Stochastic bilevel programming in structural optimization

Snorre Christiansen,[*][†] Michael Patriksson[‡] and Laura Wynter[*]

December 2, 1999

Abstract

We consider the mathematical modelling and solution of robust and cost-optimizing structural (topology) design problems. The setting is the optimal design of a linear–elastic structure, for example a truss topology, under unilateral frictionless contact, and under uncertainty in the data describing the load conditions, the material properties, and the rigid foundation. The resulting stochastic bilevel optimization model finds a structural design that responds the best to the given probability distribution in the data. This model is of special interest when a structural failure will lead to a reconstruction cost, rather than loss of life.

For the mathematical model, we provide results on the existence of optimal solutions which allow for zero lower design bounds. We establish that the optimal solution is continuous in the design bounds, a result which validates the use of small but positive values of them, and for such bounds we also establish the locally Lipschitz continuity and directional differentiability of the implicit one-level objective function. We also provide a simple algorithm for the solution of the problem, which makes use of its differentiability properties and parallelization strategies across the scenarios. A small set of numerical experiments illustrates the behaviour of the stochastic solution compared to an average-case deterministic one, establishing an increased robustness.

1 Introduction

In this paper, we consider the optimal design of a linear–elastic structure under unilateral frictionless contact. This problem incorporates the classical truss topology optimization problem as well as the (more general) problem of finding the optimal design of a two-dimensional variable thickness sheet.

Some of the recent research in structural optimization has centered on formulating models that take into account multiple load cases (e.g., Kočvara et al., 1998; and Ben-Tal and Nemirovski, 1997). Another body of research in the area has involved probabilistic approaches based on reliability constraints (see, for example, Marti, 1997, for a recent collection of work in the field; and Kirjner-Neto et al., 1998). Often, the scope of the problems is limited to fit into the framework of single-level optimization problems.

Structural optimization problems have, in general, however, an inherent bilevel form. The upper-level objective function measures some performance of the structure, or its construction cost or weight, and is optimized by selecting design parameters, which may express the shape of the structure, the choice of material, and/or the amount of material being used. Further, the structure may be subject to limits on the amount of available material, and to behavioural constraints, such as bounds on the displacements, stresses and contact forces. The lower-level problem describes the behaviour of the structure given the choice of the design variables, possible contact conditions with foundations or obstacles, and the external forces acting on it. The behaviour is (for elastic structures) described by the equilibrium law of minimal potential energy, which determines the values of the state variables (displacements, stresses and contact forces) at the lower level.

It has long been recognized that when determining the design of a mechanical structure it is vital to take into account the uncertain character of some of the parameters that will determine the ultimate

---

[*]PRISM, Université de Versailles, Versailles, France

[†]Ecole Polytechnique, Palaiseau, France

[‡]Department of Mathematics, Chalmers University of Technology, Gothenburg, Sweden

design. Often, however, engineering models in shape, sizing and (more generally) topology optimization have often ignored the presence of uncertainty in the data, such as random properties of the material used and conditions that will affect the structure once it has been built, such as varying weather conditions and external forces acting on it. This may result in the construction of designs that are unstable under varying conditions.

Models that do take uncertainty into account often do so in a worst-case manner, which yield stable structures but which do not take into account the probability of occurrence of the different scenarios. While worst-case approaches are appropriate in circumstances where a failure can entail loss of life, the failure of many structures will lead only to a reconstruction cost. It is these cases for which an average-case probabilistic approach such as we provide can be of use.

We consider the problem of finding the optimal design of a linear–elastic structure under unilateral frictionless contact, in a true bilevel framework, and incorporating explicitly probabilistic features of the model. We present first a description of the bilevel model in a deterministic setting. Then, we introduce the random elements into the model, and give a brief history of probabilistic methods for structural optimization. In Section 2, we present the bilevel stochastic model, based upon the stochastic generalized bilevel optimization model introduced in Patriksson and Wynter (1998), and give new results on the existence of solutions and some mathematical properties useful for its resolution. In Section 3, we present a heuristic algorithm for the model and a parallelization strategy, as well as an exact method for solving the problem. Then, in Section 4, a series of numerical experiments is used to illustrate the benefits of the stochastic bilevel model, using the heuristic method with parallel computation.

1.1 The deterministic problem

The physical object studied is a discrete (or, discretized) mechanical structure. Potentially, at each of the M nodes of the structure, an external force is applied, where both the force magnitude and direction are specified. (In the case of a truss the forces are distributed along the bars.) Letting dim denote the dimension of the space, we define $m := \dim \cdot M - p$, where p is the number of prescribed zero displacement directions. (Non-zero prescribed displacements will not be considered here.) The forces are then collected into a constant m-dimensional vector, F. As a result of the applied forces, each node of the structure is subject to a displacement; this vector is denoted by $y \in \Re^m$. The stiffness of the structure is determined by a (symmetric and positive semi-definite) stiffness matrix $K(x) \in \Re^{m \times m}$, where $x \in \Re^n$ is the vector of design parameters. In truss topology optimization, x contains the bar volumes, in the variable thickness sheet problem it is the finite element discretized thickness function, and in shape optimization x contains coordinates of control points for the splines defining the boundaries to design. In the first two cases, the stiffness of the structure is determined by a linear relation, $K(x) := \sum_{i=1}^{n} K_i x_i$, where $K_i \in \Re^{m \times m}$, $i = 1, \ldots, n$, is the stiffness matrix for element i in the discrete structure. Each matrix K_i is symmetric and positive semi-definite; we assume that a sufficient number of zero displacement directions are specified so that the matrix $K(x)$, however, is positive definite when $x > 0^n$. In this model, the design variables x are subject to constraints on the available amount of material; in particular, we require that

$$x \in \mathcal{X} := \{ \, x \in \Re^n \mid 1^T x = V; \quad l \le x \le u \, \},$$

where $0^n \le l < u$, that is, the total volume available is equal to a constant, V. Furthermore, there are lower and upper bounds on each element volume. If the lower bound $l = 0^n$, then bars can be eliminated from the ground set of trusses (or, for more general discrete structures, holes are allowed). Thereby, a true topology optimization is possible. However, note that with a zero lower bound on x, $K(x)$ is not positive definite everywhere on $\mathcal{X}$.

Given a design parameter vector $x \in \mathcal{X}$, the state of the structure is determined by its minimal potential energy,

$$\Pi(x, y) := \frac{1}{2} y^T K(x) y - F^T y.$$

We may also cover situations arising in contact mechanics by including a set of unilateral contact conditions, enforced by the presence of r rigid obstacles. The initial (normal) distance from a potential contact node of the structure to its obstacle point is collected in the vector $b \in \Re^r$; the corresponding normal

directions $\nu_i \in \Re^m$ are collected in the kinematic transformation matrix $C \in \Re^{r \times m}$. Hence, the admissible displacements y are given by the polyhedral set

$$\mathcal{Y} := \{\, y \in \Re^m \mid Cy \leq b \,\}.$$

For a given design vector $x \in \mathcal{X}$, the equilibrium state is determined by solving the convex quadratic program

$$y \in \arg\min_{z \in \mathcal{Y}} \Pi(x, z). \tag{1}$$

As stated above, since $K(x)$ is positive definite in general only if $x > 0^n$, for designs that include bar elimination (or, holes) this problem may not have an optimal solution for every feasible x.

The cost of the structure can in general depend on both the design and displacement variables, as well as on the contact force magnitude $\lambda \in \Re_+^r$, which is given by the vector of Lagrange multipliers for the contact conditions in (1). A general (deterministic) structural optimization problem is thus (see, for example, Bendsøe, 1995; Petersson, 1995; and Haslinger and Neittaanmäki, 1996):

$$\underset{(x,y)\in\mathcal{Z}}{\text{minimize}} \quad f(x, y), \tag{2a}$$

$$\text{subject to} \quad y \in \arg\min_{z \in \mathcal{Y}} \Pi(x, z), \tag{2b}$$

where $\mathcal{Z} = \mathcal{X} \times \Re^m$, and $f : \mathcal{Z} \mapsto \Re$ is the cost function of the structure. Assumptions on this function are introduced in the following section.

Most truss topology optimization models rely upon simpler forms of the primary objective function which enable one to reduce the model to a one-level optimization problem. If the primary objective is chosen as the minimum extended compliance, that is, $f(x, y, \lambda) := F^T y + b^T \lambda$, where F is the vector of external forces, y the displacements, b the vector of initial distances to the given rigid obstacles, and λ the vector of contact force magnitudes, then the bilevel problem reduces to that of finding the stiffest structure. The problem can then be equivalently written as a convex–concave saddle-point problem. Likewise, if contact conditions are not present (that is, $r = 0$) and the topology is fixed (that is, $l > 0^n$), then the lower-level problem can (in principle) be eliminated entirely from the structural optimization problem through the equilibrium conditions $K(x)y = F$. For other, more general, measures such as cost minimization, and when a true topology optimization is permitted, however, the problem is inherently bilevel and non-convex. Such more general measures can be included in our framework.

One way to take random events into account would be to design structures based on worst-case scenarios (replacing the average-case scenario considered in the deterministic model); this would mean building stable structures, but based in some cases on random events with very low probability. This may lead to designing excessively costly structures. To overcome this, we consider the use of probabilistic methods.

Jóźwiak (1989) considers the problem of finding the minimum average weight of a truss subject to a probabilistic constraint on the risk of failure of the structure, where the strength of the bars is assumed to have a (known) log-normal distribution. Tada and Seguchi (1989) consider the maximum stiffness of a finite-element structure under uncertain loading conditions; unilateral contact conditions not being present, they are able to state a system of nonlinear equations representing the average stationarity of the system under a constant volume condition. Marti (1992) considers a stochastic structural design optimization model under behavioural constraints, formulated as a variety of one-level design optimization problems with probabilistic constraints on the structural response; the (essentially) unconstrained model is analyzed via stochastic quasi-gradient methods. The special issue of Marti (1997) presents a number of papers with probabilistic approaches to reliability constraints. All of these models are one-level optimization models.

Recently, Kirjner-Neto et al. (1998) presented a semi-infinite bilevel optimization program, where the lower-level problem represents a reformulation of a set of probabilistic constraints ensuring a minimum reliability level. Their approach is different and complementary to that presented here, in which we consider the entire set of performance states and the probability of each, as opposed to setting a maximum failure probability.

The model of Tada and Seguchi (1989) is similar in spirit to a classical approach in engineering design to dealing with varying external forces, which is known as the *multi-load* case of the structural optimization problem (e.g., Pedersen and Jørgensen, 1984). Such an approach seeks to take different load cases into

account in the model by, essentially, considering the worst-case scenario for each set of design parameter values. The model of Kočvara et al. (1998) [see also Bendsøe, 1995] extends the max–min problem of Petersson (1995) to the multi-load case as follows: a finite number of load cases F_i, $i = 1, \ldots, L$, are to be considered; in our notation, then, the problem (2b) would be extended to

$$y \in \arg\min_{z \in \mathcal{Y}} \min_{i=1,\ldots,L} \Pi_i(x, z) := \frac{1}{2} z^T K(x) z - F_i^T z,$$

that is, among the possible load cases the one which gives the minimum potential energy is chosen. The multi-load case clearly does not take into account the probability with which each load case appears. Zowe et al. (1997) extend the model of Kočvara et al. (1998) by including material choice in the optimization, along with weighting factors on each load case; the use of these weighting factors could easily be modified to include the probability of each load case. Ben-Tal and Nemirovski (1997) present a model for *robust* structural optimization, which can be viewed as an extension of the multi-load model to a continuum of loads identified by an ellipse around certain nodes; this model differs from ours in that it remains a worst-case analysis and does not consider the probability with which each event can occur. Furthermore, as previously mentioned, the models summarized above are all single-level programs by the use of a particular primary objective–that of minimal extended compliance.

2   The stochastic bilevel model

To simplify the discussion, we consider a stochastic programming extension of (2) in which only the load F is subject to random variation. Suppose therefore that $F = F_\omega$, where ω belongs to a probability space $(\Omega, \mathcal{A}, P)$.

In the most general case, we would consider the following extension to (2):

$$\underset{x \in \mathcal{X}}{\text{minimize}} \quad E_\omega[f(x, y(\omega))] := \int_\Omega f(x, y(\omega)) \, dP(\omega), \tag{3a}$$

$$\text{subject to} \quad y(\omega) \in \arg\min_{z \in \mathcal{Y}} \Pi_\omega(x, z) := \frac{1}{2} z^T K(x) z - F_\omega^T z, \qquad \omega \in \Omega. \tag{3b}$$

Here, E_ω denotes the expected value, and P denotes the cumulative distribution function, assumed to be continuously differentiable, with $\int_\Omega dP(\omega) = 1$ by construction.

We make the assumption (or, approximation) that the random variables $\omega \in \Omega$ are discrete or that the joint distribution function can be adequately discretized; in this case, the expected value is reduced to a sum over a discrete set $\mathcal{L}$ of random events, numbered from $\ell = 1, \ldots, |\mathcal{L}|$, the probability of each such event being given by $\rho_\ell > 0$, with $\sum_{\ell \in \mathcal{L}} \rho_\ell = 1$ by construction. We can then reformulate (3) as the following discretely-distributed stochastic bilevel optimization problem:

$$\underset{x \in \mathcal{X}}{\text{minimize}} \quad E_\ell[f(x, y_\ell)] := \sum_{\ell \in \mathcal{L}} \rho_\ell f(x, y_\ell), \tag{4a}$$

$$\text{subject to} \quad y_\ell \in \arg\min_{z \in \mathcal{Y}} \Pi_\ell(x, z) := \frac{1}{2} z^T K(x) z - F_\ell^T z, \qquad \ell \in \mathcal{L}. \tag{4b}$$

As opposed to the multi-load model, in this model, the design x is chosen so that the average performance of the structure is good in probability, taking into account the equilibrium displacements of each different load case. We also note that the formulation (4) requires the lower-level problem (4b) to be solvable for each realization (that is, a static equilibrium must exist for all of the forces specified), since otherwise an infinite penalty would be added to f. We conclude that the proposed model (potentially) provides a robust (as well as cost-optimal) structure.

We note, finally, that if the number of discrete scenarios is large (which it needs to be in order to represent random events realistically), then the problem (4) is considerably larger than its deterministic origin, (2), and must be attacked by specialized methods, involving decomposition and parallelism.

2.1 Properties of the model

This section establishes the existence of an optimal solution to the problem (4), and provides sufficient conditions for the problem to have an equivalent one-level representation that permits utilizing a descent algorithm for its solution.

Proposition 2.2 below follows directly from Corollary 2.4 of Patriksson and Wynter (1998), but we provide an independent proof for the sake of completeness. Let, for each $\ell \in \mathcal{L}$,

$$\operatorname{gr} \mathcal{S}_\ell := \{ (x, y_\ell) \in \Re^{n+m} \mid y_\ell \text{ solves (4b)} \} \tag{5}$$

denote the graph of the set $\mathcal{S}_\ell(x)$ of solutions to the lower-level problem (4b). The next result relies on the following conditions on the upper-level objective function.

ASSUMPTION 2.1 (Inf-compactness). *The upper-level objective function $f : \mathcal{Z} \mapsto \Re$ is* (i) *lower semicontinuous (l.s.c.) and proper, and* (ii) *has bounded lower level sets on $\mathcal{Z} \cap \operatorname{gr} \mathcal{S}_\ell$, $\ell \in \mathcal{L}$.*

Part (i) of this assumption is satisfied if f is continuous on $\mathcal{Z}$, and part (ii) is equivalent to the weak coercivity of f on $\mathcal{Z} \cap \operatorname{gr} \mathcal{S}_\ell$, $\ell \in \mathcal{L}$, that is,

$$\lim_{\substack{(x, y_\ell) \in \mathcal{Z} \cap \operatorname{gr} \mathcal{S}_\ell \\ \|y_\ell\| \to +\infty}} f(x, y_\ell) = +\infty, \qquad \ell \in \mathcal{L}.$$

This last property has the interpretation of the minimization of f striving against very large equilibrium displacements. The following result is based in part on an existence result of Zhang (1994).

PROPOSITION 2.2 (Existence of an optimal solution under inf-compactness). *Under the assumptions stated for the problem (2) in Section 1.1 and Assumption 2.1, the problem (4) has at least one optimal solution.*

Proof: First, we show that for all $\ell \in \mathcal{L}$, $\mathcal{Z} \cap \operatorname{gr} \mathcal{S}_\ell$ is a closed set. Letting $\{(x^t, y_\ell^t)\}$ be a sequence of pairs in $\operatorname{gr} \mathcal{S}_\ell$, and $\{(x^t, y_\ell^t)\} \to (x, y_\ell)$, then by the closedness of $\mathcal{Y}$, and the continuity of Π_ℓ, it follows immediately that $(x, y_\ell) \in \operatorname{gr} \mathcal{S}_\ell$, and hence $\operatorname{gr} \mathcal{S}_\ell$ is closed. The conclusion is then obvious.

Next, we note that for all $\ell \in \mathcal{L}$, the set $\mathcal{Z} \cap \operatorname{gr} \mathcal{S}_\ell$ is non-empty, since there exists an $x \in \mathcal{X}$ such that $x > 0^n$, in which case the equilibrium problem (4b) has a unique solution for all $\ell \in \mathcal{L}$. Because f is proper, such a choice of x also confirms that for all $\ell \in \mathcal{L}$, there exists a pair $(x, y_\ell) \in \operatorname{gr} \mathcal{S}_\ell$ such that $f(x, y_\ell) < +\infty$. Since f is inf-compact (l.s.c. and coercive) by assumption, Weierstrass' Theorem then establishes the existence of a solution. □

In a number of applications, the coercivity condition imposed on f in Proposition 2.2 is overly restrictive. One such example is the case of friction-less sliding on a horizontal surface: consider a structure in unilateral contact on a horizontal surface subject to an external downward force. In this case, if no nodes are fixed, then the horizontal equilibrium displacement y may be infinite. However, for a number of common functional forms, the value of the upper-level cost function f will remain constant under such horizontal translations in equilibrium; the coercivity condition is therefore not satisfied in such a situation.

For a certain class of problems we can show the existence of an optimal solution without imposing the coercivity of f; in particular, we consider the class of upper-level objectives that are quadratic in y.

Consider first the following assumptions.

ASSUMPTION 2.3 (Quadratic objective). *The upper-level objective function $f : \mathcal{Z} \mapsto \Re$ is lower bounded on $\mathcal{Z} \cap \operatorname{gr} \mathcal{S}_\ell$, $\ell \in \mathcal{L}$, and quadratic in y for each $x \in \mathcal{X}$.*

This assumption will ensure that the value of the *implicit* objective function,

$$\hat{f}(x) := \sum_{\ell \in \mathcal{L}} \rho_\ell \min_{y_\ell : (x, y_\ell) \in \operatorname{gr} \mathcal{S}_\ell} f(x, y_\ell), \tag{6}$$

is finite for every $x \in \mathcal{X}$ for which an equilibrium solution exists for all $\ell \in \mathcal{L}$. We define this set as

$$\mathcal{S} := \{ x \in \mathcal{X} \mid y_\ell \text{ solves (4b) for some } y_\ell \in \mathcal{Y}, \ell \in \mathcal{L} \}.$$

As before, this set is nonempty, since there exists an $x \in \mathcal{X}$ such that $x > 0^n$, whence (4b) has a unique solution y_ℓ for all $\ell \in \mathcal{L}$. In addition, $\mathcal{S}$ is bounded since $\mathcal{S} \subseteq \mathcal{X}$, where $\mathcal{X}$ is bounded. However, $\mathcal{S}$ is not closed in general when the lower design bound is zero. It is therefore not suitable as a feasible set over which to minimize $\hat{f}$. For this purpose, we introduce a subset, $\mathcal{W}$, of $\mathcal{S}$, which we assume is closed.

ASSUMPTION 2.4 (Closed feasible set). *For some $\nu \in \mathfrak{R}$, the set $\mathcal{W} \subseteq \mathcal{S}$,*

$$\mathcal{W} := \left\{ x \in X \;\middle|\; \hat{f}(x) \leq \nu \right\}$$

is closed.

Later, in Assumption 2.7 we introduce a condition which is equivalent to Assumption 2.4 but which has a more intuitive physical interpretation.

PROPOSITION 2.5 (Existence of an optimal solution without inf-compactness). *Under the assumptions stated for the problem (2) in Section 1.1 and Assumptions 2.3 and 2.4, the problem (4) has at least one optimal solution.*

Proof: The proof proceeds as follows. We first consider a fixed design, $\bar{x}$, in $\mathcal{S}$, and the restriction of the problem (4) to this value of x. For this fixed value of x, and under Assumption 2.3, we can use the results of Frank and Wolfe (1956) and Eaves (1971) for quadratic programming problems to show that the restricted problem, defined only in terms of y, has an optimal solution. This is used to establish that $\hat{f}$ is lower semicontinuous on the set $\mathcal{W}$. Then, by the compactness of the set $\mathcal{W}$, the desired result is established.

The lower-level problem can be written in a KKT formulation as follows: for each $\ell \in \mathcal{L}$,

$$K(x)y_\ell - F_\ell + C^T \lambda_\ell \;=\; 0^m, \tag{7a}$$
$$Cy_\ell \;\leq\; b, \tag{7b}$$
$$\lambda_\ell^T (Cy_\ell - b) \;=\; 0, \tag{7c}$$
$$\lambda_\ell \;\geq\; 0^r. \tag{7d}$$

Note that the definition of the graph of $\mathcal{S}_\ell$ in (5) can be equivalently expressed as

$$\mathrm{gr}\, \mathcal{S}_\ell := \{\, (x, y_\ell) \in \mathfrak{R}^{n+m} \mid (x, y_\ell, \lambda_\ell) \text{ solves (7) for some } \lambda_\ell \in \mathfrak{R}^r \,\}.$$

We then follow the proof of Theorem 1.4.3, p. 60, in Luo et al. (1997), and for each subset $\mathcal{I}$ of $\{1, 2, \ldots, r\}$ and $\ell \in \mathcal{L}$, let

$$\mathcal{S}_{\ell, \mathcal{I}} := \{\, (x, y_\ell, \lambda_\ell) \in \mathcal{X} \times \mathfrak{R}^{m+r} \mid K(x)y_\ell - F_\ell + C^T \lambda_\ell = 0^m;$$
$$\lambda_{\ell, i} = 0, \; \nu_i^T y_\ell \leq b_i, \; i \in \mathcal{I};$$
$$\lambda_{\ell, i} \geq 0, \; \nu_i^T y_\ell = b_i, \; i \notin \mathcal{I} \,\};$$

and let

$$\mathcal{S}'_{\ell, \mathcal{I}} := \{\, (x, y_\ell) \in \mathfrak{R}^{n+m} \mid (x, y_\ell, \lambda_\ell) \in \mathcal{S}_{\ell, \mathcal{I}} \text{ for some } \lambda_\ell \in \mathfrak{R}^r_+ \,\}.$$

Next, consider any fixed $\bar{x}$ in the set $\mathcal{S}$ of designs for which an equilibrium exists for every $\ell \in \mathcal{L}$, and the corresponding restriction of the problem (4). For each such $\bar{x} \in \mathcal{S}$, the restriction of $\mathcal{S}'_{\ell, \mathcal{I}}$ to $\bar{x}$ is a (possibly empty) polyhedron. Furthermore, the feasible set of the restriction of the problem (4) to $\bar{x}$ is in fact the product of the sets $\cup_{\mathcal{I}} \mathcal{S}'_{\ell, \mathcal{I}}$, $\ell \in \mathcal{L}$ (with $x := \bar{x}$). By assumption, f is lower bounded on $\mathcal{Z} \cap \mathrm{gr}\, \mathcal{S}_\ell$, $\ell \in \mathcal{L}$. It must therefore in particular be lower bounded on each nonempty polyhedral set $\mathcal{S}'_{\ell, \mathcal{I}}$ (with $x := \bar{x}$). Since $f(\bar{x}, \cdot)$ is also quadratic, we can utilize the results of Frank and Wolfe (1956) and Eaves (1971) to conclude that $f(\bar{x}, \cdot)$ attains its minimum on each nonempty polyhedral set $\mathcal{S}'_{\ell, \mathcal{I}}$ (with $x := \bar{x}$) and, since $\mathcal{I}$ is finite, also on the feasible sets $\cup_{\mathcal{I}} \mathcal{S}'_{\ell, \mathcal{I}}$, $\ell \in \mathcal{L}$ (with $x := \bar{x}$).

By these arguments, the minimum in (6) is attained. We next show that $\hat{f}$ is lower semicontinuous on $\mathcal{W}$. Let $\bar{x} \in \mathcal{W}$, and consider any sequence $\{x^t\} \subset \mathcal{W}$ which converges to $\bar{x}$. Let $\{y_\ell^t\} \subset \mathcal{Y}$, $\ell \in \mathcal{L}$, be sequences of the corresponding equilibrium solutions which define $\hat{f}(x^t)$. Let also $\bar{y}_\ell$, $\ell \in \mathcal{L}$, be equilibrium

solutions which define the value $\hat{f}(\bar{x})$. For each $\ell \in \mathcal{L}$, let $\hat{y}_\ell$ be an arbitrary accumulation point of the sequence $\{y_\ell^t\}$. Then, by the continuity of f, in these subsequences,

$$\{\hat{f}(x^t)\} = \left\{\sum_{\ell \in \mathcal{L}} \rho_\ell f(x^t, y_\ell^t)\right\} \to \sum_{\ell \in \mathcal{L}} \rho_\ell f(\bar{x}, \hat{y}_\ell) \geq \sum_{\ell \in \mathcal{L}} \rho_\ell f(\bar{x}, \bar{y}_\ell) = \hat{f}(\bar{x}).$$

Since the accumulation points were arbitrary, it follows that

$$\liminf_{t \to \infty} \hat{f}(x^t) \geq \hat{f}(\bar{x}),$$

whence we conclude that $\hat{f}$ is lower semicontinuous on $\mathcal{W}$. Then, by Weierstrass' Theorem, it attains its minimum on $\mathcal{W}$. Any such minimizer clearly solves the problem (4). This completes the proof. □

REMARK 2.6 (Existence of an optimal solution under extended compliance). Consider the upper-level objective given by the extended compliance function, that is, $f(x, y_\ell, \lambda_\ell) := F_\ell^T y_\ell + b^T \lambda_\ell$. As previously discussed, the bilevel problem reduces in this case to that of finding the stiffest structure, and can be equivalently written as a convex–concave saddle-point problem in terms of the given function f. Existence of a solution to this problem can be deduced from Proposition 2.5 as follows. We first note that the above result is easily extended to cover the case where also the multiplier values λ_ℓ enter into the upper-level objective function such that it is quadratic in both y_ℓ and λ_ℓ simultaneously (for each fixed x). Since the primary objective here is linear in y_ℓ and λ_ℓ, it suffices to show that f is lower bounded on $\operatorname{gr} \mathcal{S}_\ell$, $\ell \in \mathcal{L}$, to establish that Assumption 2.3 is satisfied. In general, $b \geq 0^r$. It then follows directly from (7) and the positive semi-definiteness of $K(x)$ on $\mathcal{X}$ that $f(x, y_\ell, \lambda_\ell) \geq 0$ on $\operatorname{gr} \mathcal{S}_\ell$, $\ell \in \mathcal{L}$. To prove that Assumption 2.4 is fulfilled, we utilize Proposition 4.1 of Achtziger (1998), which states that in this case the implicit objective function, $\hat{f}$, is l.s.c. on $\Re^n$. Then it follows that the set $\mathcal{W}$ is closed. □

Note that when $\hat{f}$ is not l.s.c. on $\Re^n$, Assumption 2.4 may not be easy to verify. It may, however, be possible to bound the minimum (equilibrium) displacements for a particular volume vector. We show that such a condition exists and is equivalent to that of a closed feasible set.

ASSUMPTION 2.7 (Bounding the graph). *For some $\bar{x} \in \mathcal{S}$ and $r > 0$, it holds that for every $x \in \mathcal{S}$ with $\hat{f}(x) \leq \hat{f}(\bar{x})$,*

$$\min \left\{ \|y_\ell\| \,\middle|\, f(x, y_\ell) = \min_{z_\ell : (x, z_\ell) \in \operatorname{gr} \mathcal{S}_\ell} f(x, z_\ell), \right\} \leq r, \qquad \ell \in \mathcal{L}. \tag{8}$$

In other words, we assume that it is possible to obtain a bounded graph by considering only the equilibrium displacements y_ℓ of minimum norm and which minimize the upper-level objective for each feasible design, if one includes only those designs with an upper-level objective value $\hat{f}(x)$ lower than some reference value $\hat{f}(\bar{x})$. We next show that Assumptions 2.4 and 2.7 are equivalent.

PROPOSITION 2.8 (Equivalence of Assumptions 2.4 and 2.7). *Under the assumptions stated for the problem (2) in Section 1.1 and Assumption 2.3, the Assumptions 2.4 and 2.7 are equivalent.*

Proof: Let Assumption 2.4 be satisfied. We first show that Assumption 2.7 is also satisfied. Let $\bar{x} \in \mathcal{S}$ be such that $\hat{f}(\bar{x}) \leq \nu$. Then, by the compactness of the set $\mathcal{W}$, we set

$$r_\ell := \max_{x \in \mathcal{W}} \min \left\{ \|y_\ell\| \,\middle|\, f(x, y_\ell) = \min_{z_\ell : (x, z_\ell) \in \operatorname{gr} \mathcal{S}_\ell} f(x, z_\ell), \right\} < \infty, \qquad \ell \in \mathcal{L}.$$

Let $r := \max_{\ell \in \mathcal{L}} r_\ell$. Hence, Assumption 2.7 is fulfilled.

Conversely, let Assumption 2.7 hold. Let $\nu = \hat{f}(\bar{x})$. Clearly, $\mathcal{W}$ is nonempty and bounded. Let $\{x^t\} \subset \mathcal{W}$ converge to x. Let $\{y_\ell^t\}$, $\ell \in \mathcal{L}$, be the corresponding sequences of equilibrium solutions fulfilling (8). By Assumption 2.7, these sequences are bounded. Let $\hat{y}_\ell$, $\ell \in \mathcal{L}$, be arbitrary accumulation points of the respective sequence. Then, by the closedness of $\operatorname{gr} \mathcal{S}_\ell$, $(x, \hat{y}_\ell) \in \operatorname{gr} \mathcal{S}_\ell$, $\ell \in \mathcal{L}$. Hence, $x \in \mathcal{S}$. Moreover,

by the definition of $\hat{f}$, the arbitrariness of the choice of accumulation points $\hat{y}_\ell$, $\ell \in \mathcal{L}$, and the fact that $\{x^t\} \subset \mathcal{W}$,

$$\hat{f}(x) \le \sum_{\ell \in \mathcal{L}} \rho_\ell f(x, \hat{y}_\ell) \le \limsup_{t \to \infty} \sum_{\ell \in \mathcal{L}} \rho_\ell f(x^t, y_\ell^t) \le \nu.$$

Hence, $x \in \mathcal{W}$, and we conclude that $\mathcal{W}$ is closed. This completes the proof. □

Note that the preceding existence results allow for zero designs, and are therefore general existence results for topology optimization problems. The results clearly allow for the incorporation of constraints on both levels. Furthermore, they permit the use of more general upper-level objectives than previous models which restrict the upper-level function to that of extended compliance and result in a saddle-point formulation.

The remaining results concern the continuity and differentiability properties of the implicitly defined upper-level objective function $x \mapsto \tilde{f}(x) := \sum_{\ell \in \mathcal{L}} \rho_\ell f(x, y_\ell(x))$. To effectively solve the stochastic topology optimization model, we will require the behavioural response of the structure to be unique, whence we, from here on, fix the topology by assuming that $l > 0^n$. However, we can still treat the true topology optimization problem by considering a sequence $\{l^t\}$ of positive lower bounds which decrease to zero. Proposition 2.11 on the continuity of the optimal value of the upper-level function justifies this approach.

PROPOSITION 2.9 (Properties of the implicit upper-level objective function). *Consider the conditions of either Proposition 2.2 or Proposition 2.5. Let the following three additional assumptions hold:* (i) $l > 0^n$; (ii) *no more than one contact condition is imposed for each node of the structure; and* (iii) *the upper-level objective function $f : \mathcal{Z} \mapsto \Re$ is continuously differentiable on $\mathcal{Z}$. Then, the lower-level problem problem* (4b) *has a unique solution, $y_\ell(x)$, for every $x \in \mathcal{X}$ and $\ell \in \mathcal{L}$, and the implicit function $x \mapsto \tilde{f}(x) := \sum_{\ell \in \mathcal{L}} \rho_\ell f(x, y_\ell(x))$ is locally Lipschitz continuous on $\mathcal{X}$ and directionally differentiable in any direction.*

Proof: By the sensitivity analysis of Robinson (1980, 1991), the result holds if, for each $\ell \in \mathcal{L}$, the lower-level objective $\Pi_\ell(x, \cdot)$ is strongly convex for each $x \in \mathcal{X}$, the gradients of the binding lower-level constraints at the solution to the lower-level problem (4b) for any given $x \in \mathcal{X}$ are linearly independent, and the function f is continuously differentiable and $\Pi_\ell(x, \cdot)$ and the functions defining the set $\mathcal{Y}$ are twice continuously differentiable. But the first property follows from (i), the second property follows from (ii) because the rows of C then will have no non-zero components at the same positions (Petersson and Patriksson, 1997), and the third follows from (iii). □

Proposition 2.11 extends Corollary 6.1 of Achtziger (1998) on the continuity of the optimal value of the minimax compliance function with respect to the lower bound l, to the more general case of the stochastic bilevel model of (4). In the proof of this result, we consider the following equivalent reformulation of the problem (4), which explicitly accounts for the possible non-uniqueness of the displacement vector y in the presence of zero lower design bounds:

$$\min_{x \in \mathcal{X}} \hat{f}(x), \tag{9}$$

where the definition of $\hat{f}$ is taken from (6).

Furthermore, let $\{l^t\}$ be a sequence of lower design bounds such that $0^n < l^{t+1} < l^t$ and $\{l^t\} \to 0^n$, define

$$\mathcal{X}^t := \{ x \in \Re^n \mid 1^T x = V; \quad l^t \le x \le u \},$$

and let

$$\bar{\mathcal{X}} := \{ x \in \Re^n \mid 1^T x = V; \quad 0^n \le x \le u \},$$

which is also the limit set of the sequence $\{\mathcal{X}^t\}$. For the simplicity of notation, we shall also define $y := \prod_{\ell \in \mathcal{L}} y_\ell \in \Re^{|\mathcal{L}|m}$. Let further $\bar{\mathcal{X}}^*$ be the set of optimal designs for the limit problem.

LEMMA 2.10 (Existence of a feasible sequence). *Let $x \in \bar{\mathcal{X}}$. Then, there exists a sequence $\{\tilde{x}^t\}$ with $\tilde{x}^t \in \mathcal{X}^t$ such that $\{\tilde{x}^t\} \to x$.*

Proof: Follows from Proposition 1.1.2 of Aubin and Frankowska (1990). ∎

PROPOSITION 2.11 (Continuity of the optimal solution). *Consider the conditions of either Proposition* 2.2 *or Proposition* 2.5. *Assume that f is continuous on $\bar{\mathcal{X}} \times \Re^m$. For each t, let (x^t, y^t) be an optimal solution to the problem (4) for $\mathcal{X} := \mathcal{X}^t$, and let (x^*, y^*) be an optimal solution to the problem (4) for $\mathcal{X} := \bar{\mathcal{X}}$. Then, the sequence $\{x^t\}$ is bounded and converges to the set $\bar{\mathcal{X}}^*$ in the sense that $\{\min_{x \in \bar{\mathcal{X}}^*} \|x - x^t\|\} \to 0$ holds, and each limit point $(\bar{x}, \bar{y})$ of the sequence $\{(x^t, y^t)\}$ is an optimal solution to the problem (4). Moreover, $\{f(x^t, y^t)\} \to f(x^*, y^*)$.*

Proof: By the conditions of Proposition 2.2 or Proposition 2.5, an optimal solution exists to the problem (4) for $\mathcal{X} := \mathcal{X}^t$ for all t, as well as for $\mathcal{X} := \bar{\mathcal{X}}$. Further, since $x^t \in \mathcal{X}^t \subset \bar{\mathcal{X}}$, and $\bar{\mathcal{X}}$ is compact, the sequence $\{x^t\}$ is bounded. The sequence $\{f(x^t, y^t)\}$ is descending, since $\mathcal{X}^t \subset \mathcal{X}^{t+1}$ for all t. It is furthermore lower bounded, either by the conditions of Proposition 2.2 since the sequence $\{y^t\}$ then is bounded by the coercivity assumption on f, or by the lower boundedness assumption in Proposition 2.5. Hence, the sequence $\{f(x^t, y^t)\}$ converges.

Let $\bar{x} \in \bar{\mathcal{X}}$ be an arbitrary limit point of the sequence $\{x^t\}$. Let $(x^*, y^*) \in \bar{\mathcal{X}} \times \prod_{\ell \in \mathcal{L}} \mathcal{Y}$ be an arbitrary optimal solution to the problem (4) for $\mathcal{X} := \bar{\mathcal{X}}$. By Lemma 2.10, there exists a sequence $\{(\tilde{x}^t, \tilde{y}^t)\}$ with $(\tilde{x}^t, \tilde{y}^t) \in \mathcal{X}^t \times \prod_{\ell \in \mathcal{L}} \mathcal{Y}$ that converges to (x^*, y^*). By the optimality of (x^*, y^*) in the problem (4) for $\mathcal{X} := \bar{\mathcal{X}}$, the descending property of the sequence $\{f(x^t, y^t)\}$, the continuity of f and the definition of $\hat{f}(\bar{x})$ in (9), we obtain that

$$f(x^*, y^*) \leq \hat{f}(\bar{x}) := \sum_{\ell \in \mathcal{L}} \rho_\ell \min_{y_\ell : (\bar{x}, y_\ell) \in \operatorname{gr} \mathcal{S}_\ell} f(\bar{x}, y_\ell) \leq \lim_{t \to \infty} f(x^t, y^t). \tag{10}$$

Further, by the optimality of (x^t, y^t) in the problem (4) for $\mathcal{X} := \mathcal{X}^t$ and the continuity of f, we obtain that

$$\lim_{t \to \infty} f(x^t, y^t) \leq \lim_{t \to \infty} f(\tilde{x}^t, \tilde{y}^t) = f(x^*, y^*). \tag{11}$$

Combining (10) and (11), it follows that $\{f(x^t, y^t)\} \to f(x^*, y^*)$, and that $\bar{x}$ solves the problem (9). If further $\{(x^t, y^t)\}$ has a limit point $(\bar{x}, \bar{y})$, then in the corresponding convergent subsequence, $\{f(x^t, y^t)\} \to f(\bar{x}, \bar{y})$ holds, and by the above we therefore must have that $f(\bar{x}, \bar{y}) = f(x^*, y^*)$. By the closedness of $\operatorname{gr} \mathcal{S}_\ell$, $(\bar{x}, \bar{y}_\ell) \in \operatorname{gr} \mathcal{S}_\ell$ holds for all $\ell \in \mathcal{L}$, and hence $(\bar{x}, \bar{y})$ is optimal in the problem (4). ∎

The above proposition shows that the optimal solutions to (4) change continuously with the lower bound on the design constraints. This property justifies the algorithmic method presented in the next section for solving the topology optimization problem—that is, the use of very small values of l in order to approximate the case in which some bars $i \in \{1, 2, \ldots, n\}$ may be eliminated (and thus $x_i = 0$).

3 Algorithms

In this section, we present an easy-to-implement algorithm for the solution of the problem (4), along with an effective parallelization and load balancing strategy.

The algorithm utilizes the local Lipschitz continuity and directional differentiability of the implicit function $\tilde{f}$, established in Proposition 2.9.

3.1 Calculating subgradients of $\tilde{f}$

For a given $x \in \mathcal{X}$ and $\ell \in \mathcal{L}$, let y_ℓ be the (unique) solution to the lower-level problem (4b), and let $\mathcal{I}(x, \ell)$ denote the subset of $\{1, \ldots, r\}$ for which $\nu_i^T y_\ell = b_i$. We also introduce the subsets $\mathcal{I}_+(x, \ell)$ and $\mathcal{I}_0(x, \ell)$ of $\mathcal{I}(x, \ell)$ for which $\lambda_i > 0$ and $\lambda_i = 0$ holds, respectively. In the event that $\mathcal{I}_0(x, \ell)$ is nonempty (that is, strict complementarity does not hold at y_ℓ) we further introduce a subset $\mathcal{J}(x, \ell)$ of $\mathcal{I}(x, \ell)$ for which we require only that $\mathcal{I}(x, \ell) \supseteq \mathcal{J}(x, \ell) \supseteq \mathcal{I}_+(x, \ell)$ holds. Let $C_\mathcal{I}$ and $d_\mathcal{I}$ denote the sub-matrix of C and sub-vector of d where only rows $i \in \mathcal{I}$ are included.

A subgradient of $\tilde{f}$ at $x \in \mathcal{X}$ can be calculated directly according to the formula proposed by Outrata and Zowe (1995), since their conditions A1–A3 and LI are automatically satisfied by our model. We obtain the following:

For each $\ell \in \mathcal{L}$, solve the equilibrium problem (4b) and the following linear system of equations in order to obtain d_{y_ℓ}:

$$
\begin{bmatrix} K(x) & -C_{\mathcal{J}(x,\ell)}^T \\ C_{\mathcal{J}(x,\ell)} & 0^{m \times |\mathcal{J}(x,\ell)|} \end{bmatrix} \begin{pmatrix} d_{y_\ell} \\ d_{\lambda_{\mathcal{J}(x,\ell)}} \end{pmatrix} = \begin{pmatrix} -\nabla_y f(x, y_\ell) \\ 0^m \end{pmatrix}. \tag{12}
$$

Then,

$$
\xi_{\tilde{f}}(x) := \sum_{\ell \in \mathcal{L}} \rho_\ell \left\{ \nabla_x f(x, y_\ell) + [K_1 y_\ell, \dots, K_n y_\ell]^T d_{y_\ell} \right\} \in \partial \tilde{f}(x). \tag{13}
$$

The subgradient can be used in an algorithm for the heuristic solution of the problem (4) or be embedded within a more sophisticated algorithm. The following subgradient projection algorithm utilizes the fact that $\tilde{f}$ is differentiable almost everywhere and in particular so when y_ℓ, $\ell \in \mathcal{L}$, are strictly complementary:

Subgradient projection method. Given $x \in \mathcal{X}$, an initial step in the direction of an arbitrary element $-\xi_{\tilde{f}}(x) \in -\partial \tilde{f}(x)$ is taken, followed by a Euclidean projection onto $\mathcal{X}$; an Armijo-like backtracking line search in this step length can be made so that the resulting feasible solution has a sufficiently lower objective value. Note, however, that the Armijo-type search requires repeated evaluation of the implicit objective function, which can be overly costly in a large-scale application. In this case, simpler step length rules may be preferable.

We note that at points of nondifferentiability, this method may break down because the negative of the subgradient may then not be a descent direction; in order to obtain a well-defined iteration also at such points, we can envisage utilizing a step length which is the maximum of the one supplied by the backtracking line search and the result of a predetermined step length formula used in traditional subgradient optimization techniques.

Outrata and Zowe (1995) utilize the analog of the above analysis for calculating subgradients in a bundle method for the solution of a deterministic bilevel problem. This can be viewed as a more complex technique than the above which however will ensure convergence to a stationary point. For other algorithms based on the implicit function $\tilde{f}$, see Luo et al. (1997) and Outrata et al. (1998).

3.2 Parallel resolution strategy for the problem (4)

The calculation of a subgradient $\xi_{\tilde{f}}(x)$ of $\tilde{f}$ at a given $x \in \mathcal{X}$ amounts to solving the quadratic lower-level programs (4b), followed by the calculation of a subgradient through the system (13), each term of which is calculated through the system (12). It is clear that the vast majority of this work can be done in parallel, provided that the necessary number of processors are available, whereas if the number of scenarios is large, this is not possible. Consider, however, the case where we may identify a cluster of scenarios corresponding to vectors F_ℓ with similar values. Allocating these scenarios to the same processor, we may solve the corresponding lower-level problems (4b) by utilizing efficient reoptimization procedures given that any of them have been solved to optimality, since the optimal solution to any one of them is feasible and near-optimal to all the others. Furthermore, we remark that the matrix appearing in the linear system (12) does not depend on the value of y_ℓ, only on the set $\mathcal{J}(x, \ell)$ at y_ℓ, and those sets will differ only slightly within the cluster. For any subset of scenarios with the same set $\mathcal{J}(x, \ell)$, the corresponding systems (12) will hence only differ in their right-hand sides, facilitating their efficient solution through factorization techniques.

Also for scenarios with slightly differing sets $\mathcal{J}(x, \ell)$, efficient computational schemes can be devised; for example, sorting the scenarios such that $\mathcal{J}(x, \ell_1) \subset \mathcal{J}(x, \ell_2)$, and so on, one may solve the systems in sequence, expanding the matrix with the necessary rows and columns and utilizing the solution to the former system as a starting point in the search for the next. The fact that the choice of $\mathcal{J}(x, \ell)$ is arbitrary in the range of active constraints may be used to minimize the number of scenarios with distinct values of $\mathcal{J}(x, \ell)$. (The reader may note that when no unilateral contact conditions are imposed, the system (12) simplifies significantly, since it then reduces to a linear system with $K(x)$ as system matrix and $|\mathcal{L}|$ different right-hand sides.)

4 Numerical experiments

In this section we provide an evaluation of a stochastic truss topology design problem along with numerical results of the subgradient projection algorithm.

4.1 Implementational details

The subgradient optimization algorithm of Section 3.1 was implemented as follows. A very small positive lower design bound l was used, in accordance with the results of Proposition 2.11. The initial design is the uniform design solution $x_j^0 := V/n$, $j = 1, \ldots, n$; that is, for a given total volume V, the n initial bar volumes are the same. Then, given an iterate x^t, the independent subproblems (4b) are solved using a Gauss–Seidel (coordinate-wise minimization) method, in which we sequentially search for the equilibrium displacement of one node while holding the remaining node displacements fixed. Each iteration of this algorithm thus amounts to solving a sequence of single-variable equations (and for nodes in potential contact: followed by a simple projection onto the admissible domain). A subgradient $\xi_{\tilde{f}}(x^t)$ is then calculated using the formula (13), where the linear system (12) is solved using the same approach as for the quadratic programs (4b). In order to avoid having to calculate the implicit objective function $\tilde{f}$ repeatedly, we chose to replace the backtracking line search explained earlier with the classic step length formula for subgradient optimization,

$$s_t := \frac{a}{b+t}, \quad a > 0, \quad b \geq 0, \qquad t = 0, 1, \ldots \qquad (14)$$

The design is updated by

$$x^{t+1} := P_{\mathcal{X}}[x^t - s_t \xi_f(x^t)],$$

where $P_{\mathcal{X}}[z]$ denotes the Euclidean projection of the vector $z \in \Re^n$ onto the set $\mathcal{X}$. Note that this projection is performed in linear time by utilizing a median search among the breakpoints of the derivative of the Lagrangean dual function corresponding to the equivalent quadratic program. A reliable stopping criterion for the subgradient algorithm is not easily constructed, and we chose to accept using either a fixed number of iterations or to terminate when the objective value did not decrease sufficiently after a fixed number of successive iterations. In the experiments performed, a maximum of 200 main iterations are allowed, whereas in the quadratic programs (4b) and linear system (12) a maximum of 700 iterations of the Gauss–Seidel procedure is allowed per main iteration.

The algorithm utilizes the fact that the main computational effort is easily parallelizable. Indeed, both the quadratic program (4b) and the linear system (12) separate into $|\mathcal{L}|$ independent problems over the different independent external forces. The parallelization was realized through the use of the PVM library, over a network of workstations; the speed-up obtained was very nearly linear.

It is important to mention that the efficient performance of the implementation was greatly enhanced by the reoptimization of the equilibrium problems utilized within each processor, as described in the preceding section.

The algorithm is sensitive to the choice of the values of a and b in (14), which were necessary to calibrate for each problem instance. We found experimentally that a proper value of b was inversely proportional to the size of the problem.

To test the functionality of the algorithm, it was first run on a classic deterministic problem in stiffness optimization: the ten-bar truss (e.g., Kirsch, 1993, pp. 156–157). In about 20 iterations, the algorithm terminated with a design solution equivalent to that provided in Kirsch (1993).

We next turn to a larger three-dimensional structure that illustrates issues pertaining to stochasticity.

4.2 A stochastic truss topology optimization problem

We consider a 5 (width) by 3 (depth) by 3 (height) truss on a rectangular grid, as shown in Figure 1. The bars of the ground structure are all those connecting pairs of nodes separated by at most one node. The four nodes (A, B, C and the hidden node D) at the left-hand corners of the truss are fixed, node E is in unilateral contact from below, and node F, on the right-hand side of the truss, is subject to external forces.

In the following figures, the widths of the bars are proportional to the volumes as given by the vector x. Positions are illustrated through use of a grey-scale, in which a bar is darker than a bar behind it.

[**Place of Figure 1**]

Scenarios of upward and downward forces lead to highly different deterministic optimal designs, as seen in Figures 2 and 3. The objective function is here and in the following chosen as the (squared) norm of the displacement vector y.

[**Place of Figure 2**]

[**Place of Figure 3**]

We now consider a situation in which the structure is subject to a Gaussian distributed force with zero mean and a given variance.

To discretize Ω as defined in Section 2, we first initialize a discretization of the standard Gaussian one-dimensional distribution, on N points. For each independent force F_k, $k \in \{1, \ldots, K\}$, acting on the truss in the respective coordinate direction, this discretization will give N forces F_k^i. With this discretization Ω has $|\mathcal{L}| = N^K$ elements. In the experiments, $N = 9$ and $K = 2$ were chosen. Note that it is not necessary to keep all the different forces F_k^i in memory from one sweep through Ω to another.

In Figure 4, we consider the effect of a random vertical Gaussian force that interpolates between the two deterministic scenarios of Figures 2 and 3, with zero mean and a non-zero variance. Note that the response of the structure to forces will be nonlinear in the presence of obstacles in contact. The resulting structure represents a compromise between the structures of Figures 2 and 3, neither of which was able to respond well to a variation in the force direction.

[**Place of Figure 4**]

The principal flaw in the structure of Figure 3 is that for any force having a non-zero horizontal component, the structure is unable to attain a position of static equilibrium: nodes A and B are insufficient to fix the structure in such a setting. Figures 5 and 6 show stochastic optimal designs with respect to a random force with a mean vertical component equal to that in Figure 3 but with a Gaussian horizontal component having zero mean and variance equal to one half the value of the downward force. Finally, in Figure 6, the variance on the horizontal component was doubled. As might be expected, the resulting design is able to handle a wider range of forces.

[**Place of Figure 5**]

[**Place of Figure 6**]

The computational effort needed to solve the latter model is reasonable, as can be seen from the Figures 7 and 8. Indeed, both the implicit upper-level objective function value and the number of non-zero volume bars decrease quite rapidly initially, and the two values are reduced by a factor of more than 7 and 30, respectively, compared to their initial values in the uniform, unoptimized, structure.

[**Place of Figure 7**]

[**Place of Figure 8**]

5   Conclusions

The present paper serves to introduce a stochastic topology optimization problem, which explicitly takes into account the variability of the input data and provides a robust structural design. The mathematical properties of the model are presented, some of which generalize recently published results. We also contribute a very simple yet efficient subgradient optimization algorithm for this stochastic bilevel optimization problem and evaluate it numerically on a stochastic truss topology optimization problem.

References

Achtziger, W. (1998). Multiple-load truss topology and sizing optimization: Some properties of minimax compliance. *Journal of Optimization Theory and Applications*, **98**, pp. 255–280.

Aubin, J.-P. and Frankowska, H. (1990). *Set-Valued Analysis*. Birkhäuser, Boston.

Bendsøe, M. P. (1995). *Optimization of Structural Topology, Shape, and Material*. Springer-Verlag, Berlin.

Ben-Tal, A. and Nemirovski, A. (1997). Robust truss topology design via semidefinite programming. *SIAM Journal on Optimization*, **7**, pp. 991–1016.

Brandt, A., Jendo, S. and Marks, W. (1984). Probabilistic approach to reliability-based optimum structural design. *Engineering Transactions*, **32**, pp. 57–74.

Eaves, B. C. (1971). On quadratic programming. *Management Science*, **17**, pp. 698–711.

Frank, M. and Wolfe, P. (1956). An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, **3**, pp. 95–110.

Haslinger, J. and Neittaanmäki, P. (1996). *Finite Element Approximation for Optimal Shape, Material and Topology Design*. John Wiley & Sons, Chichester.

Jóźwiak, S. F. (1989). Probability-based optimization of truss structures. *Computers & Structures*, **32**, pp. 87–91.

Kirjner-Neto, C. Polak, E. and Der Kiureghian, A. (1998). An outer approximations approach to reliability-based optimal design of structures. *Journal of Optimization Theory and Applications*, **98**, pp. 1–16.

Kirsch, U. (1993). *Structural Optimization*. Springer-Verlag, Berlin.

Kočvara, M., Zibulevsky, M. and Zowe, J. (1998). Mechanical design problems with unilateral contact. *Mathematical Modelling and Numerical Analysis*, **32**, pp. 255–281.

Luo, Z.-Q. Pang, J.-S. and Ralph, D. (1997). *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, UK.

Marti, K. (1992). Stochastic optimization in structural design. *Zeitschrift für angewandte Mathematik und Mechanik*, **72**, pp. 452–464.

Marti, K., ed. (1997). Structural Reliability and Stochastic Structural Optimization. Special issue of *Mathematical Methods of Operations Research*, **46**, no. 3.

Outrata, J., Kočvara, M. and Zowe, J. (1998). *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Outrata, J. and Zowe, J. (1995). A numerical approach to optimization problems with variational inequality constraints. *Mathematical Programming*, **68**, pp. 105–130.

Patriksson, M. and Wynter, L. (1998). Stochastic mathematical programs with equilibrium constraints, PRISM report 97/017. PRISM, Université de Versailles, Versailles, France. (Revised for *Operations Research Letters*.)

Pedersen, P. and Jørgensen (1984). Minimum mass design of elastic frames subjected to multiple load cases. *Computers & Structures*, **18**, pp. 147–157.

Petersson, J. (1995). Optimization of Structures in Unilateral Contact. Dissertation No. 397, Department of Mechanical Engineering, Linköping Institute of Technology, Linköping, Sweden.

Petersson, J. and Patriksson, M. (1997). Topology optimization of sheets in contact by a subgradient method. *International Journal of Numerical Methods in Engineering*, **40**, pp. 1295–1321.

Robinson, S. M. (1980). Strongly regular generalized equations. *Mathematics of Operations Research*, **5**, pp. 43–62.

Robinson, S. M. (1991). An implicit-function theorem for a class of nonsmooth functions. *Mathematics of Operations Research*, **16**, pp. 282-309.

Shimizu, K., Ishizuka, Y. and Bard, J. (1996). *Nondifferentiable and Two-Level Mathematical Programming*. Kluwer Academic Publishers, Boston.

Tada, Y. and Seguchi, Y. (1989). Structural shape determination under uncertain loading conditions. *Computational Mechanics*, **4**, pp. 219–228.

Zhang, R. (1994). Problems of hierarchical optimization in finite dimension. *SIAM Journal on Optimization*, **4**, pp. 521–536.

Zowe, J., Kočvara, M. and Bendsœ, M. P. (1997). Free material optimization via mathematical programming. *Mathematical Programming*, **79**, pp. 445–466.
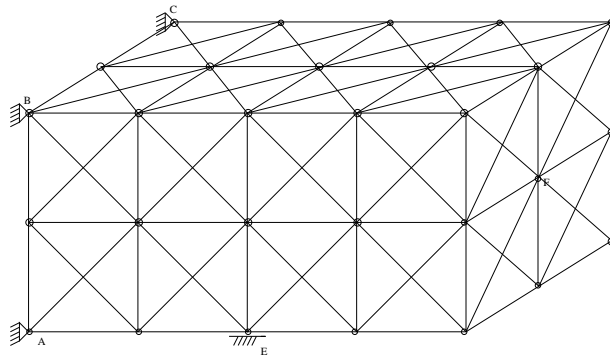
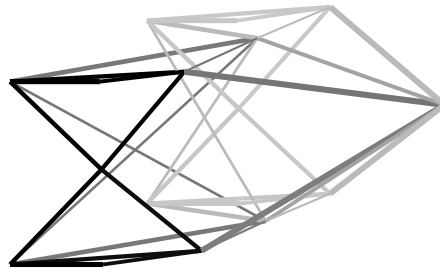FIGURE 1: Ground structure, no force applied

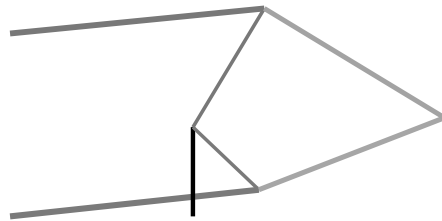FIGURE 2: Optimal structure, force directed upward
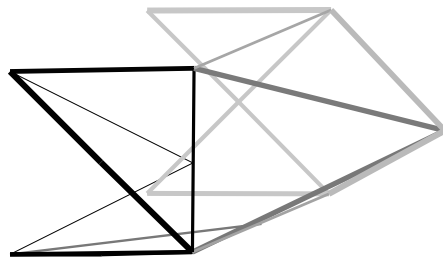
FIGURE 3: Optimal structure, force directed downward

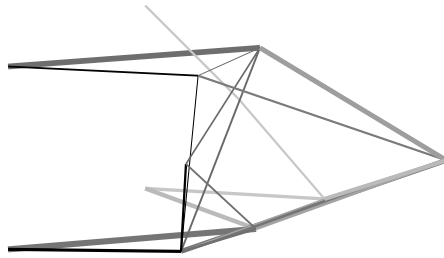FIGURE 4: Optimal structure, random vertical force

FIGURE 5: Optimal structure, random vertical force with downward mean
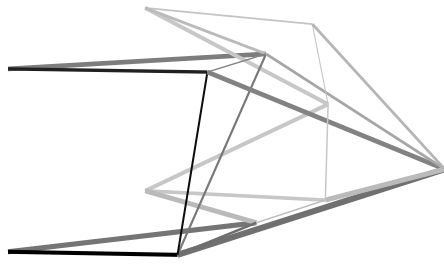
FIGURE 6: Optimal structure, random vertical force with downward mean, double variance
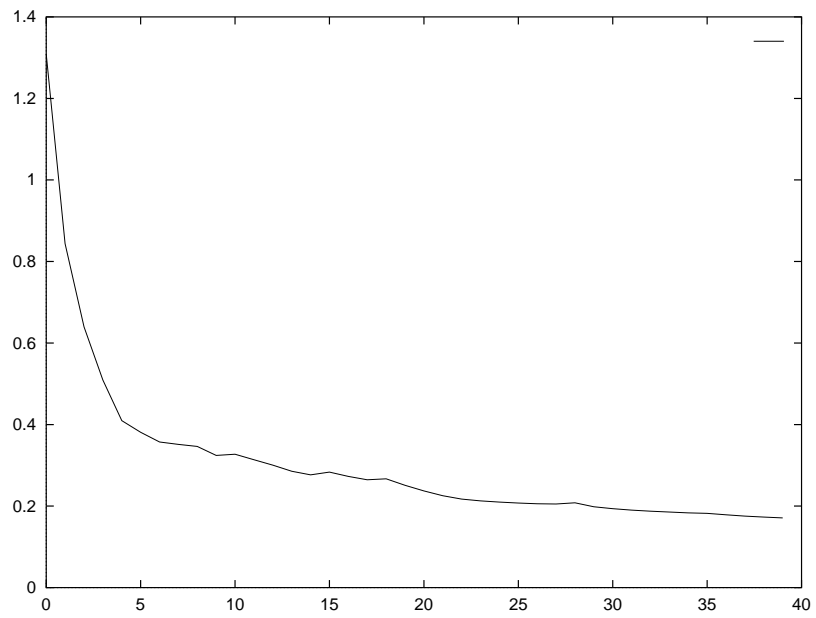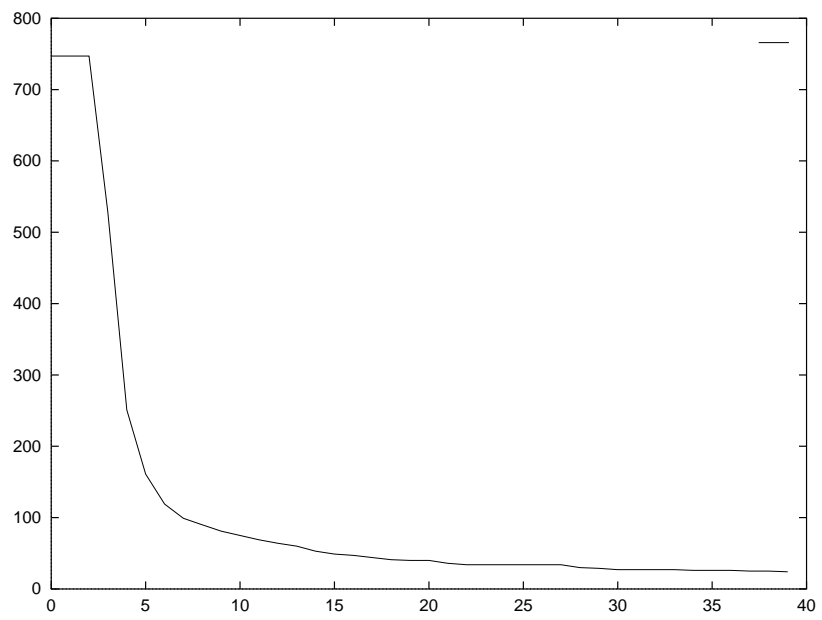

FIGURE 7: Value of the objective function versus iteration

FIGURE 8: Number of non-zero volume bars versus iteration