Multi-fidelity bayesian optimization using model-order reduction for viscoplastic structures

Stéphane Nachar, Pierre-Alain Boucard, David Néron, Christian Rey

Multi-fidelity bayesian optimization using model-order reduction for viscoplastic structures

Stéphane Nachar[a], Pierre-Alain Boucard[a], David Néron[a,*], Rey Christian[b]

[a] *LMT/ENS Paris-Saclay/CNRS/Université Paris-Saclay, 61 avenue du Président Wilson, 94230 Cachan, France*
[b] *SAFRAN Tech, CS 80112, 78772 Magny-les-Hameaux, France*

Abstract

The numerical optimization of a mechanical part requires a balance between computation time and model accuracy. The work presented herein aims at accelerate global optimization problem by using the framework of Bayesian optimization on a quantity of interest with multiple levels of fidelity. These multi-fidelity data are generated from a quality-driven model-order reduction framework: the LATIN Proper Generalized Decomposition. Within this framework, a reduced-order basis is generated on-the-fly and re-exploited to reduce the computational cost of observations. This strategy is tested on two elasto-viscoplastic test cases: a rocket damper and an aircraft blade and gives significant speedups.

Keywords: Bayesian optimization, Multi-fidelity kriging, Reduced-order models

Introduction

Structural optimization with high-fidelity computer experiments has a long tradition [1, 2, 3], using solver responses to build and optimize a quantity of interest (QoI) y for a given design space $\mathcal{D}$. There has been a growing interest in recent years in Bayesian optimization [4], which has proven to be effective on a number of difficult reference functions [5]. Bayesian optimization typically works by modeling the costly objective function y by a cheaper surrogate one $\hat{y}$ build as a Gaussian process regressor [6, 7, 8]. The uncertainly of the approximation is used to construct an acquisition function which estimates the probability of improving the global optimum obtained and quantify the exploitation–exploration trade-off. This acquisition function is maximized to determine the next query point. This approch try to reduce the number of calculated points.

This technique is commonly used in machine learning to optimize parameters of the model [9, 10] but, in a mechanical case, each resolution of the linked partial

*Corresponding author
Email address: david.neron@ens-paris-saclay.fr (David Néron)

differential equations (PDE) requires several hours of computation to obtain the converged value of the QoI. If the point is far from the accurate optimum zone, the QoI might have been computed with an unnecessary accuracy. A solution to reduce the duration of the optimization process is to use further data sources and generate a multi-fidelity metamodel. Low-fidelity but low-cost calculations are initially performed to estimate the optimal area. After a certain number of points, low-fidelity and high-fidelity information are both requested for the next query points to ensure the optimum obtained, and to estimate the optimal additive correction on the low-fidelity data previously computed. Further developments have already made on this subject [11, 12, 13] and have already been applied on structural optimization [14, 15].

Model reduction techniques constitute an efficient way to generate low-fidelity data by seeking the solution of the problem in a reduced-order basis (ROB), whose dimension is much lower than the original vector space. Using a model-order reduction database as a low-fidelity data is quite recent [16] and uses *a posteriori* approch which usually consists of defining this ROB by the decomposition of the solution of a surrogate model relevant to the initial model. *A priori* methods like Proper Generalized Decomposition (PGD) [17, 18] follow a different path by building progressively an approximation of the solution in a separated-variable form, without assuming any basis. This approch seems more appropriate without *a priori* information on the design space and was hitherto not tested in this multi-fidelity process.

The purpose of this paper is to show how to couple multi-fidelity kriging and model-order reduction to speedup the global optimization of the QoI. This strategy has already been performed for virtual charts generation [19]. Optimization is made here by calculating the QoI from the PGD modes generated on-the-fly with the LATIN-PGD framework [20] to deal with nonlinear problems. This framework generates the solution $\mathbf{s}$ from an initial solution $\mathbf{s}_0$ by computing at each step a spatio-temporal correction in a separated variable form. Since the method characterizes the solution over its entire domain at each iteration, it is possible to compute an approximation of the QoI before convergence. This property is particularly suitable for a global optimization process requiring an exploration phase in the design space. Low-fidelity fields are obtained by stopping the solver before convergence, and high-fidelity information is obtained with converged fields. In addition, the solver ability to reuse information from previously calculated PGD modes is exploited on a what we called "multiparametric strategy" [21, 22] and accelerate more and more computation as the process is moving forward. The coupling of these methods is done on viscoplastic test cases and an $8.4\times$ speedup is obtained, allowing a first demonstration of this approach.

The bayesian optimization strategy for multi-fidelity data is described on Section 1, with a focus on multi-fidelity kriging methods, a modified Evofusion method and a non-usual acquisition function which take into account multi-fidelity data. Observations used in this optimization process are computed thanks to the LATIN-PGD framework presented in Section 2. This framework is used in the context of viscoplastic problems and exploit the multiparametric

strategy to start computation from fields interpolated from previously calculated solutions. This multi-fidelity bayesian optimization process using the LATIN-PGD framework gives significant speedup and results of this strategy is visible on Section 3 with a presentation of the test cases.

1. Bayesian optimization strategy for multi-fidelity data

1.1. Bayesian optimization

Let us consider here a spatio-temporal mechanical problem defined by some input parameters $\underline{x}$ in a design space $\mathcal{D}$. For each value $\underline{x} \in \mathcal{D}$, the solution of the problem allows to compute a given quantity of interest (QoI), denoted $y(\underline{x}) \in \mathbb{R}$, which is assumed to be scalar. The associated objective function $\underline{x} \mapsto y(\underline{x})$ is considered as a black-box function without analytical expression of y nor its derivatives.

In the case of minimization, this problem is formulated concisely as follows:

$$\text{Find } (\underline{x}^*, \mathrm{y}^*) \in \mathcal{D} \times \mathbb{R}, \text{solution of } \underline{x}^* = \underset{\underline{x} \in \mathcal{D}}{\operatorname{argmin}}\, y(\underline{x}) \text{ and } \mathrm{y}^* = y(\underline{x}^*) \quad (1)$$

Global optimisation are without constraints here, but they can be taken into account with some adjustments on the strategy.

A non-exhaustive list of global optimization methods are presented in [23]. Standard gradient-based optimizers and heuristics are are particularly appreciated for their convergence properties but needs many observed points. Among them, Bayesian optimization is commonly used in the case of black-box expensive optimization [24, 5, 25]. The black-box function y is modeled herein as a Gaussian process $\hat{y}$ conditioned on observations, using gaussian process regression [8, 6, 7]. Initial observations are chosen from near-random sampling method like Latin Hypercube Sampling (LHS) [26] to prevent large areas without points. If needed, training set is sequentially enriched by the global maximum of an acquisition function which estimates the probability of improving the global optimum. This function is build from the posterior distribution of $\hat{y}$ and is a cheap function which can be globally optimized unlike y.

1.2. Gaussian Process Regression (Kriging)

N points $\underline{x}_i$ are observed with $\forall i \in [\![1, \mathrm{N}]\!]$, $\mathrm{y}_i = y(\underline{x}_i)$. The set of observed points is noted as $\boldsymbol{X}$ and their responses $\underline{\mathrm{y}}$. They are used to create the training set $(\boldsymbol{X}, \underline{\mathrm{y}})$. The objective is to build a regressor from this training set. In the case of linear regression methods, the output $y(\underline{x})$ can be mapped by a linear combination of the input $\hat{y}$:

$$\hat{y}(\underline{x}) = \underline{x}^T \underline{\beta} + Z(\underline{x}) \quad (2)$$

$$\forall i \in [\![1, \mathrm{N}]\!], \mathrm{y}_i = \underline{x}_i^T \underline{\beta} + Z_i \quad (3)$$

The term $\underline{x}^T \underline{\beta}$ represents the scalar product between the input vector and the weights associated $\underline{\beta}$. The term Z is the estimated residual error between the linear prediction and the true response.

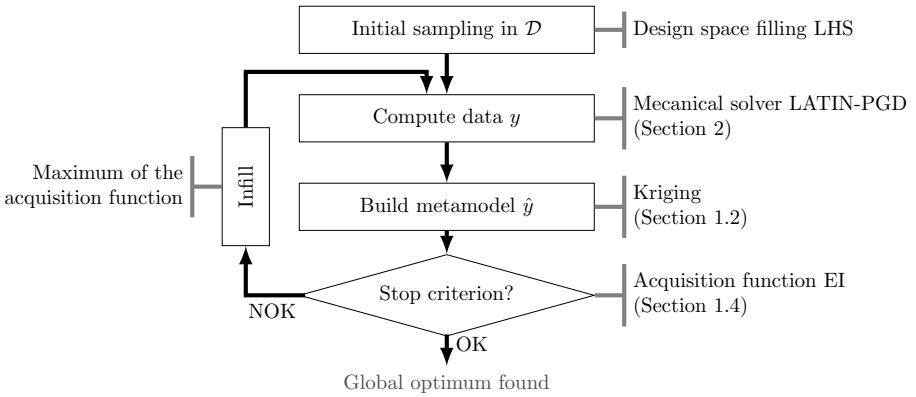


Figure 1: Classical bayesian optimization

Kriging [6, 7], also called Gaussian process regression [8], can be presented as an extension to linear regression and can also be understood as a form of Bayesian inference [27] by considering Z as a Gaussian process:

$$\hat{y}(\underline{x}) = \underline{f}(\underline{x})^{\mathrm{T}}\underline{\beta} + Z(\underline{x}) \tag{4}$$

The covariance between any two samples is defined by the covariance function (or kernel) $k(\underline{x}, \underline{x}'; \underline{\theta})$, (hyper)parametrized by $\underline{\theta}$. The Gaussian process $\hat{y}$ is conditional on observations, which means that by assuming the responses y_i are random variables, a joint distribution between the prediction and the observations is defined by the equation (5):

$$\begin{pmatrix} \hat{y}(\underline{x}) \\ \underline{y} \end{pmatrix} \sim N\left(\begin{pmatrix} \underline{f}(\underline{x})^{\mathrm{T}}\underline{\beta} \\ \mathbf{F}^{\mathrm{T}}\underline{\beta} \end{pmatrix}, \begin{pmatrix} \mathbf{1} & \underline{\mathrm{k}}^{\mathrm{T}}(\underline{x}) \\ \underline{\mathrm{k}}(\underline{x}) & \mathbf{C} \end{pmatrix} \right) \tag{5}$$

With $\mathbf{F}$ regression matrix such as $\mathbf{F}_{ij} = \underline{f}_i(\underline{x}_j)$, $\mathbf{C}$ covariance matrix between observed data such as $\mathbf{C} = k(\underline{x}_i, \underline{x}_j; \underline{\theta})$ and $\underline{\mathrm{k}}(\underline{x})$ covariance vector between observed and predicted data such as $\underline{\mathrm{k}}_j(\underline{x}) = k(\underline{x}_j, \underline{x}; \underline{\theta})$.

The predictive distribution is defined by $\left[\hat{y}(\underline{x}) | \hat{y}(\boldsymbol{X}) = \underline{y}, \underline{\beta}, \underline{\theta}\right]$ where the notation [A|B] stands for the distribution of A conditionally to B. So, conditionally to $\underline{\beta}$ and $\underline{\theta}$, the distribution $\underline{x} \longrightarrow \left[\hat{y}(\underline{x}) | \hat{y}(\boldsymbol{X}) = \underline{y}, \underline{\beta}, \underline{\theta}\right]$ is Gaussian $N\left(\mu(\underline{x}), \sigma^2(\underline{x})\right)$ with:

$$\begin{cases} \mu(\underline{x}) &= \underline{f}(\underline{x})^{\mathrm{T}}\underline{\beta} + \underline{\mathrm{k}}(\underline{x})^{\mathrm{T}}\mathbf{C}^{-1}(\underline{\mathrm{y}} - \mathbf{F}\underline{\beta}) \\ \sigma^2(\underline{x}) &= \left(k(\underline{x}, \underline{x})^2 - \underline{\mathrm{k}}(\underline{x})^{\mathrm{T}}\mathbf{C}^{-1}\underline{\mathrm{k}}(\underline{x})\right) \end{cases} \tag{6}$$

The kriging mean $\underline{x} \mapsto \mu(\underline{x})$ is the surrogate model that is used to approximate the objective function y and the kriging variance $\underline{x} \mapsto \sigma^2(\underline{x})$ represents the model mean squared error. The great interest of using kriging is the definition of the kriging variance as an error indicator of the metamodel on the whole

design space. A 95% confidence interval can be thus computed corresponding to $\mu(\underline{x}) \pm 3\sigma(\underline{x})$ as shown in figure 2.

Parameters $\underline{\beta}$ and hyperparameters $\underline{\theta}$ can be obtained by integrating the parameter posterior distributions, or by maximizing the likelihood function associated with the Gaussian process. Here, the Leave-One-Out cross-validation method is used. The idea behind is to generate n sub-training sets by extracting the observation $(\underline{x}_i, \mathrm{y}_i)$, $i \in [|1, n|]$ and using this one as a validation set to monitor performance. Hyperparameters are used to build metamodel liked with each sub-training sets and the LOO predictive log probability when leaving the training case i is computed :

$$\log p(y_i|\boldsymbol{X}, \underline{\mathrm{y}}_{-i}, \underline{\theta}) = \frac{1}{2}\log(\sigma_i^2) + \frac{(\mathrm{y}_i - \mu_i)}{\sigma_i^2}. \tag{7}$$

where notation $\underline{\mathrm{y}}_{-i}$ means all targets except number i, and μ_i and σ_i^2 are respectively the mean and the variance of the metamodel build without the training case i on $\underline{x}_i$. Accordingly, the LOO log predictive probability is :

$$L_{LOO} = \sum_{i=1}^{n} \log p(y_i|\boldsymbol{X}, \underline{\mathrm{y}}_{-i}, \underline{\theta}) \tag{8}$$

Objective is to minimize this mesure to obtain the optimal hyperparameters. It can be done by using bayesian optimization on hyperparameters [9].

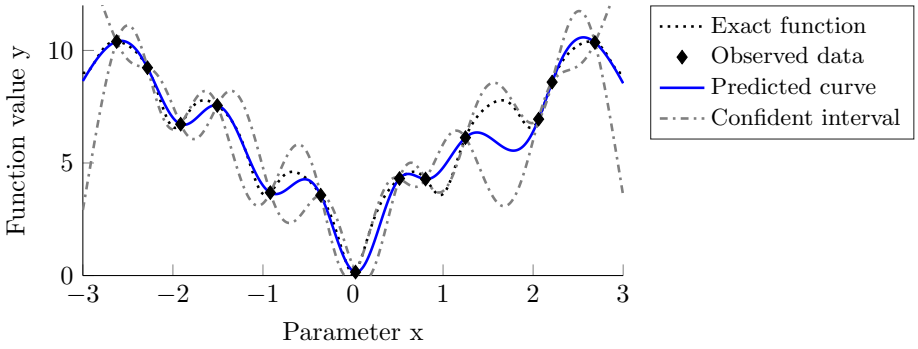


Figure 2: Kriging based on observations of the Ackley function
$(x, y) \mapsto -20\exp\left(-0.2\sqrt{0.5(x^2+y^2)}\right) - \exp\left(0.5\left(\cos 2\pi x + \cos 2\pi y\right)\right) + 20$

1.3. Data fusion with enhanced multi-fidelity kriging

A way to reduce the design step is to reuse this process for kriging by considering multiple solvers from the coarse one to the finest one and merge data observations. The reader can refer to [28] for a review of surrogate modeling and multi-fidelity approach.

Coming back to Equation (5), a possibility is to build joint distribution with multiple random observation vectors which include low-fidelity observations $(\boldsymbol{X}_1, \underline{\mathrm{y}}_1)$ and high-fidelity one $(\boldsymbol{X}_2, \underline{\mathrm{y}}_2)$. This approach is called "cokriging". Some hypotheses are needed for defining interactions between observations and can be found in [13, 29, 30, 31]. All these methods modify covariance matrix to take into account solver quality. So they allow expert judgement, but they are intrusive.

Other methods based on recursive metamodeling exist: Hierarchical Kriging [32] which replaces regression function $\underline{f}$ by a Gaussian process conditioned to low-fidelity data $\hat{y}_1$, and Evofusion [11]. Evofusion method is an simple one and has given good results in a multi-fidelity benchmark [33]. Details of Evofusion algorithm can be seen in Algorithm 1.

Algorithm 1: Evofusion Algorithm

Input: low-fidelity and high-fidelity observations $(\boldsymbol{X}_1, \underline{\mathrm{y}}_1)$, $(\boldsymbol{X}_2, \underline{\mathrm{y}}_2)$
Build a low-fidelity metamodel with only low-fidelity data
$(\boldsymbol{X}_1, \underline{\mathrm{y}}_1) \longrightarrow \hat{y}_1$
Compute the gap between the low-fidelity metamodel $\hat{y}_1$ and high-fidelity observations $\underline{\mathrm{y}}_2$ on high-fidelity points $\boldsymbol{X}_2$:
$y_{corr} = y_2 - \hat{y}_1(\boldsymbol{X}_2)$
Build a correction metamodel with correction data $(\boldsymbol{X}_2, \underline{\mathrm{y}}_{corr}) \longrightarrow \hat{y}_{corr}$
Modify low-fidelity observations with the correction metamodel:
$\underline{\mathrm{y}}_{1c} = \underline{\mathrm{y}}_1 + \hat{y}_{corr}(\boldsymbol{X}_1)$
Build the fused metamodel with corrected data and high-fidelity data:
$(\boldsymbol{X}_1, \underline{\mathrm{y}}_{1c}) \bigcup (\boldsymbol{X}_2, \underline{\mathrm{y}}_2) \longrightarrow \hat{y}$
Output: Fused metamodel: $\hat{y}$

A modified version will be used here. Corrected low-fidelity points were interpolated on the classical method. The correction made by the $\hat{y}_{corr}$ error metamodel can be somewhat inaccurate due to a lack of information. Adding variance on these points will release the exploration component of the acquisition function (detailed in the next section) around LF points which are far away from a high-fidelity observation (and therefore the correction is potentially false in this area).

So the improvement is to consider an estimated variance $\sigma_{\boldsymbol{\varepsilon}}$ on low-fidelity data. This noise can be quantified by the variance associated with the error metamodel. Indeed, the sum of the random variables $\tilde{\underline{\mathrm{y}}}_1$ and $\hat{y}_c(\boldsymbol{X}_1)$ gives an associated variance $\sigma_{\boldsymbol{\varepsilon}}(\boldsymbol{X}_1)$ such that

$$\sigma_{\boldsymbol{\varepsilon}}(\boldsymbol{X}_1)^2 = \sigma(\tilde{\underline{\mathrm{y}}}_1)^2 + \sigma(\hat{y}_c(\boldsymbol{X}_1))^2 + 2\,\mathrm{Cov}(\tilde{\underline{\mathrm{y}}}_1, \hat{y}_c(\boldsymbol{X}_1)) \tag{9}$$

The independence of the random variables is assumed: $\mathrm{Cov}(\tilde{\underline{\mathrm{y}}}_1, \hat{y}_c(\boldsymbol{X}_1)) = 0$. In addition, no measurement noise is considered on the LF observations initially: $\sigma(\tilde{\underline{\mathrm{y}}}_1) = 0$. We will therefore consider in the following $\sigma_{\boldsymbol{\varepsilon}}(\boldsymbol{X}_1) = \sigma(\hat{y}_c(\boldsymbol{X}_1))$. This analysis is also appropriate for HF data with $\sigma(\hat{y}_c(\boldsymbol{X}_2)) = 0$ at these points.

Figure 3 presents the method through the next functions [11]:

$$\begin{aligned} y_1 &: x \mapsto 0.5(6x-2)^2 + \sin(12x-4) + 10(x-0.5) - 5 \\ y_2 &: x \mapsto 2y_1(x) - 20x + 20 \end{aligned} \tag{10}$$

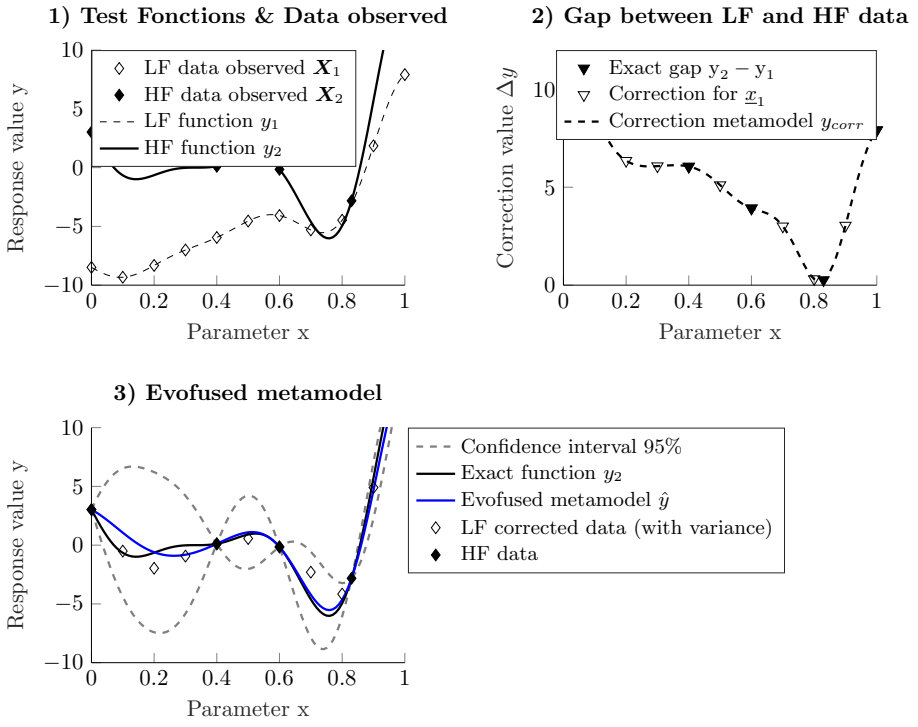


Figure 3: Illustration of the Evofusion method

1.4. Adapted acquisition function for multi-fidelity kriging

The role of the acquisition function is to guide the search for the optimum by considers a good exploitation–exploration trade-off. The new point will be around already known optimum zones (exploitation), or in wide unobserved areas (exploration). To do that, the expected improvement function [4] is the most widely used in the literature. It consists in calculating not the probability but the average of the improvement function by integrating the density function of the improvement function I:

$$\mathrm{I}(\underline{x}) = \langle \hat{y}(\underline{x}) - y^* \rangle_+ \tag{11}$$

with $\hat{y}$ the Gaussian process $N\left(\mu(\underline{x}), \sigma(\underline{x})\right)$ which modelizes y and y^* the maximum of observed values.

The new query point is found by maximizing the expected improvement:

$$\underline{x} = \text{argmax}\, \mathbb{E}[\mathrm{I}(\underline{x})] \tag{12}$$

The likelihood of improvement $\mathbb{E}[\mathrm{I}(\underline{x})]$ can be computed by integrating the density function of I:

$$\mathbb{E}[\mathrm{I}(\underline{x})] = \int_{I=0}^{I=\infty} I \frac{1}{\sqrt{2\pi}\sigma(\underline{x})} \exp\left(-\frac{(\mu(\underline{x}) - y^* - I)^2}{2\,\sigma^2(\underline{x})}\right) dI \tag{13}$$

and can be evaluated analytically:

$$\mathbb{E}[\mathrm{I}(\underline{x})] = (\mu(\underline{x}) - y^*)\Phi\left(\frac{\mu(\underline{x}) - y^*}{\sigma(\underline{x})}\right) + \sigma(\underline{x})\phi\left(\frac{\mu(\underline{x}) - y^*}{\sigma(\underline{x})}\right) \tag{14}$$

where Φ and ϕ denote respectively the probability density function and the cumulative distribution function of the standard normal distribution. There are two terms in the analytical expression (14). The first one is the exploration term whose contributions will be significant in the unobserved areas of $\mathcal{D}$. The second one is the exploitation term whose contributions will be significant in area around the observed optimum.

In the case of kriging without variance associated with observations, $\mathrm{EI}(\underline{x}) = 0$. As the function is positive, the points already sampled will never be recomputed. For this reason, it is essential that the low-fidelity data have an associated variance. In addition, the observed minimum is not necessarily accurate if it is derived from low-fidelity data. To address this last point, it is possible to not consider the observed minimum $y^\star$ but the minimum of the mean or a specific quantile of the Gaussian process at the observed points, noted $\mu^\star$ [34]. This method is called "Expected improvement with plug-in":

$$\mathrm{PI}(\underline{x}) = (\mu(\underline{x}) - \mu^\star)\Phi\left(\frac{\mu(\underline{x}) - \mu^\star}{\sigma(\underline{x})}\right) + \sigma(\underline{x})\phi\left(\frac{\mu(\underline{x}) - \mu^\star}{\sigma(\underline{x})}\right) \tag{15}$$

This method does not take into account the noise of the future observation: the improvement is defined and its expectation is calculated as if the next evaluation would be deterministic. Nevertheless, it already gives better results than EI [34].

2. Fast quality-driven QoI computations using the LATIN-PGD framework

2.1. Reference problem

For the sake of simplicity, let us consider the quasi-static and iso-thermal evolution of a viscoelastic structure defined over the time–space domain $I \times \Omega$, under the assumption of small perturbations. The continuum body is submitted to surface forces $\underline{\mathrm{F}}_d$ on a portion of its boundary $\partial_2\Omega$ and to body forces $\underline{\mathrm{f}}_d$ in its interior Ω. The continuum body is also submitted to prescribed displacements $\underline{\mathrm{U}}_d$ over $\partial_1\Omega$ with $\partial_1\Omega \cap \partial_2\Omega = \emptyset$.

Movement within the structure results in a displacement field with respect to the initial configuration $\underline{\mathrm{U}} : I \times \Omega \mapsto \mathbb{R}^3$. Infinitesimal strain theory is assumed and allows the linearization of the strain tensor $\boldsymbol{\varepsilon} : I \times \Omega \mapsto \mathbb{R}^3$ with $\boldsymbol{\varepsilon} = \frac{1}{2}\left[\nabla\underline{\mathrm{U}} + \nabla^T\underline{\mathrm{U}}\right]$. Internal forces are expressed by the stress tensor $\boldsymbol{\sigma} : I \times \Omega \mapsto \mathbb{R}^3$ verifying the following boundary value problem:

$$\underline{\mathrm{div}}(\boldsymbol{\sigma}) + \underline{\mathrm{f}}_d = \underline{0} \text{ on } \Omega, \quad \boldsymbol{\sigma}.\underline{n} = \underline{\mathrm{F}}_d \text{ on } \partial_2\Omega, \quad \underline{\mathrm{U}} = \underline{\mathrm{U}}_d \text{ on } \partial_1\Omega \tag{16}$$

The nonlinear behavior laws are taken into account with the equation (17) and will be in details on the following section, but formelly, one has:

$$\forall (M,t) \in \Omega \times I, \boldsymbol{\sigma}(M,t) = \mathcal{B}(\boldsymbol{\varepsilon}(M,\tau), \tau \in [0,t]) \tag{17}$$

In the context of elasto-viscoplasticity, solving the boundary value problem given by (16) is equivalent to minimizing the energy functional $\mathcal{E} : \mathcal{U} \mapsto \mathbb{R}$ defined as:

$$\mathcal{E}(\underline{\mathrm{U}}) = \int_\Omega \mathcal{B}(\boldsymbol{\varepsilon}) : \boldsymbol{\varepsilon}\, \mathrm{d}\Omega - \int_\Omega \underline{f_d} \cdot \underline{\mathrm{U}}\, \mathrm{d}\Omega - \int_{\partial_2\Omega} \underline{F_d} \cdot \underline{\mathrm{U}}\, \mathrm{dS} \tag{18}$$

where $\mathcal{U}$ is a functional space of the form:

$$\mathcal{U} = \left\{\underline{\mathrm{U}} \in \mathrm{H}^{1,p}(\Omega) \,/\, \mathcal{E}(\underline{\mathrm{U}}) < +\infty,\ \underline{\mathrm{U}} = \underline{\mathrm{U}}_d \text{ sur } \partial_1\Omega\right\}$$

with $p \geq 1$ and $\mathrm{H}^{1,p}(\Omega)$ the Sobolev space of functions in $L^p(\Omega, \mathbb{R}^3)$ with partial derivatives in $L^p(\Omega, \mathbb{R}^3)$.

A weak formulation of the boundary value problem given by (18) can be written *i.e.* find $\underline{\mathrm{U}} \in \mathcal{U} \times \mathcal{I}$ such that

$$\forall t \in I,\ \forall \underline{\mathrm{U}}^* \in \mathcal{U}^*,\ \mathcal{R}\left(\underline{\mathrm{U}}(M,t); \underline{\mathrm{U}}^*\right) = 0 \tag{19}$$

with $\mathcal{U}^*$ the space of functions in $\mathcal{U}$ that vanish on the boundary $\partial_1\Omega$ with $\underline{\mathrm{U}}_d = \underline{0}$, and the residual function $\mathcal{R}$ is define by:

$$\int_{I\times\Omega} \mathcal{B}(\boldsymbol{\varepsilon}[\underline{\mathrm{U}}]) : \boldsymbol{\varepsilon}(\underline{\mathrm{U}}^*)\, \mathrm{d}\Omega\mathrm{dt} - \int_{I\times\Omega} \underline{f_d} \cdot \underline{\mathrm{U}}^*\, \mathrm{d}\Omega\mathrm{dt} - \int_{I\times\partial_2\Omega} \underline{F_d} \cdot \underline{\mathrm{U}}^*\, \mathrm{dsdt} \tag{20}$$

The resolution of this nonlinear problem will be done using the LATIN method explained on section 2.2.

2.2. The Chaboche elasto-viscoplasticity behavior law

Unified viscoplastic framework previously presented in [35] is considered. In this constitutive law, strain $\boldsymbol{\varepsilon}$ is splited between elastic reversible strain $\boldsymbol{\varepsilon_e}$ and plastic strain $\boldsymbol{\varepsilon_p}$ with $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon_p} + \boldsymbol{\varepsilon_e}$. Stress is driven by $\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\varepsilon_e}$ where $\mathbf{C}$ is the Hooke matrix. Plastic behavior zone appears when stress goes over an elastic limit f. This limit is usually represented by an ellipsis in deviatoric stress principal component space also called yield surface. Size and origin of

the ellipsis are driven respectively by isotropic hardening R (drag effect) and a unique linear kinematic hardening X:

$$f = (\boldsymbol{\sigma} - R)_{eq} - \sigma_0 \tag{21}$$

where $J_2 = (\boldsymbol{\sigma} - R)_{eq}$ is the Von Mises equivalent stress and $\sigma_0 = \sigma_y - R$ is the yield surface size. We define the plastic strain p and the primal field linked with X, $\boldsymbol{\alpha}$. Primal fields $(p, \boldsymbol{\alpha})$ are associated with (R, X).

The Norton-Hoff law drives the plastic strain p:

$$\dot{p} = \left\langle \frac{f}{k} \right\rangle_+^N \tag{22}$$

where k, N are material dependant scalars, $\langle \cdot \rangle_+$ are Macaulay brackets.

State laws are:

$$\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\varepsilon_e} \tag{23}$$

$$X = \frac{2}{3} C \boldsymbol{\alpha} \tag{24}$$

$$R = R_\infty (1 - e^{-bp}) \tag{25}$$

where C, R_∞, b are material dependant scalars.

By the definition of a pseudo-dissipation potential F, the evolution equations can be expressed:

$$\frac{d}{dt} \begin{bmatrix} \boldsymbol{\varepsilon_p} \\ -\boldsymbol{\alpha} \\ -p \end{bmatrix} = \left\langle \frac{f}{k} \right\rangle_+^N \begin{pmatrix} \sqrt{\frac{3}{2}}\mathbb{N} \\ -\sqrt{\frac{3}{2}}\mathbb{N} + \frac{3\gamma}{2C} X \\ -1 \end{pmatrix} \tag{26}$$

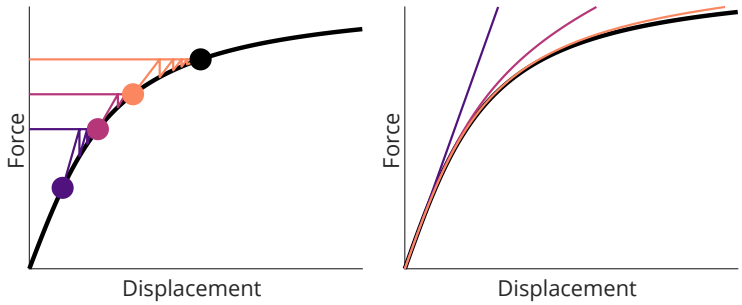
with $\mathbb{N}$ the unitary normal vector defined as

$$\mathbb{N} = \sqrt{\frac{3}{2}} \frac{\boldsymbol{\sigma}_D - X}{(\boldsymbol{\sigma}_D - X)_{eq}}, \qquad (\mathbb{N})_{eq} = 1 \tag{27}$$

2.3. LATIN algorithm for nonlinear quality-driven resolution

The LATIN framework [20] is a powerful method to obtain the linear equilibrium of the structure with the respect of nonlinear behavior laws. Here we recall the main principles of the method, the details of which can be found in [36].

Equations are separated into two sub-problem which defines the two manifolds A_d and Γ. In this case, A_d contains solutions which satisfy kinematic and static admissibility and state laws. These linear equations are global in time and space. On the other hand, the variety Γ contains the nonlinear behaviour



(a) Modified Newton-Raphson (b) LATIN

Figure 4: Schematic representation of nonlinear problem solving from iterative solvers

equations. $A_{d,0}$ is similar to A_d with boundary conditions set equal to zero. Unlike Newton-like techniques, this approach provides a complete solution at each iteration of the algorithm. The schematic representation in Figure 4 illustrates differences between the two solvers.

Lets denote $\mathbf{s} = (\boldsymbol{\varepsilon}, \boldsymbol{\sigma})$ and $\hat{\mathbf{s}} = (\hat{\boldsymbol{\varepsilon}}, \hat{\boldsymbol{\sigma}})$ set of fields respectively describing the state of the structure in A_d and in Γ. The solution is alternately searched in each manifold using search directions $\mathbf{E}^+$ and $\mathbf{E}^-$ which link the two subproblems. The physical quantities involved are written in the form of primary mechanical fields and dual mechanical fields associated with the problem. The solution is obtained at convergence and is in $A_d \cap \Gamma$. The method can be also apply to multiscale problems [37], for which micro and macro problems are solved alternately; domain decomposition [38], for which separation operates on subdomains on the one hand and interfaces on the other hand; or multi-physics [39]. This strategy has also been applied to other applications such as composite damage [40], rapid dynamic [41] or contact problems [42].

Steps of the method are briefly described below and figure 5 gives a schematic representation.

- **Elastic initialization:** The algorithm is classically initialized by computing the elastic solution of the problem, such that $\mathbf{s}_0 \in A_d$ but $\mathbf{s}_0$ can be enriched by adding $A_{d,0}$ fields interpolated from previous computations.

- **Local stage:** Knowing a solution $\mathbf{s}$ in A_d, local stage consists in finding a local solution $\hat{\mathbf{s}}$ in Γ using search direction $(\hat{\mathbf{s}} - \mathbf{s}) \in \mathbf{E}^+$.

$$\mathbf{E}^+(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}) + (\hat{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon}) = \mathbf{0} \tag{28}$$

  The local problem can be solved at each time steps and at each Gauss point. Here, the search direction is chosen to have $\hat{\boldsymbol{\varepsilon}} = \boldsymbol{\varepsilon}$.

- **Global stage:** The linear stage consists in finding $\mathbf{s} \in A_d$ knowing $\hat{\mathbf{s}} \in \Gamma$ and using the search direction $\mathbf{E}^-$ (29).

$$\mathbf{E}^-(\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}}) - (\boldsymbol{\varepsilon} - \hat{\boldsymbol{\varepsilon}}) = \mathbf{0} \tag{29}$$

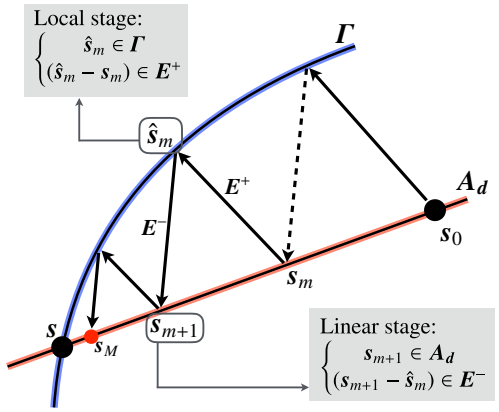


Figure 5: Schematic representation of the LATIN strategy

The corresponding sub-problem is linear but global in space and time. This problem is equivalent to the search of a correction $\delta\mathbf{s} \in A_{d,0}$ of the solution $\tilde{\mathbf{s}} \in A_d$ calculated at the previous convergence iteration. Equation 29 can be written as:

$$\mathbf{E}^-((\tilde{\boldsymbol{\sigma}} + \delta\boldsymbol{\sigma}) - \hat{\boldsymbol{\sigma}}) - ((\tilde{\boldsymbol{\varepsilon}} + \delta\boldsymbol{\varepsilon}) - \hat{\boldsymbol{\varepsilon}}) = 0 \tag{30}$$

The chosen error indicator is a stagnation criterion on consecutive solutions $\mathbf{s}$ and $\hat{\mathbf{s}}$:

$$\nu_{i+1} = \frac{||\mathbf{s}_{i+1} - \hat{\mathbf{s}}_{i+\frac{1}{2}}||_{\mathbf{C}}}{\frac{1}{2}||\mathbf{s}_{i+1} + \hat{\mathbf{s}}_{i+\frac{1}{2}}||_{\mathbf{C}}}, \quad ||\mathbf{s}||_{\mathbf{C}} = \int_{\Omega\times I} \frac{1}{2}\left(\boldsymbol{\varepsilon} : \mathbf{C}\,\boldsymbol{\varepsilon} + \boldsymbol{\sigma} : \mathbf{C}^{-1}\boldsymbol{\sigma}\right) \mathrm{d}\Omega\mathrm{d}t \tag{31}$$

where $||\mathbf{s}||_{\mathbf{C}}$ is the norm associated to the Hooke operator.

This method is particulary adapted to our process:

- An interrupted Newton-Raphson resolution gives the converged solution up to the time step i_t. In the case of the LATIN method, an approximation of the complete spatio-temporal solution is available. This property allows the solver to be considered as a quality-driven one and will be used to generate multi-fidelity QoI data.

- It is possible to approximate the resolution at the global stage by searching fields correction $\delta\mathbf{s}$ like separated variable functions. The computational cost of the global stage is reduced and the set of fields $\mathbf{s}$ can be written as a PGD reduced-order model. This aspect will be explained on section 2.4.

- The initialization of the Newton-Raphson method by an already computed spatio-temporal approximation is complex. By its structure, the LATIN

method naturally allows the initialization of the calculation by a complete pre-calculated spatio-temporal solution. This property allows a enhanced multiparametric strategy detailled on section 2.5.

2.4. On-the-fly model-order reduction with the PGD method

The Proper Generalized Decomposition is a ROM method. Initiated under the name of radial approximation [43] as part of the LATIN framework for space-time decomposition, it is also used for rheology problems [44], for an approximate resolution of the Navier-Stokes equation [45] or in image correlation [46]. The method has been generalized with multiparametric use [47] — or even in large dimensions [48] — with possible use of hyper-reduction methods [49]. The method is also particularly suitable for application cases such as optimization [50], uncertainty quantification [51] in dynamic data-driven application systems (DDDAS) [52] or in haptics [53].

The objective is to compute the best correction of an approximation of the solution in a separated variable form. Modes database is no longer imposed as in Proper Orthogonal Decomposition and the solution of the problem is sought at the same time as the best basis for its representation. The approximation performed is enriched until a given error criterion is reached.

We assume that a PGD approximation of order $m-1$ has already been computed:

$$\underline{\mathrm{U}}(t,M) \approx \underline{\mathrm{U}}_0(t,M) + \underbrace{\sum_{i=1}^{m-1} \lambda_i(t)\cdot\underline{\Lambda}_i(M)}_{\underline{\mathrm{U}}_c(t,M)} \tag{32}$$

with $\underline{\mathrm{U}}_0$ the elastic solution such as $\underline{\mathrm{U}}_0 \in A_d$ and $\forall i, (\underline{\Lambda}_i, \lambda_i) \in \mathcal{U}^* \times \mathcal{I}$. In the PGD approach, neither functions $\underline{\Lambda}_k$ nor functions λ_k are given initially, and both families are computed on-the-fly. At the global linear stage of the LATIN method, the correction of fields $\delta\underline{\mathrm{U}}$, $\delta\boldsymbol{\varepsilon}$, $\delta\boldsymbol{\sigma}$ can be find with a separated-variable form:

$$\delta\underline{\mathrm{U}} = \underline{\Lambda}\cdot\lambda \tag{33}$$

And can be obtained by minimizing the global search direction (19). It can be expressed as [18],

$$\text{Find } \delta\underline{\mathrm{U}}(t,M) \in \mathcal{U}^* \text{ such that } \forall \underline{\mathrm{U}}^* \in \mathcal{U}^*, a(\delta\underline{\mathrm{U}}, \underline{\mathrm{U}}^*) = \ell(\underline{\mathrm{U}}^*) \tag{34}$$

with $\mathcal{I} = L^2(I,\mathbb{R})$ the space of the sumable square functions in I, a is a continuous coercive bilinear operator according to $H_0^1(\Omega)\otimes\mathcal{I}$ and ℓ is a continuous linear operator according to $H_0^1(\Omega)\otimes\mathcal{I}$.

The new pair $(\underline{\Lambda},\lambda) \in \mathcal{U}^*\times\mathcal{I}$ is defined as the one that verifies the following double Galerkin orthogonality criterion:

$$a(\tilde{u}_{m-1} + \underline{\Lambda}\lambda, \underline{\Lambda}\lambda^* + \underline{\Lambda}^*\lambda) = \ell(\underline{\Lambda}\lambda^* + \underline{\Lambda}^*\lambda), \quad \forall\lambda^* \in \mathcal{I}^*, \forall\underline{\Lambda}^* \in \mathcal{I} \tag{35}$$

We can thus define the two following applications:

- $S_m : \mathcal{I} \mapsto \mathcal{U}$ is the application that maps a time function λ into a space function $\underline{\Lambda} = S_m(\lambda)$ defined as:

$$a(\tilde{u}_{m-1} + \underline{\Lambda}\lambda, +\underline{\Lambda}^*\lambda) = \ell(\underline{\Lambda}^*\lambda), \quad \forall \underline{\Lambda}^* \in \mathcal{U}^* \tag{36}$$

  It is associated to a space problem.

- $T_m : \mathcal{U} \mapsto \mathcal{I}$ is the application that maps a space function $\underline{\Lambda}$ into a time function $\lambda = T_m(\underline{\Lambda})$ defined as:

$$a(\tilde{u}_{m-1} + \underline{\Lambda}\lambda, +\lambda^*\underline{\Lambda}) = \ell(\lambda^*\underline{\Lambda}), \quad \forall \lambda^* \in \mathcal{I} \tag{37}$$

  It is associated to a time problem (Scalar ODE).

A pair $(\underline{\Lambda}, \lambda)$ verifies (35) if and only if $\underline{\Lambda} = S_m(\lambda)$ and $\lambda = T_m(\underline{\Lambda})$, which is a nonlinear problem which could be solved with a fixed-point algorithm. The interested reader can refer to [54] for a review of the different algorithms to solve a linear problem with the PGD and [55] for the special case of the viscoplastic LATIN algorithm.

So with the LATIN-PGD framework, at each convergence step, a reduced-order approximation of the complete spatio-temporal solution is obtained and allow us to compute an approximation of the QoI firstly. By this aspect, a space and time basis can also be used to start future computations.

2.5. Fast computation of multiple solutions with LATIN multiparametric strategy and PGD interpolation

For surrogate modelling, the computation of a quantity of interest yi is done sequentially on certain points $\underline{x}i$ in the design space $\mathcal{D}$. The computation of multiple solutions is fully compatible with a major feature of the LATIN framework: its multiparametric strategy [56] which allows to start a new calculation with fields created from previous computations. Its objective is to provide very quickly the solution of nonlinear evolution problem for several parameter values of the model and reduce the number of iterations to reach the required error estimator level.

We assume that for $i \in [\![1, j-1]\!]$, displacement solution $\underline{\mathrm{U}}_i$ linked with the problem parametred by $\underline{x}_i$ has already been computed. Each solution are expressed from its space and time basis generated at the same time as the solution.

For the calculation of the query parameter $\underline{x}_j$, an approximation of $\mathbf{s} \in A_d$ from previous computations is considered by interpolate the displacement field $\underline{\mathrm{U}}_j$. Inverse distance weighting is considered, using euclidian distance in the design space $\mathcal{D}$:

$$\underline{\mathrm{U}}_j(t, M) \approx \underline{\mathrm{U}}_{j,0}(t, M) + \sum_{k=0}^{j-1} \omega_k \, \underline{\mathrm{U}}_{k,c}(t, M) \tag{38}$$

with $\forall k \in [\![1, j-1]\!]$, $\omega_k = \phi(||\underline{x}_j - \underline{x}_k||_{\mathcal{D}})$, $||\bullet||_{\mathcal{D}}$ the euclidian norm of $\mathcal{D}$ and $\phi : \mathrm{d} \mapsto \dfrac{1}{\mathrm{d}^p}$.

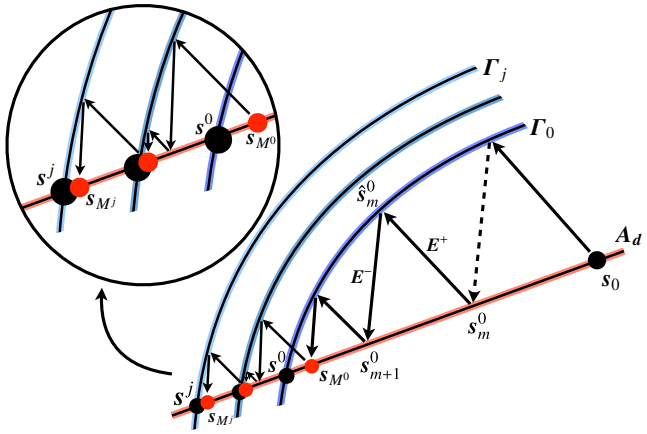


Figure 6: Schematic representation of the multiparametric strategy [57]. For computing solution parametred by $\underline{x}_j$, the solution of a similar problem parametred by $\underline{x}_{j-1}$ can be used to reduce the iteration number.

Furthermore, a global space basis $\mathbf{U} = [\underline{\Lambda}_1, \underline{\Lambda}_2, \ldots, \underline{\Lambda}_{ms}] \in \mathbb{R}^{d_s \times m_s}$ can be considered by orthonormalizing the concatenation of space bases previously obtained, such as $\forall k \in [\![1, j-1]\!]$:

$$\underline{\mathrm{U}}_{k,c}(t, M) = \sum_{i=0}^{m_s} \lambda_i^{(k)}(t) \Lambda_i(M) \tag{39}$$

So, by considering (38) and (39),

$$\underline{\mathrm{U}}_j(t, M) \approx \underline{\mathrm{U}}_{j,0}(t, M) + \sum_{i=0}^{m_s} \tilde{\lambda}_i(t) \Lambda_i(M) \tag{40}$$

with $\forall i \in [\![1, m_s]\!]$ $\tilde{\lambda}_i = \sum_{k=0}^{j-1} \omega_k \lambda_i^{(k)}$. Algorithm 2 summarizes the different steps of the multiparametric strategy implemented.

This strategy is tested on the turbine blade test case (see section 3.4). A 3D solver implementation was developed for the test using MATLAB. To reach fast and parallel computation, LATIN operators are organized as nd-arrays and a parallel implementation of Einstein summation exploits variable broadcasting to deal with time- and spatial-dependent operators. This computational paradigm allows to be as fast as commercial softwares. 200 equidistant points of the design space $[\alpha, T] = [81°, 100°] \times [881°\mathrm{C}, 900°\mathrm{C}]$ are computed without the strategy, and with the strategy by an ordered path, and 15 random paths:

With this LATIN-PGD framework, multi-fidelity QoI values can be obtained easily. Furthermore, the multiparametric strategy allows good speedups. The

Algorithm 2: Multiparametric strategy

Input: n points $\mu_1, \ldots, \mu_n \in \mathcal{D}$
For each point μ_j
  For each previous point μ_k
    Calculation of the contribution: $\omega_k = \phi(\mathrm{d}(\mu_j, \mu_k))$
  Field interpolation: $\underline{\mathrm{U}}_j(t, M) \leftarrow \sum_{k=1}^{j-1} \omega_k \underline{\mathrm{U}}_k(t, M)$
  Generation of the common spatial base
    Concatenation of spatial modes:
    $[\boldsymbol{\Xi}, \boldsymbol{\Sigma}, \mathbf{V}] \xleftarrow{\text{Thin SVD}} [\boldsymbol{\Xi}_1, \ldots, \boldsymbol{\Xi}_{j-1}]$
  Calculation of the associated elastic field: $\underline{\mathrm{U}}_{j,0}(t, M)$
  For each spatial mode $\Lambda_i^{(\Xi)}$
    Compute correction of time functions $\bar{\lambda}_i \leftarrow \left\langle \Lambda_i^{(\Xi)}, \underline{\tilde{\mathrm{U}}}_j - \underline{\mathrm{U}}_{j,0} \right\rangle_S$
  LATIN Correction: $\underline{\mathrm{U}}_j \in \mathbf{A}_{\mathrm{d}j} \cap \Gamma_j$
Output: Mechanical solution for the n points

Points computed	**2**	**5**	**10**	**20**	**50**	**100**	**200**
Ordered path	4.25×	4.24×	4.46×	4.86×	4.95×	4.87×	5.05×
Random path	2.09×	2.13×	2.22×	2.56×	3.13×	3.51×	3.78×

Table 1: Speedup provided by the multiparametric strategy on the turbine blade test case

aim here is to see if using multi-fidelity bayesian optimisation with the multi-parametric strategy is adapted and reduces time computation to obtain a good estimation of the global optimum.

3. Global optimization by coupling multi-fidelity kriging and reduced-order models

3.1. Coupling algorithm

The different algorithmic bricks presented in the previous sections are assembled in algorithm 7. This algorithm have some parameters that influence the calculation time and the quality of the optimum obtained. There are two of them: The number of low-fidelity observations n_{BF} before enrichment with high-fidelity data, and the quality of low-fidelity observations which will be driven by the LATIN indicator ν_{BF} which will be the stopping criterion of the solver. To obtain the optimal pair of parameters, an important testing campaign was carried out. Results are presented on two test cases on following sections.

3.2. Presentation of the damper test case

This second test case is freely inspired from a damper part of the Vulcain engine of the Ariane 5 launcher [18]. The typical dimensions in (x,y,z) are

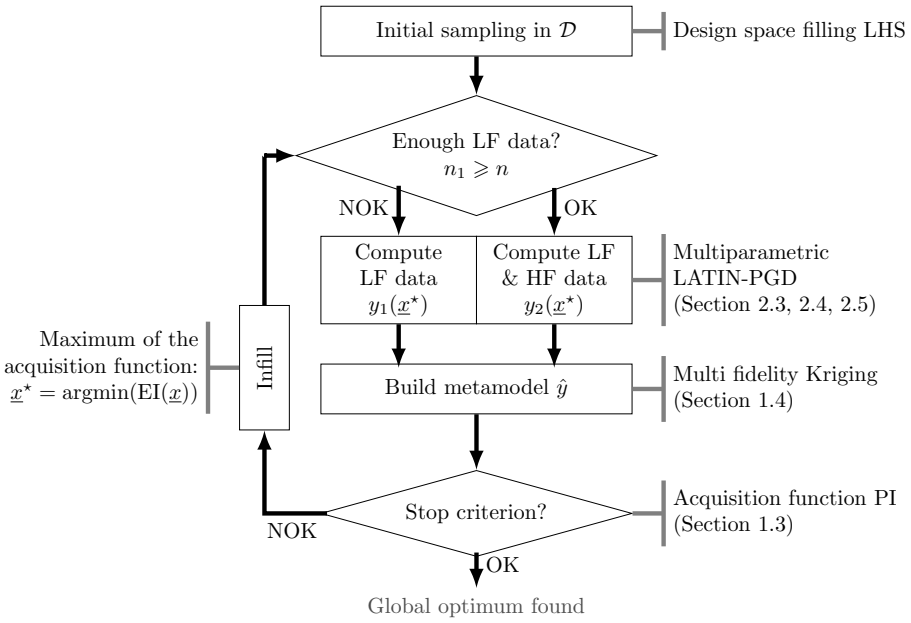


Figure 7: Enriched bayesian optimization

$45 \times 70 \times 50$ mm. The mechanical part is clamped on the upper, left and right sides (Blue parts on figure 8). The turbine nose is visible on the bottom left of the figure 8. This one is loaded on the front and rear sides (on green and red on the figure). The two loading pressures have the same intensity $P(t)$ described by figure 9 with $P_{max} = 80$ MPa, but their direction are different. Each loading direction is driven by two angles (θ, ϕ) which could go between $[0°, 90°]$. The description of the angle-driven loading conditions is shown on figure 8. The material is a 316 Steel at 600°K. Material behavior is described by the Chaboche constitutive law (see Section 2.2) and Table 3 for the material coefficients.

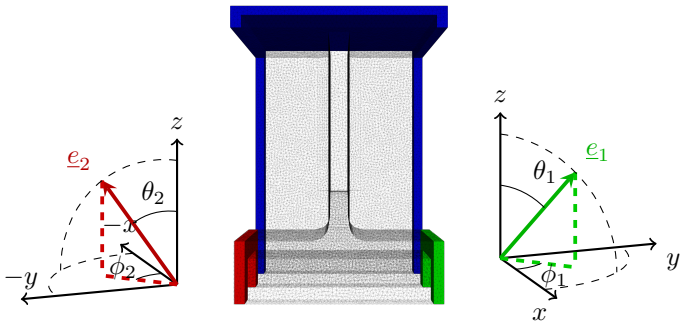
The aim here is to obtain the worse loading case for the structure by dealing with the maximum of the Von Mises stress during one loading cycle. All the loading options describe a 4D design space $\mathcal{D} = [0°, 90°]^4$.

3.3. Strategy parameters estimation on the damper test case

For each of the (n_1, ν_1) pairs tested, 100 optimizations are performed with different initial points. A budget of 100 high-fidelity points is set. The optimization success criterion corresponds to the case where the exact solution calculated at least estimated is greater than 113 MPa:

$$y_2(\underline{x}^\star) > 113\text{MPa} \tag{41}$$

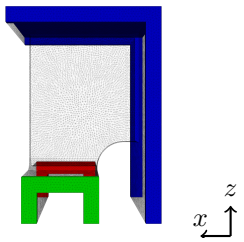
Table 4 shows the percentage of test cases that converged to the global minimum area. Table 5 give computation time to obtain the success criterion and



(a) Angle-driven loading force on the red part

(b) X-view

(c) Angle-driven loading force on the green part



(d) Y-view

Figure 8: Damper test case

Objective function	$y(\underline{x}) = \max_{I\times\Omega} \sigma_{VonMises}$
Parameters	$\underline{x} = (\theta_1, \phi_1, \theta_2, \phi_2) \in \mathcal{D} = [0°, 90°]^4$
Space element type	Linear triangular (30k DOFs)
Loading cycles	One cycle (10s — 41 time steps)
Boundary conditions	Clamped on blue side faces
	$\underline{P}_1(t) = P(t)\,\underline{e}_1$ on green
	$\underline{P}_2(t) = P(t)\,\underline{e}_2$ on red

Table 2: Second test case characteristics

E	ν	N	k	σ_0	C	γ	R_∞	b
137.6 GPa	0.3	14	150 MPa.s$^{1/n}$	20 MPa	37.2 GPa	300	80 MPa	10

Table 3: Elastic-viscoplastic constitutive coefficients for the second test case

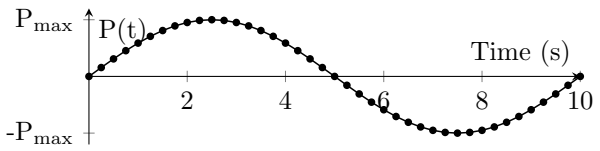


Figure 9: Pressure evolution P(t)

Table 6 gives the impact of the random choice of initial points on computation time.

ν_1 / n_1	10^{-3}	10^{-2}	10^{-1}
6/dim	56%	83%	75%
10/dim	58%	84%	74%
12/dim	63%	87%	74%
14/dim	59%	87%	77%

Table 4: Percentage of cases where the global minimum area was found

ν_1 / n_1	10^{-3}	10^{-2}	10^{-1}
6/dim	27'40	9'48	3'14
10/dim	36'02	10'09	3'23
12/dim	42'38	10'27	3'26
14/dim	43'22	11'14	3'38

Table 5: Average calculation time — Only high-fidelity data: 1h01

ν_1 / n_1	10^{-3}	10^{-2}	10^{-1}
6/dim	47%	87%	14%
10/dim	72%	73%	28%
12/dim	106%	78%	30%
14/dim	107%	91%	42%

Table 6: Normalized variance of computation time — Only HF: 34%

For reducing computation time, the best strategy is to consider firstly the computation of 4×6 low-fidelity points, obtained by considering $\nu_1 = 10^{-1}$ as the solver stopping criterion. This result is hardly surprising when looking at curve 10.

This figure is the empirical cumulative distribution function obtained by computing 20,000 QoI values from the 4D design space. The curve obtained for $\nu_1 = 10^{-1}$ allows to identify more easily the optimum area than informations given by high-fidelity values ν_2. With this choice, the gain brought by multi-fidelity is 19×. The gain provided by the multiparametric method is estimated by testing multiple random paths of 30 points in the design space. The contribution of the multiparametric strategy allows fields to be calculated 2.5× faster, which allows to estimate that the gain provided by coupling methods is around 47.5×. Nevertheless, using $\nu_1 = 10^{-2}$ seems to be a good compromise between

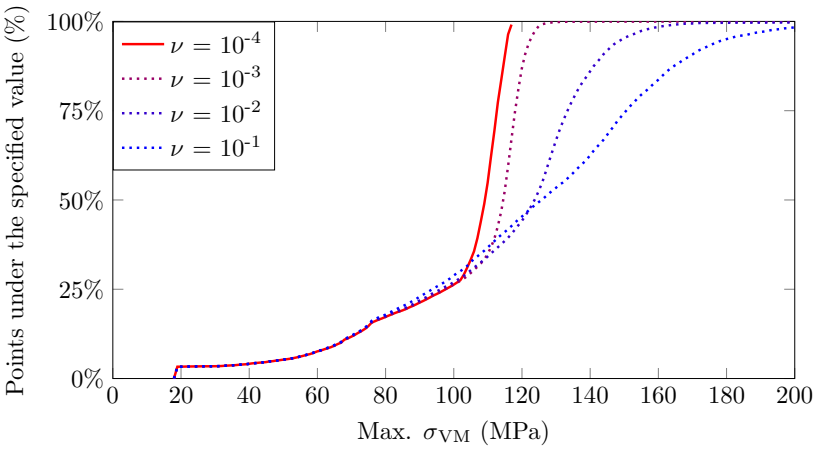


Figure 10: Empirical QoI cumulative distribution function

calculation time and success and this gain needs to be validated *a posteriori* by performing the calculation without the contribution of the multiparametric strategy, with only high-fidelity data. The next test case overcomes this last issue.

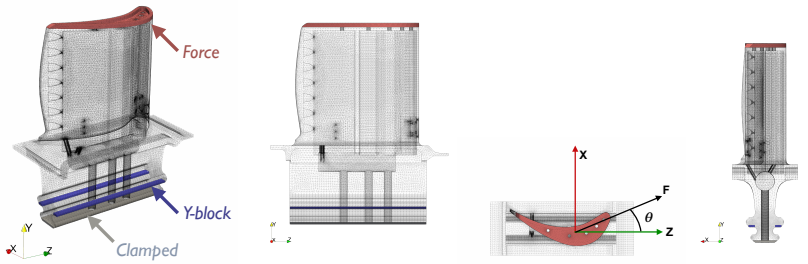
3.4. Presentation of the turbine blade test case

The second test-case is an aircraft turbine blade. The mesh comes from [58]. This blade is considered as clamped on its base, with blocked movement along the y-axis on some bottom sides (on purple in the figure 11). This one has a centrifugal load and a uniform load on the upper surface (on red in the figure 11). The direction load is in the surface plane and is driven by angle $\alpha \in [0°, 180°]$ as seen in the figure 11c. The intensity of loads are shown on figure 12. The material of this test case is an Inconel 601 at temperature $T \in [800°C, 900°C]$. Operating temperature will influence material parameters. Three sets of material parameters are given in table 7 and other values will interpolated from them.

The aim here is to obtain $\max_{I\times\Omega} \sigma_{VonMises} = 180$ MPa. To do that, a minimization is done on the 2D-design space with $(\alpha, T) \in [0°, 180°]\times[800°\mathrm{C}, 900°\mathrm{C}]$.

T	n	K	σ_0	C	γ	R_∞	b
800°C	14	630 MPa.s$^{1/n}$	80 MPa	615 GPa	1.53.10^6 s^{-1}	80 MPa	300
850°C	11	560 MPa.s$^{1/n}$	71 MPa	497 GPa	1.36.10^6 s^{-1}	70 MPa	250
900°C	9	490 MPa.s$^{1/n}$	60 MPa	362 GPa	1.2.10^{6}s^{-1}	60 MPa	200

Table 7: Elastic-viscoplastic constitutive coefficients for the first test case (with $E = 210$ GPa and $\nu = 0.28$)



(a) 3D-view (b) X-view (c) Y-view (d) Z-view

Figure 11: Different views of the structure

Objective function	$y(\underline{x}) = \lvert\max_{I\times\Omega} \sigma_{VonMises} - 180\rvert$
Parameters	$\underline{x} = (\alpha, T) \in [0°, 180°] \times [850°\text{C}, 950°\text{C}]$
Space element type	Quadratric triangular (330kDOFs)
Boundary conditions	Angle-driven load & Centrifugal load
	Clamping on bottom part
	Y-Block on bottom sides

Table 8: First test case characteristics

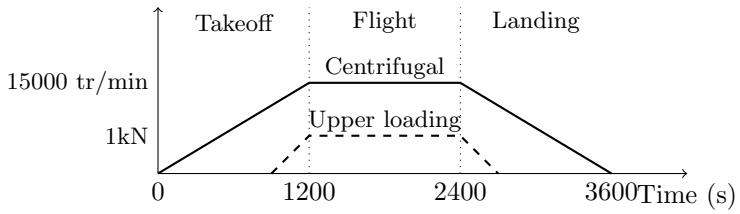
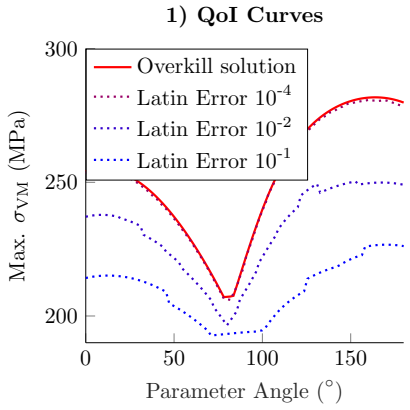


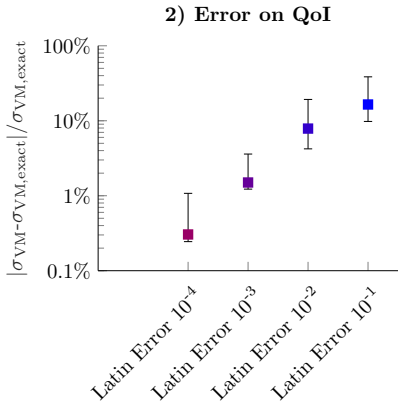
Figure 12: Loadings time evolution

181 loading cases are tested with $T = 900°\text{C}$ and $\alpha \in [0°, 180°]$ to quantify the influence of the error indicator explained in 2.3 on the computation time, on the number of PGD modes generated, and the error on the quantity of interest. Results are presented in Figure 13 with an overkill solution computed with LATIN indicator equal to 10^{-8}. Theses results allow us to consider LATIN Indicator 10^{-4} as the stopping criterion for converged field.

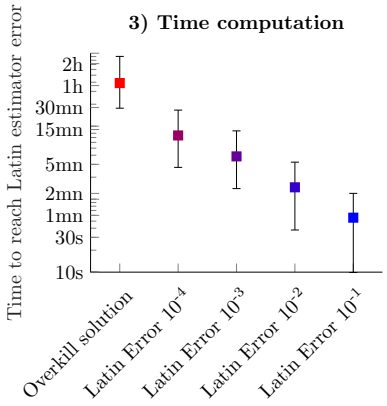
Figure 13a shows that even with low-fidelity data like $\nu_1 = 10^{-1}$, the QoI curves gives good trend and a good localization of the optimum area. A balance between QoI error and computation time can be found : if we consider low-fidelity data with $\nu_1 = 10^{-2}$, the figure 13b shows that the mean of the QoI error is around 8%, but computation time is 4.6× faster than with data obtained with $\nu_2 = 10^{-4}$. So 4 observations can be obtained faster than obtaining a single high-



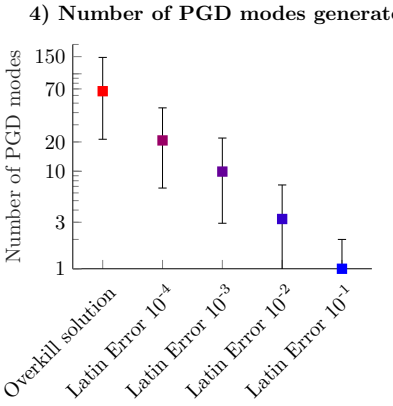
(a) Quantity of interest



(b) Error on QoI compared to overkill solution



(c) Time to reach Latin estimator error



(d) Number of PGD modes generated

Figure 13: Influence of the value of the LATIN error indicator as stopping solver criterion

fidelity observation, which allows to find faster the optimum zone. Nevertheless, high-fidelity informations are still mandatory to verify the obtained optimum.

3.5. Results of the optimization test on the turbine blade test case

To distinguish the performance gain provided by the multi-fidelity data fusion and the multiparametric strategy, three computation batchs are performed. Each batch considers 20 optimisations made with different initial sampling to estimate the average time savings. The first batch ϕ_0 is performed with only exact data without the multiparametric strategy. The second one ϕ_1 is made with only exact data and the multiparametric strategy, and the last one ϕ_2 is

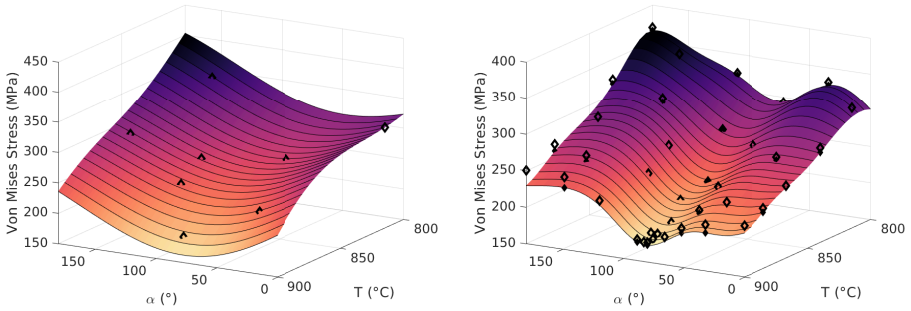
obtained by coupling methods. Each optimisation are made by initially compute 10 points per dimension with an LF stopping criterion $\nu_1 = 10^{-2}$.

The success criterion corresponds to the case where the exact solution calculated at least estimated is less than 180 MPa:

$$y_2(\underline{x}^\star) < 180\,\text{MPa} \tag{42}$$

In this case, the success criterion is met each time. The ϕ_0 phase is performed in 6h05, the ϕ_1 phase in 2h51 and finally the ϕ_2 phase in 44 minutes. The multiparametric strategy provides a gain of 2.1× and the multi-fidelity strategy a gain of 3.9× for a total gain of 8.4×. 17 high-fidelity points were requested on average by the algorithm to pass the objective (In the worst case, 24 high-fidelity points were requested).

Figure 14a shows one of the 20 initial sampling used to start the bayesian optimisation. Every low-fidelity observations are interpolated since we don't have information about QoI error on these points. Figure 14b shows the last metamodel generated by the bayesian optimisation. Low-fidelity observations were corrected by high-fidelity information and the correction metamodel. The low-fidelity corrected point in $(180°, 900°\text{C})$ is not interpolant as we consider the lack of information on the correction metamodel and so its variance. High-fidelity points are filled on the figure. On the same position in the design space, low-fidelity points are also visible to understand the gap between low-fidelity and high-fidelity data.

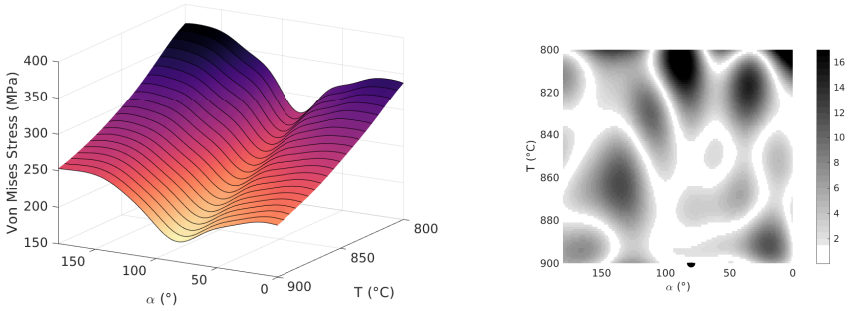


(a) The initial one (with only low-fidelity data)

(b) The final one (High-fidelity data are filled diamonds)

Figure 14: Metamodels generated during the optimization process

This figure can be compared to Figure 15a which is the QoI reference function to give Figure 15b. The QoI error between the final metamodel generated by the optimization and the reference shows that the QoI error is almost zero around the optimum, but quite important far from the optimum zone.



(a) Reference

(b) QoI error

Figure 15: Comparaison between the final metamodel generated by the optimization and the reference

Conclusion

The aim of this work is to accelerate the global optimization of mechanical structures to allow the use of more complex and accurate models within the design offices. To achieve this goal, the bayesian optimization process has been improved to the use of multi-fidelity data by considering an adapted acquisition function. The surrogate model generation step also uses an improved version of the multi-fidelity kriging method Evofusion which adding an estimated variance on low-fidelity data.

The multi-fidelity data are not obtained by calling a given quality model, but by using a solver allowing the computation of mechanical fields with an adapted level of fidelity. The proposed solver is based on the LATIN-PGD framework, which has the triple advantage of a fast computation of the spatio-temporal fields in the PGD separated-variable form, giving an approximation of the complete solution and the amount of interest at each iteration, and allowing to start computation from an approximation interpolated from previous computations with the multiparametric strategy.

The two families of methods presented will thus work together to bring significant time savings on the global optimization of a quantity of interest. The generalization of the enrichment strategy requires the determination of the number of calculated low-fidelity points, and the solver stopping criterion in this low-fidelity case. A parametric study campaign estimate these two parameters. 10 low-fidelity points per dimension with a LATIN stopping criterion $\nu_1 = 10^{-2}$ will be firstly considered before adding high-fidelity points. With these parameters, the airplane blade optimization was obtained in 44 minutes instead of 6 hours for a gain of 8.4×.

In the short term, the strategy is needed to be applied for geometric optimization problems. The main difficulty concerns the multiparametric strategy and the interpolation of fields for different geometries. This difficulty can be overcome by morphing around a single mesh [59], by using XFEM methods

[60], or by isoeometric parameterization of the structure (Isoeometrical Analysis — IGA) [61]. In this last case, tools already developed can be used using an IGA-FEM coupling method [62].

Thereupon, the strategy can also be improved by a better choice of the acquisition function. A comparison of the different acquisition functions should be made in the same way as [34] in the case of mechanical problems. Several enrichment criteria, including the AKG criterion, were not compared to the others. However, it appears that the AKG function can help to obtain the optimum faster.

References

[1] P. Duysinx, M. P. Bendsøe, Topology Optimization of Continuum Structures with Local Stress Constraints, International Journal for Numerical Methods in Engineering 43 (1998) 1453–1478 (1998).

[2] G. Allaire, Shape Optimization by the Homogenization Method, Applied Mathematical Sciences, Springer-Verlag, New York, 2002 (2002).

[3] S. Choi, J. J. Alonso, S. Kim, I. M. Kroo, M. Wintzer, Two-Level Multi-Fidelity Design Optimization Studies for Supersonic Jets, in: 43rd AIAA Aerospace Sciences Meeting and Exhibit, AIAA Paper 2005 - 0531, Reno, Nevada, 2005, p. 2 (2005).

[4] J. Mockus, V. Tiesis, A. Zilinskas, The application of Bayesian methods for seeking the extremum, Towards global optimization 2 (117-129) (1978) 2 (1978).

[5] D. R. Jones, M. Schonlau, W. J. Welch, Efficient global optimization of expensive black-box functions, Journal of Global optimization 13 (4) (1998) 455–492 (1998).

[6] D. G. Krige, A statistical approach to some mine valuation and allied problems on the Witwatersrand: By DG Krige (1951).

[7] G. Matheron, Principles of geostatistics, Economic geology 58 (8) (1963) 1246–1266 (1963).

[8] C. E. Rasmussen, C. K. Williams, Gaussian processes in machine learning, Lecture notes in computer science 3176 (2004) 63–71 (2004).

[9] F. Hutter, H. H. Hoos, K. Leyton-Brown, Sequential Model-Based Optimization for General Algorithm Configuration, in: C. A. C. Coello (Ed.), Learning and Intelligent Optimization, Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2011, pp. 507–523 (2011).

[10] K. Kandasamy, G. Dasarathy, J. Oliva, J. Schneider, B. Póczos, Gaussian Process Bandit Optimisation with Multi-fidelity Evaluations (2016) 23 (2016).

[11] A. I. Forrester, N. W. Bressloff, A. J. Keane, Optimization using surrogate models and partially converged computational fluid dynamics simulations, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 462 (2071) (2006) 2177–2204 (Jul. 2006). doi:10.1098/rspa.2006.1679.

[12] Z. Han, R. Zimmerman, S. Görtz, Alternative Cokriging Method for Variable-Fidelity Surrogate Modeling, AIAA Journal 50 (5) (2012) 1205–1210 (2012). doi:10.2514/1.J051243.

[13] R. Zimmermann, Z.-H. Han, Simplified Cross-Correlation Estimation for Multi-Fidelity Surrogate Cokriging Models, Advances and Applications in Mathematical Sciences 7 (2) (2010) 181–201 (2010).

[14] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, Journal of the Royal Statistical Society: Series B (Statistical Methodology) (2017) 301–320 (2017). doi:10.1111/j.1467-9868.2005.00503.x.

[15] H. K. Yong, L. Wang, D. J. J. Toal, A. J. Keane, F. Stanley, Multi-fidelity Kriging-assisted structural optimization of whole engine models employing medial meshes, Structural and Multidisciplinary Optimization (2019). doi:10.1007/s00158-019-02242-6.

[16] T. Benamara, P. Breitkopf, I. Lepot, C. Sainvitu, P. Villon, Multi-fidelity POD surrogate-assisted optimization: Concept and aero-design study, Structural and Multidisciplinary Optimization 56 (6) (2017) 1387–1412 (Dec. 2017). doi:10.1007/s00158-017-1730-4.

[17] F. Chinesta, R. Keunings, A. Leygue, The Proper Generalized Decomposition for Advanced Numerical Simulations, SpringerBriefs in Applied Sciences and Technology, Springer International Publishing, Cham, 2014 (2014).

[18] D. Néron, P.-A. Boucard, N. Relun, Time-space PGD for the rapid solution of 3D nonlinear parametrized problems in the many-query context, International Journal for Numerical Methods in Engineering 103 (4) (2015) 275–292 (2015). doi:10.1002/nme.4893.

[19] S. Nachar, P.-A. Boucard, D. Néron, F. Bordeu, Coupling multi-fidelity kriging & model-order reduction for the construction of virtual charts, Computational Mechanics (2019).

[20] P. Ladevèze, Nonlinear Computational Structural Mechanics: New Approaches and Non-Incremental Methods of Calculation, Mechanical Engineering Series, Springer Verlag, 1999 (1999).

[21] P.-A. Boucard, P. Ladevèze, A multiple solution method for non-linear structural mechanics, Mechanical Engineering 50 (5) (1999) 317–328 (1999).

[22] C. Heyberger, P.-A. Boucard, D. Néron, Multiparametric analysis within the proper generalized decomposition framework, Computational Mechanics 49 (3) (2012) 277–289 (Mar. 2012). doi:10.1007/s00466-011-0646-x.

[23] J. Nocedal, S. J. Wright, Numerical Optimization, 2nd Edition, Springer, 2006 (2006).

[24] J. Močkus, On Bayesian Methods for Seeking the Extremum, in: G. I. Marchuk (Ed.), Optimization Techniques IFIP Technical Conference: Novosibirsk, July 1–7, 1974, Lecture Notes in Computer Science, Springer Berlin Heidelberg, Berlin, Heidelberg, 1975, pp. 400–404 (1975). doi:10.1007/978-3-662-38527-2.

[25] D. R. Jones, A Taxonomy of Global Optimization Methods Based on Response Surfaces, Journal of Global Optimization 21 (4) (2001) 345–383 (Dec. 2001). doi:10.1023/A:1012771025575.

[26] M. D. McKay, R. J. Beckman, W. J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics 42 (1) (2000) 55–61 (2000).

[27] C. K. I. Williams, Prediction with Gaussian processes: From linear regression to linear prediction and beyond, 1998 (1998). doi:10.1007/978-94-011-5014-9_23.

[28] M. De Lozzo, Substitution de modèle et approche multifidélité en expérimentation numérique, Journal de la Société Française de Statistique 156 (4) (2015) 21–55 (2015).

[29] Z.-H. Han, R. Zimmermann, S. Görtz, A New Cokriging Method for Variable-Fidelity Surrogate Modeling of Aerodynamic Data, in: 48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, AIAA 2010-1225, Orlando,FL, 2010 (Jan. 2010).

[30] M. C. Kennedy, A. O'Hagan, Predicting the output from a complex computer code when fast approximations are available, Biometrika 87 (1) (2000) 1–13 (Mar. 2000). doi:10.1093/biomet/87.1.1.

[31] L. Le Gratiet, Bayesian Analysis of Hierarchical Multifidelity Codes, SIAM/ASA Journal on Uncertainty Quantification 1 (1) (2013) 244–269 (Jan. 2013). doi:10.1137/120884122.

[32] Z.-H. Han, S. Görtz, Hierarchical kriging model for variable-fidelity surrogate modeling, AIAA journal 50 (9) (2012) 1885–1896 (2012).

[33] N. Courrier, P.-A. Boucard, B. Soulier, Variable-fidelity modeling of structural analysis of assemblies, Journal of Global Optimization 64 (3) (2016) 577–613 (Mar. 2016). doi:10.1007/s10898-015-0345-9.

[34] V. Picheny, T. Wagner, D. Ginsbourger, A benchmark of kriging-based infill criteria for noisy optimization, Structural and Multidisciplinary Optimization 48 (3) (2013) 607–626 (Sep. 2013). doi:10.1007/s00158-013-0919-4.

[35] J. Lemaitre, J.-L. Chaboche, Mechanics of Solid Materials, Cambridge university press, 1994 (1994).

[36] N. Relun, D. Néron, P. A. Boucard, A model reduction technique based on the PGD for elastic-viscoplastic computational analysis, Computational Mechanics 51 (1) (2013) 83–92 (Jan. 2013). doi:10.1007/s00466-012-0706-x.

[37] P. Ladevèze, A. Nouy, O. Loiseau, A multiscale computational approach for contact problems, Computer Methods in Applied Mechanics and Engineering 191 (43) (2002) 4869–4891 (Sep. 2002). doi:10.1016/S0045-7825(02)00406-1.

[38] L. Champaney, J. Y. Cognard, D. Dureisseix, P. Ladevèze, Large scale applications on parallel computers of a mixed domain decomposition method, Computational Mechanics 19 (4) (1997) 253–263 (Mar. 1997). doi:10.1007/s004660050174.

[39] D. Dureisseix, P. Ladeveze, D. Neron, B. Schrefler, A Multi-Time-Scale Strategy for Multiphysics Problems: Application to Poroelasticity, International Journal for Multiscale Computational Engineering 1 (4) (2003). doi:10.1615/IntJMultCompEng.v1.i4.50.

[40] O. Allix, P. Ladevèze, Interlaminar interface modelling for the prediction of delamination, Composite Structures 22 (4) (1992) 235–242 (Jan. 1992). doi:10.1016/0263-8223(92)90060-P.

[41] H. Lemoussu, P. A. Boucard, P. Ladevèze, A 3D shock computational strategy for real assembly and shock attenuator, Advances in Engineering Software 33 (7) (2002) 517–526 (Jul. 2002). doi:10.1016/S0965-9978(02)00074-1.

[42] L. Champaney, Une nouvelle approche modulaire pour l'analyse d'assemblages de structures tridimentionnelles, Ph.D. thesis, ENS-Cachan, Cachan, France (1996).

[43] P. Ladevèze, Sur une famille d'algorithmes en mécanique des structures, Comptes-rendus des séances de l'Académie des sciences. Série 2, Mécanique-physique, chimie, sciences de l'univers, sciences de la terre 300 (2) (1985) 41–44 (1985).

[44] F. Chinesta, P. Ladeveze, E. Cueto, A Short Review on Model Order Reduction Based on Proper Generalized Decomposition, Archives of Computational Methods in Engineering 18 (4) (2011) 395 (Oct. 2011). doi:10.1007/s11831-011-9064-7.

[45] A. Dumon, C. Allery, A. Ammar, Proper general decomposition (PGD) for the resolution of Navier–Stokes equations, Journal of Computational Physics 230 (4) (2011) 1387–1407 (Feb. 2011). doi:10.1016/j.jcp.2010.11.010.

[46] J.-C. Passieux, J.-N. Périé, High resolution digital image correlation using proper generalized decomposition: PGD-DIC, International Journal for Numerical Methods in Engineering 92 (6) (2012) 531–550 (2012). doi:10.1002/nme.4349.

[47] C. Heyberger, P. A. Boucard, D. Néron, A rational strategy for the resolution of parametrized problems in the PGD framework, Computer Methods in Applied Mechanics and Engineering 259 (2013) 40–49 (Jun. 2013). doi:10.1016/j.cma.2013.03.002.

[48] C. Paillet, D. Néron, P. Ladevèze, A door to model reduction in high-dimensional parameter space, Comptes Rendus Mécanique 346 (7) (2018) 524–531 (Jul. 2018). doi:10.1016/j.crme.2018.04.009.

[49] M. Capaldo, P. A. Guidault, D. Néron, P. Ladevèze, The Reference Point Method, a "hyperreduction" technique: Application to PGD-based nonlinear model reduction, Computer Methods in Applied Mechanics and Engineering 322 (2017) 483–514 (Aug. 2017). doi:10.1016/j.cma.2017.04.033.

[50] F. Schmidt, N. Pirc, M. Mongeau, F. Chinesta, Efficient mold cooling optimization by using model reduction, International Journal of Material Forming 4 (1) (2011) 73–82 (Mar. 2011). doi:10.1007/s12289-010-0988-5.

[51] A. Nouy, A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations, Computer Methods in Applied Mechanics and Engineering 196 (45) (2007) 4521–4537 (Sep. 2007). doi:10.1016/j.cma.2007.05.016.

[52] C. Ghnatios, F. Masson, A. Huerta, A. Leygue, E. Cueto, F. Chinesta, Proper Generalized Decomposition based dynamic data-driven control of thermal processes, Computer Methods in Applied Mechanics and Engineering 213-216 (2012) 29–41 (Mar. 2012). doi:10.1016/j.cma.2011.11.018.

[53] F. Bordeu, F. Chinesta, A. Leygue, E. Cueto, S. Niroomandi, Réduction de modèle par PGD appliqué la simulation en temps réel de solide déformables, in: 10e Colloque National En Calcul Des Structures, Giens, France, 2011 (2011).

[54] A. Nouy, A priori model reduction through Proper Generalized Decomposition for solving time-dependent partial differential equations, Computer Methods in Applied Mechanics and Engineering 199 (23-24) (2010) 1603–1626 (2010). doi:10.1016/j.cma.2010.01.009.

[55] M. Bhattacharyya, A. Fau, U. Nackenhorst, D. Néron, P. Ladevèze, A multi-temporal scale model reduction approach for the computation of fatigue damage, Computer Methods in Applied Mechanics and Engineering 340 (2018) 630–656 (Oct. 2018). doi:10.1016/j.cma.2018.06.004.

[56] P.-A. Boucard, L. Champaney, Approche multirésolution pour l'étude paramétrique d'assemblages par contact et frottement., Revue Européenne des Eléments Finis 13 (5/7) (2004) 437–448 (2004).

[57] M. Vitse, D. Néron, P.-A. Boucard, Virtual charts of solutions for parametrized nonlinear equations, Computational Mechanics 54 (6) (2014) 1529–1539 (Dec. 2014). doi:10.1007/s00466-014-1073-6.

[58] M. Blanchard, Méthode global/local non-intrusive pour les simulations cycliques non-linéaires, Ph.D. thesis, Université Paris-Saclay (Jan. 2018).

[59] A. Courard, D. Néron, P. Ladevèze, L. Ballere, Integration of PGD-virtual charts into an engineering design process, Computational Mechanics 57 (4) (2016) 637–651 (Apr. 2016). doi:10.1007/s00466-015-1246-y.

[60] L. Noël, L. V. Miegroet, P. Duysinx, Analytical sensitivity analysis using the extended finite element method in shape optimization of bimaterial structures, International Journal for Numerical Methods in Engineering 107 (8) (2016) 669–695 (2016). doi:10.1002/nme.5181.

[61] T. Hughes, J. Cottrell, Y. Bazilevs., Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement., Computer Methods in Applied Mechanics and Engineering 194 (39-41) (2005) 4135–4195 (2005).

[62] M. Tirvaudey, R. Bouclier, J.-C. Passieux, L. Chamoin, Non-invasive implementation of nonlinear Isogeometric Analysis in an industrial FE software, Engineering Computations In press (2019).