Multi-fidelity Methods for Improved Efficiency in Multi-disciplinary Optimization Problems

Margarida Cosme Gomes Pereira
margarida.cosme@tecnico.ulisboa.pt

Instituto Superior Técnico, Universidade de Lisboa, Portugal

January 2021

Abstract

In the pursuit of designing complex systems accurately and with affordable computational cost, multi-fidelity Super Efficient Global Optimization (SEGO) is one of the most recent approaches. Furthermore, various systems involve multiple disciplines interaction that must be considered during the optimization. These disciplines directly influence each other, hence information between them needs to be exchanged during optimization through transfer schemes. Since the transfer schemes are a significant part of the computational cost in the optimization process, new configurations can be used to try to reduce this time. Therefore, this work considers the extension of a multi-disciplinary optimization problem to multi-fidelity using SEGO. An aircraft wing is desired to be optimized with the aim of consuming the minimum fuel during the flight mission. The problem is defined using a low fidelity multi-disciplinary tool, OpenAeroStruct (OAS), that considers models for the aerodynamics and structures. In this work, we verify that the transfer schemes implemented in OAS do not fulfill the conservation requirement. Then, an extension of the transfer schemes is developed to enable different discipline discretization. Multi-fidelity SEGO with fidelity levels varied from one to three and two types of design of experiments is performed. We verify that the best approach is the one that uses two fidelity levels with only nested samples. Then, the transfer schemes extension is employed in the best approach of multi-fidelity SEGO for the same problem. We verify that very similar results are obtained with a computational cost reduction.

Keywords: Surrogate modeling, Bayesian optimization, Efficient Global Optimization, Aero-structural design

1. Introduction

When designing complex systems, such as an aircraft, there are various models of the system that can be used to simulate the system behavior. This models can have different degrees of accuracy, usually designated as fidelity levels. A model that accurately represents the system is a high fidelity model and the contrary is a low fidelity model. Unfortunately, designing complex systems with high fidelity models can be prohibitive and using low fidelity models may signify a big loss of accuracy. In this case, we may use a surrogate model, a model that mimics the behavior of the system as closely as possible while being computationally cheaper to evaluate. This model is constructed based on a dataset of limited intelligently chosen sample points. When using this approach, we can feed the surrogate model with samples from one fidelity level (single-fidelity) or we can combine models with different fidelity levels (multi-fidelity). Bayesian optimization [1] is an optimization that from the approximation model, determines what is the best sample to query. One particular algorithm is the Efficient Global Optimization (EGO) that uses a kriging model as approximation function.

Another important topic when designing systems is the correct consideration of the various disciplines that interact with each other [2]. An aircraft, for instance evolves structures, aerodynamics, control, propulsion among others. Traditionally, each of the disciplines are tackled individually by independent teams. However, the disciplines are coupled and a change in the aircraft's design affects more than one of the teams. This sequential approach produces non-optimal results, which can be overpassed by Multi-disciplinary Design Optimization (MDO). One of the most common applications of MDO in aircraft design is coupled aerodynamic and structural optimization due to the strong interaction between the two disciplines. Information between these two disciplines needs to be continuously exchanged during the optimization process through transfer schemes. This is a demanding task in computationally time terms. Therefore, the

way the transfer schemes are implemented, that is related with the discretization used for each discipline, greatly influences the computational time spent.

This work aims to explore the potentiality of an extension of a multi-disciplinary optimization problem to multi-fidelity. We study the performance of single and multi-fidelity optimization (fidelity levels varying from one to three) using EGO in an aircraft wing considering the aerodynamic and structural disciplines. To integrate the disciplines in the optimization, a low fidelity tool, OpenAeroStruct (OAS) [3], is used. OAS uses the same spanwise discipline discretization when modeling the system. We extend the OAS transfer schemes to handle non-identical spanwise discipline discretization. We implement the new transfer schemes to EGO and compare the results and computational time.

2. Background

2.1. Surrogate Modeling

A surrogate model or metamodel mimics the behavior of a function, constructed on a finite set of samples, the Design of Experiments (DOE), while being computationally cheaper. In general, a surrogate assumes that the outcome of interest $y(\mathbf{x})$ can be expressed as $\hat{y}(\mathbf{x}) + \epsilon$, where $\hat{y}(\mathbf{x})$ is the model and ϵ are the residuals, that are independent identically distributed. The kriging model assumes that the residuals are not independent, but rather a function of the samples location. This model takes the form $\hat{y}(\mathbf{x}) = m(\mathbf{x}) + Z(\mathbf{x})$, where $m(\mathbf{x})$ is the regression term and $Z(\mathbf{x})$ is the functional departure from the regression [4]. The regression term for the universal kriging is formulated as

$$m(\mathbf{x}) = \sum_{j=1}^{r} \beta_j f_j(\mathbf{x}), \tag{1}$$

where $f_j(\mathbf{x})$ are the basis functions and β_j are the corresponding coefficients. $Z(\mathbf{x})$ is a Gaussian process with covariance function expressed as

$$cov(Z(\mathbf{x}), Z(\mathbf{w})) = \sigma_z^2 R(\mathbf{x}, \mathbf{w}), \tag{2}$$

where σ_z^2 is the process variance, $\mathbf{x}$ and $\mathbf{w}$ are two design points in $\mathbb{R}^d$ and $R(\mathbf{x}, \mathbf{w})$ is the spatial correlation function, also known as the process kernel. The choice of the kernel determines how the metamodel fits the data. One of the most commonly used kernels is the squared exponential correlation kernel

$$R(\mathbf{x}, \mathbf{w}) = \exp(-\sum_{b=1}^{d} \theta_b (\mathbf{x}_b - \mathbf{w}_b)^2), \tag{3}$$

where $\boldsymbol{\theta} \in \mathbb{R}^d$ and $\boldsymbol{\theta} > 0$ is a vector of hyperparameters of the kriging model. The vector of hyperparameters is fitted using a set of sample data $X = \{x_1, ..., x_n\}$ with $x_i \in \mathbb{R}^d$, with observed responses $Y = \{y_1, ..., y_n\}$ with $y_i \in \mathbb{R}$. Once the kernel is defined, the mean and variance of the kriging model can be formulated as

$$\mu(\mathbf{x}) = f(\mathbf{x})'\boldsymbol{\beta} + r(\mathbf{x})' R^{-1}(Y - F\boldsymbol{\beta}) \tag{4}$$

and

$$\sigma^2(\mathbf{x}) = \sigma_z^2[1 - r(\mathbf{x})' R^{-1} r(\mathbf{x}) + (f(\mathbf{x})' - r(\mathbf{x})' R^{-1} F) (F' R^{-1} F)(f(\mathbf{x})' - r(\mathbf{x})' R^{-1} F)], \tag{5}$$

where $\mathbf{x} \in \mathbb{R}^d$ is the prediction point, R is the matrix of correlations among training points, $r(\mathbf{x})$ is the vector of correlations between the prediction point and the sample data, $\boldsymbol{\beta}$ is the vector of coefficients presented in equation (1), F is the matrix of values of the regression basis function at the positions of the training points and $f(\mathbf{x})$ is the vector of values of these functions at the prediction point. For multi-fidelity surrogate models, Kennedy and O'Hagan [5] proposed a formulation that links the high and low fidelity through a scaling factor ρ and a discrepancy function $\delta(\mathbf{x})$ as

$$f_{HF}(\mathbf{x}) = \rho f_{LF}(\mathbf{x}) + \delta(\mathbf{x}), \tag{6}$$

with $f_{LF}(\mathbf{x})$ and $\delta(\mathbf{x})$ independent. Le Gratiet [6] proposed a formulation inspired in [5] that takes the lower fidelity model as a basis function and thus reformulates equation (1) as

$$m(\mathbf{x}) = \sum_{j=1}^{r} \beta_j f_j(\mathbf{x}) + \beta_\rho f_{LF}(\mathbf{x}), \tag{7}$$

where β_ρ is an estimation of ρ performed by the likelihood maximization [7]. Based on equation (6), the mean and variance of a high fidelity model (only two fidelity levels) can be expressed respectively as

$$\mu_{HF} = \rho \mu_{LF} + \mu_\delta \tag{8}$$

$$\sigma_{HF}^2 = \rho^2 \sigma_{LF}^2 + \sigma_\delta^2. \tag{9}$$

This approach can be applied to l levels of fidelity. Let us denote $f_1, ..., f_l$ the hierarchically ranked fidelity levels, being f_1 the lowest and f_l the highest. Equations (8) and (9) can then be rewritten as

$$\mu_k = \rho_{k-1} \mu_{k-1} + \mu_{\delta_k} \tag{10}$$

$$\sigma_k^2 = \rho_{k-1}^2 \sigma_{k-1}^2 + \sigma_{\delta_k}^2. \tag{11}$$

Le Gratiet's formulation requires nested DOE, that is, $X_l \subseteq X_{l-1} ... \subseteq X_1$, where X_k represents the sample data to train the model of fidelity level 1 to l. We express the uncertainty contribution of the fidelity level k at design point $\mathbf{x}$ to the model (corrected from page 163, [6]) as

$$\sigma_{cont}^2(k, \mathbf{x}) = \sigma_{\delta_k}^2(x) \prod_{j=k}^{l-1} \rho_j^2. \tag{12}$$

Equation (12) shows that the uncertainty of the fidelity level k is computed by the variance contribution of the same level scaled using ρ_j until we get to the highest fidelity level.

2.2. Multi-fidelity Super Efficient Global Optimization (MFSEGO)

In 1975, Močkus [1] defined Bayesian optimization as an optimization technique based upon the minimization of the expected deviation from the extremum of the studied function. This optimization approach places a prior over the objective through a surrogate model which is constructed with an initial set of samples. The prior captures an estimated behavior of the function. This estimation is used to construct an acquisition function, often also referred as Infill Sampling Criterion (ISC), that determines what is the most appropriate query point to add.

Efficient Global Optimization (EGO) is a Bayesian optimization approach developed by Jones [8]. Let y be the objective function. We start by getting a set of samples $X = \{x_1, x_2, ..., x_n\}$ yielding the responses $Y = \{y_1, y_2, ..., y_n\}$. Using this set of samples, a kriging model is built with a mean and variance function μ and σ^2 respectively, as described in section 2.1. The information provided by the kriging model is used to establish the Expected Improvement function (EI), the employed ISC in EGO. This function will determine the next point to query. First, let

$$y_{\min} = \min\{y_1, y_2, ..., y_n\} \tag{13}$$

be the current best function value. The EI is simply given by

$$E[I(x)] = E[\max(y_{\min} - Y(\mathbf{x}), 0)], \tag{14}$$

where $Y(\mathbf{x})$ is a random variable following the distribution $\mathcal{N}(\mu(\mathbf{x}), \sigma^2(\mathbf{x}))$. By expressing the right-hand side of equation (14) as an integral, and applying some tedious integration by parts, the EI can be expressed in closed form as

$$\begin{aligned} E[I(\mathbf{x})] = (y_{\min} - \mu(\mathbf{x}))\Phi\Big(\frac{y_{\min} - \mu(\mathbf{x})}{\sigma(\mathbf{x})}\Big) \\ + \quad \sigma(\mathbf{x})\phi\Big(\frac{y_{\min} - \mu(\mathbf{x})}{\sigma(\mathbf{x})}\Big), \end{aligned} \tag{15}$$

where $\Phi(.)$ and $\phi(.)$ are respectively the cumulative distribution function and the probability density function of $\mathcal{N}(0, 1)$. Note that for equation (15) $E[I] = 0$ when $\sigma(\mathbf{x}) = 0$. Then, we determine the next point to be sampled as

$$\mathbf{x}_{n+1} = \underset{\mathbf{x}}{\operatorname{argmax}}\,(E[I(\mathbf{x})]), \tag{16}$$

and the respective response y_{n+1} is computed. The model is updated with the new information and this process is repeated for a certain number of times. Equation (15) is a balance between seeking promising areas of the design space (Exploitation) and choosing something from where we can better learn the design space (Exploration). This is the most important property of EGO.

More sophisticated ISC functions have been developed over the EI criterion, such as Watson and Barnes (WB2) criterion [9] formulated as

$$WB2(\mathbf{x}) = -\mu(\mathbf{x}) + E[I(\mathbf{x})], \tag{17}$$

that gives slightly more merit to local search instead of design space exploration.

Sasena [4] proposed an extension of the EGO algorithm to handle constrained problems: the Super Efficient Global Optimization (SEGO). SEGO constructs kriging models of the objective function and the n constraints. Each new point to sample is in the feasible domain, defined by the constraints kriging models.

SEGO can be extended to multi-fidelity problems. Following the approach presented by Meliani [10], MFSEGO algorithm handles with a two-stage decision process: (1) the search of the most promising sample and (2) the choice of the level of enrichment of that sample. The first problem is tackled by the ISC. The second problem analysis the uncertainty of the chosen point and decides the fidelity level where this point should be queried. Generally, low fidelity models are used to exploration and high fidelity models are used to exploitation.

Let $f_1, ..., f_l$ be the lowest to highest fidelity model of the objective function, with the querying costs $c_1, ..., c_l$. Recalling the recursive formulation of Le Gratiet, we known that fidelity levels relate as

$$f_k = \rho_{k-1}\, f_{k-1} + \delta_k \quad \text{for} \quad k \in \{1, ..., l\} \tag{18}$$

$$\rho_{k-1} = \operatorname{corr}(f_k, f_{k-1}) \frac{\operatorname{std}(f_k)}{\operatorname{std}(f_{k-1})} \tag{19}$$

$$\sigma_k^2 = \rho_{k-1}^2 \sigma_{k-1}^2 + \sigma_{\delta k}^2. \tag{20}$$

Recalling equation (12) and due to the necessity of nested DOEs (the fidelity levels lower than k of the sampling point $\mathbf{x}^*$ are enriched as well) the uncertainty reduction becomes

$$\sigma_{\text{red}}^2(k, \mathbf{x}^*) = \sum_{i=0}^{k} \sigma_{\delta i}^2(k, \mathbf{x}^*) \prod_{j=k}^{l-1} \rho_j^2. \tag{21}$$

The corresponding cost of the enrichment through fidelity level k yields

$$\text{cost}_{\text{total}}(k) = \sum_{i=0}^{k} c_i. \tag{22}$$

Mostafa [10] proposed the following enrichment level criterion

$$t = \underset{k\in(0,\dots,l)}{\text{argmax}} \quad \frac{\sigma^2_{\text{red}}(k, \mathbf{x}^*)}{\text{cost}_{\text{total}}(k)^2 + \text{offset}}, \tag{23}$$

where t is the highest fidelity level to be added and the offset is the scalar that takes into account the cost of building the surrogate model and maximization of the ISC.

2.3. Multi-disciplinary Design Analysis Optimization (MDAO) and Optimization Algorithms

The architecture of a MDAO system describes how a method handles the coupling of several disciplines. Figure 1 depicts the eXtended Design Structure Matrix (XDSM) diagram of Multi-Disciplinary Feasible (MDF) similar to the one employed at the core of this work.

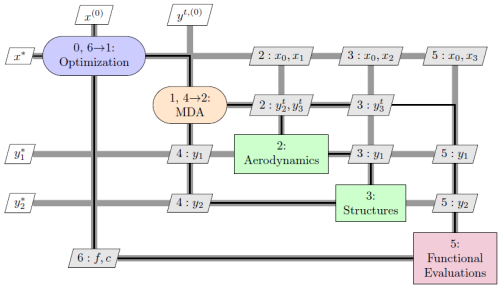


Figure 1: XDSM for a MDF architecture with a Gauss-Seidel MDA solver.

Following the sequential numeral flow in figure 1, for each optimization iteration the MDA solver iterates through all the discipline analysis until convergence guaranteeing a set of feasible coupling variables. Then, these variables are used to compute the objective function and constraints values and the next point is determined.

To determine the next point we can use gradient-free algorithms, such as EGO, or gradient-based algorithms. Gradient-based algorithms define the search direction from one iteration to the next based on evaluations of the objective function and its derivatives with respect to the design parameters. In this work, we use the Sequential Least-Square Quadratic Programming (SLSQP) [11] that simplifies the original problem, a non-linear problem, into a sequence of simpler quadratic sub-problems. This algorithm is used to maximize the WB2 criterion and to obtain the reference solution.

2.4. Aerodynamics Model

The vortex-lattice method (VLM), commonly used to model incompressible potential flow, is an extension of the lifting-line theory [12]. In this method, several lifting lines are superimposed along the chordwise direction to enable two-dimensional discretization. The wing is modeled as a set of panels, each of them containing a single horseshoe vortex with a control point. For each of the control points, a velocity $\mathbf{V}$ is induced by the discretized vortexes. These velocities can be obtained using the Biot-Savat law,

$$d\mathbf{V} = \frac{\Gamma}{4\pi}\frac{\mathbf{dl}\times\mathbf{r}}{||\mathbf{r^3}||}, \tag{24}$$

where Γ is the circulation of the vortex, $\mathbf{dl}$ moves along the vortex filament at hand and $\mathbf{r}$ is the distance from said filament to the control point. A system can then be assemble to compute the vortex velocities $\mathbf{u}_n$ through all the vortex circulations Γ_n.

$$\mathbf{u}_n = \sum_{j=1}^{N} A_{i,j}\Gamma_j, \tag{25}$$

where N represents the total number of vortexes defined and $A_{i,j}$ represents a row of the aerodynamic influence coefficient matrix. The influence coefficient $A_{i,j}$ represents the induced flow on panel i due to the vortex on panel j. From equation (25), one can obtains the velocity at control point n considering the circulations of the N vortexes.

2.5. Structures Model

We now define the structural model to obtain the wing stiffness. The finite element method (FEM) is employed. This method divides the lifting surface into smaller parts with the aim to simplify the model computations. In this work, a spatial beam element with 12 degrees of freedom (six at each node) is utilized, illustrated in figure 2.

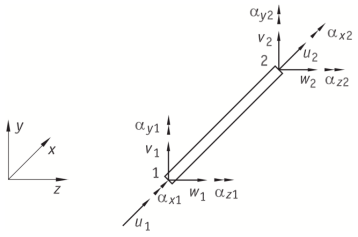


Figure 2: 12-DOF beam element.

The relevant equations are solved for each element, in a local reference frame, and then the results are transformed to a global reference frame, obtaining the global stiffness matrix $[K]$ of the entire structure. The aerodynamic analysis provides a set of loads that are applied to the lifting surfaces. Now, we can define the Hook's law used to obtain the wing stiffness

$$[K]\times\mathbf{u} = \mathbf{F}, \tag{26}$$

where $\mathbf{u}$ is the displacement/rotation vector and $\mathbf{F}$ is the force/moment vector. The vector of displacements and rotations produce a deformed updated

mesh. Then, we can iterate upon the architecture using the updated mesh on the aerodynamics analysis, obtaining a new set of loads.

2.6. Fluid-Structure Interaction

The previously described disciplines, aerodynamics and structures, are constantly exchanging information during the optimization process, such as aerodynamic pressures and structural displacements. Thus, a fluid-structure interaction is implemented.

The fluid-structure interaction must satisfy the requirements of consistency and conservation [2]. Consistency states that resultant forces and moments due to the pressure field, must be transferred into an equivalent set of nodal forces and moments, thereby satisfying the load conservation. However, there is an infinite number of sets of nodal forces that satisfy this requirement. The conservation requirement states that the virtual work performed by the load vector, over virtual displacements must be equal to the work performed by the pressure field, undergoing equivalent displacements [3].

3. Implementation

As referred, OpenAeroStruct is the low fidelity tool employed to model and analyzed the aerodynamic and structural disciplines. This tool models both disciplines with the same spanwise discretization as depicted in figure 3

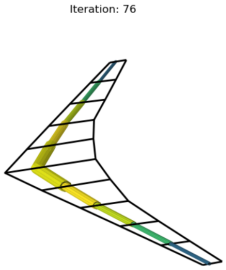
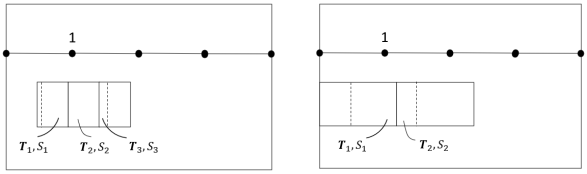


Figure 3: Aerodynamic grid and beam finite elements.

In this section, the developed transfer schemes to handle non-identical spanwise discretization are presented. The transfer schemes presented in [3] fulfill the consistency and conservation requirements, but the implemented ones do not fulfill the conservation requirement. We informed the MDO lab that developed OAS, however the extended schemes are based on the implemented ones, since the new version of OAS has not yet been released.

3.1. Load Transfer

The load transfer considers that the force and moment of a structural node is computed through the area bounded by half length of the left beam element to half length of the right beam element of that node. The content of that area can vary, from multiple sections of aerodynamic panels to just one panel section according to the applied discretization. Figure 4 illustrates the area corresponding to node 1 for two different discipline discretizations. For figure 4 (a), the aerodynamic discretization (AD) is more refined than the structural discretization (SD) and for figure 4 (b), the SD is more refined than the AD. The aerodynamic panels are represented by the continuous line and the area associated to node 1 is bounded by the dashed line. In figure 4 (a), node 1 has associated 3 panel slices and in figure 4 (b) only 2 panel slices. Each of these parts has an associated area S_i and a load per unit area $\mathbf{T}_i$.



(a) AD more refined than SD

(b) SD more refined than AD

Figure 4: Panels sections associated to node 1 to perform load transfer.

The equations to compute the force and moment of a structural node yields

$$\begin{cases} \mathbf{F}_s = \sum_{i=1}^{n} \mathbf{T}_i S_i \\ \mathbf{M}_s = \sum_{i=1}^{n} \mathbf{r}_{cp,i} \times \mathbf{T}_i S_i, \end{cases} \tag{27}$$

where $\mathbf{T}_i$ and S_i are, respectively, the load and area associated to panel slice i. The vector $\mathbf{r}_{cp,i}$ points from the point on the beam element aligned with the center of pressure to the aerodynamic center of pressure of the panel slice.

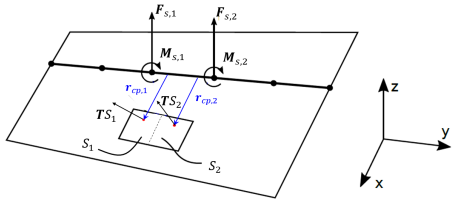


Figure 5: Scheme for transferring the load $\mathbf{T}$ to adjacent structural nodes for method 1 with same discipline discretization (adapted from [3]).

To illustrate this method, let us analyze how the load is transferred to the structural nodes using the same discipline discretization, depicted in figure 5.

The forces and moments at the nodes are defined as

$$\begin{cases} \mathbf{F}_{s,1} = \mathbf{T}S_1 \\ \mathbf{F}_{s,2} = \mathbf{T}S_2 \end{cases} \tag{28}$$

and

$$\begin{cases} \mathbf{M}_{s,1} = \mathbf{r}_{cp,1} \times \mathbf{T}S_1 \\ \mathbf{M}_{s,2} = \mathbf{r}_{cp,2} \times \mathbf{T}S_2. \end{cases} \tag{29}$$

3.2. Displacement Transfer

Let us now introduce the displacement transfer for non-identical discipline discretization. The edges (at constant y) of the panels are aligned with the structural nodes for the same discipline discretization. The displacement transfer algorithm compares the y location of the structural nodes with the aerodynamic mesh nodes. In case this y coordinate of the aerodynamic mesh nodes does not match any of the structural nodes, a linear regression of the displacement of the two surrounding structural nodes and the y coordinates is performed. Thus, we obtain displacements on a structural mesh with the same discretization as the aerodynamic mesh.

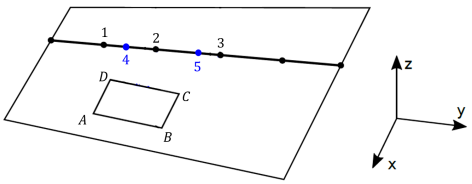


Figure 6: Scheme for transferring the displacement of the structural mesh to the aerodynamic mesh.

To clarify the scheme, let us introduce an example illustrated in figure 6. Since there are no structural nodes aligned with edge $[A, D]$, a linear interpolation with nodes 1 and 2 is computed to obtain the displacement of node 4 yielding

$$\mathbf{u}_{s,4} = \frac{\mathbf{u}_{s,2} - \mathbf{u}_{s,1}}{y_{s,2} - y_{s,1}} \cdot y_{s,4} + \mathbf{u}_{s,1} - y_{s,1}\left(\frac{\mathbf{u}_{s,2} - \mathbf{u}_{s,1}}{y_{s,2} - y_{s,1}}\right), \tag{30}$$

where $\mathbf{u}_{s,i}$ is the structural displacement associated with node i and $y_{s,i}$ is the y coordinate of node i. The same idea is applied in nodes 2 and 3 to produce node 5.

4. Results

4.1. Optimization Problem Definition and Initial Setup

Let us now define the optimization problem in table 1. The quantity we are looking to minimize is the fuel consumed using the Breguet range equation. The wing twist and spar thickness are controlled by a B-pline with 5 points each, so our problem has a total of 11 design variables. For the failure constraint, the aggregation method based on the Kreisselmeier-Steinhauser function [13] is employed.

Table 1: Optimization problem definition.

	Function/variable	Bounds
minimize	fuel consumed	
w.r.t.	angle of attack	[8, 12]°
	wing twist	[-6, 3]°
	spar thickness	[0.0015, 0.05] m
subject to	lift = weight	
	$\sigma_{\text{von Mises}} \leq \frac{\sigma_y}{2.5}$	

The mechanical properties of Aluminum-7075 are considered for the wing spar and additionally the following group of parameters.

Parameter	Value
Mach number	0.84
Altitude (m)	10.7×10^3
Air density (kg/m^3)	0.38
Range (m)	11×10^6
Empty weight (kg)	1.2×10^5
C_T $(kg/W/s)$	1.54×10^{-5}

Table 2: Cruise flight conditions and thrust specific fuel consumption (C_T).

Next, we present the employed numerical tolerances. The first two lines of table 3 are the stopping criterion for each design point, that is, the MDA solution. The last line of table 3 is a relative tolerance used to converge the optimization problem when using the SLSQP optimizer.

Parameter	Value
MDA absolute error tolerance	10^{-7}
MDA relative error tolerance	10^{-30}
Optimizer tolerance	10^{-3}

Table 3: Gauss-Seidel MDA and SLSQP optimizer numerical tolerances.

4.2. Fidelity Levels and Associated Cost

Once we will perform optimizations up to three fidelity levels, we need to define the mesh discretization and cost of each level. To define the discretizations, we use the mesh convergence study of the structural failure quantity shown in figure 7. The variables num_y and num_x in figure 7 are the spanwise and chordwise wing discretization, respectively. As the mesh is progressively refined, the

time to run the MDA increases and the model converges to -0.84. The lower the fidelity level of the model, the less accurate results it provides. Thus, we choose the less refined mesh of figure 7 to be the low fidelity mesh. As medium fidelity mesh, we choose the second less refined mesh. The second most refined mesh of figure 7 is set as high fidelity mesh, since it presents excellent convergence results.

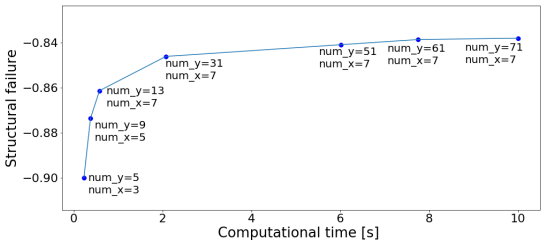


Figure 7: Structural failure mesh convergence study.

The discretization associated to each fidelity level is summarized in table 4.

Table 4: Discretization associated to high, medium and low fidelity levels.

Fidelity levels	High	Medium	Low
num_y	61	9	5
num_x	7	5	3

Normally, the cost ratios imposed to a sample of each fidelity level are linked to the computational time needed to run the analysis. Here, as we are using the OAS for all fidelity levels, the time difference to run the analysis between the fidelity levels is not very significant. Once we are testing an optimization algorithm, we impose the costs of each fidelity level to be more related with real problems. The cost of the high fidelity samples is always set to 1 and the other fidelity levels have smaller associated costs. The cost selected to the low fidelity samples is based on previous internal studies. To select the cost of the medium fidelity samples, a linear interpolation between the selected costs for high and low samples and the structural failure values was performed from figure 7. The cost ratios imposed to a sample of each fidelity level are summed up in table 5.

Table 5: Cost ratios associated to each fidelity level.

Fidelity cost	High	Medium	Low
Normalized cost	1	50/125	1/125

4.3. Design of Experiments Sampling Size

The sampling size of the employed DOEs are summarized in table 6. The DOEs are partitioned in two different groups - modified DOE and complete DOE - each of them compound by two DOEs of two and three fidelity levels (FL) optimization (2 F and 3 F). Additionally, there is the DOE corresponding to one fidelity level optimization (1 F).

Table 6: Design of experiments sampling size.

FL	Number of fidelities				
	1 F	Modified DOE		Complete DOE	
		2 F	3 F	2 F	3 F
L	-	40	40	80	160
M	-	-	40	-	80
H	40	40	40	40	40

Please note that high, medium and low had been shortened to H, M and L. The complete DOE group considers that the number of samples of the FL k doubles the number of samples of the most accurate consecutive FL $k+1$. Thus, half of the samples of each FL are equal to the samples of the most accurate consecutive FL (nested samples), and the other half of samples is extra information to the model. The DOE from the complete group of 3 F is the only one that is computed while the remaining four DOEs are obtained by filtering from it.

4.4. Multi-fidelity Parameters Definition

The optimization budget is a parameter to control the desired computational time to solve the problem and it is managed by summing progressively the associated cost of each point to be sampled during the optimization and stopping the optimization process when the budget has been reached. The chosen optimization budget is a trade-off between having acceptable computational time for several optimization runs and acceptable objective function results. The objective function results are considered to be acceptable when they are close or smaller than those of the reference solution presented in next section. After some experiments, the budget of 110 was defined. It is important to note that this budget does not include the cost associated to each DOE.

The Root Square Constraint Violation (RSCV) is used to evaluate the results in terms of constraint violation and it is given by

$$RSCV = \sqrt{C_1^2 + C_2^2}, \tag{31}$$

where C_1 is the equality constraint of lift=weight, and C_2 is the wing failure constraint. The common logarithmic log (RSCV) is employed during the results analysis for simplicity.

4.5. Reference Solution

For each of the DOE types, we run 100 run tests and filter them based on a reference solution. This solution, shown in table 7 along with the corresponding log (RSCV) value, is obtained using the SLSQP optimizer, the numerical tolerances presented in table 3 and the high fidelity mesh discretization. Our objective is, based on this reference solution, to define the minimum requirements that the solutions of the 100 run tests must satisfy. These limits are defined based on a trade-off relationship. On one hand a significant amount of run tests should fulfill the requirements to allow data statistical studies, on the other hand the filtered results should be close to the reference solution. After some tests, we chose the threshold presented in table 7.

Table 7: Reference solution and maximum acceptable values to filter the results.

Parameter	Value
Reference objective function	1.0680×10^5 Kg
log (RSCV)	-5.23
Maximum objective function	1.0787×10^5 Kg
Maximum log (RSCV)	-2.5

We are now ready to perform the filtering process. We are seeking for the best solution of each test. This solution is the one that has the lowest objective function (fuel consumption) value and fulfills the established requirements. Table 8 shows the number of optimizations for each DOE type that is conserved after the filtering process. The number of solutions that fulfilled the requirements has the same order of magnitude, except for the 3 F optimizations. Let us now observe the results of each optimization approach.

Table 8: Number of optimizations conserved after the filtering process.

1 F	Modified DOE		Complete DOE	
	2 F	3 F	2 F	3 F
74	72	32	78	43

4.6. One Fidelity Level Optimization

Figure 8 shows the scatter plot where each circle represents a best filtered solution associated with the cost and the fuel consumption value in the vertical and horizontal axis, respectively. Moreover, the color of the circle provides a qualitative representation of the log (RSCV) value of that solution. The more negative the log (RSCV) value is, the less the solution violates the constraints. Figure 8 depicts a cluster of circles in the upper left corner corresponding to the lowest values of fuel consumption associated with the highest costs.

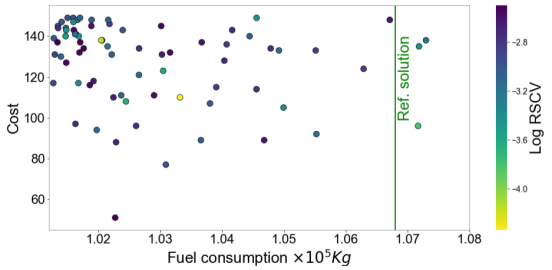


Figure 8: Scatter plot for the best solutions of 1 fidelity level optimization.

4.7. Two Fidelity Levels Optimization

Next, the acceptable result of optimizations with two fidelity levels using complete and modified DOE are shown and compared. Let us start by analyzing the scatter plots in figures 8 and 9.

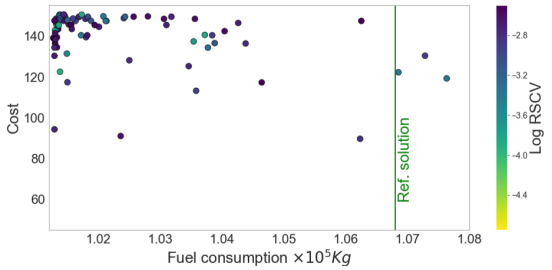


Figure 9: Scatter plot of the best solutions of 2 fidelity levels optimization using complete DOE.

Observing the log (RSCV) values, no clear conclusion can be taken between the log (RSCV) and the other two quantities. The first impression when observing both plots is the data clustering in the upper left corner of the 2 F complete DOE solutions in contrast to the more scattered solutions of 1 F.

The 2 F approach not only starts with a more informed model, but also usually starts with the exploration of the design space by querying low fidelity (LF) samples and then the exploitation phase, where mainly high fidelity (HF) samples are queried. In the first stage, the optimizer reduces the uncertainty of the ISC, allowing a more efficient decision when querying HF samples. The data clustering in the scatter plot shows that this approach manages most of the times to find low fuel consumption values with different initial DOE, providing robustness to the method.

Thereby, we can conclude that the 2 F complete DOE approach is the one that provides the best

solutions more often and so, we proceed the comparison between 2 F optimizations for complete and modified DOE using the same type of plots.

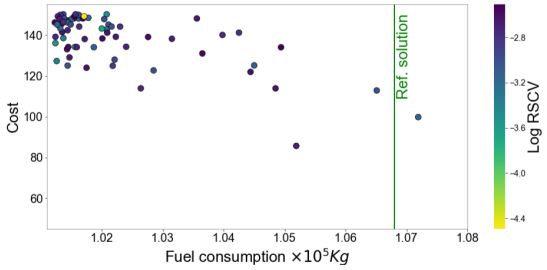


Figure 10: Scatter plot of the best solutions for 2 fidelity levels optimization using modified DOE.

Observing the scatter plots in figures 9 and 10, we start by noting that, again, it is difficult to detected any relation between the log (RSCV) values and the other two quantities. Both figures have a data clustering in the upper left corner of the distributions, that is, the lower fuel consumption values are associated with the maximum costs. We also observe that for the 2 F complete DOE there are more scattered solutions across the domain. Those scattered solutions are also more often associated with the maximum costs.

Thereby, the modified DOE approach is more likely to provide better solutions outcomes, which appears to be an odd tendency, once the problem is initiated with less information. In fact, the quality of the 40 non-nested LF samples included in the complete DOE data set can be so low that they are erratically informing the kriging model and inducing the optimizer into less promising areas of the design space where the budget is spent.

4.8. Three Fidelity Levels Optimization

Next, we present the acceptable result of the optimizations with three fidelity levels using complete and modified DOE using the same type of plots.

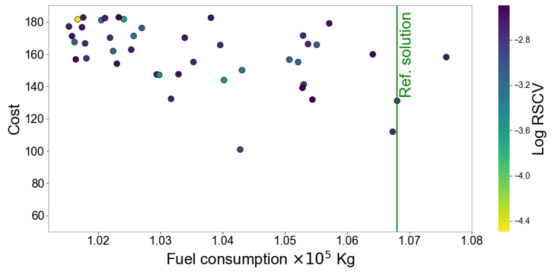


Figure 11: Scatter plot of the best solutions for 3 fidelity level optimization using complete DOE.

Observing figures 11 and 12, we conclude that the results of both approaches are less clustered and have higher costs than the previous 1 F and 2 F optimizations. The 3 F optimizations not only tend

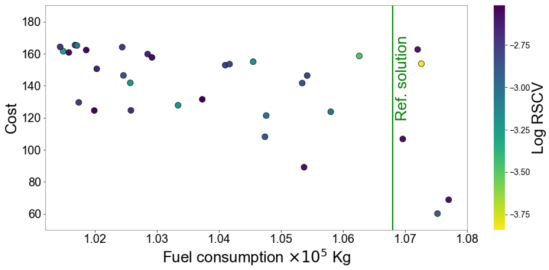


Figure 12: Scatter plot of the best solutions for 3 fidelity level optimization using modified DOE.

to provide worst results with higher costs, but also have more solutions that failed to fulfill the requirements, as noted in table 8. The extra fidelity level and the necessity of nested DOE increases the cost of HF samples, making the optimizer choosing less of them. This degrades the quality of the results.

4.9. Two Fidelity Levels Optimization with Non-identical Spanwise Discretization

Finally, let us introduce the optimization with the developed transfer schemes for non-identical discipline discretization. The 2 F modified DOE approach is employed. Table 9 shows the new samples discretization where num_s represents the structural discretization. The cost of the LF samples is set to 1/200.

Table 9: Discretization of high and low fidelity levels for non-identical discipline discretization.

FL	num_y	num_x	num_s
High	61	7	31
Low	5	3	3

Again, we run 100 run tests for this approach and we obtain 86 acceptable solutions. Observing figure 9, we see that this optimization provides very similar results to the 2 F approaches.

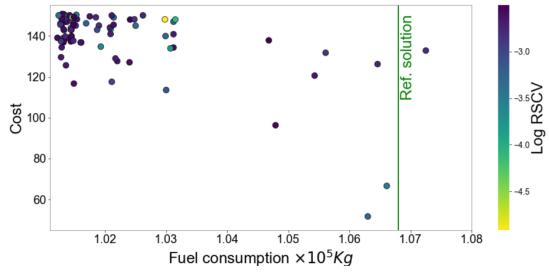


Figure 13: Scatter plot for the best solutions of 2 fidelity levels optimization using modified DOE for non-identical discipline discretization.

4.10. Summary and Overview

Next, we present the time to run 10 optimizations of each of the approaches. We conclude that higher

the number of fidelity levels used, higher the time needed. Thus, when choosing the number of fidelity levels, we must be aware that adding a fidelity level to the optimization implies a time penalty, so the decision of use an extra fidelity level must provide a significantly improvement of the results, as observed between 1 F and 2 F. Comparing the time of the 2 F approaches with the non-identical discipline discretization approach, we observe a slight reduction, as expected.

Table 10: Mean computational time to run 10 optimizations of each fidelity level approach.

1F	Modified DOE		Complete DOE		NI
	2 F	3 F	2 F	3 F	
0.63	1.70	3.15	1.67	3.27	1.60

Finally, the approaches can be organized by performance improvement as: 3 F modified DOE, 3 F Complete DOE, 1 F, 2 F complete DOE, 2 F modified DOE and 2 F modified DOE with NI discretization.

5. Conclusions

In this work, we extended the transfer schemes of OAS to non-identical discipline discretization. We performed a study to explore and compare the results of SEGO using a different number of fidelity levels, varying from one to three, and different DOE. We concluded that the best approach is the multi-fidelity optimization using 2 fidelity levels with only nested samples in the initial dataset, which proves the superiority of multi-fidelity optimization over single-fidelity optimization using SEGO. Then, the optimization approach that provided the best results was used to test the new transfer schemes. We selected new mesh discretizations and associated costs and performed again 100 run tests. We concluded that the results were very similar to the ones of 2 fidelity levels with a reduction of the computational time needed to run the optimization. Thereby, we can deduce that using a different spanwise discretization in the aerodynamic and structural disciplines is a viable solution to reduce the computational time of an optimization process while keeping the quality of the result.

Acknowledgements

The author would like to thank to ONERA for all the support and for making this project possible.

References

[1] Jonas Močkus. *On bayesian methods for seeking the extremum*, pages 400–404. Springer, Berlin, Heidelberg, 1975.

[2] Joaquim R Martins. *A coupled-adjoint method for high-fidelity aero-structural optimization*. PhD thesis, Stanford University, October 2002, USA.

[3] John Jasa, John Hwang, and Joaquim R Martins. Open-source coupled aerostructural optimization using python. *Structural and Multidisciplinary Optimization*, 02 2018.

[4] Michael J Sasena. *Flexibility and Efficiency Enhancements for Constrained Global Design Optimization with Kriging Approximations*. PhD thesis, University of Michigan, 2002.

[5] Marc C Kennedy and Anthony O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464.

[6] Loic Le Gratiet. *Multi-fidelity Gaussian process regression for computer experiments*. PhD thesis, University Paris-Diderot, 2013.

[7] Alexander J Forrester, András Sóbester, and Andy J Keane. *Engineering Design via Surrogate Modelling*, chapter 2, pages 51–54. John Wiley Sons, Ltd, 2008.

[8] Donald R Jones, Matthias Schonlau, and William J Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, 1998.

[9] Alan G Watson and Randal J Barnes. Infill sampling criteria to locate extremes. *Mathematical Geology*, 27(5):589–608, 1995.

[10] Mostafa Meliani, Nathalie Bartoli, Thierry Lefebvre, Mohamed-Amine Bouhlel, Joaquim R Martins, and Joseph Morlier. Multi-fidelity efficient global optimization: Methodology and application to airfoil shape design. In *AIAA Aviation 2019 Forum*, page 3236, 2019.

[11] Dieter Kraft. A software package for sequential quadratic programming. 1988. DFVLR Obersfaffeuhofen, Germany.

[12] John D Anderson. *Fundamentals of aerodynamics*. Tata McGraw-Hill Education, 2010.

[13] Joaquim R Martins and Nicholas M Poon. On structural optimization using constraint aggregation. In *VI World Congress on Structural and Multidisciplinary Optimization WCSMO6, Rio de Janeiro, Brasil*. Citeseer, 2005.