DISTRIBUTIONALLY ROBUST SHAPE AND TOPOLOGY OPTIMIZATION: OVERVIEW AND DEPLOYMENT FOR GEOMETRIC UNCERTAINTIES

C. DAPOGNY[1], F. IUTZELER[2], J. PRANDO[3] AND B. THIBERT[3]

[1] *Sorbonne Université, Université Paris Cité, CNRS, Inria, Laboratoire Jacques-Louis Lions, LJLL, F-75005 Paris, France.*
[2] *Université de Toulouse, CNRS, UPS, F-31062 Toulouse, France.*
[3] *Univ. Grenoble Alpes, CNRS, Grenoble INP, LJK, 38000 Grenoble, France.*

Abstract. This article contributes to the development of shape optimization formulations that take into account the presence of uncertainty in the parameters of the underlying physical model. More precisely, it aims to apply the idea of distributional robustness in the context of the minimization of a cost function evaluating the performance of the shape of a mechanical structure. This concern is motivated by the fact that in practice, the law of the uncertain parameters itself is often unknown, making unefficient the usual practice of minimizing the expectation of the cost under a necessarily approximate law. One then wises to minimize the worst-case expected value of the cost over probability laws that are "close" to a nominal law, reconstructed from a few observed samples, which is the only available information. This idea is relatively recent in convex analysis and operations research, where the problems at stake typically enjoy a rich mathematical structure. It was adapted to the setting of optimal design in a series of recent works, where tractable versions of distributionally robust optimal design problems were proposed, leveraging recent findings in robust optimization and optimal transport. The first part of this article reviews these contributions. Its second part is devoted to a new development, about the adaptation of this framework to the treatment of geometric uncertainties: the optimized shape itself is subject to uncertain perturbations, instead of a set of independent parameters, which requires a suitable reformulation of the problem. Two numerical examples are discussed to assess the main features of the proposed methods.

Contents

1. Introduction

Shape optimization aims to find the best-performing design of a physical system. Unfortunately, this very process often makes the resulting design extremely sensitive to modeling errors and uncertain operating conditions – a general, well-documented fact known as the optimizer's curse [73].

Mathematically, shape optimization arises as the minimization of a cost function $\mathcal{C}(\Omega, \xi)$ depending on the shape Ω of a system under scrutiny, and on a collection ξ of physical parameters. For instance, in mechanical engineering, Ω typically represents a structure optimized with respect to its elastic compliance and the parameter ξ refers to the loads applied to Ω or to the Young's modulus of its constituent material. In fluid mechanics, Ω could be a duct conveying a fluid and ξ may represent its viscosity, density, or the incoming velocity profile; the cost $\mathcal{C}(\Omega, \xi)$ may then account for the energy dissipated within the pipe under the effect of viscous forces.

In realistic situations, the parameters ξ are rarely known with certainty and, consistent with the optimizer's curse, a design optimized under the assumption of an ideal value ξ^0 for ξ may have poor performance when its actual value deviates, even slightly, from ξ^0. A dramatic illustration of such scenario was proposed in [19], where the best microstructure of a material to withstand a perfectly vertical load is also the worst possible microstructure as soon as the actual load deviates from this ideal configuration. As shape and topology optimization methods are increasingly striving for deployability in industry, it is essential to account for uncertainties in these model parameters.

Under this impetus, the treatment of uncertainties in topology optimization has been extensively considered in the literature; we refer for instance to [68] for a general discussion about robust optimization in engineering and to [53] for a survey focusing on topology optimization. As reviewed briefly in Section 3 below, two broad classes of approaches are available to incorporate awareness of uncertainties about physical parameters ξ into the optimal design process, depending on the information at hand. When only a maximum bound about the difference between their actual and ideal values ξ and ξ^0 is available, it is natural to optimize the worst-case (i.e. the largest) value of the cost function $\mathcal{C}(\Omega, \xi)$ over all possible values of ξ. A richer context, described with the help of probability theory, features ξ as random vector with known law $\mathbb{P}_{\text{true}}$: one then optimizes the expectation of $\mathcal{C}(\Omega, \xi)$, or a probability of failure. Both paradigms suffer from major limitations; while worst-case approaches tend to be overly conservative, probabilistic formulations usually assume that the law $\mathbb{P}_{\text{true}}$ of the uncertain parameters is known exactly, which is rarely the case in practice: often, only an approximation $\mathbb{P}$ reconstructed from a few observed samples is available, e.g. as their empirical distribution.

The paradigm of distributional robustness was originally proposed in [67] as a principled means to overcome this conceptual flaw of probabilistic formulations. It consists in minimizing the worst value of the expectation of the cost, when the law $\mathbb{Q}$ for ξ belongs to an ambiguity set $\mathcal{A}$ of probability measures that are "close" to an available, approximate nominal law $\mathbb{P}$; typically, $\mathbb{P}$ is the aforementioned empirical distribution attached to a collection of observed samples. This idea has recently raised much enthusiasm in operations research and machine learning, as reviewed in e.g. [13, 44, 47]. The highly structured character of the problems featured in these fields – for instance, the convex dependence of the cost function with respect to the design variable, or its piecewise linear dependence on the parameter ξ – allows to derive rigorous guarantees about this approach, for instance in terms of generalization, i.e. of how the optimized variable remains efficient under conditions that are not present among the samples of $\mathbb{P}$. On the contrary, distributionally robust approaches have received little attention in mechanical engineering and even less in optimal design, where no such favorable structure can be expected. To the best of our knowledge, it was only very recently introduced in this field in [42], in a quite restricted setting though. Our subsequent contributions [25, 26] have proposed a general framework, applicable to realistic structures, based on recent findings in robust optimization and convex analysis.

The purpose of this article is twofold. In a first part, we review the recent implementation of distributional robustness in optimal design proposed in the works [25, 26]. For clarity, we focus on the specific framework

of geometric shape optimization, where the optimized design is an elastic structure, represented by a domain Ω in $\mathbb{R}^d$ ($d = 2$ or 3) and ξ is a physical parameter entering the mathematical formulation of the elasticity system, such as the applied loads or the material coefficients. Moreover, we only consider one particular type for the ambiguity set $\mathcal{A}$ accounting for the probability measures $\mathbb{Q}$ that are close to the nominal law $\mathbb{P}$ for ξ. It is based on the notion of Wasserstein distance from optimal transport theory, which faithfully captures the difference between the "geometry" of probability distributions while enjoying rigorous mathematical guarantees. Although this setting is quite restricted, our methodology is much more general: as described in [26], it can be applied to a wide variety of physical problems; it also extends to the framework of density-based topology optimization, and it can deal with ambiguity sets of different natures, e.g. based on the moments of the probability measures at play. The second part of this article is devoted to an original development: we adapt our distributionally robust shape optimization framework to the case where the uncertainty bears on the geometry of the optimized shape itself, symbolically $\xi = \Omega$. Such geometric uncertainties are a major concern in applications, since the shape of a structure is often altered due to manufacturing imperfections or operational wear. Their mathematical treatment significantly departs from the previous situations, that are predominantly considered in the literature, where ξ is a physical parameter independent of the optimized shape. The fact that it is now the optimized shape itself destroys the structure where this parameter pertains to a fixed, finite-dimensional vector space where sampling and differentiation are "simple". To circumvent this difficulty, we propose a linearization strategy, in the spirit of our previous work [3], where the cost function is replaced by a linearized counterpart with respect to smooth (Hadamard-type) perturbations of the geometry of the optimized shape.

The remainder of the present article is organized as follows. The next Section 2 introduces the physical setting of interest, as well as the shape optimization problem under scrutiny. It also introduces the necessary background about the theoretical and numerical treatments of such problems. Then, Section 3 introduces uncertainties into the considered shape optimization problem; the prevailing approaches to handle them are reviewed, and the idea of distributional robustness is sketched. Section 4 reviews our previous work about the distributionally robust treatment of uncertainty about the parameters of the physical model: after a short reminder of basic concepts of optimal transport theory, we present the main ingredient from convex analysis, making it possible to rewrite intricate min-max distributionally robust problems as single level minimization problems. Section 5 is then devoted to the novel distributionally robust treatment of a peculiar source of uncertainties, namely, uncertainties over the geometry of the optimized shape itself. Eventually, a conclusion and a few perspectives for future work are outlined in Section 6.

2. Shape optimization and its numerical implementation

This section presents the shape optimization framework of the article and its numerical implementation. After introducing the mathematical model of elastic structures in Section 2.1, we recall in Section 2.2 the classical Hadamard's boundary variation method, underlying the notion of shape derivative for a function depending on the domain. Eventually, Section 2.3 broaches the methods involved in our numerical resolution of shape optimization problems.

2.1. Shape optimization of elastic structures

This article takes place in the physical context of linearly elastic structures, see e.g. [37]. A shape Ω is a bounded Lipschitz domain in $\mathbb{R}^d$ ($d = 2$ or 3) accounting for a mechanical structure. As depicted in Fig. 1, its boundary is made of three disjoint open pieces:

$$\partial\Omega = \overline{\Gamma_D} \cup \overline{\Gamma_N} \cup \overline{\Gamma},$$

where

- The shape Ω is clamped on Γ_D;
- Surface loads $g \in L^2(\Gamma_N)^d$ are applied on the region Γ_N;
- The remaining part Γ is traction-free.

Both regions Γ_D and Γ_N are imposed by the context, and only the free boundary Γ is subject to optimization. In this setting, we consider shape optimization problems of the form:

$$\min_{\Omega} \; J(\Omega), \tag{2.1}$$

where $J(\Omega)$ is the objective function. For simplicity of the exposition, we omit constraints in this formulation, although their treatment does not incur additional subsequent difficulty, as exemplified by the numerical experiments of Sections 4.4 and 5.3.

In our applications, the criterion $J(\Omega)$ depends on Ω via the elastic displacement u_Ω, which is the $H^1(\Omega)^d$ solution to the linear elasticity system:

$$\text{(2.2)} \qquad \begin{cases} -\mathrm{div}(Ae(u_\Omega)) = 0 & \text{in } \Omega, \\ u_\Omega = 0 & \text{on } \Gamma_D, \\ Ae(u_\Omega)n = g & \text{on } \Gamma_N, \\ Ae(u_\Omega)n = 0 & \text{on } \Gamma, \end{cases}$$

where body forces are omitted for simplicity. Here, $n \equiv n_\Omega$ is the unit normal vector to $\partial\Omega$, pointing outward Ω; $e(u) = \frac{1}{2}(\nabla u + \nabla u^T)$ is the strain tensor associated to a displacement field $u : \Omega \to \mathbb{R}^d$, and A is the Hooke's tensor, characterizing the material behavior:

$$\text{For all } \ d \times d \text{ symmetric matrix } e, \quad Ae = 2\mu e + \lambda \mathrm{tr}(e)\mathrm{I},$$

where λ and μ are the Lamé parameters of the material.

2.2. **Hadamard's method of boundary variations**

The treatment of shape optimization problems of the form (2.1) usually rests on the derivative of $J(\Omega)$ with respect to the domain. This notion can be given different acceptions and in this work, we rely on the classical boundary variation method of Hadamard [7, 41, 56, 74]. Briefly, variations of a given shape Ω are considered under the form:

$$\Omega_\theta := (\mathrm{Id} + \theta)(\Omega), \ \text{ where } \theta \in W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d), \ \ ||\theta||_{W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)} < 1,$$

i.e. Ω_θ is obtained from Ω via perturbation of its points according to a "small" vector field θ, see Fig. 1 for an illustration.

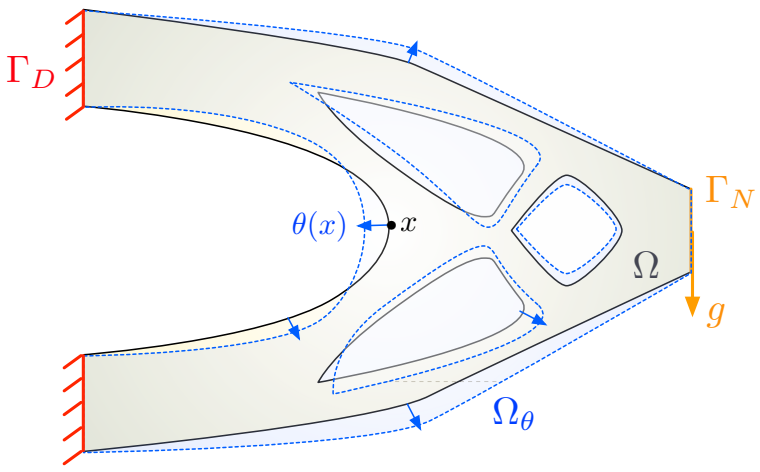

FIGURE 1. *Variation Ω_θ of a reference shape Ω in the framework of the boundary variation method of Hadamard presented in Section 2.1.*

A function $J(\Omega)$ of the domain is then called shape differentiable at Ω if the underlying mapping $\theta \mapsto J(\Omega_\theta)$, defined from a neighborhood of 0 in $W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)$ into $\mathbb{R}$, is Fréchet differentiable at $\theta = 0$. The corresponding derivative $J'(\Omega)(\theta)$, the shape derivative of $J(\Omega)$ at Ω, then satisfies the following expansion:

$$J(\Omega_\theta) = J(\Omega) + J'(\Omega)(\theta) + \mathrm{o}(\theta), \ \text{ where } \frac{\mathrm{o}(\theta)}{||\theta||_{W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)}} \xrightarrow{\theta \to 0} 0.$$

Crucially, the knowledge of the shape derivative allows to identify a descent direction for $J(\Omega)$ i.e. a vector field θ such that $J'(\Omega)(\theta) < 0$, thus ensuring that:

$$\text{For "small" } \tau > 0, \quad J(\Omega_{\tau\theta}) \approx J(\Omega) + \tau J'(\Omega)(\theta) < J(\Omega),$$

i.e. the perturbed shape $\Omega_{\tau\theta}$ has better performance than Ω. The calculation of shape derivatives, based on the adjoint method [48, 61], and the identification of descent directions θ with the help of the so-called "Hilbertian trick" [15, 27] are non trivial, but classical issues in shape optimization. For brevity, we simply point out that these tasks require the solution of one or several boundary value problems posed on Ω, including that (2.2) for the elastic displacement u_Ω, and refer to the above mentioned works for more details.

2.3. Numerical solution of the shape optimization problem

The mathematical framework recalled in Section 2.2 paves the way to a steepest-descent strategy for the numerical minimization of $J(\Omega)$. Starting from an initial guess Ω^0, a series Ω^n of shapes is produced through iterations of the form:

$$\Omega^{n+1} = (\mathrm{Id} + \tau^n \theta^n)(\Omega^n), \text{ where } \begin{cases} \theta^n \text{ is a descent direction for } J \text{ from } \Omega^n; \\ \tau^n \text{ is a "small enough" pseudo-time step.} \end{cases} \tag{2.3}$$

The numerical implementation of this strategy raises the need for a numerical representation of the shape Ω which lends itself to accurate mechanical computations (e.g. via the Finite Element Method), which leaving the room for a robust capture of the updates (2.3). Here, we rely on the numerical method introduced in our previous works [4, 5, 6], see also [23] for an open-source implementation. In a nutshell, each shape $\Omega = \Omega^0, \dots$ arising in the course of the iterative optimization process is contained in a large, fixed computational domain D, where it is represented by two complementary means:

- (Meshed representation) Ω is meshed exactly: a mesh $\mathcal{T}$ of the total domain D is available, which is valid and conforming, and Ω is accounted for by a sub-mesh $\mathcal{T}_{\mathrm{int}}$ of $\mathcal{T}$, i.e. a subcollection of the elements of $\mathcal{T}$, see Fig. 2 (a).
- (Level set representation) Ω is described via the Level Set Method, see [59, 58, 71] and [8, 76] about its introduction in shape optimization. Precisely, Ω is the negative region of a scalar, "level set" function $\phi : D \to \mathbb{R}$:

$$\forall x \in D, \quad \begin{cases} \phi(x) < 0 & \text{if } \ x \in \Omega, \\ \phi(x) = 0 & \text{if } x \in \partial\Omega, \\ \phi(x) > 0 & \text{otherwise.} \end{cases}$$

  In practice, ϕ is discretized at the vertices of $\mathcal{T}$, see Fig. 2 (b).

The most suited of both representations is used whenever an operation is performed that involves Ω: the meshed representation allows for accurate mechanical computations, carried out by a non intrusive use of any numerical solver on the explicit mesh $\mathcal{T}_{\mathrm{int}}$ of Ω; in our implementation, we use the open-source environment `FreeFem` [40]. On the other hand, the update $\Omega \mapsto (\mathrm{Id} + \tau\theta)$ of Ω via a vector field $\theta : \mathbb{R}^d \to \mathbb{R}^d$ for a small pseudo-time step τ is conveniently realized from a level set representation $\phi : D \to \mathbb{R}$ for Ω. Indeed, it is enough to this end to solve the following advection equation on the total mesh $\mathcal{T}$ of D:

$$\begin{cases} \frac{\partial \psi}{\partial t}\ (t,x) + \theta(x) \cdot \nabla \psi(t,x) = 0 & \text{for } t \in (0,\tau), \ \ x \in D, \\ \psi(0,x) = \phi(x) & \text{for } x \in D, \end{cases}$$

and the new shape $\Omega_{\tau\theta}$ is accounted for by the level set function $\psi(\tau, \cdot)$. This operation is realized by the open-source library `advection` [14].

This strategy crucially hinges on efficient algorithms for passing from a meshed to a level set representation of a shape, and conversely. The computation of a level set function from a meshed representation of Ω is achieved by the Fast Marching Method [69, 70], which is implemented in the open-source library `mshdist` [24]. On the other hand, a meshed representation of Ω, as a sub-mesh $\mathcal{T}_{\mathrm{int}}$ of the mesh $\mathcal{T}$ of D, is obtained from a level set representation $\phi : D \to \mathbb{R}$ thanks to the remeshing algorithm `mmg` [11, 22].

3. Modelling and classical approaches to uncertainties in shape optimization

In realistic contexts, some of the parameters of the physical model introduced in Section 2.1 are uncertain. For instance, the loads $g : \Gamma_N \to \mathbb{R}^d$ applied to the shape Ω are typically subject to measurement errors. Another manifestation of this phenomenon concerns the Lamé coefficients (λ, μ) (or the Young's modulus) of the material constituting the shapes, which are often altered through time.

Let ξ denote the uncertain physical parameters under scrutiny; these are assumed to belong to a fixed subset Ξ of a finite-dimensional space $\mathbb{R}^k$, that we assume to be compact. We also introduce a cost function

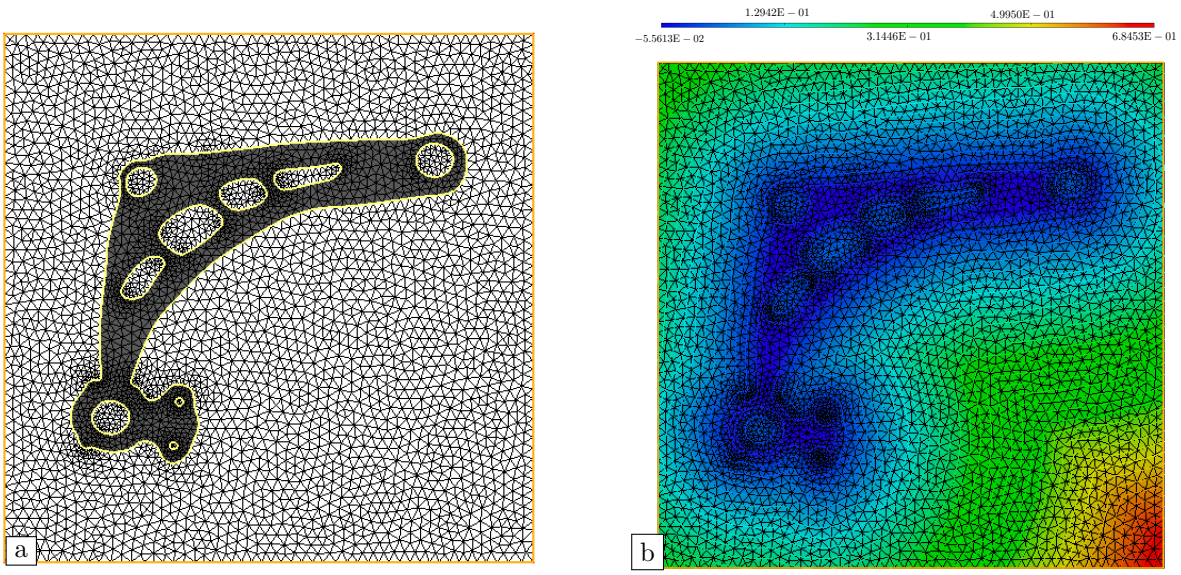

FIGURE 2. *Two complementary numerical representations of the optimized shape Ω involved in the shape optimization workflow of Section 2.3; (a) Meshed representation; (b) Level set representation.*

$\mathcal{C}(\Omega, \xi)$ evaluating the performance of the optimized shape Ω; the latter plays the role of the objective function $J(\Omega)$ in (2.1) but its refers explicitly to the uncertain parameters ξ. Then, in ideal conditions, where ξ takes the known value ξ^0, the optimization problem reads:

$$\min_{\Omega} J_{\text{det}}(\Omega), \text{ where } J_{\text{det}}(\Omega) := \mathcal{C}(\Omega, \xi^0). \tag{3.1}$$

In the next Sections 3.1 and 3.2, we describe the two most common paradigms used to take into account uncertainties about the parameters ξ in the treatment of such a shape optimization problem, namely the worst-case and probabilistic approaches. These formulations pave the way to the idea of distributional robustness, that we introduce in Section 3.3.

Remark 3.1. *We have deliberately omitted one important source of uncertainties in applications, namely uncertainties about the geometry of the optimized variable Ω itself. This quite specific situation will be considered in the next Section 5.*

Remark 3.2. *The assumption that ξ should belong to a finite-dimensional space may seem restrictive at first glance. In applications, uncertainties often show up as a random field v, i.e. for each point $x \in D$ of the hold-all domain, $v(x, \cdot)$ is a random variable. Classical model order reduction techniques, notably the Karhunen-Loève expansion, allow to approximate such objects with a finite-dimensional structure:*

$$v(x, \cdot) \approx \sum_{i=1}^{N} v_i(x)\xi_i(\cdot),$$

where the v_i are deterministic functions on D and the ξ_i are uncorrelated random variables. We refer to Section 5.3 for an example of this practice.

3.1. The worst-case approach

Let us first discuss the situation where no information is available about the uncertain parameters ξ, except for a maximum bound m on their deviation to a known "ideal" value $\xi^0 \in \Xi$. In this context, one is naturally led to minimize the worst (i.e. largest) value of the cost $\mathcal{C}(\Omega, \xi)$ over all possible perturbations. This worst-case design version of the original problem (2.1) then reads:

$$\min_{\Omega} J_{\text{wc}}(\Omega), \text{ where } J_{\text{wc}}(\Omega) := \sup_{|\xi - \xi^0| \leq m} \mathcal{C}(\Omega, \xi). \tag{3.2}$$

In rare, very particular situations, this bi-level optimization problem can be given a tractable reformulation. Notably when ξ stands for the applied loads and the cost function $\mathcal{C}(\Omega, \xi)$ is the compliance of Ω, the worst-case function $J_{\rm wc}(\Omega)$ can be rewritten as an eigenvalue of a boundary-value problem, see [19, 28]; we also refer to [9] for another approach to worst-case compliance minimization, based on the "stability radius", i.e. the largest load level such that the cost stay below a given safety threshold. Unfortunately, in general, the rigorous treatment of (3.2) resorts to costly bi-level programming techniques, see [12] for a general discussion. Let us also mention formal, approximate methods based on a "linearization" of the inner supremum, see [2, 34, 38] and [50] for an adaptation of this idea to express robustness to the emergence of porosity within the shape Ω.

Beyond the computational effort of their treatment, worst-case approaches are often reproached to be overly "pessimistic", insofar as they tend to predict optimized designs with poor performance in ideal conditions, for the sake of anticipating an unlikely worst-case scenario.

3.2. **The probabilistic approach**

We now assume that a little more information is available about the uncertain parameters ξ: they now show up as a random vector with law $\mathbb{P}_{\rm true}$ in the set $\mathcal{P}(\Xi)$ of probability measures on Ξ. This means that, for each measurable subset $A \subset \Xi$, the probability that ξ belongs to A equals:

$$\mathbb{P}_{\rm true}(A) = \int_A \mathrm{d}\mathbb{P}_{\rm true}(\xi).$$

In this setting, robust shape optimization approaches consist in minimizing probabilistic quantities of the cost $\mathcal{C}(\Omega, \xi)$, such as its mean value $J_{\rm mean}(\Omega)$ or its variance $J_{\rm var}(\Omega)$, defined by:

$$J_{\rm mean}(\Omega) = \int_\Xi \mathcal{C}(\Omega, \xi)\, \mathrm{d}\mathbb{P}_{\rm true}(\xi) \text{ and } J_{\rm var}(\Omega) = \int_\Xi \Big(\mathcal{C}(\Omega, \xi) - J_{\rm mean}(\Omega)\Big)^2\, \mathrm{d}\mathbb{P}_{\rm true}(\xi).$$

Alternatively, so-called reliability-based approaches are based on a probability of failure $J_{\rm fail}(\Omega)$ of the system, appraising the events where the cost of the shape Ω exceeds a safety threshold C_T:

$$J_{\rm fail}(\Omega) := \mathbb{P}_{\rm true}\Big\{\xi \in \Xi \text{ s.t. } \mathcal{C}(\Omega, \xi) > C_T\Big\}.$$

Central to the use of probabilistic approaches is the challenge to evaluate high-dimensional integrals. This task can be achieved e.g. by (costly) Monte-Carlo sampling or stochastic Galerkin methods [16, 18, 39, 51, 52]. Alternatively, linearization strategies can be thought of, based on the assumption that perturbations have "small" amplitude in a suitable sense, see e.g. [3]. We generally refer to [72] for a more detailed account of the probabilistic treatment of uncertainties in optimization.

As we have mentioned in the introduction, these approaches suffer from a major conceptual drawback: in practice, the law $\mathbb{P}_{\rm true}$ of ξ is not known exactly. This calls for taking into account this additional source of uncertainty in the probabilistic formulation of the problem (2.1), which is the starting acknowledgement of distributionally robust approaches.

3.3. **Towards distributionally robust optimization**

As we just mentioned, the true law $\mathbb{P}_{\rm true} \in \mathcal{P}(\Xi)$ of the uncertain parameters ξ at stake is often unknown in practice. At best, an estimate $\mathbb{P}$ of $\mathbb{P}_{\rm true}$ is available, which is often reconstructed as the empirical distribution of a few observed samples $\xi^1, \dots, \xi^N \in \Xi$, that is:

$$\mathbb{P} = \frac{1}{N} \sum_{i=1}^N \delta_{\xi^i}, \text{ where } \delta_{\xi^i} \in \mathcal{P}(\Xi) \text{ is the Dirac mass at } \xi^i.$$

Having the optimizer's curse in mind, one may expect that the use of a "classical" probabilistic formulation such as those described in Section 3.2, based on the assumption of the law $\mathbb{P}$ for ξ results in a design with poor performance in actual conditions.

To circumvent this drawback, the distributionally robust version of our model shape optimization problem (3.1) consists in minimizing the worst (i.e. largest) possible value $J_{\rm dr}(\Omega)$ of the expected cost when the actual

law $\mathbb{Q}$ of ξ is "close" to the nominal law $\mathbb{P}$, that is:

$$\min_{\Omega} J_{\mathrm{dr}}(\Omega), \text{ where } J_{\mathrm{dr}}(\Omega) := \sup_{\mathbb{Q}\in\mathcal{A}} \int_{\Xi} \mathcal{C}(\Omega,\xi)\,\mathrm{d}\mathbb{Q}(\xi). \tag{3.3}$$

This formulation features a so-called ambiguity set $\mathcal{A} \subset \mathcal{P}(\Xi)$, gathering those probability laws that are "close" to $\mathcal{P}$, whose definition may depend on the available information about the true law $\mathbb{P}_{\mathrm{true}}$.

This distributionally robust approach is more carefully detailed in the next Section 4.

4. Wasserstein distributionally robust shape optimization

This section gets into the mathematical details and the numerical treatment of distributionally robust shape optimization problems, of the form (3.3). At first glance, (3.3) is an intricate min-max problem, similar to the worst-case formulation (3.2). The key contribution of our previous works [25, 26] is to propose a physically meaningful approximation inspired from convex analysis, paving the way to a tractable, single-level reformulation of (3.3). Because of their rich mathematical structure and physical relevance, we focus on one particular type of ambiguity sets $\mathcal{A}$ in (3.3).

After providing in Section 4.1 the necessary background about optimal transport and the notion of Wasserstein distance, we formulate in Section 4.2 the distributionally robust problems considered in our applications. The next Section 4.3 details the reformulation of these difficult, bi-level min-max problems into single-level problems. A numerical example is eventually discussed in Section 4.4.

4.1. A brief reminder of optimal transport and the Wasserstein distance

This section broaches a few facts from optimal transport theory, referring to the books [20, 60, 66] for further details. We recall that Ξ is a compact subset of the finite-dimensional space $\mathbb{R}^k$, and that $\mathcal{P}(\Xi)$ stands for the set of probability measures on Ξ.

The Wasserstein distance between two elements $\mathbb{P}$ and $\mathbb{Q}$ in $\mathcal{P}(\Xi)$ is defined by:

$$W(\mathbb{P},\mathbb{Q}) = \inf_{\substack{\pi\in\mathcal{P}(\Xi\times\Xi)\\ \pi_1=\mathbb{P},\ \pi_2=\mathbb{Q}}} \int_{\Xi\times\Xi} c(\xi,\zeta)\,\mathrm{d}\pi(\xi,\zeta). \tag{4.1}$$

This expression features couplings, i.e. probability measures $\pi \in \mathcal{P}(\Xi\times\Xi)$, whose first and second marginals π_1 and $\pi_2 \in \mathcal{P}(\Xi)$ are defined by, for any continuous function $\varphi : \Xi \to \mathbb{R}$:

$$\int_{\Xi\times\Xi} \varphi(\xi)\,\mathrm{d}\pi(\xi,\zeta) = \int_{\Xi} \varphi(\xi)\,\mathrm{d}\pi_1(\xi), \text{ and } \int_{\Xi\times\Xi} \varphi(\zeta)\,\mathrm{d}\pi(\xi,\zeta) = \int_{\Xi} \varphi(\zeta)\,\mathrm{d}\pi_2(\zeta).$$

Intuitively, for all measurable subsets $A, B \subset \Xi$, $\pi(A\times B)$ is the amount of mass transferred from A to B, and $\pi_1(A)$ and $\pi_2(B)$ represent the mass removed from A and arriving at B in the process, respectively, see Fig. 3. The "ground cost" $c(\xi,\zeta)$ evaluates the effort of moving one unit of mass from ξ to ζ; for technical convenience, in this work, it is defined as the quadratic function $c(\xi,\zeta) := |\xi-\zeta|^2$. Hence, $W(\mathbb{P},\mathbb{Q})$ accounts for the minimum cost of transferring the mass of $\mathbb{P}$ onto $\mathbb{Q}$.

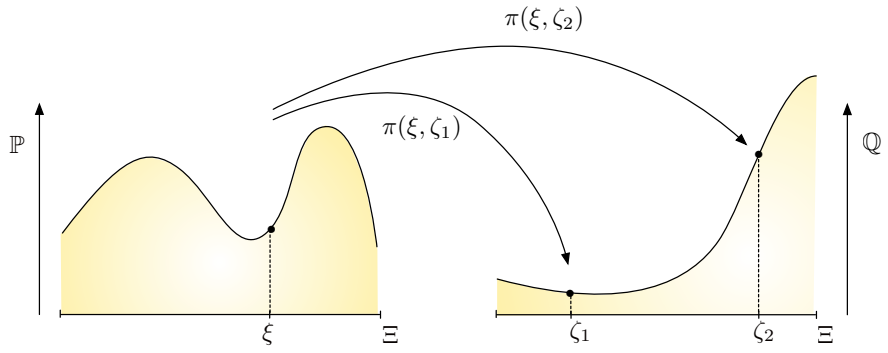

Figure 3. *A coupling between two probability measures on a subset Ξ of the real line $\mathbb{R}$.*

Remark 4.1. *Strictly speaking, in the present situation where the ground cost $c(\xi,\zeta)$ is quadratic, the Wasserstein distance over $\mathcal{P}(\Xi)$ is the square root of the quantity $W(\mathbb{P},\mathbb{Q})$ in (4.1). To keep the terminology simple, we abusively refer to $W(\mathbb{P},\mathbb{Q})$ as the Wasserstein distance in the sequel.*

Since the seminal work [21], the Wasserstein distance (4.1) is often used under a regularized version $W_\varepsilon(\cdot,\cdot)$ in practice:

$$W_\varepsilon(\mathbb{P},\mathbb{Q}) = \inf_{\substack{\pi\in\mathcal{P}(\Xi\times\Xi)\\ \pi_1=\mathbb{P},\ \pi_2=\mathbb{Q}}} \left(\int_{\Xi\times\Xi} c(\xi,\zeta)\,\mathrm{d}\pi(\xi,\zeta) + \varepsilon H(\pi) \right), \tag{4.2}$$

where ε is a "small" parameter and the entropy $H(\pi)$ of a coupling π is the convex function defined by:

$$H(\pi) = \begin{cases} \int_{\Xi\times\Xi} \log \frac{\mathrm{d}\pi}{\mathrm{d}\pi_0}\,\mathrm{d}\pi & \text{if } \pi \text{ is absolutely continuous w.r.t. } \pi_0, \\ \infty & \text{otherwise.} \end{cases} \tag{4.3}$$

The entropy $H(\pi)$ – also referred to as Kullback-Leibler divergence in the literature – evaluates the discrepancy between π and a fixed reference coupling $\pi_0 \in \mathcal{P}(\Xi\times\Xi)$. Here, a probability measure $\pi \in \mathcal{P}(\Xi\times\Xi)$ is called absolutely continuous with respect to π_0 if there exists an integrable function $\alpha \in L^1(\Xi\times\Xi;\mathrm{d}\pi_0)$ such that:

$$\pi = \alpha(\xi,\zeta)\pi_0.$$

The regularization (4.2) was originally perceived as a mere trick to ease the numerical computation of optimal transport plans. However, the quantity $W_\varepsilon(\mathbb{P},\mathbb{Q})$ has recently gained popularity as a "blurred" version of the Wasserstein distance $W(\mathbb{P},\mathbb{Q})$, which is less sensitive to the presence of noise in its arguments $\mathbb{P}$ and $\mathbb{Q}$ [33, 36]. We refer to [57] for a thorough account of the subject of entropy-regularized optimal transport.

According to [10], a judicious choice about π_0, enjoying desirable statistical guarantees, is

$$\pi_0(\xi,\zeta) = \mathbb{P}(\xi)\otimes \mathrm{d}\nu_\xi(\zeta), \quad \text{with} \quad \mathrm{d}\nu_\xi(\zeta) := \alpha_\xi e^{-\frac{c(\xi,\zeta)}{2\sigma^2}} \mathbb{1}_\Xi(\zeta)\mathrm{d}\zeta, \tag{4.4}$$

where the parameter σ^2 controls the variance of the Gaussian in the variable ζ and α_ξ is a normalization factor. Precisely, this definition means that, for any continuous function $\varphi : \Xi\times\Xi \to \mathbb{R}$, it holds:

$$\int_{\Xi\times\Xi} \varphi(\xi,\zeta)\,\mathrm{d}\pi_0(\xi,\zeta) = \int_\Xi \left(\int_\Xi \varphi(\xi,\zeta)\,\mathrm{d}\nu_\xi(\zeta) \right) \mathrm{d}\mathbb{P}(\xi).$$

Returning to the above intuition about couplings, π_0 accounts for a smearing of the local mass of the distribution $\mathbb{P}$ over a characteristic length σ. When σ is very small, π_0 resembles the "identity" coupling, that lays each atom of $\mathbb{P}$ in place; the entropy penalty term $H(\pi)$ in (4.2) then imposes that the only couplings acting in the minimization problem defining W_ε should be "close" to this trivial coupling. On the contrary, when σ gets large, couplings with increasingly smeared second marginal become relevant candidates in (4.2). Loosely speaking, the parameter σ plays the role of a degree of confidence in the nominal law $\mathbb{P}$.

4.2. Formulation of the distributionally robust shape optimization problem

As presented in Section 3.3, the distributionally robust version of the ideal shape optimization problem (2.1) amounts to minimize the worst-case expectation of the cost $\mathcal{C}(\Omega,\xi)$ when the probability law $\mathbb{Q}$ of the uncertain parameters ξ is "close" to the nominal law $\mathbb{P}$:

$$\min_\Omega J_{\mathrm{dr}}(\Omega), \ \text{where } J_{\mathrm{dr}}(\Omega) := \sup_{\mathbb{Q}\in\mathcal{A}} \int_\Xi \mathcal{C}(\Omega,\xi)\,\mathrm{d}\mathbb{Q}(\xi). \tag{4.5}$$

Following [35, 54], in this article, we consider ambiguity sets $\mathcal{A}$ made of laws $\mathbb{Q}$ that are close to $\mathbb{P}$ in terms of the (regularized or non regularized) Wasserstein distance, i.e. $\mathcal{A} = \mathcal{A}_{\mathrm{W}}$ or $\mathcal{A}_{\mathrm{W}_\varepsilon}$, where:

$$\mathcal{A}_{\mathrm{W}} := \Big\{ \mathbb{Q}\in\mathcal{P}(\Xi) \text{ s.t. } W(\mathbb{P},\mathbb{Q}) \le m \Big\}, \ \mathcal{A}_{\mathrm{W}_\varepsilon} := \Big\{ \mathbb{Q}\in\mathcal{P}(\Xi) \text{ s.t. } W_\varepsilon(\mathbb{P},\mathbb{Q}) \le m \Big\}, \tag{4.6}$$

and $m \ge 0$ is a user-defined bound.

The ambiguity set $\mathcal{A}_{\mathrm{W}_\varepsilon}$ built from the entropy-regularized Wasserstein distance is of particular interest, as it lends itself to a very convenient reformulation of the intricate bilevel problem (4.5) for general cost functions $\mathcal{C}(\Omega,\xi)$, as we shall present in the next Section 4.3. Before proceeding, let us point out at a few important extensions of this model, that are addressed in [26]:

- Ambiguity sets different from those $\mathcal{A}_{\mathrm{W}}$ or $\mathcal{A}_{\mathrm{W}_\varepsilon}$ in (4.6) could be considered. For instance, following [29], when only information about the first- and second-order moments of $\mathbb{P}$ is available, it is natural to rely on an ambiguity set $\mathcal{A}_{\mathrm{M}}$ gathering laws whose first- and second-order moments coincide with, or are close to those of $\mathbb{P}$.
- Other measures of risk of the cost $\mathcal{C}(\Omega,\xi)$ than the mean value $\int_\Xi \mathcal{C}(\Omega,\xi)\,\mathrm{d}\mathbb{P}(\xi)$ can be given distributionally robust counterparts, notably its conditional value at risk (or superquantile), see [63, 64] about this notion, and [30, 43] about its use in optimal design.

4.3. A single level formulation for the entropy-regularized Wasserstein distributionally robust problem

The distributionally robust problem (4.5) consists in the minimization of the supremum functional $J_{\mathrm{dr}}(\Omega)$. This min-max problem is awkward from the numerical viewpoint: typically, at each iteration of an outer minimization loop, one has to compute the inner supremum by a global optimization algorithm.

In this section, we show that when the ambiguity set $\mathcal{A}$ in (4.5) is that $\mathcal{A}_{\mathrm{W}_\varepsilon}$ in (4.6), built from the entropy-regularized Wasserstein distance, this problem rewrites as a single-level minimization problem, featuring an additional optimization variable.

The cornerstone of this treatment is a duality formula derived in [10, 75] which expresses the worst-case functional $J_{\mathrm{dr}}(\Omega)$ as an infimum. This formula is used as a systematical tool in the optimal design framework of our works [25, 26]. For simplicity, we state it at the formal level, and we provide a sketch of proof which can be adapted to other contexts, see notably Section 5.

Proposition 4.1. *Let Ξ be a compact subset of $\mathbb{R}^k$ and let $\mathbb{P}\in\mathcal{P}(\Xi)$ be a probability measure. Then, for any continuous function $f:\Xi\to\mathbb{R}$ satisfying "mild" additional properties, it holds:*

$$\sup_{\mathbb{Q}\in\mathcal{A}_{\mathrm{W}_\varepsilon}} \int_\Xi f(\xi)\,\mathrm{d}\mathbb{Q}(\xi) = \inf_{\lambda\geq 0}\left(\lambda m + \lambda\varepsilon\int_\Xi \log\left(\int_\Xi e^{\frac{f(\zeta)-\lambda c(\xi,\zeta)}{\lambda\varepsilon}}\,\mathrm{d}\nu_\xi(\zeta)\right)\mathrm{d}\mathbb{P}(\xi)\right). \tag{4.7}$$

Sketch of proof. Let A denote the left-hand side of (4.7). We first express the constraint featured in this quantity with the help of a Lagrange multiplier:

$$A = \sup_{\mathbb{Q}\in\mathcal{P}(\Xi)}\inf_{\lambda\geq 0}\left(\int_\Xi f(\xi)\,\mathrm{d}\mathbb{Q}(\xi) + \lambda\Big(m - W_\varepsilon(\mathbb{P},\mathbb{Q})\Big)\right).$$

Indeed, the inner infimum in the above expression equals $-\infty$ when $W_\varepsilon(\mathbb{P},\mathbb{Q}) > m$ and 0 otherwise. We now exchange the infimum and supremum in this expression; according to [31], this is possible under "mild" assumptions since the mapping

$$(\mathbb{Q},\lambda) \;\mapsto\; \int_\Xi f(\xi)\,\mathrm{d}\mathbb{Q}(\xi) + \lambda\Big(m - W_\varepsilon(\mathbb{P},\mathbb{Q})\Big)$$

is concave with respect to $\mathbb{Q}$ (by virtue of the properties of the entropy H) and convex (actually linear) with respect to λ. This yields:

$$A = \inf_{\lambda\geq 0}\left(\lambda m + \sup_{\mathbb{Q}\in\mathcal{P}(\Xi)}\left(\int_\Xi f(\xi)\,\mathrm{d}\mathbb{Q}(\xi) - \lambda W_\varepsilon(\mathbb{P},\mathbb{Q})\right)\right).$$

Now recalling the definition (4.2) of the entropy-regularized Wasserstein distance $W_\varepsilon(\mathbb{P},\mathbb{Q})$, we obtain:

$$\begin{aligned}
A &= \inf_{\lambda\geq 0}\left(\lambda m + \sup_{\mathbb{Q}\in\mathcal{P}(\Xi)}\ \sup_{\substack{\pi\in\mathcal{P}(\Xi\times\Xi),\\ \pi_1=\mathbb{P},\ \pi_2=\mathbb{Q}}}\left(\int_\Xi f(\xi)\,\mathrm{d}\mathbb{Q}(\xi) - \lambda\int_{\Xi\times\Xi} c(\xi,\zeta)\mathrm{d}\pi(\xi,\zeta)\ - \lambda\varepsilon H(\pi)\right)\right)\\
&= \inf_{\lambda\geq 0}\left(\lambda m + \sup_{\mathbb{Q}\in\mathcal{P}(\Xi)}\ \sup_{\substack{\pi\in\mathcal{P}(\Xi\times\Xi),\\ \pi_1=\mathbb{P},\ \pi_2=\mathbb{Q}}}\left(\int_{\Xi\times\Xi}\Big(f(\zeta) - \lambda c(\xi,\zeta)\Big)\,\mathrm{d}\pi(\xi,\zeta)\ - \lambda\varepsilon H(\pi)\right)\right)\\
&= \inf_{\lambda\geq 0}\left(\lambda m + \sup_{\substack{\pi\in\mathcal{P}(\Xi\times\Xi),\\ \pi_1=\mathbb{P}}}\left(\int_{\Xi\times\Xi}\Big(f(\zeta) - \lambda c(\xi,\zeta)\Big)\,\mathrm{d}\pi(\xi,\zeta) - \lambda\varepsilon H(\pi)\right)\right).
\end{aligned}$$

Here, we have used the fact that $\mathbb{Q}$ is the second marginal of the candidate couplings involved in the inner supremum to pass from the first line to the second one. In order to obtain the third line, we have rewritten the double supremum over $\mathbb{Q}$ and transport plans π with $\mathbb{Q}$ as second marginal as a single supremum over transport plans with free second marginal. We now recall that the entropy $H(\pi)$ of a coupling $\pi \in \mathcal{P}(\Xi \times \Xi)$ is infinite when π is not absolutely continuous with respect to π_0, so that the inner supremum, that we call $B(\lambda)$, rewrites:

$$B(\lambda) := \sup_{\substack{\pi \in \mathcal{P}(\Xi \times \Xi), \\ \pi_1 = \mathbb{P}}} \left(\int_{\Xi \times \Xi} \Big(f(\zeta) - \lambda c(\xi, \zeta) \Big) \, \mathrm{d}\pi(\xi, \zeta) - \lambda \varepsilon H(\pi) \right) \tag{4.8}$$

$$= \sup_{\substack{\alpha \in L^1(\Xi \times \Xi; \mathrm{d}\pi_0), \, \alpha \geq 0 \\ \int_\Xi \alpha(\xi, \zeta) \, \mathrm{d}\nu_\xi(\zeta) = 1 \text{ for } \mathbb{P}\text{-a.e. } \xi \in \Xi}} \int_{\Xi \times \Xi} \Big(f(\zeta) - \lambda c(\xi, \zeta) - \lambda \varepsilon \log \alpha(\xi, \zeta) \Big) \alpha(\xi, \zeta) \, \mathrm{d}\pi_0(\xi, \zeta).$$

This quantity can now be calculated explicitly, as the constrained supremum of a strictly concave functional. Indeed, by writing the first order optimality conditions associated to the maximization program in (4.8), we see that there exists a function $\mu(\xi)$ such that:

$$\text{For } \pi_0 - \text{a.e. } (\xi, \zeta) \in \Xi \times \Xi, \quad f(\zeta) - \lambda c(\xi, \zeta) - \lambda \varepsilon \log \alpha(\xi, \zeta) - \lambda \varepsilon + \mu(\xi) = 0.$$

It easily follows from this characterization and the constraint $\int_\Xi \alpha(\xi, \zeta) \, \mathrm{d}\nu_\xi(\zeta) = 1$ that:

$$\alpha(\xi, \zeta) = \frac{1}{\int_\Xi e^{\frac{f(\zeta) - \lambda c(\xi, \zeta)}{\lambda \varepsilon}} \mathrm{d}\nu_\xi(\zeta)} e^{\frac{f(\zeta) - \lambda c(\xi, \zeta)}{\lambda \varepsilon}}.$$

Eventually, inserting this expression into (4.8) and rearranging, we arrive at the desired result. □

Applying Proposition 4.1 to the function $f(\xi) = \mathcal{C}(\Omega, \xi)$, the distributionally robust problem (4.5) rewrites:

$$\min_{\substack{\Omega \subset D, \\ \lambda \geq 0}} \mathcal{D}(\Omega, \lambda), \text{ where } \mathcal{D}(\Omega, \lambda) = \lambda m + \lambda \varepsilon \int_\Xi \log \left(\int_\Xi e^{\frac{\mathcal{C}(\Omega, \zeta) - \lambda c(\xi, \zeta)}{\lambda \varepsilon}} \, \mathrm{d}\nu_\xi(\zeta) \right) \mathrm{d}\mathbb{P}(\xi). \tag{4.9}$$

This program is a single-level minimization problem, whose variables are the shape and the additional real variable λ. The minimization of $\mathcal{D}(\Omega, \lambda)$ can be achieved by an alternating descent strategy: minimization steps with respect to Ω at fixed λ – which are realized along the lines of Section 2.3 – are intertwined with minimization steps with respect to the real parameter λ; the latter are very simple, since the partial mapping $\mathcal{D}(\Omega, \cdot)$ is strictly convex.

The numerical solution of (4.9) raises the need to evaluate the integrals over the parameter set Ξ involved in the expressions of $\mathcal{D}(\Omega, \lambda)$ and its derivatives. To achieve this, we rely on Monte-Carlo sampling to approximate every integral of the form $\int_\Xi f(\xi) \, \mathrm{d}\mathbb{P}(\xi)$ as:

$$\int_\Xi f(\xi) \, \mathrm{d}\mathbb{P}(\xi) \approx \frac{1}{N_{\mathrm{MC}}} \sum_{i=1}^{N_{\mathrm{MC}}} f(\xi^i), \tag{4.10}$$

where the ξ^i are N_{MC} independent and identically distributed samples drawn from the probability distribution $\mathbb{P}$. This operation is inevitably costly, as each evaluation of $\mathcal{C}(\Omega, \xi)$ requires the solution to at least one boundary-value problem of the form (2.2). Advanced variance reduction techniques could be used to accelerate this process by keeping the needed number of samples relatively low, see e.g. [55].

Remark 4.2. *According to the proof of Proposition 4.1, the additional variable λ in the single-level optimization problem (4.9) can be interpreted as the Lagrange multiplier for the Wasserstein distance constraint in the original problem (4.5).*

Remark 4.3. *A duality formula similar to (4.7) holds true when the ambiguity set $\mathcal{A}_{\mathrm{W}_\varepsilon}$, based on the regularized Wasserstein distance $W_\varepsilon(\cdot, \cdot)$, is replaced by that $\mathcal{A}_{\mathrm{W}}$ featuring the non regularized distance $W(\cdot, \cdot)$. An argument analogous to the proof of Proposition 4.1 indeed reveals that:*

$$\sup_{\mathbb{Q} \in \mathcal{A}_{\mathrm{W}}} \int_\Xi f(\xi) \, \mathrm{d}\mathbb{Q}(\xi) = \inf_{\lambda \geq 0} \left(\lambda m + \int_\Xi \sup_{\zeta \in \Xi} \left(f(\zeta) - \lambda c(\xi, \zeta) \right) \, \mathrm{d}\mathbb{P}(\xi) \right). \tag{4.11}$$

This formula is consistent with (4.7) insofar as the "log-sum-exp" integrand featured in there is smooth and approximates the supremum in the right-hand side of (4.11) in the limit $\varepsilon \to 0$ [33, 65]. However, (4.7) can be easily evaluated by Monte-Carlo sampling, whereas (4.11) is difficult to compute without additional assumptions about the function f or the set Ξ, see nevertheless Section 5 for an interesting situation where this is possible.

4.4. **A numerical example**

This example deals with the distributionally robust optimization of an L-shaped beam. The situation is illustrated on Fig. 4 (a): the optimized structure Ω is contained in a 2d computational domain D with size 2×2, equipped with a mesh comprising approximately $7,000$ vertices and $13,000$ triangles. The optimized structure is clamped on the upper side Γ_D of ∂D, and a load ξ is applied on a small region Γ_N in the middle of its right-hand side.

We first assume that the load ξ is known perfectly: its ideal value is $\xi^0 := (0, -1)$. We then wish to minimize the stress within the structure Ω under a volume constraint:

$$\min_{\Omega \subset D} \mathcal{C}(\Omega, \xi^0) \text{ s.t. } \mathrm{Vol}(\Omega) = V_T. \tag{4.12}$$

The stress and volume functionals $\mathcal{C}(\Omega, \xi)$ and $\mathrm{Vol}(\Omega)$ featured in this formulation are defined by:

$$\mathcal{C}(\Omega, \xi) = \int_\Omega k(x) ||\sigma(u_{\Omega,\xi})||^2 \, \mathrm{d}x \text{ and } \mathrm{Vol}(\Omega) = \int_\Omega \mathrm{d}x, \tag{4.13}$$

involving a weight function $k(x)$ which equals 1 everywhere except on two small neighborhoods of the regions Γ_D where Ω is clamped and Γ_N where traction loads are applied. The target volume is set to $V_T = 0.7$.

We apply the shape optimization strategy described in Section 2.3 to solve this problem; the volume constraint is handled with the `nullspace` algorithm [32]. The resolution requires about 200 iterations, for a computational time of about 15mn. The shape Ω^*_{det} resulting from this optimization under ideal conditions is depicted on Fig. 4 (b); the stress level within Ω^*_{det} equals $\mathcal{C}(\Omega^*_{\mathrm{det}}, \xi^0) = 1.15732$.

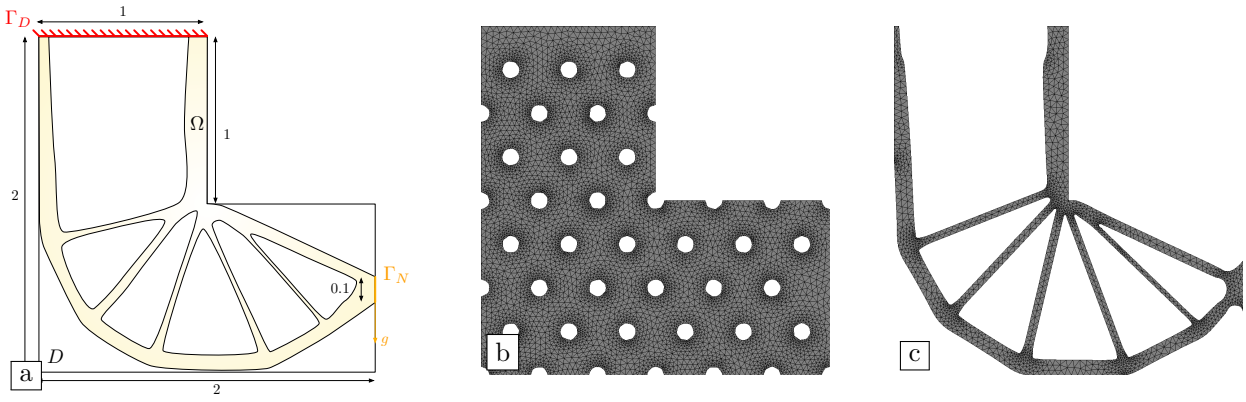

FIGURE 4. *Optimal design of an L-shaped beam considered in Section 4.4; (a) Setting of the test-case; (b) Initial shape; (c) Optimized shape Ω^*_{det} under the assumption of a perfectly known load ξ^0.*

We now assume that the load vector ξ is not known perfectly and that its law itself is subject to uncertainty: we have only access to its empirical reconstruction $\mathbb{P} = \delta_{\xi^0}$, based on the single observation of the ideal load $\xi^0 = (0, -1)$. As proposed in Section 4.2, we thus consider the Wasserstein distributionally version of this problem, using the ambiguity set $\mathcal{A}_{\mathrm{W}_\varepsilon}$ defined in (4.6):

$$\min_{\Omega \subset D} \sup_{\mathbb{Q} \in \mathcal{A}_{\mathrm{W}_\varepsilon}} \int_\Xi \mathcal{C}(\Omega, \xi) \, \mathrm{d}\mathbb{Q}(\xi) \text{ s.t. } \mathrm{Vol}(\Omega) = V_T; \tag{4.14}$$

here, the entropy regularization parameter is set to $\varepsilon = 0.01$. According to Section 4.3, this problem admits the following single-level equivalent formulation:

$$\min_{\substack{\Omega\subset D,\\ \lambda\geq 0}} \mathcal{D}_{\mathrm{W}}(\Omega,\lambda) \;\text{ s.t. }\; \mathrm{Vol}(\Omega) = V_T, \;\text{ where } \mathcal{D}_{\mathrm{W}}(\Omega,\lambda) := \lambda m + \lambda\varepsilon \log\left(\int_{\Xi} e^{\frac{\mathcal{C}(\Omega,\zeta)-\lambda c(\xi^0,\zeta)}{\lambda\varepsilon}} \,\mathrm{d}\nu_{\xi^0}(\zeta)\right).$$

We solve this problem for several values of the bound m on the regularized Wasserstein distance involved in the ambiguity set $\mathcal{A}_{\mathrm{W}_\varepsilon}$ and of the variance σ^2 featured in the reference coupling π_0, see (4.4). The resulting optimized structures are presented in Fig. 5, and the corresponding values of the stress-based functional $\mathcal{C}(\Omega,\xi^0)$ under ideal load conditions are reported in Table 1.

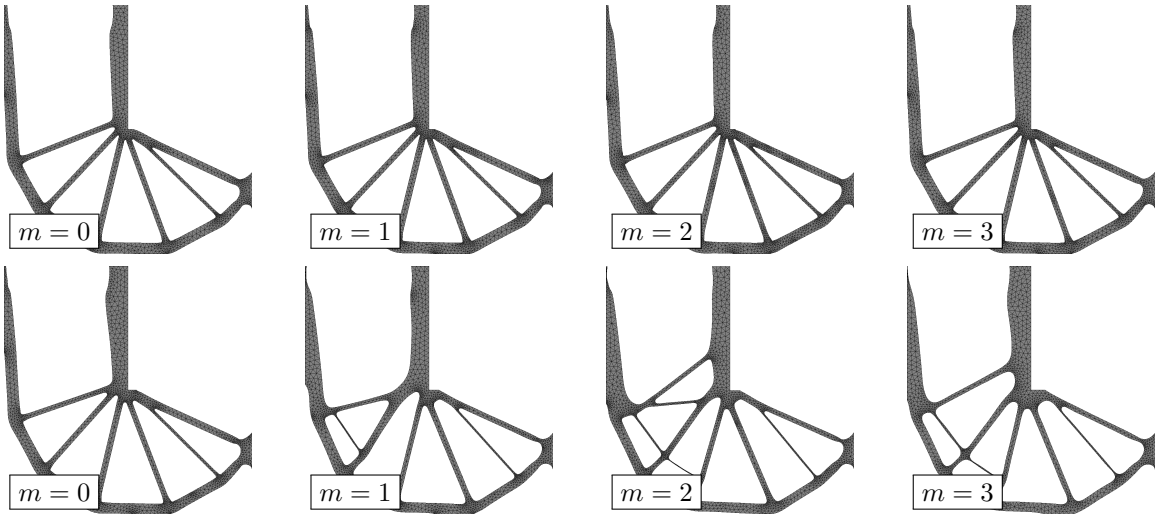

FIGURE 5. *Distributionally robust optimized designs obtained in the L-shaped beam example of Section 4.4, associated to various values of the parameters m and σ^2; (upper row) $\sigma^2 = 1e{-}3$; (lower row) $\sigma^2 = 1e{-}0.5$.*

σ^2 \ m	0	1	2	3
$10^{-0.5}$	1.18047	1.27367	1.30444	1.29508
10^{-3}	1.16097	1.1524	1.17945	1.16396

TABLE 1. *Values of the stress-based cost functional $\mathcal{C}(\Omega,\xi^0)$ for the distributionally robust designs of the L-shaped beam obtained in Section 4.4 when the ideal load ξ^0 is applied.*

When the variance σ^2 in the reference coupling π_0 is low, the optimized structures associated to different values of the Wasserstein radius m of $\mathcal{A}_{\mathrm{W}_\varepsilon}$ are similar to one another, and so are their performances. This observation is in line with the interpretation of σ^2 provided in Section 4.1, as a degree of confidence in the nominal law $\mathbb{P}$: in this situation, the entropy penalization in the definition (4.2) of $W_\varepsilon(\cdot,\cdot)$ imposes the optimal coupling to resemble much that induced by the identity mapping. On the contrary, for higher variance σ^2, the designs optimized for various values of m differ significantly. Their performance in ideal conditions get degraded as m increases. On the contrary, unreported out-of-sample analyses reveal that, on average, the performance under load conditions that differ from ξ^0 get better and better, see [62] for more details on this point. This highlights the ability of distributionally robust formulations to predict efficient designs in situations that are not captured (but "close to") a known ideal situation.

5. Distributionally robust shape optimization under geometric uncertainties

This section adapts the previous ideas to the distributionally robust treatment of geometric uncertainties in shape optimization. In a nutshell, one aims to minimize a cost function $\mathcal{C}(\Omega)$ while taking into account the possibility that the actual shape be a perturbed version of the "blueprint" optimization variable Ω. This concern is ubiquitous in practice, as the geometry of a shape may be unpredictably altered by multiple phenomena, such as imperfections of the manufacturing process, or wear in the course of its use. From the technical viewpoint, the present situation is more delicate than those addressed in Section 4: the uncertainty now bears on the optimized variable itself, whose perturbations are not easily accounted for.

In the literature, geometric uncertainties have been comparatively less often considered than uncertainties over physical parameters such as loads or material coefficients. In the framework of density-based methods, the contributions [45, 46] place the uncertainty in the parameters of a Heaviside filter for the density function; the works [2, 3, 17], dealing with geometric shape optimization, use an avatar of Hadamard's boundary variation method to define perturbations of a shape, and the contribution [77] assumes uncertain parameters in the level set function representing the optimized shape.

After describing the considered mathematical model for geometric uncertainties in Section 5.1, we formulate the considered distributionally robust shape optimization problem under geometric uncertainties and derive a tractable formulation of the latter in Section 5.2. A numerical example is eventually presented in Section 5.3.

5.1. Statement of the optimization problem under geometric uncertainties

Throughout this section, we assume that the parameters of the linear elasticity model (2.2) for the optimized structure Ω, i.e. the loads and the material coefficients, are known perfectly. We consider the minimization of a cost function $\mathcal{C}(\Omega)$:

$$\min_{\Omega} \mathcal{C}(\Omega), \tag{5.1}$$

where, again, constraints are omitted for simplicity of the presentation.

We now wish to modify this ideal formulation by distinguishing the "blueprint" optimized shape Ω and its actual, randomly perturbed version through which it appears in the problem. Following [2, 3], we consider uncertain perturbations of the geometry of a shape Ω inspired by the method of Hadamard recalled in Section 2.2:

$$(\mathrm{Id} + \eta v(x,\xi) n_\Omega)(\Omega), \tag{5.2}$$

where:

- The unit normal vector n_Ω to $\partial\Omega$ is extended to the whole space $\mathbb{R}^d$,
- The parameter η accounts for the "smallness" of the considered geometric perturbations,
- The random field $v(x,\xi)$ encoding the profile of these perturbations has the following finite-dimensional structure:

$$v(x,\xi) = \sum_{i=1}^{k} v_i(x)\xi_i, \tag{5.3}$$

  made of known, deterministic functions $v_i : \mathbb{R}^d \to \mathbb{R}$ and a random vector $\xi \in \Xi$. For simplicity, we also assume that Ξ is the closed ball in $\mathbb{R}^k$ with center 0 and radius R.

In this setting, the cost $\mathcal{C}(\Omega,\xi)$ of a shape Ω undergoing perturbations parametrized by $\xi \in \Xi$ is:

$$\mathcal{C}(\Omega,\xi) = \mathcal{C}((\mathrm{Id} + \eta v(x,\xi) n_\Omega)(\Omega)).$$

Note the small abuse of notations, where we denote in the same way $\mathcal{C}(\Omega) \equiv \mathcal{C}(\Omega, 0)$.

Remark 5.1. *In practice, the finite-dimensional structure (5.3) for the random perturbation field $v(x,\xi)$ of the boundary of shapes results from a truncated Karhunen-Loève decomposition of the "true" infinite-dimensional random field, from the datum of its covariance kernel, see [49], and [1] for a comprehensive introduction to this practice.*

Remark 5.2.

- *The model introduced in this section for geometric uncertainties is quite specific to the setting of this article, where the optimized design is a domain Ω in $\mathbb{R}^d$. However, we believe that it could be adapted to density-based topology optimization, where the optimized design is represented by a "gray level" function $h : D \to [0,1]$, defined a large, fixed hold-all domain D.*
- *The present model of geometric uncertainties is admittedly questionable. Notably, one would naturally expect that the random perturbation field $v(x,\xi)$ should depend on the shape Ω itself.*

5.2. **Distributionally robust treatment of geometric uncertainties**

Like in the previous Section 4, we place ourselves in the context where the true law $\mathbb{P}_{\text{true}} \in \mathcal{P}(\Xi)$ of the random perturbation vector ξ featured in (5.3) is unknown: only an approximate nominal law $\mathbb{P} \in \mathcal{P}(\Xi)$ is available. The considered distributionally robust version of the problem (5.1) under geometric uncertainties then reads:

$$\min_{\Omega} J_{\text{dr}}(\Omega), \text{ where } J_{\text{dr}}(\Omega) := \sup_{\mathbb{Q}\in\mathcal{A}} \int_{\Xi} \mathcal{C}(\Omega,\xi)\, \mathrm{d}\mathbb{Q}(\xi), \tag{5.4}$$

where $\mathcal{A}$ stands for one of the Wasserstein-type ambiguity sets $\mathcal{A}_{\text{W}}$ or $\mathcal{A}_{\text{W}_\varepsilon}$ defined in (4.6).

5.2.1. *Ambiguity set built from the regularized Wasserstein distance*

We first investigate the instance of (5.4) where the considered ambiguity set $\mathcal{A}$ is that $\mathcal{A}_{\text{W}_\varepsilon}$, built from the regularized Wasserstein distance $W_\varepsilon(\cdot,\cdot)$ in (4.2). Then, the same analysis as in Section 4, based on the application of the duality formula of Proposition 4.1, shows that the distributionally robust problem (5.4) admits the following single-level formulation:

$$\min_{\substack{\Omega,\\ \lambda\geq 0}} \mathcal{D}(\Omega,\lambda), \text{ where } \mathcal{D}(\Omega,\lambda) = \lambda m + \lambda\varepsilon \int_{\Xi} \log\left(\int_{\Xi} e^{\frac{\mathcal{C}(\Omega,\zeta)-\lambda c(\xi,\zeta)}{\lambda\varepsilon}}\, \mathrm{d}\nu_\xi(\zeta)\right) \mathrm{d}\mathbb{P}(\xi).$$

Contrary to the situation addressed in Section 4 however, the numerical evaluation of the single-level objective function $\mathcal{D}(\Omega,\lambda)$ is awkward. Let us indeed recall that the integrals in its expression are computed by Monte-Carlo sampling (4.10). Thus, for each sample $\xi^i \in \Xi$ in this formula, one has first to identify the perturbed domain $(\mathrm{Id} + \eta v(x,\xi^i))(\Omega)$ (i.e. to mesh it), before solving the boundary-value problem (2.2) posed on the latter. Hence, the evaluation of $\mathcal{D}(\Omega,\lambda)$ by the Monte-Carlo formula (4.10) not only demands to solve N_{MC} different boundary value problems, but also to create the N_{MC} meshes of the perturbed versions of the shape where they take place.

To circumvent this dramatic increase in CPU cost, we replace $\mathcal{C}(\Omega,\xi)$ with a linearized version, taking advantage of the "smallness" of the normal amplitude η of geometric perturbations:

$$\mathcal{C}(\Omega,\xi) \approx \mathcal{C}(\Omega) + \eta\, \mathcal{C}'(\Omega)(v(x,\xi)n_\Omega).$$

Doing so, introducing the notation

$$G(\Omega) = (G_1(\Omega),\dots,G_k(\Omega)) \in \mathbb{R}^k \text{ with } G_i(\Omega) = \eta\mathcal{C}'(\Omega)(v_i n_\Omega), \tag{5.5}$$

the functional $\mathcal{D}(\Omega,\lambda)$ is replaced by the approximate version:

$$\mathcal{D}(\Omega,\lambda) \approx \lambda m + \lambda\varepsilon \int_{\Xi} \log\left(\int_{\Xi} e^{\frac{\mathcal{C}(\Omega)+G(\Omega)\cdot\zeta-\lambda c(\xi,\zeta)}{\lambda\varepsilon}}\, \mathrm{d}\nu_\xi(\zeta)\right) \mathrm{d}\mathbb{P}(\xi).$$

This last expression can be conveniently evaluated by using a single mesh for the current shape Ω.

5.2.2. *Ambiguity set based on the non regularized Wasserstein distance*

Retaining the idea of approximating the cost function with its linearized version, we now show that an explicit expression can be derived for the version of the distributionally robust problem (5.4) featuring an ambiguity set $\mathcal{A}_{\text{W}}$ based on the non regularized Wasserstein distance.

Indeed, recalling the duality formula (4.11) associated to the use of the non regularized Wasserstein distance $W(\cdot,\cdot)$, the distributionally robust functional $J_{\text{dr}}(\Omega)$ takes the following form:

$$J_{\text{dr}}(\Omega) = \mathcal{C}(\Omega) + \inf_{\lambda\geq 0}\left(\lambda m + \int_{\Xi} B(\xi,\lambda)\, \mathrm{d}\mathbb{P}(\xi)\right), \text{ where } B(\xi,\lambda) := \sup_{\zeta\in\Xi}\left(G(\Omega)\cdot\zeta - \lambda c(\xi,\zeta)\right),$$

and $G(\Omega) \in \mathbb{R}^k$ is vector defined in (5.5). Here, contrary to the general situation discussed in Remark 4.3, the innermost supremum $B(\xi, \lambda)$ features a linear "cost function" $\zeta \mapsto G(\Omega) \cdot \zeta$, and it can be computed in closed form. Indeed, recalling that Ξ is the unit ball in $\mathbb{R}^k$ with center 0 and radius R, a simple calculation shows that this supremum is attained at the point:

$$\zeta^*(\xi, \lambda) = \begin{cases} \frac{1}{2\lambda} G(\Omega) + \xi & \text{if } \left| \frac{1}{2\lambda} G(\Omega) + \xi \right| \leq R, \\ R \frac{G(\Omega) + 2\lambda\xi}{|G(\Omega) + 2\lambda\xi|} & \text{otherwise.} \end{cases}$$

Another elementary calculation then shows that:

$$B(\xi, \lambda) = \begin{cases} G(\Omega) \cdot \xi + \frac{1}{4\lambda} |G(\Omega)|^2 & \text{if } \left| \frac{1}{2\lambda} G(\Omega) + \xi \right| \leq R, \\ -\lambda(R^2 + |\xi|^2) + R|G(\Omega) + 2\lambda\xi| & \text{otherwise.} \end{cases}$$

All things considered, the distributionally robust problem (5.4) to geometric uncertainties reads:

$$\min_{\substack{\Omega, \\ \lambda \geq 0}} \mathcal{D}(\Omega, \lambda), \text{ where } \mathcal{D}(\Omega, \lambda) := \mathcal{C}(\Omega) + \lambda m + \int_\Xi B(\xi, \lambda) \, \mathrm{d}\mathbb{P}(\xi). \tag{5.6}$$

Again, this single-level shape optimization problem featuring the additional variable $\lambda \geq 0$ can be readily addressed by the numerical methods presented in Section 2.3.

Remark 5.3. *The expression in (5.6) can be interpreted as a form of robustification by regularization; see Section 8 in [44] for a general treatment and discussion.*

5.3. **A numerical example**

In this section, we consider the distributionally robust optimal design of a two-dimensional wheel under geometric uncertainties. The physical situation is described in the reference frame of the wheel; the latter is attached at the boundary of the central circular hole Γ_D, and the other boundary regions, denoted by Γ, are subject to traction-free (homogeneous Neumann) boundary conditions, see Fig. 6 (a). The structure undergoes an inertial force $f : \mathbb{R}^2 \to \mathbb{R}^2$ induced by the rotation of the frame:

$$\forall x = (x_1, x_2) \in \mathbb{R}^2, \quad f(x) = x^\perp |x|, \text{ where } x^\perp := (-x_2, x_1).$$

In this situation, we aim to optimize the compliance of the wheel under a volume constraint. In the ideal situation where the actual geometry of the wheel coincides with its "blueprint" version, the problem at hand reads:

$$\min_\Omega \mathcal{C}(\Omega) \text{ s.t. } \mathrm{Vol}(\Omega) = V_T, \text{ where } \mathcal{C}(\Omega) := \int_\Omega Ae(u_\Omega) : e(u_\Omega) \, \mathrm{d}x.$$

This problem is solved for the two different values of the volume target $V_T = 1.35$ and $V_T = 1.6$; the corresponding optimized shapes are presented on Fig. 6 (b) and (c).

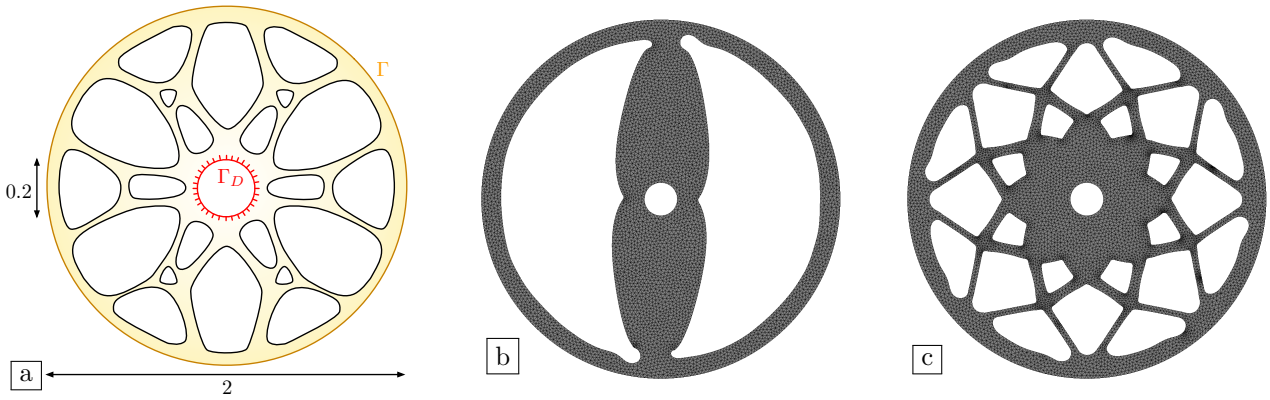

FIGURE 6. *(a) Setting of the wheel optimization example of Section 5.3; (b) Optimized shape of the wheel under ideal conditions with volume target $V_T = 1.35$; (c) Optimized shape under ideal conditions with volume target $V_T = 1.6$.*

Let us now slip into the context where the geometry of the shape Ω is subjected to perturbations, whose law is not known with certainty. As considered in Section 5.1, the optimized shape Ω may in reality be produced, or altered with time, under the perturbed form (5.2). Here, the parameter η accounting for the magnitude of perturbations equals $\eta = 1e-3$, the normal perturbation field $v(x, \xi)$ has the finite-dimensional structure (5.3), resulting from the truncation after the first $k = 10$ modes of a Karhunen-Loève expansion of the covariance kernel defined by:

$$\forall x, y \in D, \quad \mathrm{Cov}(x, y) = \beta^2 e^{-\frac{|x-y|}{l_{\mathrm{cor}}}}, \tag{5.7}$$

involving the amplitude $\beta = 10$ and correlation length $l_{\mathrm{cor}} = 4e - 2$. The eigenfunctions v_i featured in this decomposition are depicted on Fig. 7.

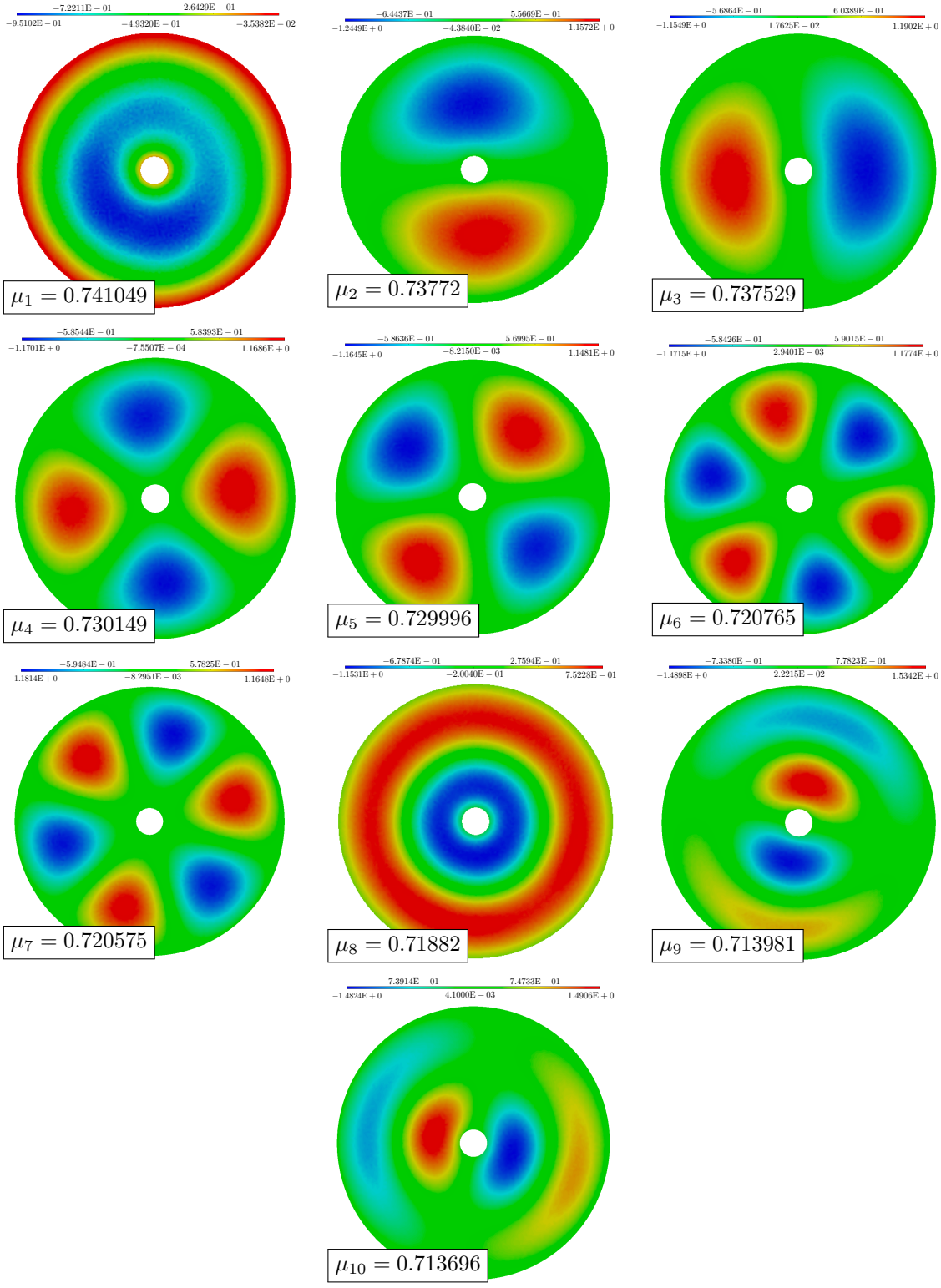

FIGURE 7. *Representation of the first 10 eigenpairs* (v_i, μ_i) *of the covariance kernel* $\mathrm{Cov}(x, y)$ *defined in* (5.7) *in the wheel optimization example of Section 5.3.*

We apply the numerical strategy of Section 2.3 to the resolution of the approximate distributionally robust optimization problem (5.6), based on the use of the non regularized Wasserstein distance, see Section 5.2.2. The results, associated to the volume targets $V_T = 1.35$ and 1.6 and various values of the parameters m and R are depicted on Figs. 8 and 9, respectively.

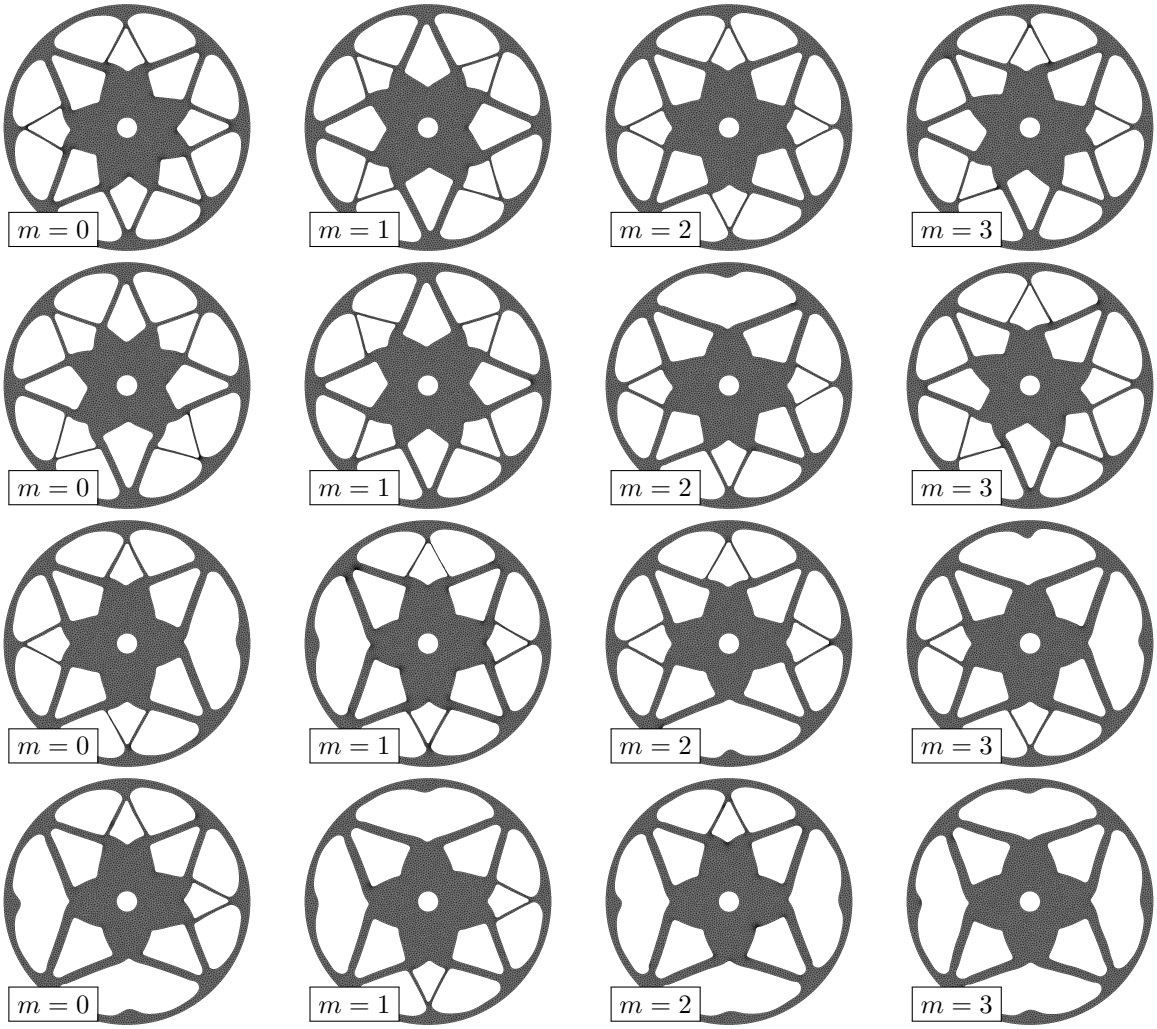

FIGURE 8. *Distributionally robust optimized designs obtained in the wheel example of Section 5.3 associated to the volume target $V_T = 1.35$ and various values of the Wasserstein distance parameter m and the radius R of the set Ξ; (1^{st} row) $R = 0.1$, (2^{nd} row) $R = 1$; (3^{rd} row) $R = 5$; (4^{th} row) $R = 50$.*

The first series of examples, featuring a volume target $V_T = 1.35$ and represented on Fig. 8 reveals a high sensitivity to the parameters R and m. As a general trend, when the parameters m and R take large values, the structures tend to adopt a configuration made of fewer thin bars and four main diagonal connections. By contrast, the designs obtained by imposing a volume target $V_T = 1.6$, represented on Fig. 9, are much less sensitive to these parameters, suggesting that the shape in Fig. 6 (c), optimized under ideal conditions, is already quite robust to geometric uncertainties.

6. Conclusion and perspectives

In this article, we have outlined the introduction of the idea of distributional robustness into the field of shape and topology optimization proposed in our previous works [25, 26]. In the situation where some of the

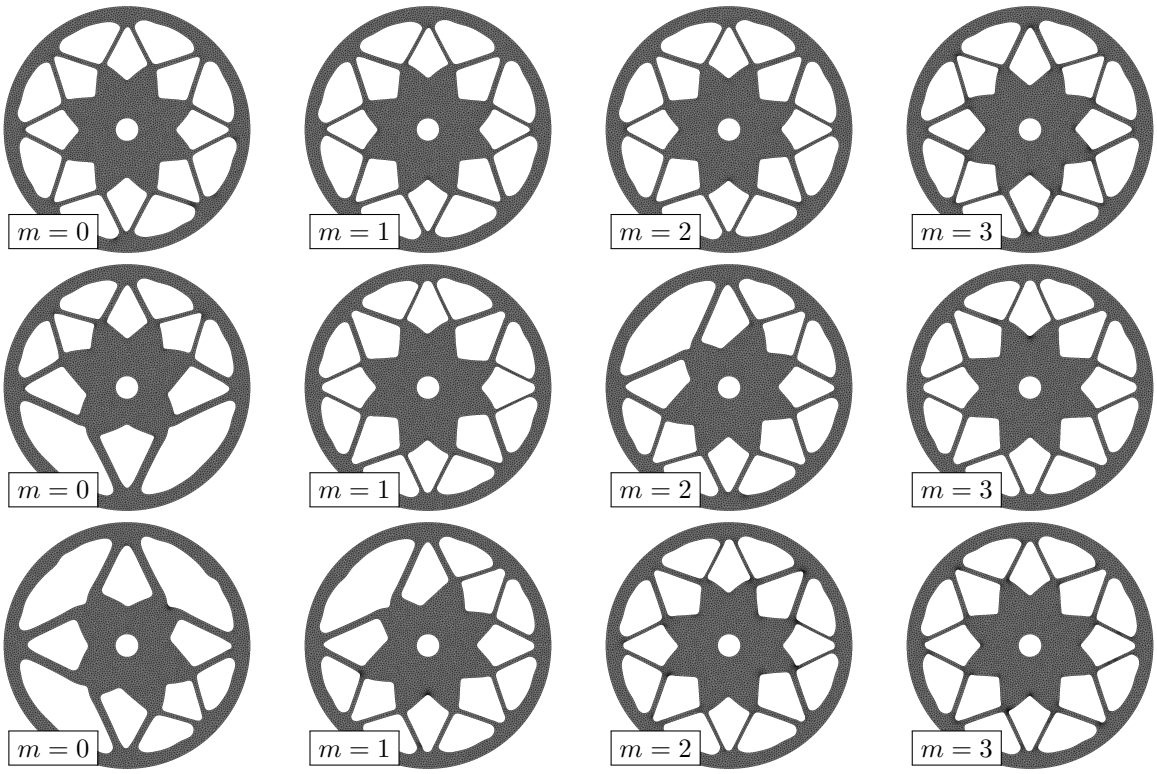

FIGURE 9. *Distributionally robust optimized designs obtained in the wheel optimization example of Section 5.3 associated to the volume target $V_T = 1.6$ and various values of the Wasserstein distance parameter m and the radius R of Ξ; (1^{st} row) $R = 1$; (2^{nd} row) $R = 5$; (3^{rd} row) $R = 50$.*

parameters of the physical model are uncertain, the latter allows to take into account that, often in reality, the probability law of these parameters is itself unknown: at best, it can be approximated by a nominal law reconstructed from a few observed samples. The distributionally robust viewpoint consists in minimizing the worst-case expectation of a physical cost, when the law of the uncertain parameters is "close" to the nominal one. Focusing on a particular situation for clarity, we have exemplified a very general procedure, based on convex duality, to derive tractable single-level formulations of the difficult min-max problems raised by this viewpoint. In a second part of this article, we have adapted this framework to handle geometric uncertainties: the optimization variable itself is subject to uncertainty, rather than an independent parameter of the physical model. The model for geometric uncertainties used in there is admittedly not completely convincing, as it ignores the dependence of the geometric uncertainties on the actual shape. In the near future, we wish to explore another formulation of such geometric uncertainties, that leverages optimal transport as a means to appraise when shapes are "close" from one another.

Acknowledgements. The work of C.D, J.P. and B.T. is partially supported by the projects ANR-24-CE40-2216 STOIQUES and ANR-22-CE46-0006 StableProxies and by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01), financed by the French Agence Nationale de la Recherche (ANR). F.I. acknowledges the support of the AI Cluster ANITI (ANR-23-IACL-0002).

REFERENCES

[1] A. ALEXANDERIAN, *A primer on the karhunen-loève expansion*, arXiv preprint arXiv:2605.08959, (2026).
[2] G. ALLAIRE AND C. DAPOGNY, *A linearized approach to worst-case design in parametric and geometric shape optimization*, Mathematical Models and Methods in Applied Sciences, 24 (2014), pp. 2199–2257.
[3] ———, *A deterministic approximation method in shape optimization under random uncertainties*, SMAI Journal of computational mathematics, 1 (2015), pp. 83–143.

[4] G. Allaire, C. Dapogny, and P. Frey, *Topology and geometry optimization of elastic structures by exact deformation of simplicial mesh*, Comptes Rendus Mathematique, 349 (2011), pp. 999–1003.

[5] ———, *A mesh evolution algorithm based on the level set method for geometry and topology optimization*, Structural and Multidisciplinary Optimization, 48 (2013), pp. 711–715.

[6] ———, *Shape optimization with a level set based mesh evolution method*, Computer Methods in Applied Mechanics and Engineering, 282 (2014), pp. 22–53.

[7] G. Allaire, C. Dapogny, and F. Jouve, *Shape and topology optimization*, in Geometric partial differential equations, part II, A. Bonito and R. Nochetto eds., Handbook of Numerical Analysis, vol. 22, (2021), pp. 1–132.

[8] G. Allaire, F. Jouve, and A.-M. Toader, *Structural optimization using sensitivity analysis and a level-set method*, Journal of computational physics, 194 (2004), pp. 363–393.

[9] S. Amstutz and M. Ciligot-Travain, *A notion of compliance robustness in topology optimization*, ESAIM: Control, Optimisation and Calculus of Variations, 22 (2016), pp. 64–87.

[10] W. Azizian, F. Iutzeler, and J. Malick, *Regularization for wasserstein distributionally robust optimization*, ESAIM: Control, Optimisation and Calculus of Variations, 29 (2023), p. 33.

[11] G. Balarac, F. Basile, P. Bénard, F. Bordeu, J.-B. Chapelier, L. Cirrottola, G. Caumon, C. Dapogny, P. Frey, A. Froehly, et al., *Tetrahedral remeshing in the context of large-scale numerical simulation and high performance computing*, to appear in Maths in Action, (2021).

[12] A. Ben-Tal, A. Nemirovski, and L. El Ghaoui, *Robust optimization*, (2009).

[13] J. Blanchet, J. Li, S. Lin, and X. Zhang, *Distributionally robust optimization and robust statistics*, arXiv preprint arXiv:2401.14655, (2024).

[14] C. Bui, C. Dapogny, and P. Frey, *An accurate anisotropic adaptation method for solving the level set advection equation*, International Journal for Numerical Methods in Fluids, 70 (2012), pp. 899–922.

[15] M. Burger, *A framework for the construction of level set methods for shape optimization and reconstruction*, Interfaces and Free boundaries, 5 (2003), pp. 301–329.

[16] E. L. Cardoso, G. Da Silva, and A. T. Beck, *Robust topology optimization of compliant mechanisms with uncertainties in output stiffness*, International Journal for Numerical Methods in Engineering, 119 (2019), pp. 532–547.

[17] S. Chen and W. Chen, *A new level-set based approach to shape and topology optimization under geometric uncertainty*, Structural and Multidisciplinary Optimization, 44 (2011), pp. 1–18.

[18] S. Chen, W. Chen, and S. Lee, *Level set based robust shape and topology optimization under random field uncertainties*, Structural and Multidisciplinary Optimization, 41 (2010), pp. 507–524.

[19] A. Cherkaev and E. Cherkaeva, *Optimal design for uncertain loading condition*, in Homogenization: In Memory of Serguei Kozlov, World Scientific, 1999, pp. 193–213.

[20] S. Chewi, J. Niles-Weed, and P. Rigollet, *Statistical optimal transport.*, Springer Nature, 2025.

[21] M. Cuturi, *Sinkhorn distances: Lightspeed computation of optimal transport*, Advances in neural information processing systems, 26 (2013).

[22] C. Dapogny, C. Dobrzynski, and P. Frey, *Three-dimensional adaptive domain remeshing, implicit domain meshing, and applications to free and moving boundary problems*, Journal of computational physics, 262 (2014), pp. 358–378.

[23] C. Dapogny and F. Feppon, *Shape optimization using a level set based mesh evolution method: an overview and tutorial*, Comptes Rendus. Mathématique, 361 (2023), pp. 1267–1332.

[24] C. Dapogny and P. Frey, *Computation of the signed distance function to a discrete contour on adapted triangulation*, Calcolo, 49 (2012), pp. 193–219.

[25] C. Dapogny, F. Iutzeler, A. Meda, and B. Thibert, *Entropy-regularized wasserstein distributionally robust shape and topology optimization*, Structural and Multidisciplinary Optimization, 66 (2023), p. 42.

[26] C. Dapogny, J. Prando, and B. Thibert, *Distributionally robust shape and topology optimization*, to appear in J. Comput. Phys., (2025).

[27] F. De Gournay, *Velocity extension for the level-set method and multiple eigenvalues in shape optimization*, SIAM journal on control and optimization, 45 (2006), pp. 343–367.

[28] F. De Gournay, G. Allaire, and F. Jouve, *Shape and topology optimization of the robust compliance via the level set method*, ESAIM: Control, Optimisation and Calculus of Variations, 14 (2008), pp. 43–70.

[29] E. Delage and Y. Ye, *Distributionally robust optimization under moment uncertainty with application to data-driven problems*, Operations research, 58 (2010), pp. 595–612.

[30] M. Eigel, J. Neumann, R. Schneider, and S. Wolf, *Risk averse stochastic structural topology optimization*, Computer Methods in Applied Mechanics and Engineering, 334 (2018), pp. 470–482.

[31] I. Ekeland and R. Temam, *Convex analysis and variational problems*, SIAM, 1999.

[32] F. Feppon, G. Allaire, and C. Dapogny, *Null space gradient flows for constrained optimization with applications to shape optimization*, ESAIM: Control, Optimisation and Calculus of Variations, 26 (2020), p. 90.

[33] J. Feydy, *Analyse de données géométriques, au delà des convolutions*, PhD thesis, Université Paris-Saclay, 2020.

[34] P. Gangl, T. Komann, N. Krenn, and S. Ulbrich, *Robust topology optimization of electric machines using topological derivatives*, arXiv preprint arXiv:2504.05070, (2025).

[35] R. Gao and A. J. Kleywegt, *Distributionally robust stochastic optimization with wasserstein distance*, `arXiv` preprint `arXiv:1604.02199`, (2016).

[36] A. Genevay, *Entropy-regularized optimal transport for machine learning*, PhD thesis, Université Paris sciences et lettres, 2019.

[37] P. L. Gould and Y. Feng, *Introduction to linear elasticity*, Springer, 1994.
[38] X. Guo, W. Zhang, and L. Zhang, *Robust structural topology optimization considering boundary uncertainties*, Computer Methods in Applied Mechanics and Engineering, 253 (2013), pp. 356–368.
[39] K. M. Hamdia, H. Ghasemi, X. Zhuang, and T. Rabczuk, *Multilevel monte carlo method for topology optimization of flexoelectric composites with uncertain material properties*, Engineering Analysis with Boundary Elements, 134 (2022), pp. 412–418.
[40] F. Hecht, *New development in freefem++*, Journal of numerical mathematics, 20 (2012), pp. 251–266.
[41] A. Henrot and M. Pierre, *Shape Variation and Optimization*, EMS Tracts in Mathematics Vol. 28, 2018.
[42] M. G. Kapteyn, K. E. Willcox, and A. B. Philpott, *Distributionally robust optimization for engineering design under uncertainty*, International Journal for Numerical Methods in Engineering, 120 (2019), pp. 835–859.
[43] D. P. Kouri and T. M. Surowiec, *Risk-averse pde-constrained optimization using the conditional value-at-risk*, SIAM Journal on Optimization, 26 (2016), pp. 365–396.
[44] D. Kuhn, S. Shafiee, and W. Wiesemann, *Distributionally robust optimization*, Acta Numerica, 34 (2025), p. 579–804.
[45] B. S. Lazarov, M. Schevenels, and O. Sigmund, *Topology optimization considering material and geometric uncertainties using stochastic collocation methods*, Structural and Multidisciplinary optimization, 46 (2012), pp. 597–612.
[46] B. S. Lazarov, M. Schevenels, and O. Sigmund, *Topology optimization with geometric uncertainties by perturbation techniques*, International Journal for Numerical Methods in Engineering, 90 (2012), pp. 1321–1336.
[47] F. Lin, X. Fang, and Z. Gao, *Distributionally robust optimization: A review on theory and applications*, Numerical Algebra, Control & Optimization, 12 (2022), p. 159.
[48] J. Lions, *Optimal control of systems governed by partial differential equations*, Grundlehren der mathematischen Wissenschaften, Springer-Verlag, 1971.
[49] M. Loève, *Probability theory*, Courier Dover Publications, 2017.
[50] J. Martínez-Frutos, G. Allaire, C. Dapogny, and F. Periago, *Structural optimization under internal porosity constraints using topological derivatives*, Computer Methods in Applied Mechanics and Engineering, 345 (2019), pp. 1–25.
[51] J. Martínez-Frutos and D. Herrero-Pérez, *Large-scale robust topology optimization using multi-gpu systems*, Computer Methods in Applied Mechanics and Engineering, 311 (2016), pp. 393–414.
[52] J. Martínez-Frutos, D. Herrero-Pérez, M. Kessler, and F. Periago, *Robust shape optimization of continuous structures via the level set method*, Computer Methods in Applied Mechanics and Engineering, 305 (2016), pp. 271–291.
[53] K. Maute, *Topology optimization under uncertainty*, in Topology optimization in structural and continuum mechanics, (2014), pp. 457–471.
[54] P. Mohajerin Esfahani and D. Kuhn, *Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations*, Mathematical Programming, 171 (2018), pp. 115–166.
[55] S. Mohamed, M. Rosca, M. Figurnov, and A. Mnih, *Monte carlo gradient estimation in machine learning*, Journal of Machine Learning Research, 21 (2020), pp. 1–62.
[56] F. Murat and J. Simon, *Sur le contrôle par un domaine géométrique*, Pré-publication du Laboratoire d'Analyse Numérique,(76015), (1976).
[57] M. Nutz, *Introduction to entropic optimal transport*, (2021).
[58] S. Osher and R. Fedkiw, *Level set methods and dynamic implicit surfaces*, vol. 153, Springer Science & Business Media, 2006.
[59] S. Osher and J. A. Sethian, *Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations*, Journal of computational physics, 79 (1988), pp. 12–49.
[60] G. Peyré and M. Cuturi, *Computational optimal transport: With applications to data science*, Foundations and Trends in Machine Learning, 11 (2019), pp. 355–607.
[61] R.-E. Plessix, *A review of the adjoint-state method for computing the gradient of a functional with geophysical applications*, Geophysical Journal International, 167 (2006), pp. 495–503.
[62] J. Prando, *Distributionally robust shape and topology optimization*, PhD thesis, Université Grenoble Alpes, 2020.
[63] R. T. Rockafellar and S. Uryasev, *Conditional value-at-risk for general loss distributions*, Journal of banking & finance, 26 (2002), pp. 1443–1471.
[64] R. T. Rockafellar, S. Uryasev, et al., *Optimization of conditional value-at-risk*, Journal of risk, 2 (2000), pp. 21–42.
[65] R. T. Rockafellar and R. J. Wets, *Variational analysis*, Springer, 1998.
[66] F. Santambrogio, *Optimal transport for applied mathematicians*, Birkäuser, 2015.
[67] H. Scarf, *A min-max solution to an inventory problem*, Studies in Mathematical Theory of Inventory and Production, (1958), pp. 201–209.
[68] G. I. Schuëller and H. A. Jensen, *Computational methods in optimization considering uncertainties–an overview*, Computer Methods in Applied Mechanics and Engineering, 198 (2008), pp. 2–13.
[69] J. A. Sethian, *A fast marching level set method for monotonically advancing fronts.*, proceedings of the National Academy of Sciences, 93 (1996), pp. 1591–1595.
[70] ———, *Fast marching methods*, SIAM review, 41 (1999), pp. 199–235.
[71] J. A. Sethian, *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, vol. 3, Cambridge university press, 1999.
[72] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on stochastic programming: modeling and theory*, SIAM, 2021.

[73] J. E. Smith and R. L. Winkler, *The optimizer's curse: Skepticism and postdecision surprise in decision analysis*, Management Science, 52 (2006), pp. 311–322.
[74] J. Sokolowski and J.-P. Zolésio, *Introduction to shape optimization*, Springer, 1992.
[75] J. Wang, R. Gao, and Y. Xie, *Sinkhorn distributionally robust optimization*, arXiv preprint arXiv:2109.11926, (2021).
[76] M. Y. Wang, X. Wang, and D. Guo, *A level set method for structural topology optimization*, Computer methods in applied mechanics and engineering, 192 (2003), pp. 227–246.
[77] W. Zhang and Z. Kang, *Robust shape and topology optimization considering geometric uncertainties with stochastic level set perturbation*, International Journal for Numerical Methods in Engineering, 110 (2017), pp. 31–56.