Design optimization under uncertainties of a mesoscale implant in biological tissues using a probabilistic learning algorithm

Christian Soize

Design optimization under uncertainties of a mesoscale implant in biological tissues using a probabilistic learning algorithm

C. Soize

Abstract This paper deals with the optimal design of a titanium mesoscale implant in a cortical bone for which the apparent elasticity tensor is modeled by a non-Gaussian random field at mesoscale, which has been experimentally identified. The external applied forces are also random. The design parameters are geometrical dimensions related to the geometry of the implant. The stochastic elastostatic boundary value problem is discretized by the finite element method. The objective function and the constraints are related to normal, shear, and von Mises stresses inside the cortical bone. The constrained nonconvex optimization problem in presence of uncertainties is solved by using a probabilistic learning algorithm that allows for considerably reducing the numerical cost with respect to the classical approaches.

1 Introduction

This paper concerns the optimal design of a titanium mesoscale implant in a cortical bone (biological tissue), which requires to solve a Constrained Nonconvex Optimization Problem in Presence of Uncertainties (CNOPPU) that means that there are an underlying stochastic operator and stochastic constraints. These stochastic aspects are due to the heterogeneous elastic random medium (the cortical bone) at the microscale and also to the random forces that are applied to the titanium implant. This paper deals with the optimization of geometrical parameters (design parameters) of the implant in order to minimize some stresses in a given region located at the interface between the titanium implant and the cortical bone (objective function), and also to limit stresses in some regions of the cortical bone (nonlinear constraint functions). For each given value of the design parameters, the evaluation of the objective and constraints functions is numerically expensive because a stochastic boundary value problem must be solved. Due to the nonconvexity of such an optimization problem, random search algorithms or genetic algorithm must, for instance, be used, inducing a very high numerical cost. The use of an algorithm based on a probabilistic learning methodology can thus help to circumvent these difficulties.

In order to well explain the existing approaches, the difficulties encountered for solving the design optimization of a mesoscale implant in biological tissues, and the methodology proposed for circumvent these difficulties, the following items are briefly presented:

- Design optimization of implants in bone.
- Stochastic modeling of biological tissues.
- Uncertainty quantification in computational mechanics.
- Algorithms for solving optimization problems under uncertainties.
- Probabilistic learning algorithm used in this paper.
- Main objectives and novelties of the methodology proposed.
- Organization of the paper.

1. Design optimization of implants in bone. The biomechanical aspects play an important role in the techniques related to implants in bones and many research have been per-

C. Soize
Université Paris-Est, Modelisation et Simulation Multi-Echelle, MSME UMR 8208 CNRS, 5 bd Descartes, 77455 Marne-la-Vallée Cedex 02, France
Tel.: +33-160957661
Fax: +33-160957799
E-mail: christian.soize@univ-paris-est.fr

formed, in particular by using experimental investigations and by developing computational mechanical models for understanding the aging and the implants stability as a function of the geometric shape and of the dimensions of implants (see for instance [2, 4, 9, 17, 19, 32, 37] and also [44, 49, 58, 60, 71, 78, 92, 95]). Design optimization of implants in bone at macroscale by using deterministic computational models has also received increased attention (see for instance [36, 43, 50, 53, 56, 74]). However, there are only a few works in the framework of the optimization under uncertainties in biomechanics (see for instance [34, 65, 67, 73]). Nevertheless, this type of works does not exactly correspond to the framework of the present paper that concerns the optimization of the geometry of a mesoscale implant in a cortical bone that must be modeled by a heterogeneous random elastic medium at this scale. Consequently, the CNOPPU that has to be solved, is very expensive because it requires the computation of quantities of interest for each realization of the random elasticity field.

2. Stochastic modeling of biological tissues. When the random microstructure of a heterogeneous material can be described in terms of its constituents, the homogenization method allows for calculating the effective mechanical properties and their bounds at the macroscale. Often, the probabilistic model of a random microstructure (such as a composite constituted of several constituents) can directly be constructed from the geometry and mechanical properties of its constituents. This is the case for the class of random heterogeneous materials whose microstructures can be modeled as a distribution of inclusions or cavities of well-defined geometry in a given matrix. This is also the case of random heterogeneous materials having a cellular structure for which the probabilistic model is directly constructed using cell statistics, random field models, percolation, and clustering (see for instance [18, 20, 42, 47, 38–40, 59, 62, 66, 76, 77, 94]). However, at microscale, a biological tissue such as the cortical bone appears as a random medium and its elastic heterogeneous microstructure cannot be described in terms of its constituents. Consequently, its elasticity field is modeled by a tensor-valued random field at mesoscale (see for instance [86]) and therefore, the derived computational model exhibits uncertainties. That is the case considered in this paper. In addition, since the implant is at the mesoscale, it is necessary to use a stochastic modeling of the random elasticity field adapted to this scale. Note that the problem under consideration is not relevant of stochastic homogenization because the coupling of the implant with the cortical bone is at the mesoscale.

3. Uncertainty quantification in computational mechanics. Concerning uncertainties, it is recognized that uncertainty quantification plays an increasing role in computational mechanics and more generally, for large-scale computational models in sciences and engineering. The reader will be able to find an excellent extended, complete, and advanced presentation in the handbook of uncertainty quantification [26] recently published in which the UQ ideas, the useful mathematical tools, and applications in many fields can be found, and which corresponds to the present state of the art. Relatively recently, it has been recognized that multiscale modelings and simulations with uncertain computational model require adapted algorithms and very powerful computers [93]. In this framework, advanced multiscale modelings under uncertainties have been developed (see for instance [13, 52, 55, 83, 84, 101]). When these works were published, design optimization with stochastic computational model was not still really possible for large scale multiscale computational models. The applications that have been published (see for instance [1, 5, 11, 12, 16, 22, 29] and also [35, 54, 57, 68, 70, 72, 80, 99, 96]) were devoted to optimization problems under uncertainties for which the computational models had a reasonable number of degrees of freedom, for which the optimizers were based on the use of relatively classical optimization algorithms and/or the introduction of approximations such as surface responses and surrogate models.

4. Algorithms for solving optimization problems under uncertainties. The algorithms for solving optimization problems without uncertainties have extensively been developed for several decades and are today really efficient (see for instance [102]). Nevertheless, for optimization with large-scale nonlinear computational models, the introduction of parametric reduced-order models are often necessary (see for instance [3, 24]). Concerning the algorithms for solving optimization problems under uncertainties, many methods have been proposed in the literature such as the gradient-based learning that is adapted to convex problems [48, 89], the global search algorithms such as the stochastic algorithms, the genetic algorithm, and the evolutionary algorithms [14, 46]. The statistical learning methods have also been developed for solving deterministic optimization problems that are interpreted as representatives of stochastic problem families with the benefit of enabling statistical learning [41]. The learning process is often implemented by using surrogate models that allow for constructing approximations of the functions whose evaluations are numerically expensive [10, 75]. The resulting error and its repercussions on the optimal solution depend on the various algorithms that are proposed [33], in particular for calculating the mathematical expectations that must be estimated for evaluating the expensive functions for each design point and that require the computation of a large number of statistical realizations of the model outputs. The global character of the surrogate models is usually satisfied either by using a deterministic interpolation process or by using a stochastic model whereby the complex dependencies between model outputs and design parameters are expressed by computing statis-

tical correlations. Gaussian process models are most commonly used in this context [45,90]. However, more robust alternatives based on Bayesian optimization [28,41,97,98] have also proven useful. For expensive evaluations of functions that depend on uncertainties, computational difficulties still remain important enough to require the introduction of simplifying assumptions in the form of surrogate models and approximations for estimating probabilities [10,21,23, 100].

5. Probabilistic learning algorithm used in this paper. In this paper, the probabilistic learning used for solving the optimization problem under uncertainties refers to an algorithm that is based on the use of the two following methodologies.

(i) The first one is the novel methodology that is proposed in [25,88] for solving the CNOPPU, which consists in using

- an optimizer that is adapted to the nature of the optimization problem under consideration,
- a dataset that is made up of a small number of points generated by the optimizer for which only a small number of expensive evaluations of functions is carried out.

It is well known that the choice of the optimizer depends on the number of design parameters and on the nature of the optimization problem. A possible optimizer can, *a priori*, be a grid search algorithm [8], a gradient-based algorithm [7], a bayesian optimization algorithm [81], a random search algorithm [8] or a genetic algorithm [27]. It should be noted that, for the nonconvex optimization problem under consideration, a random search algorithm or the genetic algorithm is better adapted. Nevertheless, if the number of the design parameters is 1, 2 or 3, a grid search algorithm is efficient for solving a nonconvex optimization problem. In this paper, the number of design parameters being 2, we will use the grid search algorithm.

(ii) The second one is also a new methodology recently proposed in [85,87], which is devoted to the probabilistic learning on manifold from a dataset made up of a small number of points and, which allows

- for identifying, from a database made up of realizations of a $\mathbb{R}^\nu$-valued random variable, its non-Gaussian probability distribution that is unknown and that is concentrated on an unknown subset $\mathcal{S}_\nu$ (a manifold) of $\mathbb{R}^\nu$.
- for generating additional realizations that follow the unknown probability distribution by preserving the concentration on $\mathcal{S}_\nu$ and consequently, by avoiding the scattering of the realizations when they are generated by the classical Markov Chain Monte Carlo (MCMC) method.

Such a probabilistic learning algorithm allows for estimating the objective function and the nonlinear constraint functions for each design point proposed by the optimizer without using the stochastic computational model.

6. Main objectives and novelties of the methodology proposed. The two main objectives and the novelties of this paper are the following.

(i) The optimization problem is formulated for the design of the geometry of a mesoscale titanium implant in a cortical bone by introducing a realistic probabilistic model of the heterogeneous random elasticity field of the cortical bone at the mesoscale. This non-Gaussian random field has experimentally been identified by using an experimental multiscale measurements on the basis of a digital image correlation technique [64]. Note that the heterogeneous medium at mesoscale is not modeled by using only a few real-valued random variables related to the elasticity tensor, but is modeled by using a complete non-Gaussian tensor-valued anisotropic random field for the elasticity field [82,83,87]. Note also that the two spatial correlation lengths of this field in the plane perpendicular to the mean direction of the Haversian canals of the osteons of the cortical bone is of the same order that the thickness of the titanium implant. This means that there is a strong effect of the statistical fluctuations in the cortical bone at mesoscale on the optimal design of the implant.

(ii) As previously explained, the CNOPPU is solved with a combination of the algorithms recently introduced in [25, 85,87], which will allow us to solve the design optimization of a mesoscale implant in a cortical bone with a reasonable numerical cost. The algorithm proposed is different from those existing and does not introduce any approximations such as surface responses, surrogate models, etc. The considerable gain that is obtained with the algorithm proposed opens a way for design optimization under uncertainties of more complex biomechanical systems.

It should be noted that, although the objective of the paper is not to validate the used probabilistic learning algorithm that has already been published, one obtains an additional validation. The main objective of this paper is above all to demonstrate how the optimization of the design of a mesoscale implant in a biological tissue can effectively be performed in a realistic framework by considering the cortical bone as a random medium at the mesoscale. Such an objective can be reached thanks to the use of an advanced algorithm that allows for solving this very difficult constrained nonconvex optimization problem in presence of uncertainties in biomechanics, problems that generate extremely high computational costs without the use of this type of algorithm.

7. Organization of the paper. Section 2 deals with the definition of the optimum design problem with a stochastic computational model. In Section 3, the stochastic modeling of the cortical bone is presented. Section 4 is devoted to the construction of the stochastic computational model. In Section 5, we first discuss the major difficulties encountered for solving the nonconvex optimization problem under uncertainties. Then, we present the method that is pro-

posed for solving the optimization design problem with an advanced probabilistic learning algorithm. Section 6 is devoted to the construction of the reference optimal solution with the stochastic computational model for which a stochastic convergence analysis has been performed. Section 7 deals with the computation of the optimal solution by using the probabilistic learning algorithm and Section 8 presents an analysis of the CPU-time gain. Additional details related to the construction of the non-Gaussian tensor-valued random field of the cortical bone medium are given in Appendix A. In order to confer a good readability to the paper, the probabilistic learning algorithm is summarized in Appendices B and C.

Notations

A lower case letter such as y, η, or u, is a real deterministic variable.

A boldface lower case letter such as $\mathbf{y}$, $\boldsymbol{\eta}$, or $\mathbf{u}$ is a real deterministic vector.

An upper case letter such as Y, H, or U, is a real random variable.

A boldface upper case letter, $\mathbf{Y}$, $\mathbf{H}$, or $\mathbf{U}$, is a real random vector.

A lower case letter between brackets such as $[y]$, $[\eta]$, or $[u]$), is a real deterministic matrix.

A boldface upper case letter between brackets such as $[\mathbf{Y}]$, $[\mathbf{H}]$, or $[\mathbf{U}]$, is a real random matrix.

$\mathbb{R}$: set of all the real numbers.
$\mathbb{R}^n$: Euclidean vector space on $\mathbb{R}$ of dimension n.
$\|\mathbf{x}\|$: usual Euclidean norm in $\mathbb{R}^n$.
$\mathbb{M}_{n,N}$: set of all the $(n \times N)$ real matrices.
$\mathbb{M}_n$: set of all the square $(n \times n)$ real matrices.
$\mathbb{M}_n^+$: set of all the positive-definite symmetric $(n \times n)$ real matrices.
$[y]_{kj}$: entry of matrix $[y]$.
$[y]^T$: transpose of matrix $[y]$.
$[I_n]$: identity matrix in $\mathbb{M}_n$.
δ: dispersion parameter $\delta_{kk'}$: Kronecker's symbol such that $\delta_{kk'} = 0$ if $k \neq k'$ and $= 1$ if $k = k'$.
E: Mathematical expectation.
pdf: probability density function.
CNOPPU: Constrained Nonconvex Optimization Problem in Presence of Uncertainties.
ISDE: Itô Stochastic Differential Equation.
MCMC: Markov Chain Monte Carlo.
PCA: Principal Component Analysis.
SCM: Stochastic Computational Model.

2 Definition of the optimum design problem under uncertainties

In this paper, we have voluntary chosen a geometry relatively simple in order that all data can be described and that the computation can be reproduced by any people. Nevertheless, the case proposed for the analysis contains all the difficulties that can be encountered in such a problem, which requires solving a CNOPPU.

2.1 Mechanical framework associated with a mesoscale implant in a cortical bone

The physical space $\mathbb{R}^3$ is referred to a cartesian reference system $Ox_1x_2x_3$ and the generic point of $\mathbb{R}^3$ is denoted by $\mathbf{x} = (x_1, x_2, x_3)$. For any function $g(\mathbf{x})$, the notation $g_{,j}$ designates the partial derivative with respect to x_j. The classical convention for summations over repeated Latin indices is used. The mean direction of the Haversian canals of the osteons of the cortical bone is the longitudinal direction Ox_1. In the cortical bone medium at mesoscale, since the spatial correlation length λ_1 in the longitudinal direction Ox_1 is much more larger than the spatial-correlation lengths λ_2 and λ_3 in the transversal directions Ox_2 and Ox_3, we consider a 2.5D-problem that is a layer with thickness $L_1 = 10^{-4}\, m$, which is analyzed as a 3D-problem. At mesoscale, we thus consider a 3D open bounded domain Ω that is constituted of a titanium implant occupying the open bounded domain Ω_{Ti}, inserted in a cortical bone (biological tissue) occupying the open bounded domain Ω_{cb}. Domain Ω is a rectangular parallelepiped for which the length sides are $L_1 = 10^{-4}\, m$, $L_2 = L_3 = 10^{-3}\, m$ (see Fig. 1). At mesoscale, the mechanical properties of the two media are the following.

- The titanium domain Ω_{Ti} is assumed to be a linear elastic homogeneous isotropic material for which the elasticity tensor is a constant tensor that is denoted by $\mathbb{C}^{\text{Ti}}$ and for which the Young modulus is $E^{\text{Ti}} = 110 \times 10^9\, Pa$ and the Poisson coefficient is $\nu^{\text{Ti}} = 0.35$.

- At microscale (see Fig. 2), the cortical bone that occupies domain Ω_{cb} is a linear elastic heterogeneous material that cannot be described in terms of its constituents. Therefore, its microstructure is described by a random medium at the mesoscale, for which its apparent elasticity field is modeled by a non-Gaussian fourth-order tensor-valued random field, $\mathbb{C}^{\text{cb}} = \{\mathbb{C}^{\text{cb}}(\mathbf{x}), \mathbf{x} \in \Omega_{\text{cb}}\}$ defined on a probability space $(\Theta, \mathcal{T}, \mathcal{P})$. Note that due to the positiveness property of the elasticity tensor, the random elasticity field cannot be a Gaussian random field [82].
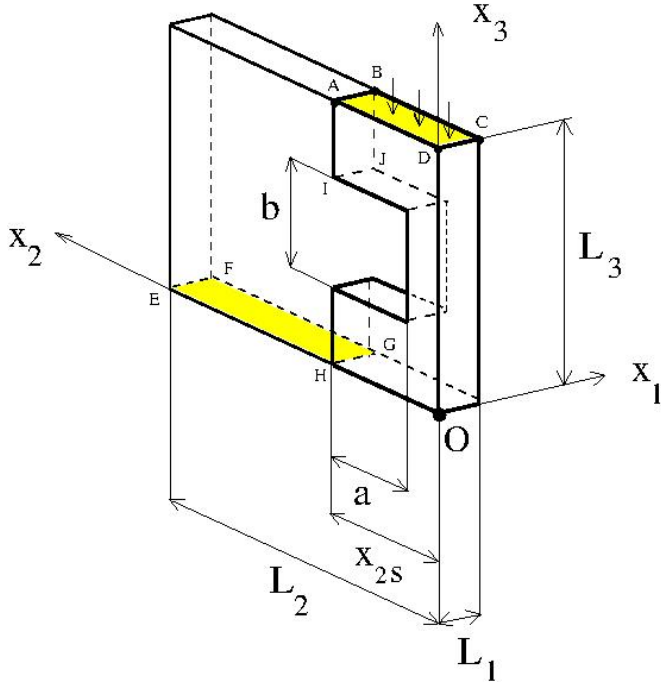
Fig. 1 Parameterization of the geometry for a tinanium implant occupying domain Ω_{Ti} (right part of the scheme) in a biological tissue occupying domain Ω_{cb} (left part of the scheme).
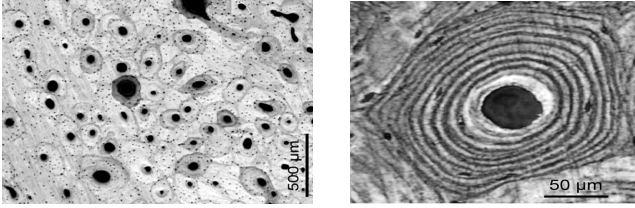
Fig. 2 Microstructure of a cortical bone at scale 5×10^{-4} m (left) and one osteon at scale 5×10^{-5} m. [Photo from Julius Wolff Institute, Charité - Universitatsmedizin Berlin]

2.2 Formulation of the constrained nonconvex optimal design problem under uncertainties

Boundary value problem. The elastostatic system is assumed to be in linear deformation around a static equilibrium state taken as a natural state at rest. At equilibrium, the microstructure occupies the three-dimensional open bounded domain Ω and the displacement field defined on Ω is denoted by $\mathbf{u}(\mathbf{x}) = (u_1(\mathbf{x}), u_2(\mathbf{x}), u_3(\mathbf{x}))$. The boundary $\partial\Omega$ of Ω is written as $\Gamma \cup \Gamma_0 \cup \Gamma_1$. The outward unit normal to $\partial\Omega$ is denoted by $\mathbf{n} = (n_1, n_2, n_3)$. The part Γ_0 is the plane surface $EFGH$ belonging to the boundary of Ω_{cb}, on which there is a zero Dirichlet condition for $\mathbf{u}$ (see Fig. 1). The part Γ_1 is the plane surface $ABCD$ belonging to the boundary of Ω_{Ti}, on which a random surface force field $\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}), F_3(\mathbf{x}))$ is applied (see Fig. 1). On the complementary part $\Gamma = \partial\Omega \backslash (\Gamma_0 \cup \Gamma_1)$, there is neither applied force field nor Dirichlet condition. Let $\sigma = \{\sigma_{ij}\}_{ij}$ be the symmetric stress tensor and let $\Gamma_c = \partial\Omega_{\mathrm{cb}} \cap \partial\Omega_{\mathrm{Ti}}$ be the coupling interface between the cortical bone and the titanium implant. It is assumed that the displacements $\mathbf{u}$ and the surface-forces interaction $\sigma\mathbf{n} = \{\sigma_{ij}n_j\}_i$ are continuous through Γ_c. In order to be able to define the objective function, we introduce $\Gamma_2 \subset \Gamma_c$ as the plane surface $ABIJ$ that is a part of the coupling interface Γ_c (see Fig. 1).

The stochastic boundary value problem is written, for $i = 1, 2, 3$, as

$$-\sigma_{ij,j} = 0 \quad \text{in} \quad \Omega, \tag{1}$$

$$\sigma_{ij}n_j = F_j \quad \text{on} \quad \Gamma_1, \tag{2}$$

$$u_i = 0 \quad \text{on} \quad \Gamma_0. \tag{3}$$

The stress tensor is written as $\sigma_{ij} = \mathbb{C}_{ijkh}\,\varepsilon_{kh}$ in which the symmetric strain tensor ε_{kh} is such that $\varepsilon_{kh} = \frac{1}{2}(u_{k,h}(\mathbf{x}) + u_{h,k}(\mathbf{x}))$. The elasticity field $\mathbb{C}$ is such that

$$\mathbb{C}(\mathbf{x}) = \mathbb{C}^{\mathrm{Ti}} \quad, \quad \forall \mathbf{x} \in \Omega_{\mathrm{Ti}}, \tag{4}$$

$$\mathbb{C}(\mathbf{x}) = \mathbb{C}^{\mathrm{cb}}(\mathbf{x}) \quad, \quad \forall \mathbf{x} \in \Omega_{\mathrm{cb}}. \tag{5}$$

Definition of the random surface force field $\mathbf{F}$. The random field $\{\mathbf{F}(\mathbf{x}), \mathbf{x} \in \Gamma_1\}$ is also defined on probability space $(\Theta, \mathcal{T}, \mathcal{P})$, is statistically independent of random elasticity field $\mathbb{C}^{\mathrm{cb}}$, and is defined as follows. Let $\mathbf{f} = (f_1, f_2, f_3)$ be the $\mathbb{R}^3$-valued random variable independent of $\mathbb{C}^{\mathrm{cb}}$ such that

$$f_1 = 0 \quad, \quad f_2 = -2\cos\Phi \quad, \quad f_3 = -2\sin\Phi,$$

in which $\Phi = (\pi/2)\,\mathcal{U}$ where $\mathcal{U}$ is a uniform random variable on $[0, 1]$. Let $\|\mathbf{f}\|$ be the Euclidean norm of $\mathbf{f}$. The random surface force field $\mathbf{F}$ indexed by Γ_1 is such that

$$\mathbf{F}(\mathbf{x}) = \frac{100}{|\Gamma_1|\|\mathbf{f}\|}\,\mathbf{f} \quad, \quad \mathbf{x} \in \Gamma_1,$$

which is a random force expressed in Newton (N) per square meter (m^2) and where $|\Gamma_1| = \int_{\Gamma_1} ds(\mathbf{x})$ is the area of surface Γ_1. It can be seen that

$$\int_{\Gamma_1} \|\mathbf{F}(\mathbf{x})\|\, ds(\mathbf{x}) = 100 \quad N \quad a.s.$$

Note that random surface force field $\mathbf{F}$ is independent of $\mathbf{x}$, will be independent of the design parameters (because x_{2s} will be fixed), belongs to plane $O\,x_2\,x_3$, has a random angle with values in $[0, \pi/2]$ in this plane, and finally, $F_2(\mathbf{x})$ and $F_3(\mathbf{x})$ are negative almost surely. This random surface force field induces a traction of the normal stress through the part Γ_2 of coupling interface Γ_c.

Definition of the design parameters. We refer the reader to Fig. 1 for the definition of the parameterization of the geometry. The length x_{2s} is fixed to the value $0.4 \times L_2$. There are two design parameters, the length a (channel depth) that belongs to the admissible set $\mathcal{C}_a$ and the length b (thread or

channel width) that belongs to the admissible set $\mathcal{C}_b$ such that

$$\mathcal{C}_a = [0.1\,,0.298] \times L_2 \quad , \quad \mathcal{C}_b = [0.85\,,0.898] \times L_3\,. \quad (6)$$

For any given value of a and b, the geometry of domain Ω_{Ti} and the geometry of domain Ω_{cb} are symmetric with respect to the plane of equation $x_3 = L_3/2$. Let $\mathbf{w} = (w_1, w_2)$ be the two-dimensional design parameter, such that $w_1 = a$ and $w_2 = b$, with values in the admissible set $\mathcal{C}_{\mathbf{w}}$ such that

$$\mathcal{C}_{\mathbf{w}} = \mathcal{C}_a \times \mathcal{C}_b \subset \mathbb{R}^2\,. \quad (7)$$

Definition of the objective function. The objective function is related to the normal stress to the part Γ_2 of coupling interface Γ_c between the cortical bone and the titanium implant. For all $\mathbf{w}$ fixed in $\mathcal{C}_{\mathbf{w}}$, the objective function $J(\mathbf{w})$ is defined by

$$J(\mathbf{w}) = \sqrt{E\{\mathcal{Q}(\mathbf{w})^2\}}\,, \quad (8)$$

in which the positive-valued random variable $\mathcal{Q}(\mathbf{w})$ is defined on $(\Theta, \mathcal{T}, \mathcal{P})$ and is written as

$$\mathcal{Q}(\mathbf{w}) = \max_{\mathbf{x} \in \Gamma_2} |\sigma_{22}(\mathbf{x}; \mathbf{w})|\,, \quad (9)$$

where $\sigma_{22}(\mathbf{x}; \mathbf{w})$ is the normal component along x_2 of the random stress tensor at point $\mathbf{x}$ in Γ_2 (which depends on design parameter $\mathbf{w}$) and where E is the mathematical expectation.

Definition of the constraint function. Let Ω_{cb0} be the open subset of Ω_{cb}, which is defined as the set of all the points $\mathbf{x} = (x_1, x_2, x_3)$ such that $0 < x_1 < L_1$, $x_{2s} - a < x_2 < x_{2s}$, and $(L_3 - b)/2 < x_3 < (L_3 + b)/2$. The constraint function $\mathbf{w} \mapsto \mathbf{c}(\mathbf{w}) = (c_1(\mathbf{w}), c_2(\mathbf{w}))$ is the mapping from $\mathcal{C}_{\mathbf{w}}$ into $\mathbb{R}^2$, such that $\mathbf{c}(\mathbf{w}) < 0$ that means $c_1(\mathbf{w}) < 0$ and $c_2(\mathbf{w}) < 0$. The two components of $\mathbf{c}(\mathbf{w})$ are related to the maximum of the absolute value $|\sigma_{13}(\mathbf{x}; \mathbf{w})|$ of the shear stress for $\mathbf{x} \in \Omega_{\mathrm{cb0}}$ and to the maximum of the Von Mises stress $\sigma_{\mathrm{VM}}(\mathbf{x}; \mathbf{w}) = 2^{-1/2}\{(\kappa_1 - \kappa_2)^2 + (\kappa_2 - \kappa_3)^2 + (\kappa_3 - \kappa_1)^2\}^{1/2}$ for $\mathbf{x} \in \Omega_{\mathrm{cb0}}$ in which κ_α is a notation for $\kappa_\alpha(\mathbf{x}; \mathbf{w})$ that are the three eigenvalues of the stress matrix $[\sigma(\mathbf{x}; \mathbf{w})]_{ij} = \sigma_{ij}(\mathbf{x}; \mathbf{w})$. For all $\mathbf{w}$ fixed in $\mathcal{C}_{\mathbf{w}}$, the constraint functions are thus defined by

$$c_k(\mathbf{w}) = \sqrt{E\{\mathcal{B}_k(\mathbf{w})^2\}} - c_k^{\mathrm{lim}} \quad , \quad k \in \{1, 2\}\,, \quad (10)$$

in which $c_1^{\mathrm{lim}} = 5.0 \times 10^9\,Pa$ and $c_2^{\mathrm{lim}} = 1.8 \times 10^{10}\,Pa$. The positive-valued random variables $\mathcal{B}_1(\mathbf{w})$ and $\mathcal{B}_2(\mathbf{w})$ are defined on probability space $(\Theta, \mathcal{T}, \mathcal{P})$ and are such that

$$\mathcal{B}_1(\mathbf{w}) = \max_{\mathbf{x} \in \Omega_{\mathrm{cb0}}} |\sigma_{13}(\mathbf{x}; \mathbf{w})|\,, \quad (11)$$

$$\mathcal{B}_2(\mathbf{w}) = \max_{\mathbf{x} \in \Omega_{\mathrm{cb0}}} \sigma_{\mathrm{VM}}(\mathbf{x}; \mathbf{w})\,. \quad (12)$$

Formulation of the optimization problem under uncertainties. We formulate the optimization problem under uncertainties for any dimension m_w of the design parameter $\mathbf{w} = (w_1, \ldots, w_{m_w})$ that belongs to the admissible set $\mathcal{C}_{\mathbf{w}} \subset \mathbb{R}^{m_w}$ and for any value m_c of the number of components of constraint function $\mathbf{c}(\mathbf{w}) = (c_1(\mathbf{w}), \ldots, c_{m_c}(\mathbf{w}))$ from $\mathcal{C}_{\mathbf{w}}$ into $\mathbb{R}^{m_c}$. In our case, we have $m_w = 2$ and $m_c = 2$. We consider the following constrained nonconvex optimization problem in presence of uncertainties with nonlinear constraints,

$$\mathbf{w}^{\mathrm{opt}} = \arg \min_{\substack{\mathbf{w} \in \mathcal{C}_{\mathbf{w}} \\ \mathbf{c}(\mathbf{w}) < 0}} J(\mathbf{w})\,. \quad (13)$$

Remark on the existence of a second-order solution. Let $\mathcal{Q} = \{\mathcal{Q}(\mathbf{w}), \mathbf{w} \in \mathcal{C}_{\mathbf{w}}\}$ and $\mathcal{B} = \{\mathcal{B}(\mathbf{w}) = (\mathcal{B}_1(\mathbf{w}), \ldots, \mathcal{B}_{m_c}(\mathbf{w})), \mathbf{w} \in \mathcal{C}_{\mathbf{w}}\}$ be the dependent stochastic processes on $(\Theta, \mathcal{T}, \mathcal{P})$, indexed by $\mathcal{C}_{\mathbf{w}}$, with values in $\mathbb{R}$ and $\mathbb{R}^{m_c}$ respectively, defined by Eqs. (9), (11), and (12). The stochastic modeling of non-Gaussian random field $\mathbb{C}^{\mathrm{cb}}$, presented in Section 3, is constructed in order that the stochastic elliptic boundary value problem defined by Eqs. (1) to (5) has a unique second-order stochastic solution such that the non-Gaussian dependent stochastic processes $\mathcal{Q}$ and $\mathcal{B}$ are second-order stochastic processes, that is to say, for all $\mathbf{w}$ in $\mathcal{C}_{\mathbf{w}}$,

$$E\{\mathcal{Q}(\mathbf{w})^2\} = \int_{\Theta} \mathcal{Q}(\mathbf{w}; \theta)^2\,d\mathcal{P}(\theta) < +\infty\,,$$

$$E\{\|\mathcal{B}(\mathbf{w})\|^2\} = \int_{\Theta} \|\mathcal{B}(\mathbf{w}; \theta)\|^2\,d\mathcal{P}(\theta) < +\infty\,.$$

Consequently, for all $\mathbf{w}$ in $\mathcal{C}_{\mathbf{w}}$, $J(\mathbf{w})$ and $\mathbf{c}(\mathbf{w})$ defined by Eqs. (8) and (10) are finite.

3 Stochastic modeling of the biological tissue

In order to complete the probabilistic modeling of the boundary value problem, we have to properly define the stochastic model of the elasticity field of the cortical bone at mesoscale.

3.1 Random matrix representation of the fourth-order tensor-valued random elasticity field for the cortical bone medium

At mesoscale, the probabilistic model of the apparent elasticity field $\{\mathbb{C}^{\mathrm{cb}}(\mathbf{x}), \mathbf{x} \in \Omega_{\mathrm{cb}}\}$ is carried out by introducing a prior stochastic model for which the construction is based on the following hypotheses.

1. Its statistical mean value $\underline{\mathbb{C}}^{\mathrm{cb}} = E\{\mathbb{C}^{\mathrm{cb}}(\mathbf{x})\}$ is assumed to be independent of $\mathbf{x}$ and is also assumed to belong to the material symmetry class of the transversely isotropic elastic materials for which the symmetry axis is $O\,x_1$.

2. Around the mean value, the statistical fluctuations field, $\{\mathbb{C}^{\mathrm{cb}}(\mathbf{x}) - \underline{\mathbb{C}}^{\mathrm{cb}}, \mathbf{x} \in \Omega_{\mathrm{cb}}\}$ is a non-Gaussian second-order centered fourth-order tensor-valued random field that belongs to the material symmetry class of the anisotropic elastic materials.

3. Random field $\{\mathbb{C}^{\mathrm{cb}}(\mathbf{x}), \mathbf{x} \in \Omega_{\mathrm{cb}}\}$ is the restriction to bounded domain Ω_{cb} of a statistically homogeneous (stationary) random field $\{\mathbb{C}^{\mathrm{cb}}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ indexed by $\mathbb{R}^d$ with $d = 3$ (we keep the same notation for the random field indexed by $\mathbb{R}^d$ and its restriction to Ω_{cb}).

In order to simplify the presentation, we introduce the random elasticity field $\{[\mathbf{K}(\mathbf{x})], \mathbf{x} \in \mathbb{R}^d\}$, indexed by $\mathbb{R}^d$ with $d = 3$, with values in $\mathbb{M}_n^+$ with $n = 6$, such that the fourth-order elasticity tensor $\{\mathbb{C}^{\mathrm{cb}}(\mathbf{x}), \mathbf{x} \in \Omega^{\mathrm{cb}}\}$ is written as

$$\mathbb{C}_{ijkh}^{\mathrm{cb}}(\mathbf{x}) = [\mathbf{K}(\mathbf{x})]_{IJ} \quad , \quad \forall \mathbf{x} \in \Omega^{\mathrm{cb}}, \tag{14}$$

in which the indices $I = (i, j)$ and $J = (k, h)$ belong to the set $\{1, \ldots, 6\}$ (Kelvin matrix representation). Eq. (14) means that random field $\mathbb{C}^{\mathrm{cb}}$ indexed by $\Omega^{\mathrm{cb}} \subset \mathbb{R}^d$ is represented by the restriction to Ω^{cb} of the statistically homogeneous random field $[\mathbf{K}]$ indexed by $\mathbb{R}^d$.

3.2 Prior stochastic model of the elasticity field for the cortical bone medium

In this section, in order to give a sufficient readability of this paper, the prior stochastic model that is used is briefly summarized and additional details are given in Appendix A. For the mesoscale stochastic modeling of apparent elasticity field $\{[\mathbf{K}(\mathbf{x})], \mathbf{x} \in \mathbb{R}^d\}$, we use the prior stochastic model initially introduced in [82], for which the parameterization consists of the spatial-correlation lengths, a dispersion parameter, and a positive-definite lower bound [83,86]. This model has been used for modeling and identifying cortical bone in [61,63]. Note that extensions of this model can be found in [30,31,86] for some positive-definite lower and upper bounds introduced as constraints and also for random fields with any material symmetry property. Let $[\underline{K}] = E\{[\mathbf{K}(\mathbf{x})]\}$ be the mean matrix in $\mathbb{M}_n^+$, which corresponds to the positive-definite fourth-order tensor $\underline{\mathbb{C}}^{\mathrm{cb}}$ such that $\underline{\mathbb{C}}_{ijkh}^{\mathrm{cb}} = [\underline{K}]_{IJ}$. The Cholesky factorization of matrix $[\underline{K}]$ is written as

$$[\underline{K}] = [\underline{\mathcal{L}}]^T [\underline{\mathcal{L}}]. \tag{15}$$

Introduction of an adapted representation. The random field $\{[\mathbf{K}(\mathbf{x})], \mathbf{x} \in \mathbb{R}^d\}$, defined on probability space $(\Theta, \mathcal{T}, \mathcal{P})$, indexed by $\mathbb{R}^d$, with values in $\mathbb{M}_n^+$, is a non-Gaussian homogeneous second-order random field such that

$$[\mathbf{K}(\mathbf{x})] = [K_\ell] + [\underline{\mathcal{L}}_\varepsilon]^T [\mathbf{G}_0(\mathbf{x})] [\underline{\mathcal{L}}_\varepsilon] \quad , \quad \forall \mathbf{x} \in \mathbb{R}^d. \tag{16}$$

In Eq. (16), the positive-definite symmetric $(n \times n)$ real matrix $[K_\ell]$ that is independent of $\mathbf{x}$ is the lower bound.

For the heterogeneous material that we are considering at mesoscale, there is no available information concerning such a lower bound. Therefore, the following model is used for preserving the uniform ellipticity of the elasticity differential operator: the upper triangular $(n \times n)$ real matrix $[\underline{\mathcal{L}}_\varepsilon]$ and the positive-definite symmetric $(n \times n)$ real matrix $[K_\ell]$, which are independent of $\mathbf{x}$, are written as

$$[\underline{\mathcal{L}}_\varepsilon] = \frac{1}{\sqrt{1 + \varepsilon}} [\underline{\mathcal{L}}] \quad , \quad [K_\ell] = \frac{\varepsilon}{1 + \varepsilon} [\underline{K}], \tag{17}$$

in which $\varepsilon > 0$ is any fixed positive real number (for instance, $\varepsilon = 10^{-6}$). In Eq. (16), the random field $\{[\mathbf{G}_0(\mathbf{x})], \mathbf{x} \in \mathbb{R}^d\}$ is defined on $(\Theta, \mathcal{T}, \mathcal{P})$, is indexed by $\mathbb{R}^d$ with values in $\mathbb{M}_n^+$, is homogeneous on $\mathbb{R}^d$, and is a non-Gaussian second-order random field such that, for all $\mathbf{x}$ in $\mathbb{R}^d$,

$$E\{[\mathbf{G}_0(\mathbf{x})]\} = [I_n] \quad , \quad [\mathbf{G}_0(\mathbf{x})] > 0 \quad \text{a.s.} \tag{18}$$

The construction of the prior stochastic model of the non-Gaussian random field $[\mathbf{G}_0]$ and its generator of realizations are detailed in Appendix A. Finally, it can easily be seen that, for all $\mathbf{x}$ in $\mathbb{R}^d$,

$$E\{[\mathbf{K}(\mathbf{x})]\} = [\underline{K}] \in \mathbb{M}_n^+ \ , \ [\mathbf{K}(\mathbf{x})] - [K_\ell] > 0 \quad \text{a.s.}, \tag{19}$$

in which the notation, $[\mathbf{A}] > 0$ a.s., means that all the random eigenvalues of the symmetric real random matrix $[\mathbf{A}]$ are strictly positive almost surely.

Summarizing the hyperparameters of the prior stochastic model. Let $\mathbf{s} = (s_1, \ldots, s_\mu)$ be the vector of the hyperparameters of the prior stochastic model $\{[\mathbf{K}(\mathbf{x}; \mathbf{s})], \mathbf{x} \in \mathbb{R}^d\}$ that has been constructed for a transversely isotropic mean model with anisotropic statistical fluctuations. Vector $\mathbf{s}$ belongs to $\mathcal{C}_{\mathrm{ad}} \subset \mathbb{R}^\mu$ and its components are constituted of

1. the reshaping of $[\underline{K}] \in \mathbb{M}_n^+(\mathbb{R})$. Since the mean value is assumed to be in the transversely isotropic class for which the symmetry is with respect to a rotation around symmetry axis Ox_1, matrix $[\underline{K}]$ depends only on the five mechanical parameters of the constitutive equation: E_1, ν_{12}, and G_{12} for the longitudinal direction Ox_1 and, E_2 and ν_{23} for the transverse direction Ox_2.

2. the three spatial-correlation lengths, $\lambda_1, \lambda_2, \lambda_3$, which are defined in Appendix A.

3. the dispersion parameter δ that allows the statistical fluctuations to be controlled in the anisotropic class, which must be such that $0 < \delta < \sqrt{(n+1)/(n+5)}$, and which is defined in Appendix A.

3.3 Experimental identification of the prior stochastic model of the cortical bone medium at mesoscale

For the prior stochastic model of the apparent elasticity field of the cortical bone medium at mesoscale, we reuse the mechanical properties and the values of the hyperparameters

that have experimentally been identified in [63,64] by the means of a multiscale experiment based on a digital image correlation technique (see Fig. 3). For such an experiment, only a single specimen was measured, was submitted to a given load applied at the macroscale, and was tested. The displacement field was measured at the macroscale (with a spatial resolution $10^{-3}\,m$), and simultaneously, was measured at the mesoscale (with a spatial resolution $10^{-5}\,m$). The measured strain fields were directly deduced from the measured displacement fields. The experimental identification of the hyperparameters of the prior stochastic model has given the following values for the cortical bone at mesoscale:
- for the transversely isotropic parameters of the mean elasticity tensor: $E_1 = 16.6 \times 10^9\,Pa$, $\nu_{12} = 0.37$, $G_{12} = 4.7 \times 10^9\,Pa$, $E_2 = 9.5 \times 10^9\,Pa$, and $\nu_{23} = 0.37$.
- for the 3 spatial-correlation lengths: $\lambda_1 = 10^{-2}\,m$, $\lambda_2 = \lambda_3 = 5 \times 10^{-5}\,m$.
- for the dispersion parameter $\delta = 0.28$.

4 Stochastic computational model

The stochastic computational model (SCM) is obtained by the finite element discretization of the stochastic boundary value problem defined by Eqs. (1) to (5). This finite element model is constructed using a regular cartesian mesh of $11 \times 101 \times 101 = 112{,}211$ nodes and $10 \times 100 \times 100 = 100{,}000$ finite elements that are 8-nodes solid elements with $2 \times 2 \times 2$ Gauss-Legendre quadrature points. This mesh is fixed and then independent of the value of the design parameter $\mathbf{w}$. For any given value of the design parameter $\mathbf{w}$ belonging to its admissible set $\mathcal{C}_{\mathbf{w}}$, the cortical bone domain $\Omega_{\text{Ti}}(\mathbf{w})$ and the titanium implant domain $\Omega_{\text{cb}}(\mathbf{w})$ depend on $\mathbf{w}$ and are defined in order that the coupling interface $\Gamma_c(\mathbf{w}) = \partial\Omega_{\text{cb}}(\mathbf{w}) \cap \partial\Omega_{\text{Ti}}(\mathbf{w})$ between these two domains coincides with the mesh faces of the finite element mesh of Ω. Before applying the zero Dirichlet condition on Γ_0, there are $336{,}333$ degrees of freedom and $800{,}000$ Gauss-Legendre quadrature points.

Random field $\{[\mathbf{K}(\mathbf{x})]\,,\mathbf{x}\in\Omega_{\text{cb}}(\mathbf{w})\}$ is discretized at all the Gauss-Legendre quadrature points $\{\mathbf{x}^\alpha(\mathbf{w})\}_\alpha$ belonging to $\Omega_{\text{cb}}(\mathbf{w})$. In the three directions, the mesh sizes are sufficient small with respect to the spatial correlation lengths of

Fig. 3 The specimen of cortical bone is a cube with dimensions $0.01 \times 0.01 \times 0.01\,m^3$ (left figure). Measuring bench at LMS of Ecole Polytechnique (right figure). [Figures from [64]].

the random elasticity field for guarantying the mean-square convergence of the integrals related to the computation of the random elementary stiffness matrices (see [83]).

For any value $\mathbf{w}^\ell$ of $\mathbf{w}$ and for θ'_ℓ and θ''_ℓ in Θ, let $\mathbf{F}(\cdot;\theta'_\ell)$ be a realization of the random surface force field $\mathbf{F}$ (that is independent of $\mathbf{w}$) and let $\{[\mathbf{K}(\mathbf{x}^\alpha(\mathbf{w}^\ell));\theta''_\ell)]\}_\alpha$ be a realization of the discretized random elasticity field. The corresponding realization $q^\ell = \mathcal{Q}(\mathbf{w}^\ell;\theta'_\ell,\theta''_\ell)$ of random variable $\mathcal{Q}(\mathbf{w}^\ell)$ and the corresponding realization $\mathbf{b}^\ell = \boldsymbol{\mathcal{B}}(\mathbf{w}^\ell;\theta'_\ell,\theta''_\ell)$ of random variable $\boldsymbol{\mathcal{B}}(\mathbf{w}^\ell)$ are then computed with the SCM by solving a linear matrix equation for the discretized displacement field and by using Eqs. (9), (11), and (12).

5 Solving the optimization design problem under uncertainties with a probabilistic learning algorithm

5.1 Major difficulties encountered for solving the optimization problem under uncertainties

We start this section in describing the most classical method that will allow for constructing the reference solution but for which the numerical cost is very high. Such a method has been evoked in Section 1. The optimization problem defined by Eq. (13) is solved by using an optimization algorithm such as the grid search. For each value of $\mathbf{w}$ given in $\mathcal{C}_{\mathbf{w}}$, which is proposed by the optimization algorithm, the mathematical expectations $E\{\mathcal{Q}(\mathbf{w})^2\}$ and $E\{\mathcal{B}_k(\mathbf{w})^2\}$ are estimated with the Monte Carlo method as follows. For every $\mathbf{w}$ given in $\mathcal{C}_{\mathbf{w}}$, N'_s independent realizations $\mathcal{Q}(\mathbf{w};\theta'_\ell,\theta''_\ell)$, $\mathcal{B}_1(\mathbf{w};\theta'_\ell,\theta''_\ell)$, and $\mathcal{B}_2(\mathbf{w};\theta'_\ell,\theta''_\ell)$ are computed with the SCM. For N'_s sufficiently large, an accurate estimation of the mathematical expectations can be computed according to,

$$E\{\mathcal{Q}(\mathbf{w})^2\} \simeq \tfrac{1}{N'_s} \sum_{\ell=1}^{N'_s} \mathcal{Q}(\mathbf{w};\theta'_\ell,\theta''_\ell)^2\,, \qquad (20)$$

$$E\{\mathcal{B}_k(\mathbf{w})^2\} \simeq \tfrac{1}{N'_s} \sum_{\ell=1}^{N'_s} \mathcal{B}_k(\mathbf{w};\theta'_\ell,\theta''_\ell)^2\,. \qquad (21)$$

It can be seen that, if the optimization algorithm requires N_s evaluations of the objective and constraint functions, then the SCM must be used $N'_s \times N_s$ times, which is generally very prohibitive. The probabilistic learning proposed in Section 5.2 will allow drastically for reducing the number of calls to the SCM by using only N calls with $N \ll N'_s \times N_s$.

5.2 Probabilistic learning algorithm proposed for circumvent the major difficulties

In this section, we summarize the methodology that we propose to use. The ingredients of the probabilistic learning algorithm and smoothing technique are the following.

(i) The first ingredient is the probabilistic learning on manifold that has been presented in [85,87]. This algorithm allows for generating additional realizations concentrated on
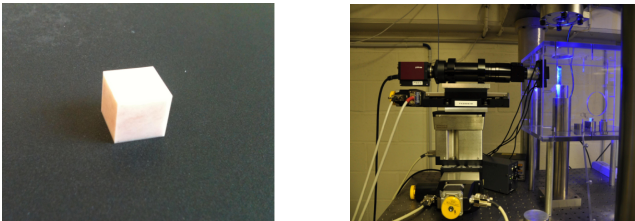
the manifold that is identified by using a dataset, without performing additional function evaluations by using the SCM.

(ii) The second ingredient is a smoothing technique [25, 88] for estimating $J(\mathbf{w}^0)$ and $\mathbf{c}(\mathbf{w}^0)$ at any point $\mathbf{w}^0$ in $\mathcal{C}_{\mathbf{w}}$, by using only the additional realizations generated from the given dataset. The points $\mathbf{w}^0$ are those that are generated for the optimization algorithm.

5.3 Probabilistic learning algorithm and smoothing technique for avoiding a large number of functions evaluations with the SCM

The algorithm is constituted of 3 steps.

Step 1: Construction of a dataset by using only N functions evaluations. Such a number N is defined as follows. Let us consider a fixed number N_0 of values $\widetilde{\mathbf{w}}^1, \ldots, \widetilde{\mathbf{w}}^{N_0}$ in $\mathcal{C}_{\mathbf{w}}$ of vector $\mathbf{w}$, which are the values of $\mathbf{w}$ generated by the optimization algorithm as it explores the admissible set $\mathcal{C}_{\mathbf{w}}$. For $\ell' = 1, \ldots, N_0$, for $\ell'' = 1, \ldots, N_{sd}$, for $\ell = (\ell', \ell'')$ in $\{1, \ldots, N\}$ with $N = N_0 \times N_{sd}$, and for $\theta'_{\ell''}$ and $\theta''_{\ell''}$ in Θ, let

$$\mathbf{w}^\ell = \widetilde{\mathbf{w}}^{\ell'} \ , \ q^\ell = \mathcal{Q}(\mathbf{w}^{\ell'}; \theta'_{\ell''}, \theta''_{\ell''}) \ , \ \mathbf{b}^\ell = \mathcal{B}(\mathbf{w}^{\ell'}; \theta'_{\ell''}, \theta''_{\ell''}),$$

(22)

be the realizations of the dependent random variables $\mathcal{Q}(\mathbf{w}^{\ell'})$ and $\mathcal{B}(\mathbf{w}^{\ell'})$ computed using the SCM, where N_{sd} is a number of realizations. In general, N_{sd} is chosen as 1 but, for increasing the quality of the geometry identification of the manifold on which is concentrated the probability distribution, N_{sd} can sometimes be chosen as a few units (for instance, $N_{sd} = 5$ or 10, but stays such that $N_{sd} \ll N'_s$. Let be $n = m_w + 1 + m_c$ and let us now introduce the N data points $\mathbf{y}^1, \ldots, \mathbf{y}^N$ in $\mathbb{R}^n$ such that, for $\ell = 1, \ldots, N$,

$$\mathbf{y}^\ell = (\mathbf{w}^\ell, q^\ell, \mathbf{b}^\ell) \in \mathbb{R}^n = \mathbb{R}^{m_w} \times \mathbb{R} \times \mathbb{R}^{m_c} \ .$$

(23)

Step 2: Construction of the diffusion-maps basis and generator of additional realizations. We introduce the random variable

$$\mathbf{Y} = (\mathbf{W}, Q, \mathbf{B}) \ ,$$

(24)

with values in $\mathbb{R}^n$ such that $\mathbf{y}^\ell = (\mathbf{w}^\ell, q^\ell, \mathbf{b}^\ell)$ are N independent realizations of $\mathbf{Y}$. Note that mathematical symbol Q must carefully be distinguished from $\mathcal{Q}$ defined by Eq. (9). The diffusion-maps basis is constructed by using $\{\mathbf{y}^\ell\}_{\ell=1, \ldots, N}$. Then, without performing additional function evaluations with the SCM, $\nu_{\text{sim}} \gg N$ additional realizations

$$\mathbf{y}^\ell_{\text{ar}} = (\mathbf{w}^\ell_{\text{ar}}, q^\ell_{\text{ar}}, \mathbf{b}^\ell_{\text{ar}}) \quad , \quad \ell = 1, \ldots, \nu_{\text{sim}}$$

(25)

are computed by solving a nonlinear Itô stochastic differential equation. The algorithm is summarized in Appendix B.

Step 3: Smoothing technique by using a nonparametric statistical estimation. For any point $\mathbf{w}^0$ in $\mathcal{C}_{\mathbf{w}}$, the values $J(\mathbf{w}^0)$ and $\mathbf{c}(\mathbf{w}^0)$ of J and $\mathbf{c}$ are calculated by using the following smoothing technique of the mathematical expectations, $E\{\mathcal{Q}(\mathbf{w}^0)^2\}$ and $E\{\mathcal{B}_k(\mathbf{w}^0)^2\}$:

$$E\{\mathcal{Q}(\mathbf{w}^0)^2\} \simeq E\{Q^2 \,|\, \mathbf{W} = \mathbf{w}^0\} \ ,$$

(26)

$$E\{\mathcal{B}_k(\mathbf{w}^0)^2\} \simeq E\{B_k^2 \,|\, \mathbf{W} = \mathbf{w}^0\} \ , \ k = 1, \ldots, m_c \ .$$

(27)

The calculation of the conditional mathematical expectations appearing in the right-hand sides of Eqs. (26) and (27) is carried out by using the kernel density estimation method with the additional realizations $\{\mathbf{y}^\ell_{\text{ar}} = (\mathbf{w}^\ell_{\text{ar}}, q^\ell_{\text{ar}}, \mathbf{b}^\ell_{\text{ar}}), \ell = 1, \ldots, \nu_{\text{sim}}\}$ computed in Step 2 without performing additional function evaluations with the SCM. For such a calculation, two explicit numerical formulas have been developed and are given in Appendix C.

5.4 Remark on the efficiency of the probabilistic learning algorithm with respect to the number of design parameters

There is no theoretical limit to the value of the number m_w of design parameters in the proposed probabilistic learning method. However, the larger the m_w number is and the larger the computing resource should be, but this is true for any optimization algorithm that is adapted for solving nonconvex problems, especially when a genetic algorithm is used. However, for the probabilistic learning algorithm to be effective, the number N of points in the dataset for which each point $\mathbf{y}^\ell$ is computed using the SCM must be adapted to the dimension m_w of the $\mathbf{w}$ parameter, so that the algorithm can actually learn. In general, if the number m_w of design parameters increases, then the number N of points will have to increase, but a genetic algorithm coupled to the proposed probabilistic learning algorithm, will always show a very important gain. For example, it would be quite feasible to do topological optimization with the proposed methodology.

6 Computation of the reference optimal solution with the SCM and convergence analysis of the stochastic solver

In this section, we compute the reference optimal solution by using Eqs. (20) and (21) with N'_s calls to the SCM as explained in Section 5.1. A convergence analysis with respect to N'_s is carried out in order to obtain the converged value of the reference optimal solution.

Optimizer. The optimization problem defined by Eq. (13) is solved by using the grid search algorithm. The admissible set $\mathcal{C}_{\mathbf{w}} = \mathcal{C}_a \times \mathcal{C}_b$ defined by Eq. (7) is replaced by the finite

subset $\mathcal{C}_{\mathbf{w}}^{\text{grid}}$ of $\mathcal{C}_{\mathbf{w}}$, which has been constructed by an adapted training procedure, such that

$$\mathcal{C}_{\mathbf{w}}^{\text{grid}} = \mathcal{C}_a^{\text{grid}} \times \mathcal{C}_b^{\text{grid}} \subset \mathcal{C}_{\mathbf{w}} \subset \mathbb{R}^2 \, ,$$

in which $\mathcal{C}_a^{\text{grid}} = L_2 \times \{0.1, 0.13, 0.16, 0.19, 0.197, 0.2, 0.203,$ $0.206, 0.209, 0.212, 0.216, 0.218, 0.221, 0.224, 0.227, 0.230,$ $0.232, 0.234, 0.237, 0.24, 0.25, 0.26, 0.28, 0.298\}$ is made up of 24 points and where $\mathcal{C}_b^{\text{grid}} = L_3 \times \{0.850, 0.862, 0.874,$ $0.886, 0.898\}$ is made up of 5 points. The number of points in $\mathcal{C}_{\mathbf{w}}^{\text{grid}}$ is thus $120 = 24 \times 5$ and has voluntary been limited so that the computation of the optimal reference solution can be effectively performed with a feasible elapse time (see Section 8). Note that pre-calculations have been done in order to study the shape of the objective function in order to validate the training procedure that has been done for defining the points in $\mathcal{C}_{\mathbf{w}}^{\text{grid}}$. The reference optimal solution is denoted by $\mathbf{w}_{\text{ref}}^{\text{opt}}$ and is obtained by solving the approximate optimization problem,

$$\mathbf{w}_{\text{ref}}^{\text{opt}} = \arg \min_{\substack{\mathbf{w} \in \mathcal{C}_{\mathbf{w}}^{\text{grid}} \\ \mathbf{c}(\mathbf{w}) < 0}} J(\mathbf{w}) \, .$$

Stochastic solver and convergence analysis of the reference optimal solution. The stochastic solver used for constructing the solution with the SCM is the Monte Carlo method. For each $\mathbf{w}$ belonging to the grid $\mathcal{C}_{\mathbf{w}}^{\text{grid}} \subset \mathcal{C}_{\mathbf{w}}$, Eqs. (20) and (21) are computed with N_s' independent realizations such that $N_s' \in [5, 175]$ (the closed interval is made of integers 5 to 175 by step of 5). The convergence analysis of the optimal solution $\mathbf{w}^{\text{opt}}(N_s')$ shows that the component $w_1^{\text{opt}}(N_s')$ of $\mathbf{w}^{\text{opt}}(N_s')$ is converged for $N_s' \geq 145$ while the component $w_2^{\text{opt}}(N_s')$ of $\mathbf{w}^{\text{opt}}(N_s')$ is converged as soon as $N_s' \geq 5$. Fig. 4 displays the graph of the function $N_s' \mapsto w_1^{\text{opt}}(N_s')$ for $N_s' \in [5, 175]$.

The reference optimal solution is thus defined as the optimal solution for $N_s' = 145$ and is denoted by $\mathbf{w}_{\text{ref}}^{\text{opt}} = (a_{\text{ref}}^{\text{opt}}, b_{\text{ref}}^{\text{opt}})$ with $a_{\text{ref}}^{\text{opt}} = w_1^{\text{opt}}(N_s')$ and $b_{\text{ref}}^{\text{opt}} = w_2^{\text{opt}}(N_s')$. We have obtained,

$$a_{\text{ref}}^{\text{opt}} = 2.60 \times 10^{-4} \, m \quad , \quad b_{\text{ref}}^{\text{opt}} = 8.50 \times 10^{-4} \, m \, .$$

The corresponding value of the objective function is

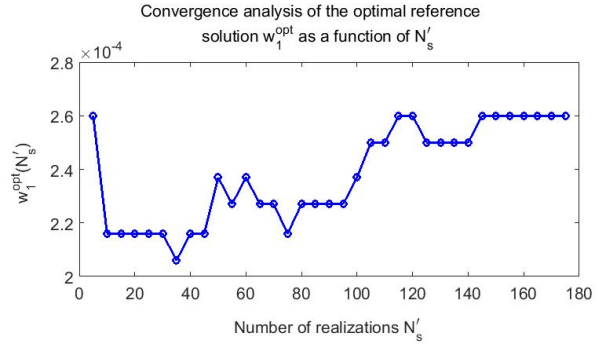$$J(\mathbf{w}_{\text{ref}}^{\text{opt}}) = 6.014 \times 10^9 \, Pa \, .$$

The constraint c_1 defined by Eqs. (10) and (11) is saturated and is such that

$$c_1(\mathbf{w}_{\text{ref}}^{\text{opt}}) = 5.00 \times 10^9 - 5.00 \times 10^9 = 0 \, Pa \, ,$$

while the constraint c_2 defined by Eqs. (10) and (12) is almost saturated and is such that

$$c_2(\mathbf{w}_{\text{ref}}^{\text{opt}}) = 1.77 \times 10^{10} - 1.80 \times 10^{10} = -0.03 \times 10^{10} \, Pa \, .$$

Remark. In Fig. 4, the "oscillations" that appear in the curve,

Fig. 4 Converge analysis of the reference optimal solution. Graph of $N_s' \mapsto w_1^{\text{opt}}(N_s')$ in which w_1 is the design parameter a of the implant and where $N_s' \in [5, 175]$ is the number of realizations used by the stochastic solver (Monte Carlo numerical method).

$N_s' \mapsto w_1^{\text{opt}}(N_s')$, are not directly due the mean-square convergence with respect to N_s' of the random solution $\{\mathbf{u}(\mathbf{x}; \mathbf{w}), \mathbf{x} \in \Omega\}$ for a fixed value of $\mathbf{w}$, but are relative to the function $N_s' \mapsto w_1^{\text{opt}}(N_s')$ that is the solution of the optimization problem defined by Eq. (13), which involves the family of random fields, $\{\{\mathbf{u}(\mathbf{x}; \mathbf{w}), \mathbf{x} \in \Omega\}, \mathbf{w} \in \mathcal{C}_{\mathbf{w}}, \mathbf{c}(\mathbf{w}) < 0\}$. For every fixed number N_s' of realizations for which the convergence is not reached, there is no reason for that the values of $\mathbf{w}$ that are proposed by the optimizer correspond to the same rate of convergence. Consequently, before the convergence be reached, the graph of function $N_s' \mapsto w_1^{\text{opt}}(N_s')$ can oscillated a lot before convergence with respect to N_s'. The convergence is stabilized for $N's \geq 110$ and is converged for $N's \geq 150$. On the other hand, the oscillations that appear in Figs. 5 to 7, for which the value of N_s' is fixed to a value for which the convergence of the optimal solution is reached, are not due the smoothing technique used for plotting the graphs. These oscillations are due to the variability of the random stress field in the cortical bone in the neighborhood of the implant when the geometry of the implant changes with $\mathbf{w}$.

3D visualization of the reference objective function and of the reference constraint functions. For the 3D visualization, we introduce a grid of points in $\mathcal{C}_{\mathbf{w}}$, which is defined as follows. Let $\mathcal{C}_g$ be the finite subset of $\mathcal{C}_{\mathbf{w}}$ made up of $N_g = 10{,}000$ points, $\{\mathbf{w}_g^i = (w_{g,1}^i, w_{g,2}^i), i = 1, \ldots, N_g\}$, obtained by a uniform cartesian mesh of 100×100 points of $\mathcal{C}_w = \mathcal{C}_a \times \mathcal{C}_b \subset \mathbb{R}^2$. For $N_s' = 145$, Fig. 5 displays the graph of function $\mathbf{w} \mapsto J(\mathbf{w})$ restricted to $\mathcal{C}_g$ while Figs. 6 and 7 display the graphs of functions $\mathbf{w} \mapsto c_1(\mathbf{w}) + c_1^{\text{lim}}$ and $\mathbf{w} \mapsto c_2(\mathbf{w}) + c_2^{\text{lim}}$ also restricted to $\mathcal{C}_g$. Note that for every point $\mathbf{w}_g^i$ in $\mathcal{C}_g$, $J(\mathbf{w}_g^i)$ and $\mathbf{c}(\mathbf{w}_g^i)$ are computed by using a smoothing technique similar to the one presented in Appendix C. This smoothing technique will not be detailed here, because this 3D visualization is only performed for a qualitative analysis of the geometrical surfaces that are rela-
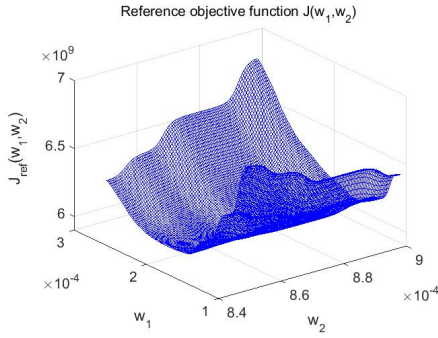
Fig. 5 Graph of the reference objective function $\mathbf{w} \mapsto J(\mathbf{w}) = \sqrt{E\{\mathcal{Q}(\mathbf{w})^2\}}$ computed with $N'_s = 145$ realizations for the stochastic solver (Monte Carlo numerical method) and for $\mathbf{w}$ belonging to the finite subset $\mathcal{C}_g$ of $\mathcal{C}_\mathbf{w}$ made up of $N_g = 10{,}000$ points $\mathbf{w}_g^i$.

tively complex and that show that the optimization problem is effectively nonconvex.

Remark 1. The reference optimal solution has been solved on $\mathcal{C}_\mathbf{w}^{\text{grid}}$ and not on $\mathcal{C}_g$ for the following reasons. The number of points in $\mathcal{C}_g$ is too large ($N_g = 10{,}000$) for being able to do the computation with a reasonable elapse time (see Section 8). If the optimal solution had been constructed as the minimum over the set $\{J(\mathbf{w}_g^i), i = 1, \ldots, N_g\}$, we would have constructed an approximation of the optimal solution due to the use of the smoothing technique and not the reference optimal solution.

Remark 2. The convergence analysis with respect to N'_s displayed in Fig. 4 shows that a significant number of realizations ($N'_s \geq 150$) is required for obtaining the convergence of the optimal solution. This means that the statistical fluctuations of $\mathcal{Q}(\mathbf{w})$ (related to the solution of the stochastic boundary value problem), induced by the statistical fluctuations of the apparent elasticity field of the heterogeneous medium at mesoscale ($\mathbf{x}$-dependent coefficients of the stochastic elliptic operator), are significant. This relatively large value of N'_s that is required for reaching convergence is the major reason for which the optimization problem is numerically expensive with a classical optimization algorithm (as explained at the end of Section 5.1), because, for each value $\mathbf{w}$ proposed by the optimizer, the stochastic computational model must be called N'_S times.

7 Computation of the optimal solution using the probabilistic learning algorithm

In this section, we compute an approximation of the optimal solution by using the probabilistic learning algorithm presented in Section 5.3 and we compare the approximation computed with a low numerical cost to the reference optimal
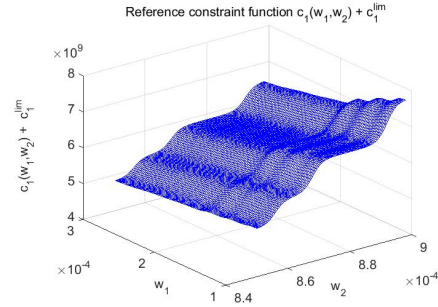
Fig. 6 Graph of the reference constraint function $\mathbf{w} \mapsto c_1(\mathbf{w}) + c_1^{\text{lim}}$ computed with $N'_s = 145$ realizations for the stochastic solver (Monte Carlo numerical method) and for $\mathbf{w}$ belonging to the finite subset $\mathcal{C}_g$ of $\mathcal{C}_\mathbf{w}$ made up of $N_g = 10{,}000$ points $\mathbf{w}_g^i$.
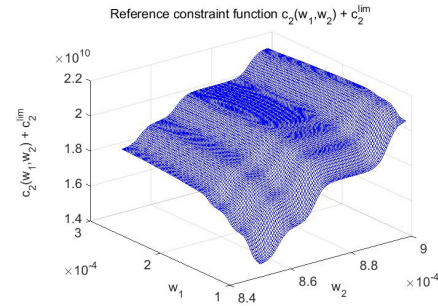
Fig. 7 Graph of the reference constraint function $\mathbf{w} \mapsto c_2(\mathbf{w}) + c_2^{\text{lim}}$ computed with $N'_s = 145$ realizations for the stochastic solver (Monte Carlo numerical method) and for $\mathbf{w}$ belonging to the finite subset $\mathcal{C}_g$ of $\mathcal{C}_\mathbf{w}$ made up of $N_g = 10{,}000$ points $\mathbf{w}_g^i$.

solution computed with a very high numerical cost.

Optimizer. As for the computation of the reference optimal solution presented in Section 6, the optimization problem defined by Eq. (13) is solved by using the grid search algorithm for which the admissible set $\mathcal{C}_\mathbf{w} = \mathcal{C}_a \times \mathcal{C}_b$ defined by Eq. (7) is replaced by the finite subset $\mathcal{C}_g$ of $\mathcal{C}_\mathbf{w}$ made up of $N_g = 10{,}000$ points $\{\mathbf{w}_g^i, i = 1, \ldots, N_g\}$ defined in Section 6.

Construction of the dataset by using only the N evaluations of the SCM. We consider $N = 600$ points $\mathbf{w}^1, \ldots, \mathbf{w}^N$ that are generated as follows. For $N_0 = 120$, let $\widetilde{\mathbf{w}}^1, \ldots, \widetilde{\mathbf{w}}^{N_0}$ be the points that are randomly generated with a uniform distribution on the subset $[0.1, 0.298] \times [0.8075, 0.898]$ of $\mathbb{R}^2$. For $N_{sd} = 5$, we generate the $N = N_0 \times N_{sd} = 120 \times 5 = 600$ points $\mathbf{w}^1, \ldots, \mathbf{w}^N$ as explained in Step 1 of Section 5.3 (see Eq. (22)). The N data points $\mathbf{y}^1, \ldots, \mathbf{y}^N$ with $\mathbf{y}^\ell = (\mathbf{w}^\ell, q^\ell, \mathbf{b}^\ell)$ in $\mathbb{R}^n$ defined by Eq. (23) are then computed by using the SCM with $\mathbf{w}^\ell$, q^ℓ, and $\mathbf{b}^\ell$ defined by Eq. (22).

Remark. As explained in Step 1 of Section 5.3, the value of N_{sd} must be chosen small enough, a few units. Since the geometry of the surface defined by the objective function is relatively complex (see Fig. 5), the grid search algorithm needs sufficient information for identifying the optimal solution. The minimum value, $N_{sd} = 1$, has been tried, but this value of N_{sd} is not sufficient for that the probabilistic learning algorithm sufficiently learns from the corresponding $N = N_0 \times N_{sd}$ points. This is the reason why, the value $N_{sd} = 5$ has been retained.

Diffusion-maps basis and generator of additional realizations. Step 2 of Section 5.3, for which the details of the algorithm are summarized in Appendix B, is used for constructing the additional realizations. In order to analyze the convergence of the optimal solution with respect to the value of parameter n_{MC}, the calculations have been carried out for $n_{\mathrm{MC}} \in \{50, 100, 150, 200, 250, 300\}$. Therefore, the number $\nu_{\mathrm{sim}} = n_{\mathrm{MC}} \times N$ of additional realizations vary in the interval $[3, 18] \times 10^4$. The additional realizations are generated without using the SCM as follows.

1. The $N = 600$ data points are normalized by using the PCA presented in Paragraph 1 of Appendix B. As $n = 5$ is small and as all the eigenvalues are strictly positive, ν is chosen equal to n and therefore, $\nu = 5$.
2. The diffusion-maps basis is computed as explained in Paragraph 2 of Appendix B. The smoothing parameter is chosen as $\varepsilon = 10$. Figure 8 displays the graph $\alpha \mapsto \Lambda_\alpha$ of the 20 largest eigenvalues of the transition matrix $[\mathbb{P}]$. This graph shows that the first six eigenvalues are significant. Eliminating the first eigenvalue $\Lambda_1 = 1$ that correspond to the constant vector basis, the final dimension of the diffusion-maps basis is $m = 5$ and the vectors $\mathbf{g}^1, \dots, \mathbf{g}^m$ are such that $\mathbf{g}^\alpha = \Lambda_{\alpha+1} \boldsymbol{\psi}^{\alpha+1} \in \mathbb{R}^N$.
3. For each given value of n_{MC}, $\nu_{\mathrm{sim}} = n_{\mathrm{MC}} \times 600$ additional realizations $\mathbf{y}_{\mathrm{ar}}^\ell = (\mathbf{w}_{\mathrm{ar}}^\ell, q_{\mathrm{ar}}^\ell, \mathbf{b}_{\mathrm{ar}}^\ell)$ with $\ell = 1, \dots, \nu_{\mathrm{sim}}$ are generated (without using the SCM) by using Paragraphs 4 and 5 of Appendix B with the following values for the parameters: $f_0 = 150$, $\Delta r = 0.12352$, $M_0 = 1$, and $M = M_0 \times n_{\mathrm{MC}}$. For instance, for $N = 600$ points $(w_1^\ell, w_2^\ell, q^\ell)$ and $n_{\mathrm{MC}} = 250$, Fig. 9 displays the $\nu_{\mathrm{sim}} = 150{,}000$ additional realizations $(w_{\mathrm{ar},1}^\ell, w_{\mathrm{ar},2}^\ell, q_{\mathrm{ar}}^\ell)$ generated with the probabilistic learning algorithm.

Remark. The diffusion-maps basis depends on ε but is not very sensitive to ε. Value $\varepsilon = 10$ has been estimated by constructing the graph of the eigenvalues shows in Fig. 8 for several values of ε (the numerical cost of such an analysis is negligible). The value retained corresponds to the one for which the graph of the eigenvalues is not continuously decreasing but exhibits a clear discontinuity as shown in Fig. 8 (see [85]).
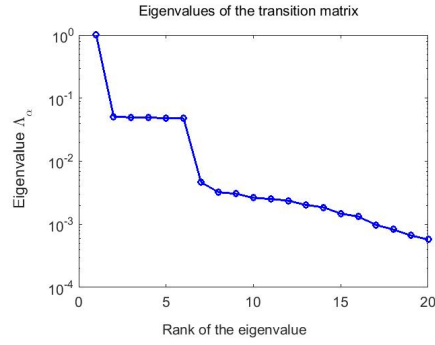
Fig. 8 For $N = 600$ (number of points in the dataset used as the input for the probabilistic learning algorithm and constructed by using the stochastic computational model), the figure displays the first twenty largest eigenvalues (in $\log_{10}$-scale) of the transition matrix $[\mathbb{P}] \in \mathbb{M}_N$, which allows for constructing the diffusion basis as explained in Paragraph 2 of Appendix B.
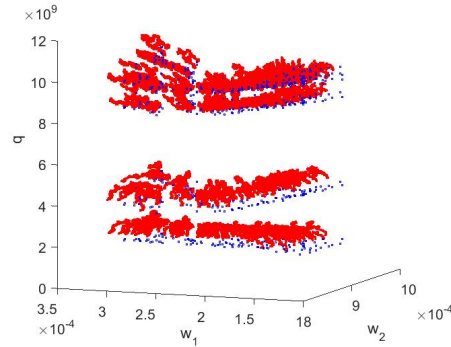
Fig. 9 Dataset of $N = 600$ points $(w_1^\ell, w_2^\ell, q^\ell)$ (blue symbols) and $\nu_{\mathrm{sim}} = n_{\mathrm{MC}} \times N = 250 \times 600 = 150{,}000$ additional points $(w_{\mathrm{ar},1}^\ell, w_{\mathrm{ar},2}^\ell, q_{\mathrm{ar}}^\ell)$ (red symbols) generated with the probabilistic learning algorithm.

Smoothing technique and solving the optimization problem. For a given value of n_{MC} yielding a value of ν_{sim}, the optimal solution computed with the probabilistic learning algorithm is denoted by $\mathbf{w}_{\mathrm{pla}}^{\mathrm{opt}}(n_{\mathrm{MC}})$ and is obtained by solving the approximate optimization problem,

$$\mathbf{w}_{\mathrm{pla}}^{\mathrm{opt}}(n_{\mathrm{MC}}) = \arg \min_{\substack{\mathbf{w} \in \mathcal{C}_g \\ \mathbf{c}(\mathbf{w}) < 0}} J(\mathbf{w}),$$

which is equivalent to

$$\mathbf{w}_{\mathrm{pla}}^{\mathrm{opt}}(n_{\mathrm{MC}}) = \min_{\substack{i=1,\dots,N_g \\ \mathbf{c}(\mathbf{w}_g^i) < 0}} J(\mathbf{w}_g^i).$$

For every point $\mathbf{w}_g^i$ in $\mathcal{C}_g$, $J(\mathbf{w}_g^i)$ and $\mathbf{c}(\mathbf{w}_g^i)$ are computed by using Appendix C in which $\mathbf{w}^0$ has to be chosen as $\mathbf{w}_g^i$.

Convergence analysis of the optimal solution $\mathbf{w}_{\mathrm{pla}}^{\mathrm{opt}}(n_{\mathrm{MC}})$ with respect to n_{MC}. For n_{MC} in $[50, 300]$, the optimal solution is rewritten as $\mathbf{w}_{\mathrm{pla}}^{\mathrm{opt}}(n_{\mathrm{MC}}) = (a_{\mathrm{pla}}^{\mathrm{opt}}(n_{\mathrm{MC}}), b_{\mathrm{pla}}^{\mathrm{opt}}(n_{\mathrm{MC}}))$. Figure 10 displays the graph of function $n_{\mathrm{MC}} \mapsto a_{\mathrm{pla}}^{\mathrm{opt}}(n_{\mathrm{MC}})$ for $n_{\mathrm{MC}} \in$

$[50\,,300]$ while $n_{\rm MC} \mapsto b_{\rm pla}^{\rm opt}(n_{\rm MC})$ is constant and equal to $8.50 \times 10^{-4}\,m$.

Optimal solution for $n_{\rm MC} = 250$. Figure 10 shows that the optimal solution is converged for $n_{\rm MC} \geq 250$. The converged optimal solution, which is denoted by $\mathbf{w}_{\rm pla}^{\rm opt} = (a_{\rm pla}^{\rm opt}, b_{\rm pla}^{\rm opt})$, is such that

$$a_{\rm pla}^{\rm opt} = 2.56 \times 10^{-4}\,m \quad , \quad b_{\rm pla}^{\rm opt} = 8.50 \times 10^{-4}\,m\,.$$

The corresponding value of the objective function is

$$J(\mathbf{w}_{\rm pla}^{\rm opt}) = 7.29 \times 10^9\,Pa\,.$$

The constraint c_1 defined by Eqs. (10) and (11) is saturated and is such that

$$c_1(\mathbf{w}_{\rm pla}^{\rm opt}) = 4.61 \times 10^9 - 5.00 \times 10^9 = -0.39 \times 10^9\,Pa\,,$$

while the constraint c_2 defined by Eqs. (10) and (12) is almost saturated and is such that

$$c_2(\mathbf{w}_{\rm pla}^{\rm opt}) = 1.67 \times 10^{10} - 1.80 \times 10^{10} = -0.13 \times 10^{10}\,Pa\,.$$

The comparison of this solution $\mathbf{w}_{\rm pla}^{\rm opt}$ computed with the probabilistic algorithm, with the reference optimal solution $\mathbf{w}_{\rm ref}^{\rm opt}$ is very good although a small bias induced by the statistical mean of the additional realizations appears, and lightly shifts the values of the objective and constraints functions.

8 Analysis of the CPU-time gain

In this section, we present the CPU-time gain obtained by using the probabilistic learning algorithm with respect to the usual approach. The computation have been performed with a Linux work station using parallel computing with 40 workers (cores). The parallelization is carried out with respect to the values of the design parameters (each worker is devoted to a given value of the design parameter). The CPU time is equal to the elapse time multiplied by the number of workers (40).
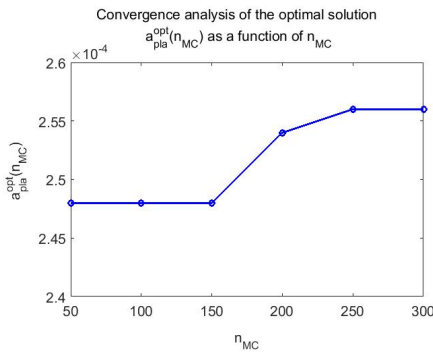
Fig. 10 Convergence analysis of the optimal solution: graph of function $n_{\rm MC} \mapsto a_{\rm pla}^{\rm opt}(n_{\rm MC})$ for $n_{\rm MC} \in [50\,,300]$.

– For the construction of the reference optimal solution using the SCM with $N_s' = 145$ corresponding to the convergence, the elapse time is 64 hours (CPU time 2,560 hours).
– For the construction of the optimal solution using the probabilistic learning algorithm with $N_{sd} = 5$, the elapse time is 2.6 hours (CPU time 104 hours).

It can be seen that the CPU-time gain factor is N_s'/n_{sd}, which is coherent. For the case analyzed the gain factor is about 25.

9 Conclusion

We have presented an efficient methodology for the design optimization of an implant in biological tissue at the mesoscale for which it is necessary to use a realistic stochastic model of the elasticity tensor field, because a separation of the scales cannot be obtained and consequently, the stress and the strain fields used for calculating the objective and the constraint functions strongly depend on the statistical fluctuations in the biological tissue. The numerical model used has been chosen sufficiently simple to be in capability to describe all data so that the results could be reproduced by anyone. To solve the problem of constrained nonconvex optimization in the presence of uncertainties corresponding to design optimization, an advanced algorithm of probabilistic learning has been used to reduce the computation time. The two ingredients of this algorithm have already been published in three papers, but we have preferred synthesized and summarized the algorithm in the appendices in order to confer a readability to the paper. The results and the gain obtained would seem open an interesting way for making the design optimization of implants in biological tissues using computers of intermediate power and certainly, to perform the design optimization with very high fidelity computational models of implants having a very complex 3D geometry by using powerful computers. The methodology presented could also help to solve design optimization problems for large-scale uncertain computational models in computational mechanics. Finally, it should be noted that the probabilistic learning algorithm proposed for solving the optimization problem is not intrusive. The dataset used by the probabilistic learning algorithm can be constructed with any commercial software that has the capability to analyze stochastic computational models and that can be considered as a black box. The optimization performed with the probabilistic learning algorithm is a post processing, independent of the black box, for which the input of the algorithm is the dataset constructed with the black box.

A Construction of the non-Gaussian random field [G₀] and its generator of realizations

The non-Gaussian random field $\{[\mathbf{G}_0(\mathbf{x})], \mathbf{x} \in \mathbb{R}^d\}$ is defined on $(\Theta, \mathcal{T}, \mathcal{P})$, is indexed by $\mathbb{R}^d$ with values in $\mathbb{M}_n^+(\mathbb{R})$, and is constructed by using a family of Gaussian random fields $\{\mathcal{U}_{jk}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ for $1 \le j \le k \le n$, which will be defined after. The spatial-correlation structure of non-Gaussian random field $[\mathbf{G}_0]$ is induced by the spatial-correlation structure of Gaussian random fields $\{\mathcal{U}_{jk}\}_{jk}$ and its statistical fluctuations will be controlled by a positive hyperparameter δ that is taken independent of $\mathbf{x}$ and such that

$$0 < \delta < \sqrt{(n+1)/(n+5)} < 1. \tag{28}$$

The construction of random field $\{[\mathbf{G}_0]$ is the following and gives a generator of independent realizations.

— For all $\mathbf{x}$ fixed in $\mathbb{R}^d$, the random matrix $[\mathbf{G}_0(\mathbf{x})]$ is written as

$$[\mathbf{G}_0(\mathbf{x})] = [\mathbf{L}(\mathbf{x})]^T [\mathbf{L}(\mathbf{x})], \tag{29}$$

in which $[\mathbf{L}(\mathbf{x})]$ is the upper $(n \times n)$ real triangular random matrix.
— For $1 \le j \le k \le n$, the random fields $\{[\mathbf{L}(\mathbf{x})]_{jk}, \mathbf{x} \in \mathbb{R}^d\}$ are independent.
— For $j < k$, the real-valued random field $\{[\mathbf{L}(\mathbf{x})]_{jk}, \mathbf{x} \in \mathbb{R}^d\}$ is defined by $[\mathbf{L}(\mathbf{x})]_{jk} = \sigma_n \mathcal{U}_{jk}(\mathbf{x})$ in which σ_n is such that $\sigma_n = \delta/\sqrt{n+1}$.
— For $j = k$, the positive-valued random field $\{[\mathbf{L}(\mathbf{x})]_{jj}, \mathbf{x} \in \mathbb{R}^d\}$ is defined by $[\mathbf{L}(\mathbf{x})]_{jj} = \sigma_n \sqrt{2\,h(\mathcal{U}_{jj}(\mathbf{x}), a_j)}$ in which $a_j = (n+1)/(2\delta^2) + (1-j)/2$. Function $u \mapsto h(u, \alpha)$ is such that $\Gamma_\alpha = h(U, \alpha)$ is a gamma random variable with parameter α when U is a normalized Gaussian random variable.

Concerning the construction of the family of Gaussian random fields, a more general formulation can be found in [83,86] than the one presented hereinafter and that corresponds to the stochastic model that has been experimentally identified. Such a family $\{\mathcal{U}_{jk}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}_{1 \le j \le k \le n}$ of random fields can be viewed as the Gaussian stochastic germs of the non-Gaussian random field $[\mathbf{G}_0]$ and is constructed as follows. Each random field $\mathcal{U}_{jk}$ of the family is an independent copy of a unique Gaussian, homogeneous, second-order, real-valued random field, $\mathcal{U} = \{\mathcal{U}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d\}$ that is defined on probability space $(\Theta, \mathcal{T}, \mathcal{P})$ and that is such that

$$E\{\mathcal{U}(\mathbf{x})\} = 0, \quad E\{\mathcal{U}(\mathbf{x})^2\} = 1. \tag{30}$$

Random field $\mathcal{U}$ is thus completely and uniquely defined by its autocorrelation function

$$\zeta \mapsto R_\mathcal{U}(\zeta) = E\{\mathcal{U}(\mathbf{x}+\zeta)\mathcal{U}(\mathbf{x})\}, \tag{31}$$

from $\mathbb{R}^d$ into $\mathbb{R}$, such that $R_\mathcal{U}(0) = 1$. Let $\zeta = (\zeta_1, \ldots, \zeta_d)$ be in $\mathbb{R}^d$. The spatial-correlation lengths $\lambda_1, \ldots, \lambda_d$ of random field $\mathcal{U}$ are defined, for $\alpha = 1, \ldots d$, by

$$\lambda_\alpha = \int_0^{+\infty} |R_\mathcal{U}(0, \ldots, \zeta_\alpha, \ldots, 0)|\, d\zeta_\alpha, \tag{32}$$

and are chosen as hyperparameters. The autocorrelation function is written as

$$R_\mathcal{U}(\zeta) = \rho_1(\zeta_1) \times \ldots \times \rho_d(\zeta_d), \tag{33}$$

$$\rho_\alpha(\zeta_\alpha) = \{4(\lambda_\alpha)^2/(\pi^2 \zeta_\alpha^2)\} \sin^2(\pi\zeta_\alpha/(2\lambda_\alpha)). \tag{34}$$

Consequently, random field $\mathcal{U}$ is mean-square continuous on $\mathbb{R}^d$ and its power spectral density function defined on $\mathbb{R}^d$ has a compact support that is written as,

$$[-\pi/\lambda_1, \pi/\lambda_1] \times \ldots \times [-\pi/\lambda_d, \pi/\lambda_d]. \tag{35}$$

It should be noted that any other autocorrelation function could be chosen in such a construction. Nevertheless, the great interest of this simple choice is briefly summarized as follows: (1) the spatial-correlation lengths are directly controlled for $\mathcal{U}$ and consequently, are controlled for $[\mathbf{G}_0]$ and $[\mathbf{K}]$, (2) the support of the spectral measure is compact that implies the regularity of the sample paths and in addition, allows for controlling the space sampling of the random field $[\mathbf{K}]$ for constructing the finite element discretization by using the Shannon theorem, (3) this model exhibit only d real parameters that is well adapted for the experimental identification that requires to solve a statistical inverse problem related to a boundary value problem. The details concerning the generation of realizations of random field $\mathcal{U}$ can be found in [82, 83]. One possible method is based on the usual numerical simulation of homogeneous Gaussian vector-valued random field constructed with the stochastic integral representation of homogeneous stochastic fields (see [69,79]).

B Construction of the diffusion-maps basis and generator of additional realizations

In this Appendix, we summarize the probabilistic learning algorithm [85,87] that we use. This appendix gives the algorithm for constructing the diffusion-maps basis and gives the algorithm for generating additional realizations by solving a nonlinear Itô stochastic differential equation.

Let $\{\mathbf{y}^\ell = (\mathbf{w}^\ell, q^\ell, \mathbf{b}^\ell), \ell = 1, \ldots, N\}$ be the set of the N data points in $\mathbb{R}^n = \mathbb{R}^{m_w} \times \mathbb{R} \times \mathbb{R}^{m_c}$ that have been computed using the SCM as explained in Section 5.3. Let $\mathbf{Y}$ be the random variable

$$\mathbf{Y} = (\mathbf{W}, Q, \mathbf{B}), \tag{36}$$

with values in $\mathbb{R}^n = \mathbb{R}^{m_w} \times \mathbb{R} \times \mathbb{R}^{m_c}$, with $\mathbf{W} = (W_1, \ldots, W_{m_w})$ and $\mathbf{B} = (B_1, \ldots, B_{m_c})$, introduced in Step 2 of Section 5.3, for which $\{\mathbf{y}^\ell, \ell = 1, \ldots, N\}$ constitutes N independent realizations. The objective of this Appendix is to summarize the probabilistic learning algorithm that allows for generating $\nu_{\text{sim}} \gg N$ additional realizations,

$$\mathbf{y}_{\text{ar}}^\ell = (\mathbf{w}_{\text{ar}}^\ell, q_{\text{ar}}^\ell, \mathbf{b}_{\text{ar}}^\ell), \quad \ell = 1, \ldots, \nu_{\text{sim}},$$

without performing additional function evaluations with the SCM.

1. Data normalization. Let $\mathbf{Y}$ be the $\mathbb{R}^n$-valued second-order random vector defined by Eq. (36) for which the N independent realizations are the N data points in $\mathbb{R}^n$, represented by the matrix $[y_d] = [\mathbf{y}^1 \ldots \mathbf{y}^N]$ in $\mathbb{M}_{n,N}$. Let $[\mathbf{Y}] = [\mathbf{Y}^1, \ldots, \mathbf{Y}^N]$ be the random matrix with values in $\mathbb{M}_{n,N}$, whose columns are N independent copies of random vector $\mathbf{Y}$. The normalization of random matrix $[\mathbf{Y}]$ is attained with random matrix $[\mathbf{H}] = [\mathbf{H}^1, \ldots, \mathbf{H}^N]$ with values in $\mathbb{M}_{\nu,N}$, whose columns are N independent copies of a random vector $\mathbf{H}$, with $\nu \le n$, obtained by using the principal component analysis that allows us to write $[\mathbf{Y}]$ as

$$[\mathbf{Y}] = [\underline{y}] + [\varphi]\,[\lambda]^{1/2}\,[\mathbf{H}], \tag{37}$$

in which $[\lambda]$ is the $(\nu \times \nu)$ diagonal matrix of the ν positive eigenvalues of the empirical estimate of the covariance matrix of $\mathbf{Y}$ (computed using $\mathbf{y}^1, \ldots, \mathbf{y}^N$), where $[\varphi]$ is the $(n \times \nu)$ matrix of the associated eigenvectors such $[\varphi]^T [\varphi] = [I_\nu]$, and where $[\underline{y}]$ is the matrix in $\mathbb{M}_{n,N}$ with identical columns, each one being equal to the empirical estimate $\underline{\mathbf{y}} \in \mathbb{R}^n$ of the mean value of random vector $\mathbf{y}$ (computed using $\mathbf{y}^1, \ldots, \mathbf{y}^N$). The sample $[\eta_d] = [\boldsymbol{\eta}^1 \ldots \boldsymbol{\eta}^N] \in \mathbb{M}_{\nu,N}$ of $[\mathbf{H}]$ (associated with the sample $[y_d]$ of $[\mathbf{Y}]$) is computed by

$$[\eta_d] = [\lambda]^{-1/2}[\varphi]^T ([y_d] - [\underline{y}]). $$

When n is small, ν can be chosen as n. If some eigenvalues are zero, they must be eliminated and then $\nu < n$. When n is high, a statistical reduction can be done as usual and therefore $\nu < n$ in such a case.

2. Diffusion-maps basis. For $1 < m+1 \leq N$, let $[g] = [\mathbf{g}^1 \ldots \mathbf{g}^m] \in \mathbb{M}_{N,m}$ be the "diffusion-maps basis" that is constructed by using the diffusion maps proposed in [15]. Let $[\mathbb{b}]$ be the positive-definite diagonal real matrix in $\mathbb{M}_N$ such that $[\mathbb{b}]_{ij} = \delta_{ij} \sum_{j'=1}^N [K]_{ij'}$ in which $[K]_{ij'} = \exp(-\frac{1}{4\varepsilon} \|\boldsymbol{\eta}^i - \boldsymbol{\eta}^{j'}\|^2)$, depending on a real smoothing parameter $\varepsilon > 0$. Let $[\mathbb{P}]$ be the transition matrix in $\mathbb{M}_N$ such that $[\mathbb{P}] = [\mathbb{b}]^{-1} [K]$. Let $\boldsymbol{\psi}^1, \ldots, \boldsymbol{\psi}^{m+1}$ be the right eigenvectors of $[\mathbb{P}]$, associated with the eigenvalues $1 = \Lambda_1 > \Lambda_2 > \ldots > \Lambda_{m+1}$, such that

$$[\mathbb{P}] \, \boldsymbol{\psi}^\alpha = \Lambda_\alpha \, \boldsymbol{\psi}^\alpha \, .$$

The normalization of the right eigenvectors of $[\mathbb{P}]$ is such that the matrix $[\psi] = [\boldsymbol{\psi}^1 \ldots \boldsymbol{\psi}^{m+1}]$ is chosen for that $[\psi]^T [\mathbb{b}] [\psi] = [I_{m+1}]$. The eigenvector $\boldsymbol{\psi}^1$ associated with the largest eigenvalue $\Lambda_1 = 1$ is a constant vector (all its components are equal). The "diffusion-maps basis" is defined by $[g] = [\mathbf{g}^1 \ldots \mathbf{g}^m] \in \mathbb{M}_{N,m}$ such that

$$\mathbf{g}^\alpha = \Lambda_{\alpha+1} \, \boldsymbol{\psi}^{\alpha+1} \in \mathbb{R}^N \, .$$

3. Reduced-order representation of random matrices $[\mathbf{H}]$ and $[\mathbf{Y}]$. The diffusion-maps vectors $\mathbf{g}^1, \ldots, \mathbf{g}^m \in \mathbb{R}^N$ span a subspace of $\mathbb{R}^N$ that characterizes the local geometry structure of the dataset concentrated in the neighborhood of a subset of $\mathbb{R}^N$. The reduced-order representation is obtained in projecting each column of the $\mathbb{M}_{N,\nu}$-valued random matrix $[\mathbf{H}]^T$ on the subspace of $\mathbb{R}^N$, spanned by $\{\mathbf{g}^1 \ldots \mathbf{g}^m\}$. Let $[\mathbf{Z}]$ be the random matrix with values in $\mathbb{M}_{\nu,m}$ such that $[\mathbf{H}] = [\mathbf{Z}] [g]^T$. As the matrix $[g]^T [g] \in \mathbb{M}_m$ is invertible, the least squares approximation of $\mathbf{Z}$ is written as $[\mathbf{Z}] = [\mathbf{H}] [a]$ in which

$$[a] = [g] \, ([g]^T [g])^{-1} \in \mathbb{M}_{N,m} \, ,$$

and the realization $[z_d] \in \mathbb{M}_{\nu,m}$ of $[\mathbf{Z}]$ is written as

$$[z_d] = [\eta_d] \, [a] \in \mathbb{M}_{\nu,m} \, .$$

The representation of random matrix $[\mathbf{Y}]$ as function of random matrix $[\mathbf{Z}]$ is then given by

$$[\mathbf{Y}] = [\underline{y}] + [\varphi] \, [\lambda]^{1/2} \, [\mathbf{Z}] \, [g]^T \, . \tag{38}$$

The dimension m of the reduced-order representation is estimated by analyzing the convergence of the representation with respect to m (see [85,87]). It should be noted that the reduced-order representation constructed with the diffusion-maps basis is not a reduction of the dimension of random vector $\mathbf{Y}$ with respect to the number of its components (reduction that has already been made with the PCA as explained in Paragraph 1 above), but is a reduction with respect to dimension N of data.

4. Generation of additional realizations $\{\mathbf{y}_{\text{ar}}^\ell, \ell = 1, \ldots, \nu_{\text{sim}}\}$ of random vector $\mathbf{Y}$. The generation of additional realizations $[z_{\text{ar}}^1], \ldots, [z_{\text{ar}}^{n_{\text{MC}}}]$ of random matrix $[\mathbf{Z}]$ is carried out by using an unusual MCMC method based on a reduced-order Itô stochastic differential equation (ISDE) that is constructed as the projection on the diffusion-maps basis of the ISDE related to a dissipative Hamiltonian dynamical system for which the invariant measure is the pdf of random matrix $[\mathbf{H}]$ constructed with the Gaussian kernel-density estimation method and $[\eta_d]$. This method preserves the concentration of the probability measure and avoids the scatter phenomenon. The constructed reduced-order ISDE is then used for generating n_{MC} additional realizations, $[z_{\text{ar}}^1], \ldots, [z_{\text{ar}}^{n_{\text{MC}}}]$ in $\mathbb{M}_{\nu,m}$, of random matrix $[\mathbf{Z}]$, and therefore, for deducing the n_{MC} additional realizations, $[\eta_{\text{ar}}^1], \ldots, [\eta_{\text{ar}}^{n_{\text{MC}}}]$ in $\mathbb{M}_{\nu,N}$ of random matrix $[\mathbf{H}]$, such that $[\eta_{\text{ar}}^\ell] = [z_{\text{ar}}^\ell] [g]^T$ for $\ell = 1, \ldots, n_{\text{MC}}$. Let $\{([\boldsymbol{\mathcal{Z}}(r)], [\boldsymbol{\mathcal{Y}}(r)]), r \in \mathbb{R}^+\}$ be the unique asymptotic (for $r \to +\infty$) stationary and ergodic diffusion stochastic process with values in $\mathbb{M}_{\nu,m} \times \mathbb{M}_{\nu,m}$, of the following reduced-order ISDE (stochastic nonlinear second-order dissipative Hamiltonian dynamical system), for $r > 0$,

$$d[\boldsymbol{\mathcal{Z}}(r)] = [\boldsymbol{\mathcal{Y}}(r)] \, dr \, ,$$

$$d[\boldsymbol{\mathcal{Y}}(r)] = [\mathcal{L}([\boldsymbol{\mathcal{Z}}(r)])] \, dr - \frac{1}{2} f_0 \, [\boldsymbol{\mathcal{Y}}(r)] \, dr + \sqrt{f_0} \, [d\boldsymbol{\mathcal{W}}(r)] \, ,$$

with the initial condition

$$[\boldsymbol{\mathcal{Z}}(0)] = [z_d] \quad , \quad [\boldsymbol{\mathcal{Y}}(0)] = [\boldsymbol{\mathcal{N}}] \, [a] \quad a.s.$$

(i) The random matrix $[\mathcal{L}([\boldsymbol{\mathcal{Z}}(r)])]$ with values in $\mathbb{M}_{\nu,m}$ is such that $[\mathcal{L}([\boldsymbol{\mathcal{Z}}(r)])] = [L([\boldsymbol{\mathcal{Z}}(r)] [g]^T)] [a]$. For all $[u] = [\mathbf{u}^1 \ldots \mathbf{u}^N]$ in $\mathbb{M}_{\nu,N}$ with $\mathbf{u}^\ell = (u_1^\ell, \ldots, u_\nu^\ell)$ in $\mathbb{R}^\nu$, the matrix $[L([u])]$ in $\mathbb{M}_{\nu,N}$ is defined, for all $k = 1, \ldots, \nu$ and for all $\ell = 1, \ldots, N$, by

$$[L([u])]_{k\ell} = \frac{1}{p(\mathbf{u}^\ell)} \left\{ \boldsymbol{\nabla}_{\mathbf{u}^\ell} \, p(\mathbf{u}^\ell) \right\}_k \, ,$$

$$p(\mathbf{u}^\ell) = \frac{1}{N} \sum_{j=1}^N \exp\{-\frac{1}{2\widehat{s}_\nu^2} \|\frac{\widehat{s}_\nu}{s_\nu} \boldsymbol{\eta}^j - \mathbf{u}^\ell\|^2\} \, ,$$

$$\boldsymbol{\nabla}_{\mathbf{u}^\ell} \, p(\mathbf{u}^\ell) = \frac{1}{\widehat{s}_\nu^2} \frac{1}{N} \sum_{j=1}^N (\frac{\widehat{s}_\nu}{s_\nu} \boldsymbol{\eta}^j - \mathbf{u}^\ell) \exp\{-\frac{1}{2\widehat{s}_\nu^2} \|\frac{\widehat{s}_\nu}{s_\nu} \boldsymbol{\eta}^j - \mathbf{u}^\ell\|^2\} \, ,$$

$$s_\nu = \left\{ \frac{4}{N(2+\nu)} \right\}^{1/(\nu+4)} \quad , \quad \widehat{s}_\nu = \frac{s_\nu}{\sqrt{s_\nu^2 + \frac{N-1}{N}}} \, .$$

(ii) $[d\boldsymbol{\mathcal{W}}(r)] = [d\mathbb{W}(r)] [a]$ where $[d\mathbb{W}(r)]$ is the $\mathbb{M}_{\nu,N}$-valued normalized Wiener stochastic process.
(iii) $[\boldsymbol{\mathcal{N}}]$ is the $\mathbb{M}_{\nu,N}$-valued normalized Gaussian random matrix.
(iv) The free parameter $f_0 > 0$ allows the dissipation term of the nonlinear second-order dynamical system (dissipative Hamiltonian system) to be controlled in order to kill the transient part induced by the initial conditions.
(v) We then have $[\mathbf{Z}] = \lim_{r \to +\infty} [\boldsymbol{\mathcal{Z}}(r)]$ in probability distribution, which allows for generating the additional realizations, $[z_{\text{ar}}^1], \ldots, [z_{\text{ar}}^{n_{\text{MC}}}]$, and then, generating the additional realizations $[y_{\text{ar}}^1], \ldots, [y_{\text{ar}}^{n_{\text{MC}}}]$ by using Eq. (38), which are reshaped in order to deduce the $\nu_{\text{sim}} = n_{\text{MC}} \times N$ additional realizations, $\{\mathbf{y}_{\text{ar}}^\ell, \ell = 1, \ldots, \nu_{\text{sim}}\}$.

5. Algorithm for solving the reduced-order ISDE. The algorithm for solving the reduced-order ISDE is detailed in [85] and is summarized hereinafter. The Störmer-Verlet scheme is used. Let $M = n_{\text{MC}} \times M_0$ be the positive integer in which n_{MC} and M_0 are integers. The reduced-order ISDE is solved on the finite interval $\mathcal{R} = [0, M \Delta r]$, in which Δr is the sampling step of the continuous index parameter r. The integration scheme is based on the use of the $M + 1$ sampling points $r_{\ell'}$ such that $r_{\ell'} = \ell' \Delta r$ for $\ell' = 0, \ldots, M$. The following notations are introduced: $[\boldsymbol{\mathcal{Z}}_{\ell'}] = [\boldsymbol{\mathcal{Z}}(r_{\ell'})]$, $[\boldsymbol{\mathcal{Y}}_{\ell'}] = [\boldsymbol{\mathcal{Y}}(r_{\ell'})]$, and $[\boldsymbol{\mathcal{W}}_{\ell'}] = [\boldsymbol{\mathcal{W}}(r_{\ell'})]$, for $\ell' = 0, \ldots, M$, with $[\boldsymbol{\mathcal{Z}}_0] = [z_d]$, $[\boldsymbol{\mathcal{Y}}_0] = [\boldsymbol{\mathcal{N}}] [a]$, and $[\boldsymbol{\mathcal{W}}_0] = [0_{\nu,m}]$. For $\ell' = 0, \ldots, M - 1$, let $[\Delta\boldsymbol{\mathcal{W}}_{\ell'+1}] = [\Delta\mathbb{W}_{\ell'+1}] [a]$ be the sequence of random matrices with values in $\mathbb{M}_{\nu,m}$, in which $[\Delta\mathbb{W}_{\ell'+1}] = [\mathbb{W}_{\ell'+1}] - [\mathbb{W}_{\ell'}]$. The increments $[\Delta\mathbb{W}_1], \ldots, [\Delta\mathbb{W}_M]$ are M independent random matrices with values in $\mathbb{M}_{\nu,N}$. For all $k = 1, \ldots, \nu$ and for all $j = 1, \ldots, N$, the real-valued random variables $\{[\Delta\mathbb{W}_{\ell'+1}]_{kj}\}_{kj}$ are independent, Gaussian, second-order, and centered random variables such that

$$E\{[\Delta\mathbb{W}_{\ell'+1}]_{kj} [\Delta\mathbb{W}_{\ell'+1}]_{k'j'}\} = \Delta r \, \delta_{kk'} \, \delta_{jj'} \, .$$

For $\ell' = 0, \ldots, M - 1$, the Störmer-Verlet scheme applied to the reduced-order ISDE yields

$$[\boldsymbol{\mathcal{Z}}_{\ell'+\frac{1}{2}}] = [\boldsymbol{\mathcal{Z}}_{\ell'}] + \frac{\Delta r}{2} [\boldsymbol{\mathcal{Y}}_{\ell'}] \, ,$$

$$[\boldsymbol{\mathcal{Y}}_{\ell'+1}] = \frac{1-b}{1+b}[\boldsymbol{\mathcal{Y}}_{\ell'}] + \frac{\Delta r}{1+b}[\boldsymbol{\mathcal{L}}_{\ell'+\frac{1}{2}}] + \frac{\sqrt{f_0}}{1+b}[\Delta\boldsymbol{\mathcal{W}}_{\ell'+1}],$$

$$[\boldsymbol{\mathcal{Z}}_{\ell'+1}] = [\boldsymbol{\mathcal{Z}}_{\ell'+\frac{1}{2}}] + \frac{\Delta r}{2}[\boldsymbol{\mathcal{Y}}_{\ell'+1}],$$

with the initial condition defined before, where $b = f_0\,\Delta r\,/4$, and where $[\boldsymbol{\mathcal{L}}_{\ell'+\frac{1}{2}}]$ is the $\mathbb{M}_{\nu,m}$-valued random variable such that

$$[\boldsymbol{\mathcal{L}}_{\ell'+\frac{1}{2}}] = [\mathcal{L}([\boldsymbol{\mathcal{Z}}_{\ell'+\frac{1}{2}}])] = [L([\boldsymbol{\mathcal{Z}}_{\ell'+\frac{1}{2}}]\,[g]^T)]\,[a]\,.$$

C Smoothing technique by using a nonparametric statistical estimation

In this Appendix, we summarize the smoothing algorithm [25, 88] that is used for estimating $J(\mathbf{w}^0)$ and $\mathbf{c}(\mathbf{w}^0)$ at any point $\mathbf{w}^0$ in $\mathcal{C}_\mathbf{w}$, by using only the additional realizations $\{\mathbf{y}_{ar}^\ell, \ell = 1,\ldots,\nu_{sim}\}$ that have been generated with the probabilistic learning algorithm summarized in Appendix B. The points $\mathbf{w}^0$ are those that are generated by the optimization algorithm. The smoothing algorithm is derived from a nonparametric statistical estimate of the probability density function $p_\mathbf{Y}$ of $\mathbf{Y}$. The construction of this algorithm is detailed in [25, 88] and is summarized hereinafter.

The steps of the algorithm for estimating the objective function $J(\mathbf{w}^0)$ and the constraints function $\mathbf{c}(\mathbf{w}^0)$ at a point $\mathbf{w}^0$ given in $\mathcal{C}_\mathbf{w}$ using only the ν_{sim} additional realizations are the following.

1. From the $\nu_{sim} = n_{MC} \times N$ additional realizations, $\{\mathbf{y}_{ar}^\ell, \ell = 1,\ldots,\nu_{sim}\}$, extracting:
   (a) the ν_{sim} realizations $(\mathbf{w}_{ar}^\ell, q_{ar}^\ell) \in \mathbb{R}^{m_w} \times \mathbb{R}$.
   (b) the ν_{sim} realizations $(\mathbf{w}_{ar}^\ell, \mathbf{b}_{ar}^\ell) \in \mathbb{R}^{m_w} \times \mathbb{R}^{m_c}$.

2. Computing
   (a) $\underline{q}_{ar} = \frac{1}{\nu_{sim}}\sum_{\ell=1}^{\nu_{sim}} q_{ar}^\ell$.
   (b) $\sigma_{q_{ar}}^2 = \frac{1}{\nu_{sim}}\sum_{\ell=1}^{\nu_{sim}} (q_{ar}^\ell - \underline{q}_{ar})^2$.
   (c) $\underline{b}_{ar,k} = \frac{1}{\nu_{sim}}\sum_{\ell=1}^{\nu_{sim}} b_{ar,k}^\ell$, $k = 1,\ldots,m_c$.
   (d) $\sigma_{b_{ar,k}}^2 = \frac{1}{\nu_{sim}}\sum_{\ell=1}^{\nu_{sim}} (b_{ar,k}^\ell - \underline{b}_{ar,k})^2$, $k = 1,\ldots,m_c$.

3. Normalizing the additional realizations: for $j = 1,\ldots,m_w$, let $\underline{w}_{ar,j}$ and $\sigma_{ar,j}$ be the empirical estimates of the mean value and of the standard deviation of random variable W_j constructed using the ν_{sim} independent additional realizations $\{w_{ar,j}^\ell, \ell = 1,\ldots,\nu_{sim}\}$. Let R be the real-valued random variable that denotes either Q, either B_1, or B_2. Let $\underline{r}_{ar}$ and σ_{ar} be the empirical estimates of the mean value and of the standard deviation of random variable R constructed with the ν_{sim} independent additional realizations $\{r_{ar}^\ell, \ell = 1,\ldots,\nu_{sim}\}$. We then introduce the normalized random variables $\widehat{W}_j$ for $j = 1,\ldots,m_w$ and $\widehat{R}$ defined by

$$\widehat{W}_j = (W_j - \underline{w}_{ar,j})/\sigma_{ar,j} \quad, \quad \widehat{R} = (R - \underline{r}_{ar})/\sigma_{ar},$$

for which the ν_{sim} independent normalized realizations are written, for $\ell = 1,\ldots,\nu_{sim}$, as

$$\widehat{w}_{ar,j}^\ell = (w_{ar,j}^\ell - \underline{w}_{ar,j})/\sigma_{ar,j} \quad, \quad \widehat{r}_{ar,j}^\ell = (r_{ar}^\ell - \underline{r}_{ar})/\sigma_{ar}\,.$$

The ν_{sim} independent normalized realizations are written, for $\ell = 1,\ldots,\nu_{sim}$, as
   (a) $(\widehat{\mathbf{w}}_{ar}^\ell, \widehat{q}_{ar}^\ell) \in \mathbb{R}^{m_w} \times \mathbb{R}$.
   (b) $(\widehat{\mathbf{w}}_{ar}^\ell, \widehat{b}_{ar,k}^\ell) \in \mathbb{R}^{m_w} \times \mathbb{R}$ for $k = 1,\ldots,m_c$.

4. Computing s by

$$s = \left\{\frac{4}{\nu_{sim}(2 + m_w + 1)}\right\}^{1/(4+m_w+1)}.$$

5. For $\mathbf{w}^0$ given in $\mathcal{C}_\mathbf{w}$, the normalization $\widehat{\mathbf{w}}^0$ of $\mathbf{w}^0$ is computed by

$$\widehat{w}_j^0 = (w_j^0 - \underline{w}_{ar,j})/\sigma_{ar,j} \quad, \quad j = 1,\ldots,m_w\,.$$

Introducing the joint pdf $p_{\widehat{\mathbf{W}},\widehat{R}}(\widehat{\mathbf{w}},\widehat{r})$ of $\widehat{\mathbf{W}} = (\widehat{W}_1,\ldots,\widehat{W}_{m_w})$ and $\widehat{R}$, and introducing the pdf $p_{\widehat{\mathbf{W}}}(\widehat{\mathbf{w}}) = \int_\mathbb{R} p_{\widehat{\mathbf{W}},\widehat{R}}(\widehat{\mathbf{w}},\widehat{r})\,d\widehat{r}$ of $\widehat{\mathbf{W}}$, the conditional mathematical expectation $E\{R\,|\,\mathbf{W} = \mathbf{w}^0\}$ is written as

$$E\{R\,|\,\mathbf{W} = \mathbf{w}^0\} \simeq \underline{r}_{ar} + \sigma_{ar}\frac{1}{p_{\widehat{\mathbf{W}}}(\widehat{\mathbf{w}}^0)}\int_\mathbb{R} \widehat{r}\,p_{\widehat{\mathbf{W}},\widehat{R}}(\widehat{\mathbf{w}}^0,\widehat{r})\,d\widehat{r}\,.$$

For each point $\mathbf{w}^0$ in $\mathcal{C}_\mathbf{w}$, $J(\mathbf{w}^0)$ and $\mathbf{c}(\mathbf{w}^0)$ are computed by using the following algorithm:
   (a) Computing $e^\ell(\widehat{\mathbf{w}}^0) = \exp\{-\frac{1}{2s^2}\|\widehat{\mathbf{w}}_{ar}^\ell - \widehat{\mathbf{w}}^0\|^2\}$ for $\ell = 1,\ldots,\nu_{sim}$.
   (b) Computing $\gamma(\widehat{\mathbf{w}}^0) = \sum_{\ell=1}^{\nu_{sim}} e^\ell(\widehat{\mathbf{w}}^0)$.
   (c) Computing the estimate of $J(\mathbf{w}^0)$ by

$$J(\mathbf{w}^0) \simeq \underline{q}_{ar} + \frac{\sigma_{q_{ar}}}{\gamma(\widehat{\mathbf{w}}^0)}\sum_{\ell=1}^{\nu_{sim}} \widehat{q}_{ar}^\ell\,e^\ell(\widehat{\mathbf{w}}^0)\,.$$

   (d) For $k = 1,\ldots,m_c$, computing the estimate of $c_k(\mathbf{w}^0)$ by

$$c_k(\mathbf{w}^0) \simeq \underline{b}_{ar,k} + \frac{\sigma_{b_{ar,k}}}{\gamma(\widehat{\mathbf{w}}^0)}\sum_{\ell=1}^{\nu_{sim}} \widehat{b}_{ar,k}^\ell\,e^\ell(\widehat{\mathbf{w}}^0)\,.$$

References

1. Agarwal H, Renaud JE, Preston EV, Padmanabhan D (2004) Uncertainty quantification using evidence theory in multidisciplinary design optimization. Reliability Engineering & System Safety 85(1-3):281294. doi:10.1016/j.ress.2004.03.017.
2. Ammar HH, Ngan P, Crout RJ, Mucino VH, Mukdadi OM (2011) Three-dimensional modeling and finite element analysis in treatment planning for orthodontic tooth movement. American Journal of Orthodontics and Dentofacial Orthopedics 139(1):e59e71. doi:10.1016/j.ajodo.2010.09.020.
3. Amsallem D, Zahr M, Choi Y, Farhat C (2015) Design optimization using hyper-reduced-order models. Structural and Multidisciplinary Optimization 51(4):919940. doi:10.1007/s00158-014-1183-y
4. Anitua E, Tapia R, Luzuriaga F, Orive G (2010) Influence of implant length, diameter, and geometry on stress distribution: A finite element analysis. International Journal of Periodontics & Restorative Dentistry 30(1):89-95.
5. Apley DW, Liu J, Chen W (2005) Understanding the effects of model uncertainty in robust design with computer experiments. Journal of Mechanical Design - ASME 128(4), 945-958 doi:10.1115/1.2204974.
6. Baggi L, Cappelloni I, Girolamo MD, Maceri F, Vairo G (2008) The influence of implant diameter and length on stress distribution of osseointegrated implants related to crestal bone geometry: A three-dimensional finite element analysis. The Journal of Prosthetic Dentistry 100(6):422431. doi:10.1016/S0022-3913(08)60259-0
7. Bengio Y (2000) Gradient-based optimization of hyperparameters. Neural Computation 12(8):1889-1900. doi:10.1162/089976600300015187
8. Bergstra J, Bengio Y (2012) Random search for hyper-parameter optimization. Journal of Machine Learning Research 13:281305.
9. Bianco RJ, Arnoux JP, Wagnac E, Mac-Thiong JM, Aubin CE (2017) Minimizing pedicle screw pullout risks: A detailed biomechanical analysis of screw design and placement. Clinical Spine Surgery 30(3):E226E232. doi:10.1097/BSD.0000000000000151.

10. Byrd RH, Chin GM, Neveitt W, Nocedal J (2011) On the use of stochastic hessian information in optimization methods for machine learning. SIAM Journal of Optimization 21(3):977-995. doi:10.1137/10079923X.

11. Calandra R, Seyfarth A, Peters J, Deisenroth MP (2016) Bayesian optimization for learning gaits under uncertainty. Annals of Mathematics and Artificial Intelligence 76(1):523. doi:10.1007/s10472-015-9463-9.

12. Capiez-Lernout E, Soize C (2008) Robust design optimization in computational mechanics. Journal of Applied Mechanics - Transactions of the ASME 75(2):1-11. doi:10.1115/1.2775493.

13. Chen W, Yin X, Lee S, Liu WK (2010) A multiscale design methodology for hierarchical systems with random field uncertainty. Journal of Mechanical Design - ASME 132(4):041006. doi:10.1115/1.4001210.

14. Coello C, Carlos A (2006) Evolutionary multi-objective optimization: A historical view of the field. IEEE Computational Intelligence Magazine 1(1):28-36. doi:10.1109/MCI.2006.1597059.

15. Coifman RR, Lafon S, Lee AB, Maggioni M, Nadler B, Warner F, Zucker SW (2005) Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. PNAS 102(1):7426-7431.

16. Desceliers C, Soize C, Zarroug M (2013) Computational strategy for the crash design analysis using an uncertain computational mechanical model. Computational Mechanics 52(2):453462. doi:10.1007/s00466-012-0822-7.

17. Ding X, Zhu XH, Liao SH,Zhang XH (2009) Implantbone interface stress distribution in immediately loaded implants of different diameters: A three-dimensional finite element analysis. Journal of Prosthodontics 18(5):393402. doi:10.1111/j.1532-849X.2009.00453.x.

18. Dirrenberger J, Forest S, Jeulin D (2012) Towards gigantic RVE sizes for 3D stochastic fibrous networks. International Journal of Solids and Structures 51(2):359-376. doi:10.1016/j.ijsolstr.2013.10.011.

19. Dorogoy A, Rittel D, Shemtov-Yona K, Korabi R (2017) Modeling dental implant insertion. Journal of the Mechanical Behavior of Biomedical Materials 68:4250. doi:10.1016/j.jmbbm.2017.01.021.

20. Drugan WJ, Willis JR (1996) A micromechanics-based nonlocal constitutive equation and estimates of representative volume element size for elastic composites. Journal of the Mechanics and Physics of Solids 44(4):497-524. doi:10.1016/0022-5096(96)00007-5

21. Du X, Chen W (2004) Sequential optimization and reliability assessment method for efficient probabilistic design. ASME Journal of Mechanical Design 126(2):225-233. doi:10.1115/1.1649968.

22. Dubourg V, Sudret B, Bourinet JM (2011) Reliability-based design optimization using kriging surrogates and subset simulation. Structural and Multidisciplinary Optimization 44(5):673690. doi:10.1007/s00158-011-0653-8.

23. Eldred M (2011) Design under uncertainty employing stochastic expansion methods. International Journal for Uncertainty Quantification 1(2):119-146. doi:10.1615/Int.J.UncertaintyQuantification.v1.i2.20.

24. Farhat C, Chapman T, Avery P (2015) Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of nonlinear finite element dynamic models. International Journal for Numerical Methods in Engineering 102(5):10771110. doi:10.1002/nme.4820.

25. Ghanem R, Soize C (2017) Probabilistic nonconvex constrained optimization with fixed number of function evaluations. International Journal for Numerical Methods in Engineering, published online 15 September 2017:1-25. doi: 10.1002/nme.5632.

26. Ghanem R, Higdon D, Owhadi H (Eds.) (2017) Handbook of Uncertainty Quantification. Springer, New York. doi:10.1007/978-3-319-11259-6.

27. Goldberg DE (1989) Genetic Algorithms in Search. Optimization & Machine Learning, Addison-Wesley. ISBN:0201157675.

28. Goury O, Amsallem D, Bordas SPA, Liu WK, Kerfriden P (2016) Automatised selection of load paths to construct reduced-order models in computational damage micromechanics: from dissipation-driven random selection to Bayesian optimization. Computational Mechanics 58(2):213234. doi:10.1007/s00466-016-1290-2.

29. Gu X, Renaud JE, Batill SM, Brach RM, Budhiraja AS (2000) Worst case propagated uncertainty of multidisciplinary systems in robust design optimization. Structural and Multidisciplinary Optimization 20(3):190213. doi:10.1007/s001580050148.

30. Guilleminot J, Noshadravan A, Soize C, Ghanem R (2011) A probabilistic model for bounded elasticity tensor random fields with application to polycrystalline microstructures. Computer Methods in Applied Mechanics and Engineering 200(17-20):1637-1648. doi:10.1016/j.cma.2011.01.016.

31. Guilleminot J, Soize C (2013) Stochastic model and generator for random fields with symmetry properties: application to the mesoscopic modeling of elastic random media. Multiscale Modeling and Simulation (A SIAM Interdisciplinary Journal) 11(3):840-870. doi:10.1137/120898346.

32. Haiat G, Wang HL, Brunski JB (2014) Effects of biomechanical properties of the bone-implant interface on dental implant stability: from in silico approaches to the patient's mouth. Annual Review of Biomedical Engineering 11(16):187-213. doi:10.1146/annurev-bioeng-071813-104854.

33. Homem-de-Mello T, Bayraksan G (2014) Monte Carlo sampling-based methods for stochastic optimization. Surveys in Operations Research and Management Science 19(1):56-85. doi:10.1016/j.sorms.2014.05.001.

34. Hu N (2016) Advances in Multiscale Methods with Applications in Optimization, Uncertainty Quantification and Biomechanics. Columbia University, ProQuest Dissertations Publishing, 10151391.

35. Huang HZ, Zhang X (2009) Design optimization with discrete and continuous variables of aleatory and epistemic uncertainties. Journal of Mechanical Design 131(3):031006 doi:10.1115/1.3066712

36. Ichim PI, Hu X, Bazen JJ, Yi W (2015) Design optimization of a radial functionally graded dental implant. Journal of Biomedical Materials Research 104(1):5866. doi:10.1002/jbm.b.33345.

37. Jasmine MIF, Yezdani AA, Tajir F, Venu RM (2012) Analysis of stress in bone and microimplants during en-masse retraction of maxillary and mandibular anterior teeth with different insertion angulations: A 3-dimensional finite element analysis study. American Journal of Orthodontics and Dentofacial Orthopedics 141(1):7180. doi:10.1016/j.ajodo.2011.06.031.

38. Jeulin D (2000) Random texture models for material structures. Statistics and Computing 10(2):121132. doi:10.1023/A:1008942325749

39. Jeulin D (2012) Morphology and effective properties of multiscale random sets: A review. Comptes Rendus Mcanique 340(4-5):219-229. doi:10.1016/j.crme.2012.02.004

40. Jeulin D, Ostoja-Starzewski M (Eds.) (2001) Mechanics of Random and Multiscale Microstructures. Springer, Vienna. doi:10.1007/978-3-7091-2780-3

41. Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. Journal of Global Optimization 13(4):455-492. doi:10.1023/A:1008306431147.

42. Kachanov M, Sevostianov I, Shafiro B (2001) Explicit cross-property correlations for porous materials with anisotropic microstructures. Journal of the Mechanics and Physics of Solids 49(1):1-25. doi:10.1016/S0022-5096(00)00033-8

43. Khanoki SA, Pasini D (2012) Multiscale design and multiobjective optimization of orthopedic hip implants with functionally graded cellular material. Journal of Biomechanical Engineering 134(3):031004. doi:10.1115/1.4006115.

44. Kitaruma E, Stegaroiu R, Nomura S, Miyakawa O (2004) Biomechanical aspects of marginal bone resorption around osseointegrated implants: considerations based on a three-dimensional finite

element analysis. Clinical Oral Implants Research 15(4):401412. doi:10.1111/j.1600-0501.2004.01022.x.

45. Kleijnen JPC, van Beers W, van Nieuwenhuyse I (2010) Constrained optimization in expensive simulation: Novel approach. European Journal of Operational Research 202(1):164-174. doi:10.1016/j.ejor.2009.05.002.

46. Konak A, Coit DW, Smith AE (2006) Multi-objective optimization using genetic algorithms: A tutorial. Reliability Engineering and System Safety 91(9):992-1007. doi:10.1016/j.ress.2005.11.018.

47. Kröner E (1977) Bounds for effective elastic moduli of disordered materials. Journal of the Mechanics and Physics of Solids 25(2):137-155. doi:10.1016/0022-5096(77)90009-6

48. Lecun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proceedings of the IEEE 86(11):2278-2324. doi:10.1109/5.726791.

49. Lee HK, Bayome M, Ahn CS, Kim SH, Kim KB, Mo SS, Kook YA (2014) Stress distribution and displacement by different bone-borne palatal expanders with micro-implants: a three-dimensional finite-element analysis European Journal of Orthodontics 36(5):531-540. doi:10.1093/ejo/cjs063.

50. Lewis MB, Klineberg I (2011) Prosthodontic considerations designed to optimize outcomes for single-tooth implants. A review of the literature. Australian Dental Journal 56(2):181192. doi:10.1111/j.1834-7819.2011.01322.x.

51. Li M, Azarm S (2008) Multiobjective collaborative robust optimization with interval uncertainty and interdisciplinary uncertainty propagation. Journal of Mechanical Design - ASME 130(8):081402. doi:10.1115/1.2936898.

52. Li Y, Stroberg W, Lee TR, Kim HS, Man H, Ho D, Decuzzi P, Liu KM (2014) Multiscale modeling and uncertainty quantification in nanoparticle-mediated drug/gene delivery. Computational Mechanics 53(3):511537. doi:10.1007/s00466-013-0953-5.

53. Lin D, Li Q, Li W, Zhou S, Swain MV (2009) Design optimization of functionally graded dental implant for bone remodeling. Composites: Part B 40(7):668675. doi:10.1016/j.compositesb.2009.04.015.

54. Lin PT, Lin SP (2016) An effective approach to solve design optimization problems with arbitrarily distributed uncertainties in the original design space using ensemble of Gaussian reliability analyses. Journal of Mechanical Design - ASME 138(7):071403. doi:10.1115/1.4033548.

55. Liu Y, Greene MS, Chen W, Dikin DA, Liu WK (2013) Computational microstructure characterization and reconstruction for stochastic multiscale material design. Computer-Aided Design 45(1):6576. doi:10.1016/j.cad.2012.03.007.

56. Lu S, Li T, Zhang Y, Lu C, Sun Y, Zhang J, Xu D (2013) Biomechanical optimization of the diameter of distraction screw in distraction implant by three-dimensional finite element analysis. Computers in Biology and Medicine 43(11):19491954. doi:10.1016/j.compbiomed.2013.08.019.

57. Ma J, Wriggers P, Gao W, Chen JJ, Sahraee S (2011) Reliability-based optimization of trusses with random parameters under dynamic loads Computational Mechanics 47(6):627-640. doi:10.1007/s00466-010-0561-6.

58. Michel A, Bosc R, Mathieu V, Hernigou P, Haiat, G (2015) In vitro evaluation of the acetabular cup primary stability by impact analysis. Journal of Biomechanical Engineering 137(3):031011. doi:10.1115/1.4029505.

59. Milton GW (2002) The theory of Composites. Cambridge University Press, Cambridge.

60. Miyamoto I, Tsuboi Y, Wada E, Suwa H, Iizuka T (2005) Influence of cortical bone thickness and implant length on implant stability at the time of surgeryclinical, prospective, biomechanical, and imaging study. Bone 37(6):776780. doi:10.1016/j.bone.2005.06.019.

61. Naili S, Nguyen VH, Vu MB, Desceliers C, Soize C (2015) Modeling of transient wave propagation in a heterogeneous solid layer coupled with fluid: application to long bones. Journal of the Acoustical Society of America 137(2):668-678. doi:10.1121/1.4906838.

62. Nemat-Nasser S, Hori M (1999) Micromechanics: overall properties of heterogeneous materials. Second revised edition. Elsevier, Amsterdam.

63. Nguyen MT, Desceliers C, Soize C, Allain JM, Gharbi H (2015) Multiscale identification of the random elasticity field at mesoscale of a heterogeneous microstructure using multiscale experimental observations. International Journal for Multiscale Computational Engineering 13(4): 281-295. doi:10.1615/IntJMultCompEng.2015011435.

64. Nguyen MT, Allain JM, Gharbi H, Desceliers C, Soize C (2016) Experimental multiscale measurements for the mechanical identification of a cortical bone by digital image correlation. Journal of the Mechanical Behavior of Biomedical Materials 63:125-133. doi:10.1016/j.jmbbm.2016.06.011.

65. Nicolella DP, Thackera BH, Katoozian H, Davy DT (2006) The effect of three-dimensional shape optimization on the probabilistic response of a cemented femoral hip prosthesis. Journal of Biomechanics 39(7):12651278. doi:10.1016/j.jbiomech.2005.03.010

66. Ostoja-Starzewski M (1998) Random field models of heterogeneous materials. International Journal of Solids and Structures 35(19):2429-2455. doi:10.1016/S0020-7683(97)00144-3

67. Petrie CS, Williams JL (2007) Probabilistic analysis of peri-implant strain predictions as influenced by uncertainties in bone properties and occlusal forces. Clinical Oral Implants Research 18(5):611619. doi:10.1111/j.1600-0501.2007.01384.x.

68. Pishvaee MS, Rabbani M, Torabi SA (2011) A robust optimization approach to closed-loop supply chain network design under uncertainty. Applied Mathematical Modelling 35(2):637649. doi:10.1016/j.apm.2010.07.013.

69. Poirion F, Soize C (1995) Numerical methods and mathematical aspects for simulation of homogeneous and non homogeneous Gaussian vector fields. pp. 17-53, in Probabilistic Methods in Applied Physics, edited by P. Krée and W. Wedig, Springer-Verlag, Berlin. doi:10.1007/3-540-60214-3-50.

70. Rao S (1984) Multiobjective optimization in structural design with uncertain parameters and stochastic processes. AIAA Journal 22(11):1670-1678. doi:10.2514/3.8834.

71. Ridwan-Pramana A, Marcián P, Borák L, Narra N, Forouzanfar T, Wolff J (2016) Structural and mechanical implications of PMMA implant shape and interface geometry in cranioplasty A finite element study. Journal of Cranio-Maxillofacial Surgery 44(1):34-44. doi:10.1016/j.jcms.2015.10.014.

72. Ritto TG, Soize C, Sampaio R (2010) Robust optimization of the rate of penetration of a drill-string using a stochastic nonlinear dynamical model. Computational Mechanics 45(5):415-427 (2010). doi:10.1007/s00466-009-0462-8.

73. Rossi L, Linares JM, Chaves-Jacob J, Mailhé J, Sprauel JM (2014) Design optimization using Statistical Confidence Boundaries of response surfaces: Application to robust design of a biomedical implant. CIRP Annals - Manufacturing Technology 63(1):141144. doi:10.1016/j.cirp.2014.03.088.

74. Rungsiyakull C, Li QL, Sun G, Li W, Swain MV (2010) Surface morphology optimization for osseointegration of coated implants. Biomaterials 31(27):71967204. doi:10.1016/j.biomaterials.2010.05.077.

75. Queipo NV, Haftka RT, Shyy W, Goel T, Vaidyanathan R, Tucker PK (2005) Surrogate-based analysis and optimization. Progress in Aerospace Science 41(1):1-28. doi:10.1016/j.paerosci.2005.02.001

76. Sab K (1992) On the homogenization and the simulation of random materials. European Journal of Mechanics, A/Solids 11(5):585-607.

77. Sab K, Nedjar B (2005) Periodization of random media and representative volume element size for linear composites. Comptes Rendus Mécanique 333(2):187-195. doi:10.1016/j.crme.2004.10.003

78. Santiago Junior JF, Pellizzer EP, Verri FR, Perri de Carvalho PS (2013) Stress analysis in bone tissue around single implants

with different diameters and veneering materials: A 3-D finite element study. Materials Science and Engineering: C 33(8):47004714. doi:10.1016/j.msec.2013.07.027.

79. Shinozuka M (1971) Simulations of multivariate and multidimensional random processes. Journal of Acoustical Society of America 39(1):357-367. doi:10.1121/1.1912338.

80. Simpson TW, Mauery TM, Korte JJ, Mistree F (2001) Kriging models for global approximation in simulation-based multidisciplinary design optimization. AIAA Journal 39(12):2233-2241. doi:10.2514/2.1234.

81. Snoek J, Larochelle H, Adams RP (2012) Practical Bayesian optimization of machine learning algorithms. Advances in Neural Information Processing Systems 25:2960-2968.

82. Soize C (2006) Non Gaussian positive-definite matrix-valued random fields for elliptic stochastic partial differential operators. Computer Methods in Applied Mechanics and Engineering 195(1-3):26-64. doi:10.1016/j.cma.2004.12.014.

83. Soize C (2008) Tensor-valued random fields for mesoscale stochastic model of anisotropic elastic microstructure and probabilistic analysis of representative volume element size. Probabilistic Engineering Mechanics 23(2-3):307-323. doi:10.1016/j.probengmech.2007.12.019.

84. Soize C (2010) Identification of high-dimension polynomial chaos expansions with random coefficients for non-Gaussian tensor-valued random fields using partial and limited experimental data. Computer Methods in Applied Mechanics and Engineering 199(33-36):2150-2164. doi:10.1016/j.cma.2010.03.013.

85. Soize C, Ghanem R (2016) Data-driven probability concentration and sampling on manifold. Journal of Computational Physics 321:242-258. doi:10.1016/j.jcp.2016.05.044.

86. Soize C (2017) Uncertainty Quantification. An Accelerated Course with Advanced Applications in Computational Engineering. Springer, New York. doi: 10.1007/978-3-319-54339-0.

87. Soize C, Ghanem R (2017) Polynomial chaos representation of databases on manifolds. Journal of Computational Physics 335:201-221. doi:10.1016/j.jcp.2017.01.031.

88. Soize C, Ghanem R (2017) Probabilistic learning on manifold for optimization under uncertainties, pp. 1-15 in The Proceedings of UNCECOMP 2017, 2nd ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering, M. Papadrakakis, V. Papadopoulos, G. Stephanou (eds.), Rhodes Islands, Greece, 15-17 June 2017.

89. Spall JC (2003) Introduction to Stochastic Search and Optimization. John Wiley and Sons, Hoboken, New Jersey.

90. Sui Y, Gotovos A, Burdick JW, Krause A (2015) Safe Exploration for Optimization with Gaussian Processes. Proceedings of the 32nd International Conference on Machine Learning, Lille, France, JMLR: W&CP Volume 37.

91. Suquet P (Ed.) (2014) Continuum Micromechanics. Springer-Verlag, Wien.

92. Taylor M, Prendergast PJ (2015) Four decades of finite element analysis of orthopaedic devices: Where are we now and what are the opportunities? Journal of Biomechanics 48(5):767778. doi:10.1016/j.jbiomech.2014.12.019.

93. To AC, Liu WK, Olson GB, Belytschko T, Chen W, Shephard MS, Chung YW, Ghanem R, Voorhees PM, Seidman DN, Wolverton C, Chen JS, Moran B, Freeman AJ, Tian R, Luo X, Lautenschlager E, Challoner AD (2008) Materials integrity in microsystems: a framework for a petascale predictive-science-based multiscale modeling and simulation system. Computational Mechanics 42(4):485510. doi:10.1007/s00466-008-0267-1.

94. Torquato S (2002) Random Heterogeneous Materials, Microstructure and Macroscopic Properties. Springer-Verlag, New York.

95. Vayron R, Nguyen VH, Bosc R, Naili S, Haiat G (2015) Finite element simulation of ultrasonic wave propagation in a dental implant for biomechanical stability assessment. Biomechanics and Modeling in Mechanobiology 14(5):1021-1032. doi:10.1007/s10237-015-0651-7.

96. Wang H, Li E, Li GY (2011) Probability-based least square support vector regression metamodeling technique for crashworthiness optimization problems. Computational Mechanics 47(3):251263. doi:10.1007/s00466-010-0532-y.

97. Wang Z, Zoghi M, Hutter F, Matheson D, de Freitas N (2016) Bayesian optimization in a billion dimensions via random embeddings. Journal of Artificial Intelligence Research 55(1):361-387.

98. Xie J, Frazier PI, Chick SE (2016) Bayesian optimization via simulation with pairwise sampling and correlated pair beliefs. Operations Research 64(2):542-559. doi:10.1287/opre.2016.1480.

99. Youn BD, Choi KK, Park YH (2003) Hybrid analysis method for reliability-based design optimization. Journal of Mechanical Design - ASME 125(2):221-232. doi:10.1115/1.1561042.

100. Yao W, Chen X, Luo W, vanTooren M, Guo J (2011) Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles. Progress in Aerospace Sciences 47(6):450-479. doi:10.1016/j.paerosci.2011.05.001.

101. Yin X, Lee S, Chen W, Liu WK (2009) Efficient random field uncertainty propagation in design using multiscale analysis. Journal of Mechanical Design - ASME 131(2):021006. doi:10.1115/1.3042159.

102. Zelinka I, Snasel V, Abraham A (Eds.) (2013) Handbook of Optimization. From Classical to Modern Approach. Springer, New York.

103. Zienkiewicz OC, Taylor RL (2005) The Finite Element Method For Solid and Structural Mechanics. Elsevier, Butterworth-Heinemann, Sixth edition, Amsterdam