HAL
open science

Shape optimization for quadratic functionals and states with random right-hand sides

Marc Dambrine, Charles Dapogny, Helmut Harbrecht

To cite this version:

Marc Dambrine, Charles Dapogny, Helmut Harbrecht. Shape optimization for quadratic functionals and states with random right-hand sides. SIAM Journal on Control and Optimization, 2015, 53 (5), pp.2725-3317. 10.1137/15M1017041 . hal-01143319

HAL Id: hal-01143319

https://hal.science/hal-01143319

Submitted on 28 May 2015

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

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

SHAPE OPTIMIZATION FOR QUADRATIC FUNCTIONALS AND STATES WITH RANDOM RIGHT-HAND SIDES

M. DAMBRINE[1], C. DAPOGNY[2], H. HARBRECHT[3].

[1] *Departement of Mathematics, Université de Pau et des Pays de l'Adour, Avenue de l'Université, BP 1155, 64013 Pau Cedex, France.*
[2] *Laboratoire Jean Kuntzmann, CNRS, Université Joseph Fourier, Grenoble INP, Université Pierre Mendès France, BP 53, 38041 Grenoble Cedex 9, France.*
[3] *Departement Mathematik und Informatik, Universität Basel, Spiegelgasse 1, 4051 Basel, Switzerland.*

Abstract. In this work, we investigate a particular class of shape optimization problems under uncertainties on the input parameters. More precisely, we are interested in the minimization of the expectation of a quadratic objective in a situation where the state function depends linearly on a random input parameter. This framework covers important objectives such as tracking-type functionals for elliptic second order partial differential equations and the compliance in linear elasticity. We show that the robust objective and its gradient are completely and explicitly determined by low-order moments of the random input. We then derive a cheap, deterministic algorithm to minimize this objective and present model cases in structural optimization.

Contents

1. **Introduction** 1
2. **Formal presentation of the main idea** 3
3. **Shape optimization setting** 5
   3.1. Differentiation with respect to the domain: Hadamard's method 5
   3.2. Tensor products and Hilbert spaces 5
   3.3. First and second moments analysis 7
4. **Applications in concrete situations** 8
   4.1. Quadratic shape functionals in the context of the Poisson equation 8
   4.2. Quadratic functionals in the context of linear elasticity 10
5. **Numerical realization** 11
   5.1. Computing second moments 12
   5.2. Numerical calculation of a low-rank approximation of $\mathrm{Cor}(f)$ 12
   5.3. Low-rank approximation of the shape functional and its gradient 13
6. **Numerical examples** 14
   6.1. Presentation of the numerical algorithm 15
   6.2. Comparison between correlated and uncorrelated loads 15
   6.3. An example with a more complex correlation function 17
References 18

1. Introduction

Over the last decades, shape optimization has been developed as an efficient method for designing devices which are optimized with respect to a given purpose. Many practical problems in engineering lead to boundary value problems for an unknown function which needs to be computed to obtain a real quantity of interest. For example, in structural mechanics, the equations of linear elasticity are usually considered and

solved to compute e.g. the leading mode of a structure or its compliance. Shape optimization is then applied to optimize the workpiece under consideration with respect to this objective functional. We refer the reader to [1, 17, 20, 23, 29] and the references therein for an overview on the topic of shape optimization, which falls into the general setting of optimal control of partial differential equations.

Usually, the input parameters of the model, like the applied loads, the material's properties (typically the value of the Young modulus or Poisson ratio) or the geometry of the involved shapes itself are assumed to be perfectly known. However, convenient for the analysis of shape optimization problems, this assumption is unrealistic with respect to applications. In practice, a manufactured device achieves its nominal geometry only up to a tolerance, the material parameters never match the requirements perfectly and applied forces can only be estimated. Therefore, shape optimization under uncertainty is of great practical importance but started only recently to be investigated, see e.g. [2, 7, 8, 9, 12, 19, 26] for related results.

Two approaches are available in the context of optimization under uncertainty, depending on the actual knowledge of the uncertain parameters. On the one hand, if no a priori information is available short of an upper bound on their magnitude, one usually considers a worst-case approach. On the other hand, if some statistical information on the distribution of the unknown parameters is given, one can study the statistics of the objective, which depends on the random parameters through the state equation. Notice that in this case the state function is a random field and so the objective itself becomes random – the value of the objective depends on the design variables and on the random variable.

One is usually primarily interested in stochastic quantities of the objective such as its expectation. When this crude average is not sufficient, one may consider a weighted combination of the expectation and the standard deviation in order to limit the dispersion of the objective values around its expectation. Finally, one sometimes also considers the probability that the objective exceeds a given threshold. This last objective usually stands for constraints.

In the present article, we address the following problem: *given a partial statistical description of the random loading, design an efficient algorithm to minimize the expectation of the objective.*

We restrict ourselves to a special class of problems, namely those involving a quadratic shape functional for the state function which is defined by a state equation with a random right-hand side. This in particular means that the random state depends linearly on the random input parameter. Our theory covers important shape functionals like the Dirichlet energy and quadratic tracking-type functionals in the context of the Laplace operator. Besides, the compliance functional in linear elasticity belongs to the important members of the class of functionals under consideration.

Our main message is the following: *for objectives of the class under consideration, whose expectation is to be minimized, all quantities for performing a gradient-based shape optimization algorithm can be expressed deterministically, i.e., without any approximation. We only need access to the random parameter's first and second moment.* The main object is the two-point correlation function $\mathrm{Cor}(u)$ of the state function u. It is the solution of a tensor-product type boundary value problem with the random right-hand side's two-point correlation as right hand side. As a consequence, both, the expectation of shape functional, as well as the related shape gradient can explicitly be determined and efficiently be computed just from the knowledge of the random right hand side's expectation and two-point correlation function. This fact is of tremendous importance for applications: it is completely unrealistic to have access to the law of the random loadings, whereas the knowledge of its expectation and of its two-point correlation function seem to be much more reasonable assumptions. We therefore end up with a fully deterministic, efficient algorithm of similar cost as for classical shape optimization when no uncertainties are taken into account.

This article is organized as follows. First, we present in Section 2 the leading idea to reduce the stochastic problem to a deterministic one, relying on a very simplified model in finite dimension. We introduce the tensor formulation that is the keystone of the subsequent calculations. Then, in Section 3, we present the shape calculus which we shall use and adapt the idea of Section 2 to this more complex setting. In particular, we recall definitions and properties of tensor products on Hilbert spaces. We then apply in Section 4 the obtained method to three significant examples in the context of the Laplace operator and the equations of linear elasticity. Finally, we explain in Section 5 how to design efficient numerical methods to solve the corresponding optimization problem. The main point is to remark that the numerical resolution of the high-dimensional boundary value problem which defines $\mathrm{Cor}(u)$ can be avoided if desired. An appropriate low-rank approximation scheme allows to reduce this computation to the resolution of some classical boundary values

that can be solved thanks to a standard toolbox. This leads to a non-intrusive implementation of the proposed method. We conclude in Section 6 with numerical examples concerning the robust optimization of the compliance of a mechanical structure.

2. Formal presentation of the main idea

In this section, we formally outline the main idea of our approach in a finite-dimensional setting where calculations can be performed in an intuitive way by using only elementary algebra. To this end, let $\mathcal{H}$ be a vector space of designs h, whose performances are evaluated by a cost function $\mathcal{C}(h,\omega)$ which depends on h via the solution $u(h,\omega) = (u_i(h,\omega))_{i=1,\dots,N}$ of the N-dimensional system

$$\mathcal{A}(h)u(h,\omega) = f(\omega). \tag{2.1}$$

In this formula, $\mathcal{A}(h) \in \mathbb{R}^{N^2}$ is an invertible matrix of dimension $N \times N$, $f(\omega)$ is a (random) vector in $\mathbb{R}^N$, and $\omega \in \Omega$ is an event, belonging to a complete probability space $(\Omega, \Sigma, \mathbb{P})$. The cost function $\mathcal{C}$ is assumed to be *quadratic*, i.e., of the form

$$\mathcal{C}(h,\omega) = \langle \mathcal{B}u(h,\omega), u(h,\omega)\rangle = \mathcal{B} : (u(h,\omega) \otimes u(h,\omega)),$$

where $\mathcal{B} \in \mathbb{R}^{N^2}$ is independent of the design for the sake of simplicity. In this formula, the tensor product $v \otimes w$ of two vectors $v, w \in \mathbb{R}^N$ is the $(N \times N)$-matrix with entries $(v \otimes w)_{i,j} = v_i w_j$, $i, j = 1, \dots, N$, and $:$ stands for the Frobenius inner product over matrices.

The objective function of interest is the mean value of the cost $\mathcal{C}(h,\omega)$:

$$\mathcal{M}(h) = \int_\Omega \mathcal{C}(h,\omega)\, \mathbb{P}(d\omega) = \mathcal{B} : \mathrm{Cor}(u)(h). \tag{2.2}$$

Here, $\mathrm{Cor}(u)(h)$ is the $N \times N$ *correlation matrix*[1] of $u(h,\omega)$, whose entries read

$$\mathrm{Cor}(u)(h)_{i,j} = \int_\Omega u_i(h,\omega) u_j(h,\omega)\, \mathbb{P}(d\omega), \quad i,j = 1,\dots,N.$$

This matrix can be calculated as the solution to the N^2-dimensional system

$$(\mathcal{A}(h) \otimes \mathcal{A}(h))\, \mathrm{Cor}(u)(h) = \mathrm{Cor}(f). \tag{2.3}$$

At this point, let us recall that $\mathcal{A}(h) \otimes \mathcal{A}(h) : \mathbb{R}^{N^2} \to \mathbb{R}^{N^2}$ is the unique linear mapping such that

$$\forall u, v \in \mathbb{R}^N,\ \ (\mathcal{A}(h) \otimes \mathcal{A}(h))(u \otimes v) = (\mathcal{A}(h)u) \otimes (\mathcal{A}(h)v).$$

Let us now calculate the gradient of $\mathcal{M}(h)$. Denoting with $'$ the differentiation with respect to h, we differentiate (2.1,2.2) in an arbitrary direction $\widehat{h}$ to obtain

$$\mathcal{M}'(h)(\widehat{h}) = 2\int_\Omega \langle \mathcal{B}u'(h,\omega)(\widehat{h}), u(h,\omega)\rangle\, \mathbb{P}(d\omega), \tag{2.4}$$

where

$$\mathcal{A}(h)u'(h,\omega)(\widehat{h}) = -\mathcal{A}'(h)(\widehat{h})u(h,\omega).$$

Introducing the adjoint state $p(h,\omega)$, which is the solution to the (N-dimensional) system

$$\mathcal{A}(h)^T p(h,\omega) = -2\mathcal{B}^T u(h,\omega), \tag{2.5}$$

we derive successively

$$2\langle \mathcal{B}u'(h,\omega)(\widehat{h}), u(h,\omega)\rangle = -\langle u'(h,\omega)(\widehat{h}), \mathcal{A}(h)^T p(h,\omega)\rangle = \langle \mathcal{A}'(h)(\widehat{h})u(h,\omega), p(h,\omega)\rangle.$$

Hence, we arrive at

$$\mathcal{M}'(h)(\widehat{h}) = \Big(\mathcal{A}'(h)(\widehat{h}) \otimes I\Big)\, \mathrm{Cor}(u,p)(h).$$

In this last formula, the correlation matrix $\mathrm{Cor}(u,p)(h)$ can be calculated as the solution to an N^2-dimensional system; indeed, using (2.1,2.5), one has for any event ω that

$$\left(\mathcal{A}(h) \otimes \mathcal{A}(h)^T\right)(u(h,\omega) \otimes p(h,\omega)) = -\left(\mathcal{A}(h) \otimes \mathcal{B}\right)(u(h,\omega) \otimes u(h,\omega)),$$

[1] Unfortunately, the literature is not consistent as far as the notions of *covariance* and *correlation* are concerned. In this article, the correlation between two square-integrable random variables X and Y is defined as $\mathbb{E}(XY)$.

whence the following system for $\mathrm{Cor}(u,p)$:

$$\left(\mathcal{A}(h)\otimes\mathcal{A}(h)^T\right)\mathrm{Cor}(u,p)(h) = -\left(\mathcal{A}(h)\otimes\mathcal{B}\right)\mathrm{Cor}(u)(h). \tag{2.6}$$

These considerations show that both, the objective function $\mathcal{M}(h)$ and its gradient, can be exactly calculated from the sole datum of the correlation matrix of f (and not of its law!).

At this point, one may wonder about the practical interest of the foregoing computations since the systems (2.3,2.6) seem difficult to solve (see however [27]). The main idea consists in performing a low-rank approximation of the correlation matrix $\mathrm{Cor}(f)$:

$$\mathrm{Cor}(f) \approx \sum_{i=1}^{m} f_i \otimes f_i, \quad m \ll N.$$

Then, formula (2.3) leads to the calculation of a convenient approximation of $\mathrm{Cor}(u)(h)$ in accordance with

$$\mathrm{Cor}(u)(h) \approx \sum_{i=1}^{m} u_i(h)\otimes u_i(h),$$

where $u_i(h)$ arises as the solution of the system

$$\mathcal{A}(h)u_i(h) = f_i. \tag{2.7}$$

Similarly, by (2.6), one has

$$\mathrm{Cor}(u,p)(h) \approx \sum_{i=1}^{m} u_i(h)\otimes p_i(h)$$

with

$$\mathcal{A}(h)^T p_i(h) = -\mathcal{B}^T u_i(h). \tag{2.8}$$

Hence, calculating $\mathcal{M}(h)$ and its derivative $\mathcal{M}'(h)$ amounts to solving (only) m systems of the form (2.7) and m systems of the form (2.8).

Remark 1.

- Cost functionals $\mathcal{C}$ of the design involving a linear term of the form $\ell(u(h,\omega))$ can be considered in the same way (see Section 4.1.2). The corresponding mean value also involves the mean value of u, $\mathbb{E}(u)(h) := \int_\Omega u(h,\omega)\,\mathbb{P}(d\omega)$.
- Formulae (2.2,2.4) show explicit expressions of $\mathcal{M}(h)$ and $\mathcal{M}'(h)$ only in terms of the correlation $\mathrm{Cor}(f)$, which is quite appealing for at least two reasons. First, $\mathrm{Cor}(f)$ may be imperfectly known (in realistic applications, it is often reconstructed from observations by statistical methods). Second, as we have just discussed, it is often desirable to approximate it, so to ease numerical computations. In either situation, these formulae allow to measure directly the impact of an approximation of $\mathrm{Cor}(f)$ on $\mathcal{M}(h)$ and $\mathcal{M}'(h)$.
- An alternative approach for the calculation of $\mathcal{M}(h)$ and $\mathcal{M}'(h)$ consists in computing a truncated Karhunen-Loève expansion of $f(\omega)$, i.e., $f(\omega) \approx \sum_{i=1}^m f_i\xi_i(\omega)$ with $\{f_i\}$ being orthogonal vectors in $\mathbb{R}^N$ and $\{\xi_i\}$ being uncorrelated random variables. Injecting this expression in (2.1) yields an approximation of u in accordance with $u(h,\omega) \approx \sum_{i=1}^m u_i(h)\xi_i(\omega)$, where the $u_i(h)$ are given by (2.7). Then, using the quadratic structure of the cost $\mathcal{C}$ allows to conveniently approximate $\mathcal{M}(h)$ and $\mathcal{M}'(h)$, leading to similar formulae. Doing so is however less efficient than the proposed approach for at least two reasons. On the one hand, calculating the Karhunen-Loève expansion of a random field is rather involved in terms of computational cost. In contrast, the proposed method in this article relies on *any* low-rank approximation of the correlation $\mathrm{Cor}(f)$. On the other hand, estimating the error entailed by such a process is awkward, since it does not rely on the direct connection between $\mathcal{M}(h)$, $\mathcal{M}'(h)$ and $\mathrm{Cor}(u)$, $\mathrm{Cor}(u,p)$ (it passes through the approximation of $u(h,\omega)$ itself, which is not directly involved in their expressions).

3. Shape optimization setting

Extending the framework presented in the previous section to the infinite dimensional setting, and more specifically to that of shape optimization, demands adequate extensions of the notions of random variables, correlation matrices, etc. At the centre of these generalizations lies the notion of tensor product between Hilbert spaces, about which this section gathers some useful material for the reader's convenience.

3.1. **Differentiation with respect to the domain: Hadamard's method.**

Several notions of differentation with respect to the domain are available in the literature. In this article, we rely on Hadamard's boundary variation method (see e.g. [1, 17, 20]), which advocates to describe variations of a bounded, Lipschitz domain $D \subset \mathbb{R}^d$ of the form (see Figure 1):

$$D_\theta = (I+\theta)(D), \;\; \theta \in W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d), \;\; ||\theta||_{W^{1,\infty}(\mathbb{R}^d,\mathbb{R}^d)} \leq 1. \tag{3.1}$$

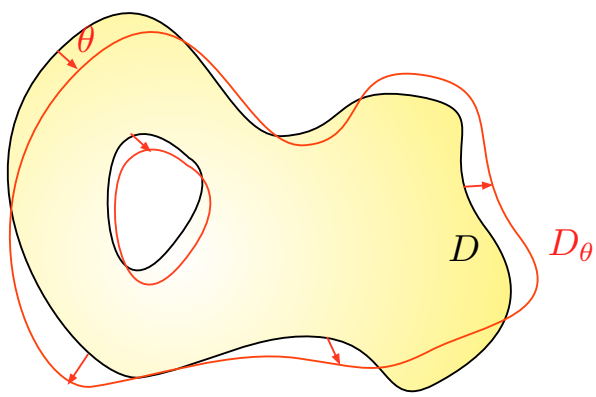

Figure 1. One variation D_θ of a shape D in the framework of Hadamard's method.

Definition 1. *A function $J(D)$ of the domain is said to be shape differentiable at D if the underlying functional $\theta \mapsto J(D_\theta)$ which maps $W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$ into $\mathbb{R}$ is Fréchet-differentiable at $\theta = 0$. The shape derivative $\theta \mapsto J'(D)(\theta)$ of J at D is the corresponding Fréchet derivative, so that the following asymptotic expansion holds in the vicinity of $\theta = 0$:*

$$J(D_\theta) = J(D) + J'(D)(\theta) + o(\theta), \;\; \textit{where} \;\; \frac{|o(\theta)|}{||\theta||_{W^{1,\infty}(\mathbb{R}^d,\mathbb{R}^d)}} \overset{\theta \to 0}{\longrightarrow} 0.$$

In practice, shape optimization problems are defined only over a set $\mathcal{U}_{ad}$ of *admissible shapes* (which e.g. satisfy volume or regularity constraints). To ensure that variations (3.1) of admissible shapes remain admissible, one usually imposes that the deformations θ lie in a subset $\Theta_{ad} \subset W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$ of admissible deformations.

In the following, we implicitly and systematically assume that the sets $\mathcal{U}_{ad}$ and Θ_{ad} contain shapes or deformations with sufficient regularity to legitimate the use of the classical formulae for the shape derivatives of the considered functionals. We refer e.g. to [17] for precise statements on these issues.

3.2. **Tensor products and Hilbert spaces.**

In this subsection, we collect some definitions and properties around the notion of tensor product of Hilbert spaces. A more thorough exposition can be found in [18, 25]. In what follows, we consistently assume all the Hilbert spaces to be separable. This is merely a matter of convenience and most of the forthcoming definitions and results hold also in the general context.

Definition 2. *Let $(H_1, \langle\cdot\rangle_{H_1})$, $(H_2, \langle\cdot\rangle_{H_2})$ be two (separable) Hilbert spaces. Then, for any $h_1 \in H_1$, $h_2 \in H_2$, the pure tensor $h_1 \otimes h_2$ is the bilinear form acting on $H_1 \times H_2$ as*

$$\forall(\varphi_1, \varphi_2) \in H_1 \times H_2,\ \ (h_1 \otimes h_2)(\varphi_1, \varphi_2) = \langle h_1, \varphi_1\rangle_{H_1} \langle h_2, \varphi_2\rangle_{H_2}. \tag{3.2}$$

The vector space of all pure tensors

$$\mathcal{H} = \operatorname{span}\{h_1 \otimes h_2,\ \ h_1 \in H_1,\ h_2 \in H_2\}$$

has a well-defined inner product $\mathcal{H} \times \mathcal{H} \to \mathbb{R}$ which is determined by its action on pure tensors

$$\forall \varphi_1 \otimes \varphi_2,\ \psi_1 \otimes \psi_2 \in \mathcal{H},\ \ \langle \varphi_1 \otimes \varphi_2, \psi_1 \otimes \psi_2\rangle = \langle \varphi_1, \psi_1\rangle_{H_1} \langle \varphi_2, \psi_2\rangle_{H_2} \tag{3.3}$$

and extended to $\mathcal{H}$ by bilinearity. In particular, this inner product does not depend on the choice of the decomposition of elements of $\mathcal{H}$ as finite sums of pure tensors used to calculate it. The tensor product $H_1 \otimes H_2$ is finally the Hilbert space which is defined as the completion of $\mathcal{H}$ for $\langle\cdot,\cdot\rangle$.

Remark 2. Alternative definitions can be found in the literature, e.g. relying on the notion of Hilbert-Schmidt operators.

Proposition 1. *Let H_1, H_2 be Hilbert spaces and let $\{\phi_i\}_{i\in\mathbb{N}}$, $\{\psi_j\}_{j\in\mathbb{N}}$ be associated orthonormal bases.*

(1) The set $\{\phi_i \otimes \psi_j\}_{i,j\in\mathbb{N}}$ is an orthonormal basis of $H_1 \otimes H_2$ for the inner product defined in (3.3).
(2) For any $h \in H_1 \otimes H_2$, there exists a unique family $\{u_i\}_{i\in\mathbb{N}}$ of elements in H_1 such that

$$h = \sum_{i=1}^{\infty} u_i \otimes \psi_i.$$

The main purpose of tensor algebra is to transform multilinear expressions into linear ones. Before stating a version of the universal property of tensor products in the Hilbert space context, let us recall that a *weak Hilbert-Schmidt mapping* $b : H_1 \times H_2 \to K$ between Hilbert spaces H_1, H_2, $(K, \langle\rangle_K)$ is a bilinear mapping with the property that there is a constant $c > 0$ such that

$$\forall u \in K,\ \ \sum_{i,j=1}^{\infty} |\langle b(\phi_i, \psi_j), u\rangle_K|^2 \leq c||u||_K^2$$

for given (then any) orthonormal bases $\{\phi_i\}_{i\in\mathbb{N}}$ of H_1 and $\{\psi_j\}_{j\in\mathbb{N}}$ of H_2, respectively.

Proposition 2. *Let H_1, H_2, K be Hilbert spaces, and $b : H_1 \times H_2 \to K$ be a weak Hilbert-Schmidt mapping. Then, there exists a unique bounded operator $\ell : H_1 \otimes H_2 \to K$ such that the following diagram is commutative*

$$\begin{tikzcd} H_1 \times H_2 \arrow[d] \arrow[dr, "b"] & \\ H_1 \otimes H_2 \arrow[r, "\ell"] & K \end{tikzcd}$$

where the mapping $H_1 \times H_2 \to H_1 \otimes H_2$ is simply $(h_1, h_2) \mapsto h_1 \otimes h_2$.

We now come to the following very important identification of L^2 spaces taking values in a Hilbert space.

Proposition 3. *Let (Ω, μ) be a measured space, and H be a Hilbert space. Then, the mapping*

$$L^2(\Omega, \mu) \times H \ni (\xi, h) \mapsto \xi h \in L^2(\Omega, \mu, H)$$

induces a natural isomorphism $L^2(\Omega, \mu) \otimes H \simeq L^2(\Omega, \mu, H)$ between Hilbert spaces.

Example 1. In the particular case that $H = L^2(D,\nu)$, where (D,ν) is another measured space, Proposition 3 yields the isomorphism $L^2(\Omega,\mu)\otimes L^2(D,\nu)\approx L^2(\Omega\times D,\mu\otimes\nu)$, where $\mu\otimes\nu$ stands for the usual product measure of μ and ν on $\Omega\times D$ and the above identification is supplied by

$$\forall u\in L^2(\Omega,\mu),\ v\in L^2(D,\nu),\ \ (u\otimes v)(x,y)=u(x)v(y),\ \ x\in\Omega,\ y\in D.$$

In the following, we consistently employ this identification.

3.3. **First and second moments analysis.**

In this section, we slip into the probabilistic context, so to speak, relying on the framework of [27, 28]. Hence, let $(\Omega,\Sigma,\mathbb{P})$ be a complete probability space and let H be a (separable) Hilbert space. To keep notation simple, we omit to mention the measure $\mathbb{P}$ on Ω when the context is clear.

Definition 3. *(1) The mean value $\mathbb{E}: L^2(\Omega)\otimes H\to H$ is the unique linear and bounded operator which satisfies*

$$\forall\xi\in L^2(\Omega),\ u\in H,\ \ \mathbb{E}(\xi\otimes u)=\left(\int_\Omega\xi\right)u.$$

(2) The correlation $\mathrm{Cor}(u,v)\in H\otimes H$ between two elements $u,v\in L^2(\Omega)\otimes H$ is defined as

$$\mathrm{Cor}(u,v)=\sum_{i=1}^{\infty}u_i\otimes v_i,$$

where $u=\sum_{i=1}^{\infty}\xi_i\otimes u_i$ and $v=\sum_{i=1}^{\infty}\xi_i\otimes v_i$ are the decompositions of u and v supplied by Proposition 1, according to an orthonormal basis $\{\xi_i\}_{i\in\mathbb{N}}$ of $L^2(\Omega)$. $\mathrm{Cor}(u,v)$ is moreover independent of the basis $\{\xi_i\}_{i\in\mathbb{N}}$ used to perform the above construction.

(3) The function $\mathrm{Cor}(u,u)$ is simply denoted as $\mathrm{Cor}(u)$ and called the (two-point) correlation of u.

This terminology is consistent with the usual definitions of the mean and correlation of random fields. Indeed, if $D\subset\mathbb{R}^d$ is a domain (equipped with the usual Lebesgue measure) and $H=L^2(D)$, then it is easily seen that the mean value $\mathbb{E}(u)\in L^2(D)$ of an element $u\in L^2(\Omega)\otimes L^2(D)\simeq L^2(\Omega\times D)$ is

$$\mathbb{E}(u)=\int_\Omega u(\cdot,\omega)\,\mathbb{P}(d\omega),$$

and that the correlation $\mathrm{Cor}(u,v)\in L^2(D)\otimes L^2(D)\simeq L^2(D\times D)$ of $u,v\in L^2(\Omega)\otimes L^2(D)$ is given by

$$\mathrm{Cor}(u,v)(x,y)=\int_\Omega u(x,\omega)v(y,\omega)\mathbb{P}(d\omega)\ \ a.e.\ (x,y)\in D\times D.$$

For further use, we have to give a sense to expressions of the form $\mathrm{Cor}(u,v)(x,x)$, where $u,v\in L^2(\Omega)\otimes L^2(D)$ and $x\in D$. Note that this is not completely straightforward since $\mathrm{Cor}(u,v)$ is only an element in $L^2(D\times D)$; hence, it is a priori not defined on null measure subsets of $D\times D$. Doing so is the purpose of the next lemma:

Lemma 4. *(1) Let $\mathcal{F}\subset L^2(D\times D)$ be the subspace defined as*

$$\mathcal{F}=\mathrm{span}\left\{u\otimes v,\ \ u,v\in L^2(D)\right\},$$

equipped with the nuclear norm

$$||h||_*=\inf\left\{\sum_{i=1}^{N}||u_i||_{L^2(D)}||v_i||_{L^2(D)},\ \ h=\sum_{i=1}^{N}u_i\otimes v_i,\ u_i,v_i\in L^2(D)\right\}.$$

There is a unique linear and continuous operator $\gamma:\mathcal{F}\to L^1(D)$ such that, for any functions $u,v\in\mathcal{D}(D)$, it holds that

$$\gamma(u\otimes v)(x)=u(x)v(x)\ \ a.e.\ x\in D.$$

(2) Let $\mathcal{F}_c \subset L^2(D \times D)$ be the subspace defined as

$$\mathcal{F}_c = \operatorname{span}\left\{\operatorname{Cor}(u,v),\ \ u,v \in L^2(\Omega)\otimes L^2(D)\right\},$$

equipped with the norm

$$||h||_{c,*} = \inf\left\{\sum_{i=1}^N ||u_i||_{L^2(\Omega\times D)}||v_i||_{L^2(\Omega\times D)},\ \ h=\sum_{i=1}^N \operatorname{Cor}(u_i,v_i),\ u_i,v_i\in L^2(\Omega)\otimes L^2(D)\right\}.$$

There is a unique linear and continuous operator $\gamma_c : \mathcal{F}_c \to L^1(D)$ such that, for any functions $u,v \in \mathcal{D}(\Omega\times D)$, it holds that

$$\gamma_c(\operatorname{Cor}(u,v))(x) = \operatorname{Cor}(u,v)(x,x)\ \ a.e.\ x\in D.$$

(3) The following diagram is commutative:

$$\begin{array}{ccc} L^2(\Omega\times D)^2 & \xrightarrow{\operatorname{Cor}} & L^2(D\times D) \\ \downarrow{\scriptstyle\gamma} & & \downarrow{\scriptstyle\gamma_c} \\ L^1(\Omega\times D) & \xrightarrow{\int_\Omega} & L^1(D) \end{array}$$

Proof. 1. First, note that $||\cdot||_*$ does define a norm on $\mathcal{F}$ since, for arbitrary $h\in\mathcal{F}$, one has that $||h||_{L^2(D\times D)} \leq ||h||_*$. What is more, the subspace $\mathcal{G} := \operatorname{span}\{u\otimes v,\ \ u,v\in\mathcal{D}(D)\}$ is obviously dense in $\mathcal{F}$ for the norm $||\cdot||_*$. Define now the mapping $\gamma : \mathcal{G}\to L^1(D)$ by

$$\forall h = \sum_{i=1}^N u_i\otimes v_i,\ \ u_i,v_i\in\mathcal{D}(D),\quad \gamma(h)(x) = \sum_{i=1}^N u_i(x)v_i(x)\ \text{ a.e. } x\in D.$$

This mapping is obviously well-defined and continuous provided that $\mathcal{G}$ is endowed with the norm $||\cdot||_*$. Thus, it is uniquely extended into a linear and bounded operator $\gamma:\mathcal{F}\to L^1(D)$, which fulfills the desired properties.

2. The proof is completely analogous to that of *1.*

3. The commutation relations obviously hold for smooth functions $\varphi,\psi\in\mathcal{D}(\Omega\times D)$, and the general result follows by continuity of the mappings at play. □

Remark 3.

- To keep notations simple, we will often omit to mention explicitly the operators γ and γ_c, and, e.g. we write expressions such as $(u\otimes v)(x,x)$ instead of $\gamma(u\otimes v)(x,x)$.
- Analogous definitions and commutation relations, involving different operators (derivatives, etc.), hold and can be proved in an identical way when the space H used in the Definition 3 of the correlation function is, for instance, $H=[L^2(D)]^d$ or $H=H^1(D)$. We do not specify these natural relations so to keep the exposition simple.

4. Applications in concrete situations

4.1. Quadratic shape functionals in the context of the Poisson equation.

Let D be a smooth domain and $f\in L^2(\Omega, L^2(\mathbb{R}^d))$ be a random source term. For almost any event $\omega\in\Omega$, let $u_D(\cdot,\omega)\in H^1_0(D)$ be the unique solution to the Poisson equation

$$\left\{\begin{array}{ll} -\Delta u(\cdot,\omega) = f(\omega) & \text{in } D, \\ u(\cdot,\omega) = 0 & \text{on } \partial D. \end{array}\right. \tag{4.1}$$

By the standard Lax-Milgram theory, this equation has a unique solution in $H^1_0(D)$ for almost all events $\omega\in\Omega$, and it is easily seen that $u_D\in L^2(\Omega,H^1_0(D))$. Moreover, owing to the standard elliptic regularity theory (see e.g. [5]), it holds $u(\cdot,\omega)\in H^2(D)$ for almost all events $\omega\in\Omega$.

4.1.1. *The Dirichlet energy as cost function.*

The first functional under consideration is the mean value $\mathcal{M}(D)$ of the Dirichlet energy:

$$\mathcal{M}(D) = \int_\Omega \mathcal{C}(D,\omega)\mathbb{P}(d\omega), \quad \mathcal{C}(D,\omega) = -\frac{1}{2}\int_D ||\nabla u_D(x,\omega)||^2\, dx.$$

The result of interest is as follows.

Theorem 5. *The objective function $\mathcal{M}(D)$ can be rewritten as*

$$\mathcal{M}(D) = -\frac{1}{2}\int_D \langle \nabla \otimes \nabla \rangle \operatorname{Cor}(u_D)(x,x)\, dx. \tag{4.2}$$

This functional is shape differentiable over $\mathcal{U}_{ad}$ and its shape gradient is given by

$$\forall \theta \in \Theta_{ad},\ \mathcal{M}'(D)(\theta) = -\frac{1}{2}\int_{\partial D} \left(\frac{\partial}{\partial n} \otimes \frac{\partial}{\partial n}\right) \operatorname{Cor}(u_D)(x,x)\, (\theta \cdot n)(x)\, ds(x). \tag{4.3}$$

In the last two formulae, $\langle \nabla \otimes \nabla \rangle : H_0^1(D) \otimes H_0^1(D) \to L^2(D) \otimes L^2(D)$ and $(\frac{\partial}{\partial n} \otimes \frac{\partial}{\partial n}) : H^2(D) \otimes H^2(D) \to L^2(\partial D) \otimes L^2(\partial D)$ stand for the linear forms induced by the respective bilinear mappings

$$(u,v) \mapsto \nabla u \cdot \nabla v \text{ and } (u,v) \mapsto \frac{\partial u}{\partial n}\frac{\partial v}{\partial n}.$$

The correlation function $\operatorname{Cor}(u_D) \in H_0^1(D) \times H_0^1(D) \simeq H_0^1(D \times D)$ can be obtained as the unique solution of the boundary value problem

$$\begin{cases} -(\Delta \otimes \Delta)(\operatorname{Cor}(u_D)) = \operatorname{Cor}(f) & \text{in } D \times D, \\ \operatorname{Cor}(u_D) = 0 & \text{on } \partial(D \times D). \end{cases} \tag{4.4}$$

Proof. We find

$$\begin{aligned} \mathcal{M}(D) &= -\frac{1}{2}\int_\Omega \int_D ||\nabla u_D(x,\omega)||^2\, dx\, \mathbb{P}(d\omega) \\ &= -\frac{1}{2}\int_\Omega \int_D \gamma(\langle \nabla \otimes \nabla \rangle (u_D(\cdot,\omega) \otimes u_D(\cdot,\omega)))(x,x)\, dx\, \mathbb{P}(d\omega). \end{aligned}$$

Now using Lemma 4 (see also Remark 3), we obtain

$$\begin{aligned} \mathcal{M}(D) &= -\frac{1}{2}\int_D \gamma_c \left(\int_\Omega \langle \nabla \otimes \nabla \rangle (u_D(\cdot,\omega) \otimes u_D(\cdot,\omega))(x,x)\, \mathbb{P}(d\omega) \right) dx \\ &= -\frac{1}{2}\int_D \gamma_c \left((\nabla \otimes \nabla) \left(\int_\Omega u_D \otimes u_D\, \mathbb{P}(d\omega) \right) \right)(x,x)\, dx, \end{aligned}$$

which implies the desired expression (4.2).

To prove (4.3), we follow the same analysis, starting from the classical formula for the shape derivative, for a given event $\omega \in \Omega$ (see e.g. [17]):

$$\mathcal{C}'(D,\omega)(\theta) = -\frac{1}{2}\int_{\partial D} \left| \frac{\partial u_D(\cdot,\omega)}{\partial n} \right|^2 \theta \cdot n\, ds.$$

Finally, that u_D is the unique solution to the system (4.4) follows from tensorizing the state equation (4.1) and applying the usual Lax-Milgram theory (again, see [27] for details). □

4.1.2. *L^2-tracking type cost function.*

Still in the setting of the Poisson equation outlined above, we are now interested in the L^2-tracking type cost function

$$\mathcal{C}(D,\omega) = \frac{1}{2}\int_B |u_D(x,\omega) - u_0(x)|^2\, dx,$$

where u_D is the solution to (4.1), $B \Subset D$ is a fixed subset of D, and $u_0 \in L^2(B)$ is a prescribed function. Again, we aim at minimizing the mean value of the cost $\mathcal{C}$, that is

$$\mathcal{M}(D) = \int_\Omega \mathcal{C}(D,\omega)\,\mathbb{P}(d\omega).$$

The main result in the present context is the following theorem:

Theorem 6. *The functional $\mathcal{M}(D)$ defined above can be rewritten as*

$$\mathcal{M}(D) = \frac{1}{2}\int_D \big(\mathrm{Cor}(u_D)(x,x) - 2u_0(x)\mathbb{E}(u_D)(x) + u_0^2(x)\big)\,dx. \tag{4.5}$$

It is shape differentiable at any shape $D \in \mathcal{U}_{ad}$ with shape derivative given by

$$\forall \theta \in \Theta_{ad},\ \ \mathcal{M}'(D)(\theta) = -\int_{\partial D}\left(\frac{\partial}{\partial n}\otimes\frac{\partial}{\partial n}\right)\mathrm{Cor}(p_D,u_D)(x,x)(\theta\cdot n)(x)\,ds(x). \tag{4.6}$$

In this formula, the adjoint state $p_D \in L^2(\Omega)\otimes H_0^1(D)$ satisfies the boundary value problem

$$a.e.\ \omega\in\Omega,\ \begin{cases} -\Delta p(\cdot,\omega) = -\chi_B(u_D(\cdot,\omega)-u_0) & \text{in } D,\\ p(\cdot,\omega) = 0 & \text{on } \partial D,\end{cases} \tag{4.7}$$

where χ_B stands for the characteristic function of B. Moreover, the mean value $\mathbb{E}(u_D) \in H_0^1(D)$ which enters (4.5) is the unique solution of

$$\begin{cases} -\Delta\mathbb{E}(u) = \mathbb{E}(f) & \text{in } D,\\ \mathbb{E}(u) = 0 & \text{on } \partial D,\end{cases} \tag{4.8}$$

the two-point correlation $\mathrm{Cor}(u_D)$ is the solution of (4.4), and the correlation function $\mathrm{Cor}(p_D,u_D) \in H_0^1(D)\otimes H_0^1(D)$ can be calculated by solving the boundary value problem

$$\begin{cases} -(\Delta\otimes I)\mathrm{Cor}(p,u) = -(\chi_B\otimes I)(\mathrm{Cor}(u) - u_0\otimes\mathbb{E}(u)) & \text{in } D\times D,\\ \mathrm{Cor}(p,u) = 0 & \text{on } \partial(D\times D).\end{cases} \tag{4.9}$$

Proof. The proof is essentially identical to that of Theorem 5 and hence is not repeated here. Nevertheless, let us just point out that (4.9) stems from the computation

$$\begin{aligned} -(\Delta\otimes I)\mathrm{Cor}(p_D,u_D)(x,y) &= \int_\Omega -\Delta p_D(x,\omega)u_D(y,\omega)\mathbb{P}(d\omega)\\ &= -\chi_B(x)\int_\Omega (u_D(x,\omega)-u_0(x))u_D(y,\omega)\mathbb{P}(d\omega)\\ &= -\chi_B(x)\left(\mathrm{Cor}(u)(x,y) - u_0(x)\mathbb{E}(u_D)(y)\right)\end{aligned}$$

for almost all $(x,y)\in D\times D$. □

4.2. **Quadratic functionals in the context of linear elasticity.**

We now slip into the context of the linear elasticity system. The shapes $D\subset\mathbb{R}^d$ under consideration are filled with a linear elastic material with Hooke's law A given by

$$\forall e\in\mathcal{S}(\mathbb{R}^d),\ \ Ae = 2\mu e + \lambda\,\mathrm{tr}\,e\,I,$$

where the *Lamé coefficients* λ and μ satisfy $\mu>0$ and $\lambda+2\mu/d>0$.

The admissible shapes $D\in\mathcal{U}_{ad}$ are clamped on a fixed subset Γ_D of their boundaries, surface loads are applied on another fixed, disjoint part $\Gamma_N\subset\partial D$, so that only the *free boundary* $\Gamma := \partial D\setminus(\Gamma_D\cup\Gamma_N)$ is subject to optimization. Accordingly, we shall assume that all the deformation fields $\theta\in\Theta_{ad}$ vanish on $\Gamma_D\cup\Gamma_N$. Omitting body forces for simplicity, the displacement u_D of D belongs to the space $[H^1_{\Gamma_D}(D)]^d$, where

$$H^1_{\Gamma_D}(D) = \left\{u\in H^1(D),\ \text{s.t. } u=0 \text{ on } \Gamma_D\right\},$$

and is the unique solution in this space to the boundary value problem

$$
\begin{cases}
-\mathrm{div}(Ae(u)) = 0 & \text{in } D, \\
u = 0 & \text{on } \Gamma_D, \\
Ae(u)n = g & \text{on } \Gamma_N, \\
Ae(u)n = 0 & \text{on } \Gamma,
\end{cases}
\tag{4.10}
$$

where $e(u) = (\nabla u + \nabla u^T)/2$ stands for the linearized strain tensor.

The cost function at stake is the *compliance* of shapes

$$
\mathcal{C}(D,\omega) = \int_D Ae(u_D)(x,\omega) : e(u_D)(x,\omega)\, dx = \int_D g(x,\omega)\cdot u_D(x,\omega)\, ds(x),
$$

and we still aim at optimizing its mean value $\mathcal{M}(D) = \int_\Omega \mathcal{C}(D,\omega)\,\mathbb{P}(d\omega)$. Arguing as in the previous subsection, we obtain the following result.

Theorem 7. *The above functional $\mathcal{M}(D)$ can be rewritten in accordance with*

$$
\mathcal{M}(D) = \int_D ((Ae_x : e_y)\mathrm{Cor}(u))(x,x)\, dx,
$$

where $(Ae_x : e_y) : [H^1_{\Gamma_D}(D)]^d \otimes [H^1_{\Gamma_D}(D)]^d \to L^2(D)\otimes L^2(D)$ is the linear operator induced by Proposition 2 from the bilinear mapping

$$
(u,v) \mapsto Ae(u) : e(v).
$$

This functional is differentiable at any shape $D \in \mathcal{U}_{ad}$ and its derivative reads

$$
\forall \theta \in \Theta_{ad},\ \ \mathcal{M}'(D)(\theta) = -\int_\Gamma ((Ae_x : e_y)\mathrm{Cor}(u))(x,x)(\theta\cdot n)(x)\, ds(x).
$$

Here, the two-point correlation function $\mathrm{Cor}(u) \in [H^1_{\Gamma_D}(D)]^d \otimes [H^1_{\Gamma_D}(D)]^d$ is the unique solution to the following boundary value problem:

$$
\begin{cases}
(\mathrm{div}_x \otimes \mathrm{div}_y)(Ae_x \otimes Ae_y)\,\mathrm{Cor}(u) = 0 & \text{in } D\times D, \\
\mathrm{Cor}(u) = 0 & \text{on } \Gamma_D\times\Gamma_D, \\
(\mathrm{div}_x \otimes I_y)(Ae_x \otimes I_y)\,\mathrm{Cor}(u) = 0 & \text{on } D\times\Gamma_D, \\
(I_x \otimes \mathrm{div}_y)(I_x \otimes Ae_y)\,\mathrm{Cor}(u) = 0 & \text{on } \Gamma_D\times D, \\
(Ae_x \otimes Ae_y)\,\mathrm{Cor}(u)(n_x\otimes n_y) = \mathrm{Cor}(g) & \text{on } \Gamma_N\times\Gamma_N, \\
(\mathrm{div}_x \otimes I_y)(Ae_x \otimes Ae_y)\,\mathrm{Cor}(u)(I_x\otimes n_y) = 0 & \text{on } D\times(\Gamma_N\cup\Gamma), \\
(I_x \otimes \mathrm{div}_y)(Ae_x \otimes Ae_y)\,\mathrm{Cor}(u)(n_x\otimes I_y) = 0 & \text{on } (\Gamma_N\cup\Gamma_N)\times D, \\
(Ae_x \otimes Ae_y)\,\mathrm{Cor}(u)(n_x\otimes n_y) = 0 & \text{on } \big((\Gamma_N\cup\Gamma)\times(\Gamma_N\cup\Gamma)\big)\setminus(\Gamma_N\times\Gamma_N), \\
(Ae_x \otimes I_y)\,\mathrm{Cor}(u)(n_x\otimes I_y) = 0 & \text{on } (\Gamma_N\times\Gamma)\times\Gamma_D, \\
(I_x \otimes Ae_y)\,\mathrm{Cor}(u)(I_x\otimes n_y) = 0 & \text{on } \Gamma_D\times(\Gamma_N\times\Gamma).
\end{cases}
$$

Remark 4. All the involved mappings in the foregoing expressions are naturally produced by Proposition 2, and we do not make the underlying functional spaces explicit. The subscripts $_x$ and $_y$ refer to operators acting respectively on the first and second component of a pure tensor in a tensor product space.

5. Numerical realization

In this section, we now focus on how the previous formulae for the objective functions of interest and their derivatives pave the way to efficient calculations in numerical practice.

5.1. **Computing second moments.**

Without loss of generality, we focus the discussion on the setting of the Poisson equation, as discussed in Section 4.1. The expressions (4.2,4.3,4.5,4.6) involve the mean value $\mathbb{E}(u)$ and the correlation $\mathrm{Cor}(u)$ of the solution $u(\cdot,\omega)$ to (4.1) and the correlation $\mathrm{Cor}(u,p)$ between u and the solution $p(\cdot,\omega)$ to (4.7).

The quantity $\mathbb{E}(u)$ is fairly straightforward to calculate once the mean of the data $\mathbb{E}(f)$ is known; indeed, it arises as the solution to the boundary value problem (4.8) which can be solved owing to any standard finite element method.

It is however more complicated to compute $\mathrm{Cor}(u)$ (or $\mathrm{Cor}(u,p)$) since, in accordance with (4.4), a fairly unusual boundary value problem for the tensorized Laplace operator needs to be solved on the product domain $D \times D$. This moderately high-dimensional problem can be solved in essentially the same complexity as (4.8) if a sparse tensor product discretization is employed as proposed in e.g. [13, 15, 16]. However, the implementation of this approach is highly intrusive, insofar as it demands a dedicated solver.

One way to get past this difficulty, which is also much simpler to implement, consists in relying on an expansion of the two-point correlation function of $f(\cdot,\omega)$ of the form

$$\mathrm{Cor}(f) = \sum_k f_k \otimes f_k. \tag{5.1}$$

Then, the two-point correlation function $\mathrm{Cor}(u)$ can be expressed as

$$\mathrm{Cor}(u) = \sum_k u_k \otimes u_k, \tag{5.2}$$

where each function u_k is the solution to the Poisson equation (4.1) with data f_k. In other terms, the solution's two-point correlation function can be determined from solving (possibly infinite) standard boundary value problems. In practice, the expansion (5.1) is truncated so that this process becomes feasible.

Several possibilities are available when it comes to decomposing $\mathrm{Cor}(f)$ as in (5.1). For example, in the situation that $\mathrm{Cor}(f) \in L^2(D \times D)$, the most natural idea is to perform a spectral decomposition, as an application of Mercer's theorem

$$\mathrm{Cor}(f) = \sum_k \lambda_k (\phi_k \otimes \phi_k), \tag{5.3}$$

where (λ_k, ϕ_k) are the eigenpairs of the associated Hilbert-Schmidt operator

$$L^2(D) \ni \phi \mapsto \int_D \mathrm{Cor}(f)(\cdot, y)\phi(y)\, dy \in L^2(D).$$

Another way to obtain the decomposition (5.1) is a (possibly infinite) Cholesky decomposition of the two-point correlation function. Pivoting the Cholesky decomposition yields an extremely efficient approximation method, see e.g. [13, 14].

5.2. **Numerical calculation of a low-rank approximation of** $\mathrm{Cor}(f)$**.**

In general, the expansion (5.1) is infinite and has to be appropriately truncated for numerical computations. Let $V := \mathrm{span}\{\varphi_i : i = 1, 2, \ldots, N\} \subset L^2(D)$ be a suitable discretization of $L^2(D)$, e.g. a finite element space associated with a mesh of D. We are looking for a low-rank approximation of $\mathrm{Cor}(f)$ in the tensor product space $V \otimes V$:

$$\mathrm{Cor}(f)(x,y) \approx \sum_{k=1}^{m} \left(\sum_{i=1}^{n} \ell_{k,i}\varphi_i(x) \right) \left(\sum_{j=1}^{n} \ell_{k,j}\varphi_j(y) \right) \in V \otimes V. \tag{5.4}$$

The unknown coefficient vectors in (5.4) can be computed as follows. Define the discrete correlation matrix $C \in \mathbb{R}^{N^2}$ as

$$C_{i,j} = \int_D \int_D \mathrm{Cor}(f)(x,y)\varphi_i(x)\varphi_j(y)\, dxdy, \;\; i,j = 1, \ldots, N.$$

and the mass matrix $G \in \mathbb{R}^{N^2}$ as

$$G_{i,j} = \int_{\partial D} \varphi_i(x)\varphi_j(x)dx, \;\; i,j = 1,\dots,N.$$

Then, it is easily seen that searching for a decomposition of the form (5.4) translates, in terms of matrices, into the search of an approximation C_m such that

$$C \approx C_m = \sum_{k=1}^{m} \widetilde{\ell}_k \widetilde{\ell}_k^T \quad \text{with} \quad \ell_k = (\ell_{k,i})_{i=1,\dots,N} = G^{-1}\widetilde{\ell}_k,$$

in such a way that the truncation error $\|C - C_m\|$ is rigorously controlled (in a way yet to be defined).

The best low-rank approximation in $L^2(D \times D)$ is known to be the truncated spectral decomposition (5.3) (see e.g. [28]). In the discrete setting, this corresponds to the spectral decomposition of C, which is a very demanding task. In particular, the decay of the eigenvalues $\{\lambda_k\}$ and thus the rank m to be reached for an accurate decomposition depend heavily on the smoothness of the underlying two-point correlation function $\mathrm{Cor}(f)$. Related decay rates have been proved in [28].

We suggest instead to employ the pivoted Cholesky decomposition in order to compute a discrete low-rank approximation of $\mathrm{Cor}(f)$, as originally proposed in [14]. It is a purely algebraic approach which is quite simple to implement. It produces a low-rank approximation of the matrix C for any given precision $\varepsilon > 0$ where the approximation error is rigorously controlled in the trace norm. A rank-m approximation is computed in $\mathcal{O}(m^2 n)$ operations. Exponential convergence rates in m hold under the assumption that the eigenvalues of C exhibit a sufficiently fast exponential decay, see [14] for details. Nevertheless, numerical experiments suggest that, in general, the pivoted Cholesky decomposition converges optimally in the sense that the rank m is uniformly bounded with respect to the truncation error ε by the number of terms required for the spectral decomposition of $\mathrm{Cor}(f)$ to get the same error ε.

5.3. **Low-rank approximation of the shape functional and its gradient.**

The basic idea is now to insert the state's expansion (5.2) into the expressions of the expectation of the random shape functional and the associated shape gradient to derive computable expressions. In fact, it turns out that only standard solvers for boundary value problems need to be provided. Loosely speaking, this implies that, if one can compute the shape functional and its gradient for a deterministic right hand side, then one can also evaluate the expectation of the shape functional and its gradient for a random right-hand side. We illustrate this idea with the three examples introduced in Section 4:

- Having decomposed $\mathrm{Cor}(u)$ as $\mathrm{Cor}(u) = \sum_k u_k \otimes u_k$, with each function u_k satisfying

$$\left\{ \begin{array}{rl} -\Delta u_k = f_k & \text{in } D, \\ u_k = 0 & \text{on } \partial D, \end{array} \right. \tag{5.5}$$

  the mean value (4.2) of the Dirichlet energy can be computed as:

$$\mathcal{M}(D) = -\frac{1}{2}\sum_k \int_D \|\nabla u_k\|^2 \, dx.$$

  Moreover, its shape derivative is given by

$$\forall \theta \in \Theta_{ad}, \;\; \mathcal{M}'(D)(\theta) = -\frac{1}{2}\int_{\partial D} \left(\sum_k \left| \frac{\partial u_k}{\partial n} \right|^2 \right) \theta \cdot n \, ds(x).$$

- The mean value (4.5) of the L^2-tracking type functional considered in Section 4.1.2 can be computed as

$$\mathcal{M}(D) = \frac{1}{2}\int_B \sum_k \left(u_k - u_0\right)^2 dx$$

with the u_k given by (5.5). As for the calculation of the shape gradient (4.6), we have to introduce the adjoint states $p_k \in H_0^1(D)$, defined by

$$\begin{cases} -\Delta p_k = -(u_k - u_0) & \text{in } D, \\ p_k = 0 & \text{on } \partial D, \end{cases}$$

Thus, in view of the fact that $\mathrm{Cor}(p_D, u_D) = \sum_k p_k \otimes u_k$, we are led to the following formula for the shape gradient:

$$\forall \theta \in \Theta_{ad},\ \mathcal{M}'(D)(\theta) = -\int_{\partial D} \left(\sum_k \frac{\partial u_k}{\partial n} \frac{\partial p_k}{\partial n} \right) \theta \cdot n \, ds(x).$$

- Last but not least, in the linear elasticity setting of Section 4.2, expanding the correlation function of the surface loads $\mathrm{Cor}(g) = \sum_k g_k \otimes g_k$, the two-point correlation $\mathrm{Cor}(u)$ satisfies the expansion $\mathrm{Cor}(u) = \sum_k u_k \otimes u_k$ with u_k given by

$$\begin{cases} -\mathrm{div}(Ae(u_k)) = 0 & \text{in } D, \\ u_k = 0 & \text{on } \Gamma_D, \\ Ae(u_k)n = g_k & \text{on } \Gamma_N, \\ Ae(u_k)n = 0 & \text{on } \Gamma. \end{cases}$$

  Hence, the mean value $\mathcal{M}(D)$ of the compliance is given by

$$\mathcal{M}(D) = \sum_k \int_D Ae(u_k) : e(u_k) \, dx,$$

  while its shape derivative reads

$$\forall \theta \in \Theta_{ad},\ \mathcal{M}'(D)(\theta) = -\int_\Gamma \left(\sum_k Ae(u_k) : e(u_k) \right) \theta \cdot n \, dx.$$

Remark 5.

- In the last example, one may be interested in the case that random body forces $f(x, \omega)$ are also applied to the system (which are, e.g. assumed to be uncorrelated with the surface loads g). Then, one has to perform two low-rank expansions $\mathrm{Cor}(f) = \sum_k f_k \otimes f_k$ and $\mathrm{Cor}(g) = \sum_l g_l \otimes g_l$ which leads to an expansion of u of the form $\mathrm{Cor}(u) = \sum_{k,l} u_{k,l} \otimes u_{k,l}$ with the $u_{k,l}$ given by

$$\begin{cases} -\mathrm{div}(Ae(u_{k,l})) = f_k & \text{in } D, \\ u_{k,l} = 0 & \text{on } \Gamma_D, \\ Ae(u_{k,l})n = g_l & \text{on } \Gamma_N, \\ Ae(u_{k,l})n = 0 & \text{on } \Gamma. \end{cases}$$

- The above formulae coincide with those for the multiple load objective functions (and their derivatives) proposed, e.g. in [3]. In contrast to this work, here, the different load cases are not known a priori, but originate from a low-rank approximation of the correlation function of the data.
- Hitherto, we have only been considering low-rank approximations of $\mathrm{Cor}(f)$ of the form $\mathrm{Cor}(g) \approx \sum_k g_k \otimes g_k$, where the g_k are deterministic data functions, as they are naturally produced by the pivoted Cholesky decomposition. Notice however that the above discussion straightforwardly extends to the case of a low-rank decomposition of the kind $\mathrm{Cor}(g) \approx \sum_k (g_k \otimes \widetilde{g}_k + \widetilde{g}_k \otimes g_k)$ (see the example of Section 6.2 for an application of this remark).

6. Numerical examples

We eventually propose two numerical examples which illustrate the main features of this article; both of them take place in the setting of linear elasticity as considered in Section 4.2.

6.1. **Presentation of the numerical algorithm.**

When it comes to the numerical implementation of shape optimization algorithms, one main difficulty lies in the robust representation of shapes and their evolution. To achieve this, we rely on the level set method, initially introduced in [22] and brought to the context of shape optimization in [4, 30].

The basic idea is to consider a shape $D \subset \mathbb{R}^d$ as the negative subdomain of an auxiliary 'level set' function $\phi : \mathbb{R}^d \to \mathbb{R}$, i.e.

$$\forall x \in \mathbb{R}^d, \quad \begin{cases} \phi(x) < 0 & \text{if } x \in D, \\ \phi(x) = 0 & \text{if } x \in \partial D, \\ \phi(x) > 0 & \text{if } x \in {}^c\overline{D}. \end{cases}$$

Thus, the motion of a domain $D(t)$, $t \in [0, T]$, induced by a velocity field with normal amplitude $V(t, x)$, translates in terms of an associated level set function $\phi(t, \cdot)$ as a *Hamilton-Jacobi* equation:

$$\frac{\partial \phi}{\partial t} + V|\nabla \phi| = 0, \; t \in (0, T), \; x \in \mathbb{R}^d. \tag{6.1}$$

Hence, a (difficult) domain evolution problem is replaced by a (hopefully easier) PDE problem. Note that, in the present situation, V stems from the analytical formula for the shape derivative of the considered objective function $\mathcal{M}(D)$, which enjoys the structure (see Theorem 7):

$$\forall \theta \in \Theta_{ad}, \;\; \mathcal{M}'(D)(\theta) = \int_\Gamma \mathcal{D}_D \, \theta \cdot n \, ds,$$

where $\mathcal{D}_D$ is a scalar function.

In numerical practice, the whole space $\mathbb{R}^d$ is reduced to a large computational box D_0, equipped with a fixed triangular mesh $\mathcal{T}$. Each shape $D \subset D_0$ is represented by means of a level set function ϕ, discretized at the vertices of $\mathcal{T}$. In this context the elastic displacement u_D, solution to the linear elasticity system (4.10), which is involved in the computation of $\mathcal{D}_D$, cannot be calculated exactly since no mesh of D is available. Therefore, we employ the Ersatz material approach [4] to achieve this calculation approximately: the problem (4.10) is transferred to a problem on D_0 by filling the void part $D_0 \setminus \overline{D}$ with a very soft material, whose Hooke's law is εA with $\varepsilon \ll 1$.

All our finite element computations are performed within the `FreeFem++` environment [24], and we rely on algorithms from our previous works [6, 10], based on the method of characteristics, when it comes to *redistancing* level set functions or solving equation (6.1).

6.2. **Comparison between correlated and uncorrelated loads.**

This first example is aimed at appraising the influence of correlation between different sets of loads applied to the shapes. Let us consider the situation depicted on Figure 2 (left): a bridge is clamped on its lower part, and two sets of loads $g_a = (1, -1)$ and $g_b = (-1, 1)$ are applied on its superior part, which may or may not be correlated. The actual loads are then modelled as a random field $g(x, \omega)$ of the form

$$g(x, \omega) = \xi_1(\omega) g_a(x) + \xi_2(\omega) g_b(x),$$

where the random variables ξ_1 and ξ_2 are normalized so that

$$\int_\Omega \xi_i \, \mathbb{P}(d\omega) = 0, \quad \int_\Omega \xi_i^2 \, \mathbb{P}(d\omega) = 1, \quad i = 1, 2.$$

The degree of correlation between g_a and g_b is measured by $\alpha := \int_\Omega \xi_1 \xi_2 \, \mathbb{P}(d\omega)$ where the case $\alpha = 0$ corresponds to uncorrelated loads. In this context, the correlation function $\mathrm{Cor}(g) \in [L^2(\Gamma_N)]^d \times [L^2(\Gamma_N)]^d$ naturally arises as a finite sum of pure tensor products (so that no low-rank approximation is necessary), and reads

$$\mathrm{Cor}(g) = g_a \otimes g_a + g_b \otimes g_b + \alpha \left(g_a \otimes g_b + g_b \otimes g_a \right).$$

The mean value $\mathcal{M}(D)$ of the compliance of shapes and its derivative can be calculated explicitly, along the lines of Section 5.3 (see also Remark 5).

We run several examples, associated to different values of the degree of correlation $|\alpha| \leq 1$. In each situation, an equality constraint $\mathrm{Vol}(D) = \int_D dx = 0.35$ on the volume of shapes is enforced owing to a standard Augmented Lagrangian procedure (see [21], §17.4). Starting from the initial shape of Figure 2 (right), 250

iterations of the algorithm outlined in Section 6.1 are performed. The mesh $\mathcal{T}$ of the computational domain is composed of 12 141 vertices and 23 800 triangles; the CPU time for each example is approximately 12 min on a MacBook Air with a 1.8 GHz Intel Core i5 with 4 GB of RAM. The resulting optimal shapes are represented in Figure 3, and the evolution of the objective function $\mathcal{M}(D)$ and the volume $\mathrm{Vol}(D)$ can be appraised on the histories of Figure 4.

A huge difference in trends can be observed, depending on the degree of correlation between the loads (observe also the values of the objective function $\mathcal{M}(D)$ in Figure 4). Roughly speaking, as α gets closer and closer to -1, the shapes have to withstand the 'worst-case' of the situations entailed by g_a and g_b.

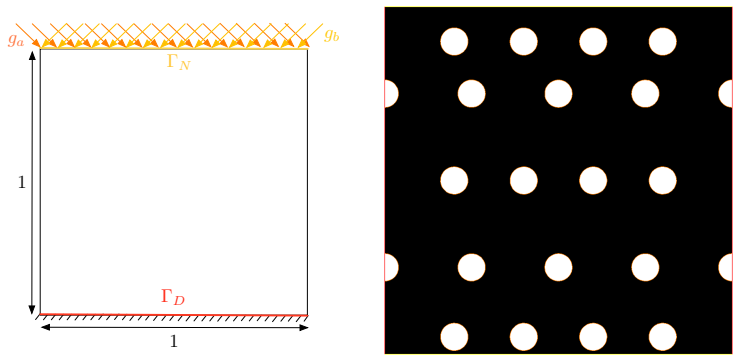

FIGURE 2. Details of the test case of Section 6.2 (left) and initial shape (right).

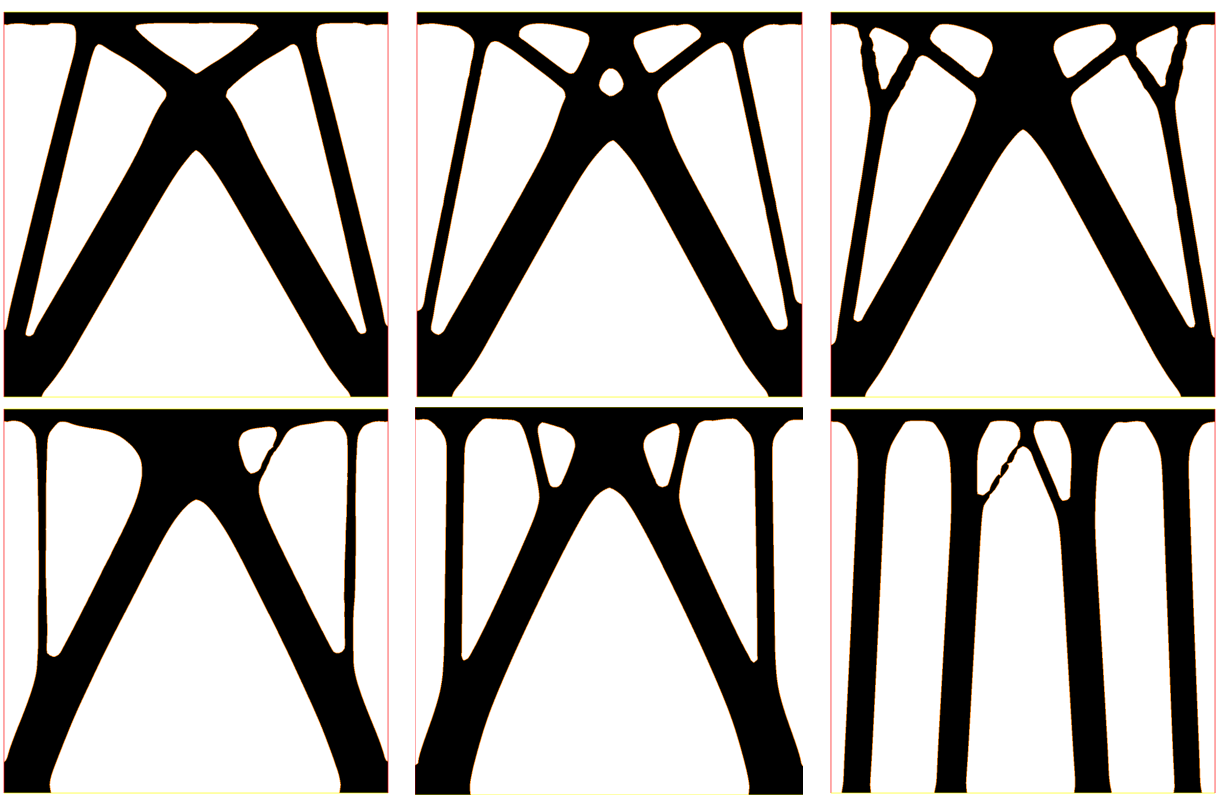

FIGURE 3. Optimal shapes obtained in the test case of Section 6.2, associated to degrees of correlation $\alpha = -1, -0.7, 0, 0.5, 0.8, 1$ (from left to right, top to bottom).

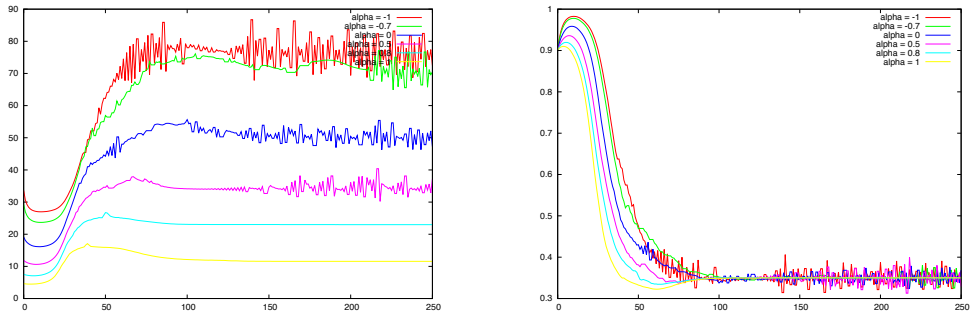

FIGURE 4. Convergence histories for the mean value (left) and the volume (right) in the test case of Section 6.2.

6.3. **An example with a more complex correlation function.**

Let us turn to an example where the correlation function of the data is no longer trivial (i.e., it cannot be written as a finite sum of pure tensor products). The situation at stake is depicted on Figure 6: a bridge is clamped on its lower part and (random) surface loads $g = (g_1, g_2)$ are applied on its top.

We study three different scenarii, corresponding to surface loads $g^i = (g_1^i, g_2^i)$, $i = 1, 2, 3$. For the sake of simplicity, in all three cases, the horizontal and vertical components g_1^i and g_2^i are uncorrelated; the associated correlation functions are given by

$$\forall x, y \in \Gamma_N, \quad \begin{cases} \mathrm{Cor}(g_1^i)(x,y) = 10^5\, h_i^+\left(\dfrac{x_2+y_2}{2}\right) e^{-\frac{|x_1-y_1|}{l}}, \\ \mathrm{Cor}(g_2^i)(x,y) = 10^6\, k_i^+\left(\dfrac{x_2+y_2}{2}\right) e^{-\frac{|x_1-y_1|}{l}}, \end{cases}$$

where the superscript $^+$ stands for the positive part, the characteristic length l is taken as $l = 0.1$, and the functions h_i and k_i $(i = 1, 2, 3)$ are defined as (see also the graphs in Figure 5)

$$h_1(t) = 1 - 4\left(t - \frac{1}{2}\right)^2, \quad k_1(t) = \begin{cases} 16\left(t - \frac{1}{4}\right)^2, & \text{if } t \le \frac{1}{2}, \\ 16\left(t - \frac{3}{4}\right)^2, & \text{else,} \end{cases}$$

$$h_2(t) = 2t(1-t) + \frac{1}{2}, \qquad k_2(t) = \begin{cases} 24\left(t - \frac{1}{4}\right)\left(t - \frac{1}{3}\right), & \text{if } t \le \frac{1}{2}, \\ 24\left(t - \frac{3}{4}\right)\left(t - \frac{2}{3}\right), & \text{else,} \end{cases}$$

$$h_3(t) = 1, \qquad k_3(t) = \begin{cases} 24\left(t - \frac{1}{4}\right)\left(t - \frac{1}{6}\right), & \text{if } t \le \frac{1}{2}, \\ 24\left(t - \frac{3}{4}\right)\left(t - \frac{5}{6}\right), & \text{else.} \end{cases}$$

Loosely speaking, these correlation functions show a decreasing dependence on the distance $|x - y|$ between two points $x, y \in \Gamma_N$, and the factors h_i, k_i, which depend only on the average position $(x + y)/2$, mimic a variable intensity of the loads according to the spatial location. Note that the pivoted Cholesky decomposition, as described in Section 5, is used to obtain low-rank approximations of these correlation functions of the form (5.1), and we retain the first five terms in each expansion.

The objective function of interest is, again, the mean value $\mathcal{M}(D)$ of the compliance of the structure. A constraint $\mathrm{Vol}(D) = 0.75$ is imposed on the volume of shapes and 250 iterations of the algorithm outlined in Section 6.1 are performed in each situation on a computational mesh composed of 5 752 vertices and 11 202 triangles which requires a CPU time of approximately 15 min. The resulting optimal shapes and convergence histories are reported on Figures 7 and 8, respectively, showing very different trends depending on the particular situation.

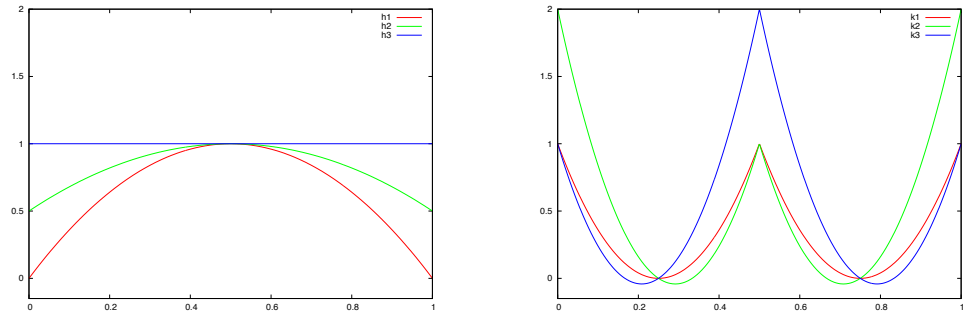

FIGURE 5. Graphs of the functions h_i (left) and graphs of the functions k_i (right).

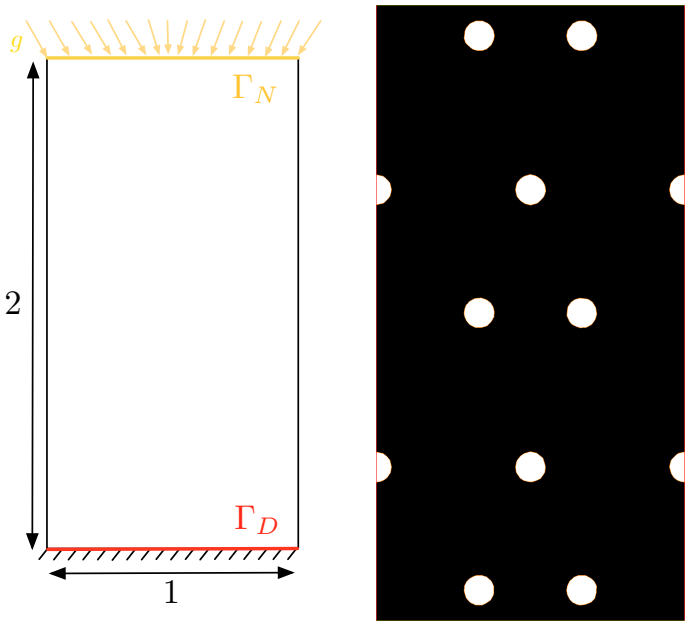

FIGURE 6. The setup of the test case of Section 6.3 (left) and initial shape (right).

REFERENCES

[1] G. ALLAIRE, *Conception optimale de structures*. Vol. 58 of Mathématiques et Applications, Springer, Berlin, 2007.

[2] G. ALLAIRE AND C. DAPOGNY, A linearized approach to worst-case design in parametric and geometric shape optimization. *Math. Models Meth. Appl. Sciences*, 24(11):2199–2257, 2014

[3] G. ALLAIRE AND F. JOUVE, A level-set method for vibration and multiple loads structural optimization. *Comput. Meths. Appl. Mech. Engrg.*, 194:3269–3290, 2005.

[4] G. ALLAIRE, F. JOUVE, AND A.M. TOADER, Structural optimization using shape sensitivity analysis and a level-set method. *J. Comput. Phys.*, 194:363–393, 2004.

[5] H. BREZIS, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York, 2011.

[6] C. BUI, C. DAPOGNY AND P. FREY, An accurate anisotropic adaptation method for solving the level set advection equation. *Int. J. Numer. Methods in Fluids*, 70(7):899–922, 2012.

[7] S. CHEN AND W. CHEN, A new level-set based approach to shape and topology optimization under geometric uncertainty. *Struct. Multidisc. Optim.*, 44(1):1–18, 2011.

[8] S. CHEN, W. CHEN, AND S. LEE, Level set based robust shape and topology optimization under random field uncertainties. *Struct. Multidisc. Optim.*, 41(4):507–524, 2010.

[9] S. CONTI, H. HELD, M. PACH, M. RUMPF, AND R. SCHULTZ, Shape optimization under uncertainty. A stochastic programming approach. *SIAM J. Optim.*, 19(4):1610–1632, 2009.

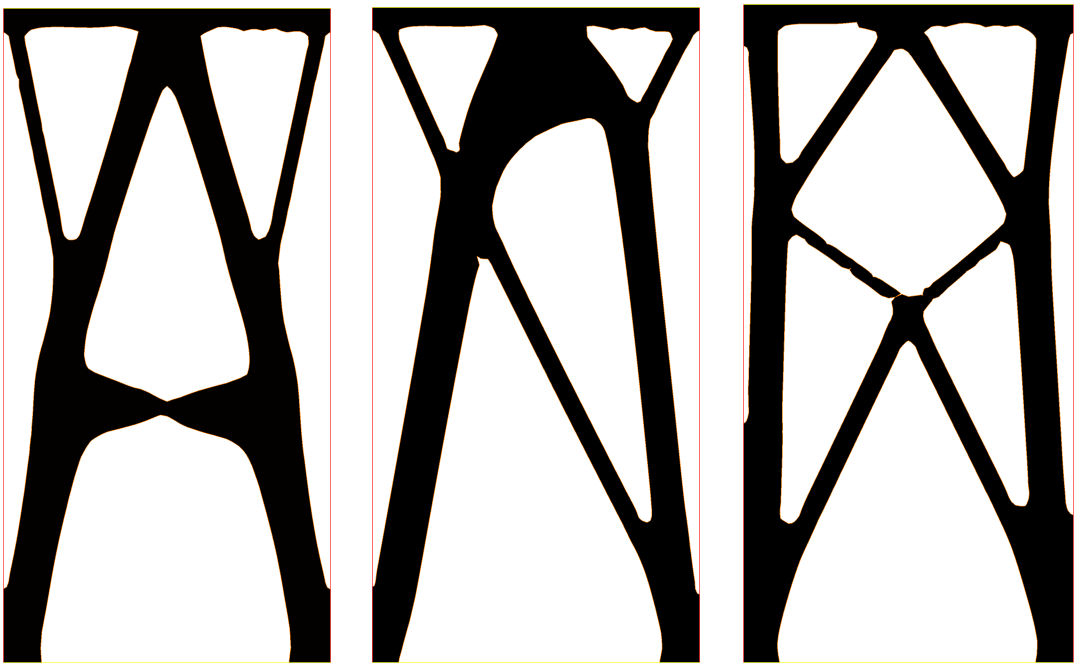

FIGURE 7. Optimal shapes obtained in the situations $1, 2, 3$ in the test case of Section 6.3 (from left to right).

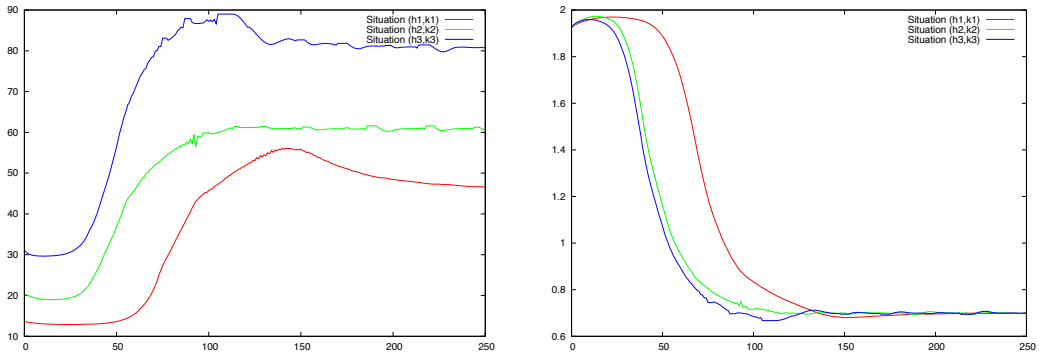

FIGURE 8. Convergence histories for the mean value (left) and the volume (right) in the test case of Section 6.3.

[10] C. DAPOGNY AND P. FREY, *Computation of the signed distance function to a discrete contour on adapted triangulation. Calcolo*, 49(3):193–219, 2012.

[11] M. DELFOUR AND J.-P. ZOLESIO, *Shapes and Geometries*. SIAM, Philadelphia, 2001.

[12] P.D. DUNNING AND H.A. KIM, Robust topology optimization. Minimization of expected and variance of compliance. *AIAA Journal*, 51(11):2656–2664, 2013.

[13] H. HARBRECHT, Second moment analysis for Robin boundary value problems on random domains. In M. Griebel, editor, *Singular Phenomena and Scaling in Mathematical Models*, pages 361–382, Springer, Berlin, 2013.

[14] H. HARBRECHT, M. PETERS, AND R. SCHNEIDER, On the low-rank approximation by the pivoted Cholesky decomposition. *Appl. Numer. Math.*, 62(4):428–440, 2012.

[15] H. HARBRECHT, M. PETERS, AND M. SIEBENMORGEN, Combination technique based k–th moment analysis of elliptic problems with random diffusion. *J. Comput. Phys.*, 252:128–141, 2013.

[16] H. HARBRECHT, R. SCHNEIDER, AND C. SCHWAB, Sparse second moment analysis for elliptic problems in stochastic domains. *Numer. Math.*, 109(3):385–414, 2008.

[17] A. HENROT AND M. PIERRE, *Variation et optimisation de formes, une analyse géométrique*. Vol. 48 of Mathématiques et Applications, Springer, Berlin, 2005.

[18] R.V. Kadison and J.R. Ringrose, *Fundamentals of the theory of operator algebras. Vol. I. Elementary theory.* Vol. 100 of Pure and Applied Mathematics, Academic Press, New York, 1983.

[19] B.S. Lazarov, M. Schevenels, and O. Sigmund, Topology optimization with geometric uncertainties by perturbation techniques. *Int. J. Numer. Meth. Engng.*, 90:1321–1336, 2012.

[20] F. Murat and J. Simon, Étude de problèmes d'optimal design. In J. Céa, editor, *Optimization Techniques, Modeling and Optimization in the Service of Man*, pages 54–62. Vol. 41 of Lect. Notes Comput. Sci., Springer, Berlin, 1976.

[21] J. Nocedal and S.J. Wright, *Numerical optimization.* Springer Science+Business Media, New York, 2006.

[22] S.J. Osher and J.A. Sethian, Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79:12–49, 1988.

[23] O. Pironneau, *Optimal shape design for elliptic systems.* Springer, New York, 1983.

[24] O. Pironneau, F. Hecht, and A. Le Hyaric, *FreeFem++ version 2.15-1*, `http://www.freefem.org/ff++/`.

[25] M.Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. I Functional Analysis*, Academic Press, New York, 1972.

[26] C. Schillings, S. Schmidt, and V. Schulz, Efficient shape optimization for certain and uncertain aerodynamic design. *Computers & Fluids*, 46(1):78–87, 2011.

[27] C. Schwab and R.A. Todor, Sparse Finite Elements for Elliptic Problems with Stochastic Data. *Numer. Math.*, 95:707–734, 2003.

[28] C. Schwab and R.A. Todor, Karhunen-Loéve approximation of random fields by generalized fast multipole methods. *J. Comput. Phys.*, 217:100–122, 2006.

[29] J. Sokolowski and J.-P. Zolesio, *Introduction to Shape Optimization.* Springer, Berlin, 1992.

[30] M. Y. Wang, X. Wang, and D. Guo, A level set method for structural topology optimization. *Comput. Methods. Appl. Mech. Engrg.*, 192:227–246, 2003.