FULL LENGTH PAPER

A pessimistic bilevel stochastic problem for elastic shape optimization

Johanna Burtscheidt[1] · Matthias Claus[1] · Sergio Conti[2] · Martin Rumpf[3] · Josua Sassen[3] · Rüdiger Schultz[1]

Abstract

We consider pessimistic bilevel stochastic programs in which the follower maximizes over a fixed compact convex set a strictly convex quadratic function, whose Hessian depends on the leader's decision. This results in a random upper level outcome which is evaluated by a convex risk measure. Under assumptions including real analyticity of the lower-level goal function, we prove the existence of optimal solutions. We discuss an alternate model, where the leader hedges against optimal lower-level solutions, and show that solvability can be guaranteed under weaker conditions in both, a deterministic and a stochastic setting. The approach is applied to a mechanical shape optimization problem in which the leader decides on an optimal material distribution to minimize a tracking-type cost functional, whereas the follower chooses forces from an admissible set to maximize a compliance objective. The material distribution is considered to be stochastically perturbed in the actual construction phase. Computational results illustrate the bilevel optimization concept and demonstrate the interplay of follower and leader in shape design and testing.

1 Introduction

Bilevel programs arise from the interplay of two decision makers on different levels of a hierarchy: The leader decides first and passes the upper-level decision to the

---

Extended author information available on the last page of the article

⚫ Springer

follower. Incorporating the leader's decision as a parameter, the follower then returns an optimal solution of the lower-level problem. The leader's outcome depends on both, their decision and the solution that is picked by the follower. While the first formulation of a bilevel problem dates back to a monograph on duopoly market models published in 1934 (cf. [37]), these problems have not received extensive attention until the 1970s (for more details, we refer to [11]).

In this paper, we study a class of pessimistic bilevel stochastic programs, where the lower level problem has a strictly convex quadratic objective function and a fixed feasible set. As an application, we study a mechanical shape optimization problem in which the leader (the designer) minimizes a tracking functional over the set of feasible material distributions, whereas the follower (the test engineer) chooses forces from an admissible set to maximize a compliance objective. The safety of the construction is then evaluated pessimistically with the choice of the worst possible response. Randomness comes into play via manufacturing errors that stochastically perturb the material parameters in the actual construction phase.

In what follows, let us briefly review related work. *Bilevel programs* are nonconvex, nondifferentiable and NP-hard (non-deterministic polynomial-time-hard) [1]. Moreover, conceptual difficulties arise if the lower level problem has more than a single optimal solution. In this setting, one typically considers the so-called optimistic formulation, where cooperation of the follower is assumed, or takes a pessimistic stance and hedges against the worst possible outcome [23]. It is well-known that pessimistic bilevel programs have weaker analytical properties than their optimistic counterparts. In general, the existence of optimal solutions to a pessimistic bilevel program can only be assured under restrictive conditions including weak analyticity for the lower level objective function and strong assumptions on the structure of the lower level feasible set (cf. [28, Theorem 4.1]). These difficulties can be overcome by considering a modified setting, where the leader hedges against solutions that are almost optimal for the lower level problem. Sufficient conditions for convergence of the modified optimal values to the original one have been established in [26]. A systematic analysis of more inner regularization techniques has been recently provided by Lignola and Morgan in [24,25].

A *bilevel stochastic program* arises if the problem depends on an additional random parameter, that only the follower can observe before making their decision. In contrast, the leader has to decide nonanticipatorily, but is aware of the underlying probability distribution. In this setting, the upper level objective function can be understood as a random variable, which allows the leader to base their decision on some statistical functional. The expected value is for instance considered in the very first paper on bilevel stochastic optimization [32]. In a linear setting, more general models incorporating a variety of convex risk measures have been recently studied in [2]. The control of a vibrating string with stochastic data has been investigated in [12] for the case of the excess-probability as a goal function. A level-set based approach for solving risk-averse structural topology optimization problems with random field loading and material uncertainty is given in [29].

Already in 2001, Christiansen et al. [5] studied a *stochastic bilevel programming perspective in shape optimization*. They assume that the lower level deals with the deformation of the structure for a given shape and given forces subject to different

constraints, while on the upper level the shape is decided based on an optimization of weight or a global stiffness measure. Assuming that the lower level is uniquely solvable, the authors provide sufficient conditions for the existence of optimal solutions and discuss algorithmic aspects. Herskovits et al. [21] reformulated an elastic shape optimization problem with constraints as a bilevel optimization problem. They investigate a contact problem with non-penetration constraints on the lower level and stress constraints on the upper level. In [39], Zuo investigated shape optimization of thin shells in car design as an optimistic bilevel optimization problem, where on the lower level the mass distribution along the body frame of the vehicle and on the upper level the shape of shell segments of the hull of the vehicle are optimized. Sinha et al. [36] recently presented a general overview on bilevel optimization also covering optimal design problems. In this context, they considered weight or cost optimization of a structure on the upper level and, on the lower level, the computation of displacements and stresses via minimization of the governing physical variational problem. To the best of our knowledge, pessimistic hierarchical optimization in shape optimization with an objective functional differing from the physical energy of the system on the upper two levels has not been investigated so far.

The approach presented here is based on our previous work in [8–10], which grew out of the aspiration to mobilize methodology from mainly economy-driven decision making under (stochastic) uncertainty in order to study PDE-constrained optimization with an emphasis on engineering-related topics such as shape optimization. The risk-neutral models and models with risk aversion in the objective or the constraints were treated with the classical expectation, with risk measures, or by invoking comparisons using stochastic dominance relations. In the spirit of this experience, the present paper is heading for models with the above bilevel features coming to the fore in the presence of uncertainty.

The present work is organized as follows: In Sect. 2, we introduce a bilevel programming formulation, and in Sect. 3, the extension to a bilevel problem under stochastic uncertainty, which will be placed in the context of elastic shape optimization later in the paper. Based on this, we analyze both problem formulations and investigate their solvability. The application to a mechanical shape optimization problem via discrete shells is considered in Sect. 4 as well as its numerical optimization and the results of our numerical analysis. Finally, in Sect. 5, we draw conclusions and discuss possible future extensions of our work.

2 Bilevel problem formulation

Before formally introducing the bilevel problem, we briefly present the key objects. At the lowest level, $y[u, f]$ is the elastic displacement of the discrete shell which depends nonlinearly on the material parameters u and linearly on the applied forces f. The lower-level optimal solution set $\Psi[u]$, depending on the material parameter u, is the set of values of f which maximize a quadratic functional in y. In the upper-level of our pessimistic bilevel problem, we finally minimize the worst-case cost J of the lower-level optimization with respect to the material parameters u.

In detail, this pessimistic bilevel problem reads as

$$\min_{u \in \mathcal{U}} \left\{ \max_{f \in \Psi[u]} J[u, f] \right\}, \tag{1}$$

where $\mathcal{U} \subseteq (0, \infty)^n$ is a nonempty closed set, and $J : \mathcal{U} \times \mathbb{R}^N \to \mathbb{R}$ denotes the cost functional of the leader, which we assume to be continuous. In our application, this will be a tracking-type objective for a discrete shell with thickness/stiffness parameters u in an admissible set $\mathcal{U}$ and applied forces f. Moreover, we let the lower level optimal solution set mapping $\Psi : \mathcal{U} \rightrightarrows \mathbb{R}^N$ be given by

$$\Psi[u] := \arg\max_{f \in \mathcal{F}} \left\{ y[u, f]^\top H[u] y[u, f] \right\} \tag{2}$$

with a nonempty, low-dimensional, convex and compact set of admissible forces $\mathcal{F} \subset \mathbb{R}^N$, a function $H : \mathbb{R}^n \to \mathbb{R}^{N \times N}$ such that the the restriction $H|_\mathcal{U}$ is continuous and takes values in the cone of symmetric positive definite matrices $\mathcal{S}_{++}^N$. Throughout the paper, the notation $g : X \rightrightarrows Y$ is used for a multifunction g that maps the elements of some set X to subsets of some set Y. The displacement y depends on a vector u of thickness/stiffness parameters and the forces f. In fact, the mapping $y : \mathcal{U} \times \mathbb{R}^N \to \mathbb{R}^N$ in (2) is defined by the condition

$$\{y[u, f]\} = \arg\min_{y \in \mathbb{R}^N} \left\{ \frac{1}{2} y^\top H[u] y - y^\top M f \right\} \tag{3}$$

for some fixed matrix $M \in \mathcal{S}_{++}^N$, where uniqueness follows from $H[u] \in \mathcal{S}_{++}^N$. In our application, we consider a discrete shell with n triangular facets subject to a force distribution f in a set of admissible forces in $\mathbb{R}^N$, with N being three times the number of vertices. For this case, the elastic displacement $y[u, f]$ is given as the minimizer of the total free energy of a linearized elasticity model with $H[u]$ denoting the Hessian of an originally nonlinear elastic energy and M the mass matrix for the discrete reference shell.

The above hierarchical problem Eqs. (1–3) can also be understood as a three-level program. However, as $H[u] \in \mathbb{R}^{N \times N}$ is symmetric and positive definite for any admissible material parameter $u \in \mathcal{U}$, the third-level problem in Eq. (3) is uniquely solvable. Invoking first-order optimality conditions, we obtain the explicit representation

$$y[u, f] = H[u]^{-1} M f. \tag{4}$$

Plugging this solution into the lower level problem yields a bilevel problem. Moreover, Eq. (4) leads to a simple expression for the lower level optimal value function $\psi : \mathcal{U} \to \mathbb{R}$,

$$\psi[u]:=\max_{f\in\mathcal{F}}\left\{f^{\top}MH[u]^{-1}Mf\right\}, \tag{5}$$

and to the reformulation of the definition of Ψ in Eq. (2) as

$$\Psi[u]=\left\{f\in\mathcal{F}\mid f^{\top}MH[u]^{-1}Mf=\psi[u]\right\}. \tag{6}$$

Lemma 1 *The lower level optimal value function ψ defined by (5) is well-defined and continuous. In addition, the multifunction Ψ is closed.*

Proof For fixed u, the argument in (5) is quadratic in f, and in particular continuous. Since $\mathcal{F}$ is nonempty and compact, the maximum exists.

For any $f\in\mathcal{F}$, the argument in (5) is a continuous function of u. Moreover, ψ is the pointwise supremum of a familiy of continuous functions and thus lower semicontinuous by [34, Proposition 1.26(a)].

We assume that ψ is not upper semicontinuous, which yields $u_*\in\mathcal{U}$ and a sequence $u_j\to u_*$ such that $\lim_{j\to\infty}\psi[u_j]>\psi[u_*]$. For any j, we choose $f_j\in\mathcal{F}$ such that $\psi[u_j]=f_j^{\top}MH[u_j]^{-1}Mf_j$. Passing to a subsequence, we can assume that f_j converges to some $f_*\in\mathcal{F}$, since $\mathcal{F}$ is compact. By continuity of $f^{\top}MH[u]^{-1}Mf$ we obtain

$$\psi[u_*]\geq f_*^{\top}MH[u_*]^{-1}Mf_* = \lim_{j\to\infty}f_j^{\top}MH[u_j]^{-1}Mf_j = \lim_{j\to\infty}\psi[u_j],$$

a contradiction. Hence ψ is continuous.

Then the graph of the solution set mapping

$$\operatorname{gph}\Psi = \left\{[u,f]\in\mathcal{U}\times\mathcal{F}\mid\psi[u]-f^{\top}MH[u]^{-1}Mf=0\right\}$$

is the intersection of the closed set $\mathcal{U}\times\mathcal{F}$ and the level set of a continuous function and thus closed. $\qquad\square$

Proposition 1 *The mapping $\Phi:\mathcal{U}\to\mathbb{R}$ defined by*

$$\Phi[u]:=\max_{f\in\Psi[u]}J[u,f]$$

is well-defined and upper semicontinuous. Moreover, Φ is continuous at any $u\in\mathcal{U}$ for which $\Psi[u]$ is a singleton.

Proof For any fixed $u\in\mathcal{U}$, by (6) and Lemma 1 the lower-level set solution mapping $\Psi[u]$ is a nonempty, closed subset of the compact set $\mathcal{F}$, and hence compact. Thus, Φ is well-defined by continuity of the upper-level cost functional J on $\mathcal{U}\times\mathbb{R}^N$. Consider any sequence $\{u_k\}_{k\in\mathbb{N}}\subseteq\mathcal{U}$ that converges to $u\in\mathcal{U}$. By the previous considerations, there exists a sequence $\{f_k\}_{k\in\mathbb{N}}$ such that $[u_k,f_k]\in\operatorname{gph}\Psi$ and

$$\Phi[u_k]=J[u_k,f_k]$$

holds for any $k \in \mathbb{N}$. As $\mathcal{F}$ is compact, we may assume without loss of generality that $\{f_k\}_{k \in \mathbb{N}}$ converges to some $f \in \mathcal{F}$. By Lemma 1, we have $[u, f] \in \mathrm{gph}\, \Psi$ and thus

$$\limsup_{k \to \infty} \Phi[u_k] = \limsup_{k \to \infty} J[u_k, f_k] = J[u, f] \leq \max_{\bar{f} \in \Psi[u]} J[u, \bar{f}] = \Phi[u].$$

Hence, Φ is upper semicontinuous. If $\Psi[u]$ is a singleton, we also have

$$\liminf_{k \to \infty} \Phi[u_k] = \liminf_{k \to \infty} J[u_k, f_k] = J[u, f] = \Phi[u],$$

which completes the proof. $\qquad\square$

Remark 1 To understand the significance of Proposition 1 it is useful to compute these quantities explicitly in a simple low-dimensional example. Assume $n = 1$, $N = 2$, $\mathcal{U} = [\frac{1}{2}, \frac{3}{2}]$, $M = Id$, $H[u] = \begin{pmatrix} 1 & u-1 \\ u-1 & 1 \end{pmatrix}^{-1}$, $\mathcal{F} = [-1, 1] \times [0, 1]$. Then one computes $f^T M H[u]^{-1} M f = f_1^2 + f_2^2 + 2(u-1) f_1 f_2$. Maximizing this quantity as in (5) we see that only the two points $\{\pm 1, 1\}$ of $\mathcal{F}$ are relevant, and in particular $\psi[u] = 2 + 2|u-1|$. Further, from (2) (or, equivalently, (6)) we obtain the set of extremal forces

$$\Psi[u] = \begin{cases} \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} & \text{if} \quad u > 1, \\[2ex] \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\} & \text{if} \quad u < 1, \\[2ex] \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\} & \text{if} \quad u = 1. \end{cases}$$

Choosing for example $J[u, f] = u f_1$ one obtains

$$\Phi[u] = \begin{cases} J[u, 1] = u & \text{if} \quad u \geq 1, \\ J[u, -1] = -u & \text{if} \quad u < 1. \end{cases}$$

In particular, it is clear that on $\{u \neq 1\}$ the set-valued function Ψ is a singleton and Φ is continuous, whereas $\Psi[1]$ contains two elements and Φ is not continuous, and not lower semicontinuous, at $u = 1$.

As this example shows, Φ arises as the objective function of a pessimistic bilevel program, where the lower level problem may have more than a single optimal solution and can thus not be expected to be lower semicontinuous in general, a fact that was already observed in [11, example on pages 30–31]. Note that this may prevent the bilevel program (1) from having an optimal solution even if $\mathcal{U}$ is compact.

To overcome the difficulties detailed above, we consider a model where the leader also hedges against η-optimal lower level solutions (cf. [24]). Specifically, we replace

Ψ with the mapping $\Psi_\eta : \mathcal{U} \rightrightarrows \mathbb{R}^N$ defined by

$$\Psi_\eta[u] := \left\{ f \in \mathcal{F} \mid \psi[u] - f^\top M H[u]^{-1} M f < \eta \right\}$$

for some positive constant η. This results in the modified upper level problem

$$\min_{u \in \mathcal{U}} \left\{ \sup_{f \in \Psi_\eta[u]} J[u, f] \right\}. \tag{7}$$

As $\Psi[u] \subseteq \Psi_\eta[u]$ holds for any $\eta > 0$ and $u \in \mathcal{U}$, the optimal value in (7) yields an upper bound for the optimal value in (1).

Proposition 2 *The mapping $\Phi_\eta : \mathcal{U} \to \mathbb{R}$ defined by*

$$\Phi_\eta[u] := \sup_{f \in \Psi_\eta[u]} J[u, f]$$

is well-defined and lower semicontinuous for any $\eta > 0$. In particular, (7) is solvable whenever $\mathcal{U}$ is nonempty and compact.

Proof First, note that Φ_η is well-defined and real-valued as, by continuity of J, for any $u \in \mathcal{U}$

$$\Phi_\eta[u] \leq \max_{f \in \mathcal{F}} J[u, f] < \infty.$$

To prove semicontinuity, we consider a sequence $\{u_k\}_{k \in \mathbb{N}} \subseteq \mathcal{U}$ converging to some $u_* \in \mathcal{U}$. We select a sequence $\{f_j\}_{j \in \mathbb{N}} \subseteq \Psi_\eta[u_*]$, i.e. $\psi(u_*) < \eta + f_j^\top H^{-1}[u_*] f_j$, such that $J[u_*, f_j] \to \Phi_\eta[u_*]$. By continuity of ψ and H, there is K_j such that for all $k \geq K_j$ we have $\psi(u_k) < \eta + f_j^\top H^{-1}[u_k] f_j$, which is the same as $f_j \in \Psi_\eta[u_k]$. Therefore

$$J[u_*, f_j] = \lim_{k \to \infty} J[u_k, f_j] \leq \liminf_{k \to \infty} \Phi_\eta[u_k].$$

Since j was arbitrary, taking the limit $j \to \infty$ we conclude

$$\Phi_\eta[u_*] = \lim_{j \to \infty} J[u_*, f_j] \leq \liminf_{k \to \infty} \Phi_\eta[u_k].$$

$\square$

Remark 2 In [26], the alternate model

$$\min_{u \in \mathcal{U}} \left\{ \max_{f \in \bar{\Psi}_\eta[u]} J[u, f] \right\}$$

with

$$\bar{\Psi}_\eta[u] := \left\{ f \in \mathcal{F} \mid \psi[u] - f^\top M H[u]^{-1} M f \leq \eta \right\}$$

is considered. Under the present assumptions it can be shown that

$$\lim_{\eta \downarrow 0} \inf_{u \in \mathcal{U}} \left\{ \max_{f \in \bar{\Psi}_\eta[u]} J[u, f] \right\} = \inf_{u \in \mathcal{U}} \left\{ \max_{f \in \bar{\Psi}[u]} J[u, f] \right\}.$$

However, the function

$$\bar{\Phi}_\eta[u] := \sup_{f \in \bar{\Psi}_\eta[u]} J[u, f]$$

is not lower semicontinuous in general, which is why we rather use formulation (7).

3 Stochastic model

A bilevel stochastic program arises if a random vector enters the upper or lower levels as a parameter, with the information constraint that only the follower can observe the realization of the randomness before making their decision. In contrast, the leader has to decide nonanticipatorily, but is aware of the distribution of the randomness, which is independent of the leader's decision.

In the following, we shall study a setting where the leader's decision u is subject to a random perturbation. To become more specific, let $\Upsilon : \Omega \to \mathbb{R}^n$ be a random vector (i.e., a $\mathcal{B}$-Borel measurable function) on some probability space $(\Omega, \mathcal{B}, \mathbb{P})$. We obtain the following pattern of decision and observation:

Leader decides $u \to$ Realization of $\Upsilon \to$ Follower decides f.

In our model, the randomness results from manufacturing errors and has the following effect: Throughout (1)-(3), the leader's decision u is replaced with the perturbed material vector $u \odot \upsilon$, where $\odot$ denotes the componentwise multiplication and υ is the realization of Υ. In this setting, the leader seeks to ensure that the resulting material parameters are feasible regardless of the realization of the randomness. In order for the perturbed material vector to be almost-surely admissible, the leader has to choose the design parameter u in the induced feasible set

$$\mathcal{U}_\Upsilon := \{u \mid u \odot \upsilon \in \mathcal{U} \ \ \forall \upsilon \in \mathrm{supp}\,\mu_\Upsilon\},$$

where $\mu_\Upsilon := \mathbb{P} \circ \Upsilon^{-1}$ is the induced Borel probability measure on $\mathbb{R}^n$. Note that the set $\mathcal{U}_\Upsilon$ is closed as the intersection of closed sets. Typically, we think of a situation where $\mathrm{supp}\,\mu_\Upsilon \subseteq [a, b]^n$ holds for some $0 < a < b$, possibly both close to 1.

We will consider the stochastic extension of the classical pessimistic bilevel program (1)–(3) as well as the modified version (7). In both situations, we will assume the following assumption:

(A1) The support of μ_Υ is bounded.

In the classical setting, we will need the following additional assumptions:

(A2) $\mathcal{F}$ is a nonempty, bounded polyhedron, i.e. the convex hull of its nonempty and finite set of extreme points $\mathcal{P} \subseteq \mathcal{F}$.

(A3) μ_Υ is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}^n$.

(A4) There exists an open and connected set $\tilde{\mathcal{U}} \subseteq \mathbb{R}^n$, such that $\mathcal{U} \subseteq \tilde{\mathcal{U}}$, $H|_{\tilde{\mathcal{U}}}$ is real analytical, and it takes values in a closed subset of $\mathcal{S}_{++}^N$.

From the leader's point of view, the material vector that will be passed down to the lower level after the stochastic perturbation has occurred can be understood as a random vector $u \odot \Upsilon : \mathcal{U} \odot \Omega \to \mathbb{R}^n$ which is parameterized by the decision u. Similarly, the upper level outcome is a random variable $\Phi[u \odot \Upsilon] \in L^0(\Omega, \mathcal{B}, \mathbb{P})$ for any fixed u by Proposition 1. Here and in the subsequent analysis, we denote the associated classical L^p-spaces with $p \in [1, \infty]$ by $L^p(\Omega, \mathcal{B}, \mathbb{P})$ and use $L^0(\Omega, \mathcal{B}, \mathbb{P})$ for the space of real-valued measurable functions.

Theorem 1 *Assume (A1)-(A4), then the mapping* $\mathbb{F} : \mathcal{U}_\Upsilon \to L^\infty(\Omega, \mathcal{B}, \mathbb{P})$ *given by*

$$\mathbb{F}[u] := \Phi[u \odot \Upsilon]$$

is well-defined and continuous with respect to any L^p-norm with $p \in [1, \infty)$.

The proof of Theorem 1 requires some preliminary work.

Lemma 2 *Assume (A3) and (A4), then the set of discontinuities of Φ is a Lebesgue null set.*

Proof As the lower level goal function is strictly convex, we have $\Psi[u] \subseteq \mathcal{P}$ for any $u \in \mathcal{U}$. By (A4), for any pair $(f, \tilde{f}) \in \mathcal{P} \times \mathcal{P}$ the function $G_{(f,\tilde{f})} : \tilde{\mathcal{U}} \to \mathbb{R}$ defined by

$$G_{(f,\tilde{f})}[u] := f^\top M H[u]^{-1} M f - \tilde{f}^\top M H[u]^{-1} M \tilde{f}$$

is well-defined and real analytic. Consequently, the set

$$B[f, \tilde{f}] := \left\{ u \in \mathcal{U} \mid f, \tilde{f} \in \Psi[u] \right\} \subseteq \left\{ u \in \mathcal{U} \mid G_{(f,\tilde{f})}[u] = 0 \right\}$$

of parameters for which f and $\tilde{f}$ are optimal for the lower level problem is a Lebesgue null set, or we have

$$G_{(f,\tilde{f})}[u] = 0$$

for any $u \in \mathcal{U}$ by [30, Proposition 1].

Now we start from the case that $B[f, \tilde{f}]$ is a Lebesgue null set for any $f, \tilde{f} \in \mathcal{P}$ satisfying $f \neq \tilde{f}$. Let $u \in \mathcal{U}$. If $\Psi[u]$ is a singleton, then by Proposition 1, Φ is continuous at u. Consequently, the set of discontinuity points of Φ is contained in

$$\bigcup_{f, \tilde{f} \in \mathcal{P}, \, f \neq \tilde{f}} B[f, \tilde{f}],$$

which is a Lebesgue null set by the previous considerations.

To take care of the general case, let us consider the following relation on $\mathcal{P} \times \mathcal{P}$:

$$f \sim \tilde{f} :\Leftrightarrow G_{(f, \tilde{f})}[u] = 0 \quad \text{for any } u \in \mathcal{U}.$$

It is easy to verify that $\sim$ defines an equivalence relation and that the equivalence class of any extreme point $\tilde{f} \in \mathcal{P}$ is given by

$$E[\tilde{f}] := \left\{ f \in \mathcal{P} \mid G_{(f, \tilde{f})}[u] = 0 \quad \forall u \in \mathcal{U} \right\}.$$

By (6), $E[\tilde{f}] \subseteq \Psi[u]$ if $\tilde{f} \in \Psi[u] \cap \mathcal{P}$. Let $\tilde{\mathcal{P}} \subseteq \mathcal{P}$ contain exactly one element from each equivalence class, then Φ admits the representation

$$\Phi[u] = \max_{\tilde{f} \in \tilde{\mathcal{P}} \cap \Psi[u]} \left\{ \max_{f \in E[\tilde{f}]} J[u, f] \right\}.$$

As $\mathcal{P}$ is finite, for any $\tilde{f} \in \tilde{\mathcal{P}}$ the mapping

$$u \mapsto \max_{f \in E[\tilde{f}]} J[u, f]$$

is continuous. By the same argument as in the proof of Proposition 1, Φ is continuous on each set

$$S[\tilde{f}] := \left\{ u \in \mathcal{U} \mid \{\tilde{f}\} = \Psi[u] \cap \tilde{\mathcal{P}} \right\}$$

of parameters for which $\tilde{f}$ is the only representative that is optimal for the lower level problem. Thus, the set of discontinuities of Φ is contained in the set

$$N_B := \bigcup_{f, \tilde{f} \in \tilde{\mathcal{P}}, \, f \neq \tilde{f}} B[f, \tilde{f}],$$

which is a Lebesgue null set by construction of $\tilde{\mathcal{P}}$. For later reference we remark that we obtained

$$\mathcal{U} = N_B \cup \bigcup_{\tilde{f} \in \tilde{\mathcal{P}}} S[\tilde{f}] \quad \text{with } \mathcal{L}^n(N_B) = 0, \tag{8}$$

and that the sets N_B and $S[\tilde{f}]$ for $\tilde{f} \in \tilde{\mathcal{P}}$ in the right-hand side of (8) are pairwise disjoint. $\qquad\square$

Throughout the subsequent analysis, we will use the notation introduced in the proof of Lemma 2.

Proof of Theorem 1 Let $u \in \mathcal{U}_\Upsilon$. As any upper semicontinuous function is Borel measurable, $\mathbb{F}[u] \in L^0(\Omega, \mathcal{B}, \mathbb{P})$ follows directly from Proposition 1. Moreover, we have

$$\text{ess sup } \mathbb{F}[u] \le \max_{f \in \mathcal{F}} \sup_{v \in \text{supp}\,\mu_\Upsilon} J[u \odot v, f] < \infty \qquad (9)$$

by continuity of J, (A1) and (A3).

Consider any sequence $\{u_k\}_{k \in \mathbb{N}} \subseteq \mathcal{U}_\Upsilon$ that converges to some $u \in \mathcal{U}_\Upsilon$. We write

$$\lim_{k \to \infty} \|\mathbb{F}[u] - \mathbb{F}[u_k]\|^p_{L^p(\Omega, \mathcal{B}, \mathbb{P})} = \lim_{k \to \infty} \int_{\text{supp}\,\mu_\Upsilon} |\Phi[u \odot v] - \Phi[u_k \odot v]|^p \, \mu_\Upsilon(dv).$$

The set $\{u\} \cup \{u_k \mid k \in \mathbb{N}\}$ is compact, so that by continuity of J and (9) we obtain a uniform bound on the integrand. With (A1) and dominated convergence we see that it suffices to prove pointwise convergence almost everywhere.

Let N_B be as in the proof above, and consider the set

$$\hat{N}_B := \{v \in \mathcal{U}_\Upsilon \mid u \odot v \in N_B\}.$$

By the change-of-variables formula we obtain $0 = \mathcal{L}^n(N_B) = \prod_{i=1}^n u_i \mathcal{L}^n(\hat{N}_B)$ and, since $u_i > 0$ for all i, $\mathcal{L}^n(\hat{N}_B) = 0$. By (A2), $\mu_\Upsilon(\hat{N}_B) = 0$.

Fix some $v \in \mathcal{U}_\Upsilon \setminus \hat{N}_B$. Then by (8) we have $u \odot v \in S[\tilde{f}]$ for some $\tilde{f} \in \tilde{\mathcal{P}}$, so that in particular Φ is continuous at $u \odot v$. From $u_k \to u$ with $k \to \infty$ we obtain $u_k \odot v \to u \odot v$ and therefore $\Phi[u_k \odot v] - \Phi[u \odot v] \to 0$. This proves pointwise convergence almost everywhere and concludes the proof. $\qquad\square$

Remark 3 The assertion of Theorem 1 does not hold for $p = \infty$. To see this, we consider the example of Remark 1 and extend it to the stochastic setting taking $\Omega = [\frac{9}{10}, \frac{11}{10}]$, $\mathbb{P}$ proportional to the Lebesgue measure restricted to Ω, and Υ to be the identity, so that $\text{supp}\,\mu_\Upsilon = \Omega$. Then $\mathbb{F}[u](v) = \Phi[uv] = \pm uv$, with the positive sign if and only if $uv \ge 1$. For the sequence $u_k := 1 - \frac{1}{k} \to 1$ we have $\mathbb{F}[u_k](v) = u_k v$ for $v \ge 1/u_k$, and $\mathbb{F}[u_k](v) = -u_k v$ for $v < 1/u_k$. In particular for all $v \in (1, 1/u_k)$ we have $\mathbb{F}[u_k](v) - \mathbb{F}[1](v) = -u_k v - v$. Taking the supremum over all such v we obtain $\|\mathbb{F}[u_k] - \mathbb{F}[1]\|_{L^\infty(\Omega, \mathcal{B}, \mathbb{P})} \ge 1 + \frac{1}{u_k} \to 2$, hence $\mathbb{F}[u_k]$ does not converge to $\mathbb{F}[1]$ in $L^\infty(\Omega, \mathcal{B}, \mathbb{P})$.

As a generic first choice, the leader might assess the random upper level cost based on its expected value, i.e. consider the risk neutral bilevel stochastic program

$$\min_{u \in \mathcal{U}_\Upsilon} \{\mathbb{E}[\mathbb{F}[u]]\}, \qquad (10)$$

which is well-defined by Theorem 1. More in general, to allow for varying degrees of risk aversion, we take into account a mapping $\mathcal{R} : \mathcal{X} \to \mathbb{R}$ with

$$L^\infty(\Omega, \mathcal{B}, \mathbb{P}) \subseteq \mathcal{X} \subseteq L^0(\Omega, \mathcal{B}, \mathbb{P})$$

and consider the bilevel stochastic program

$$\min_{u \in \mathcal{U}_\Upsilon} \{\mathcal{R}\,[\mathbb{F}[u]]\}. \tag{11}$$

$\mathcal{R}$ will typically be a **monetary risk measure** in the sense of [15, Definition 4.1] meaning it satisfies the following conditions:

– Monotonicity: $\mathcal{R}[Y_1] \leq \mathcal{R}[Y_2]$ for all $Y_1, Y_2 \in \mathcal{X}$ satisfying $Y_1 \leq Y_2$ $\mathbb{P}$-almost surely.
– Translation equivariance: $\mathcal{R}[Y + m] = \mathcal{R}[Y] + m$ for all $Y \in \mathcal{X}$ and $m \in \mathbb{R}$.

Moreover, we will assume the following:

(A5) $\mathcal{R} : L^p(\Omega, \mathcal{B}, \mathbb{P}) \to \mathbb{R}$ with some $p \in [1, \infty)$ is convex and nondecreasing as defined above.

Remark 4 (A5) holds for any **convex risk measure** in the sense of [13] and [14], *i.e.* for any monetary risk measure that is convex. In particular, this includes the expectation, the mean-upper semideviation of any order and the Conditional Value-at-Risk. However, as we do not assume translation equivariance, the assumption is also fulfilled for the expected excess of arbitrary order (cf. [35, Chapter 6]).

The following result is well-known in the literature, see for example [4, Theorem 4.1]. For the convenience of the reader, we provide a short self-contained proof.

Lemma 3 *Assume (A5), then the mapping $\mathcal{R}$ is continuous.*

Proof For $f \in L^p(\Omega, \mathcal{B}, \mathbb{P})$ we denote by $|f| \in L^p(\Omega, \mathcal{B}, \mathbb{P})$ the function obtained taking the pointwise absolute value, so that $f \leq |f|, -f \leq |f|$ $\mathbb{P}$-almost everywhere.

It suffices to prove that $\mathcal{R}$ is continuous in 0, and we can assume that $\mathcal{R}(0) = 0$ (otherwise we replace $\mathcal{R}$ by $\hat{\mathcal{R}}(f) := \mathcal{R}(g_* + f) - \mathcal{R}(g_*)$). If $\mathcal{R}$ is not continuous, there is $\delta > 0$ such that for any j there is $f_j \in L^p(\Omega, \mathcal{B}, \mathbb{P})$ with $\|f_j\|_{L^p(\Omega, \mathcal{B}, \mathbb{P})} < 4^{-j}$ and $|\mathcal{R}(f_j)| \geq \delta$. By convexity, $0 = \mathcal{R}(0) \leq \frac{1}{2}\mathcal{R}(f_j) + \frac{1}{2}\mathcal{R}(-f_j)$, which implies $\mathcal{R}(-f_j) \geq -\mathcal{R}(f_j)$. By monotonicity,

$$\mathcal{R}(|f_j|) \geq \max\{\mathcal{R}(f_j), \mathcal{R}(-f_j)\} \geq \max\{\mathcal{R}(f_j), -\mathcal{R}(f_j)\} = |\mathcal{R}(f_j)| \geq \delta.$$

Let $f_* := \sum_j 2^j |f_j| \in L^p(\Omega, \mathcal{B}, \mathbb{P})$. Using first monotonicity and then convexity, we obtain $\mathcal{R}(f_*) \geq \mathcal{R}(2^j |f_j|) \geq 2^j \mathcal{R}(|f_j|) \geq 2^j \delta$ for any j, which contradicts the boundedness of $\mathcal{R}(f_*)$. □

Theorem 2 *Assume (A1)–(A5), then the function $\mathcal{Q}_\mathcal{R} : \mathcal{U}_\Upsilon \to \mathbb{R}$ defined by*

$$\mathcal{Q}_\mathcal{R}[u] := \mathcal{R}\,[\mathbb{F}[u]] = \mathcal{R}\,[\Phi[u \odot \Upsilon]]$$

is continuous. In particular, the bilevel stochastic problem (11) has an optimal solution whenever the induced feasible set $\mathcal{U}_\Upsilon$ is nonempty and compact.

Proof As $\mathcal{R}$ is continuous by Lemma 3, the result follows from Theorem 1. $\quad\square$

Let us now consider the stochastic version of the modified problem (7), where the leader hedges against all η-optimal lower level solutions. For this, we will use the notion of law-invariant risk measure:

$$\mathcal{R}[Y_1] = \mathcal{R}[Y_2] \quad \text{for all } Y_1, Y_2 \in \mathcal{X} \text{ with } \mathbb{P} \circ Y_1^{-1} = \mathbb{P} \circ Y_2^{-1},$$

i.e. for all Y_1, Y_2 which induce the same Borel probability measure. The following existence result is obtained for law-invariant, convex risk measures under weaker assumptions, where we no longer restrict the analysis to polyhedral $\mathcal{F}$ and real analytic H.

Theorem 3 *Assume (A1) and (A5) and let $\mathcal{R}$ be translation equivariant as well as law-invariant. Then the mapping $\mathcal{Q}_{\mathcal{R},\eta} : \mathcal{U}_\Upsilon \to \mathbb{R}$ given by*

$$\mathcal{Q}_{\mathcal{R},\eta}[u] := \mathcal{R}\left[\Phi_\eta\left[u \odot \Upsilon\right]\right]$$

is well-defined and lower semicontinuous. In particular, the bilevel stochastic program

$$\min_{u \in \mathcal{U}_\Upsilon} \left\{\mathcal{Q}_{\mathcal{R},\eta}[u]\right\}$$

is solvable, whenever $\mathcal{U}_\Upsilon$ is nonempty and compact.

Proof First, note that $\Phi_\eta\left[u \odot \Upsilon\right] \in L^0(\Omega, \mathcal{B}, \mathbb{P})$ and

$$\left\|\Phi_\eta\left[u \odot \Upsilon\right]\right\|_{L^\infty(\Omega,\mathcal{B},\mathbb{P})} \leq \sup_{\upsilon \in \mathrm{supp}\,\mu_\Upsilon} \Phi_\eta\left[u \odot \upsilon\right] \leq \sup_{\upsilon \in \mathrm{supp}\,\mu_\Upsilon} \sup_{f \in \mathcal{F}} J[u \odot \upsilon, f] < \infty$$

hold for any $u \in \mathcal{U}_\Upsilon$ by Proposition 2 and (A1). Thus, $\mathcal{Q}_{\mathcal{R},\eta}$ is well-defined.

Let $\mathcal{R}^*$ denote the convex conjugate of $\mathcal{R}$ (cf. [22, Theorem 2.1]), then $\mathcal{R}$ admits a robust representation as

$$\mathcal{R}[\mathcal{Y}] = \sup_{\mathbb{P}' \in \mathrm{Env}} \left\{\mathbb{E}_{\mathbb{P}'}[\mathcal{Y}] - \mathcal{R}^*[\mathbb{P}']\right\} \quad \forall \mathcal{Y} \in L^p(\Omega, \mathcal{B}, \mathbb{P}),$$

where the risk envelope Env is a subset of the normed positive part of the dual space of $L^p(\Omega, \mathcal{B}, \mathbb{P})$ by [22, Corollary 2.3, Theorem 2.4]. Fix any $\mathbb{P}' \in$ Env. With a slight abuse of notation, we shall identify $\mathbb{P}'$ with the $\mathbb{P}$-continuous probability measure $d\mathbb{P}'/d\mathbb{P}$ and show that the mapping $u \mapsto \mathbb{E}_{\mathbb{P}'}\left[\Phi_\eta\left[u \odot \Upsilon\right]\right]$ is lower semicontinuous. The result then follows because the pointwise supremum of lower semicontinuous functions is lower semicontinuous (cf. Lemma 1 and [34, Proposition 1.26 (a)]).

Consider any sequence $\{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{U}_\Upsilon$ that converges to some $u \in \mathcal{U}_\Upsilon$. Without loss of generality, we assume that $u_k \in B_1(u) \cap \mathcal{U}_\Upsilon$ holds for any $k \in \mathbb{N}$, where $B_1(u)$

denotes the open Euclidean unit ball around u (We denote its closure by $\overline{B}_1(u)$). By definition,

$$\Phi_\eta[u_k \odot \Upsilon] \geq \min_{\upsilon \in \text{supp} \mu_\Upsilon} \min_{u \in \overline{B}_1(u) \cap \mathcal{U}_\Upsilon} \min_{f \in \mathcal{F}} J[u \odot \upsilon, f] =: \underline{J}$$

holds for any $k \in \mathbb{N}$. As J is continuous and $\text{supp}\,\mu_\Upsilon$, $\overline{B}_1(u) \cap \mathcal{U}_\Upsilon$, and $\mathcal{F}$ are nonempty and compact, we have $\underline{J} \in \mathbb{R}$. Thus, Fatou's Lemma yields

$$\liminf_{k \to \infty} \mathbb{E}_{\mathbb{P}'}\left[\Phi_\eta[u_k \odot \Upsilon]\right] \geq \int_\Omega \liminf_{k \to \infty} \Phi_\eta[u_k \odot \Upsilon]\ \mathbb{P}'(d\omega) \geq \mathbb{E}_{\mathbb{P}'}\left[\Phi_\eta[u \odot \Upsilon]\right],$$

which completes the proof. □

4 Application: discrete shells

In this section, we will apply bilevel optimization to a mechanical shape optimization problem. Our aim is to determine the optimal elastic design of curved roof-type constructions. The leader in this setup is the construction engineer who aims at minimizing a tracking-type functional via optimizing the distribution of material on a prescribed roof geometry. Due to production errors, the material distribution is considered to be stochastically perturbed in the actual construction phase. The follower is a test engineer, who is performing a worst-case analysis and considers within a given set of possible forces—for example wind and roof load—those that maximize the compliance functional.

4.1 General setting and problem formulation

Our model problem is taken from the literature on geometric design [38], but our mechanical perspective is not self-supporting structures but instead architectural structures composed of discrete thin shells. Indeed, we model the mechanical properties of a roof construction using an adaptation of the discrete elastic shell model by Grinspun et al. [17], in which the geometry is a triangular surface and each triangle is considered as a construction panel, with joints at the edges. The membrane distortion deforms the individual panels, whereas the bending distortion leads to a change of the dihedral angle between pairs of panels that share an edge. Let us emphasize that the discrete shell approach is a design tool and does not act as a computational tool for the full elastostatic modeling in a later planning stage. In fact, we consider the discrete shell model mainly as a testbed for the proposed bilevel optimization approach. We underline this by reporting all physical quantities without units.

Comparing with the notation in the previous section, the design parameter u will represent the thickness of the shell, f the applied forces, and y the resulting displacement of the shell. The minimization in Eq. (3) then corresponds to the solution of a linear elasticity problem in (13), with $H[u]$ representing the elastic energy. The problem in (2) corresponds to the follower optimizing compliance. The leader's cost

functional J in (1) measures the deviation from the prescribed shape and is defined in (14) below.

We consider the simplicial mesh of a discrete shell $\mathcal{S}_h = (\mathcal{V}, \mathcal{E}, \mathcal{T})$ consisting of sets of vertices $\mathcal{V}$, edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ and triangular faces $\mathcal{T} \subset \mathcal{V} \times \mathcal{V} \times \mathcal{V}$. In what follows, we use maps defined on the different elements of such a mesh instead of vectors used in the theoretical considerations above. For example, a map $w : \mathcal{V} \to \mathbb{R}^k$ assigning each vertex a value in this section corresponds to a vector $\mathbb{R}^{k|\mathcal{V}|}$ from the previous sections and similarly for functions defined on edges and faces. We denote evaluations $w(v)$ of such a map also via indexing to simplify notation, *i.e.* $w_v := w(v) \in \mathbb{R}^k$.

The geometry of a discrete shell is given by a map $x : \mathcal{V} \to \mathbb{R}^3$ subject to the constraint that for each face there is no straight line in $\mathbb{R}^3$ containing all three vertices, *i.e.* no triangle degenerates to a line. Thus, each triangle $t \in \mathcal{T}$ with vertices v_0, v_1, v_2 can be parametrized over the reference triangle in $\mathbb{R}^2$ with vertices $(0, 0), (1, 0)$ and $(0, 1)$ via the affine map x_t interpolating $x(v_0), x(v_1), x(v_2)$. We denote by Dx_t the differential of this affine map for face t, so that the associated metric tensor in the same face is

$$G[x_t] := (Dx_t)^\top Dx_t.$$

We denote by $\hat{x} : \mathcal{V} \to \mathbb{R}^3$ the fixed stress-free reference configuration of the discrete shell, and parametrize the deformed configuration $x = \hat{x} + y$ in terms of the elastic displacement of the vertices $y : \mathcal{V} \to \mathbb{R}^3$. We denote by l_e the length of an edge $e \in \mathcal{E}$ and by a_t the area of a face $t \in \mathcal{T}$ in the reference configuration. Then, $a_e := \frac{1}{3}(a_t + a_{t'})$ is a corresponding edge-associated area, where t and t' are the two faces adjacent to the interior edge $e \in \mathcal{E}$; correspondingly $a_v := \frac{1}{3} \sum_{t \in \mathcal{T}_v} a_t$ a vertex-associated area for the ring of faces $\mathcal{T}_v$ around a vertex $v \in \mathcal{V}$.

The design variable is the material thickness parameter, which is assumed to be constant on each of the triangles and is denoted by $u : \mathcal{T} \to (0, \infty)$. In order to evaluate the bending contribution to the energy, see (12) below, we shall use on an interior edge e the averaged thickness $u_e := \frac{1}{2}(u_t + u_{t'})$ of the two triangles t and t' sharing the edge e.

Variational Formulation of Discrete Shells. In the modeling of thin shells, the elastic stored energy is typically the sum of two terms: the stored energy caused by in-plane membrane distortion and the stored energy reflecting bending distortion [7,27]. The two terms scale linearly and cubically, respectively, in the thickness of the shell.

For a displacement y, the Cauchy-Green strain tensor measuring the change of lengths, and consequently area, of a face t is given by

$$\mathcal{G}[y] := \big(G[\hat{x}_t]\big)^{-1} G[(\hat{x} + y)_t].$$

Then, the membrane energy depends on this tensor and is defined as

$$\mathcal{W}_{\text{mem}}[u, y] := \sum_{t \in \mathcal{T}} a_t\, u_t\, W_{\text{mem}}(\mathcal{G}[y]|_t),$$

where we use the neo-Hookean energy density

$$W_{\text{mem}}(A) := \frac{\mu}{2} \text{tr}\, A + \frac{\lambda}{4} \det A - \left(\mu + \frac{\lambda}{2} \right) \log \det A - \mu - \frac{\lambda}{4}.$$

The linearization of this energy coincides with the planar, isotropic, linearized elasticity model with Lamé-Navier coefficients μ and λ [6,27]. In the following, we use $\mu = \lambda = 1$.

For the bending energy, we follow [19] and use an adaptation of the discrete shell bending energy introduced in [17]. It measures the change of the dihedral angles between a pair of neighboring triangles t and t' due to the displacement y in the configuration x. The angle is computed as $\theta_e(x) := \arccos(n_t(x)^\top n_{t'}(x))$, where $n_t(x)$ and $n_{t'}(x)$ are the unit normals generated by the deformation x, and the energy takes the form

$$\mathcal{W}_{\text{bend}}[u, y] := \gamma \sum_{e \in \mathcal{E}} u_e^3 \cdot \frac{(\theta_e(\hat{x} + y) - \theta_e(\hat{x}))^2}{a_e} l_e^2 \tag{12}$$

for some constant $\gamma > 0$, which in continuum models can be expressed in terms of λ and μ. We use $\gamma = 1$.

The stored elastic energy $\mathcal{W}[u, y]$ is the sum of these two energies,

$$\mathcal{W}[u, y] := \mathcal{W}_{\text{mem}}[u, y] + \mathcal{W}_{\text{bend}}[u, y],$$

so that the total free energy in the presence of external forces $f : \mathcal{V} \to \mathbb{R}^3$ reads as

$$\mathcal{I}[u, f, y] = \mathcal{W}[u, y] - f^\top M y, \tag{13}$$

where M is a diagonal mass matrix in $\mathbb{R}^{3|\mathcal{V}| \times 3|\mathcal{V}|}$ with entries a_v at positions (i, i) with $i = 3j - k$ for $j = 1, \dots, |\mathcal{V}|$ and $k = 0, 1, 2$. The elastic displacements resulting from applying the forces to the reference configuration are the minimizers of this energy.

In what follows, we restrict ourselves to the linearization of this model. We denote by $H[u] := \partial_{yy}^2 \mathcal{W}[u, 0]$ the Hessian of the stored elastic energy, and obtain the linearized stored elastic energy

$$\mathcal{W}^{\text{lin}}[u, y] := \frac{1}{2} y^\top H[u] y$$

as well as the linearized total free energy

$$\mathcal{I}^{\text{lin}}[u, f, y] := \mathcal{W}^{\text{lin}}[u, y] - f^\top M y,$$

whose minimization corresponds to the innermost problem introduced in (3). Prescribing suitable boundary data $y_v = 0$ on a set of at least three vertices $v \in \mathcal{V}$, which do not lie on a line, one can deduce (cf. [19]) that $H[u]$ is a positive-definite

matrix. As written above expression (4), for every u and f the energy $\mathcal{I}^{\text{lin}}[u, f, \cdot]$ has a unique minimizer, which is also the unique solution of the associated Euler-Lagrange equation

$$0 = \partial_y \mathcal{I}^{\text{lin}}[u, f, y] = H[u]y - Mf.$$

The Optimization Problem. To complete our practical optimization problem, we need to specify the admissible set of material parameters $\mathcal{U}$, the admissible set of force parameters $\mathcal{F}$, and the cost functional of the leader J. The objective of the lower level optimal value function ψ is already completely defined in (5) and equals the compliance functional evaluated for the displacement $y[u, f]$, *i.e.*

$$\psi[u] = \max_{f \in \mathcal{F}} \left\{ f^\top My[u, f] \right\} = \max_{f \in \mathcal{F}} \left\{ f^\top MH[u]^{-1}Mf \right\}.$$

The admissible set of force parameters $\mathcal{F}$ is assumed to consist of linear combinations of a small number of different load scenarios. We assume that the forces are of the type $f = BF$, where $F \in \mathbb{R}^d$ for some $d \ll 3|\mathcal{V}|$ are the coefficients, and the columns B_j of the matrix $B \in \mathbb{R}^{3|\mathcal{V}| \times d}$ are the basis of the d-dimensional subspace of forces. Therefore, each $B_j \in \mathbb{R}^{3|\mathcal{V}|}$ represents a force distribution on the reference configuration $\hat{x}$ which is then scaled with $F_j \in \mathbb{R}$, for $j = 1, \ldots, d$. The components of these basis vectors could be determined, for example, from the location of the vertex or the inclination of the triangular faces sharing a vertex. Furthermore, we consider different constraints on the values of the scale factors F_j, *i.e.* we assume that the set $\mathcal{F}$ is given by $\bigcap_{k=1}^K \mathcal{F}_k$ with

$$\mathcal{F}_k := \left\{ BF \in \mathbb{R}^{3|\mathcal{V}|} \mid F \in \mathbb{R}^d, \ \mathcal{Q}_k^F(F) \geq 0 \right\}$$

for some smooth functions $\mathcal{Q}_k^F$ for $k = 1, \ldots, K$. For example, if $\mathcal{F}$ consists of the forces which fulfill $|F| \leq \mu$ then one might choose $d = 3|\mathcal{V}|$, $B = \text{Id}$, $K = 1$, and $\mathcal{Q}_1^F(F) = \mu^2 - |F|^2$.

In the problem of the leader, we constrain the material thickness parameter u elementwise from below and from above, and we assume that the total volume of material, determined via the discrete integral of u, is below some fixed positive parameter.

Lastly, the upper level cost functional is considered to be of tracking-type and measures the squared discrete L^2-norm of the displacement on a predefined tracking subset of the whole shell,

$$J[u, f] := y[u, f]^\top \chi \odot My[u, f] = \sum_{v \in \mathcal{V}} \chi_v M_{vv} |y_v[u, f]|^2. \tag{14}$$

Here $\chi : \mathcal{V} \to \{0, 1\}$ is a discrete characteristic function with value 1 at vertices in the tracking set and 0 elsewhere.

In the stochastic setting, we restrict ourselves to the expected value $\mathbb{E}[\mathbb{F}[u]]$ as the risk measure for the optimization (cf. (10)). Furthermore, the stochastic perturbation of the distribution of the thickness parameter u is given by i.i.d. normal

distributions for each parameter, *i.e.* we consider the perturbed material $u \odot \Upsilon$ for $\Upsilon \sim \mathcal{TN}(1, \sigma^2, \upsilon_{\min}, \upsilon_{\max})^{|\mathcal{T}|}$, where $\mathcal{TN}(1, \sigma^2, \upsilon_{\min}, \upsilon_{\max})$ is the truncated normal distribution with average 1 and standard deviation σ, truncated to the interval $[\upsilon_{\min}, \upsilon_{\max}]$. In practice, we take $\sigma \leq 0.2$, $\upsilon_{\min} = 10^{-2}$ and $\upsilon_{\max} = 2$, so that the truncation has little effect and σ is almost identical to the standard deviation of Υ.

We further fix constants $0 < u^- < u^+$ and $V^+ > 0$ and define implicitly $\mathcal{U}$ by the condition

$$\mathcal{U}_\Upsilon = \left\{ u : \mathcal{T} \to \mathbb{R} \mid u^- \leq u_t \leq u^+ \quad \forall t \in \mathcal{T}, \sum_{t \in \mathcal{T}} a_t u_t \leq V^+ \right\} \subset (0, \infty)^{|\mathcal{T}|}.$$

4.2 Numerical optimization

To numerically solve the bilevel problem (1) in the presented setting, it is convenient to replace the restriction of u and f to admissible sets $\mathcal{U}_\Upsilon$ and $\mathcal{F}$ by smooth approximations and then to deal with a differentiable problem. In our implementation, we achieve this by using logarithmic barrier functions, as commonly used in interior point methods (see e.g. textbook [31]). Hence, with the structural assumptions on the set of admissible forces introduced above, we define the smoothed follower problem by

$$\Psi_\alpha[u] := \underset{F \in \mathbb{R}^d}{\arg\max} \left\{ y[u, BF]^\top H[u] \, y[u, BF] + \alpha^F \sum_{k=1}^{K} \log\left(\mathcal{Q}_k^F(F)\right) \right\}, \quad (15)$$

where $\alpha^F > 0$ is an appropriate scaling factor for the barrier terms.

To compute the minimizers in (15), we do not aim at a global minimization approach but rather use an ascent method (see below) to compute isolated local minimizers. Thus, we assume in the numerical optimization of the leader problem, that the solution of the follower problem is of such type. This allows us to apply conventional nonlinear optimization algorithms. In this framework, the maximizer and the set Ψ_α be interchangeable. In the examples considered below, this assumption is justified by the use of asymmetric triangulations, and additionally by the symmetry-breaking random perturbations of the material thickness. Thus, the logarithmic barrier formulation of the expected value optimization problem for the leader is

$$\min_{u \in \mathbb{R}^{|\mathcal{T}|}} \left\{ \mathbb{E}\left[J\left[u \odot \Upsilon, \Psi_\alpha[u \odot \Upsilon] \right] \right] - \alpha^u \sum_{t \in \mathcal{T}} a_t \left(\log(u_t - u^-) + \log(u^+ - u_t) \right) \right.$$

$$\left. - \alpha^V \log\left(V^+ - \sum_{t \in \mathcal{T}} a_t u_t \right) \right\}$$

for scaling factors α^u, $\alpha^V > 0$ as before.

This regular reformulation of the optimization problem can be solved numerically using a stochastic gradient method. For PDE-constrained shape optimization problems under uncertainty, this method is analyzed in [16]. In our case, the smoothed follower
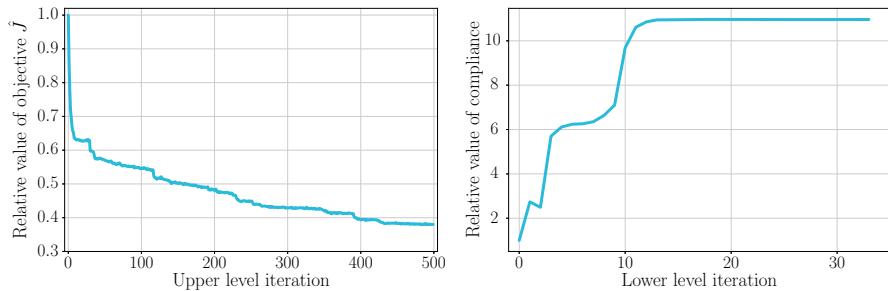
Fig. 1 Left: upper level relative cost values $\hat{J}[u^i]/\hat{J}[u^0]$ for the iterates of the stochastic gradient descent method in the example shown in the bottom row of Fig. 3. Right: corresponding lower level compliance cost $\dfrac{y[u,BF^j]^\top H[u]\, y[u,BF^j]}{y[u,BF^0]^\top H[u]\, y[u,BF^0]}$ for iterates of the Newton-type method for the follower problem in the first upper level descent step and the initial material distribution

problem is a deterministic and smooth optimization problem, and computing its first and second derivatives is straightforward. Thus, we use a Newton-type method with Armijo backtracking line search (cf. [31, Algorithm 3.2]) to compute its optimizers. The gradients of the smoothed bilevel problem can be computed via the general procedure of shape optimization calculus and thus, we employ stochastic gradient descent [33] to solve it. To this end, in each iteration of the descent algorithm, we draw finitely many samples $\upsilon^1, \ldots, \upsilon^K$ from the distribution of the material perturbation. In the experiments, we always chose $K = 128$. Using these samples, we approximate the expected value by the empirical risk $\hat{J}[u] := \frac{1}{K} \sum_{k=1}^{K} J\left[u \odot \upsilon^k, \Psi_\alpha[u \odot \upsilon^k]\right]$. Then a new iterate is computed by taking a step in the direction of the negative gradient of the combination of the empirical risk and the logarithmic barrier terms. Figure 1 depicts the decrease of the upper level cost functional over the iterations of the stochastic descent algorithm and the increase of the lower level compliance cost when solving the follower problem for the initial material distribution. Latter solves of the follower problem typically require 10 to 30 iterations of the Newton-type method per outer iteration.

We have implemented our method in C++ with the Geometric Optimization And Simulation Toolbox (GOAST) [20], where we use the Eigen library [18] for numerical linear algebra and CHOLMOD [3] from the SuiteSparse collection as direct linear solver. The code is available under https://gitlab.com/numod/bilevel-shape-optimization.

4.3 Numerical results

We applied the bilevel shape optimization method in a proof-of-concept study of discrete shells representing curved roofs. We fix an orientation so that the negative Z-axis is in the direction of gravity and the supporting ground is in the XY-plane. For each geometry, we fix a set of Dirichlet vertices near the ground plane, representing the points on which the structure is supported, and also fix the material thickness of the corresponding triangles. This removes these variables from the optimization.

The construction is exposed to two types of forces. First, there are forces emulating wind hitting the structure. For a given wind direction and strength, the force on each
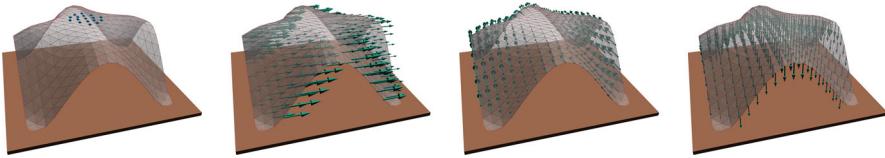
Fig. 2 The first panel shows the geometry of the roof structure, with the tracking set on the roof plateau marked with dots. The Dirichlet nodes are the vertices on the horizontal plane at the corners. The other three panels show the three basis force fields B_1 (horizontal wind in the X direction), B_2 (horizontal wind in the Y direction) and B_3 (vertical gravitational force caused by an overlay on the roof). The scale of the force arrows is arbitrary

part of the roof depends on the local orientation. We assume that the magnitude of the force on a vertex is proportional to the absolute value of the scalar product between the vertex normal (given as the average of the normals of the triangles adjacent to the vertex) and the wind direction. For simplicity, we only consider a two-dimensional subset of possible forces, spanned by the basis vectors B_1 and B_2 which represent wind along the positive X- and Y-axis, respectively. The direction and magnitude of the wind are then controlled by the scale factors F_1 and F_2. We fix a maximal magnitude of wind-type force $F_{\max,xy}$ and use the constraint function $\mathcal{Q}_1^F(F) := F_{\max,xy}^2 - \left(F_1^2 + F_2^2 \right)$ in (15). An example of these two basis vectors demonstrating the dependence on the orientation of the normal is shown in the second and third panels of Figure 2.

Second, we consider a vertical force, which could emulate the weight of snow or water overlay on the roof. The magnitude of the corresponding basis vector B_3 on each vertex is the absolute value of the scalar product between the vertex normal and the Z-axis and is shown in Fig. 2 on the far right. The magnitude of gravitational load is controlled by the scale factor F_3, we ensure that it is pointing downward via $\mathcal{Q}_2^F(F) := F_3$ and limit its magnitude via $\mathcal{Q}_3^F(F) := F_{\max,z} - F_3$, where $F_{\max,z}$ is the maximal magnitude of the gravitational force. Therefore the admissible set $\mathcal{F}$ is a cylinder with radius $F_{\max,xy}$ and height $F_{\max,z}$.

We performed most of our investigations on the simple roof geometry shown in Fig. 2. For this problem, the basic parameters, which are used in the examples if not indicated otherwise, are as follows. The roof geometry is almost filling a box of $20 \times 20 \times 10$, the maximal horizontal load is $F_{\max,xy} = 0.0015$ and the vertical one $F_{\max,z} = 2F_{\max,xy}$. The elementwise bounds on the material thickness are $u^- = 0.01$ and $u^+ = 0.2$. The volume of the material is bounded by $V^+ = 60$ and the strength of the stochastic variation is fixed by $\sigma = 0.1$. The weights of the barrier terms were $\alpha^F = 10^{-4}$, $\alpha^u = 1$, and $\alpha^V = 10^{-5}$. For the leader, we consider a tracking set restricted to the central region of the roof plateau as shown in the first panel of Fig. 2.

In Fig. 3, we show the deformed configuration, the optimized distribution of the material thickness, and the magnitude of displacements in case of the leader minimizing a tracking functional once with global support (top row) using $\chi \equiv 1$ and once restricted to the region of the roof plateau (bottom row). As for all examples presented here, in the follower problem, the maximal compliance is attained for a force F representing an extremal point of the cylinder of admissible forces. For the tracking cost domain centered on the roof plateau, one observes a concentration of mass in the central region accompanied by a significant reduction of the thickness close to
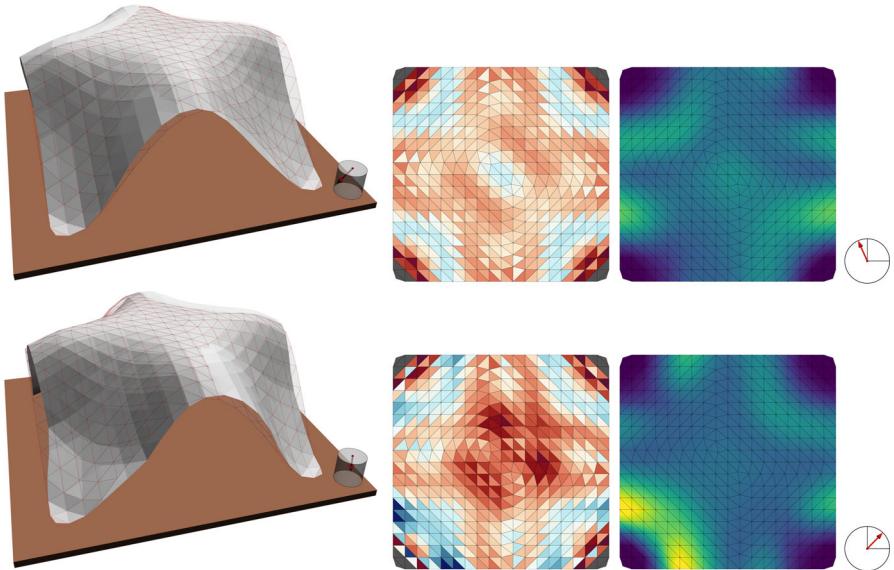
Fig. 3 Comparison of results for full vertex tracking set (top) and plateau tracking set (bottom) on the simple roof-type geometry already shown in Fig. 2. On the left, we show the deformed configurations as gray surfaces, while the undeformed surfaces are shown as translucent surfaces overlaid with red edges. Next to the surfaces, we visualize the direction of the force (F_1, F_2, F_3) chosen by the follower in the cylinder of admissible values. In the middle, we show the resulting material distributions with color map 0 ▬ 0.2, where boundary triangles with all three vertices subject to Dirichlet boundary conditions are shown in gray. On the right, we show the magnitude of the deformation y using the color map 0 ▬ ≥ 1.5. Additionally, on the far right, we show the direction of the horizontal forces (F_1, F_2)

the four corners where Dirichlet boundary conditions apply. The concentration and corresponding reduction break the symmetry of the configuration w.r.t. the diagonal from the upper left to the lower right. Due to the asymmetric reduction, the follower chooses a force pointing to the upper right and one observes a kink line connecting the two arcs in the front at approximately half of the total height. This is accompanied by large displacements, which are however outside of the tracking region on the plateau. In contrast, for the tracking with global support, no such kink with strong displacements occurs, however, the deformation exhibits a larger displacement in the central region. Finally, beyond the mass concentration in the middle, one also observes the onset of curved "beam" like structures connecting the middle region and the four arcs of the roof. In the example with localized tracking, and most of the following ones, the elementwise bounds u^+ and u^- are nearly attained for at least some triangles.

Figure 4 shows for the same geometry the impact of the upper bound on the total material volume.

As the total permitted mass is increased, the elongated curved "beams" connecting the tracking region in the center with the four arcs become thicker. Once the maximal thickness is reached in the central region and along these "beams", further mass is invested to reinforce the regions close to the Dirichlet boundaries. The curved carrier "beams" and the central region are again designed asymmetrically w.r.t. the diagonal
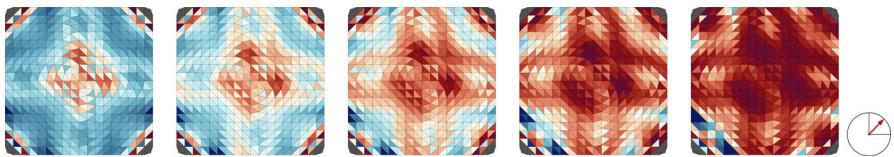
Fig. 4 A comparison of the material distribution when varying the maximal allowed material volume V^+ while keeping the other parameters fixed. The allowed volume was $V^+ = 40, 50, 60, 70, 80$ from left to right. Material thickness is shown using the color map 0 ▬ 0.2. On the far right, we show the direction of the horizontal forces, which was the same for all parameters, while the vertical force was always chosen maximal
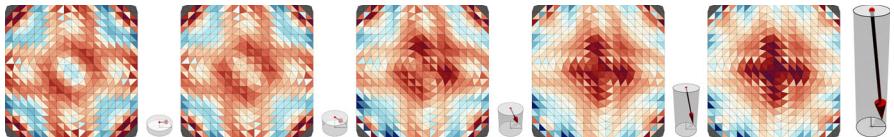
Fig. 5 A comparison of the material distribution when varying the ratio of vertical to horizontal force $\frac{F_{\max,z}}{F_{\max,xy}}$, *i.e.* the shape of the cylinder, while keeping the other parameters, especially the maximal magnitude of horizontal force, fixed. The ratio of vertical to horizontal force was $\frac{F_{\max,z}}{F_{\max,xy}} = \frac{1}{2}, 1, 2, 4, 8$ from left to right. The material thickness is shown using the color map 0 ▬ 0.2. On the right of each material distribution, we show the force in the cylinder of admissible values

from the upper left to the lower right leading the follower to push towards the upper right.

We next investigate the effect of the parameters characterizing the strength of the forces, $F_{\max,z}$ and $F_{\max,xy}$, while keeping the total amount of material constant. By scaling invariance, it is natural to focus on the ratio $\frac{F_{\max,z}}{F_{\max,xy}}$. In Fig. 5, we show that with increasing strength of the vertical force, the "beams" become thinner and instead more material is concentrated in the central region.

Interestingly, for small values of the ratio between the two forces the material distribution is nearly symmetrical w.r.t. the diagonal from the upper left to the lower right, while it is asymmetric for mid-range ratios and then becomes more symmetric again for large ratios.

Figure 6 shows the impact of the strength of the stochastic perturbation of the material thickness, as measured by the standard deviation, again for the tracking region on the roof plateau.

With increasing strength of the stochastic perturbation, the optimal structure becomes more diffuse. Indeed, in a deterministic setting, the leader could aim for a finely-structured design, but very imprecise manufacturing is likely to render it ineffective. In order to understand this effect, we describe an idealized situation: If the leader concentrates mass on a single row of k elements, then a large negative fluctuation in the thickness of any single one of them is sufficient to destroy the strength of the construction. If instead, the leader distributes the mass on k^2 elements filling a square, then at least a number of order k of those (with a specific geometry, for example, a column) must have a large negative fluctuation before the structure loses significantly in strength.
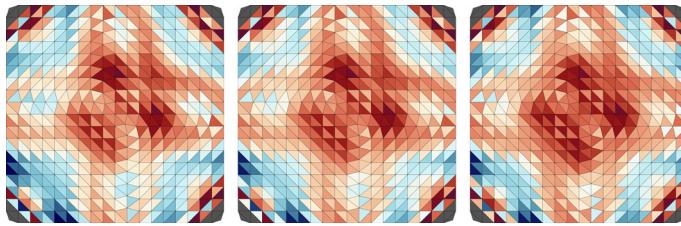
Fig. 6 Comparison of material distribution when varying the standard deviation σ of the material perturbation while keeping the other parameters fixed. The standard deviation was $\sigma = \frac{5}{100}, \frac{1}{10}, \frac{2}{10}$ from left to right. Material thickness is shown using the color map 0 ▬ 0.2
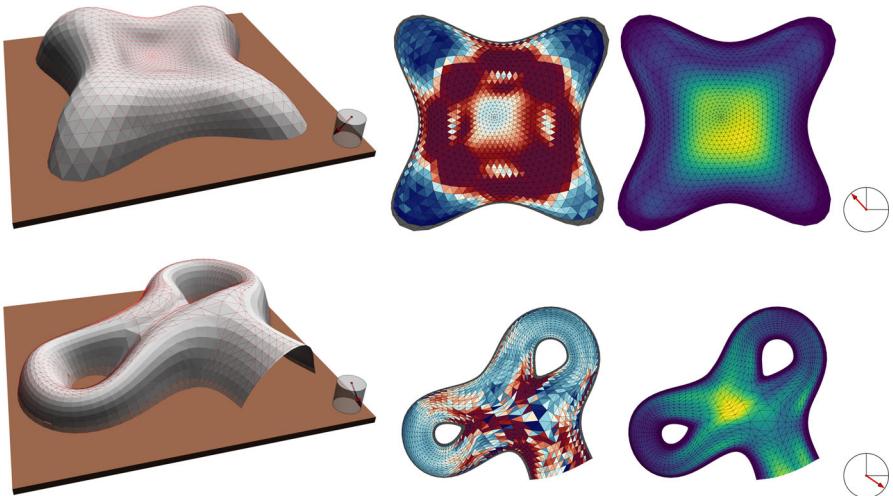
Fig. 7 Results for two geometrically more complex examples. In both cases, we used tracking on the entire domain. On the left, we show the deformed configuration as a gray surface with the undeformed surfaces as a translucent overlay. Furthermore, we visualize the direction of the force leading to the maximal deformation in the cylinder. In the middle, we see the resulting material distributions using the color map 0 ▬ 0.2. Boundary triangles for which all vertices are Dirichlet nodes are shown in gray. On the right, the magnitude of the deformation y is displayed using the color map 0 ▬ 0.7. Additionally, on the far right, we show the 2D direction of the horizontal forces

Lastly, in Fig. 7, we show two more complex examples of architectural designs of roof structures, inspired by [38]. In the top row, we use a closed hall as the reference geometry for our bilevel optimization problem, which fills a box of approximately $20 \times 20 \times 5$. We limit the horizontal load with $F_{\max,xy} = 0.005$ and the vertical load with $F_{\max,z} = 2F_{\max,xy}$. The elementwise bounds on the material thickness are $u^- = 0.01$ and $u^+ = 0.2$. The volume of the material is bounded by $V^+ = 50$ and the stochastic variation is $\sigma = 0.05$. The weights of the barrier terms are $\alpha^F = 10^{-4}$, $\alpha^u = 1$, and $\alpha^V = 10^{-3}$. In the bottom row, we use a reference geometry resembling a double torus cut in half, which fills a box of approximately $70 \times 50 \times 15$. Again, we limit the horizontal load with $F_{\max,xy} = 0.005$ and the vertical load with $F_{\max,z} = 2F_{\max,xy}$. The elementwise bounds on the material thickness are again $u^- = 0.01$ and $u^+ = 0.2$.

The volume of the material is bounded by $V^+ = 330$ and the stochastic variation is $\sigma = 0.05$. The weights of the barrier terms are $\alpha^F = 10^{-4}$, $\alpha^u = 1$, and $\alpha^V = 10^{-1}$. In both cases, we use the full domain as tracking set. The main weakness of both structures is the concavity in the central part, which can be easily deformed by the vertical force. Hence, in both optimized solutions, the material is redistributed to prevent this. In the first case, this is done by building a stabilized ledge around the center, while in the second case beam-like structures from the two "holes" and another beam from the curve in the bottom emerge. Furthermore, in the second one, also the "entrance" is stabilized by adding material at the ends of its arch.

5 Discussion

The findings in this article draw a line from curved roof-type constructions via modeling and shape optimization of discrete thin shells to pessimistic formulations of bilevel stochastic programs. The challenge is that even in the deterministic case, it is well-known that standard compactness assumptions fail to ensure the existence of optimal solutions.

Assuming that the support of the underlying probability measure is compact, we have considered stochastic parameters and assessed the random upper-level outcome based on some (law-invariant) convex risk measure. For the pessimistic model, we have shown continuity of the resulting risk functional if the random perturbation admits a Lebesgue density, the set of potential forces is a polyhedron and the lower level goal function is real analytic. Alternatively, we have investigated a regularized model where the leader also hedges against lower level solutions that are close to optimality. The risk functionals emerging from this regularized problem are automatically lower semicontinuous. In both situations, the existence of optimal solutions can be guaranteed under a compactness condition. We have developed a proof-of-concept numerical implementation that applies a pessimistic bilevel strategy to a mechanical optimal design problem, using a stochastic gradient descent approach to compute locally optimal solutions of the pessimistic model.

In closing, we would like to point out several possible directions for future research. In the numerical optimization, it would be interesting to consider interior point methods to solve the "original" leader's and follower's problem which incorporate hard constraints instead of the regularization used here. From the point of view of elasticity, it would be interesting to study the nonlinear model (13) instead of its linearized equivalent and investigate the associated nonuniqueness issue in the lower level problem. In fact, this would lead to a proper trilevel problem and bring new challenges for theoretical and numerical investigations. Furthermore, it would be worthwhile to investigate the infinite-dimensional variational problem of thin shell or volume elasticity with appropriate function spaces and from the perspective of optimization with continuous PDE constraints. In the present paper, "risky" decisions can be penalized in the objective function, but the leader ensures that the perturbed material parameters are feasible regardless of the realization, via a restriction of the design variable to the set $\mathcal{U}_\Upsilon$. Models, where this robust constraint is replaced with a system of chance or stochastic dominance constraints, can be expected to produce less conservative solu-

tions, which improve the values of the leader's cost functional at the cost of some residual risk.

References

1. Bard, J.F.: Some properties of the bilevel programming problem. J. Optimiz. Theory Appl. **68**(2), 371–378 (1991). https://doi.org/10.1007/BF00941574
2. Burtscheidt, J., Claus, M., Dempe, S.: Risk-averse models in bilevel stochastic linear programming. SIAM J. Optimiz **30**(1), 377–406 (2020). https://doi.org/10.1137/19M1242240
3. Chen, Y., Davis, T.A., Hager, W.W., Rajamanickam, S.: Algorithm 887: Cholmod, supernodal sparse cholesky factorization and update/downdate. ACM Trans. Math. Softw. **35**(3) (2008)
4. Cheridito, P., Li, T.: Risk measures on Orlicz hearts. Mathematical Finance **19**(2), 189–214 (2009). https://doi.org/10.1111/j.1467-9965.2009.00364.x
5. Christiansen, S., Patriksson, M., Wynter, L.: Stochastic bilevel programming in structural optimization. Struct Multidisc Optim **21**(5), 361–371 (2001). https://doi.org/10.1007/s001580100115
6. Ciarlet, P.G.: Mathematical Elasticity, Vol. I: three-dimensional elasticity. North-Holland (1988)
7. Ciarlet, P.G.: Mathematical Elasticity, Vol. III: Theory of shells. North-Holland (2000)
8. Conti, S., Held, H., Pach, M., Rumpf, M., Schultz, R.: Shape optimization under uncertainty: A stochastic programming perspective. SIAM J. Optim. **19**(4), 1610–1632 (2009). https://doi.org/10.1137/070702059
9. Conti, S., Held, H., Pach, M., Rumpf, M., Schultz, R.: Risk averse shape optimization. SIAM J. Control Optim. **49**(3), 927–947 (2011). https://doi.org/10.1137/090754315
10. Conti, S., Rumpf, M., Schultz, R., Tölkes, S.: Stochastic dominance constraints in elastic shape optimization. SIAM J. Control Optim. **56**(4), 3021–3034 (2018). https://doi.org/10.1137/16M108313X
11. Dempe, S.: Foundations of bilevel programming. Kluwer Acad. Publication, Dordrecht (2002)
12. Farshbaf-Shaker, M.H., Gugat, M., Heitsch, H., Henrion, R.: Optimal Neumann boundary control of a vibrating string with uncertain initial data and probabilistic terminal constraints. SIAM J. Control Optim. **58**(4), 2288–2311 (2020). https://doi.org/10.1137/19M1269944
13. Frittelli, M., Rosazza Gianin, E.: Putting order in risk measures. J. Bank. Fin. **26**(7), 1473–1486 (2002). https://doi.org/10.1016/S0378-4266(02)00270-4
14. Föllmer, H., Schied, A.: Convex measures of risk and trading constraints. Fin. Stoch. **6**(4), 429–447 (2002). https://doi.org/10.1007/s007800200072
15. Follmer, H., Schied, A.: Stochastic Finance an introduction in discrete time, 3. rev. and extended eds. edn. De Gruyter, Berlin (2011)
16. Geiersbach, C., Loayza-Romero, E., Welker, K.: Stochastic approximation for optimization in shape spaces. SIAM J. Optimiz. **31**(1), 348–376 (2021)
17. Grinspun, E., Hirani, A.N., Desbrun, M., Schroder, P.: Discrete shells. In: Proc. of ACM SIGGRAPH/Eurographics Symposium on Computer animation, pp. 62–67 (2003)
18. Guennebaud, G., Jacob, B., Others: Eigen v3. http://eigen.tuxfamily.org (2010)
19. Heeren, B., Rumpf, M., Schröder, P., Wardetzky, M., Wirth, B.: Exploring the geometry of the space of shells. Comput. Graph. Forum **33**(5), 247–256 (2014)
20. Heeren, B., Sassen, J., et al.: The Geometric Optimization And Simulation Toolbox (2020). https://gitlab.com/numod/goast

21. Herskovits, J., Leontiev, A., Dias, G., Santos, G.: Contact shape optimization: a bilevel programming approach. Struct. Multidiscipl. Optimiz. **20**, 214–221 (2000). https://doi.org/10.1007/s001580050149
22. Kaina, M., Rueschendorf, L.: On convex risk measures on L^p-spaces. Math Meth Oper Res **69**, 475–495 (2009)
23. Leitmann, G.: On generalized stackelberg strategies. J. Optimiz. Theory Appl. **26**(4), 637–643 (1978)
24. Lignola, M.B., Morgan, J.: Inner regularizations and viscosity solutions for pessimistic bilevel optimization problems. J. Optimiz. Theory Appl. **173**(1), 183–202 (2017)
25. Lignola, M.B., Morgan, J.: Further on inner regularizations in bilevel optimization. J. Optimiz. Theory Appl. **180**(3), 1087–1097 (2019)
26. Loridan, P., Morgan, J.: Weak via strong stackelberg problem: new results. J. Global Optimiz. **8**(3), 263–287 (1996)
27. Love, A.E.H.: A treatise on the mathematical theory of elasticity. hal-01307751 (1892). hal.archives-ouvertes.fr/hal-01307751
28. Lucchetti, R., Mignanego, F., Pieri, G.: Existence theorems of equilibrium points in Stackelberg games with constraints. Optimization **18**(6), 857–866 (1987)
29. Martínez-Frutos, J., Herrero-Pérez, D., Kessler, M., Periago, F.: Risk-averse structural topology optimization under random fields using stochastic expansion methods. Comp. Meth. Appl. Mech. Eng **330**, 180–206 (2018)
30. Mityagin, B.: The zero set of a real analytic function. Mathem. Notes **107**(3–4), 529–530 (2020)
31. Nocedal, J., Wright, S.J.: Numerical Optimization, 2 edn. Springer (2006)
32. Patriksson, M., Wynter, L.: Stochastic mathematical programs with equilibrium constraints. Operat. Res. Lett. **25**(4), 159–167 (1999)
33. Robbins, H., Monro, S.: A stochastic approximation method. The Annals of Mathem. Stat. **22**(3), 400–407 (1951). https://doi.org/10.1214/aoms/1177729586
34. Rockafellar, R.T., Wets, R.J.B.: Variational analysis, 3. edn. Springer, Berlin (2009). https://sites.math.washington.edu/simrtr/papers/rtr169-VarAnalysis-RockWets.pdf
35. Shapiro, A., Dentcheva, D., Ruszczyński, A.P.: Lectures on stochastic programm. modeling and theory. SIAM, Soc. for Industrial and Applied Math., Philadelphia (2009)
36. Sinha, A., Malo, P., Deb, K.: A review on bilevel optimization: from classical to evolutionary approaches and applications. IEEE Trans. Evolut. Comput. **22**(2), 276–295 (2018). https://doi.org/10.1109/TEVC.2017.2712906
37. von Stackelberg, H.: Marktform und Gleichgewicht. Julius Springer, Wien und Berlin (1934)
38. Vouga, E., Hobinger, M., Wallner, J., Pottmann, H.: Design of self-supporting surfaces. ACM Trans. Graphics (2012). Proc. SIGGRAPH
39. Zuo, W.: Bi-level optimization for the cross-sectional shape of a thin-walled car body frame with static stiffness and dynamic frequency stiffness constraints. Proc. Ins. Mechan Eng, Part D: J. Autom. Eng **229**(8), 1046–1059 (2015). https://doi.org/10.1177/0954407014551585

Affiliations

Johanna Burtscheidt[1] ⓘ · **Matthias Claus[1]** ⓘ · **Sergio Conti[2]** ⓘ ·
Martin Rumpf[3] · **Josua Sassen[3]** ⓘ · **Rüdiger Schultz[1]** ⓘ

✉ Matthias Claus
  matthias.claus@uni-due.de

  Johanna Burtscheidt
  johanna.burtscheidt@uni-due.de

  Sergio Conti
  sergio.conti@uni-bonn.de

Martin Rumpf
martin.rumpf@uni-bonn.de

Josua Sassen
josua.sassen@uni-bonn.de

Rüdiger Schultz
ruediger.schultz@uni-due.de

1   Faculty of Mathematics, University of Duisburg-Essen, Thea-Leymann-Straße 9, Essen 45127,
    Germany

2   Institute for Applied Mathematics, University of Bonn, Endenicher Allee 60, Bonn 53115,
    Germany

3   Institute for Numerical Simulation, University of Bonn, Endenicher Allee 60, Bonn 53115,
    Germany