



HAL
open science

Generation of a cokriging metamodel using a multiparametric strategy

Luc Laurent, Pierre-Alain Boucard, Bruno Soulier

► **To cite this version:**

Luc Laurent, Pierre-Alain Boucard, Bruno Soulier. Generation of a cokriging metamodel using a multiparametric strategy. *Computational Mechanics*, 2013, 51 (2), pp.151-169. 10.1007/s00466-012-0711-0 . hal-01376462

HAL Id: hal-01376462

<https://hal.science/hal-01376462>

Submitted on 16 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Generation of a Cokriging Metamodel Using a Multiparametric Strategy

Luc LAURENT · Pierre-Alain BOUCARD · Bruno SOULIER

Received: date / Accepted: date

Abstract In the course of designing structural assemblies, performing a full optimization is very expensive in terms of computation time. In order to reduce this cost, we propose a multilevel model optimization approach. This paper lays the foundations of this strategy by presenting a method for constructing an approximation of an objective function. This approach consists in coupling a multiparametric mechanical strategy based on the LATIN method with a gradient-based metamodel called a cokriging metamodel. The main difficulty is to build an accurate approximation while keeping the computation cost low. Following an introduction to multiparametric and cokriging strategies, the performance of kriging and cokriging models is studied using one- and two-dimensional analytical functions; then, the performance of metamodels built from mechanical responses provided by the multiparametric strategy is analyzed based on two mechanical test examples.

Keywords Cokriging metamodel · multiparametric strategy · LATIN method · multilevel optimization

Contents

1	Introduction	1
2	The optimization process	2
3	The multiparametric strategy	2
3.1	The problem of structural assemblies	3
3.2	The LATIN method for assemblies	3
3.3	The multiparametric strategy	3
4	The cokriging metamodel	5
4.1	Notations	5
4.2	Principle	5
4.3	The correlation function	6
4.4	Construction of the cokriging metamodel	7

L. LAURENT · P.-A. BOUCARD · B. SOULIER
LMT-Cachan, (ENS Cachan/CNRS/Université Paris 6/PRES Univer-
Sud Paris), 61 av. du Président Wilson, F-94230 Cachan, France
E-mail: {laurent,boucard,soulier}@lmt.ens-cachan.fr

4.5	Ordinary CoKriging (OCK)	7
4.6	Estimation of the parameters	8
5	Examples of the construction of metamodels	8
5.1	Analytical applications	8
5.2	Application to a contact problem with friction	12
5.3	Application to a shrink-fit problem	13
6	Conclusion	15
A	Complementary figures	17

1 Introduction

Optimization techniques have been in use in industry for many years. Structural elements, geometries, materials... are often designed and selected by means of specific industrial optimization software. The main difficulty in carrying out an optimization is that the precise localization of the optimum of an objective function often requires a large number of calculations. In the case of the design of structural assemblies, each calculation involves the resolution of very complex nonlinear problems due to the existence of contacts and friction between the parts. This makes the computation cost a major scientific stumbling block.

A technique widely used in order to make such optimizations affordable is multilevel model optimization [1]. Our approach belongs to that category in the sense that, contrary to the surrogate-based approach [2, 3], we use a metamodel only to accelerate the optimization process. The main steps are: first, the construction of a metamodel from the responses of the mechanical model; then, the localization of an approximate optimum using the metamodel; and, finally, the precise determination of the optimum using the mechanical model and the approximate optimum. This strategy leads to a reduction in computation cost and, therefore, in the time it takes to complete the optimization.

This paper focuses on the first step, *i.e.* the construction of the metamodel using:

- a cokriging metamodel;
- a dedicated computational strategy.

We will show that when gradients are available a cokriging metamodel leads to a better approximation than a kriging metamodel with the same number of sample points. Moreover, our computational strategy is capable of reusing previous calculations in order to solve new problems. When the parameters vary, this property enables one to reduce the cost of a new calculation, particularly when it comes to evaluating gradients using a finite difference method.

The first chapter of the paper is a review of multilevel optimization techniques in the context of this study. More precisely, it focuses on multilevel model optimization intended for parameter optimization. The second chapter introduces the multiparametric strategy and discusses its performance in detail. The third chapter proposes and develops a gradient-based formulation of the cokriging metamodel.

In the last chapter, numerical and analytical examples are used to show that the cokriging metamodel provides a better approximation than classical (e.g. kriging) metamodels. For mechanical examples, the coupling of a cokriging metamodel with our computational strategy leads to a significant reduction in the computation costs associated with the generation of the metamodel. The quality of the cokriging metamodel is studied based on several classical criteria and compared with that of the kriging metamodel.

2 The optimization process

Classical direct optimization, in which the optimizer is linked directly to the mechanical solver, requires huge computer resources. The cost increases not only with the complexity of the mechanical problem (*i.e.* the nonlinearities included in the model), but also with the number of degrees of freedom in the mechanical system and the number of design variables in the optimization problem. In order to reduce the cost of optimization in structural design, many works concentrate on three main aspects: the development of dedicated numerical methods to address increasingly complex mechanical problems, the improvement of the performance of the optimization algorithms and the development of strategies to coordinate the exchanges between the optimization algorithm and the mechanical solver (for example, see [3, 4]). The use of a multilevel optimization strategy provides a solution to the last aspect. Multilevel strategies were developed over the last twenty years. These approaches can be divided into two main categories: multilevel parameter optimization (which includes sequential methods [5–7] and iterative methods [8–13]) and multilevel model optimization (which includes hierarchical multilevel optimization [1, 14] and imbricated multilevel optimization [15]). Similar strategies have also been used in the context

of multidisciplinary optimization where each level of the optimization process concerns a specific discipline [9, 16]. Parallel multilevel model optimization [17, 18] can also be used in order to reduce the optimization cost and take maximum advantage of parallel computing architectures. This type of multilevel strategy calls for specific optimization algorithms, such as genetic algorithms [19–21].

The multilevel modeling strategy we propose is based on multilevel model optimization. We focus on the two-level strategy illustrated in Figure 1. On the first level, which consists in a metamodel defined using a limited amount of data, the zones where the optimum is to be sought are determined using, for example, a genetic algorithm [22]; then, this information is transferred to the second level, where a precise search for the minimum can be carried out using a gradient-based algorithm. Thus, the second level is an optimization process based on the full mechanical model.

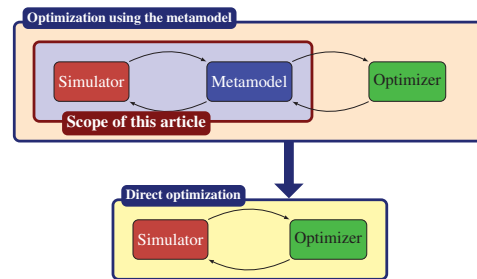


Fig. 1: Multilevel optimization

The scope of this paper is limited to the study of the cost associated with the generation of the metamodel.

Many types of metamodels can be used to find an approximate solution: they can be based on polynomial regression [23, 24], on neural networks [25, 26], on radial basis functions [27], on proper orthogonal decomposition [28], on cumulative interpolation [29], etc... We chose to use a particular class of metamodels called kriging approximations [30] and, more precisely, a cokriging metamodel [31] using derivatives [32]. The interested readers could also refer to works on metamodeling strategies in structural optimization and multidisciplinary design optimization [33, 34].

3 The multiparametric strategy

This chapter introduces the concept of multiparametric strategy, which enables one to solve similar structural problems at a greatly reduced cost. This is used for the types of structural assembly problems which are described in Section 3.1. The strategy is based on the LATIN method, which is presented in Section 3.2. The multiparametric strategy itself is

presented in Section 3.3 and its performance is discussed using the example of a 2D mechanical benchmark problem with two design variables in Section 3.3.2.

3.1 The problem of structural assemblies

The structures being considered in this paper are assemblies of linear elastic structures under the assumption of small perturbations. The only nonlinearities occur between parts of the assemblies and are due to contact and/or friction phenomena. In order to solve these problems, we use a dedicated strategy based on the LATIN algorithm introduced by P. Ladevèze [35].

This strategy is based on three main points:

- The structure being studied is divided into substructures and interfaces. This is a natural approach in the context of assemblies because treating each part of an assembly as a substructure and the contact zone(s) between two parts as (an) interface(s) constitutes the simplest decomposition. What makes this decomposition unique is the use of mixed force and displacement unknowns at the interfaces;
- A suitable iterative algorithm is used to solve the mechanical problem;
- The operators of the method remain constant and do not depend on the loading or on the parameters of the interface (friction coefficient, gap).

The resulting approach is a mixed domain decomposition method, as opposed to the primal substructuring approach [36, 37] or dual approach [38].

3.2 Short summary of the LATIN method for assemblies

Our resolution strategy is based on the LATIN algorithm developed by P. Ladevèze [35], which consists in solving two problems alternatively: one in the substructures and one at the interfaces. These problems are described by two groups of equations denoted \mathbf{A}_d and Γ : \mathbf{A}_d contains the *linear* equations related to the substructures and Γ contains the *local* equations (which can be nonlinear) related to the interfaces. The iterations between the two groups are carried out through the use of search directions which are parameters of the method. This resolution process, shown in Figure 2, leads to the solution defined as the intersection of spaces Γ and \mathbf{A}_d . In the context of elastic assemblies, the boundary conditions (forces \mathbf{F}_i and displacements \mathbf{W}_i) at each interface i between two substructures are sufficient to define the solution: the internal behavior of each substructure is the solution of a classical elastic problem. Thus, an approximate solution is described entirely by $\mathbf{s} = \bigcup_{\text{all the interfaces}} (\mathbf{W}_i, \mathbf{F}_i)$.

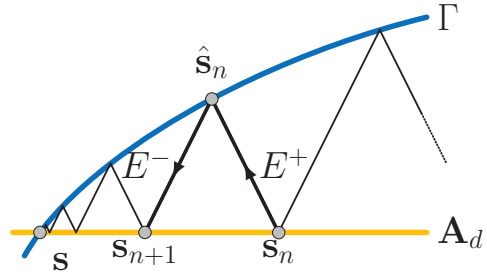


Fig. 2: Schematic representation of the LATIN process

The friction and contact conditions with their respective approaches and laws [39] are part of equation group Γ . The linear form of \mathbf{A}_d is due to the linear behavior of the substructures.

The main feature of the LATIN method is that at each iteration the resolution leads to an approximate solution over the whole loading path and in all the points of the structure. Each new iteration enriches this solution until convergence.

Additional details can be found in [35, 40] for the LATIN method and in [39, 41] for its application to assemblies.

3.3 The multiparametric strategy

The construction of the metamodel requires the evaluation of what is called in optimization an *objective function* or *cost function* at certain points of what is called the *design space*. In the case of a parametric study in the mechanical context, these evaluations lead to the resolution of many problems which are similar in the sense that only their parameters vary. In order to accelerate these resolutions, we use an approach called the MultiParametric Strategy (MPS) which was introduced in [42, 43], then studied and applied to various types of problems in [44–46]. The following presentation introduces the principle of this strategy and discusses its performance based on an academic test example.

3.3.1 Principle of the multiparametric strategy

The main idea of the multiparametric strategy is very simple and consists in initializing the LATIN algorithm using the converged solution of a previous problem. As mentioned in Section 3.2, each calculated solution, associated with a set of parameters, is entirely described by the values of the variables at the boundary. Thus, the initialization process consists in reloading the boundary values associated with the converged solution of the chosen similar problem.

In order to build a metamodel, the set of the sample points (each defined in the design space by a set of parameters) can be divided into two sets: the set of the sample points associated with calculated values of the objective function, and the set of the sample points at which the values of the

objective function are still unknown. The latter is not always defined clearly, especially in an optimization process in which the optimizer introduces new sample points gradually, but the first set constitutes a database which is available for subsequent calculations. Thus, if a new calculation associated with different values of the parameters is required, the algorithm can be initialized using a previous converged solution taken from the database. Each new evaluation of the objective function enriches the database.

Figure 3 illustrates the use of the multiparametric strategy to obtain an approximate solution starting from a previous calculation: in order to obtain the converged solution, s_3 can be initialized with one of the two available converged solutions s_1 or s_2 . The benefit depends on the choice of the starting point. Some indications for choosing the best initialization strategy were given in [47]. In our case we will use only a ‘‘closest point’’ strategy: the initial solution of the new problem is chosen to be the converged solution associated with the set of design parameters which is closest to the one being considered.

Since we are assuming that the only parameters which vary are the interface parameters, space Γ alone is affected by the change.

Thanks to this strategy, the solution converges in fewer iter-

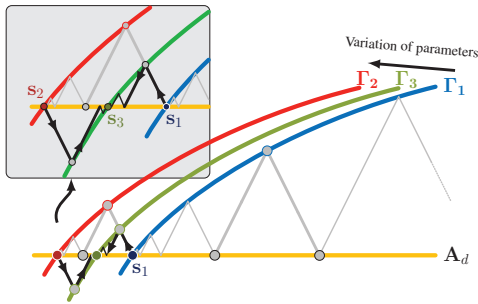


Fig. 3: The multiparametric strategy using the LATIN algorithm

ations of the algorithm and, therefore, in less time.

3.3.2 Performance of the multiparametric strategy

Let us consider the example of a quasi-static academic problem which was presented in [45]. Figure 4 shows the geometry of the problem, which consists of three square parts ($h = 50\text{mm}$, Young’s modulus $E = 2 \cdot 10^5\text{MPa}$ and Poisson’s coefficient $\nu = 0.3$) in contact with friction. Each part is represented by a single substructure discretized into 20×20 bilinear quadrangles. The parametric study consists in varying the friction coefficients μ_1 and μ_2 of the two contact interfaces. The loading consists of two stages: first, a progressive vertical pressure P_1 up to a maximum of 50MPa

applied at the top of substructure Ω_3 (the preloading stage), then a progressive horizontal load from 0 to 30MPa applied to substructure Ω_2 .

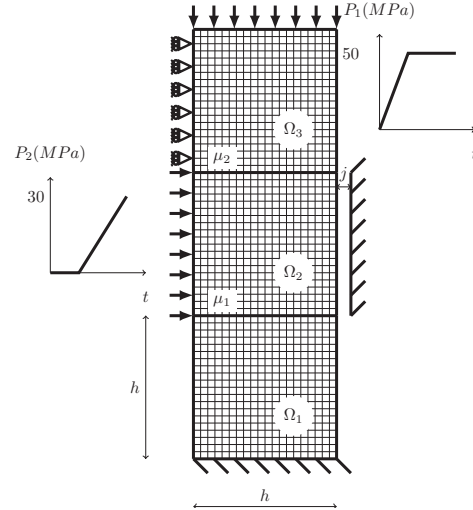


Fig. 4: The geometry of the problem

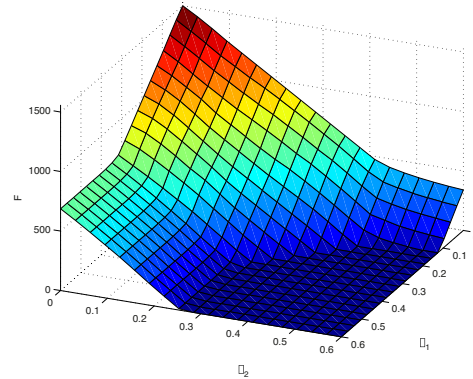


Fig. 5: The reference response surface

In this test, variations of the friction coefficients μ_1 and μ_2 between 0 and 0.6 were considered, and the function studied was the reaction force on the rigid wall. Figure 5 shows the response surface of this function obtained with 18×18 values of the friction coefficients.

In order to illustrate the performance of the multiparametric strategy, the two-variable design space was sampled with a 4×4 regular grid. For each sample, the force and its gradients were calculated (using a classical finite difference method for the gradients). Table 1 summarizes the characteristics of these calculations.

The gain obtained with our method compared to a classical calculation without the MPS was estimated using the

expression:

$$\text{Gain} = \frac{\text{CPU time without the MPS}}{\text{CPU time using the MPS}} \quad (1)$$

The simple mechanical examples used in this paper make it possible to evaluate the CPU time without the MPS, but in the general case one uses the formula:

$$\text{Gain} = \frac{\text{Number of calculations} \times \text{Reference CPU time}}{\text{CPU time using the MPS}} \quad (2)$$

where the *Reference CPU time* can be the CPU time of the first calculation, the maximum CPU time necessary to obtain one evaluation, or the average CPU time for evaluations at several points of the design space without the MPS.

	Eval.	Eval. gradients	Total
Number of calculations	16	2×16	48
Avg. CPU time per calc.	2.97s	1.53s	2.01s
Min. CPU time	0.17s	0.16s	0.16s
Max. CPU time	8.12s	4.83s	8.12s
Nb. of iterations	2,201	2,215	4,416
Avg. nb. of iterations (per calc.)	137	69.2	92
Min. nb. of iterations	5	5	5
Max. nb. of iterations	381	225	381
Total CPU time	47.5s	48.9s	96.4s
CPU time w/o MPS	135s	270s	405s
Gain	2.84	5.52	4.22

Table 1: The mechanical calculations using the multiparametric strategy

The results presented in Table 1 show that the solver can indeed accelerate the resolution of similar problems by reusing a previously converged solution. The most remarkable gain was that obtained in the evaluation of the gradients. These were calculated using a finite difference method: each gradient required the value of the response in 3 points (a sample point plus 2 neighboring points obtained with very small variations of the parameters). The results show that for the same number of evaluations the cost was lower when both the responses and the gradients were calculated than when the responses alone were calculated: in our example, 16 evaluations of the responses alone took 47.5s, but the addition of 16 gradients (32 evaluations) took only 48.9s. Thus, thanks to the ability of the method to reduce the computation time for sample points that are close together, the strategy is particularly efficient in calculating the gradients. This feature is very interesting if one wishes to use a gradient-based optimizer to achieve a significant reduction in optimization time. In this paper, we will also undertake to use these inexpensive gradients to build a gradient-enhanced metamodel.

4 The cokriging metamodel

The classical approach to building a metamodel consists in using the responses of an objective function (also called a *primary variable*) calculated at a number of sample points. The values of the objective function at the chosen sample points can be used to build a classical metamodel. However, in many cases, it may be interesting to use additional information in the form of *auxiliary variables* to build a richer metamodel. Multivariate geostatistics is a field of applied mathematics which supplies methods to handle these variables jointly. The cokriging metamodel we use [31] is an example of such methods in which the primary variable is the objective function and the auxiliary variables are its gradients.

The kriging technique was first introduced by D.G. Krige, a mining engineer [48]. Subsequently, many mathematical enhancements were proposed by G. Matheron [49–51]. Used initially in geostatistics, kriging was later coupled with calculation methods for the resolution of design problems by [52, 53]. Simpson [24] gave a review of various metamodels for multidisciplinary optimization. Additional information can be found in [54].

This chapter focuses on the gradient-enhanced cokriging metamodel [32, 55, 56]: Section 4.2 presents the principles and fundamentals of kriging and cokriging; then, Sections 4.4 and 4.5 deal with the construction of the cokriging metamodel in the case of ordinary cokriging. Finally, Section 4.6 discusses the problem of the determination of the cokriging parameters.

4.1 Notations

We will use the following notations:

1. $\mathbf{x}^{(i)}$, $i \in \llbracket 0, n_s \rrbracket$ denotes a point in the design space \mathcal{D} . $(\mathbf{x}^{(i)}, i \in \llbracket 1, n_s \rrbracket)$ is one of the n_s sample points, and $\mathbf{x}^{(0)}$ is an arbitrary point in the design space, which may or may not be a sample point.)
2. $Y(\mathbf{x}^{(i)})$ and $\tilde{Y}(\mathbf{x}^{(i)})$ denote respectively the response of the analytical function (or the response of the mechanical model) and the approximate response given by the metamodel at point $\mathbf{x}^{(i)}$, $i \in \llbracket 0, n_s \rrbracket$.
3. $R(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})$ is a correlation function expressing the correlation relation between points $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(j)}$, $(i, j) \in \llbracket 0, n_s \rrbracket^2$.

4.2 Principle

The principle of the cokriging metamodel is similar to that of the kriging metamodel. One defines a random process, associated with the deterministic response of the objective function, which is the sum of two components (Equation 3):

a linear model μ which represents the trend of the data, and a departure Z from this linear model which represents the fluctuations around the trend.

$$Y(\mathbf{x}^{(0)}) = \mu(\mathbf{x}^{(0)}) + Z(\mathbf{x}^{(0)}), \forall \mathbf{x}^{(0)} \in \mathcal{D} \quad (3)$$

with

$$\mathbb{E} [Z(\mathbf{x}^{(0)})] = 0, \text{Var} [Z(\mathbf{x}^{(0)})] \neq 0, \forall \mathbf{x}^{(0)} \in \mathcal{D} \quad (4)$$

and

$$\text{cov} [Z(\mathbf{x}^{(i)}), Z(\mathbf{x}^{(j)})] \neq 0, \forall (\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) \in \mathcal{D}^2 \quad (5)$$

\mathbb{E} , Var and cov are the classical statistical expected value, variance and covariance.

μ is a deterministic function and Z is a stationary Gaussian process with a known stationary covariance. Depending on the definition of function μ , one can build different types of kriging or cokriging metamodels (simple kriging, where μ is the average of the values of the objective function at the sample points; ordinary kriging, where μ is an unknown constant; or universal kriging, where μ is a polynomial function).

Function Z has a zero expected value and its covariance structure is a function of a generalized distance among the sampled responses. The covariance structure can be written as:

$$\begin{aligned} \forall (\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) \in \mathcal{D}^2, \\ \text{cov} [Z(\mathbf{x}^{(i)}), Z(\mathbf{x}^{(j)})] = \sigma^2 R(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = \sigma^2 c_{ij} \end{aligned} \quad (6)$$

where σ^2 is the variance of process Z . The correlation function can be a Gaussian function or a Matérn function [57, 58].

In the context of cokriging, the variables are divided into a primary variable Y and N auxiliary variables W^i . The auxiliary variables used to build a gradient-based cokriging metamodel are the components of the gradients (Equation 8). The construction of the cokriging metamodel involves Equation 7 in addition to Equation 3.

$$\forall i \in \llbracket 0, N \rrbracket, \forall \mathbf{x}^{(0)} \in \mathcal{D}, W^i(\mathbf{x}^{(0)}) = \mu_{W^i} + Q^i(\mathbf{x}^{(0)}) \quad (7)$$

where

$$\forall i \in \llbracket 0, N \rrbracket, \forall \mathbf{x}^{(0)} \in \mathcal{D}, W^i(\mathbf{x}^{(0)}) = \frac{\partial Y}{\partial x_i}(\mathbf{x}^{(0)}) \quad (8)$$

and

$$\forall i \in \llbracket 0, N \rrbracket, \forall \mathbf{x}^{(0)} \in \mathcal{D}, \mathbb{E} [W^i(\mathbf{x}^{(0)})] = \mu_{W^i} \quad (9)$$

$$\mathbb{E} [Q^i(\mathbf{x}^{(0)})] = 0 \quad (10)$$

$$\forall (\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) \in \mathcal{D}^2, \forall (k, l) \in \llbracket 0, N \rrbracket^2, \quad (11)$$

$$\text{cov} [Q^k(\mathbf{x}^{(i)}), Q^l(\mathbf{x}^{(j)})] \neq 0 \quad (12)$$

$$\text{cov} [Q^k(\mathbf{x}^{(i)}), Z(\mathbf{x}^{(j)})] \neq 0 \quad (13)$$

$$\text{Var} [Q^k(\mathbf{x}^{(i)})] = \text{Var} \left[\frac{\partial Z}{\partial x_k}(\mathbf{x}^{(i)}) \right] \quad (14)$$

Thus, in the case of a cokriging metamodel, additional covariance relations involving the different variables must be introduced [56]:

$$\begin{aligned} \text{cov} \left[\frac{\partial Z}{\partial x_k}(\mathbf{x}^{(i)}), Z(\mathbf{x}^{(j)}) \right] &= -\sigma^2 \frac{\partial R}{\partial x_k}(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) \\ &= \sigma^2 c_{kij} \end{aligned} \quad (15)$$

$$\begin{aligned} \text{cov} \left[Z(\mathbf{x}^{(i)}), \frac{\partial Z}{\partial x_k}(\mathbf{x}^{(j)}) \right] &= \sigma^2 \frac{\partial R}{\partial x_k}(\mathbf{x}^{(j)}, \mathbf{x}^{(i)}) \\ &= \sigma^2 c_{kji} \end{aligned} \quad (16)$$

$$\begin{aligned} \text{cov} \left[\frac{\partial Z}{\partial x_k}(\mathbf{x}^{(i)}), \frac{\partial Z}{\partial x_l}(\mathbf{x}^{(j)}) \right] &= -\sigma^2 \frac{\partial^2 R}{\partial x_k \partial x_l}(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) \\ &= \sigma^2 c_{klij} \end{aligned} \quad (17)$$

where

$$\forall (i, j, k, l) \in \llbracket 0, n_s \rrbracket^2 \times \llbracket 1, n_d \rrbracket^2, \forall (\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) \in \mathcal{D}^2,$$

$$c_{ij} = R(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = c_{ji} \quad (18)$$

$$c_{kij} = \frac{\partial R}{\partial x_k}(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = -\frac{\partial R}{\partial x_k}(\mathbf{x}^{(j)}, \mathbf{x}^{(i)}) = -c_{kji} \quad (19)$$

$$c_{klij} = \frac{\partial^2 R}{\partial x_k \partial x_l}(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = c_{lkij} = c_{klji} \quad (20)$$

4.3 The correlation function

In the following examples, we consider the correlation function to be the Matérn function [57], defined by:

$\forall r \in \mathbb{R}^+$,

$$\mathcal{M}(r, l) = \left(1 + \frac{\sqrt{5}}{l} r + \frac{5}{3l^2} r^2 \right) \exp \left(-\frac{\sqrt{5}}{l} r \right) \quad (21)$$

where l denotes the correlation length ($l > 0$). More details on the Matérn function can be found in [54].

In the general case, we use the correlation function defined by:

$$\forall (i, j, k, l) \in \llbracket 0, n_s \rrbracket^2 \times \llbracket 1, n_d \rrbracket^2, \forall (\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) \in \mathcal{D}^2,$$

$$R(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = \prod_{k=1}^{n_d} \mathcal{M}(|x_k^{(i)} - x_k^{(j)}|, l_k) \quad (22)$$

where n_d is the number of design variables in the problem; $x_k^{(i)}$ and l_k are the k^{th} component of point $\mathbf{x}^{(i)}$ and the associated correlation length respectively.

The first derivative of the Matérn function is non-symmetrical. Therefore, one has $c_{kij} \neq c_{kji}$ with $(k, i, j) \in \llbracket 1, n_d \rrbracket \times \llbracket 0, n_s \rrbracket^2$ (see Equation 19).

4.4 Construction of the cokriging metamodel

The method used to build the cokriging metamodel is similar to that used to build a kriging metamodel, which leads to an estimator known as the *best linear unbiased predictor* as follows:

1. The objective is to determine an estimator \tilde{Y} of the random process Y given the linear predictor defined by Equation 23 with $\mathbf{x}^{(0)} \in \mathcal{D}$; this is equivalent to determining the λ_{0i} ($\forall i \in \llbracket 1, n_s \rrbracket$) and the λ_{ij} ($\forall (i, j) \in \llbracket 1, n_s \rrbracket \times \llbracket 1, n_d \rrbracket$);
2. the linear predictor must satisfy the unbiasedness conditions of Equation 24;
3. the linear predictor must minimize the mean square error defined in Equation 25 subject to the previous unbiasedness conditions.

$$\tilde{Y}(\mathbf{x}^{(0)}) = \sum_{i=1}^{n_s} \lambda_{0i}(\mathbf{x}^{(0)}) Y(\mathbf{x}^{(i)}) + \sum_{j=1}^{n_d} \sum_{i=1}^{n_s} \lambda_{ij}(\mathbf{x}^{(0)}) \frac{\partial Y}{\partial x_j}(\mathbf{x}^{(i)}) \quad (23)$$

$$\mathbb{E}[\tilde{Y}(\mathbf{x}^{(0)})] = \mathbb{E}[Y(\mathbf{x}^{(0)})] \quad (24)$$

$$\text{MSE}[\tilde{Y}(\mathbf{x}^{(0)})] = \mathbb{E}\left[\left(\tilde{Y}(\mathbf{x}^{(0)}) - Y(\mathbf{x}^{(0)})\right)^2\right] \quad (25)$$

4.5 Ordinary CoKriging (OCK)

For simplicity's sake, as suggested in [19], we chose to limit ourselves to an ordinary cokriging metamodel. In this case, the deterministic function μ is an unknown constant. The construction method of Section 4.4 leads to the following estimator (from Equation 26)

$$\tilde{Y}(\mathbf{x}^{(0)}) = \underbrace{\hat{\beta}_c}_{\mu(\mathbf{x}^{(0)})} + \underbrace{\mathbf{r}_{c0}^T \mathbf{C}_c^{-1} (\mathbf{Y}_{sc} - \mathbf{X}_c \hat{\beta}_c)}_{z(\mathbf{x}^{(0)})} \quad (26)$$

where

$$\begin{aligned} \hat{\beta}_c &= (\mathbf{X}_c^T \mathbf{C}_c^{-1} \mathbf{X}_c)^{-1} \mathbf{X}_c^T \mathbf{C}_c^{-1} \mathbf{Y}_{sc} \\ \mathbf{r}_{c0} &= [c_{01} \ c_{02} \ \cdots \ c_{0n_s} \ c_{110} \ c_{210} \ \cdots \ c_{n_d n_s 0}] \\ \mathbf{X}_c &= [1, 1, \dots, 1, 0, 0, \dots, 0]^T \\ \mathbf{C}_c &= \begin{bmatrix} \mathbf{C} & \mathbf{C}_d \\ \mathbf{C}_d^T & \mathbf{C}_{dd} \end{bmatrix} \end{aligned}$$

and, using the notations presented in Equations 18-19-20:

$$\begin{aligned} (\mathbf{C})_{ij} &= c_{ij} \\ \mathbf{C}_d &= \begin{bmatrix} c_{111} & c_{211} & \cdots & c_{n_d 11} & c_{112} & \cdots & c_{n_d 1n_s} \\ c_{121} & c_{221} & \cdots & c_{n_d 21} & c_{122} & \cdots & c_{n_d 2n_s} \\ c_{131} & c_{231} & \cdots & & & & \\ \vdots & & \ddots & & & & \vdots \\ c_{1n_s 1} & \cdots & & & & & c_{n_d n_s n_s} \end{bmatrix} \\ \mathbf{C}_{dd} &= \begin{bmatrix} \mathbf{c}_{dd}^{11} & \mathbf{c}_{dd}^{12} & \cdots & \mathbf{c}_{dd}^{1n_s} \\ \mathbf{c}_{dd}^{21} & \mathbf{c}_{dd}^{22} & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{c}_{dd}^{n_s 1} & \mathbf{c}_{dd}^{n_s 2} & \cdots & \mathbf{c}_{dd}^{n_s n_s} \end{bmatrix} \end{aligned}$$

$$\text{with } \forall (k, l) \in \llbracket 1, n_s \rrbracket^2, \mathbf{c}_{dd}^{kl} = \begin{bmatrix} c_{11kl} & c_{12kl} & \cdots & c_{1n_d kl} \\ c_{11kl} & c_{12kl} & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ c_{n_d 1kl} & c_{n_d 2kl} & \cdots & c_{n_d n_d kl} \end{bmatrix}$$

In the cokriging case, vector \mathbf{Y}_{sc} contains both the responses and the gradients of the objective function at the sample points, whereas in the kriging case vector \mathbf{Y}_s contains only the responses of the function. Equations 15-16-17 lead to the linear estimator as a function of the correlation vector and matrices \mathbf{r}_{c0} and \mathbf{C}_c . Therefore, the resulting expression does not contain σ^2 .

Thus defined, the model can supply approximate responses of the objective function at every point of the design space. In our case, the response of the function is considered to be deterministic and we end up with a cokriging interpolation model. This type of metamodel has other advantages: for example, it provides statistical information on the process (its expected value and its variance). Due to the use of the unbiasedness condition, the expected value of \tilde{Y} is given by the trend model μ and the mean square error of \tilde{Y} [53]:

$$\text{MSE}[\tilde{Y}(\mathbf{x}^{(0)})] = \sigma^2 \left[1 - [\mathbf{1}_{n_s} \ \mathbf{r}_{c0}] \begin{bmatrix} \mathbf{0} & \mathbf{X}_c^T \\ \mathbf{X}_c & \mathbf{C}_c \end{bmatrix} \begin{bmatrix} \mathbf{1}_{n_s} \\ \mathbf{r}_{c0} \end{bmatrix} \right] \quad (27)$$

where $\mathbf{1}_{n_s}$ is a vector which contains $n_s \times 1$.

The correlation matrix and vector appear to depend on the correlation function, *i.e.* on the correlation lengths. A specific strategy can be used to estimate values of the correlation lengths and σ^2 .

4.6 Estimation of the parameters

The model's parameters (such as the characteristic correlation lengths \mathbf{l} , the variance σ of the random process Z or the regression coefficients) can be determined by maximizing the likelihood function [59]. We use this technique, which relies on the maximization of the density of the observed values \mathbf{Y}_{sc} , to determine \mathbf{l} and σ . The density can be viewed as a function L of parameters \mathbf{l} and σ :

$$L(\mathbf{l}, \sigma^2; \mathbf{Y}_{sc}) = (2\pi\sigma^2)^{-n_s(n_d+1)/2} |\mathbf{C}_c(\mathbf{l})|^{-1/2} \times \exp \left[-\frac{1}{2} (\mathbf{Y}_{sc} - \mathbf{X}_c \hat{\beta}_c)^T \mathbf{C}_c(\mathbf{l})^{-1} (\mathbf{Y}_{sc} - \mathbf{X}_c \hat{\beta}_c) \right] \quad (28)$$

where \mathbf{C}_c is the correlation matrix. This matrix depends on the correlation lengths $\{l_1, l_2, \dots, l_{n_d}\}$, which constitute a vector \mathbf{l} .

The maximization of the likelihood function can be expressed as:

$$(\hat{\mathbf{l}}, \hat{\sigma}^2) = \underset{\mathbf{l}, \sigma^2}{\operatorname{argmax}} \{L(\mathbf{l}, \sigma^2; \mathbf{Y}_{sc})\} \quad (29)$$

This problem can usually be solved by minimizing the log-likelihood.

The variance σ^2 can be determined analytically through the derivation of the likelihood function:

$$\hat{\sigma}^2 = \frac{1}{n_s} (\mathbf{Y}_{sc} - \mathbf{X}_c \hat{\beta}_c)^T \mathbf{C}_c(\mathbf{l})^{-1} (\mathbf{Y}_{sc} - \mathbf{X}_c \hat{\beta}_c) \quad (30)$$

One can also use an optimizer to determine the correlation length numerically. This method has some drawbacks [60–62]: in many cases with very few points, the log-likelihood is monotonous; the correlation matrix often suffers from conditioning problems which make it difficult to find a minimum. When such problems arise, one sets the parameters (particularly the correlation lengths) to fixed values.

5 Examples of the construction of metamodels

In this chapter, several examples based on analytical functions and mechanical test cases are presented: first, a one-dimensional analytical function is used to build metamodels and, in order to compare the quality of these metamodels, different criteria are presented in (Section 5.1.1). Then, kriging and cokriging metamodels are built using two two-dimensional analytical functions with and without anisotropy (Sections 5.1.2 and 5.1.3). Finally, two mechanical problems are used to study the performance of the coupled multiparametric/cokriging strategy approach (Sections 5.2 and 5.3). For simplicity's sake, only problems with two design variables are being considered.

5.1 Analytical applications

In this section, kriging and cokriging are applied to one- and two-dimensional analytical test functions. The abbreviations OK and OCK will be used to designate **O**rdinary **K**riging and **O**rdinary **C**okriging respectively.

5.1.1 Case of a one-dimensional test function

First, we applied the two types of metamodels to the analytical function $y(x) = \exp(-x/10) \cos(x) + x/10$. We used 6 sampled responses of the analytical function to build the OK metamodel, and an additional 6 sampled derivatives to build the OCK metamodel. The correlation function was the Matérn function. The sample points were obtained using Latin Hypercube Sampling (LHS) [63].

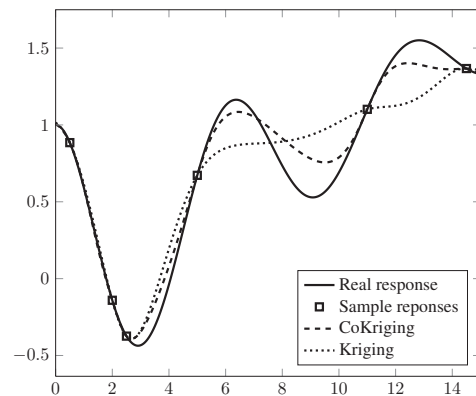


Fig. 6: Ordinary kriging and cokriging

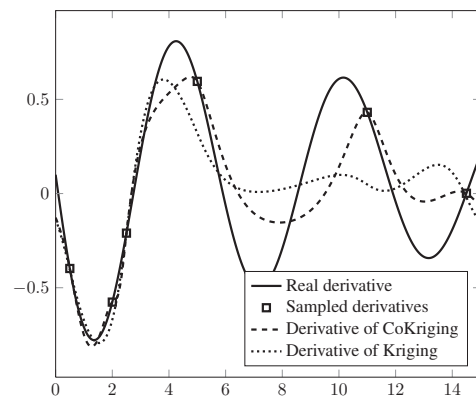


Fig. 7: OK and OCK derivatives

Figures 6 and 7 illustrate the capability of the cokriging metamodel to interpolate not only the values of the responses, but also the sampled derivatives. In this example, the cokriging metamodel performed better: Table 2 shows

that the R^2 criterion is better with cokriging than with kriging and the other criteria confirm that in this case cokriging leads to a better approximation than kriging. This statement remained true as long as we worked with only a few points. One can also observe that for a relatively smooth function the kriging metamodel converges quickly toward a good approximation of the actual function when the number of points becomes large enough.

For these two metamodels the correlation length was determined by maximizing the likelihood following the strategy introduced in Section 4.6. Figure 8 shows the log-likelihood as a function of the correlation length.

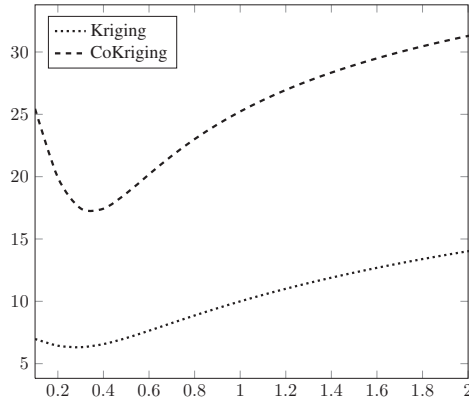


Fig. 8: The log-likelihood vs. the correlation length

Classical numerical optimization (using function `fmincon` in MATLAB) led to correlation lengths equal to $l_{OK} = 0.2807$ and $l_{OCK} = 0.3481$ for the kriging and cokriging metamodels respectively. (In all the kriging and cokriging examples presented in this paper, the data are normalized in order to build the metamodels from reduced variables.)

The characteristics of the two metamodels are summarized in Table 2.

	OK	OCK
R^2	0.8593	0.9666
RAAE	0.2949	0.1326
RMAE	0.7162	0.4427
Q_1	$7.219 \cdot 10^{-2}$	$2.758 \cdot 10^{-2}$
Q_2	5.942	1.4124
Q_3	$1.974 \cdot 10^{-2}$	$4.693 \cdot 10^{-3}$

Table 2: Characteristics of the two previous metamodels

RAAE and RMAE designate the Relative Average Absolute Error and Relative Maximum Absolute Error criteria. The criteria Q_i , which compare the actual response and the responses of the metamodels at n_c points ($n_c \gg n_s$) of a

regular grid with were calculated as follows:

$$Q_1 = \sup_{i \in \{1, 2, \dots, n_c\}} e_i = \|\mathbf{e}\|_\infty$$

$$Q_2 = \sum_{i=1}^{n_c} e_i = \|\mathbf{e}\|_1$$

$$Q_3 = \frac{Q_2}{n_c} \text{ (the average of } Q_2\text{)}$$

$$\text{where } \forall i \in \{1, 2, \dots, n_c\}, e_i = \frac{\left(Y(\mathbf{x}^{(i)}) - \tilde{Y}(\mathbf{x}^{(i)})\right)^2}{\sup_{j \in \{1, 2, \dots, n_c\}} Y(\mathbf{x}^{(j)})^2}$$

$$\text{and } \mathbf{e} = [e_1 \ e_2 \ \dots \ e_{n_c}]$$

with $\|\cdot\|_1$ and $\|\cdot\|_\infty$ being the L^1 norm and infinity norm respectively.

Based on the statistical information obtained from the kriging and cokriging metamodels, one can derive confidence intervals. The two diagrams of Figures 9 and 10 show the 95% Confidence Intervals (CIs) obtained with Expression 31. In these types of metamodels, the size of the confidence envelopes is determined mainly by the distance between each pair of neighboring points.

$$CI^\pm(\mathbf{x}) = \tilde{Y}(\mathbf{x}) \pm 2\sqrt{MSE[\tilde{Y}(\mathbf{x})]} \quad (31)$$

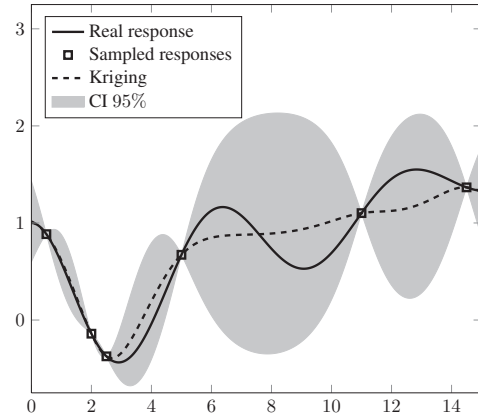


Fig. 9: The confidence envelopes for the kriging metamodel

For the same number of sample points, the cokriging metamodel provides narrower confidence intervals, especially close to the sample points. One could use the information derived from the variance or the confidence interval of random process \tilde{Y} to choose additional sample points (e.g. the points where the variance is maximum) in order to enrich the database used in constructing the metamodels. Such a strategy is not considered in this paper.

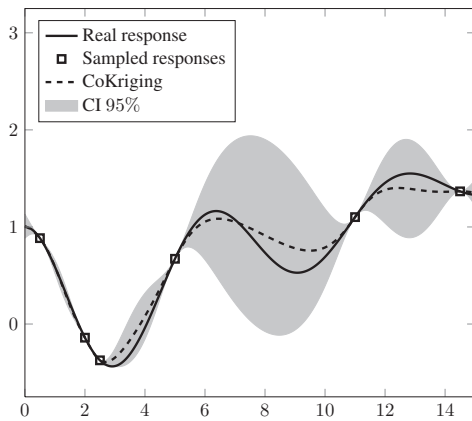


Fig. 10: The confidence envelopes for the cokriging metamodel

5.1.2 Case of a two-dimensional test function

The same two types of metamodels were used to approximate an analytical function of two variables. In order to illustrate the performance of the cokriging metamodel, we chose a very irregular function: the *six-hump camel back function* ($\forall (x_1, x_2) \in [-2, 2] \times [-1, 1]$, $f(x_1, x_2) = (4 - 2.1x_1^2 + x_1^4)x_2^2 + x_1x_2 + 4(x_2^2 - 1)x_2^2$). The two metamodels were constructed using 16 evaluations of the function for the kriging metamodel and an additional 16 evaluations of its gradients for the cokriging metamodel. In both cases, the correlation function was the Matérn function. First, the correlation length was considered to be the same for two components of the design space. Then, different correlation lengths were considered.

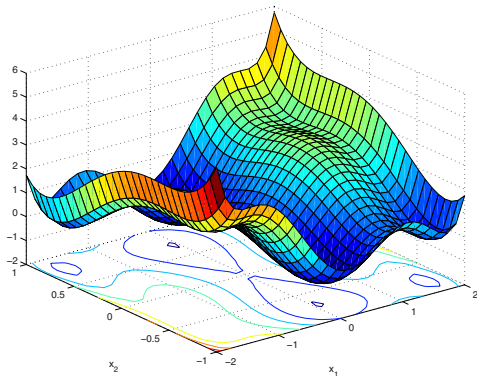
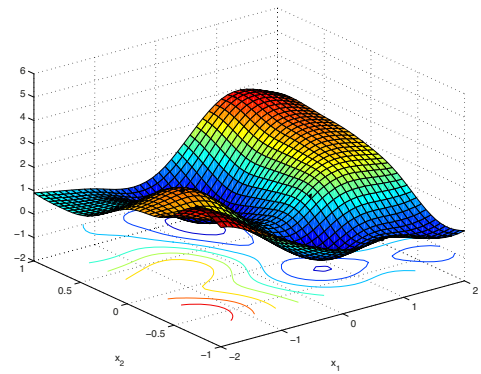
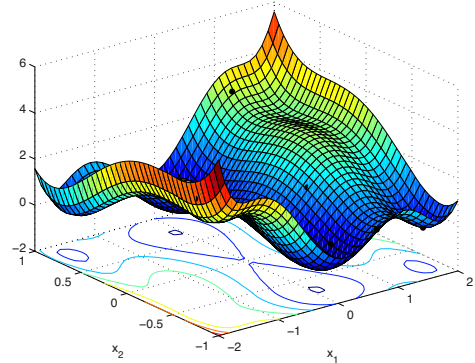


Fig. 11: The response surface of the actual function. (The isolines can be seen in Figure 20 of the appendix.)

The characteristics of these two metamodels are given in Table 3.



(a) Kriging metamodel



(b) Cokriging metamodel

Fig. 12: The kriging and cokriging metamodels obtained with 16 evaluations (plus 16 gradients for cokriging) of the analytical function with only one correlation length. (The isolines can be seen in Figure 21 of the appendix.)

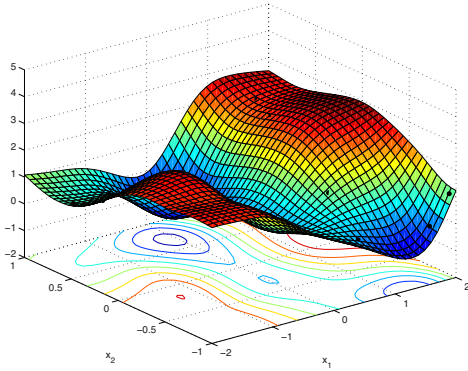
	OK	OCK
R^2	0.8236	0.9950
RAAE	0.2899	$2.938 \cdot 10^{-2}$
RMAE	2.7055	0.5604
Q_1	0.3429	$1.472 \cdot 10^{-2}$
Q_2	13.21	0.3757
Q_3	$8.257 \cdot 10^{-3}$	$2.348 \cdot 10^{-4}$

Table 3: Characteristics of the two metamodels (Figure 12a-12b) with a single correlation length

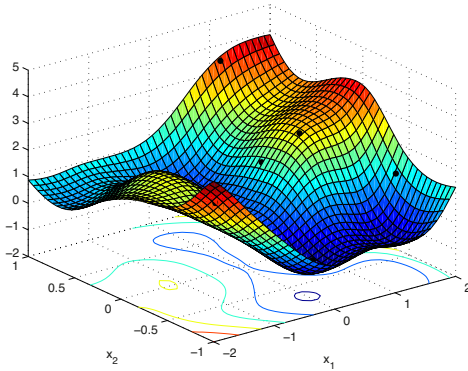
For this 2D test function, the cokriging metamodel led to a relatively accurate approximation of the actual function using only a few sample points. Taking into account the gradients, we were able to develop more efficient approximate models. However, for a problem involving the evaluation of mechanical responses, the computation cost of determining n_s responses and n_s gradients is obviously higher than for n_s responses only. (In the former case, due to the use of finite differences to calculate the gradients, the construction

of the metamodel requires $3n_s$ evaluations of the mechanical model.)

Now let us take another approach which consists in constructing metamodels using the same number of evaluations. In the following example, we used the same six-hump function, but the kriging metamodel was constructed based on the responses at 27 sample points (Figure 13a), while the cokriging metamodel was still constructed using the responses and the gradients at 9 sample points (Figure 13b).



(a) Kriging metamodel



(b) Cokriging metamodel

Fig. 13: The kriging and cokriging metamodels obtained with the 27 evaluations of the analytical function (27 responses for kriging; 9 responses and 9 gradients for cokriging) with only one correlation length. (The isolines can be seen in Figure 22 of the appendix.)

Table 4 shows that the quality of the predictions given by the two metamodels was similar, but a very important advantage of the cokriging metamodel was that it led to more zones where a minimum could be found than the kriging metamodel. If the problem involved the calculation of mechanical responses, the computation cost associated with the 9 points and their gradients would be much lower than that associated with the 27 points thanks to the multiparametric strategy, as shown in Table 1.

	OK	OCK
R^2	0.7818	0.7512
RAAE	0.2920	0.3579
RMAE	2.753	1.600
Q_1	0.3550	0.1199
Q_2	16.35	18.64
Q_3	$1.022 \cdot 10^{-2}$	$1.165 \cdot 10^{-2}$

Table 4: Characteristics of the previous metamodels obtained with the same number of evaluations (27 sample points for kriging and 9 sample points for cokriging; see Figures 13a-13b)

5.1.3 The case of two correlation lengths (anisotropy)

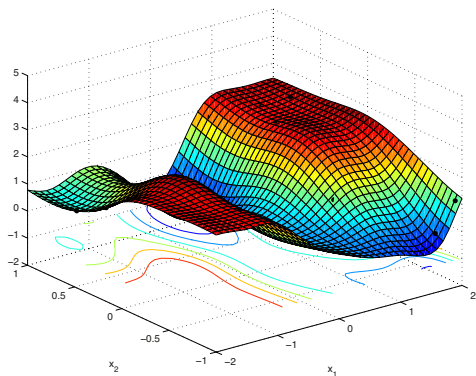
Now, let us take into account anisotropy. In the context of using correlation functions with the correlation length as a parameter, anisotropy is taken into account by allowing different correlation lengths for each design variable. We used the same two-dimensional example as in the previous section: the kriging and cokriging metamodels were constructed from the same data with the same number of function evaluations for the two models. The results were compared with those of the corresponding metamodels with a single correlation length.

Figures 14a and 14b show the kriging and cokriging metamodels whose results are to be compared with those of Figures 13a and 13b. The quality of these predictions is given in Table 5 (to be compared with the values of the quality criteria of Table 4).

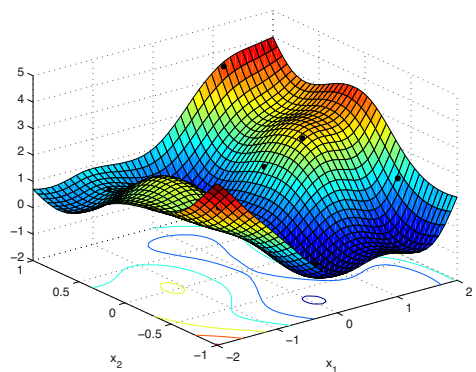
	OK	OCK
R^2	0.8163	0.7910
RAAE	0.2623	0.3208
RMAE	2.749	1.655
Q_1	0.3541	0.1283
Q_2	13.77	15.66
Q_3	$8.603 \cdot 10^{-3}$	$9.784 \cdot 10^{-3}$

Table 5: Characteristics of the two metamodels with anisotropy (Figures 14a-14b)

Table 6 shows the estimated correlation lengths obtained with the maximized likelihood with a single correlation length, then with anisotropy. For simplicity's sake, the sample points and the evaluations were normalized in order to carry out the kriging and cokriging metamodel construction process with standardized variables. The correlation lengths given in Table 6 refer to the variables after this transformation.



(a) Kriging metamodel (to be compared with the surface of Figure 13a)



(b) Cokriging metamodel (to be compared with the surface of Figure 13b)

Fig. 14: The kriging and cokriging metamodels obtained with 27 evaluations of the analytical function with anisotropy. (To be compared with Figure 13; the isolines can be seen in Figure 23 of the appendix.)

Anisotropy	No	Yes		Figures
OK (27 eval.)	0.9564	0.7106	1.5218	13a/14a
Ock (9 eval. & 9 grad.)	1.014	0.9878	1.310	13b/14b

Table 6: The correlation lengths for the construction of the metamodels of the two-dimensional analytical function

5.1.4 Discussion of the example with anisotropy

The previous results confirmed the results obtained with a single correlation length: the quality of the predictions given by the kriging and cokriging metamodels with the same amount of data (*i.e.* the same number of evaluations of the function) was very similar. The introduction of anisotropy improved the approximation for both types of metamodels (see Tables 4 and 5). This suggests that it is advisable to use a separate correlation length for each design variable.

5.2 Application to a contact problem with friction

In this section, we study the construction of the two previous metamodels (kriging and cokriging) using the multiparametric strategy presented in the first part of the paper along with the responses of mechanical models. The test cases considered are the three-squares example discussed in Section 3.3.2 plus a shrink-fit test case. The first example, which has a very smooth response surface, enables us to construct and study simple metamodels. The second example enables us to illustrate the strategy on a more complex and realistic response surface presenting local optima. In order to illustrate the performance of the coupled multiparametric/cokriging strategy, we carried out two studies: one with a fixed number of mechanical calls (Section 5.2.1), and the other with a fixed quality of the metamodels (Section 5.2.2).

5.2.1 The case of a fixed number of mechanical calls

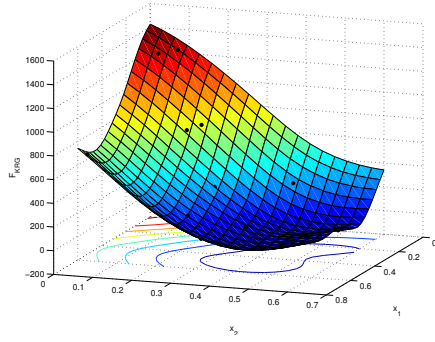
Each metamodel was constructed using 15 mechanical evaluations. The kriging metamodel was defined using 15 values of the force; the cokriging metamodel was defined using 5 values of the force and 5 gradients (with two components each, totaling 10 evaluations). The sample points were obtained through Latin hypercube sampling. Our calculations led to the two response surfaces of Figures 15a and 15b, which are to be compared with the actual response surface of Figure 5.

Tables 7 and 8 give the characteristics of the calculations and of the metamodels.

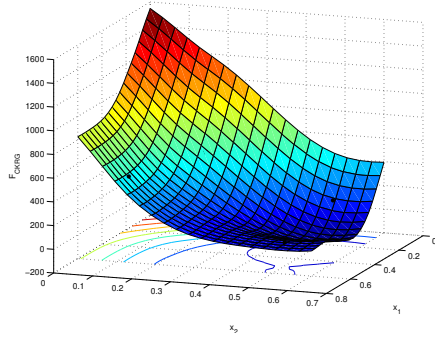
Gradients	no	yes
Avg. CPU Time (/calc)	2.34s	1.89s
Min. CPU Time	0.17s	0.18s
Max. CPU Time	8.01s	7.81s
LATIN iterations	1,636	1,311
Avg. iterations (/calc)	109	87.4
Min. iterations	5	5
Max. iterations	381	371
Total CPU Time	35.1s	28.4s
CPU Time w/o MPS	125s	126s
Gain	3.56	4.44

Table 7: Characteristics of the mechanical calculations

With this example, the cokriging metamodel led to a much better approximation of the mechanical model than the kriging metamodel. Moreover, this result was obtained at a lower computation cost.



(a) Kriging metamodel



(b) Cokriging metamodel

Fig. 15: The kriging and cokriging response surfaces of the three-squares example using 15 evaluations (15 responses for the kriging case; 5 responses and 5 gradients for the cokriging case). The isolines can be seen in Figure 24 of the appendix.

	OK	OCK
R^2	0.9880	0.9781
RAAE	$8.884 \cdot 10^{-2}$	0.1120
RMAE	0.3774	0.4720
Q_1	$8.917 \cdot 10^{-3}$	$1.395 \cdot 10^{-2}$
Q_2	0.2990	0.5467
Q_3	$7.475 \cdot 10^{-4}$	$1.367 \cdot 10^{-3}$

Table 8: Characteristics of the metamodels

5.2.2 The case of a fixed quality of the metamodels

Another way to study the cost of the metamodels consists in constructing kriging and cokriging metamodels of similar quality. In this case, we constructed the metamodels using sample points obtained through full factorial sampling. We chose to build the kriging metamodel with 9×9 samples and the cokriging metamodel with 3×3 and 5×5 samples. The results are shown in Table 9.

5×5 full factorial sampling led to a cokriging metamodel with the same quality as the 9×9 kriging metamodel. As in the previous case with a fixed number of mechanical

	OK 9×9	OCK 3×3	OCK 5×5
Avg. CPU Time (/calc)	1.79s	2.17s	1.68s
Min. CPU Time	0.16s	0.16s	0.16s
Max. CPU Time	8.01s	8.06s	8.08s
LATIN iterations	6,576	2,711	5,666
Avg. iterations (/calc)	81.2	100	75.5
Min. iterations	5	5	5
Max. iterations	381	381	381
Total CPU Time	144s	58.5s	126s
CPU Time w/o MPS	689s	225s	637s
Gain	4.78	3.84	5.06
R^2	0.9986	0.9756	0.9955
RAAE	$2.699 \cdot 10^{-2}$	0.1155	$4.468 \cdot 10^{-2}$
RMAE	0.1818	0.5352	0.3290
Q_1	$2.068 \cdot 10^{-3}$	$1.793 \cdot 10^{-2}$	$6.774 \cdot 10^{-3}$
Q_2	$3.631 \cdot 10^{-2}$	0.6087	0.1134
Q_3	$9.078 \cdot 10^{-5}$	$1.522 \cdot 10^{-3}$	$2.836 \cdot 10^{-4}$

Table 9: The metamodels constructed using full factorial sampling

calls, the cost with the cokriging metamodel was lower. The simplicity of the response surface of this test case enabled the metamodels built with the multiparametric strategy to be studied easily and visually. In order to assess the performance of the cokriging metamodel, we also introduced a more irregular response surface associated with a shrink-fit test case.

5.3 Application to a shrink-fit problem

In order to study a case with a more irregular response surface capable of illustrating the performance of the cokriging metamodel, we introduced an additional test case: a shrink-fit assembly problem in 2D. Figure 16 shows the geometry of this problem.

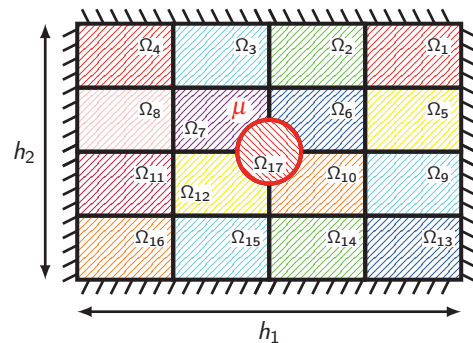


Fig. 16: The geometry of the shrink-fit assembly problem

The reference problem consists of a cylindrical shaft inserted into a perforated rectangular plate ($h_1 = 0.3m$,

$h_2 = 0.2m$, Young's modulus $E = 80GPa$, Poisson's coefficient $\nu = 0.3$). The shaft is made of a perfect material, but presents geometric defects: a diameter slightly larger than its nominal value and a small eccentricity. The plate was divided into 16 substructures, each meshed with six-node triangular elements. The shaft itself was a substructure, also meshed with six-node triangular elements, with the same material characteristics and a nominal diameter of $5cm$. Figure 17 shows the nominal geometry and the actual geometry of the shaft. The eccentricity defect was set to $e = 0.5\mu m$ and the excess diameter was set to $d_r = R - R' = 1\mu m$. The loading was applied in two stages: first, the shaft was mounted in the frame; then, a uniform vertical pressure was applied progressively to the shaft up to a maximum $250MPa$. The mechanical solution of this problem was obtained through a quasi-static resolution.

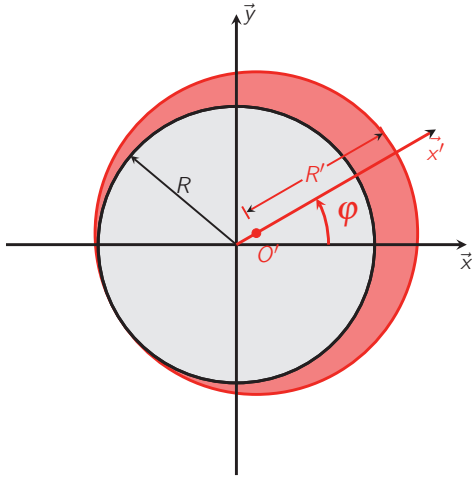


Fig. 17: The nominal and actual geometries of the shaft

We considered two design variables: the friction coefficient μ between the shaft and the frame ($\mu \in [0.02, 0.7]$) and the orientation angle of the eccentricity defect φ ($\varphi \in [0, 360]$). The objective function studied was the maximum Von Mises' stress in the structure.

Figure 18 shows the reference response surface plotted on a 20×20 regular grid: one can observe that for this problem the objective function has several local minima and is more irregular than that of the three-squares test case.

The kriging and cokriging metamodels of this test case were constructed from evaluations of the objective function at 30 and 10 sample points respectively. These sampled functions were obtained using the Latin Hypercube Sampling procedure. Figure 19 and Table 10 show the resulting metamodels and their characteristics.

These results show that the quality of the two metamodels is very similar. The cokriging metamodel was less expensive to build than the kriging metamodel, due largely to

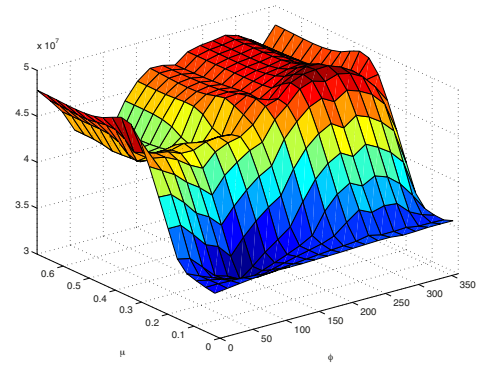


Fig. 18: The reference response surface of the shrink-fit test case

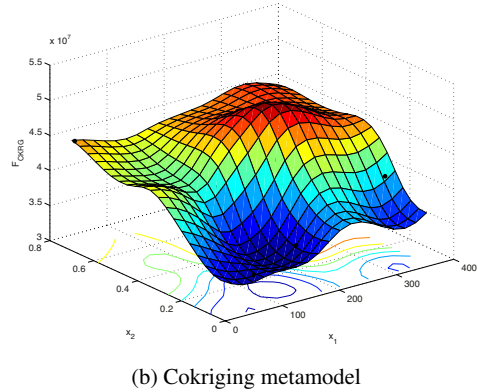
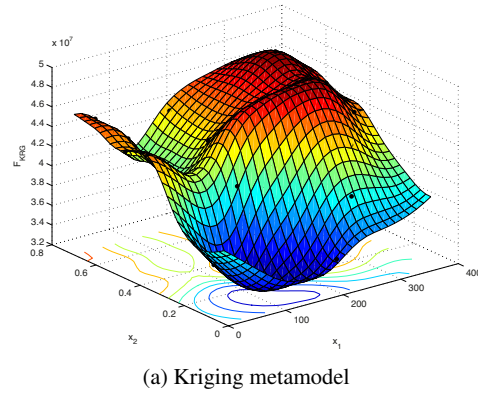


Fig. 19: The kriging and cokriging response surfaces of the shrink-fit test case using 30 evaluations (30 responses for kriging and 10 responses and 10 gradients for cokriging; the isolines are presented in Figure 25 of the appendix)

the multiparametric strategy. For this test case, the response surfaces enabled a good localization of the global minimum.

	OK	OCK
Number of evaluations	30	30
Number of responses	30	10 (10 eval.)
Number of gradients	0	10 (20 eval.)
Avg. CPU Time (/calc)	60.5s	45.6s
Min. CPU Time	33.1s	2.21s
Max. CPU Time	123.6s	146.2s
LATIN iterations	27,137	19,307
Avg. iterations (/calc)	905	644
Min. iterations	36	30
Max. iterations	3,605	2,865
Total CPU Time	1815s	1366s
CPU Time w/o MPS	3641s	3441s
Gain	2.01	2.51
R^2	0.8841	0.8158
RAAE	0.2649	0.3367
RMAE	0.9956	1.183
Q_1	$9.862 \cdot 10^{-3}$	$1.393 \cdot 10^{-2}$
Q_2	0.4601	0.7312
Q_3	$1.150 \cdot 10^{-3}$	$1.828 \cdot 10^{-3}$

Table 10: Characteristics of the metamodels

6 Conclusion

This paper proposes a new multilevel model optimization strategy based on a multiparametric approach and a cokriging metamodel and presents a detailed study of the cost of constructing this type of metamodel. Based on two tools – the Multiparametric strategy and a gradient-based cokriging metamodel – the proposed approach allows us to obtain accurate approximations of objective functions. Moreover, due to the use of the M.P.S., approximations are built with a significant reduction of the computational time in comparison with classical mechanical solver (*i.e.* without reinitialization of the solving algorithm). On a two-variables academic test example one can reach a gain about 3 to 4 in terms of computational time. The closer the sample points, the greater the gain. Obviously, the gain would be even greater in the case of 3D problems [64].

This conclusion led to the construction of a gradient-based cokriging metamodel. The final conclusion of this study is the following:

- With similar numbers of sample points, cokriging leads to more accurate results than kriging, but this is not very useful in the mechanical context due to its higher computation cost. (The construction of the cokriging metamodel requires three times as many calculations as the kriging metamodel in order to obtain the responses and the gradients.)
- For a given quality of the results, cokriging enables a significant reduction in the number of mechanical calls compared to kriging.

- For a given quality of the results, cokriging leads to a significant reduction in computation cost compared to kriging.
- Like Kriging, Cokriging provides accurate locations of the local optima.
- In all cases involving more than one design variable, the use of the anisotropy principle is crucial for optimizing the quality of the metamodels.

These results pave the way for future works on larger and more complex assembly problems in order to deal with more realistic industrial structures, such as two- and three-dimensional problems with many design variables. Besides, future developments will be also to compare cokriging metamodel with other classical models such as gradient-based Radial Basis Function [65] and Artificial Neural Network [66]. Another objective will be to complete the development of the first and second levels of the multilevel model optimization strategy, *i.e.* the exploratory phase to localize the optimum approximately and the subsequent precise optimization through new mechanical calculations using the multiparametric strategy along with the results from the first level.

Acknowledgements This work was supported by the *Agence Nationale de la Recherche* as part as ANR-08-COSI-007-10: Distributed Multi-Disciplinary Optimization.

References

1. G.M. Robinson and A.J. Keane. A case for multi-level optimization in aeronautical design. *Aeronautical Journal*, 103(1028):481–485, 1999.
2. A.J. Booker, J.E. Dennis, P.D. Frank, D.B. Serafini, V. Torczon, and M.W. Trosset. A rigorous framework for optimization of expensive functions by surrogates. *Structural and Multidisciplinary Optimization*, 17(1):1–13, 1999.
3. N.V. Queipo, R.T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P.K. Tucker. Surrogate-based analysis and optimization. *Progress in Aerospace Sciences*, 41(1):1 – 28, 2005.
4. T.W. Simpson, J.D. Poplinski, P.N. Koch, and J.K. Allen. Metamodels for computer-based engineering design: survey and recommendations. *Engineering with Computers*, 17(2):129–150, 2001.
5. S. Kravanja, S. Silih, and Z. Kravanja. The multilevel minlp optimization approach to structural synthesis: the simultaneous topology, material, standard and rounded dimension optimization. *Advances in engineering software*, 36(9):568–583, 2005.
6. S. Kravanja, A. Soršak, and Z. Kravanja. Efficient multilevel minlp strategies for solving large combinatorial problems in engineering. *Optimization and engineering*, 4(1):97–151, 2003.
7. B. Liu, R.T. Haftka, and L.T. Watson. Global-local structural optimization using response surfaces of local optimization margins. *Structural and Multidisciplinary Optimization*, 27(5):352–359, 2004.
8. N.M. Alexandrov and R.M. Lewis. Analytical and computational aspects of collaborative optimization. *NASA Technical Memorandum 210104*, 2000.
9. T.Y. Chen and C.M. Yang. Multidisciplinary design optimization of mechanisms. *Advances in engineering software*, 36(5):301–311, 2005.

10. C.A. Conceição António. A multilevel genetic algorithm for optimization of geometrically nonlinear stiffened composite structures. *Structural and Multidisciplinary Optimization*, 24(5):372–386, 2002.
11. M.P. Bendsøe. *Optimization of structural topology, shape, and material*. Springer Verlag, New York, 1995.
12. R. Le Riche and J. Gaudin. Design of dimensionally stable composites by evolutionary optimization. *Composite structures*, 41(2):97–111, 1998.
13. P.S. Theocaris and G.E. Stavroulakis. Multilevel optimal design of composite structures including materials with negative poisson's ratio. *Structural and Multidisciplinary Optimization*, 15(1):8–15, 1998.
14. A.J. Keane and N. Petruzzelli. Aircraft wing design using ga-based multi-level strategies. In *AIAA paper 2000-4937. AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Long Beach, USA, 06 - 08 Sep 2000*. American Institute of Aeronautics and Astronautics, 2000.
15. H. Engels, W. Becker, and A. Morris. Implementation of a multi-level optimisation methodology within the e-design of a blended wing body* 1. *Aerospace Science and technology*, 8(2):145–153, 2004.
16. A. Chattopadhyay, T.R. McCarthy, and N. Pagaldipti. Multilevel decomposition procedure for efficient design optimization of helicopter rotor blades. *AIAA journal*, 33(2):223–230, 1995.
17. M.E.M. El-Sayed and C.K. Hsiung. Optimum structural design with parallel finite element analysis. *Computers & structures*, 40(6):1469–1474, 1991.
18. P.K. Umesh, M.T. Venuraju, D. Hartmann, and K.R. Leimbach. Optimal design of truss structures using parallel computing. *Structural and Multidisciplinary Optimization*, 29(4):285–297, 2005.
19. B. Dunham, D. Fridshal, R. Fridshal, and J.H. North. Design by natural selection. *Synthese*, 15(1):254–259, 1963.
20. M.A. El-Beltagy and A.J. Keane. A comparison of various optimization algorithms on a multilevel problem. *Engineering Applications of Artificial Intelligence*, 12(5):639–654, 1999.
21. D. Elby, R.C. Averill, W.F. Punch, and E.D. Goodman. Evaluation of injection island ga performance on flywheel design optimisation. In *Proc. Third Int. Conf. Adaptive Computing in Design and Manufacture*, pages 121–136, 1998.
22. J.H. Holland. *Adaptation in natural and artificial systems*. Univ. Mich. Press, Ann Arbor, 1975.
23. G.E.P. Box and K.B. Wilson. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 13(1):1–45, 1951.
24. T.W. Simpson, T.M. Mauery, J.J. Korte, Multidisciplinary Optimization Branch, and F. Mistree. Comparison of response surface and kriging models for multidisciplinary design optimization. In *AIAA paper 98-4758. 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, pages 98–4755, 1998.
25. S. Haykin. *Neural networks: a comprehensive foundation*. Prentice Hall PTR Upper Saddle River, NJ, USA, 1994.
26. W. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 5(4):115–133, December 1943.
27. R.L. Hardy. Multiquadric equations of topography and other irregular surfaces. *J. Geophys. Res.*, 76:1905–1915, 1971.
28. A. Chatterjee. An introduction to the proper orthogonal decomposition. *Current Science*, 78(7):808–817, Avril 2000.
29. B. Soulier, L. Richard, B. Hazet, and V. Braibant. Crashworthiness optimization using a surrogate approach by stochastic response surface. *G. Gogu D. Coutellier P. Chedmail P. Ray Editor, Recent Advances in Integrated Design and Manufacturing in Mechanical Engineering*, pages 159–168, Mai 2003.
30. N. Cressie. The origins of kriging. *Mathematical Geology*, 22(3):239–252, 1990.
31. H. Wackernagel. *Multivariate geostatistics: an introduction with applications*. Springer Verlag, 1995.
32. H.S. Chung and J.J. Alonso. Using gradients to construct cokriging approximation models for high-dimensional design optimization problems. In *40th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada*. Citeseer, 2002.
33. J.F.M. Barthelemy and R.T. Haftka. Approximation concepts for optimum structural design - a review. *Structural and Multidisciplinary Optimization*, 5(3):129–144, 1993.
34. J. Sobieszcanski-Sobieski and R.T. Haftka. Multidisciplinary aerospace design optimization: survey of recent developments. *Structural and Multidisciplinary Optimization*, 14(1):1–23, 1997.
35. P. Ladevèze. *Nonlinear computational structural mechanics: new approaches and non-incremental methods of calculation*. Springer Verlag, 1999.
36. J. Mandel. Balancing domain decomposition. *Communications in numerical methods in engineering*, 9(3):233–241, 1993.
37. P. Le Tallec. Domain decomposition methods in computational mechanics. *Computational mechanics advances*, 1(2):121–220, 1994.
38. C. Farhat and F.X. Roux. A method of finite element tearing and interconnecting and its parallel solution algorithm. *International Journal for Numerical Methods in Engineering*, 32(6):1205–1227, 1991.
39. C. Blanzé, L. Champany, J.Y. Cognard, and P. Ladevèze. A modular approach to structure assembly computations: application to contact problems. *Engineering Computations*, 13(1):15–32, 1995.
40. P. Ladevèze. Sur une famille d'algorithmes en mécanique des structures. *Compte rendu de l'académie des Sciences*, 300(2):41–44, 1985.
41. L. Champany, JY Cognard, and P. Ladeveze. Modular analysis of assemblages of three-dimensional structures with unilateral contact conditions. *Computers & Structures*, 73(1-5):249–266, 1999.
42. P.A. Boucard and P. Ladevèze. Une application de la méthode latin au calcul multirésolution de structures non linéaires. *Revue Européenne des Eléments Finis*, 8:903–920, 1999.
43. P.A. Boucard. Application of the latin method to the calculation of response surfaces. In *Proceeding of the First MIT Conference on Computational Fluid and Solid Mechanics, Cambridge, USA*, volume 1, pages 78–81, Juin 2001.
44. O. Allix and P. Vidal. A new multi-solution approach suitable for structural identification problems. *Computer Methods in Applied Mechanics and Engineering*, 191(25-26):2727–2758, 2002.
45. P.A. Boucard and L. Champany. A suitable computational strategy for the parametric analysis of problems with multiple contact. *International Journal for Numerical Methods in Engineering*, 57(9):1259–1281, 2003.
46. L. Champany, P.A. Boucard, and S. Guinard. Adaptive multi-analysis strategy for contact problems with friction. *Computational Mechanics*, 42(2):305–315, 2008.
47. B. Soulier and P.A. Boucard. A multiparametric strategy for the large-scale multilevel optimization of structural assemblies. In *8th World Congress on Structural and Multidisciplinary Optimization, Lisbon, Portugal*, 2009.
48. D.G. Krige. A statistical approach to some mine valuation and allied problems on the Witwatersrand. Master's thesis, 1951.
49. G. Matheron. *Traité de géostatistique appliquée*, Tome I. *Memoires du Bureau de Recherches Géologiques et Minières*, 14, 1962.
50. G. Matheron. *Traité de Géostatistique Appliquée*, Tome II: Le Krigeage. *Memoires du Bureau de Recherches Géologiques et Minières*, No 24, 1962.
51. G. Matheron. Principles of geostatistics. *Economic geology*, 58(8):1246, 1963.
52. J. Sacks, S.B. Schiller, and W.J. Welch. Designs for computer experiments. *Technometrics*, 31(1):41–47, 1989.
53. J. Sacks, W.J. Welch, T.J. Mitchell, and H.P. Wynn. Design and analysis of computer experiments. *Statistical science*, 4(4):409–423, 1989.

54. C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, USA, 01 2006.
55. J.R. Koehler and A.B. Owen. Computer experiments. *Handbook of statistics*, 13:261–308, 1996.
56. M.D. Morris, T.J. Mitchell, and D. Ylvisaker. Bayesian design and analysis of computer experiments: use of derivatives in surface prediction. *Technometrics*, 35(3):243–255, 1993.
57. B. Matérn. *Spatial Variation (Lecture Notes Statist. 36)*. Springer, Berlin, 1960.
58. M.L. Stein. *Interpolation of Spatial Data: some theory for kriging*. Springer Verlag, 1999.
59. K.V. Mardia and R.J. Marshall. Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika*, 71(1):135, 1984.
60. K.V. Mardia and A.J. Watkins. On multimodality of the likelihood in the spatial linear model. *Biometrika*, 76(2):289, 1989.
61. M.J. Sasena. *Flexibility and efficiency enhancements for constrained global design optimization with kriging approximations*. PhD thesis, University of Michigan, 2002.
62. J.J. Warnes and B.D. Ripley. Problems with likelihood estimation of covariance functions of spatial gaussian processes. *Biometrika*, 74(3):640, 1987.
63. M.D. McKay, W.J. Conover, and R.J. Beckman. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.
64. V. Roulet, L. Champaney, and P.-A. Boucard. A parallel strategy for the multiparametric analysis of structures with large contact and friction surfaces. *Advances in Engineering Software*, 42(6):347 – 358, 2011.
65. S.J. Leary, A. Bhaskar, and A.J. Keane. Global approximation and optimization using adjoint computational fluid dynamics codes. *AIAA journal*, 42(3):631–641, 2004.
66. KC Giannakoglou, DI Papadimitriou, and IC Karpolis. Aerodynamic shape design using evolutionary algorithms and new gradient-assisted metamodels. *Computer methods in applied mechanics and engineering*, 195(44-47):6312–6329, 2006.

A Complementary figures

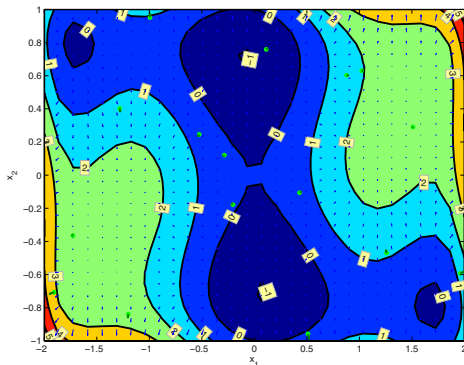
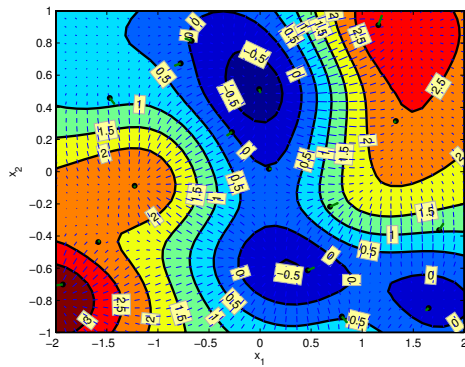
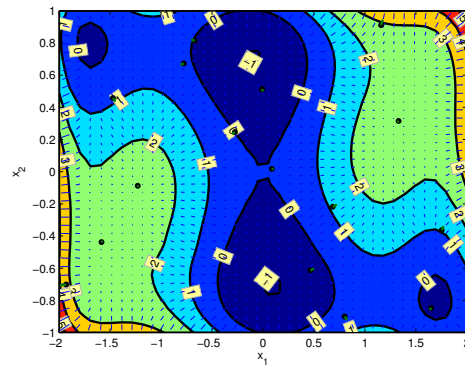


Fig. 20: Isolines and gradients of the six-hump camel back function presented in Section 5.1.2 and in Figure 11

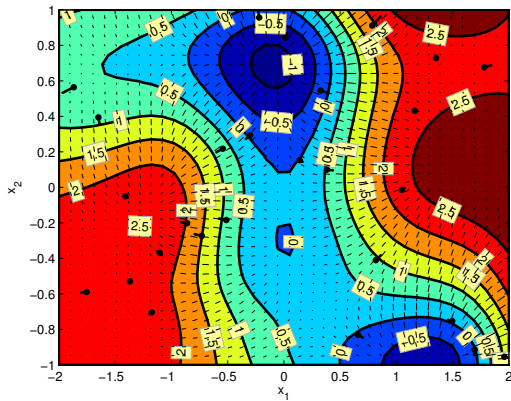


(a) Kriging metamodel (associated with the response surface of Figure 12a)

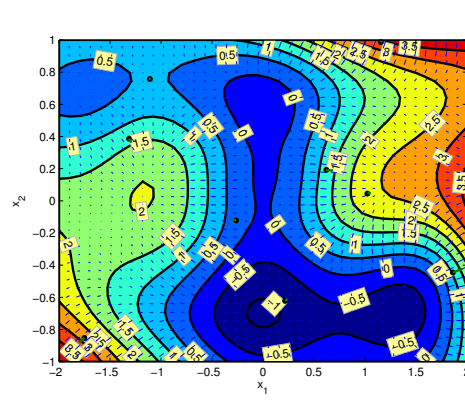


(b) Cokriging metamodel (associated with the response surface of Figure 12b)

Fig. 21: The isolines and gradients of the kriging and cokriging metamodels obtained with 16 evaluations of the analytical function and a single correlation length associated with the response surfaces presented in Figure 12 of Section 5.1.2

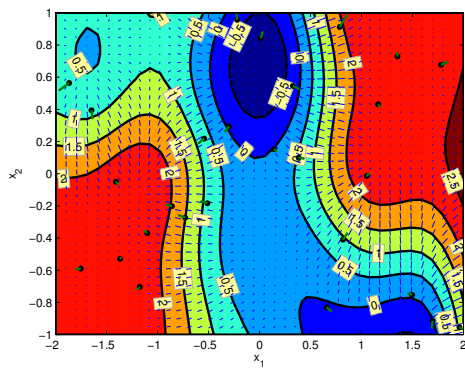


(a) Kriging metamodel (associated with the response surface of Figure 13a)

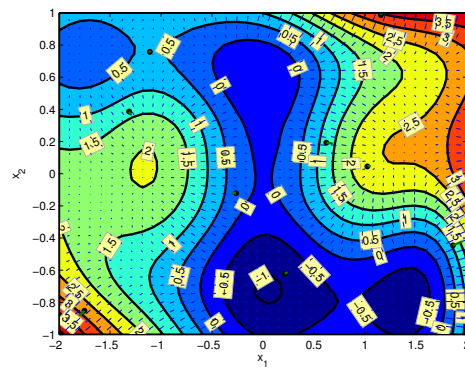


(b) Cokriging metamodel (associated with the response surface of Figure 13b)

Fig. 22: The isolines and gradients of the kriging and cokriging metamodels obtained with the 27 evaluations of the analytical function and a single correlation length associated with the response surfaces presented in Figure 13 of Section 5.1.2

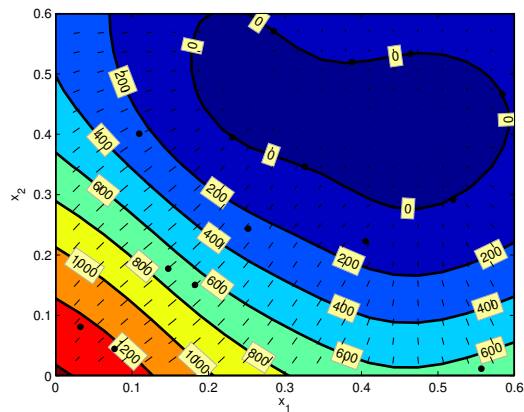


(a) Kriging metamodel (associated with the response surface of Figure 14a)

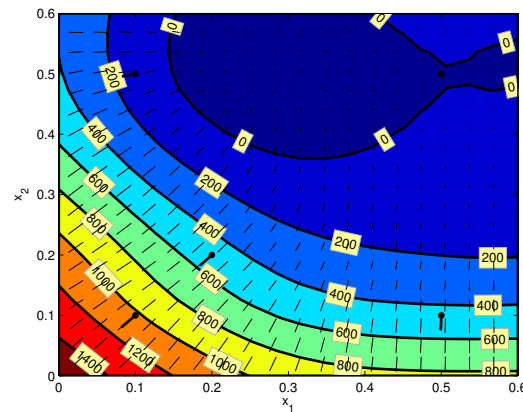


(b) Cokriging metamodel (associated with the response surface of Figure 14b)

Fig. 23: The isolines and gradients of the kriging and cokriging metamodels obtained with 27 evaluations of the function with anisotropy associated with the response surfaces presented in Figure 14 of Section 5.1.3

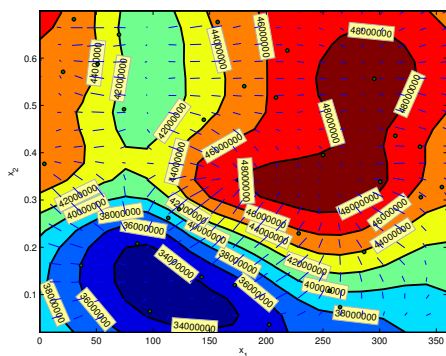


(a) Kriging isolines and gradients (associated with the response surface of Figure 15a)

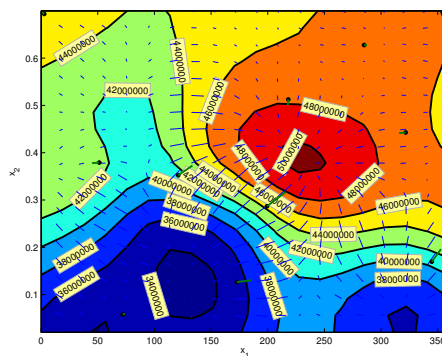


(b) Cokriging isolines and gradients (associated with the response surface of Figure 15b)

Fig. 24: The kriging and cokriging metamodels constructed with 15 mechanical evaluations associated with the response surfaces presented in Figure 15



(a) Kriging isolines and gradients (associated with the response surface of Figure 19a)



(b) Cokriging isolines and gradients (associated with the response surface of Figure 19b)

Fig. 25: The kriging and cokriging metamodels constructed with 30 mechanical evaluations associated with the response surfaces presented in Figure 19