots, n^1,\ k = 1, \dots, n_{MC} \leftarrow n^1 \times n_{MC}$ standard Gaussian samples;
$(\xi^2_j)_k,\ j = 1, \dots, n^2,\ k = 1, \dots, n_{MC} \leftarrow n^2 \times n_{MC}$ standard Gaussian samples;
for $k = 1, \dots, n_{MC}$ **do**
  $((\alpha^1_*)_i)_k \leftarrow \mu^1_i(z) + (\xi^1_i)_k \sigma^1_i(z),\ i = 1, \dots, n^1$;
  $((\alpha^2_*)_j)_k \leftarrow \mu^2_j(z) + (\xi^2_j)_k \sigma^2_j(z),\ j = 1, \dots, n^2$;
  $q^1_k \leftarrow \frac{\|\Phi^{-1}(((\alpha^1_*)_1)_k, \dots, ((\alpha^1_*)_{n^1})_k) - \Phi^{-1}((\bar{\alpha}^1_*)_1, \dots, (\bar{\alpha}^1_*)_{n^1})\|_2}{\|\Phi^{-1}((\bar{\alpha}^1_*)_1, \dots, (\bar{\alpha}^1_*)_{n^1})\|_2}$;
  $q^2_k \leftarrow \frac{\|\Phi^{-2}(((\alpha^2_*)_1)_k, \dots, ((\alpha^2_*)_{n^2})_k) - \Phi^{-2}((\bar{\alpha}^2_*)_1, \dots, (\bar{\alpha}^2_*)_{n^2})\|_2}{\|\Phi^{-2}((\bar{\alpha}^2_*)_1, \dots, (\bar{\alpha}^2_*)_{n^2})\|_2}$;
  $q_k \leftarrow q^1_k + q^2_k$;
end
$q^1_l \leftarrow \tau$-quantile of $(q^1_k)_{k=1}^{n_{MC}}$;
$q^2_l \leftarrow \tau$-quantile of $(q^2_k)_{k=1}^{n_{MC}}$;
if $q^1_l > \epsilon_q$ *or* $q^2_l > \epsilon_q$ **then**
  Then, the surrogate model is not accurate enough and the exact MDA is computed. ;
  The GPs are enriched accordingly.;
  $y^1 \leftarrow y^1_*(z_*)$;
  $y^2 \leftarrow y^2_*(z_*)$;
else
  The surrogate model is accurate enough;
  $y^1 \leftarrow \Phi^1((\bar{\alpha}^1_*)_1, \dots, (\bar{\alpha}^1_*)_{n^1})$;
  $y^2 \leftarrow \Phi^2((\bar{\alpha}^2_*)_1, \dots, (\bar{\alpha}^2_*)_{n^2})$;
end
output: y^1, y^2

where u is the displacement vector of the wing, f is the aerodynamic loads vector, z is the design variables vector described later, $\mathcal{M}_a$ is a aerodynamic solver and $\mathcal{M}_s$ a structural one.

In this study, the wing structure is assumed linear elastic and the displacement governing equation is solved using a finite element method. The wing finite element model is composed of thin plate elements for the skins, the ribs and the spars and beam elements for the skins stringers. The geometry of the wing is defined by the undeflected Common Research Model (configuration uCRM-9)[32] whose mesh is presented in Figure 10. The structure of the wing is composed of 3 spars modeled with T-beam shape with plate elements. The wing skin is stiffened by 9 stringers on the upper and the lower skin modeled by beam elements (note that in Figure 10 the upper skin of the wing is removed for sake of clarity). The total number of degrees of freedom of the structure model is $d^1 = 43416$. Finally, the black box structural solver used is Code Aster[33].

The aerodynamic is described by the theory of potential flow solved using a Vortex Lattice Method[34]. The VLM models the lifting surfaces of the wing as an infinitely thin sheet of discrete vortices to compute lift and induced drag. The influence of the thickness and viscosity is neglected. This resolution leads to a linear system of dimension $d^2 = 2100$. The aerodynamic mesh of the wing is presented in Figure 11 and the resolution is ensured by an in house black box solver.

The transfer of the forces and displacements between the aerodynamic and structure mesh is made by radial basis interpolation[35]. An interpolation matrix $H \in \mathbb{R}^{d^2 \times d^1}$ is built using the radial basis interpolation. Then, the displacement on the aerodynamic mesh u_a is obtained by $u_a = Hu$ and the aerodynamic loads on the structure mesh f_s is obtained by $f_s = H^T f$ in order to enforce the conservation of the work (see[35] for more details). The MDA obtained is described in the Figure 12.

4.2 | Parametrization of the problem

Two slightly different parametric problems are now introduced. Both problems use the disciplinary solvers previously introduced but a different number of design variables. The first one counts 4 parameters described in Table 1 and the second counts 8

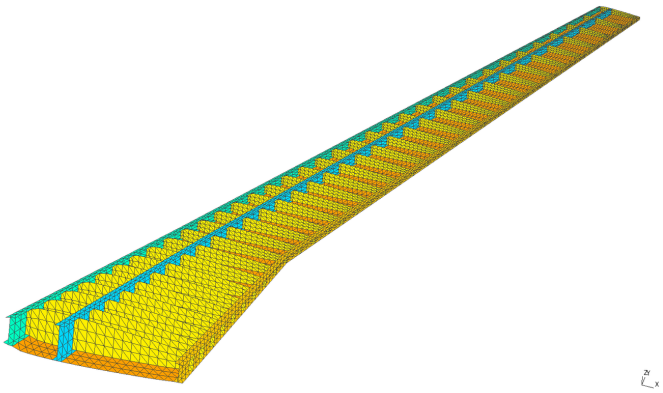

FIGURE 10 Structural mesh of the wing.

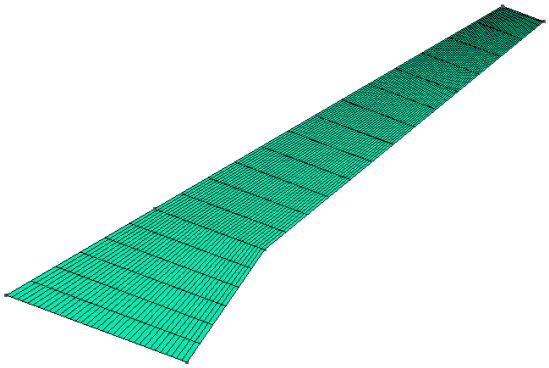

FIGURE 11 Aerodynamic mesh of the wing.

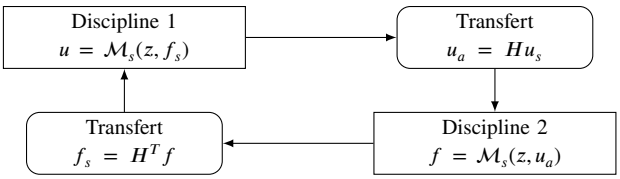

FIGURE 12 Illustration of the MDA on a aeroelastic problem described in Eq. (23) where the coupling variables are transferred from each mesh using a transfer matrix H.

parameters described in Table 2. One could note that the difference between the two problems concerns the number of parameters used to represent the thickness of the structural parts. In the 4 parameter problem the same thickness is applied to the 3 spars and the thickness of the lower and upper skins is also assumed equal. In the 8 parameter problem 4 new design variables are introduced to differentiate the thicknesses of these structural parts.

It should be noted that the design parameters are scaled to take values in $[0, 1]$ (see Table 1 and Table 2). However, the scaling values have been chosen so that the range of variation obtained with respect to the outputs of the MDA is large enough to assess the performance of the proposed method. As an example, the displacement of the wing tip obtained belongs to $[0.06, 3.95]$.

The tolerance of the MDA solver ϵ_{MDA} is set to 0.01. The starting point is given by the solution for the mean parameter. The reference for testing the method is made on a DoE containing 100 MDA solutions defined by 100 different design variables obtained through a LHS $\{z_i^{DoE},\ i = 1, \dots, 100\}$ and is denoted test set. The resolution of the MDA requires an average of 4.3 calls to each disciplinary solver for the 4 parameter problem and 3.9 for the 8 parameter problem.

variables	angle of incidence (AoI)	speed	thickness	
			spars	skins
ranges	[0,1]	[0,1]	[0,1]	[0,1]

TABLE 1 Scaled parameters of the 4 parameter problem.

variables	AoI	speed	thickness		
			leading/middle/training spars	upper/lower skins	ribs
ranges	[0,1]	[0,1]	[0,1]×[0,1]×[0,1]	[0,1]×[0,1]	[0,1]

TABLE 2 Scaled parameters of the 8 parameter problem.

4.3 | Setting parameters for the DPOD+I

The bounds λ^+ and λ^- described in Section 2.4.1 are given in this case by a physical analysis: the maximum displacement of the wing is assumed to lie between 0 and 3.6 meter. Those values are not the ones given in Section 4.1 but approximated values to ensure that accuracy on those bounds is not required. The tolerance on the mean projection error is set to the same threshold as the MDA solver: $\epsilon_{PE} = 0.01$. This parameter allows to manage the error between the exact MDA and the reduced MDA. The projection matrix is obtained by SVD using the scikit-learn library[36]. The percentage η of captured energy is set to 99.995%. The GPs are trained using the SMT library[37]. The correlation function used is a Matérn with $\nu = \frac{5}{2}$[27]. For the resolution of the MDA, two parameters have to be set : τ which defines the quantile retained and ϵ_q which is the threshold to be reached by this quantile. Here and for the rest of the study, τ is set to 0.9 and ϵ_q to 0.01. Those parameters allow to manage the error made between the reduced MDA and the surrogate MDA. To evaluate the quantile, the number n_{MC} of random MDA generated is set to 2000.

Remark:

The AoI and speed only concern the aerodynamic whereas the thicknesses only concern the structure. Consequently, for the 4 parameter problem, the number of inputs for the surrogate model of the aerodynamic solver is $n^1 + 2$ and $n^2 + 2$ for the surrogate model of the structural solver. For the 8 parameter problem, the number of inputs for the surrogate model of the aerodynamic solver is $n^1 + 2$ and $n^2 + 6$ for the surrogate model of the structural solver.

5 | 4 PARAMETER STATIC AEROELASTIC PROBLEM

5.1 | Resolution with the DPOD+I solver

First, a DPOD+I has been generated by setting r, the size of the initial sample of the coupling variables described in Section 2.4.1, to 10. The training algorithm ends after $n = 5$ iterations leading to 50 structure disciplinary calls and 60 aerodynamic disciplinary calls. The size of the POD basis for the structure is 6 and 5 for the aerodynamic. First, the quality of the POD basis obtained is analysed by mean of the relative errors between the solution of the exact MDA and its projection reconstruction:

$$\begin{aligned}(e^1_{DoE})_i &= \frac{\|y^1(z_i^{DoE}) - \Phi^{-1}(\Phi^1(y^1(z_i^{DoE})))\|_2}{\|y^1(z_i^{DoE})\|_2}\ ,\ i = 1, \dots, 100\\ (e^2_{DoE})_i &= \frac{\|y^2(z_i^{DoE}) - \Phi^{-2}(\Phi^2(y^2(z_i^{DoE})))\|_2}{\|y^2(z_i^{DoE})\|_2}\ ,\ i = 1, \dots, 100\end{aligned} \tag{24}$$

Note that the quality of the POD basis has to be analysed a posteriori on exact MDA as the set of vector used to build the POD basis does not contain any exact MDA. This analysis ensures the ability of the Algorithm 2 to create an accurate POD basis for exact MDA. Histograms of these quantities are presented in Figure 13. The relative error for the structure POD model lies between 0.16% and 3.8% with a mean of 0.7%. The relative error for the aerodynamic POD model lies between 0.03% and 0.7% with a mean of 0.16%. The POD approximation performs well for this type of disciplinary solvers and an accurate approximation with a reduced POD basis is obtained. Note that this result is important for the DPOD+I as the POD components are inputs of the reduced models that will be learned with GPs. Indeed, a large POD basis would complicate the learning of those reduced models.

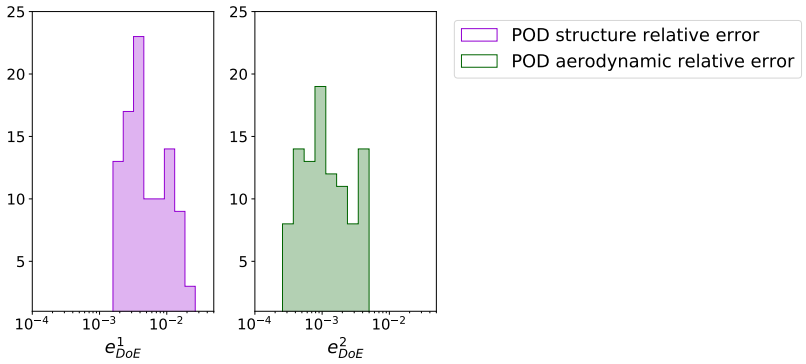

FIGURE 13 4 parameter problem, histogram of the relative error between the solution of the MDA and its projection reconstruction: e^1_{DoE} and e^2_{DoE} over the test set.

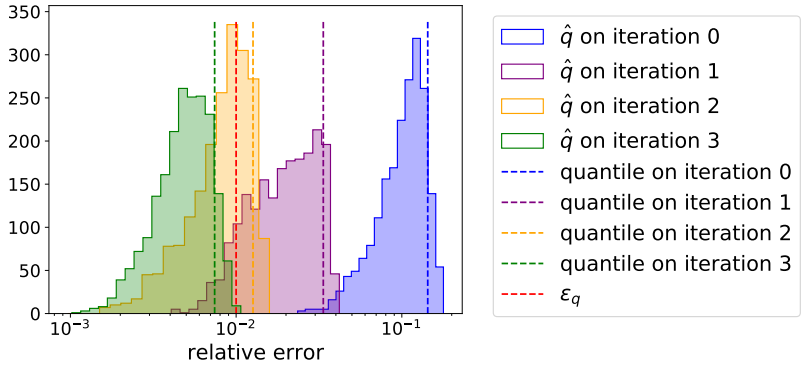

FIGURE 14 4 parameter problem, histogram and 9-th 10 quantile of $\hat{q}$ through the iterations of Algorithm 3 when solving the MDA for $z = z_1^{DoE}$ with the DPOD+I solver.

The information obtained during the training step is used to build the GPs as explained in Section 2.4.2. As a first illustration, the DPOD+I solver is used to solve the MDA at the first design point of the test set denoted by z_1^{DoE}. The histogram and τ-quantile of the quantity $\hat{q}$ defined by Eq. (19) on the 2000 random MDA solutions through the iterations are shown in the Figure 14. In this figure, the convergence of the τ-quantile of $\hat{q}$ towards the desired threshold ϵ_q is shown. The algorithm ends after 4 iterations which means 3 enrichment steps. Those enrichments were a structural one followed by an aerodynamic one and finally a structural one. This shows that the convergence to the MDA solution is faster than the usual 4 disciplinary solver calls for both disciplines with the exact MDA. Only the POD basis of the structure has been enhanced on the last iteration of the algorithm.

Then, the resolution for the 100 different parameters of the test set is achieved using Algorithm 3. The number of disciplinary calls to each solver needed for the resolution of the MDA is shown in Figure 15. The resolution of the MDA is obtained, in mean, with less than 1 disciplinary solver evaluation for each of the 100 design variables allowing to drastically reduce the computational cost. The total number of disciplinary solver calls needed for the training and the resolution is 99 for the structure and 65 for the aerodynamic.

Finally, an analysis on the solution given by the DPOD+I solver is proposed. To do so, an histogram of the relative error between the exact MDA and the solution given by the DPOD+I is given in Figure 16. The relative error for the structure lies between 0.25% and 4.8% with a mean of 1.1%. The relative error for the aerodynamic lies between 0.07% and 2.1% with a mean of 0.46%. First, the mean of the error is close to the target error which was 1%. Then, it appears that some MDA solutions are harder to approximate as their relative error is greater than 2%, this is potentially due to the error between the MDA solution projected on the POD basis and the solution that is given by the resolution of the MDA using POD projection.

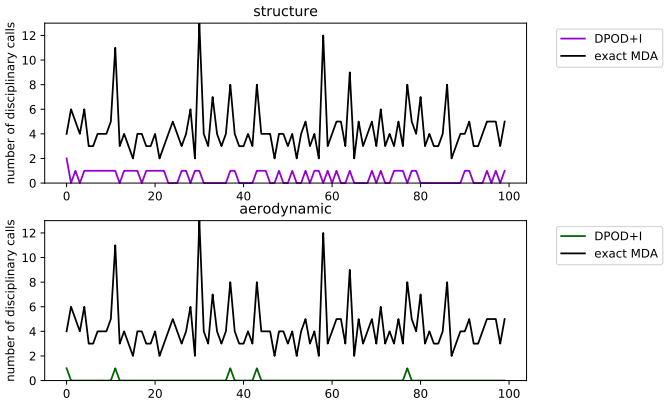

FIGURE 15 4 parameter problem, number of disciplinary solver calls needed for the resolution of the test set using the DPOD+I solver.

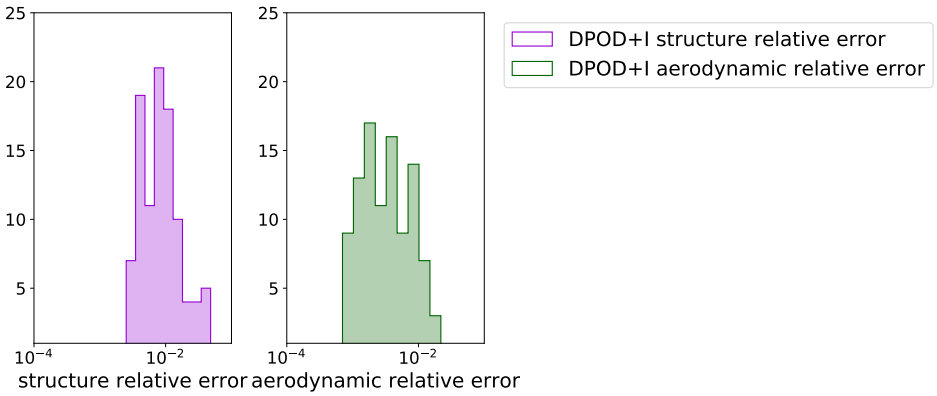

FIGURE 16 4 parameter problem, histogram of the relative error between the exact MDA solution and the solution given by the DPOD+I solver over the test set.

5.2 | Comparison with classic POD+I

A comparison with a classic POD+I is now presented to compare the accuracy of both method when an equal number of data are used to train each model. We remind that, in this context, the POD+I aims at approximating directly the exact MDA while the DPOD+I aims at creating a surrogate model of each disciplinary solver. The comparison is then made on their ability to predict the exact MDA.

From 25 exact MDA solutions, the POD basis and the GPs are trained. The number of exact MDA used (25) is chosen so that the number of disciplinary solver calls used for the DPOD+I method matches the number of disciplinary solver calls used for the resolution of the 25 exact MDA (used in the classic POD+I). Here, the 25 exact MDA resolutions involved 99 calls to each disciplinary solver (to be compared to the 50 + 49 calls to the structural solver used by DPOD+I method). Then, the relative error between the exact solution and the solution given by the classic POD+I is computed for the test set. The mean, minimum and maximum of the relative error obtained for the 100 design variables of the test set are displayed in Table 3 for the structure discipline and in Table 4 for the aerodynamic discipline. This study shows that the disciplinary surrogate models allow to get more accurate approximations. The mean error on the structure and the aerodynamic models have been reduced by a factor 3. The minimum of the relative error is lower for the DPOD+I strategy. The histogram of the relative error for both disciplines over the 100 MDA is displayed in Figure 17.

approach	min (%)	mean (%)	max (%)
DPOD+I	0.25	1.1	4.8
POD+I	0.4	3.5	28

TABLE 3 4 parameter problem, minimum, maximum and mean of the relative structure error for the DPOD+I and the classic POD+I over the test set.

approach	min (%)	mean (%)	max (%)
DPOD+I	0.07	0.46	2.2
POD+I	0.12	1.6	13

TABLE 4 4 parameter problem, minimum, maximum and mean of the relative aerodynamic error for the DPOD+I and the classic POD+I over the test set.

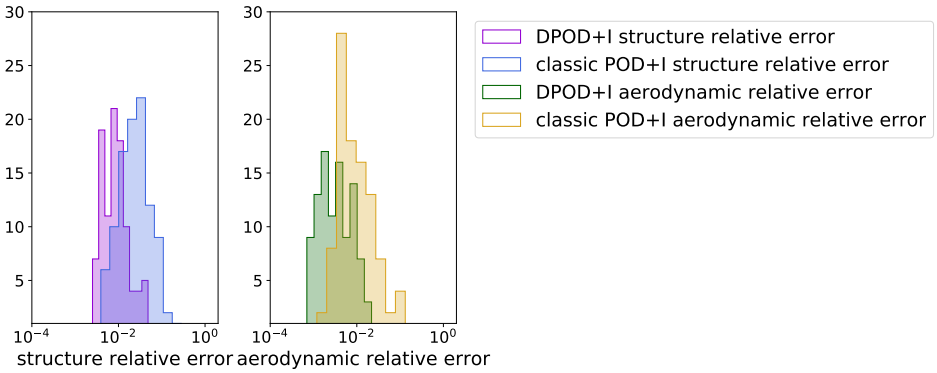

FIGURE 17 4 parameter problem, histogram of the relative error on the structure and the aerodynamic made on the test set using the classic POD+I and the DPOD+I approach.

5.3 | Comparison with enriched POD+I

The idea here is to compare the enrichment capacity for each method. Then, a similar enrichment criteria is used for both methods and they are compared on the budget needed for the enrichment. The initial POD+I approximation is given by the classic POD+I presented in Section 5.2 and is enriched using Algorithm 4.

The comparison is made on the number of disciplinary solver calls and it is presented in Table 5. For the 100 design variables, the MDA solver needed 429 calls to each disciplinary solver, the enriched POD+I allows to reduce slightly the number of disciplinary solver calls. On the other hand, DPOD+I solver reduces by three the number of structure solver calls and by five for aerodynamic solver calls compared to the enriched POD+I.

method	structure solver calls		aerodynamic solver calls		Total
	Offline	*Online*	*Offline*	*Online*	
exact MDA	0	429	0	429	858
DPOD+I	50	49	60	5	164
POD+I	99	251	99	251	700

TABLE 5 4 parameter problem, number of disciplinary solver calls needed for the resolution of the test set using the exact MDA solver, the enriched POD+I and the DPOD+I solver.

r	structure solver calls	CV (%)	aerodynamic solver calls	CV (%)
5	24.5	37	29.5	31
10	49	14	59	12
15	61.5	30	76.5	24
20	78	47	98	37

TABLE 6 4 parameter problem, influence of r on the computational cost of the training algorithm: mean and CV of the disciplinary solver calls over 10 realisations.

r	$\frac{1}{10}\sum_{j=1}^{10}(n^1)^j$	$\frac{1}{10}\sum_{j=1}^{10}(n^2)^j$	$\frac{1}{1000}\sum_{j=1}^{10}\sum_{i=1}^{100}(e^1_{DoE})_i^j$	$\frac{1}{1000}\sum_{j=1}^{10}\sum_{i=1}^{100}(e^2_{DoE})_i^j$
5	4.9	4.4	1.06%	0.23%
10	5.3	5.	0.74%	0.17%
15	5.3	5.	0.81%	0.17%
20	5.3	5.	0.78%	0.17%

TABLE 7 4 parameter problem, influence of r on the POD basis obtained with the training algorithm: mean number of POD modes and relative projection error on the test set over 10 realisations.

Finally, we have shown in this initial study that the DPOD+I is a promising method compared to the POD+I and performs well on this 4 parameter problem. However as the proposed method is probabilistic, a robustness study is performed in the following. Furthermore, the influence of the parameter r is also studied.

5.4 | Robustness of the training algorithm

In the following, it is proposed to study the influence of the number r of samples retained in Algorithm 1 and Algorithm 2. Although this parameter can be chosen by logistic constraints like the number of disciplinary solver calls that can be achieved in parallel, it influences the quality of the initial model. As the process is random, 10 realisations of the training algorithm are computed. The values tested for r are 5, 10, 15 and 20. Table 6 presents the mean number of disciplinary calls to achieve the convergence of the training steps for each different r value and their associated coefficient of variation (CV). First it appears that a high r value tends to increase the number of disciplinary calls for the training algorithm. Then, this r value should be kept as small as possible. However, if the POD basis obtained and the GPs are not trained enough, this results in more expensive enrichment phase. It should be noted that the value 10 is the best value in term of stability as it minimizes the variation of the disciplinary solver calls needed. With r set to 20, the variation is the highest due to the training phases that requires up to 200 calls for the structure solver and 210 for the aerodynamic one.

For each realisation j of the training algorithm, let $(n^1)^j$ and $(n^2)^j$ be the number of POD vectors retained for each discipline. The mean of $(n^1)^j$ and $(n^2)^j$ over the 10 realisations are displayed in Table 7. Then, the relative projection error $(e^1_{DoE})_i^j,\ i = 1,\dots,100,\ j = 1,\dots,10$ and $(e^2_{DoE})_i^j,\ i = 1,\dots,100,\ j = 1,\dots,10$ are computed for every exact MDA of the test set and for every realisation. Theirs means are exhibited in Table 7. It appears that the value of r does not influence the quality of the POD basis except for the low value $r = 5$. In that case the POD basis lacks some information to be accurate enough. The influence of this lack of initial accuracy is analysed next.

For each r value, 10 realisations of the training phase have been generated. Then, for each trained model, the resolution on the test set is achieved. For each realisation, the number of total calls of each disciplinary solver counting the calls from the training and the enrichment phases is computed. The mean and CV of the disciplinary solver calls over the 10 realisation are presented in Table 8. The total number of structure solver calls is only slightly impacted by the parameter r whereas the total number of aerodynamic solver calls clearly increases with the value of r. Indeed, it has been seen that, on this example, for the aerodynamic solver, the number of disciplinary calls for the enrichment phase is negligible compared to the training phase which is strongly influenced by the parameter r as shown by Table 6. Contrarily, for the structural solver, more points are added during the enrichment phase which smoothes the impact of the initial training phase.

Concerning the accuracy of the solution obtained by the surrogate models, the mean error over the test set is computed for each discipline and their mean and CV over the 10 realisations are displayed in Table 9. As expected, the surrogate models with

r	structure solver calls	CV (%)	aerodynamic solver calls	CV (%)
5	88.8	15	45.8	21
10	89.5	8	66	10
15	92.5	10	82.1	22
20	105.9	27	107.7	35

TABLE 8 4 parameter problem, influence of r on the total computational cost: mean number of calls and CV for each disciplinary solver after the training and the solving phases for the resolution of test set using DPOD+I over 10 realisations.

r	structure mean relative error (%)	CV(%)	aerodynamic mean relative error (%)	CV(%)
5	1.3	10	0.61	17
10	1.1	6	0.42	10
15	1.2	14	0.43	17
20	1.1	15	0.38	15

TABLE 9 4 parameter problem, influence of r on the accuracy of the DPOD+I: mean relative error for both disciplines and their associated CV on the test set using DPOD+I over 10 realisations.

r set to 5 are slightly less accurate but for the other ones, the difference remains negligible. For the coefficient of variation on the mean error made, the value 10 remains the more interesting one.

Finally, one of the remarks made on Section 3.3 concerned the fact that the random MDA does not converge when the surrogate models are poorly trained. With r set to 5, the number of failed resolution is around 1% which is acceptable. For the other ones, the fails are too rare to be noticed: less than 0.2% failed resolution.

To end this analysis, the mean number of disciplinary calls for each value of r are displayed in the Figure 18 as function of the number of MDA solution already solved. To analyse what should be the cost to solve further MDA with the DPOD+I solver, the mean number of disciplinary solver calls after 80 MDA solving is computed as it appears that this number grows almost linearly in Figure 18 after 80 MDA. Almost no further aerodynamic enrichment are needed: the surrogate model of the aerodynamic is accurate enough. For the structure, an average of 0.2 disciplinary solver calls is needed for the resolution of an MDA using DPOD+I solver. This reduces the computational cost of the MDA by a factor 20 compared to an iterative MDA solver.

6 | 8 PARAMETER STATIC AEROELASTIC PROBLEM

6.1 | Resolution with the DPOD+I solver

The same analysis is proposed for the 8 parameter problem in order to validate the result on a larger test case. The number r retained is still 10. The training algorithm ends after $n = 5$ iterations too which involves 50 structure disciplinary calls and 60 aerodynamic disciplinary calls. The size of the POD basis for the structure is 11 and 5 for the aerodynamic one. The number of POD modes for the aerodynamic remains constant but that number is increased for the structure. Indeed, the complexity of the structural output is increased due to a higher number of design variables attributed to the structure discipline. First, the quality of the POD basis obtained is analysed by means of the relative errors between the solution of the exact MDA and its projection reconstruction defined in Eq. (24). The histograms of these quantities are presented in Figure 19. The relative error for the structure POD model lies between 0.22% and 3.4% with a mean of 0.83%. The relative error for the aerodynamic POD model lies between 0.05% and 0.6% with a mean of 0.18%. The POD approximation is still accurate even with the raise of design variables.

After the training step, the DPOD+I method is illustrated on the first design point of the reference solution DoE z_1^{DoE}. The histogram and τ-quantile of the quantity $\hat{q}$ on the 2000 random MDA solution through the iterations are shown in Figure 20. In this figure, the convergence of the τ-quantile of $\hat{q}$ is showed until it is lower than ϵ_q. The algorithm ends after 3 iterations which means 2 enrichment steps. Those enrichments were a structure one followed by an aerodynamic one.

Then, the resolution for the 100 different design parameters is achieved using the DPOD+I solver. The number of disciplinary calls to each solver needed for the resolution of every MDA is shown in Figure 21. The resolution of the MDA with the DPOD+I requires less than 1 call to each disciplinary solver except one that required 2 structure solver calls. Even with a higher number

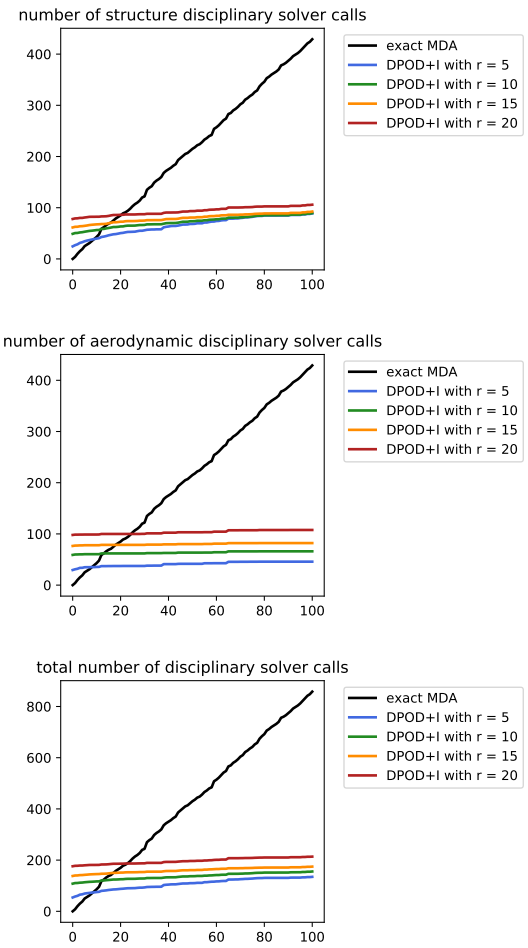

FIGURE 18 Application example, robustness study. Number of disciplinary calls through the resolution of the test set with r set to 5, 10, 15 and 20 and with exact MDA.

of design variables, the number of disciplinary calls per resolution remains low. The total number of disciplinary solver calls needed for the training and the resolutions is 128 for the structure and 66 for the aerodynamic.

To analyse the error made by the DPOD+I solver, the histogram of the relative error between the solution of the exact MDA and the solution given by the DPOD+I solver is presented in Figure 22. The relative error for the structure lies between 0.13% and 3.4% with a mean of 1.1%. The relative error for the aerodynamic lies between 0.05% and 2.2% with a mean of 0.55%. The mean of the error is still close to the target error of 1%.

6.2 | Comparison with classic POD+I

The same methodology is applied to compare with the classic POD+I. In the 8 parameter problem, 30 exact MDA are used as the total number of disciplinary solver calls needed for the DPOD+I on the 8 parameter problem is now 128. Here, the 30 exact MDA involves 121 disciplinary solver calls. The relative error between the exact solution and the solution given by the mean values of the GPs is computed for the test set. The mean, minimum and maximum of the relative error obtained for the 100 design variables of the test set are reported in Table 10 for the structure and in Table 11 for the aerodynamic. The mean error on the aerodynamic have been reduced by a factor 5 and by a factor 15 for the structure. The minimum of the relative error is lower for the DPOD+I strategy. This example shows the limits of the POD+I strategy when the number of design variables grows. Indeed, a higher number of data are needed to train accurate POD+I models. The DPOD+I is less sensitive to this phenomenon. The histogram of the relative error for both disciplines over the test set is displayed in Figure 23.

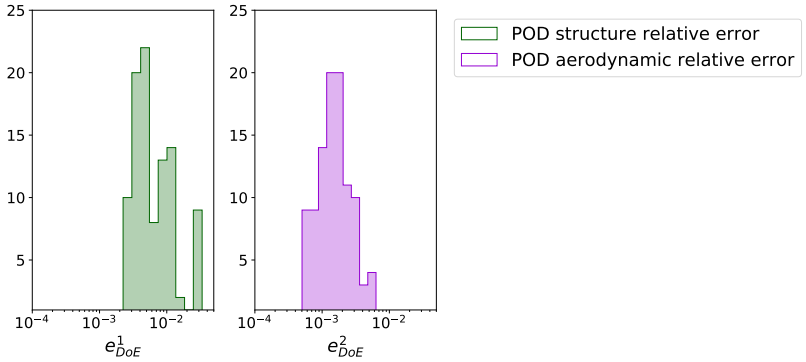

FIGURE 19 8 parameter problem, histogram of the relative error between the solution of the MDA and its projection reconstruction: e^1_{DoE} and e^2_{DoE} over the test set.

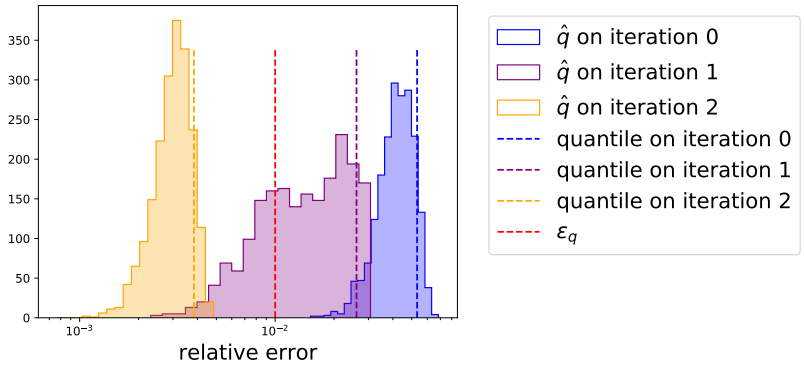

FIGURE 20 8 parameter problem, histogram and 9-th 10 quantile of $\hat{q}$ through the iterations of Algorithm 3 when solving the MDA with $z = z_1^{DoE}$ with the DPOD+I and exact MDA solver.

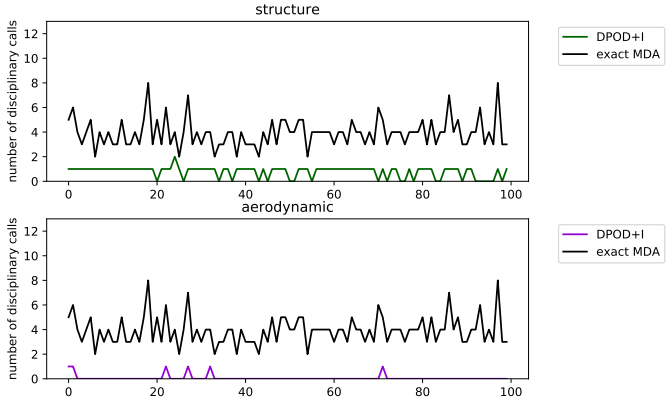

FIGURE 21 8 parameter problem, number of disciplinary solver calls needed for the resolution of the test set using the DPOD+I solver.

approach	min (%)	mean (%)	max (%)
disciplinary surrogate approach	0.13	1.1	3.4
POD+I	1.6	16	117

TABLE 10 8 parameter problem, minimum, maximum and mean of the relative structure error for the DPOD+I and the POD+I over the test set.

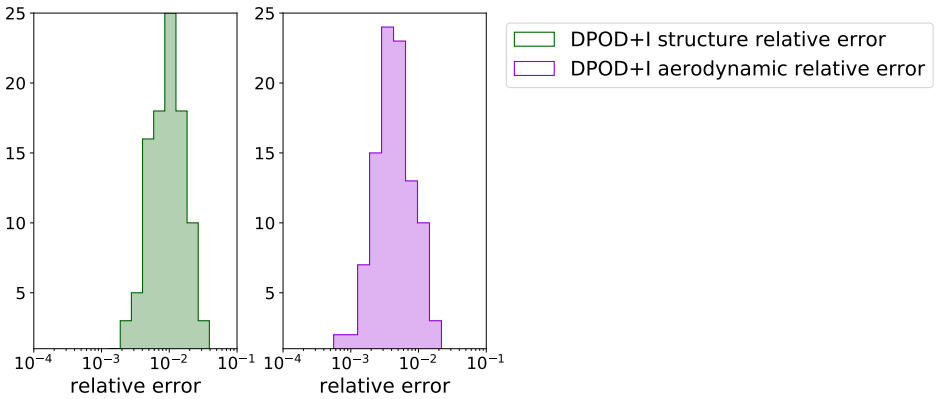

FIGURE 22 8 parameter problem, histogram of the relative error between the exact MDA solution and the solution given by the DPOD+I solver over the test set.

approach	min (%)	mean (%)	max (%)
disciplinary surrogate approach	0.05	0.55	2.2
POD+I	0.3	2.6	13

TABLE 11 8 parameter problem, minimum, maximum and mean of the relative aerodynamic error for the DPOD+I and the POD+I over the test set.

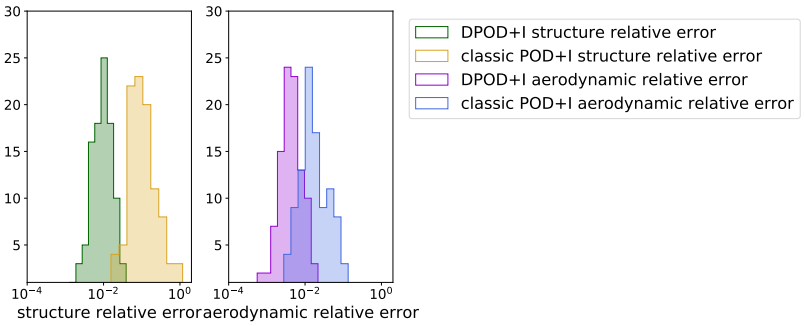

FIGURE 23 8 parameter problem, histogram of the relative error on the structure and the aerodynamic made on the test set using the POD+I and the DPOD+I approch.

6.3 | Comparison with enriched POD+I

The comparison is made on the number of disciplinary solver calls and is reported in Table 12 on a similar enrichment criterion. For the 100 design variables, the MDA solver needed 392 calls to each disciplinary solver, the POD+I solver does not allow to reduce the number of disciplinary solver calls. Indeed, there are only four times where the quantile criteria are fulfilled. Then, the total computational cost is higher due to the training phase. On the other hand, the DPOD+I solver reduces by 1.5 the number of structure solver calls and by six that of aerodynamic solver calls compared to the MDA solver. Finally, with this example involving a higher number of design variables, we have shown the limits of the POD+I and the POD+I solver compared to the DPOD+I. The accuracy is improved when using disciplinary surrogate model rather than global surrogate model. Indeed, when using POD+I strategy, each enrichment requires an exact MDA solution involving several calls to each disciplinary solver. This increases the computational cost of this method. However, the disciplinary surrogate models, with an adapted training algorithm, overpass this difficulty and show more efficiency to surrogate the MDA.

method	structure solver calls		aerodynamic solver calls		Total
	Offline	*Online*	*Offline*	*Online*	
exact MDA	0	392	0	392	784
Surrogate	50	78	60	6	194
POD+I	121	378	121	378	998

TABLE 12 8 parameter problem, number of disciplinary solver calls needed for the resolution of the test set using the exact MDA solver, the POD+I solver and the DPOD+I solver.

7 | CONCLUSION

This article presents a new adaptive surrogate based methodology to solve MDO problems when dealing with high dimensional coupling variables. The first challenge is the high dimensional coupling variables. It is proposed to use disciplinary POD approximations to reduce their dimensionalities which leads to a reduced MDA defined by problem (9). Then, the idea is to interpolate the reduced disciplinary solvers by GP surrogate models. This DPOD+I builds independent surrogate models for each disciplinary solver allowing to uncouple the MDA non linear problem, which can be advantageous in an industrial context. One of the challenges of this approach is to uncouple the training of the DPOD+I as the use of exact MDA for the training leads to poorly distributed DoE as shown in Figure 3. To do so, a new training strategy has been developed without initial knowledge on the ranges of the coupling variables. Finally, a new MDA solver is developed to solve the coupled problem by successive enrichments. This solver, includes an estimate of the error. If this error is too high, an enrichment of the discipline that has the most influence on the quantity of interest, according to Sobol sensitivity indices estimation, is achieved. Finally, some numerical applications confirm the interest of the proposed methodology. The new DPOP+I method is compared to POD+I. With a comparable number of disciplinary solver evaluations for the training, the DPOD+I outperforms the classic POD+I in term of accuracy. When the same enrichment strategy is applied to the POD+I methodology, the DPOD+I outperforms the enriched POD+I in term of disciplinary solver evaluations. Indeed, this new strategy allows the enrichment of each disciplinary surrogate model independently involving a reduction of the enrichment cost. One of the challenges is the determination of the parameters of the DPOD+I. The tolerance on the POD approximation ϵ_{PE} and the tolerance on the interpolation ϵ_q are related to the target tolerance of the solving of the MDA ϵ_{MDA}. The most difficult parameter to be set a prior is the training parameter r but we showed that its influence on the accuracy of the DPOD+I method is smoothed by the enrichment phase. Finally, the DPOD+I has shown good robustness when increasing the number of design variables contrary to POD+I. A perspective is to apply this strategy to more complex MDA problem involving non-linear disciplinary solver. However, one of the limits of this DPOD+I approach is the necessity of accurate approximation with a low number of POD modes. Indeed, the inputs of the surrogate models scales with the number of POD modes which can lead to difficulties when interpolating with GPs. As non-linear solver might be difficult to approximate with POD, dedicated non linear model order reduction could be used like ISOMAP[38].

References

1. Degroote J. Partitioned Simulation of Fluid-Structure Interaction. *Archives of Computational Methods in Engineering* 2013; 20(3): 185-238. doi: 10.1007/s11831-013-9085-5

2. Martins J, Lambe A. Multidisciplinary Design Optimization: A Survey of Architectures. *AIAA Journal* 2013; 51: 2049-2075. doi: 10.2514/1.J051895

3. Lewis R, Shubin G, Cramer E, et al. Problem Formulation for Multidisciplinary Optimization. *SIAM Journal on Optimization* 1997; 4. doi: 10.1137/0804044

4. Jones DR, Schonlau M, Welch WJ. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization* 1998; 13(4): 455–492. doi: 10.1023/A:1008306431147

5. Bartoli N, Lefebvre T, Dubreuil S, et al. *An adaptive optimization strategy based on mixture of experts for wing aerodynamic design optimization*; AIAA . 2017

6. Echard B, Gayton N, Lemaire M, Relun N. A combined Importance Sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. *Reliability Engineering and System Safety* 2013; 111: 232–240. doi: 10.1016/j.ress.2012.10.008

7. Paiva R, Crawford C, Suleman A. Comparison of Surrogate Models in a Multidisciplinary Optimization Framework for Wing Design. *AIAA Journal - AIAA J* 2010; 48: 995-1006. doi: 10.2514/1.45790

8. Zhang M, Gou W, Li L, Yang F, Yue Z. Multidisciplinary design and multi-objective optimization on guide fins of twin-web disk using Kriging surrogate model. *Structural and Multidisciplinary Optimization* 2017; 55(1): 361-373. doi: 10.1007/s00158-016-1488-0

9. Wang X, Li M, Liu Y, Sun W, Song X. Surrogate based multidisciplinary design optimization of lithium-ion battery thermal management system in electric vehicles. *Structural and Multidisciplinary Optimization* 2017; 56. doi: 10.1007/s00158-017-1733-1

10. Scholten W, Hartl D. Uncoupled method for static aeroelastic analysis. *Journal of Fluids and Structures* 2021; 101: 103221. doi: https://doi.org/10.1016/j.jfluidstructs.2021.103221

11. Dubreuil S, Bartoli N, Gogu C, Lefebvre T. Towards an efficient global multidisciplinary design optimization algorithm. *Structural and Multidisciplinary Optimization* 2020; 62(4): 1739-1765. doi: 10.1007/s00158-020-02514-6

12. Kerschen G, Golinval Jc, Vakakis AF, Bergman LA. The Method of Proper Orthogonal Decomposition for Dynamical Characterization and Order Reduction of Mechanical Systems: An Overview. *Nonlinear Dynamics* 2005; 41(1): 147-169. doi: 10.1007/s11071-005-2803-2

13. Paul-Dubois-Taine A, Amsallem D. An adaptive and efficient greedy procedure for the optimal training of parametric reduced-order models. *International Journal for Numerical Methods in Engineering* 2015; 102(5): 1262-1292. doi: https://doi.org/10.1002/nme.4759

14. Zahm O, Billaud-Friess M, Nouy A. Projection based model order reduction methods for the estimation of vector-valued variables of interest. *SIAM Journal on Scientific Computing* 2017; 39(4): A1647–A1674. doi: 10.1137/16M106385X

15. Amsallem D, Zahr M, Choi Y, Farhat C. Design optimization using hyper-reduced-order models. *Structural and Multidisciplinary Optimization* 2015; 51(4): 919-940. doi: 10.1007/s00158-014-1183-y

16. Chaturantabut S, Sorensen DC. Nonlinear Model Reduction via Discrete Empirical Interpolation. *SIAM Journal on Scientific Computing* 2010; 32(5): 2737-2764. doi: 10.1137/090766498

17. Farzam Far M, Martin F, Belahcen A, Montier L, Henneron T. Orthogonal Interpolation Method for Order Reduction of a Synchronous Machine Model. *IEEE Transactions on Magnetics* 2018; 54(2): 1-6. doi: 10.1109/TMAG.2017.2768328

18. Hesthaven J, Ubbiali S. Non-intrusive reduced order modeling of nonlinear problems using neural networks. *J. Comput. Phys.* 2018; 363: 55-78.

19. Coelho R, Breitkopf P, Knopf-Lenoir C, Villon P. Bi-level model reduction for coupled problems :Application to a 3D wing. *Structural and Multidisciplinary Optimization* 2010; 39. doi: 10.1007/s00158-008-0335-3

20. Nguyen N, Peraire J. Gaussian functional regression for output prediction: Model assimilation and experimental design. *Journal of Computational Physics* 2015; 309. doi: 10.1016/j.jcp.2015.12.035

21. Xiao M, Breitkopf P, Coelho R, Knopf-Lenoir C, Sidorkiewicz M, Villon P. Model reduction by CPOD and Kriging. *IntJStruc Multidisc Optim* 2009; 41: 555-574.

22. Brouwer KR, McNamara JJ. Generalized Treatment of Surface Deformation for High-Speed Computational Fluid Dynamic Surrogates. *AIAA Journal* 2020; 58(1): 329-340. doi: 10.2514/1.J058470

23. Filomeno Coelho R, Breitkopf P, Knopf-Lenoir C. Model reduction for multidisciplinary optimization - application to a 2D wing. *Structural and Multidisciplinary Optimization* 2008; 37(1): 29-48. doi: 10.1007/s00158-007-0212-5

24. Rozza G, Huynh DBP, Patera AT. Reduced Basis Approximation and a Posteriori Error Estimation for Affinely Parametrized Elliptic Coercive Partial Differential Equations. *Archives of Computational Methods in Engineering* 2008; 15(3): 229 - 275. doi: 10.1007/s11831-008-9019-9

25. Benner P, Gugercin S, Willcox K. A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems. *SIAM Review* 2015; 57(4): 483-531. doi: 10.1137/130932715

26. Kunisch K, Volkwein S. Control of the Burgers Equation by a Reduced-Order Approach Using Proper Orthogonal Decomposition. *J. Optim. Theory Appl.* 1999; 102(2): 345–371. doi: 10.1023/A:1021732508059

27. Rasmussen O, Luxburg vU, Rätsch G. *Gaussian Processes in Machine Learning*: 63–71; Berlin, Heidelberg: Springer Berlin Heidelberg . 2004

28. Mckay M, Beckman R, Conover W. A Comparison of Three Methods for Selecting Vales of Input Variables in the Analysis of Output From a Computer Code. *Technometrics* 1979; 21: 239-245. doi: 10.1080/00401706.1979.10489755

29. Sobol I. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* 2001; 55(1): 271-280. The Second IMACS Seminar on Monte Carlo Methodsdoi: https://doi.org/10.1016/S0378-4754(00)00270-6

30. Gogu C, Passieux JC. Efficient surrogate construction by combining response surface methodology and reduced order modeling. *Structural and Multidisciplinary Optimization* 2013; 47(6): 821-837. doi: 10.1007/s00158-012-0859-4

31. Gogu C. Improving the efficiency of large scale topology optimization through on-the-fly reduced order model construction. *International Journal for Numerical Methods in Engineering* 2014; 101. doi: 10.1002/nme.4797

32. Brooks TR, Kenway GK, Martins JRRA. *Undeflected Common Research Model (uCRM): An Aerostructural Model for the Study of High Aspect Ratio Transport Aircraft Wings*; AIAA . 2017

33. EDF . Finite element *code_aster*, Analysis of Structures and Thermomechanics for Studies and Research. Open source on www.code-aster.org; 1989–2017.

34. Katz J, Plottkin A. *Low-Speed Aerodynamics*. Press Syndicate of the Univerity of Cambridge . 2001.

35. Rendall TCS, Allen CB. Unified fluid–structure interpolation and mesh motion using radial basis functions. *International Journal for Numerics Methods in Engineering* 2007; 78: 1519-1559.

36. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 2011; 12: 2825–2830.

37. Bouhlel MA, Hwang JT, Bartoli N, Lafage R, Morlier J, Martins JRRA. A Python surrogate modeling framework with derivatives. *Advances in Engineering Software* 2019: 102662. doi: https://doi.org/10.1016/j.advengsoft.2019.03.005

38. Silva V, Tenenbaum J. Global Versus Local Methods in Nonlinear Dimensionality Reduction. *Adv Neural Inf Process Syst* 2003; 15.

8 | APPENDIX

Non linear Gauss Seidel solver

The solution of the MDA problem (1) is obtained iteratively via:

$$\begin{aligned} y^2_{i+1} &= f^2(z_0, y^1_i) \\ y^1_{i+1} &= f^1(z_0, y^2_{i+1}) \end{aligned} \tag{25}$$

where y^1_i and y^2_i are the ith approximation of the solution of the MDA. This first guess y^1_0 defines the initial point of the iterative solver and is usually obtained by the solutions of the MDA for the design variable z_0 corresponding to the center of the design

space $\mathcal{Z}$. The iterative algorithm ends when the relative errors between two iterates are low enough: $\frac{\|y^1_{i+1}-y^1_i\|_2}{\|y^1_{i+1}\|_2} \leq \epsilon_{MDA}$ and $\frac{\|y^2_{i+1}-y^2_i\|_2}{\|y^2_{i+1}\|_2} \leq \epsilon_{MDA}$ where ϵ_{MDA} is the tolerance on the MDA resolution fixed by the user.

How to cite this article: G. Berthelin, S. Dubreuil, M. Salaün, N. Bartoli, and C. Gogu (2021), Disciplinary Proper Orthogonal Decomposition and Interpolation for the resolution of parametrized Multidisciplinary Analysis