ESAIM: COCV 30 (2024) 66
https://doi.org/10.1051/cocv/2023075

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

PROBABILITY-OF-FAILURE-BASED OPTIMIZATION FOR RANDOM PDES THROUGH CONCENTRATION-OF-MEASURE INEQUALITIES

Rogelio Ortigosa-Martinez[1], Jesus Martinez-Frutos[2] and Francisco Periago[1,*]

Abstract. Control and optimization problems constrained by partial differential equations (PDEs) with random input data and that incorporate probabilities of failure in their formulations are numerically extremely challenging, since the computational cost of estimating the tails of a probability distribution is prohibitive in many situations encountered in real-life engineering problems. In addition, probabilities of failure are often discontinuous and include huge flat regions where gradients vanish. Based on the McDiarmid concentration-of-measure inequality, this paper proposes a new functional which provides a tight and smooth bound for the probability of a given random functional of exceeding a prescribed threshold parameter. Hence, this approach relieves the above-mentioned difficulties in the case where the solution map is convex with respect to the random parameter, as in the case of a deterministic differential operator and the random parameter appearing linearly in the right-hand side term. Well-posedness of the corresponding optimal control problem is established and the viability of the proposed method is numerically illustrated by two benchmarks examples arising in topology optimization and optimal control theory.

Mathematics Subject Classification. 35J20, 49J20, 49M20, 65K10.

Received May 27, 2022. Accepted October 22, 2023.

1. Introduction

During the last decades there has been an increasing interest in mathematical modelling, analysis and numerical simulation methods for optimal control and design of distributed parameter systems with random input data (*e.g.* physical parameters, initial and/or boundary conditions and modelling assumptions).

When uncertainty in those inputs is set within a probabilistic framework (typically using random variables and/or random fields), the output of the physical system (usually modelled as a PDE) is also random and so the considered objective function, which depend on that output, is a random variable. In some situations, one may be interested in obtaining an optimal control or design for each particular realization of the uncertain parameters [11, 25]. However, in many other cases of practical interest it is preferable to get a control/design

Keywords and phrases: Risk-averse control, uncertainty quantification, probability of failure, McDiarmid's inequality.

[1] Department of Applied Mathematics and Statistics, Technical University of Cartagena, Campus Muralla del Mar, 30202 Cartagena (Murcia), Spain.
[2] Computational Mechanics and Scientific Computing Group, Technical University of Cartagena, Cartagena, Spain.

[*] Corresponding author: f.periago@upct.es

independent of specific values of the random inputs and at the same time robust with respect to those random fluctuations of the parameters. In this latter case, a deterministic cost functional ought to be considered, usually involving statistical quantities of interest of the original random objective function. Within this probabilistic framework, the two following approaches are the typical models of choice in applications:

- On the one hand, if one aims at reducing the system variability under variations in the input data, then the natural choice is to average the random cost functional (see *e.g.* [5, 6, 10, 12, 29]). Since the statistical mean is risk neutral, in order to reduce the dispersion in the output, one may add to the averaged functional a measure of dispersion such as the variance of the random cost [16–18]. In that case, the tuning process for the weighting parameters plays a crucial role in order to properly balance the two terms that appear in the cost functional.
- On the other hand, if one is less concerned with the system variability and more with the quantification of expected losses related with the damages caused by extreme events, then the corresponding cost functional and/or the constraints may be defined as the probability that a certain functional exceeds a given threshold parameter (also known as probability of failure in the context of structural optimization).

The present paper is concerned with this latter situation so that a more in depth literature review related to probability-of-failure-based optimization is presented next.

Literature review. Probabilities of failure are discontinuous functions and include huge flat regions where gradients vanish. For these reasons, Monte Carlo and variants are the methods mainly used in optimization for random PDEs in this context [8, 19]. Even so, the estimation of rare events *via* Monte Carlo sampling can be computationally prohibitive in situations where a single evaluation of the cost functional is expensive. Concentration-of-measure inequalities are numerically-efficient tools to relieve the high computational cost of Monte Carlo sampling when it is used to compute probabilities of failure. The use of concentration-of-measures inequalities is not new, as it has been previously considered in several situations. For instance, in [26] the Bonferroni inequalities are used to estimate probability functions, and in [7] several linearized methods are proposed with the same purpose. Variance-based inequalities, such as Bienaymé-Tchebychev inequality, are widely used, mainly because of their easy numerical processing [4]. Jensen and Edmunson-Madansky bounds have also been used in nonlinear mathematical programming [23]. Obviously, for an upper bound to be useful in optimization it must be tight, *i.e.*, it must be close to the associated probability of failure. The well-known bounded differences inequality (also called McDiarmid's or Hoeffding–Azuma inequality) establishes sharp concentration if the involved function does not depend too much on any of its variables [21, 28]. The application of McDiarmid's inequality to certification and the theory of quantification of margins and uncertainties in Engineering has been considered in [13] where the sharpening of McDiarmid's inequality was numerically illustrated with the aid of an imploding ring example. However, its use in random PDE-constrained optimization does not appear to have been attempted to date.

Alternatively, buffered probability of exceedance (bPOE) has been proposed [14] to deal with optimization problems which involve probability functions in their formulations. bPOE counts the number of the largest losses, such that the average of these losses is equal to a given threshold. Also, it has exceptional mathematical properties and it can be easily minimized [9]. However, bPOE is more conservative compared to the measure counting the exceedances, which is the formulation this paper deals with. As a consequence, the applicability of bPOE to some specif problems of interest in Engineering (*e.g.* structural optimization) should be deeply analysed, since optimal solutions derived from bPOE could have a very poor mechanical performance. For the contrary, the tightness of the McDiarmid bound avoids this undesirable mechanical behavior, but the minimization of the corresponding objective function is mathematically more complicated, especially in the absence of suitable convexity conditions.

Contributions and organization of the paper. This article introduces and analyses, both at the theoretical and numerical viewpoints, a new functional (which is derived from McDiarmid's concentration-of-measure inequality) to deal with control and optimization problems which involve probabilities of failure (in the cost function and/or in the constraints). For clarity in the exposition and to explore its applicability in Engineering, the abstract setting is accompanied by two benchmark examples arising in topology optimization and in optimal control theory. Numerical simulation results show an excellent performance of the proposed approach, in terms of statistical accuracy and computability, in the case in which the dependence, with respect to the random parameter, is convex and smooth of the underlying random cost functionals. Indeed, as far as statistical accuracy is concerned, McDiarmid's bound has been proved as a very tight bound for the probability of failure [13]. As for computability, the main gain is achieved in the numerical approximation of the statistical quantities of interest, since the proposed descent algorithm requires, at each iteration, the computation of the expectation of the random cost function and the numerical resolution of a convex optimization problem. On the contrary, computing the tails of a probability function by using Monte-Carlo method is prohibitive in many real applications.

The paper is structured as follows. Section 2 sets out the problem and proposes the new functional. Well-posedness of the corresponding optimal control/design problem is established in Section 3. A gradient-based numerical algorithm is described in Section 4, which is complemented with numerical experiments in Section 5. Both the main strengths and shortcomings of the proposed approach are described in the final Section 6.

2. Setting of the problem

2.1. Functional framework and notation

In this paper, $D \subset \mathbb{R}^d$, $d \geq 1$ a natural number, stands for a physical domain, *i.e.*, a bounded domain with smooth boundary ∂D. The spatial variable is denoted by $x = (x_1, \cdots, x_d) \in D$. $\Gamma \subset \mathbb{R}^N$ is a parameter space. It is assumed that the parameters $z = (z_1, \cdots, z_N) \in \Gamma$ follow a probability law conducted by a density function $\rho : \Gamma \to \mathbb{R}_+$, with $\rho = \rho(z) \in L^\infty(\Gamma)$. Thus, $(\Gamma, \mathcal{B}(\Gamma), \rho(z)\,\mathrm{d}z)$ is a probability space, where $\mathcal{B}(\Gamma)$ is the σ-algebra of Borel sets on Γ and $\rho(z)\,\mathrm{d}z$ is the probability measure, with $\mathrm{d}z$ the Lebesgue measure. By V it is denoted a separable Hilbert space of real valued functions from D, and V' is the topological dual of V. Between both, a pivot Hilbert space H is considered so that $V \hookrightarrow H \hookrightarrow V'$ is a Gelfand triple. $L^2_\rho(\Gamma; V)$ is the Hilbert space composed of all (equivalence classes of) strongly measurable functions $g : \Gamma \to V$ whose norm

$$\|g\|_{L^2_\rho(\Gamma;V)} = \left(\int_\Gamma \|g(\cdot, z)\|^2_V \rho(z)\,\mathrm{d}z \right)^{1/2}$$

is finite.

Hereafter, both the divergence (div) and the gradient (∇) operators involve only derivatives with respect to the spatial variable $x \in D$.

2.2. State law

Going forward, we consider optimization problems constrained by parametric, linear PDEs of the general form

$$\mathcal{L}(y(x, z)) = \mathcal{F}(x, z, u), \quad x \in D,\ z \in \Gamma, \tag{2.1}$$

which also incorporates suitable boundary and/or initial conditions. Hereabove, $\mathcal{L}$ is a linear differential operator with derivatives only with respect to the spatial variable $x \in D$, and $\mathcal{F}$ is a forcing term, which includes a control or design variable u that typically belongs to some subspace of the pivot space H.

We shall write $y = y(z, u)$ to show the dependence of the state variable y on both the random parameter z and on the control variable u.

Our main assumption on the random field PDE solution $y(x, z)$ is:

Assumption 2.1. Set $\mathcal{F} \in L^2_\rho(\Gamma; V')$. Problem (2.1) admits a unique solution in the Hilbert space $L^2_\rho(\Gamma; V)$. Moreover, there exists $C > 0$ such that

$$\|y\|_{L^2_\rho(\Gamma;V)} \leq C\|\mathcal{F}\|_{L^2_\rho(\Gamma;V')}$$

Two typical examples that fulfill this assumption are the following.

2.2.1. Random conductivity equation

Consider the following random elliptic system

$$\begin{cases} -\mathrm{div}\,(a(x,z)\nabla y) = f(x,z), & \text{in } D \times \Gamma \\ y = 0, & \text{on } \partial D \times \Gamma. \end{cases} \tag{2.2}$$

It is very common in applications to model the diffusion coefficient $a(x,z)$ as a truncated Karhunen-Loève expansion of a log-normal random field. We refer the reader to [15], Chapter 2 for concrete examples.

When (2.2) represents a control system, the right-hand side in (2.2) may take the form

$$f(x,z) = b(x,z)u(x), \tag{2.3}$$

with $u = u(x)$ a deterministic control function, and $b(x,z)$ a random field that models uncertainty in the magnitude of the control.

Let $V = H^1_0(D)$ be the usual Sobolev space and let us assume that $f(z) \in V'$ for all $z \in \Gamma$. The variational formulation of problem (2.2) takes the form: find $y = y\,(x,z) \in L^2_\rho(\Gamma; V)$ such that

$$\int_\Gamma \rho(z) \int_D a\nabla y \cdot \nabla v \,\mathrm{d}x\mathrm{d}z = \int_\Gamma \rho(z) < f(z), v > \,\mathrm{d}z \quad \forall v \in L^2_\rho(\Gamma; V), \tag{2.4}$$

where $< \cdot, \cdot >$ denotes the duality product between V and V'. The following hypotheses on the input data are assumed to hold:

(A1) $a = a(x,z) \in L^\infty(\Gamma; L^\infty(D))$ and there exist $a_{min}, a_{max} > 0$ such that

$$0 < a_{min} \leq a(x,z) \leq a_{max} < \infty \quad \text{a.e. } x \in D \text{ and for all } z \in \Gamma.$$

(A2) $f \in L^2_\rho(\Gamma; V')$.

Under assumptions (A1) and (A2), a straightforward application of Lax-Milgram's lemma [15], Theorem. 3.1 proves that the variational problem (2.4) has a unique solution $y = y\,(x,z)$ which satisfies the estimate

$$\|y\|_{L^2_\rho(\Gamma;V)} \leq \frac{C_P}{a_{min}}\|f\|_{L^2_\rho(\Gamma;V')}, \tag{2.5}$$

where $C_P = C_P\,(D)$ is the Poincaré constant.

2.2.2. Random system of linearized elasticity

Let us assume that the boundary of the physical domain D is decomposed into three disjoint parts

$$\partial D = \partial D_D \cup \partial D_N \cup \partial D_0, \quad |\partial D_D| > 0, \tag{2.6}$$

where $|\cdot|$ stands for the Lebesgue measure. Consider the system of linearized elasticity with random input data

$$\begin{cases} -\mathrm{div}(Ae(y(x,z))) = & f \quad \text{in } D\times\Gamma, \\ y(x,z) = & 0 \quad \text{on } \partial D_D\times\Gamma, \\ (Ae(y(x,z)))\cdot n = & g \quad \text{on } \partial D_N\times\Gamma, \\ (Ae(y(x,z)))\cdot n = & 0 \quad \text{on } \partial D_0\times\Gamma, \end{cases} \tag{2.7}$$

with $y=(y_1,\cdots,y_d)$ the displacement vector field, $e(y)=\left(\nabla y^T+\nabla y\right)/2$ the strain tensor and n the unit outward normal vector to ∂D. The fourth order elasticity tensor A is defined (for isotropic linear materials) according to the Kooke's law as

$$A_{ijkl}=\mu\left(\delta_{ikjl}+\delta_{iljk}\right)+\lambda\delta_{ij}\delta_{kl};\qquad i,j,k,l\in\{1,2,3\} \tag{2.8}$$

where δ_{ij} represents the ij^{th} component of the Kronecker delta tensor. In (2.8) , $\lambda=\lambda(x,z)$ and $\mu=\mu(x,z)$ are the Lamé moduli of the material. Thus, $Ae\,(y)$ is the stress tensor. In the same vein, the volume and surface loads f and g depend on a spatial variable x and on a random parameter $z\in\Gamma$, *i.e.* $f=f(x,z)$ and $g=g(x,z)$.

The following assumptions on the uncertain input parameters of system (2.7) are considered:

(B1) $\lambda,\mu\in L^\infty(\Gamma;L^\infty(D))$ and there exist μ_{min}, μ_{max}, λ_{min}, λ_{max} such that

$$0<\mu_{min}\le\mu(x,z)\le\mu_{max}<\infty\quad\text{a.e. } x\in D,\quad\text{for all } z\in\Gamma,$$

and

$$0<\lambda_{min}\le 2\mu(x,z)+d\lambda(x,z)\le\lambda_{max}<\infty\quad\text{a.e. } x\in D,\quad\text{for all } z\in\Gamma,$$

(B2) $f=f(x,z)\in L^2_\rho\left(\Gamma;L^2(D)^d\right)$,
(B3) $g=g(x,z)\in L^2_\rho\left(\Gamma;L^2(\partial D_N)^d\right)$.

Next, the Hilbert space

$$V=\left\{v\in H^1(D)^d : v|_{\partial D_D}=0 \text{ in the sense of traces}\right\},$$

equipped with the usual $H^1(D)^d$-norm, is introduced.

Similarly to the preceding subsection, a weak solution of (2.7) is a random field $y\in L^2_\rho(\Gamma;V)$ such that

$$\int_\Gamma\int_D Ae(y(x,z)):e(v(x,z))\,\mathrm{d}x\,\rho(z)\,\mathrm{d}z=\int_\Gamma\left[\int_D f\cdot v\,\mathrm{d}x+\int_{\partial D_N} g\cdot v\,\mathrm{d}s\right]\rho(z)\,\mathrm{d}z, \tag{2.9}$$

for all $v\in L^2_\rho(\Gamma;V)$. Here above, $A:B=\sum_{i,j=1}^d a_{ij}b_{ij}$ denotes de full contraction of A and B.

By using Korn's inequality and under assumptions (B1)-(B3) it is easy to prove that there exists a unique solution to (2.9). Moreover, there exists $c>0$ such that

$$\|y\|_{L^2_{\mathbb{P}}(\Gamma;V)}\le c\left(\|f\|_{L^2_\rho(\Gamma;L^2(D)^d)}+\|g\|_{L^2_\rho(\Gamma;L^2(\partial D_N)^d)}\right).$$

Remark 2.2. As in the case of the conductivity equation, the control variable may enter the system (2.7) through the loads f and/or g. For instance, we may take $u=g$. However, in the context of Structural Optimization, the design/control variable is the characteristic function of the domain where the problem is posed, *i.e.*, $u(x) = 1_D(x)$. To facilitate the reading of the article, the setting of the optimization problem in this latter situation is presented in Section 2.3.2 below.

Remark 2.3. As in the case of the random conductivity equation, there is no additional difficulty in taking the loads $f(z) \in H^{-1}(D)$ and $g(z) \in H^{-1/2}(\partial D_N)$. This is a typical situation in structural mechanics when point-wise loads are considered.

2.3. A concentration-of-measure-based cost functional for risk-averse optimization

Let $J = J(z,u)$ be a given cost functional associated to the solution of problem (2.1). Here $u = u(x)$ denotes a deterministic control/design variable that is assumed to belong to some suitable set of admissible controls $\mathcal{U}_{ad}$. Two specific examples will be described at the end of this section.

Risk-averse optimization [15], Chapter 3 provides a conservative treatment of uncertainty. In this context, the quantity to be minimized is the probability of exceeding a prescribed threshold parameter $\eta > 0$, *i.e*,

$$\mathbb{P}\left(\{z \in \Gamma \,:\, J\left(z,u\right) \geq \eta\}\right), \tag{2.10}$$

where $\mathbb{P} = \rho\, \mathrm{d}z$.

Remark 2.4. Probabilities of failure, as given by (2.10), may also be a part of inequality constraints. A typical industrial application in structural mechanics is the minimization of mass (or volume) with risk-averse-type constraints for stress and eigenvalue of the device. The analytical and numerical treatment in this situation is very similar to the case in which probability of failure appears in the cost functional. For the sake of clarity, the latter case is considered hereafter.

As indicated in the Introduction, risk-averse optimization involves a very high computational burden. Fortunately, this difficulty may be circumvented with the help of concentration-of-measure inequalities such as McDiarmid's inequality, that is recalled and adapted to our framework next.

For any function $g : \Gamma \to \mathbb{R}$ the McDiarmid diameters of g are defined as

$$D_n(g) = \sup_{z_1,\cdots,z_n,z_n',\cdots,z_N} \left|g\left(z_1,\cdots,z_n,\cdots,z_N\right) - g\left(z_1,\cdots,z_n',\cdots,z_N\right)\right|, \quad 1 \leq n \leq N, \tag{2.11}$$

where $(z_1,\cdots,z_n,\cdots,z_N)$, $(z_1,\cdots,z_n',\cdots,z_N) \in \Gamma$. The McDiarmid diameter system is then

$$D\left(g\right) = \sqrt{\sum_{n=1}^{N} D_n^2(g)}. \tag{2.12}$$

For $g \in L^1_\rho\left(\Gamma\right)$, the McDiarmid inequality (see [21], Lem. 1.2 or [13], Thm. 2.1 for the case of independent random variables, and [22] for centering sequences) states that for every $r \geq 0$

$$\mathbb{P}\left(g - \mathbb{E}\left[g\right] \geq r\right) \leq \exp\left(-2\frac{r^2}{D^2\left(g\right)}\right), \tag{2.13}$$

where $\mathbb{E}\left[g\right] = \int_\Gamma g(z)\rho(z)\,\mathrm{d}z$ is the expectation of g.

Putting $r = \left(\eta - \mathbb{E}\left(g\right)\right)_+$, with $\eta > 0$ and $(t)_+ = \max\{0,t\}$ the positive part of $t \in \mathbb{R}$, in (2.13) yields

$$\mathbb{P}\left(g \geq \eta\right) \leq \exp\left(-2\frac{\left(\eta - \mathbb{E}\left[g\right]\right)_+^2}{D^2\left(g\right)}\right). \tag{2.14}$$

When applied to complex Engineering systems, $g = g(z)$ represents the random output of interest of a given physical system. Similarly to [13], Section 2.1, a conservative certification criterion for that system is then obtained by requiring that

$$\exp\left(-2\frac{(\eta-\mathbb{E}[g])_+^2}{D^2(g)}\right)\leq\epsilon \tag{2.15}$$

for some small certification tolerance $\epsilon > 0$. Thus, if (2.15) holds, then it is certified that, with probability $1-\epsilon$, the output g will not exceed the given threshold η.

Similarly to the theory of quantification of margins and uncertainties [13], the certification criterion (2.15) may be rewritten in the equivalent and simpler form

$$\frac{(\eta-\mathbb{E}[g])_+}{D(g)}\geq\sqrt{\log\sqrt{1/\epsilon}}. \tag{2.16}$$

For our purposes in this paper, inequality (2.14) provides a computationally affordable bound for risk-averse-type cost functionals as that bound depends on the expectation of g. However, the computation of the McDiarmid diameter $D(g)$ is a non trivial global optimization problem in the parameter space Γ. Instead of $D(g)$ one may use the trivial upper bound

$$D(g)\leq N\left(\sup_{z\in\Gamma}g(z)-\inf_{z\in\Gamma}g(z)\right). \tag{2.17}$$

Notice that equality in (2.17) holds in the case $N = 1$.

All things considered and putting in the above discussion $g(z) = J(z,u)$ we are led to consider the new optimization problem

$$\begin{cases}\text{Maximize in } u: & \mathcal{J}^{\#}(u):=\frac{(\eta-\mathbb{E}[J(z,u)])_+}{N\Delta[J(z,u)]}\\ \text{subject to} & \\ & u\in\mathcal{U}_{ad},\end{cases} \tag{2.18}$$

where the expectation $\mathbb{E}[J(z,u)]$ is calculated with respect to the random parameter z, and

$$\Delta[J(z,u)]=\sup_{z\in\Gamma}J(z,u)-\inf_{z\in\Gamma}J(z,u). \tag{2.19}$$

Remark 2.5. The following remarks concerning problem (2.18) are in order:

- Replacing the McDiarmid diameter $D(g)$ by the right-hand side in (2.17) we loss tightness in the bound (2.14). Nonetheless, this simplification is more convenient for our optimization purposes here since (2.19) retains the important information concerning the system's uncertainty and above all leads to a simpler and computationally cheaper optimal control problem.
- The meaning of problem (2.18) is as follows. In order to maximize $\mathcal{J}^{\#}(u)$, the control variable has to do a double task: on one hand, u should shift $\mathbb{E}[J(u)]$ to the left of the threshold η (numerator of $\mathcal{J}^{\#}(u)$); on the other hand, the control ought to reduce the variability (uncertainty) of the system (denominator of $\mathcal{J}^{\#}(u)$). As a consequence, u tends to concentrate the probability measure of $J(z,u)$ around its mean value. In turns, this implies the maximization of $\mathcal{J}^{\#}(u)$ and so the minimization of the desired probability of exceeding the threshold η.

- Notice that if $u^{\#}$ is an optimal control for problem (2.18), then (2.14) and (2.17) provide the estimate

$$\mathbb{P}\left(\left\{z \in \Gamma \,:\, J\left(z, u^{\#}\right) \geq \eta\right\}\right) \leq \exp\left(-2\mathcal{J}^{\#}(u^{\#})^2\right), \tag{2.20}$$

which, similarly to (2.14)-(2.15), also gives a (more conservative) certification criterion for $J\left(z, u^{\#}\right)$ of not exceeding η.

We conclude this section with two concrete examples arising in optimal control and in topology optimization, respectively.

2.3.1. Tracking-type optimal control

In the case of problem (2.2)-(2.3), and for a desired target $y_T \in L^2(D)$, the parameter-dependent cost functional $J(z, u)$ might be

$$J\left(z, u\right) = \frac{1}{2}\|y(z, u) - y_T\|^2_{L^2(D)}, \tag{2.21}$$

where $y = y(z, u)$ is the weak solution to (2.2) and the control $u \in \mathcal{U}_{ad}$, with

$$\mathcal{U}_{ad} = \left\{u \in L^2\left(D\right) : m \leq u\left(x\right) \leq M \quad \text{a.e. } x \in D\right\}, \quad -\infty < m < M < +\infty, \tag{2.22}$$

the set of admissible controls.

Thus, taking into account (2.20), problem (2.18) may be considered as a conservative version of a risk-averse approximate controllability problem.

2.3.2. Structural topology optimization

In the case of structural topology optimization, where the state law is the random system of linearized elasticity (Sect. 2.2.2), the design variable $u(x) = 1_D(x)$ is the characteristic function of the domain D. Since the domain D will change during the optimization process, λ, μ and f must be known for all possible configurations of D. Thus, a working bounded domain $\tilde{D}$ is introduced, which contains all admissible domains D and satisfies that $\partial D_D \cup \partial D_N \subset \partial\tilde{D}$. Therefore, it is assumed that λ, μ and f are defined in $\tilde{D}$. Both ∂D_D and ∂D_N are kept fixed during optimization. Thus, the only part of the boundary that is allowed to change during the optimization process is ∂D_0 (see beginning of Sect. 2.2.2). Besides, a volume constraint is usually imposed on the set of admissible shapes. Thus, the admissibility set $\mathcal{U}_{ad}$ is composed of all characteristic functions 1_D associated to subsets D that satisfy the conditions

$$D \subset \tilde{D}, \quad \partial D_D \cup \partial D_N \subset \partial\tilde{D}, \quad |D| = V_0 \quad \text{with } 0 < V_0 < |\tilde{D}|. \tag{2.23}$$

As for the cost functional, a typical choice is the so-called compliance, which is defined by

$$J(1_D, z) = \int_{\tilde{D}} 1_D(x) f(x, z) \cdot y_D(x, z)\,\mathrm{d}x + \int_{\partial D_N} g(x, z) \cdot y_D(x, z)\,\mathrm{d}s, \tag{2.24}$$

where $y = y_D$ is the solution to (2.7) posed in the domain D.

3. An existence result

This section analyses existence of solutions for the type of optimization problems considered in the preceding section. We start with the following trivial result.

Proposition 3.1. *Let $\Gamma \subset \mathbb{R}^N$ be a compact set and let us assume that for each fixed $u \in \mathcal{U}_{ad}$ the mapping*

$$z \mapsto J(z, u)$$

is continuous. Then, $\Delta[J(u)]$, as given by (2.19), is finite and non-negative.

Remark 3.2. The continuity condition of the mapping $z \mapsto J(z,u)$ is satisfied by the two examples presented in Sections 2.3.1 and 2.3.2 if one assumes enough regularity on the random input data. Indeed, for the case of tracking-type optimal control, where J is given by (2.21), in addition to conditions (A1)-(A2) one has to assume that $a = a(x,z) \in C^0(\Gamma; L^\infty(D))$ and $f \in C^0(\Gamma; L^\infty(D))$. Then, the weak solution $y(z,u)$ of (2.2) satisfies $y \in C^0\left(\Gamma; H_0^1(D)\right)$ and so the mapping

$$z \mapsto J(z, u)$$

is continuous. For a proof, we refer the reader to [1], Lemma 3.1. Similarly, for the case of structural optimization, continuity, with respect to the random parameters, of the Lamé coefficients and forces must be assumed.

Under conditions of Proposition 3.1, let us denote by

$$z_{\max}(u) = \arg\max_{z\in\Gamma} J(z,u) \quad \text{and by } J_{\max}(u) = J(z_{\max}(u), u). \tag{3.1}$$

Remark 3.3. Notice that $z_{\max}(u)$ may be not unique. However, in what follows it is assumed that the parameter space Γ is a compact subset of $\mathbb{R}^N$ and that the map $z \mapsto J(z,u)$ is continuous for each fixed u. Thus, the existence of a global maximum is guaranteed. This maximum may be not unique, yet there is no ambiguity in the definition of the optimization problem (3.2) considered here below.

In order to avoid possible singularities it is convenient to add a small positive parameter in the denominator of $\mathcal{J}(u)$, as given in problem (2.18). Moreover, if $\inf_{z\in\Gamma} J(z,u) \geq 0$, then $\Delta[J(u)] \leq J_{\max}(u)$. Notice that this condition is trivially satisfied by the two benchmark examples presented above. Both things considered lead to the following simplified version of problem (2.18) which, as shown next, enjoys nice convexity properties and so it is a well-posed optimization problem:

$$\begin{cases} \text{Maximize in } u: & \mathcal{J}^\star(u) := \frac{(\eta - \mathbb{E}[J(u)])_+}{N J_{\max}(u) + \varepsilon} \\ \text{subject to} & \\ & u \in \mathcal{U}_{ad}, \end{cases} \tag{3.2}$$

with $\varepsilon > 0$ a small regularization parameter.

Notice that if $\sup_{z\in\Gamma} J(z,u)$ and $\inf_{z\in\Gamma} J(z,u)$ range at different scales, then the solutions obtained by both problems (2.18) and (3.2) will be essentially the same. This is illustrated by the numerical simulation results in Section 5.

Analogously to the case of problem (2.18), if $u^\star$ is a solution to problem (3.2), then the bound

$$\mathbb{P}\left(\{z \in \Gamma \,:\, J(z, u^\star) \geq \eta\}\right) \leq \exp\left(-2\mathcal{J}^\star(u^\star)^2\right) \tag{3.3}$$

holds.

We are now in a position to prove our main existence result.

Theorem 3.4. *In addition to the hypotheses of Proposition 3.1 let us assume:*

(i) the admissibility set $\mathcal{U}_{ad}$ is a bounded, closed and convex subset of $L^2(D)$,

(ii) the control-to-state operator

$$S: \begin{array}{ll} L^2(D) & \to L^2_\rho(\Gamma; H) \\ u & \mapsto S(u) := y(z,u), \end{array}$$

where $y(z,u)$ is the solution to (2.1) and H is the pivot space in the Gelfand tripe $V \hookrightarrow H \hookrightarrow V'$, is linear and continuous.

(iii) the mapping $u \mapsto J(z,u)$ is convex and uniformly Lipschitz continuous, i.e., there exists $C > 0$, which doesn't depend on $z \in \Gamma$, such that

$$|J(z,u) - J(z,v)| \le C\|u - v\|_{L^2(V)} \quad \textit{for all } u, v \in \mathcal{U}_{ad}.$$

Then there exists an optimal solution for problem (3.2).

Proof. Since $(\eta - \mathbb{E}[J(u)])_+ \le \eta$ and $NJ_{\max}(u) + \varepsilon \ge \varepsilon$, the functional $\mathcal{J}^\star(u)$ is uniformly bounded above and so there exists its supremum.

Let $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}_{ad}$ be a maximizing sequence, *i.e.*

$$\lim_{j\to\infty} \mathcal{J}^\star(u_j) = \sup_{u\in\mathcal{U}_{ad}} \mathcal{J}^\star(u).$$

Since $\mathcal{U}_{ad}$ is a bounded, closed and convex subset of $L^2(D)$, there exists $u^\star \in \mathcal{U}_{ad}$ such that for some subsequence, not relabelled,

$$u_j \rightharpoonup u^\star \quad \text{weakly in } L^2(D) \quad \text{as } j \to \infty.$$

The continuity of the control-to-state operator S implies that the mapping

$$u \mapsto -\mathbb{E}[J(z,u)]$$

is continuous. By (iii), it is also concave. Thus, it is weakly upper semi-continuous and so

$$\eta - \mathbb{E}[J(u^\star)] \ge \limsup_{j\to\infty} (\eta - \mathbb{E}[J(u_j)]).$$

Moreover, since the positive part function $(\cdot)_+$ is continuous and non-decreasing,

$$(\eta - \mathbb{E}[J(u^\star)])_+ \ge \limsup_{j\to\infty} (\eta - \mathbb{E}[J(u_j)])_+. \tag{3.4}$$

Let us now analyse the denominator in $\mathcal{J}^\star(u)$. Since for each fixed $z \in \Gamma$ the function $u \mapsto J(z,u)$ is convex, for $0 < \lambda < 1$ and $u_1, u_2 \in \mathcal{U}_{ad}$ one has

$$\begin{aligned} J_{\max}(\lambda u_1 + (1-\lambda)u_2) &= J(z_{\max}(\lambda u_1 + (1-\lambda)u_2), \lambda u_1 + (1-\lambda)u_2) \\ &\le \lambda J(z_{\max}(\lambda u_1 + (1-\lambda)u_2), u_1) + (1-\lambda)J(z_{\max}(\lambda u_1 + (1-\lambda)u_2), u_2) \\ &\le \lambda \max_{z\in\Gamma} J(z,u_1) + (1-\lambda)\max_{z\in\Gamma} J(z,u_2) \\ &= \lambda J(z_{\max}(u_1), u_1) + (1-\lambda)J(z_{\max}(u_2), u_2) \\ &= \lambda J_{\max}(u_1) + (1-\lambda)J_{\max}(u_2), \end{aligned}$$

which proves that $J_{\max}(u)$ is convex. In addition, by (iii), $J_{\max}(u)$ is Lipschitz continuous since it is the supremum of a family of such functions. Thus, it is weakly lower semi-continuous and so

$$J_{\max}(u^\star) \leq \liminf_{j\to\infty} J_{\max}(u_j)\,. \tag{3.5}$$

Finally, by (3.4) and (3.5),

$$\begin{aligned}
\limsup_{j\to\infty} \mathcal{J}^\star(u_j) &= \limsup_{j\to\infty} \frac{(\eta - \mathbb{E}[J(u_j)])_+}{NJ_{\max}(u_j)+\varepsilon} \\
&\leq \limsup_{j\to\infty} (\eta - \mathbb{E}[J(u_j)])_+ \limsup_{j\to\infty} \frac{1}{NJ_{\max}(u_j)+\varepsilon} \\
&= \frac{\limsup_{j\to\infty} (\eta - \mathbb{E}[J(u_j)])_+}{\liminf_{j\to\infty} NJ_{\max}(u_j)+\varepsilon} \\
&\leq \tfrac{(\eta - \mathbb{E}[J(u^\star)])_+}{NJ_{\max}(u^\star)+\varepsilon} \\
&= \mathcal{J}^\star(u^\star)\,,
\end{aligned}$$

which completes the proof. □

Remark 3.5 (Tracking-type optimal control)**.** Let us revisit the control problem formulated in Section 2.3.1. If the right-hand side term $b(x,z)$ in (2.3) is a smooth random field and the random parameter z lies in a compact set (see [1], Sect. 3 for precise statements and examples), it is straightforward to show that hypotheses (ii)–(iii) in Theorem 3.4 are fulfilled. As a consequence, an existence of solution result is obtained in such a case.

Remark 3.6 (Structural topology optimization)**.** Consider now the optimization problem of Section 2.3.2. As before, enough regularity with respect to the random parameter $z \in \Gamma$ in the loads and compactness of the parameter space Γ must be assumed in order to show that hypotheses (ii)–(iii) in Theorem 3.4 are satisfied. As for the set of admissible designs $\mathcal{U}_{ad}$, which in this case is composed of characteristic functions associated to domains that satisfy a volume constraint, also regularity properties and a suitable topology ought to be considered in order to comply with assumptions (i)–(iii) in the above existence theorem. The ε-cone property and the convergence in the sense of characteristic functions ensure that the compactness and continuity conditions required in the proof of Theorem 3.4 are satisfied in this case. We refer the reader to [20] for more details on this passage.

4. Numerical resolution method

This section addresses the numerical approximation of the optimization problems presented in the preceding sections. The numerical resolution of problem (3.2), and similarly problem (2.18), entails a double optimization process. Indeed, for a fixed $u \in \mathcal{U}_{ad}$ the denominator in $\mathcal{J}^\star(u)$ is obtained by solving the maximization problem (3.1) in the parameter space Γ. Then, the maximization of $\mathcal{J}^\star(u)$ is performed with respect to $u \in \mathcal{U}_{ad}$.

We advocate for a gradient-based descent method. A key ingredient is to have at one disposal an accurate computation of the involved gradients. Although most of the commercial optimization routines use gradients which are approximated by finite differences methods, more accurate results are obtained when continuous gradients are supplied. Next section provides explicit computation of the required gradients for the two problems considered along the paper: the tracking-type optimal control problem of Section 2.3.1, and the topology optimization problem of Section 2.3.2. In order to fit in with the theoretical existence result presented in the preceding section, we focus on the case in which the mapping $z \to J(z,u)$ enjoys good convexity properties. This situation occurs, for instance, when uncertainty in the input parameters is restricted to the right-hand term of the state law and the random parameter appears in affine form. Two typical cases of practical interest are: (a) the control action is as in (2.3), with $b(x,z)$ a smooth random field with uncertain parameters appearing in an affine form, and (b) uncertainty enters in the loads also in affine form in the case of linearized elasticity.

Finally, we describe the numerical algorithm and provide the most relevant implementation details.

4.1. Computation of continuous gradients. Case of tracking-type control

As indicated above, throughout this section we assume that the differential operator is deterministic so that uncertainty only appears in the right-hand side term of the PDE.

4.1.1. Computation of the gradient of $J(z,u)$ with respect to $z \in \Gamma$.

Let $u \in \mathcal{U}_{ad}$ be as in (2.22). A formal computation shows that

$$\frac{\partial J}{\partial z_j}(z,u) = \int_D (y - y_T)\,\frac{\partial y}{\partial z_j}\,\mathrm{d}x, \quad 1 \le j \le N. \tag{4.1}$$

where $\frac{\partial y}{\partial z_j}$ solves

$$\int_D a\nabla \frac{\partial y}{\partial z_j}\cdot\nabla v\,\mathrm{d}x = \int_D \frac{\partial f}{\partial z_j} v\,\mathrm{d}x, \quad \forall v \in H_0^1(D) \quad \text{and } \rho - a.e. \text{ in } \Gamma. \tag{4.2}$$

4.1.2. Computation of the gradient of $\mathcal{J}^\star(u)$.

Since the function $(t)_+ = \max\{0,t\}$, $t \in \mathbb{R}$, is not differentiable, the smooth approximation is considered

$$(t)_+ \approx +\frac{1}{2}\left[\sqrt{t^2+\alpha}+t\right] := s_\alpha(t),$$

with $\alpha > 0$ a regularization parameter. Thus, $\mathcal{J}^\star(u)$ is approximated by

$$\tilde{\mathcal{J}}^\star(u) := \frac{s_\alpha\left(\eta - \mathbb{E}\left[J(z,u)\right]\right)}{NJ_{\max}(u)+\varepsilon}. \tag{4.3}$$

A straightforward application of the formal Lagrangian method [15], Section 4.2 and of the chain rule yields the gradient of $\tilde{\mathcal{J}}^\star(u)$:

$$\left(\tilde{\mathcal{J}}^\star\right)'(u) = \frac{-\left(s_\alpha'\left(\eta - \mathbb{E}\left[J(z,u)\right]\right)\right)\mathbb{E}'\left[J(z,u)\right]\left(NJ_{\max}(u)+\varepsilon\right) - s_\alpha\left(\eta-\mathbb{E}\left[J(z,u)\right]\right)NJ_{\max}'(u)}{\left(NJ_{\max}(u)+\varepsilon\right)^2}, \tag{4.4}$$

where

$$s_\alpha'\left(\eta - \mathbb{E}\left[J(z,u)\right]\right) = \frac{1}{2}\left(\frac{\eta - \mathbb{E}\left[J(z,u)\right]}{\sqrt{\left(\eta-\mathbb{E}\left[J(z,u)\right]\right)^2+\alpha}}+1\right), \tag{4.5}$$

$$\mathbb{E}'\left[J(z,u)\right] = \int_\Gamma b(x,z)p(x,z)\,\rho(z)\,\mathrm{d}z, \tag{4.6}$$

$p = p(x,z)$ being a solution to the adjoint system

$$\begin{cases} -\mathrm{div}\,(a\nabla p) = y - y_T, & \text{in } D\times\Gamma \\ p = 0, & \text{on } \partial D\times\Gamma, \end{cases} \tag{4.7}$$

and

$$J'_{\max}(u) = b(z_{\max})p_{\max}, \tag{4.8}$$

with $p_{\max} = p_{\max}(x)$ a solution to

$$\begin{cases} -\mathrm{div}\,(a\nabla p_{\max}) = y_{z_{\max}} - y_T, & \text{in } D \\ p_{\max} = 0, & \text{on } \partial D, \end{cases} \tag{4.9}$$

$y_{z_{\max}} = y_{z_{\max}}(x)$ being a solution to (2.2) for $z = z_{\max}$.

Remark 4.1. For the case of problem (2.18), in addition to solving (3.1) it is necessary to minimize $J(z,u)$ with respect to $z\in\Gamma$ for each fixed $u\in\mathcal{U}_{ad}$. Thus, denoting by

$$z_{\min}(u) = \arg\min_{z\in\Gamma} J(z,u) \quad \text{and} \quad J_{\min}(u) = J(z_{\min},u), \tag{4.10}$$

the term $J_{\max}(u)$ in (4.3) and in (4.4) has to be replaced by $J_{\max}(u) - J_{\min}(u)$, and $J'_{\max}(u)$ in (4.4) by $J'_{\max}(u) - J'_{\min}(u)$, where $J'_{\min}(u) = b(z_{\min})p_{\min}$, with $p_{\min} = p_{\min}(x)$ a solution to

$$\begin{cases} -\mathrm{div}\,(a\nabla p_{\min}) = y_{z_{\min}} - y_T, & \text{in } D \\ p_{\min} = 0, & \text{on } \partial D, \end{cases} \tag{4.11}$$

$y_{z_{\min}} = y_{z_{\min}}(x)$ being a solution to (2.2) for $z = z_{\min}$.

4.2. Computation of continuous gradients. Case of structural topology optimization

For the case of structural topology optimization we advocate for a density-based approach [2]. In this case, it is customary to introduce a pseudo-density field $u : D\to[0,1]$, where the extreme cases $u=0$ and $u=1$ represent void and solid regions, respectively. An energy interpolation scheme is used in order to interpolate the Lamé parameters $\{\lambda,\mu\}$ according to the classical Solid Isotropic Material with Penalization method (SIMP) [2, 3], *i.e.*

$$\lambda = u^p\lambda + \varepsilon\lambda(1-u^p), \quad \mu = u^p\mu + \varepsilon\mu(1-u^p), \tag{4.12}$$

where the penalizing exponent typically adopts a value of $p=3$ and where the parameter $\varepsilon\approx 10^{-7}$ is used in order to ensure an admissible condition number of the stiffness matrix associated with the underlying Finite Element discretization of (2.9). We refer the reader to [2] for a complete account of the SIMP method.

Remark 4.2. With the aim of preventing the ill-posedness of the optimization problem, a filtered pseudo-density field $\widetilde{u} : D\to[0,1]$ is introduced. A well-known filtering technique is that associated with the solution of the boundary value problem

$$\begin{cases} \widetilde{u} - \dfrac{l^2}{2}\Delta^2\widetilde{u} = u, & \text{in } D \\ \nabla\widetilde{u}\cdot\boldsymbol{n} = 0, & \text{on } \partial D_N \end{cases} \tag{4.13}$$

where $l \in \mathbb{R}^+$ represents the length scale of the filter and $\boldsymbol{n}$ denotes the unit outward normal vector.

Next, we focus on the computation of $\mathbb{E}'[J(z,u)]$, the rest of comutations being similar to the previous case. To this end, we consider the Lagrangian

$$\mathcal{L}(u,y,v) = \mathbb{E}[J(z,u)] - \Pi(u,y,v), \tag{4.14}$$

with $\Pi(u,y,v)$ defined as

$$\Pi(u,y,v) = \int_\Gamma \int_D A(u)e(y(x,z)) : e(v(x,z))\,\mathrm{d}x\,\rho(z)\,\mathrm{d}z - \underbrace{\int_\Gamma \left[\int_D f\cdot v\,\mathrm{d}x + \int_{\partial D_N} g\cdot v\,ds\right]\rho(z)\,\mathrm{d}z}_{\mathbb{E}[J(v)]}, \tag{4.15}$$

for all $v \in L^2_\rho(\Gamma; V)$. A standard computation yields

$$\mathbb{E}'[J(z,u)] = -\int_\Gamma \int_D A'(u)e(y(x,z)) : e(p(x,z))\;\mathrm{d}x\rho(z)\mathrm{d}z \tag{4.16}$$

where $p(x,z) = y(x,z)$, given the self-adjoint nature of the Lagrangian in (4.14)-(4.15). From equations (2.8) and (4.12), the term $A'(u)$ is obtained as

$$A'(u) = \mu'(u)\left(\delta_{ikjl} + \delta_{iljk}\right) + \lambda'(u)\delta_{ij}\delta_{kl} \tag{4.17}$$

with

$$\mu'(u) = p\mu u^{p-1}\left(1-\varepsilon\right); \qquad \lambda'(u) = p\lambda u^{p-1}\left(1-\varepsilon\right). \tag{4.18}$$

Finally, a straigtforward computation leads to

$$J'_{\max}(u) = -\int_D A'(u)e(y(x,z_{\max})) : e(y(x,z_{\max})\;\mathrm{d}x, \tag{4.19}$$

4.3. Numerical algorithm

Algorithm 1 summarizes the main steps for solving problem (3.2) for the case of the optimal control problem of Subsection 2.3.1, the case of topology optimization being completely analogous.

The following two remarks concerning this algorithm are in order:

Remark 4.3. The resolution of the maximization problem (3.1) is straightforward in the the case in which the parameter space Γ is a rectangle in $\mathbb{R}^N$ and the mapping $z \mapsto J(z,u)$ is convex and continuous for each $u \in \mathcal{U}_{ad}$. Indeed, in such a situation the maximum is attained at a vertex of the rectangle [24], Corollary 32.3.1. The above-mentioned conditions assumed on Γ and $J(z,u)$ are satisfied in many situations of practical interest (see, *e.g.* [15]. In particular, the two benchmark examples considered in Section 5 comply with those hypotheses.

Remark 4.4. The numerical algorithm 1 is a gradient-based descent one in the sense that: (a) the computation and use of the continuous gradient is crucial to determine a descent direction, and (b) it is not a one-shot method, but an iteration-based one. For the numerical experiments of Section 5 we rely on cutting-edge and well-established numerical optimization algorithms, precisely an interior-point method (as implemented in Matlab's routine `fmincon`) for the first experiment, and the Method of Moving Asymptotes (MMA) in the second one. These two methods follow the general scheme of our algorithm. Providing those with continuous gradients is a cornerstone of the success of both methods. Additionally, and for the sake of self-contained and completeness,

Algorithm 1: Numerical approximation of problem (3.2) by a projected gradient algorithm.

Data: Initial guess $u^0 \in \mathcal{U}_{ad}$, threshold parameter η, regularization parameters ε, α, and stopping tolerance δ
Result: optimal control $u^\star$ and optimal cost $\mathcal{J}^\star(u^\star)$
for $k \leftarrow 0$ **to** *convergence* **do**
  solve the maximization problem (3.1);
  Compute a descent direction $\overline{u}^k$ for $\tilde{\mathcal{J}}^\star(u^k)$ by using (4.4), (4.6) and (4.8);
  Choose a step size λ^k for which $\tilde{\mathcal{J}}^\star\left(P_{[m,M]}\left(u^k+\lambda^k\overline{u}^k\right)\right) < \tilde{\mathcal{J}}^\star(u^k)$, where $P_{[m,M]}(u) = \max(m, \min(u, M))$ is the projection operator onto $\mathcal{U}_{ad}$;
  Update the control $u^{k+1} = P_{[m,M]}\left(u^k+\lambda^k\overline{u}^k\right)$;
  if $\frac{|\tilde{\mathcal{J}}^\star(u^{k+1})-\tilde{\mathcal{J}}^\star(u^k)|}{\tilde{\mathcal{J}}^\star(u^{k+1})} \leq \delta$ **then**
    break ;

the numerical experiment 1 has been conducted using a simple projected gradient method (exactly the same as described in Algorithm 1) leading to the same solution as in the case of the interior-point method. In the projected gradient method, the selection of λ^k is done using the following straightforward strategy: in a first step, an acceptable step length (acceptable meaning that the cost functional decreases) is found by using the bisection method. A second stage is performed to improve this initial step length. Precisely, the acceptable step is increased by multiplying it by a scale factor while the cost functional is reduced with respect to the previous iteration. If the cost function is not decreased after a prescribed number of steps, the algorithm terminates. We refer the reader to [27], p. 95 and to the computer codes accompaning to [15], Chapter 5, where this strategy is implemented. Alternatively, more efficient line-search methods (*e.g.* Armijo's rule) may be used, yet this is not a concern in this paper. As a matter of fact, in the experiment 1 we observe that the objective function decreases at each iterates and that the algorithm stops at a stationay point. Notice that an explicit solution is known in this case.

An adaptive, anisotropic, sparse grid algorithm is used for numerical integration in the random domain, which is needed for the computation of $\mathbb{E}[J(u)]$ and its sensitivity (4.6) (see [15], Chapter 4 for details).

5. Numerical experiments

This section presents numerical simulation results for the two benchmark optimization problems considered in this paper. Experiment 1 focuses on the case of tracking-type optimal control and Experiment 2 is concerned with structural topology optimization.

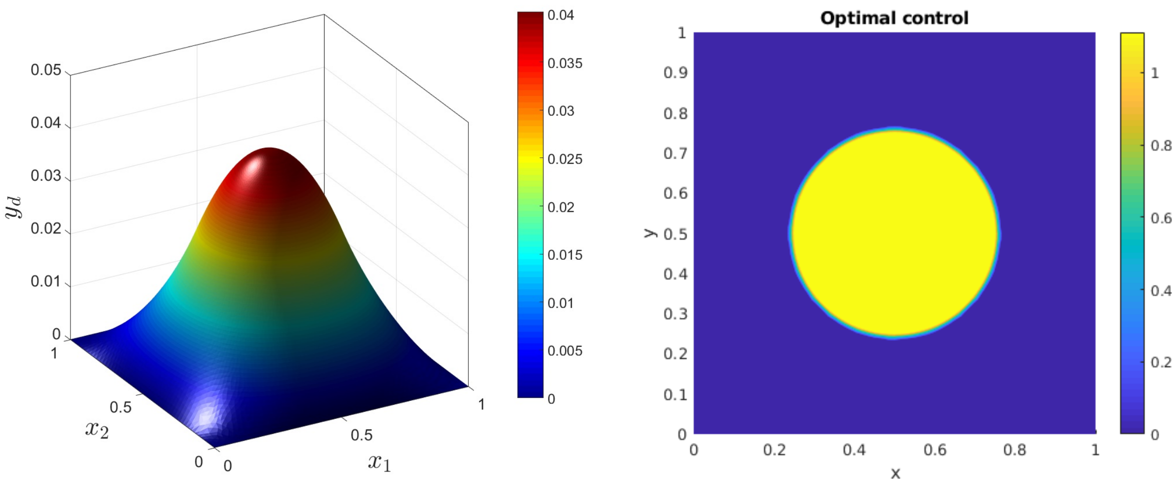
The overall objective is to study the numerical performance of the proposed approach. Precisely, we will compare numerical simulation results for both problems (2.18) and (3.2) and will check the effectiveness of McDiarmid's bound. Comparison with variance-based bounds is also considered. The computational cost of the proposed algorithm will be discussed as well.

5.1. Numerical example 1: tracking-type optimal control

Configuration of the experiment. In this experiment, the physical domain $D = (0,1)^2$ is the unit square and the control region $\mathcal{O}$ is a circle centered at the point $(0.5, 0.5)$ and of radius $r = 0.25$.

The considered target function y_T is the solution of

$$\begin{cases} -\mathrm{div}(\nabla y) = 1_{\mathcal{O}} u, & \text{in } D, \\ y = 0, & \text{on } \partial D, \end{cases} \tag{5.1}$$



(a) (b)

FIGURE 1. Experiment 1. (a) Target function y_T. (b) Optimal control for problem (2.18) and for $c = 0.1$.

with $u(x) = 1$, $x \in \mathcal{O}$. Figure 1a plots the target y_T. No constraints are imposed in the control variable.

The diffusion coefficient a is taken as constant and equal to 1. Uncertainty is incorporated in the controller device through a random variable which is uniformly distributed in the interval $[0.9 - c, 0.9 + c]$. Precisely, in (2.3) we take $b(x,z) = z 1_{\mathcal{O}}(x)$, where $z \sim \mathcal{U}(0.9 - c, 0.9 + c)$.

The cost functional $J(z,u)$ considered in this experiment is the one given by (2.21), with y_T the target function just indicated. Then, both the cost $\mathcal{J}^{\#}(u)$ and $\mathcal{J}^{\star}(u)$ are considered, as given by (2.18) and (3.2), respectively.

The input data for the gradient-based algorithms are taken as:

$$u^0(x) = 1, \quad x \in \mathcal{O}, \quad \varepsilon = \alpha = \delta = 10^{-10},$$

where δ denotes the tolerance for the stopping criterion included in `fmincon` function of Matlab.

Implementation's details. Numerical approximation in the physical domain is performed by using Lagrange $P1$-finite elements on a mesh composed of 1829 DOF. A quadrature rule composed of 100 Gauss-Legendre nodes is used for numerical approximation of $\mathbb{E}[J(z,u)]$.

Simulation's results. Due to the linear dependence of the solution of problem (5.1) with respect to the control variable it is clear that, for the above choice of the initial guess u^0, $J(z,u^0) = (1-z)^2 \|y_T\|_{L^2}^2$. Moreover, due to the choice of the random parameter z, a straightforward computation leads to

$$\mathbb{P}\left(J\left(u^0\right) \geq \eta\right) = 1 - \frac{1}{2c} \int_{1 - \frac{\sqrt{2\eta}}{\|y_T\|_{L^2}}}^{0.9+c} \mathrm{d}z.$$

Since $\|y_T\|_{L^2} = 0.018$, for $c = 0.05$ and $\eta = 2.2 \times 10^{-6}$ one gets $\mathbb{P}\left(J\left(u^0\right) \geq \eta\right) = 0.335$, which is very similar to the value in the first entry in Table 1, the latter being computed by using Monte Carlo integration with 10^5 random samples. Notice that Monte Carlo is the usual way of procceding when explicit computations of the probability cannot be carried out, as would be the case in most of the situations of practical interest. Table 1 collects simulation results for problems (2.18) and (3.2) and for two different values of the parameter c, which

TABLE 1. Experiment 1: $\eta = 2.2 \times 10^{-6}$. Summary of results for the estimates (2.20) and (3.3) for different values of the uncertain parameter. Here, u^0 denotes the initial guess, $u^{\#}$ is the optimal control of the functional $\mathcal{J}^{\#}$, and $u^{\star}$ the optimal control for $\mathcal{J}^{\star}$.

	$\mathbb{P}\left(J\left(u^0\right) \geq \eta\right)$	$\mathbb{P}\left(J\left(u^{\star}\right) \geq \eta\right)$	$\mathrm{e}^{-2\mathcal{J}^{\#}(u^{\#})^2}$	$\mathrm{e}^{-2\mathcal{J}^{\star}(u^{\star})^2}$
$c = 0.05$	0.33701	0	3.831×10^{-15}	4.506×10^{-15}
$c = 0.1$	0.41777	0.01443	0.17843	0.17859

accounts for the support of the random parameter z. (2nd column). The following conclusions are derived from the results of Table 1:

- *Comparison of problems* (2.18) *and* (3.2). On the one hand, it is observed that both problems (2.18) and (3.2) lead to essentially the same results. This is due to the fact that at each iteration k the costs $J_{\max}(u^k)$ and $J_{\min}(u^k)$ range at different scales so that $J_{\min}(u^k)$ does not play a role during the optimization process. On the other hand, for $c = 0.05$ the probability of failure for the control $u^{\star}$ equals zero. This is due to the fact that, for these specific values of the involved parameters, the optimization algorithm is able to shift the probability density function of $J(z,u)$ to the left of the threshold η. As shown in the first row and the third and forth columns of Table 1, problems (2.18) and (3.2) essentially do the same job.
- *Computational cost.* It is well-known [1] that the mapping $z \mapsto J(z,u)$ is analytic. As a consequence, the rate of convergence is exponential for the Gauss-Legendre quadrature rule used for numerical approximation in the random domain. In addition, due to the strict convexity of $z \mapsto J(z,u)$, problems (3.1) and (4.10) are solved very efficiently. In fact, for each u^k, a very few number of iterations (2 or 3 depending on the iteration for updating the control) is needed to compute the associated $z_{\max}(u^k)$ which, in turn, is a clear numerical indication that the mapping $u \mapsto z_{\max}(u)$ is continuous. Concerning the number of iterations needed to compute the optimal $u^{\star}$, it is equal to 9 for the case $c = 0.05$ and 11 for $c = 0.1$. All in all, the computational cost required for solving (2.18) and (3.2) is extremely low compared to Monte Carlo method, where rate of convergence is of the order of $\mathrm{O}(\frac{1}{\sqrt{m}})$, with m the number of samples.

The same explicit computations that we made before for the initial guess u^0 may be carried out for a constant control $u_{\text{exact}}(x) = 1.1$, $x \in \mathcal{O}$. In this case, we get that $\mathbb{P}\left(J\left(u_{\text{exact}}\right) \geq \eta\right) = 0$, which shows that u_{exact} is an optimal control for this problem. Except for numerical errors in the boundary of the control region $\mathcal{O}$, the numerical algorithm is able to capture this control (see Fig. 1b).

Remark 5.1. For the same data as before we have carried out a numerical experiment by imposing the contraints $m = 0.5$, $M = 1$ on the control variable. In this case, the numerical algorithm provides the constant control $u^*(x) = 1$, $x \in \mathcal{O}$, which is the expected one taking into acount the results of the unconstrained case.

5.2. Numerical example 2: structural topology optimization

In this numerical example we study a typical benchmark used in the context of stochastic topology optimization. As an illustration, we shall consider the case of a single surface load. The mechanical properties of the solid are deterministic and given by

$$\mu = 1\,\mathrm{Pa}; \qquad \lambda = \frac{4}{3}\,\mathrm{Pa}. \tag{5.2}$$

The experimental set up of this example is observed in Figure 2, where the solid structure is subjected to a surface load with magnitude 2.5×10^{-4} and Gaussian variation in the angle of application.

The underlying Finite Element discretization is comprised of 100 divisions along the X axis and 200 along the Y axis. Bilinear interpolation of the displacement field has been considered (*i.e.* $Q1$ interpolation), yielding a total of 40602 DOF for the displacement field. Using the anisotropic sparse grid configuration, each topology

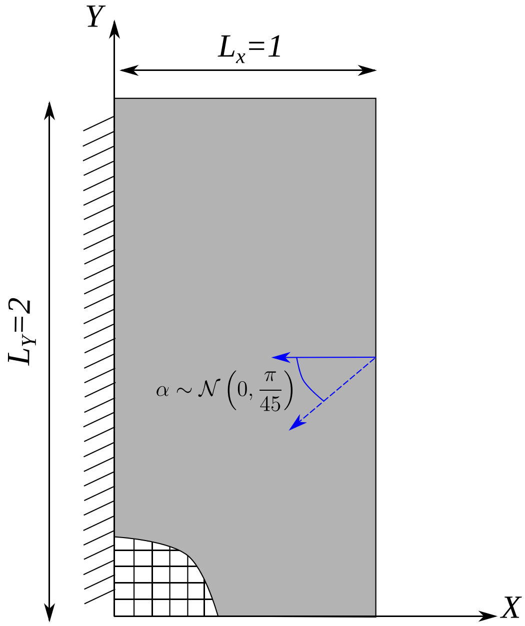


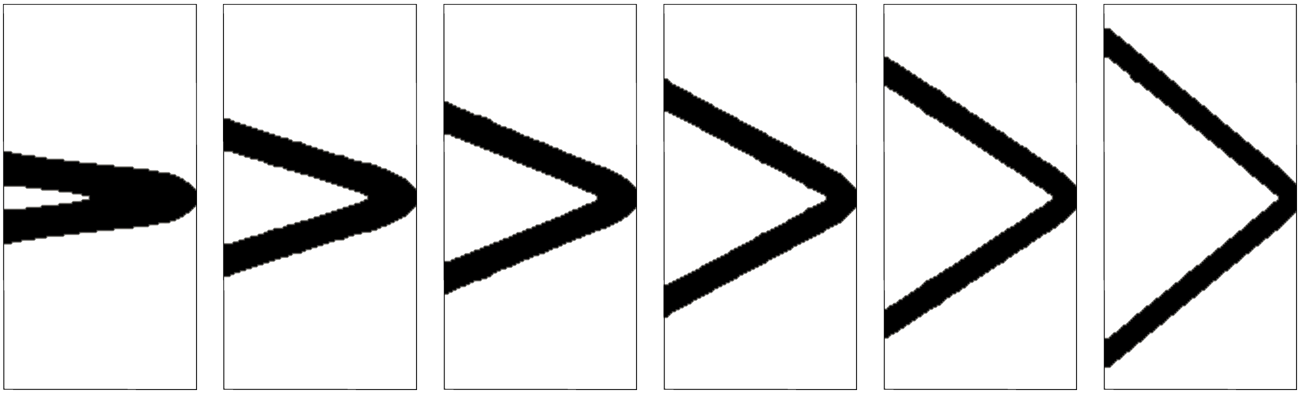
FIGURE 2. Experiment 2: problem configuration.

optimization iteration solves 4 different occurrences of the state equation, which are solved making use of parallel computing on a cluster.

Figure 3 shows the different optimal topologies obtained for the different cases considered. In the first case (*i.e.* Figure 3a), the design corresponds with the minimization of the expected value of the compliance. The remaining cases correspond with the maximization of the functional $\mathcal{J}^\star$ in (3.2) for increasing values of η. It can be observed that, in order to decrease the probability for the compliance to exceed an ever increasing threshold value η, the topologies of the optimal designs gradually decrease the thickness of the two structural bars, whilst increasing the angle of separation between them. This result is indeed in agreement with those reported in [20], where the excess probability was considered as a metric for structural robustness.

Crucially, Table 2 shows the values of the objective function $\mathcal{J}^\star$ in problem (3.2) for the six optimal designs in Figure 3 and for different values of $\frac{\eta}{\eta^\star}$. In all the cases, $\eta^\star$ corresponds with the value of $\mathbb{E}[J(z,u)]$ for the design in 3a. Clearly, for every threshold (see every row in the table), the maximum value of $\mathcal{J}^\star$ corresponds with the design that was obtained precisely for that specific threshold. This can be observed from the fact that for every row, the maximum value of $\mathcal{J}^\star$ is obtained always in the diagonal elements of the inner 6×6 sub-table within Table 2.

Figure 4 sheds some interesting additional information regarding the designs obtained in Figure 3. Specifically, Figure 4a represents the visual counterpart of Table 2, as it shows the McDiarmid bound, namely $\mathrm{e}^{-2(\mathcal{J}^\star)^2}$, for each design in Figure 3 and for different values of the threshold η. Finally, Figure 4b shows, for the design in Figure 3d, the comparison between the McDiarmid bound and the classical variance bound. It can be seen that below an approximate value of -1.34, the McDiarmid bound yields a smaller value with respect to that of the variance bound $\mathbb{P}\left(J \geq \eta\right) \leq \frac{\mathrm{Var}\ J}{(\eta-\mathbb{E}[J])^2_+}$. This indicates that for values of probability of the order of $10^{-1.34} \approx 0.046$ (*i.e.* $\mathbb{P}(J \geq \eta) < 0.046$), the McDiarmid bound yields a tighter bound than its variance-based counterpart. On the other hand, for not so strict probability values (above 4.5%), the variance-based bound seems to yield a tighter bound than the McDiarmid bound. This finding is corroborated by comparing both bounds with respect to $\mathbb{P}(J \geq \eta)$, computed by means of a Monte Carlo simulation with 10^4 random samples. This can be seen in



(a) (b) (c) (d) (e) (f)

FIGURE 3. Experiment 2: different optimal topologies for: (a) minimization of expected value of the compliance; Maximization of $\mathcal{J}^\star$ in (3.2) for (b) $\frac{\eta}{\eta^\star} = 1.2$; (c) $\frac{\eta}{\eta^\star} = 1.4$; (d) $\frac{\eta}{\eta^\star} = 1.8$; (e) $\frac{\eta}{\eta^\star} = 2.2$; (f) $\frac{\eta}{\eta^\star} = 2.8$. In all the cases, $\eta^\star$ corresponds with the value of $\mathbb{E}[J(z,u)]$ for the design in (a).

TABLE 2. Experiment 2: value of the objective function $\mathcal{J}^\star$ in (3.2) for the six optimal designs in Figure 3 and for different values of $\frac{\eta}{\eta^\star}$.

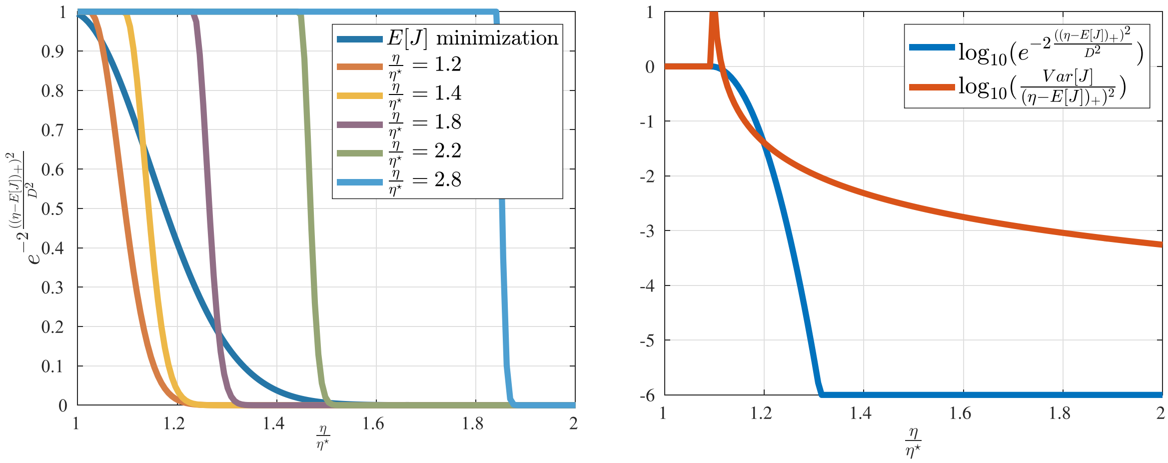
Summary of results for different cases						
	Case/Design					
	$\mathbb{E}[J]$ minimization	$\frac{\eta}{\eta^\star} = 1.2$	$\frac{\eta}{\eta^\star} = 1.4$	$\frac{\eta}{\eta^\star} = 1.8$	$\frac{\eta}{\eta^\star} = 2.2$	$\frac{\eta}{\eta^\star} = 2.8$
$\mathcal{J}^\star$ for $\frac{\eta}{\eta^\star} = 0.9$	**0.05**	0	0	0	0	0
$\mathcal{J}^\star$ for $\frac{\eta}{\eta^\star} = 1.2$	0.67	**1.48**	1.31	0	0	0
$\mathcal{J}^\star$ for $\frac{\eta}{\eta^\star} = 1.4$	1.28	3.11	**3.67**	3.02	0	0
$\mathcal{J}^\star$ for $\frac{\eta}{\eta^\star} = 1.8$	2.51	6.42	8.47	**10.23**	9.54	0
$\mathcal{J}^\star$ for $\frac{\eta}{\eta^\star} = 2.2$	3.73	9.74	13.26	17.44	**20.35**	18.78
$\mathcal{J}^\star$ for $\frac{\eta}{\eta^\star} = 2.8$	5.58	14.73	20.51	28.33	36.66	**50.52**

Table 3. It is observed that the value of $\mathbb{P}(J \geq \eta)$ is the smallest of the three, and that for low probability values, the McDiarmid bound represents the tightest approximation to the actual value of the probability.

6. CONCLUSIONS

The lack of smoothness and the presence of large flat regions where gradients vanish are very serious difficulties for the numerical approximation of control and optimization problems which involve probabilities of failure in their formulation. As a consequence, Monte Carlo (or variants) is the numerical method commonly used in this type of problems. However, its slow convergence makes Monte Carlo method useless in situations where a random realization of the objective function is expensive.

In this paper, an alternative approach has been proposed to approximate numerically optimal control problems where the objective function concerns the probability of exceeding a given threshold parameter and the state law is a partial differential equation with random input data. Making use of McDiarmid's inequality, the



(a) (b)

FIGURE 4. Experiment 2: (a) value of the McDiarmid bound for the six optimal designs in Figure 3 and for different values of $\frac{\eta}{\eta^\star}$. In all the cases, $\eta^\star$ corresponds with the value of $\mathbb{E}[J]$ for the design in 3a. (b) Comparison of the McDiarmid bound and the variance bound for different values of η for the design in Figure 3d.

TABLE 3. Experiment 2: value of $\frac{Var[J]}{((\eta-\mathbb{E}[J])_+)^2}$, $\mathrm{e}^{\frac{-2\left((\eta-\mathbb{E}[J])_+\right)^2}{(D(J))^2}}$ and $\mathbb{P}(J \geq \eta)$ the for the design in Figure 3d and for different values of $\frac{\eta}{\eta^\star}$.

Comparison between bounds and Monte Carlo-based probability for different η			
	$\frac{Var[J]}{((\eta-\mathbb{E}[J])_+)^2}$	$\mathrm{e}^{\frac{-2\left((\eta-\mathbb{E}[J])_+\right)^2}{(D(J))^2}}$	$\mathbb{P}(J \geq \eta)$
$\frac{\eta}{\eta^*} = 1.74 \times 10^{-1}$	3.4512×10^{-2}	4.33×10^{-2}	4.6×10^{-3}
$\frac{\eta}{\eta^*} = 1.77 \times 10^{-1}$	6.04×10^{-2}	4.85×10^{-3}	3.4×10^{-3}

original control problem is relaxed by considering the maximization of an analogous to the so-called confidence factor in the theory of quantification of margins and uncertainties. The new cost functional then provides a probabilistic certification criterion for the objective function of the original control problem.

At a theoretical level, well-posedness of a relaxed version of this optimization problem has been established in Theorem 3.4.

From a numerical viewpoint, the proposed formulation effectively circumvents the above mentioned difficulties associated to the approximation of probability-based objective functions and their gradients since the new functional essentially involves the computation of a statistical mean and the resolution of an optimization problem in the parameter space. Both problems may be solved very efficiently when the output of the state law is smooth and the system's variability is convex in the parameter space, as it occurs in a number of practical problems, *e.g.*, in structural optimization problems when the dominant uncertainty affects volumetric and/or surface loads [15], Chapter 6.

A comparison study between the proposed McDiarmind concentration-of-measure inequality and the classical variance-based bound has been carried out in Experiment 2. It is numerically observed that for small values of the probability of failure, McDiarmid yields a tighter bound than its variance counterpart. In order to reach an optimal tradeoff between accurate statistical modelling and computability, comparison studies with other approaches, *e.g.* with buffered probability of exceedance [9], would be in order. However, this study exceeds the scope of this paper and remains as a very interesting open problem.

The case of lack of convexity in the parameter space is not covered by the methology proposed in this paper. In this situation, the computation of the McDiarmid diameter may be overestimated as the algorithm may get stuck in local minima. This issue as well as a proper convergence analysis of the numerical algorithm in Section 4.3 are also very challenging open problems which remain beyond the scope of this paper.

Acknowledgements

Supported by grant PID2022-141957OA-C22 funded by MCIN/AEI/10.13039/501100011033 and by RDF A way of making Europe. The authors also acknowledge the support of the Autonomous Community of the Región of Murcia, Spain through the programme for the development of scientific and technical research by competitive groups (21996/PI/22), included in the Regional Program for the Promotion of Scientific and Technical Research of Fundación Séneca – Agencia de Ciencia y Tecnología de la Región de Murcia.

References

[1] I. Babuška, F. Nobile and R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Rev.* **52** (2010) 317–355.

[2] M.P. Bendsøe and O. Sigmund, Material interpolation schemes in topology Optimization. *Arch. Appl. Mech.* **69** (1999) 635–654.

[3] M.P. Bendsøe and O. Sigmund, Topology Optimization. Springer-Verlag, Berlin (2003).

[4] S. Boucheron, G. Lugosi and P. Massart, Concentration Inequalities: A Nonasymptotic Theory of Independence. Oxford University Press (2013).

[5] P. Chen, A. Quarteroni and G. Rozza, Stochastic optimal robin boundary control problems of advection-dominated elliptic equations. *SIAM J. Numer. Anal.* **51** (2013) 2700–2722.

[6] M. Dambrine, C. Dapogny and H. Harbrecht, Shape optimization for quadratic functionals and states with random right-hand sides. *SIAM J. Control Optim.* **53** (2015) 81–103.

[7] A. Gaivoronoski, Linearization methods for optimization of functionals which depend on probability measures, in Mathematical Programming Study28, Vol. 84 of *Stochastic Programming. Part II*, edited by A. Prékopa, R. J. Wets. The Mathematical Programming Society (1986) 157–181.

[8] V. Keshavarzzadeh, M. Meidani and A. Tortorelli, A gradient based design optimization under uncertainty via stochastic expansion methods. *Comput. Methods Appl. Mech. Eng.* **306** (2016) 47–76.

[9] D.P. Kouri and T.M. Surowiec, A primal-dual algorithm for risk minimization. *Math. Program. Ser. A* **1** (2022) 337–363.

[10] M. Lazar and E. Zuazua, Averaged control and observation of parameter-depending wave equations. *C. R. Mathematiques* **352** (2014) 497–502.

[11] M. Lazar and E. Zuazua, Greedy controllability of finite dimensional linear Systems. *Automatica* **74** (2016) 327–340.

[12] J. Lohéac and E. Zuazua, Averaged controllability for random evolution partial differential equations. *J. Diff. Equ.* **262** (2017) 1540–1574.

[13] L.J. Lucas, H. Owhadi and M. Ortiz, Rigorous verification, validation, uncertainty quantification and certification through concentration-of-measure inequalities. *Comput. Methods Appl. Mech. Eng.* **197** (2008) 4591–3609.

[14] A. Mafusalov and S. Uryasev, Buffered probability of exceedance: mathematical properties and optimization. *SIAM J. Optim.* **28** (2018) 1077–1103.

[15] J. Martínez-Frutos and F. Periago, Optimal Control of PDEs under Uncertainty. An introduction with application to optimal shape design., Springer Briefs in Mathematics. BCAM SpringerBriefs (2018).

[16] J. Martínez-Frutos, M. Kessler and F. Periago, Robust optimal shape design for an elliptic pde under uncertainty in its input data. *ESAIM: COCV* **21** (2015) 901–923.

[17] J. Martínez-Frutos, M. Kessler, A. Münch and F. Periago, Robust optimal Robin boundary control for the transient heat equation with random input Data. *Inter. J. Numer. Methods Eng.* **108** (2017) 116–135.

[18] F. Marín, J. Martínez-Frutos and F. Periago, Robust averaged control of vibrations for the Bernoulli–Euler beam equation. *J. Optim. Theory Appl.* **174** (2017) 428–454.

[19] F. Marín, J. Martínez-Frutos and F. Periago, A polynomial chaos approach to risk-averse piezo-electric control of random vibrations of beams. *Int. J. Numer. Methods Eng.* (2018) 1–18.

[20] J. Martínez-Frutos, D. Herrero-Pŕez, M. Kessler and F. Periago, Risk-averse structural topology optimization under random fields using stochastic expansion methods. *Comput. Methods Appl. Mech. Eng.* **330** (2018) 180–206.

[21] C. McDiarmid, On the method of bounded differences, in Surveys in Combinatorics (Norwich, 1989), London Math. Soc. Lecture Note Ser. 141. Cambridge University Press (1989) 148–188.

[22] C. McDiarmid, Centering sequences with bounded differences. *Combin. Probab. Comput.* **6** (1997) 79–86.

[23] A. Prékopa, Stochastic Programming, Vol. 324 of *Mathematics and its Applications*. Springer-Science+Business Media B.V. (1995).

[24] R.T. Rockafellar, Convex Analysis. Cambridge University Press (1970).

[25] E. Rosseel and G.N. Wells, Optimal control with stochastics PDE constraints and uncertain controls. *Comput. Methods Appl. Mech. Eng.* **213** (2012) 152–167.

[26] T. Szántai, Evaluation of a special multivariate gamma distribution function, in Mathematical Programming Study, Vol. 84 of *Stochastic Programming. Part I*, edited by A. Prékopa and R.J. Wets. The Mathematical Programming Society (1986) 1–17.

[27] F. Tröltzsch, Optimal Control of Partial Differential Equations, Graduate Studies in Mathematics. Vol. 112. American Mathematical Society (2010).

[28] L. Warnke, On the method of typical bounded differences. *Combin. Probab. Comput.* **25** (2016) 269–299.

[29] E. Zuazua, Averaged control. *Automatica* **50** (2014) 3077–3087.