



HAL
open science

Global optimality in minimum compliance topology optimization of frames and shells by moment-sum-of-squares hierarchy

Marek Tyburec, Jan Zeman, Martin Kružík, Didier Henrion

► To cite this version:

Marek Tyburec, Jan Zeman, Martin Kružík, Didier Henrion. Global optimality in minimum compliance topology optimization of frames and shells by moment-sum-of-squares hierarchy. *Structural and Multidisciplinary Optimization*, 2021, 64 (4), pp.1963-1981. 10.1007/s00158-021-02957-5 . hal-02949791

HAL Id: hal-02949791

<https://laas.hal.science/hal-02949791>

Submitted on 26 Sep 2020

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

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

Global optimality in minimum compliance topology optimization of frames and shells by moment-sum-of-squares hierarchy

Marek Tyburec · Jan Zeman · Martin Kružík · Didier Henrion

the date of receipt and acceptance should be inserted later

Abstract The design of minimum-compliance bending-resistant structures with continuous cross-section parameters is a challenging task because of its inherent non-convexity. Our contribution develops a strategy that facilitates computing all guaranteed globally optimal solutions for frame and shell structures under multiple load cases and self-weight. To this purpose, we exploit the fact that the stiffness matrix is usually a polynomial function of design variables, allowing us to build an equivalent non-linear semidefinite programming formulation over a semi-algebraic feasible set. This formulation is subsequently solved using

Marek Tyburec¹
E-mail: marek.tyburec@fsv.cvut.cz
ORCID: 0000-0003-0798-0948

Jan Zeman¹
E-mail: Jan.Zeman@cvut.cz
ORCID: 0000-0003-2503-8120

Martin Kružík²
E-mail: kruzik@utia.cas.cz
ORCID: 0000-0003-1558-5809

Didier Henrion^{3,4}
E-mail: henrion@laas.fr
ORCID: 0000-0001-6735-7715

¹Czech Technical University in Prague, Faculty of Civil Engineering, Department of Mechanics, Thákurova 7, 16629 Prague 6, Czech Republic

²Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 16629 Prague 6, Czech Republic

³Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Control Engineering, Karlovo náměstí 13, 12135 Prague 2, Czech Republic

⁴LAAS-CNRS, 7 avenue du Colonel Roche, 31400 Toulouse, France

the Lasserre moment-sum-of-squares hierarchy, generating a sequence of outer convex approximations that monotonically converges from below to the optimum of the original problem. Globally optimal solutions can subsequently be extracted using the Curto-Fialkow flat extension theorem. Furthermore, we show that a simple correction to the solutions of the relaxed problems establishes a feasible upper bound, thereby deriving a simple sufficient condition of global ε -optimality. When the original problem possesses a unique minimum, we show that this solution is found with a zero optimality gap in the limit. These theoretical findings are illustrated on several examples of topology optimization of frames and shells, for which we observe that the hierarchy converges in a finite (rather small) number of steps.

Keywords discrete topology optimization, frame structures, shell structures, semidefinite programming, polynomial optimization, global optimality

1 Introduction

Structural optimization is a research field developing concepts for the design of efficient structures. It was pioneered by [Michell \(1904\)](#), who showed that minimum-weight truss structures under a single load case are fully-stressed, and the optimal trajectories of their bars align with the principal stress directions. Hence, optimal designs can contain an infinite number of bars in general. This drawback, which hinders their manufacturability, was overcome in the work of [Dorn *et al.* \(1964\)](#) by introducing the ground structure approach, effectively discretizing the continuum into a finite set of potential nodes and their interconnections of finite elements. Because the dimensionality of these potential elements is lower than that of the continuum and the presence and sizing of each element are investigated, this setting is referred to as discrete topology optimization.

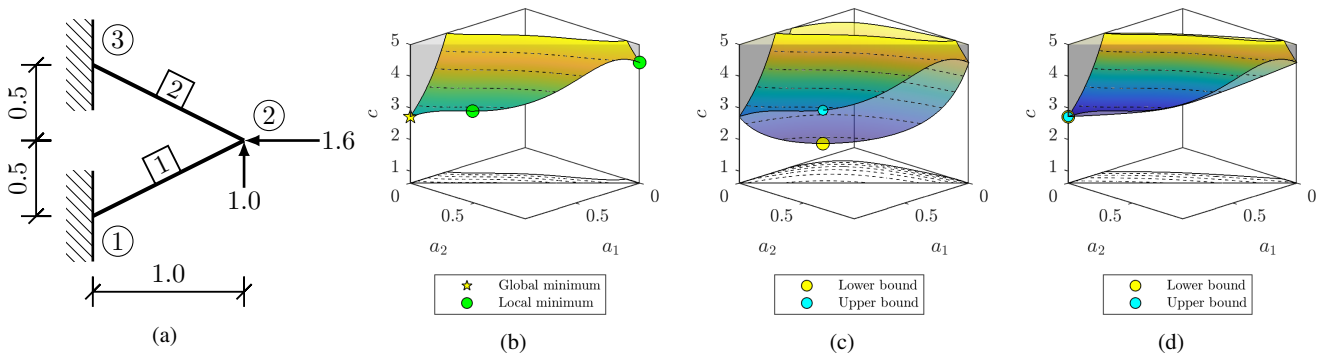


Fig. 1: (a) Boundary conditions for the motivating problem, (b) the sublevel set $c \leq 5$ of its feasible design space, and (c) first and (d) second convex outer approximations constructed by the moment-sum-of-squares approach. Variables a_1 and a_2 stand for the cross-section areas of the two elements and c denotes the corresponding compliance (assuming moments of inertia $I_i = a_i^2$, $i \in \{1, 2\}$).

A tremendous progress has been made for the case of trusses. While [Dorn et al. \(1964\)](#) developed a linear programming formulation for the single-load-case plastic design, [Bendsøe et al. \(1991\)](#) and [Achtziger et al. \(1992\)](#) introduced a convex displacement-based elastic-design quadratic program that additionally allowed for multiple load cases. Its dual, which incorporates the cross-section variables explicitly, was shown by [Lobo et al. \(1998\)](#) and [Ben-Tal and Nemirovski \(2001\)](#) to be a second-order conic program. This latter formulation handles multiple load cases with stress constraints efficiently ([Tyburec et al., forthcoming](#)). Convexity prevails even for fundamental free-vibration constraints, in which case semidefinite programming can be used ([Achtziger and Kočvara, 2008](#); [Tyburec et al., 2019](#)).

A completely different situation holds for bending-resistant structures with continuous design variables. To the best of our knowledge, no convex formulation has been established so far, and, therefore, solely local optimization techniques have been used. Among these, [Saka \(1980\)](#) developed a sequential linear programming approach to design minimum-weight framed structures, and [Wang and Arora \(2006\)](#) improved over its solution efficiency by using sequential quadratic programming instead. Another, relaxation-based sequential semidefinite programming method was proposed by [Yamada and Kanno \(2015\)](#) to deal with vibration problems. Nonlinear programming ([Fredricson et al., 2003](#)), Optimality Criteria (OC) ([Khan, 1984](#); [Chan et al., 1995](#)), the Method of Moving Assymptotes (MMA) ([Svanberg, 1987](#); [Fredricson, 2005](#)), and metaheuristics ([An and Huang, 2017](#)) are other commonly used alternatives.

Except for our earlier conference paper ([Tyburec et al., 2020a](#)), which is a very preliminary version of this manuscript, the only global approach that can be found in the literature considers a discrete setting of the problem, in which case the cross-sections are selected from a predefined cata-

log, allowing to use the branch-and-bound method ([Kanno, 2016](#)).

1.1 Motivation

This contribution investigates a conceptual design of bending-resistant structures with continuous design variables. As follows from the previous survey, only local optimization approaches have been adopted so far to tackle the frame/shell structure optimization problem. It is, therefore, not surprising that they fail to converge to globally optimal solutions even for toy problems, such as the one shown in Fig. 1a.

Consider a linear-elastic material with the dimensionless Young modulus $E = 1$, available material volume $\bar{V} = 1$, element cross-sections parameterized by their area a_i and the moment of inertia $I_i(a_i) = a_i^2$, which corresponds to rectangular cross-sections with the height-to-width ratio of 12, for example. Our goal is to find the cross-section areas a_i associated with the elements $i = \{1, 2\}$ that induce minimum compliance c within all non-negative a_i satisfying the volume bound \bar{V} . Standard local optimization techniques such as OC, MMA, and MATLAB inbuilt optimizer `fmincon` all converge¹ to the optimized compliance $c = 2.895$ and the corresponding areas $a_1 = 0.652$ and $a_2 = 0.242$. However, the globally optimal design possesses compliance $c^* = 2.719$, which requires $a_1^* = 0.4\sqrt{5}$ and $a_2^* = 0$.

This problem exhibits three local optima in particular, with the last one being $a_1 = 0$, $a_2 = 0.4\sqrt{5}$ and $c = 4.429$, see Fig. 1b. Thus, the global optimum may be reached by examining a few different starting points in the optimizers. Such a procedure, however, cannot assure global optimality and cannot neither assess quality of

¹ For OC and MMA, we adopted the commonly-used starting point of uniform mass distribution, i.e., $a_1 = a_2 = 0.2\sqrt{5}$. For `fmincon`, the default starting point was used.

the optimized designs with respect to the global optimum. When the number of structural elements increases, design-dependent loads are considered, and higher-order polynomials for the moments of inertia are used, finding globally-optimal minimum-compliance designs becomes extremely challenging.

1.2 Aims and novelty

We address this problem by exploiting the simple fact that the constraints can be formulated as polynomial functions, hence forming a (basic) semi-algebraic feasible set. Using a polynomial objective function in addition, the moment-sum-of-squares (Lasserre) hierarchy of convex outer approximations (relaxations) can be used to solve and extract all globally-optimal solutions. These relaxations provide a non-decreasing sequence of lower bounds, eventually attaining the optimal compliance, Figs 1c-1d. In addition, we show how to correct such obtained lower-bound designs, and hence generate feasible upper bounds. A comparison of these bounds then assesses the design quality, and their equality establishes a simple sufficient condition of global optimality. We further show that when a unique global optimum exists, we can expect the occurrence of a bound equality. Fortunately, this situation occurs quite often when the design domain lacks structural and boundary conditions symmetries.

This paper is organized as follows. In Section 2, we introduce polynomial optimization and the moment-sum-of-squares hierarchy. Section 3.1 develops a non-linear semidefinite programming formulation for topology optimization of frame structures, which we modify subsequently for the moment-sum-of-squares hierarchy in Section 3.2. Section 3.3 reveals how to correct the lower-bound designs generated by the hierarchy to obtain feasible upper bounds, and Section 3.4 introduces the sufficient condition of global ε -optimality as well as a zero optimality gap for the case of a unique global optimum. Section 3.5 outlines the required changes in notation to allow for thickness optimization of shell structures. These results are then illustrated on five selected optimization problems in Section 4. We finally summarize our contributions in Section 5.

2 Moment-sum-of-squares hierarchy

In this section, we briefly outline the moment-sum-of-squares hierarchy when applied to solution of problems with polynomial matrix inequalities. For more information, we refer the reader to an expository text (Henrion and Lasserre, 2006) and to the excellent books (Anjos and Lasserre, 2012; Lasserre, 2015).

Suppose we aim to solve an optimization problem of the form

$$f^* = \inf_{\mathbf{x}} f(\mathbf{x}) \quad (1a)$$

$$\text{s.t. } \mathbf{G}(\mathbf{x}) \succeq 0, \quad (1b)$$

where $f(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$ is a real polynomial function and $\mathbf{G}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{S}^m$ is a real polynomial mapping, so that $\forall i, j : G_{i,j}(\mathbf{x}) = G_{j,i}(\mathbf{x})$ are real polynomial functions of \mathbf{x} . The degree of these polynomials is less than or equal to $k \in \mathbb{N}$. The symbol \mathbb{S}^m denotes the space of real symmetric square matrices of size m , and “ \succeq ” establishes an ordering of fundamental eigenvalues, i.e., $\mathbf{G}(\mathbf{x})$ in (1b) is positive semidefinite. Hence, we call (1b) a polynomial matrix inequality (PMI) in what follows and denote its feasible set by $\mathcal{K}(\mathbf{G})$.

Clearly, the nonlinear semidefinite program (1) covers a variety of convex optimization problems as special cases, including linear and quadratic programming or linear semidefinite programming, see, e.g., (Ben-Tal and Nemirovski, 2001, Section 4.2). Although these instances can be solved in a polynomial time, hence efficiently, (1) exhibits \mathcal{NP} -hardness in general. This can be seen, for example, by a reduction from binary programming, in which case the main diagonal of $\mathbf{G}(\mathbf{x})$ contains both $x_i^2 - x_i$ and $x_i - x_i^2$ terms for all $i \in \{1 \dots n\}$.

Despite the fact that the admissible set of \mathbf{x} is generally non-convex, (1) admits an equivalent reformulation to a convex optimization problem over a finite-dimensional cone of polynomials $C_k(\mathcal{K}(\mathbf{G}))$ of degree at most k which are non-negative on $\mathcal{K}(\mathbf{G})$, i.e.,

$$\begin{aligned} f^* &= \sup_{\lambda} \{ \lambda : f(\mathbf{x}) - \lambda \geq 0, \forall \mathbf{x} \in \mathcal{K}(\mathbf{G}) \} \\ &= \sup_{\lambda} (f - \lambda), \text{ s.t. } (f - \lambda) \in C_k(\mathcal{K}(\mathbf{G})). \end{aligned} \quad (2)$$

Unfortunately, it is not known how to handle $C_k(\mathcal{K}(\mathbf{G}))$ simply and in a tractable way.

To introduce an approach that allows to solve (2), we first adopt the following notation. Let $\mathbf{x} \mapsto \mathbf{b}_k(\mathbf{x})$ be the polynomial space basis of polynomials of degree at most k ,

$$\mathbf{b}_k(\mathbf{x}) = \begin{pmatrix} 1 & x_1 & x_2 & \dots & x_n & x_1^2 & x_1x_2 & \dots & x_1x_n \\ & x_2^2 & x_2x_3 & \dots & x_n^2 & \dots & x_2^3 & \dots & x_n^k \end{pmatrix}, \quad (3)$$

Then, any polynomial $p(\mathbf{x})$ of degree at most k can be written as

$$p(\mathbf{x}) = \mathbf{q}^T \mathbf{b}_k(\mathbf{x}), \quad (4)$$

in which \mathbf{q} denotes a vector of coefficients associated with the basis $\mathbf{b}_k(\mathbf{x})$.

Definition 1 The polynomial matrix $\Sigma(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{S}^m$ is a (matrix) sum-of-squares (SOS) if there exists a polynomial matrix $\mathbf{H}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times o}$, $o \in \mathbb{N}$, such that

$$\Sigma(\mathbf{x}) = \mathbf{H}(\mathbf{x}) [\mathbf{H}(\mathbf{x})]^T, \quad \forall \mathbf{x} \in \mathbb{R}^n. \quad (5)$$

Let $\langle \cdot, \cdot \rangle$ denote the standard inner product on matrices, $\alpha \in \mathbb{N}^{|\mathbf{b}_k(\mathbf{x})|}$ with $\mathbf{1}^\top \alpha \leq k$ be the multi-index associated with the basis $\mathbf{b}_k(\mathbf{x})$, and let $\mathbf{y} \in \mathbb{R}^{|\mathbf{b}_k(\mathbf{x})|}$ be the moments (of probability measures supported on $\mathcal{K}(\mathbf{G}(\mathbf{x}))$) indexed in $\mathbf{b}_k(\mathbf{x})$. In what follows, we adopt the following notation for the elements of \mathbf{y} :

$$y_\alpha = y_{\prod_{i=1}^n x_i^{\alpha_i}} \text{ is associated with } \prod_{i=1}^n x_i^{\alpha_i}. \quad (6)$$

For example, when $\alpha = (0 \ 0 \ 1 \ 2)^\top$, $y_{0012} = y_{x_3^1 x_4^2}$ corresponds to the polynomial $x_3^1 x_4^2 \in \mathbf{b}_k(\mathbf{x})$, where $k \geq 3$.

Assumption 1. (*Henrion and Lasserre, 2006*) Assume that there exist SOS polynomials $\mathbf{x} \mapsto p_0(\mathbf{x})$ and $\mathbf{x} \mapsto \mathbf{R}(\mathbf{x})$ such that the superlevel set $\{\mathbf{x} \in \mathbb{R}^n : p_0(\mathbf{x}) + \langle \mathbf{R}(\mathbf{x}), \mathbf{G}(\mathbf{x}) \rangle \geq 0\}$ is compact.

Note that Assumption 1 is an algebraic certificate of compactness of the feasible set in problem (1). When Assumption 1 holds, then, the dual of (2) can be written equivalently as an infinite-dimensional generalized problem of moments, which is equipped with a finite-dimensional truncation:

$$f^{(r)} = \min_{\mathbf{y}} \mathbf{q}^\top \mathbf{y} \quad (7a)$$

$$\text{s.t.} \quad y_0 = 1, \quad (7b)$$

$$\mathbf{M}_k(\mathbf{y}) \succeq 0, \quad (7c)$$

$$\mathbf{M}_{k-d}(\mathbf{G}(\mathbf{x})\mathbf{y}) \succeq 0, \quad (7d)$$

in which $2r \geq k$, r is the relaxation degree, and d stands for the maximum degree of polynomials in $\mathbf{G}(\mathbf{x})$. In addition, $\mathbf{M}_k(\mathbf{y})$ and $\mathbf{M}_{k-d}(\mathbf{G}(\mathbf{x})\mathbf{y})$ are the truncated moment and localizing matrices associated with \mathbf{y} and $\mathbf{G}(\mathbf{x})$. For a precise definition of $\mathbf{M}_k(\mathbf{y})$ and $\mathbf{M}_{k-d}(\mathbf{G}(\mathbf{x})\mathbf{y})$, we refer the reader to (*Henrion and Lasserre, 2006*). These moment matrices are linear in \mathbf{y} , hence (7) is a linear semidefinite program. Because (7) is a finite-dimensional convex relaxation of (1), we have $f^{(r)} \leq f^*$, $\forall r \in \mathbb{N}$. Moreover, these relaxations are tighter with increasing r , making the sequence $(f^{(r)})_r$ monotonically increasing and converging towards f^* .

Theorem 1 (*Henrion and Lasserre, 2006, Theorem 2.2*) Let Assumption 1 be satisfied. Then, $f^{(r)} \uparrow f^*$ as $r \rightarrow \infty$ in (7).

Moreover, all globally optimal solutions of (1) can be extracted from (7) based on the flat extension theorem of *Curto and Fialkow (1996)*. Indeed, finite convergence occurs when

$$s = \text{Rank}(\mathbf{M}_k(\mathbf{y}^*)) = \text{Rank}(\mathbf{M}_{k-d}(\mathbf{y}^*)), \quad (8)$$

where \mathbf{y}^* denotes the vector of optimal moments, and s stands for the minimum number of distinct global minimizers (*Henrion and Lasserre, 2006, Theorem 2.4*).

Example We illustrate the process of building the (Lasserre) moment-sum-of-squares hierarchy on an elementary example.

$$\min_{a,c} c \quad (9a)$$

$$\text{s.t.} \quad \begin{pmatrix} c & \bar{f} \\ \bar{f} & a^2 \end{pmatrix} \succeq 0, \quad (9b)$$

$$\bar{V} - a \geq 0, \quad (9c)$$

$$a \geq 0. \quad (9d)$$

In the first relaxation, $\mathbf{y} = (y_{00} \ y_{10} \ y_{01} \ y_{20} \ y_{11} \ y_{02})^\top$ is indexed in the polynomial space basis $\mathbf{b}_1(a, c) = (1 \ c \ a \ c^2 \ ca \ a^2)^\top$. Then, the associated relaxation reads

$$\min_{\mathbf{y}} y_{10} \quad (10a)$$

$$\text{s.t.} \quad \begin{pmatrix} y_{10} & \bar{f} \\ \bar{f} & y_{02} \end{pmatrix} \succeq 0, \quad (10b)$$

$$\bar{V} - y_{01} \geq 0, \quad (10c)$$

$$y_{01} \geq 0, \quad (10d)$$

$$y_{00} = 1, \quad (10e)$$

$$\begin{pmatrix} y_{00} & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{pmatrix} \succeq 0. \quad (10f)$$

For $r = 2$, we have $\mathbf{y} = (y_{00} \ y_{10} \ y_{01} \ y_{20} \ y_{11} \ y_{02} \ y_{30} \ y_{21} \ y_{12} \ y_{03} \ y_{40} \ y_{31} \ y_{22} \ y_{13} \ y_{04})^\top$ indexed in $\mathbf{b}_2(a, c) = (1 \ c \ a \ c^2 \ ca \ a^2 \ c^3 \ c^2 a \ ca^2 \ a^3 \ c^4 \ c^3 a \ c^2 a^2 \ ca^3 \ a^4)^\top$. The corresponding relaxation is written as

$$\min_{\mathbf{y}} y_{10} \quad (11a)$$

$$\text{s.t.} \quad \begin{pmatrix} y_{10} & \bar{f} & y_{20} & \bar{f}y_{10} & y_{11} & \bar{f}y_{01} \\ \bar{f} & y_{02} & \bar{f}y_{10} & y_{12} & \bar{f}y_{01} & y_{03} \\ y_{20} & \bar{f}y_{10} & y_{30} & \bar{f}y_{20} & y_{21} & \bar{f}y_{11} \\ \bar{f}y_{10} & y_{12} & \bar{f}y_{20} & y_{22} & \bar{f}y_{11} & y_{13} \\ y_{11} & \bar{f}y_{01} & y_{21} & \bar{f}y_{11} & y_{12} & \bar{f}y_{02} \\ \bar{f}y_{01} & y_{03} & \bar{f}y_{11} & y_{13} & \bar{f}y_{02} & y_{04} \end{pmatrix} \succeq 0, \quad (11b)$$

$$\begin{pmatrix} \bar{V} - y_{01} & \bar{V}y_{10} - y_{11} & \bar{V}y_{01} - y_{02} \\ \bar{V}y_{10} - y_{11} & \bar{V}y_{20} - y_{21} & \bar{V}y_{11} - y_{12} \\ \bar{V}y_{01} - y_{02} & \bar{V}y_{11} - y_{12} & \bar{V}y_{02} - y_{03} \end{pmatrix} \succeq 0, \quad (11c)$$

$$\begin{pmatrix} y_{01} & y_{11} & y_{02} \\ y_{11} & y_{21} & y_{12} \\ y_{02} & y_{12} & y_{03} \end{pmatrix} \succeq 0, \quad (11d)$$

$$y_{00} = 1, \quad (11e)$$

$$\begin{pmatrix} y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{pmatrix} \succeq 0. \quad (11f)$$

When solved, this relaxation allows for extracting the global solution of $a^* = \bar{V}$ and $c^* = \bar{f}^2 / \bar{V}^2$.

3 Methodology

Topology optimization of discrete structures provides a natural application for the ground structure approach (Dorn *et al.*, 1964), a discretized design domain composed of a fixed set of $n_n \in \mathbb{N}$ nodes and their subsets of admissible $n_e \in \mathbb{N}$ finite elements. Here, we employ the simplest two-node Euler-Bernoulli frame elements that adopt linear shape functions to interpolate the longitudinal displacements and cubic shape functions to interpolate the lateral displacements and rotations. Another elements can be adopted though, see Section 4.4 for applications to the Timoshenko beam element and the MITC4 shell element.

Each of these finite elements (indexed with i) must be supplied with the non-negative cross-section area $a_i \in \mathbb{R}_{\geq 0}$ and the area moment of inertia $I_i \in \mathbb{R}_{\geq 0}$. These are to be found in the optimization process. In this contribution, we assume, for convenience, that the moment of inertia is a second- or third-order polynomial function of the cross-sections,

$$I_i(a_i) = c_{II}a_i^2 + c_{III}a_i^3, \quad (12)$$

with $c_{II}, c_{III} \in \mathbb{R}_{\geq 0}$ being fixed constants. When $I_i(a_i) = 0$ and $a_i = 0$, the finite element vanishes and does not contribute to the load transfer.

Different topology optimization formulations exist, accommodating specific needs of particular applications. Here, we consider the problem of searching the minimum-compliant design under multiple load cases (13a) while satisfying the linear-elastic equilibrium equation (13b) and limiting the material volume from above by $\bar{V} \in \mathbb{R}_{> 0}$ (13c). Physical admissibility of the resulting designs is ensured by the non-negative cross-section areas (13d). Combination of these ingredients establishes the basic elastic-design formulation

$$\min_{\mathbf{a}, \mathbf{u}_1, \dots, \mathbf{u}_{n_{lc}}} \sum_{j=1}^{n_{lc}} \omega_j \mathbf{f}_j(\mathbf{a})^T \mathbf{u}_j \quad (13a)$$

$$\text{s.t.} \quad \mathbf{K}_j(\mathbf{a})\mathbf{u}_j - \mathbf{f}_j(\mathbf{a}) = \mathbf{0}, \quad \forall j \in \{1 \dots n_{lc}\}, \quad (13b)$$

$$\bar{V} - \boldsymbol{\ell}^T \mathbf{a} \geq 0, \quad (13c)$$

$$\mathbf{a} \geq \mathbf{0}, \quad (13d)$$

in which $\boldsymbol{\omega} \in \mathbb{R}_{> 0}^{n_{lc}}$ are positive weights associated with n_{lc} load cases, and $\boldsymbol{\ell} \in \mathbb{R}_{\geq 0}^{n_e}$ stands for the element lengths column vector. Further, $\mathbf{f}_j(\mathbf{a}) \in \mathbb{R}^{n_{\text{dof},j}}$ and $\mathbf{u}_j \in \mathbb{R}^{n_{\text{dof},j}}$ denote the force and displacement column vectors of the j -th load case, $n_{\text{dof},j} \in \mathbb{N}$ stands for the associated number of degrees of freedom, and $\mathbf{K}_j(\mathbf{a}) \in \mathbb{R}^{n_{\text{dof},j} \times n_{\text{dof},j}}$ is the corresponding symmetric positive semidefinite stiffness matrix.

For these stiffness matrices, we require $\forall \mathbf{a} > \mathbf{0} : \mathbf{K}_j(\mathbf{a}) \succ \mathbf{0}$ to exclude rigid body motions. Using the finite element method, $\mathbf{K}_j(\mathbf{a})$ is assembled as

$$\mathbf{K}_j(\mathbf{a}) = \mathbf{K}_{j,0} + \sum_{i=1}^{n_e} \left[\mathbf{K}_{j,i}^{(1)} a_i + \mathbf{K}_{j,i}^{(2)} a_i^2 + \mathbf{K}_{j,i}^{(3)} a_i^3 \right], \quad (14)$$

with $\mathbf{K}_{j,0} \succeq \mathbf{0}$ standing for a design-independent stiffness (such as fixed structural elements), $\mathbf{K}_{j,i}^{(1)} \succeq \mathbf{0}$ being the unit-cross-section-area membrane stiffness of the i -th element in the j -th load case, and $\mathbf{K}_{j,i}^{(2)} \succeq \mathbf{0}$ with $\mathbf{K}_{j,i}^{(3)} \succeq \mathbf{0}$ are the corresponding bending stiffness counterparts associated with the unit cross-section area. The force column vector \mathbf{f}_j is assumed in the form

$$\mathbf{f}_j(\mathbf{a}) = \mathbf{f}_{j,0} + \sum_{i=1}^{n_e} \left[\mathbf{f}_{j,i}^{(1)} a_i \right], \quad (15)$$

where $\mathbf{f}_{j,0}$ stands for the design-independent load and $\mathbf{f}_{j,i}^{(1)}$ are the design-dependent loads such as self-weight. One can also add higher-order terms to (15) to handle non-zero displacement boundary conditions.

The formulation (13) is nonlinear and lacks convexity in general. The non-convexity comes not only from the polynomial entries in the stiffness matrix (14), but also from its possible singularity caused by zero cross-section areas (13d).

3.1 Semidefinite programming formulation for topology optimization of frame structures

An approach to simplify (13) relies on eliminating the displacement variables \mathbf{u}_j from the problem formulation. In the nested approach, which is commonly used in topology optimization, the cross-section areas are bounded from below by a strictly positive $\varepsilon \in \mathbb{R}_{> 0}$, allowing for a computation of $[\mathbf{K}(\mathbf{a})]^{-1}$. Recall that $\mathbf{K}(\mathbf{a}) \succ \mathbf{0}$ for all $\mathbf{a} > \mathbf{0}$. The optimization procedure then traditionally adopts, e.g., the Method of Moving Asymptotes (MMA) (Svanberg, 1987), or the Optimality Criteria (OC) method (Rozvany, 1989). Notice that $\varepsilon \rightarrow 0$ results in a high condition number of $\mathbf{K}(\mathbf{a})$ and that larger values of ε may impair quality of optimized designs, as sizing optimization is solved instead of the original topology optimization. In contrast, here, we eliminate \mathbf{u}_j and allow the cross sections to truly attain zero.

Because $[\mathbf{K}(\mathbf{a})]^{-1}$ may not exist in our case, we rely on the Moore-Penrose pseudo-inverse $[\mathbf{K}(\mathbf{a})]^\dagger$ instead. Its role in enforcing the equilibrium conditions is clarified in the next lemma.

Lemma 1 *Consider the equation $\mathbf{K}_j(\mathbf{a})\mathbf{u}_j = \mathbf{f}_j(\mathbf{a}) - \mathbf{r}_j$ with $\mathbf{u}_j = \mathbf{K}_j(\mathbf{a})^\dagger \mathbf{f}_j(\mathbf{a})$ and a residual vector $\mathbf{r}_j \in \mathbb{R}^{n_{\text{dof}}}$. Then, $\mathbf{r}_j = \mathbf{0}$ if and only if $\mathbf{f}_j(\mathbf{a}) \in \text{Im}(\mathbf{K}_j(\mathbf{a}))$.*

Proof. Let $\mathbf{f}_j(\mathbf{a}) = \mathbf{v}_j(\mathbf{a}) + \mathbf{w}_j(\mathbf{a})$, in which $\mathbf{v}_j(\mathbf{a}) \in \text{Im}(\mathbf{K}_j(\mathbf{a}))$, and $\mathbf{w}_j(\mathbf{a}) \in \text{Ker}(\mathbf{K}_j(\mathbf{a}))$. Then, because $\mathbf{K}_j(\mathbf{a})\mathbf{K}_j(\mathbf{a})^\dagger$ is an orthogonal projector onto the range of $\mathbf{K}_j(\mathbf{a})$, we obtain $\mathbf{K}_j(\mathbf{a})\mathbf{K}_j(\mathbf{a})^\dagger\mathbf{f}_j(\mathbf{a}) = \mathbf{v}_j(\mathbf{a})$. Clearly, when $\mathbf{f}_j(\mathbf{a}) \in \text{Im}(\mathbf{K}_j(\mathbf{a}))$, we have $\mathbf{v}_j(\mathbf{a}) = \mathbf{f}_j(\mathbf{a})$ with $\mathbf{w}_j(\mathbf{a}) = \mathbf{0}$, implying that $\mathbf{r}_j = \mathbf{0}$. For the case of $\mathbf{f}_j(\mathbf{a}) \notin \text{Im}(\mathbf{K}_j(\mathbf{a}))$, $\mathbf{w}_j(\mathbf{a}) \neq \mathbf{0}$, showing that $\mathbf{r}_j = -\mathbf{w}_j(\mathbf{a})$. \square

Lemma 1 allows us to eliminate the displacement variables and write the optimization problem (13) only in terms of the cross-section areas \mathbf{a} as

$$\min_{\mathbf{a}} \sum_{j=1}^{n_{lc}} \omega_j \mathbf{f}_j(\mathbf{a})^\top [\mathbf{K}_j(\mathbf{a})]^\dagger \mathbf{f}_j(\mathbf{a}) \quad (16a)$$

$$\text{s.t.} \quad \bar{V} - \ell^\top \mathbf{a} \geq 0, \quad (16b)$$

$$\mathbf{a} \geq 0, \quad (16c)$$

$$\forall j \in \{1 \dots n_{lc}\}: \quad \mathbf{f}_j(\mathbf{a}) \in \text{Im}(\mathbf{K}_j(\mathbf{a})). \quad (16d)$$

Notice that (16d) essentially eliminates the nonphysical setup when $\mathbf{K}_j(\mathbf{a}) = [\mathbf{K}_j(\mathbf{a})]^\dagger = \mathbf{0}$ produces zero compliance.

Because $[\mathbf{K}_j(\mathbf{a})]^\dagger \succeq 0, \forall j \in \{1 \dots n_{lc}\}$, and $\omega > 0$ by definition, (16a) is bounded from below by 0. Thus, we introduce slack variables $\mathbf{c} \in \mathbb{R}_{\geq 0}^{n_{lc}}$ comprising the to-be-minimized upper bounds on compliances for the load cases and rewrite (16) equivalently as

$$\min_{\mathbf{a}, \mathbf{c}} \omega^\top \mathbf{c} \quad (17a)$$

$$\text{s.t. } c_j - \mathbf{f}_j(\mathbf{a})^\top [\mathbf{K}_j(\mathbf{a})]^\dagger \mathbf{f}_j(\mathbf{a}) \geq 0, \forall j \in \{1 \dots n_{lc}\}, \quad (17b)$$

$$\bar{V} - \ell^\top \mathbf{a} \geq 0, \quad (17c)$$

$$\mathbf{a} \geq 0, \quad (17d)$$

$$\forall j \in \{1 \dots n_{lc}\}: \quad \mathbf{f}_j(\mathbf{a}) \in \text{Im}(\mathbf{K}_j(\mathbf{a})). \quad (17e)$$

To derive a nonlinear semidefinite programming formulation, let us now recall the generalized Schur complement lemma:

Lemma 2 (Gallier, 2011, Theorem 16.1) *Let \mathbf{A} and \mathbf{C} be symmetric square matrices, \mathbf{B} have appropriate dimensions, and \mathbf{I} denote an identity matrix. Then, the following conditions are equivalent:*

1. $\begin{pmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{C} \end{pmatrix} \succeq 0$,
2. $\mathbf{C} \succeq 0, \mathbf{A} - \mathbf{B}^\top \mathbf{C}^\dagger \mathbf{B} \succeq 0, (\mathbf{I} - \mathbf{C}\mathbf{C}^\dagger)\mathbf{B} = \mathbf{0}$.

Since we already have $\mathbf{K}_j(\mathbf{a}) \succeq 0$ by definition and $c_j - [\mathbf{f}_j(\mathbf{a})]^\top [\mathbf{K}_j(\mathbf{a})]^\dagger \mathbf{f}_j(\mathbf{a}) \geq 0$ in (17b), to use Lemma 2 it suffices to show that

$$(\mathbf{I} - \mathbf{K}_j(\mathbf{a}) [\mathbf{K}_j(\mathbf{a})]^\dagger) \mathbf{f}_j(\mathbf{a}) = \mathbf{0}. \quad (18)$$

Proposition 1 *The condition (18) is equivalent to $\mathbf{f}_j(\mathbf{a}) \in \text{Im}(\mathbf{K}_j(\mathbf{a}))$.*

Proof. First, consider $\mathbf{f}_j(\mathbf{a}) \in \text{Im}(\mathbf{K}_j(\mathbf{a}))$. Then, $\mathbf{f}_j(\mathbf{a}) = \mathbf{K}_j(\mathbf{a})\mathbf{u}_j$ for some displacement vector \mathbf{u}_j . After inserting it into the left-hand-side of (18), we have

$$\left(\mathbf{K}_j(\mathbf{a}) - \mathbf{K}_j(\mathbf{a}) [\mathbf{K}_j(\mathbf{a})]^\dagger \mathbf{K}_j(\mathbf{a}) \right) \mathbf{u}_j = \mathbf{0}, \quad (19)$$

which holds for all such \mathbf{u}_j as $\mathbf{K}_j(\mathbf{a}) [\mathbf{K}_j(\mathbf{a})]^\dagger \mathbf{K}_j(\mathbf{a}) = \mathbf{K}_j(\mathbf{a})$ by the definition of the Moore-Penrose pseudo-inverse (Gallier, 2011, Lemma 14.1).

Otherwise, consider $\mathbf{f}_j(\mathbf{a}) \notin \text{Im}(\mathbf{K}_j(\mathbf{a}))$ and let $\tilde{\mathbf{u}} = [\mathbf{K}_j(\mathbf{a})]^\dagger \mathbf{f}_j(\mathbf{a})$. Then, $\mathbf{K}_j(\mathbf{a})\tilde{\mathbf{u}} = \mathbf{f}_j(\mathbf{a}) - \mathbf{r}_j$ for some $\mathbf{r}_j \in \text{Ker}(\mathbf{K}_j(\mathbf{a}))$, $\mathbf{r}_j \neq \mathbf{0}$ by Lemma 1. Thus, the left-hand-side of (18) simplifies to

$$\begin{aligned} \mathbf{f}_j(\mathbf{a}) - \mathbf{K}_j(\mathbf{a}) [\mathbf{K}_j(\mathbf{a})]^\dagger \mathbf{f}_j(\mathbf{a}) &= \\ = \mathbf{f}_j(\mathbf{a}) - \mathbf{K}_j(\mathbf{a})\tilde{\mathbf{u}} &= \mathbf{r}_j \neq \mathbf{0}, \end{aligned} \quad (20)$$

which completes the proof. \square

Finally, Proposition 1 and Lemma 2 facilitate an equivalent reformulation of the optimization problem (17) as a nonlinear semidefinite program

$$\min_{\mathbf{a}, \mathbf{c}} \omega^\top \mathbf{c} \quad (21a)$$

$$\text{s.t.} \quad \begin{pmatrix} c_j & -\mathbf{f}_j(\mathbf{a})^\top \\ -\mathbf{f}_j(\mathbf{a}) & \mathbf{K}_j(\mathbf{a}) \end{pmatrix} \succeq 0, \forall j \in \{1 \dots n_{lc}\}, \quad (21b)$$

$$\bar{V} - \ell^\top \mathbf{a} \geq 0, \quad (21c)$$

$$\mathbf{a} \geq 0, \quad (21d)$$

in which only the constraint (21b) lacks convexity. Importantly, all constraints are polynomial functions of \mathbf{a} , forming therefore a semi-algebraic feasible set.

3.2 Efficient polynomial reformulation

The optimization problem (21) constitutes a minimization of a linear function over a semi-algebraic set, allowing for a solution using the moment-sum-of-squares hierarchy, as briefly discussed in Section 2. However, efficiency of the hierarchy can be improved after modifying (21) to provide a tighter feasible set of relaxed problems and to reduce numerical issues by scaling the design variables. These modifications are outlined in the following paragraphs.

3.2.1 Compactness of the feasible set

We start by enforcing compactness of the feasible set of the optimization problem (21) because of two reasons. First, compactness is required for Theorem 1, in the form of Assumption 1. Second, compactness also allows tightening the feasible sets of relaxed problems, notably improving numerical performance.

Proposition 2 Assume that \mathbf{a}^* and \mathbf{c}^* are optimal cross-section areas and compliances associated with the optimization problem (21). Then, $\forall i \in \{1 \dots n_e\} : 0 \leq a_i^* \leq \bar{a}_i$ with $\bar{a}_i = \bar{V}/\ell_i$ and $\forall j \in \{1 \dots n_{lc}\} : 0 \leq c_j^* \leq \bar{c}/\omega_j$, where $\bar{c} = \sum_{j=1}^{n_{lc}} [\omega_j \mathbf{f}(\hat{\mathbf{a}})^T \mathbf{K}(\hat{\mathbf{a}})^{-1} \mathbf{f}(\hat{\mathbf{a}})]$ with $\hat{\mathbf{a}} = \mathbf{1}\bar{V}/\sum_{i=1}^{n_e} \ell_i$.

Proof. The cross-section areas are non-negative by definition (21d). Therefore, (21c) represents a conic combination and none of the structural elements can occupy a larger volume than the volume bound \bar{V} , $\forall i \in \{1 \dots n_e\} : a_i^* \leq \bar{V}/\ell_i$.

The compliance variables are placed at the main diagonal of the polynomial matrix inequality (PMI) (21b) and are hence non-negative, $c_j^* \geq 0$. Then, because $\boldsymbol{\omega} > \mathbf{0}$, the conic combination $\boldsymbol{\omega}^T \mathbf{c}^*$ is an upper bound for its summands, $\omega_j c_j^* \leq \boldsymbol{\omega}^T \mathbf{c}^*$. Moreover, since $\hat{\mathbf{a}}$ determines uniquely the compliances $\hat{\mathbf{c}}$, $\hat{c}_j = \mathbf{f}_j(\hat{\mathbf{a}})^T \mathbf{K}_j(\hat{\mathbf{a}})^{-1} \mathbf{f}_j(\hat{\mathbf{a}})$, the pair $(\hat{\mathbf{a}}, \hat{\mathbf{c}})$ is a feasible solution to (21b)–(21d), so that we also have $\boldsymbol{\omega}^T \mathbf{c}^* \leq \bar{c} = \boldsymbol{\omega}^T \hat{\mathbf{c}}$. Consequently, $\forall j \in \{1 \dots n_{lc}\} : \omega_j c_j^* \leq \bar{c}$. \square

Among the bounds in Proposition 2, only the compliance upper bounds are not enforced in the formulation (21). Indeed, for any fixed $\mathbf{a} > \mathbf{0}$, $\mathbf{c} \rightarrow \infty$ is feasible to (21), so that Assumption 1 is not satisfied. To make the design space bounded, we add the (redundant) upper-bound compliance constraint from Proposition 2, or, eventually, an upper-bound obtained by solving the convex truss topology optimization problem instead, see Appendix A. Subsequently, we arrive at the optimization problem

$$\min_{\mathbf{a}, \mathbf{c}} \boldsymbol{\omega}^T \mathbf{c} \quad (22a)$$

$$\text{s.t.} \quad \begin{pmatrix} c_j & -\mathbf{f}_j(\mathbf{a})^T \\ -\mathbf{f}_j(\mathbf{a}) & \mathbf{K}_j(\mathbf{a}) \end{pmatrix} \succeq 0, \quad \forall j \in \{1 \dots n_{lc}\}, \quad (22b)$$

$$\bar{V} - \boldsymbol{\ell}^T \mathbf{a} \geq 0, \quad (22c)$$

$$\bar{c} - \boldsymbol{\omega}^T \mathbf{c} \geq 0, \quad (22d)$$

$$\mathbf{a} \geq \mathbf{0}, \quad (22e)$$

for which we have the following result:

Proposition 3 The feasible set of (22) is compact.

Proof. The feasible set is bounded based on Proposition 2. Moreover, \mathbf{a} and \mathbf{c} satisfying conditions (22c)–(22e) form a closed set. Thus, it suffices to show that (22b) is closed. But the elements in (22b) are polynomial functions that are continuous. Moreover, the set of semidefinite matrices is closed so \mathbf{a} and \mathbf{c} satisfying (22b) live in a closed set. Boundedness and closeness imply compactness because we are in a finite dimensional space. \square

3.2.2 Scaling and box constraints

After introducing box constraints in formulation (22), we can scale all variables domains to $[-1, 1]$. This scaling reduces numerical issues that may arise during the problem solution. To this goal, we have

$$c_j = \frac{1}{2\omega_j} (c_{s,j} + 1) \bar{c}, \quad \forall j \in \{1 \dots n_{lc}\}, \quad (23a)$$

$$a_i = 0.5 (a_{s,i} + 1) \bar{a}_i, \quad \forall i \in \{1 \dots n_e\}, \quad (23b)$$

where \mathbf{a}_s and \mathbf{c}_s are the scaled cross-section areas and compliance variables.

In addition, we explicitly insert the box constraints into the optimization problem formulation to tighten feasible sets of the relaxed problems. There are multiple options how to write these box constraints $\mathbf{a}_s, \mathbf{c}_s \in [-1, 1]$, e.g.,

$$-1 \leq a_{s,i} \leq 1, \quad \forall i \in \{1 \dots n_b\}, \quad (24a)$$

$$-1 \leq c_{s,j} \leq 1, \quad \forall j \in \{1 \dots n_{lc}\}, \quad (24b)$$

$$a_{s,i}^2 \leq 1, \quad \forall i \in \{1 \dots n_b\}, \quad (24b)$$

$$c_{s,j}^2 \leq 1, \quad \forall j \in \{1 \dots n_{lc}\}. \quad (24b)$$

$$a_{s,i}^4 \leq 1, \quad \forall i \in \{1 \dots n_b\}, \quad (24c)$$

$$c_{s,j}^4 \leq 1, \quad \forall j \in \{1 \dots n_{lc}\}. \quad (24c)$$

Despite equivalent in what they enforce, their numerical performance in the moment-sum-of-squares hierarchy varies considerably; we refer the reader to the recent note of Anjos *et al.* (2020). Here, we use the quadratic bounds (24b) when the third-order terms $\mathbf{K}_{i,j}^{(3)} a_i^3$ are absent, and quartic bounds (24c) otherwise. Then, the optimization problem reads

$$\min_{\mathbf{a}_s, \mathbf{c}_s} \sum_{j=1}^{n_{lc}} [0.5 (c_{s,j} + 1) \bar{c}] \quad (25a)$$

$$\text{s.t.} \quad \begin{pmatrix} \frac{1}{2\omega_j} (c_{s,j} + 1) \bar{c} - \mathbf{f}_j(\mathbf{a}_s)^T \\ -\mathbf{f}_j(\mathbf{a}_s) & \mathbf{K}_j(\mathbf{a}_s) \end{pmatrix} \succeq 0, \quad (25b)$$

$$\forall j \in \{1 \dots n_{lc}\}, \quad (25c)$$

$$2 - n_e - \mathbf{1}^T \mathbf{a}_s \geq 0, \quad (25c)$$

$$1 - \boldsymbol{\omega}^T \mathbf{c}_s \geq 0, \quad (25d)$$

$$\text{bound constraints (24b) or (24c)}. \quad (25e)$$

Because the feasible set of (25) is compact by Proposition 3, one may tempt to add a redundant polynomial inequality constraint to satisfy Assumption 1. However, the assumption is already satisfied in our case.

Proposition 4 The optimization problem (25) satisfies Assumption 1.

Proof. Let $\mathbf{G}(\mathbf{a}_s, \mathbf{c}_s)$ be a block-diagonal matrix with the blocks (25b)–(25e) and let \mathbf{H} be a sparse matrix of the same dimensions with the structure

$$\mathbf{H} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \quad (26)$$

in which the identity matrix $\mathbf{I} \in \mathbb{S}^{n_e+n_{1c}}$ matches the positions of (25e) in $\mathbf{G}(\mathbf{a}_s, \mathbf{c}_s)$. Clearly, \mathbf{H} is a SOS because of $\mathbf{H} = \mathbf{H}\mathbf{H}^T$, recall Definition 1. Then, if (24b) is used, $p_{(24b)} = \langle \mathbf{H}\mathbf{H}^T, \mathbf{G}(\mathbf{a}_s, \mathbf{c}_s) \rangle = n_e + n_{1c} - \sum_{i=1}^{n_e} a_{s,i}^2 - \sum_{j=1}^{n_{1c}} c_{s,j}^2$, so that the level set $\{\mathbf{a}_s \in \mathbb{R}^{n_e}, \mathbf{c}_s \in \mathbb{R}^{n_{1c}} \mid p_{(24b)} \geq 0\}$ is compact. For (24c), $p_{(24c)} = n_e + n_{1c} - \sum_{i=1}^{n_e} a_{s,i}^4 - \sum_{j=1}^{n_{1c}} c_{s,j}^4$, showing that also the level set $\{\mathbf{a}_s \in \mathbb{R}^{n_e}, \mathbf{c}_s \in \mathbb{R}^{n_{1c}} \mid p_{(24c)} \geq 0\}$ is compact. \square

Remark 1 The higher-order constraints in (25e) are tighter in the moment representation, i.e., (24b) are tighter than (24a) and (24c) are tighter than (24b).

To see this, assume that $(y_0, y_1, y_2, y_3, y_4)$ are the moments associated with the canonical basis of the vector space of polynomials of degree at most four, $(1, x, x^2, x^3, x^4)$, where one can substitute x by any element of \mathbf{a}_s or \mathbf{c}_s . Then, in the first relaxation of the moment-sum-of-squares hierarchy, the fourth-order constraint $1 - x^4 \geq 0$ becomes

$$y_0 - y_4 \geq 0, \quad (27)$$

the quadratic constraint $1 - x^2 \geq 0$ yields

$$y_0 - y_2 \geq 0, \quad (28)$$

and the box constraint $-1 \leq x \leq 1$ provides

$$-y_0 \leq y_1 \leq y_0, \quad (29)$$

with $y_0 = 1$. Moreover, the localizing matrix of the entire optimization problem contains principal submatrices

$$\begin{pmatrix} y_0 & y_1 \\ y_1 & y_2 \end{pmatrix} \succeq 0, \quad (30a)$$

$$\begin{pmatrix} y_0 & y_2 \\ y_2 & y_4 \end{pmatrix} \succeq 0, \quad (30b)$$

$$\begin{pmatrix} y_2 & y_3 \\ y_3 & y_4 \end{pmatrix} \succeq 0, \quad (30c)$$

that must be positive semi-definite as the entire localizing matrix is. Notice that (30b) and (30c) are not present in the degree-one relaxation, which can, therefore, be used only if the bending stiffness is a polynomial of degree at most two. Thus, we omit them in the reasoning behind the relaxation tightness for (24a) and (24b).

In the lowest, degree-two, relaxation the fourth-order constraints (24c) provide $y_4 \leq 1$ from Eq. (27) and $y_2 \geq 0$ with $y_4 \geq 0$ based on (30). Moreover, the determinants of (30) must be non-negative, implying that $y_1^2 \leq y_2$, $y_3^2 \leq y_2y_4$, and $y_2^2 \leq y_4$. A combination of these inequalities then results in $0 \leq y_1^4 \leq y_2^2 \leq y_4 \leq 1$ and $0 \leq y_3^2 \leq y_2y_4 \leq 1$.

For the quadratic constraints (24b), $y_2 \leq 1$ from Eq. (28) and $y_2 \geq 0$ because of Eq. (30). Writing the determinant of (30) then provides us with $y_1^2 \leq y_2$. Consequently, we

observe that $0 \leq y_1^2 \leq y_2 \leq 1$. Notice that this inequality is automatically satisfied in the preceding case.

In the case of pure box constraints (24a), we only have $0 \leq y_1^2 \leq 1$, Eq. (29), and $y_1^2 \leq y_2$, Eq. (30). Note that there is no upper bound for y_2 , which can attain arbitrarily large values in the first relaxation. From the mechanical point of view, this allows for an arbitrarily-large rotational stiffnesses $\mathbf{K}_{j,i}^{(2)} a_i^2$ of the elements.

These observations then allow us to show feasibility of the first-order moments for (25c)–(25e):

Proposition 5 *Let $\mathbf{y}_{c^1}^*$ and $\mathbf{y}_{a^1}^*$ be the first-order moments associated with the variables \mathbf{c}_s and \mathbf{a}_s obtained from a solution to any relaxation of (25) using the moment-sum-of-squares hierarchy. Then, these moments satisfy*

$$2 - n_e - \mathbf{1}^T \mathbf{y}_{a^1}^* \geq 0, \quad (31a)$$

$$1 - \boldsymbol{\omega}^T \mathbf{y}_{c^1}^* \geq 0, \quad (31b)$$

$$1 - \left(y_{a_i^1}^*\right)^2 \geq 0 \quad \text{if (24b) is used,} \quad (31c)$$

$$1 - \left(y_{a_i^1}^*\right)^4 \geq 0 \quad \text{if (24c) is used,} \quad (31d)$$

$$1 - \left(y_{c_j^1}^*\right)^2 \geq 0 \quad \text{if (24b) is used,} \quad (31e)$$

$$1 - \left(y_{c_j^1}^*\right)^4 \geq 0 \quad \text{if (24c) is used.} \quad (31f)$$

Proof. (31a) and (31b) hold trivially from construction of the hierarchy. (31c)–(31f) follow from Remark 1. \square

3.3 Recovering feasible upper-bound solutions

In Proposition 5, we have shown that the first-order moments obtained by solving any relaxation of the moment-sum-of-squares hierarchy satisfy all the constraints of (25) except for (25b). This section is therefore devoted to the question how to “correct” these moments to produce feasible upper-bounds to the original problem (21) and provide a natural sufficient condition of global optimality.

We start by proving the following essential result:

Proposition 6 *Let $\mathbf{y}_{c^1}^*$ and $\mathbf{y}_{a^1}^*$ be the first-order moments associated with the variables \mathbf{c}_s and \mathbf{a}_s obtained from a solution to any relaxation of (25) using the moment-sum-of-squares hierarchy and let $\forall i \in \{1 \dots n_e\} : \tilde{a}_i = 0.5(y_{a_i^1}^* + 1)\bar{a}$ be the corresponding cross-section areas. Then,*

$$\mathbf{f}_{j,0} + \sum_{i=1}^{n_e} \mathbf{f}_{j,i} \tilde{a}_i \in \text{Im} \left(\mathbf{K}_{j,0} + \sum_{i=1}^{n_e} \sum_{k=1}^3 \mathbf{K}_{j,i}^{(k)} \tilde{a}_i^k \right). \quad (32)$$

Proof. In the lowest relaxation of the moment-sum-of-squares hierarchy, the PMI constraint (25b) becomes

$$\begin{pmatrix} \frac{1}{2\omega_j} \left(y_{c_j^1}^* + 1\right) \bar{c} & -\mathbf{f}_j^T \left(\mathbf{y}_{a^1}^*\right)^T \\ -\mathbf{f}_j \left(\mathbf{y}_{a^1}^*\right) & \mathbf{K}_j \left(\mathbf{y}_{a^1}^*, \mathbf{y}_{a^2}^*, \mathbf{y}_{a^3}^*\right) \end{pmatrix} \succeq 0, \quad (33)$$

where $\mathbf{y}_{\mathbf{a}^2}^*$ and $\mathbf{y}_{\mathbf{a}^3}^*$ are the second- and third-order moments associated with \mathbf{a}_s and, with a slight abuse of notation, $\mathbf{K}_j(\mathbf{y}_{\mathbf{a}^1}^*, \mathbf{y}_{\mathbf{a}^2}^*, \mathbf{y}_{\mathbf{a}^3}^*)$ and $\mathbf{f}_j(\mathbf{y}_{\mathbf{a}^1}^*)$ are the stiffness matrix and force column vector constructed from the moments \mathbf{y} . Using Lemma 2 and Proposition 1, we observe that

$$\mathbf{f}(\mathbf{y}_{\mathbf{a}^1}^*) \in \text{Im}(\mathbf{K}_j(\mathbf{y}_{\mathbf{a}^1}^*, \mathbf{y}_{\mathbf{a}^2}^*, \mathbf{y}_{\mathbf{a}^3}^*)). \quad (34)$$

Because we have considered solely degree-one moments in (32) and $\forall \mathbf{a} > \mathbf{0} : \mathbf{K}_j(\mathbf{a}) \succ \mathbf{0}$ was our initial assumption, we must show that the combination of $a_i = 0$ with $I_i > 0$ cannot occur for any i , because that would result in a lower rank of $\mathbf{K}_j(\tilde{\mathbf{a}})$ when compared with $\mathbf{K}_j(\mathbf{y}_{\mathbf{a}^1}^*, \mathbf{y}_{\mathbf{a}^2}^*, \mathbf{y}_{\mathbf{a}^3}^*)$.

To this goal, let $a_i = 0$, which is equivalent to $y_{a_i^1}^* = -1$. Then, the non-negative determinant of (30a) and (30b) with the inequalities (24b) or (24c) imply that $y_{a_i^2}^* = y_{a_i^4}^* = 1$. Moreover, also the determinant of the principal submatrix

$$\begin{pmatrix} 1 & y_{a_i^1}^* & y_{a_i^2}^* \\ y_{a_i^1}^* & y_{a_i^2}^* & y_{a_i^3}^* \\ y_{a_i^2}^* & y_{a_i^3}^* & y_{a_i^4}^* \end{pmatrix} \succeq 0 \quad (35)$$

of the moment matrix must be non-negative. Inserting $y_{a_i^2}^* = y_{a_i^4}^* = 1$ with $y_{a_i^1}^* = -1$ into (35) yields a unique feasible $y_{a_i^3}^* = -1$. Thus, we write the moment of inertia in terms of the scaled cross-section areas (23b). After inserting first-order moments, we obtain

$$I_i = 0.25c_{\text{II}}\bar{a}^2 \left(y_{a_i^2}^* + 2y_{a_i^1}^* + 1 \right) + 0.125c_{\text{III}}\bar{a}^3 \left(y_{a_i^3}^* + 3y_{a_i^2}^* + 3y_{a_i^1}^* + 1 \right) = 0. \quad (36)$$

□

Using Proposition 6, we can correct \mathbf{c} based on $\mathbf{y}_{\mathbf{a}^1}^*$ to provide a feasible solution to (21).

Theorem 2 *Let $\mathbf{y}_{\mathbf{c}^1}^*$ and $\mathbf{y}_{\mathbf{a}^1}^*$ be the first-order moments associated with the variables \mathbf{c}_s and \mathbf{a}_s obtained from a solution to any relaxation of (25) using the moment-sum-of-squares hierarchy. Then,*

$$\tilde{a}_i = 0.5(y_{a_i^1}^* + 1)\bar{a}_i, \quad \forall i \in \{1 \dots n_e\}, \quad (37a)$$

$$\tilde{c}_j = [\mathbf{f}_j(\tilde{\mathbf{a}})]^T \mathbf{K}_j^\dagger(\tilde{\mathbf{a}}) \mathbf{f}_j(\tilde{\mathbf{a}}), \quad \forall j \in \{1 \dots n_{\text{lc}}\} \quad (37b)$$

is feasible (upper-bound) to (21).

Proof. Based on Proposition 5, $\tilde{\mathbf{a}}$ satisfies the constraints imposed on the cross-section areas. By correcting the compliance variables according to (37b), the equilibrium equation (and so the PMI (21b)) is satisfied due to Proposition 6. Consequently, all the constraints of (21) are feasible for the pair $\tilde{\mathbf{a}}, \tilde{\mathbf{c}}$, showing that $\boldsymbol{\omega}^T \mathbf{c}^* \leq \boldsymbol{\omega}^T \tilde{\mathbf{c}} < \infty$. □

We wish to emphasize that in Theorem 2, we have proved feasibility of the upper bounds to (21) and such upper bounds may violate the compliance bound constraints (25d). Thus, knowledge of $\boldsymbol{\omega}^T \mathbf{c}^*$ does not assure convergence of the lowest relaxation to the optimal cross-section areas.

3.4 Certificate of global ε -optimality

Because the hierarchy generates a sequence of lower bounds and we have just shown in Theorem 2 how to compute upper bounds in each relaxation, we naturally arrive at a simple sufficient condition of global ε -optimality.

Theorem 3 *Let $\mathbf{y}_{\mathbf{c}^1}^*$ and $\mathbf{y}_{\mathbf{a}^1}^*$ be the first-order moments associated with the variables \mathbf{c}_s and \mathbf{a}_s obtained from a solution to any relaxation of (25) using the moment-sum-of-squares hierarchy. Then,*

$$\boldsymbol{\omega}^T \tilde{\mathbf{c}} - \sum_{j=1}^{n_{\text{lc}}} \left[0.5(y_{c_j^1}^* + 1) \bar{c}_j \right] \leq \varepsilon \quad (38)$$

is a sufficient condition of global ε -optimality.

Theorem 3 is very simple to verify computationally, significantly simpler than the traditional rank-based certificate of global optimality (8), e.g., (Henrion and Lasserre, 2006). However, (38) fails to be a necessary condition. Indeed, the optimality gap ε may remain strictly positive even when the hierarchy converged according to (8) in the case of multiple globally optimal solutions. Then, the optimal first-order moments \mathbf{y} are not unique; for instance, they may correspond to any convex combination of the global optima, we refer to Section 4.2 for a specific example.

A stronger result holds, however, when the optimization problem possesses a unique global optimum. To show this, we first prove that, with an increasing relaxation degree r , the feasible space of relaxations converges to the convex hull of the initial (non-convex) problem.

Proposition 7 *Let $\mathcal{K}^{(r)}$ be the feasible set of the first-order moments in the r -th relaxation of the moment-sum-of-squares hierarchy of (25). Then, $\mathcal{K}^{(r)} \uparrow \text{conv}(\mathcal{K})$ as $r \rightarrow \infty$, where \mathcal{K} is the intersection of (25b)–(25e).*

Proof. Let $f(\mathbf{a}_s, \mathbf{c}_s)$ be an arbitrary affine function. Based on Proposition 4, Assumption 1 holds for (25) independently of the objective function. Hence, optimization of $f(\mathbf{a}_s, \mathbf{c}_s)$ over \mathcal{K} yields $f(\mathbf{a}_s, \mathbf{c}_s) \uparrow f^*(\mathbf{a}_s, \mathbf{c}_s)$ as $r \rightarrow \infty$ due to Theorem 1. Because $f(\mathbf{a}_s, \mathbf{c}_s)$ is arbitrary, $\mathcal{K}^{(r)} \uparrow \text{conv}(\mathcal{K})$ as $r \rightarrow \infty$. □

Finally, we can prove that the hierarchy eventually attains a zero optimality gap.

Theorem 4 *If there is a unique global solution to (25), then*

$$\boldsymbol{\omega}^T \tilde{\mathbf{c}} - \sum_{j=1}^{n_{lc}} \left[0.5(y_{c_j^1}^{(r)*} + 1)\tilde{c} \right] = 0 \quad (39)$$

as $r \rightarrow \infty$.

Proof. Assuming $r \rightarrow \infty$, optimization of (25a) over $\mathcal{K}^{(r)}$ is equivalent to optimization of (25a) over $\text{conv}(\mathcal{K})$ by Proposition 7. Because \mathcal{K} is compact, its convex hull must be also compact. Hence, it can be equivalently expressed as the convex hull of the limit points of \mathcal{K} that are denoted by $\mathbf{d}_1, \mathbf{d}_2, \dots$, i.e.,

$$\text{conv}(\mathcal{K}) = \text{conv}(\cup_{i=1}^{\infty} \{\mathbf{d}_i\}). \quad (40)$$

Because we assume there is the unique global optimum when optimizing over \mathcal{K} , there must be a unique limit point \mathbf{d}^* associated with this optimum. \square

Remark 2 Although Theorem 4 relies on $r \rightarrow \infty$, a finite (and fairly small) r is required in all our test cases to reach the zero optimality gap. Moreover, this bound equality has occurred when the hierarchy converged based on the rank test (8). It might be possible, therefore, to strengthen Theorem 4 to a finite termination result.

3.5 Global topology optimization of shell structures

Until now, solely frame structures have been considered. However, the optimization formulations (21) and (25) allow for simple modifications to optimize other discrete structures such as shells. Let $\mathbf{t} \in \mathbb{R}_{\geq 0}^{n_e}$ be the vector of shell element thicknesses. Then, the formulation (21) becomes

$$\min_{\mathbf{t}, \mathbf{c}} \boldsymbol{\omega}^T \mathbf{c} \quad (41a)$$

$$\text{s.t.} \quad \begin{pmatrix} c_j & -\mathbf{f}_j(\mathbf{t})^T \\ -\mathbf{f}_j(\mathbf{t}) & \mathbf{K}_j(\mathbf{t}) \end{pmatrix} \succeq 0, \forall j \in \{1 \dots n_{lc}\}, \quad (41b)$$

$$\bar{V} - \mathbf{s}^T \mathbf{t} \geq 0, \quad (41c)$$

$$\mathbf{t} \geq 0, \quad (41d)$$

where $\mathbf{s} \in \mathbb{R}_{>0}^{n_e}$ is a vector of the surface areas of individual shell elements, and $\mathbf{K}_j(\mathbf{t})$ is assembled as

$$\mathbf{K}_j(\mathbf{t}) = \mathbf{K}_{j,0} + \sum_{i=1}^{n_e} \left[\mathbf{K}_{j,i}^{(1)} t_i + \mathbf{K}_{j,i}^{(3)} t_i^3 \right]. \quad (42)$$

Because the design variables can be bounded very similarly to Proposition 2 and scaled, all proven results hold true.

4 Sample problems

This section investigates global topology optimization of selected small-scale structural design problems using the proposed strategy solved numerically by the MOSEK optimizer (MOSEK ApS, 2019). These examples demonstrate strengths and weaknesses of the presented approach: certificate of global ε -optimality using Theorem 3, extraction² of all guaranteed globally optimal solutions based on the flat extension theorem (Curto and Fialkow, 1996), but also higher computational demands when compared to selected local optimization techniques: OC and MMA adopting the nested approach, see, e.g., (Bendsøe and Sigmund, 2004), MATLAB's inbuilt optimizer `fmincon` solving (13) directly, and non-linear semidefinite programming (NSDP) formulation (21) solved by the PENLAB optimizer (Fiala et al., 2013). Except for the nested approaches, all optimization problems were modeled using the YALMIP toolbox (Löfberg, 2004). Our implementation and the corresponding source codes written in MATLAB can be accessed at (Tyburec et al., 2020b).

The first three examples involve two finite elements only to allow visualization of the feasible sets and provide intuition about the solution approach. In the later part, we investigate the influence of finite element types on optimal design and increase the number of elements to evaluate scalability of the approach. All computations were performed on a personal laptop with 16 GB of RAM and Intel® Core™ i5-8350U CPU. Times of individual optimization approaches are measured to allow a simple comparison of the computational demands.

4.1 Structure possessing multiple global optima

As the first problem, we consider a frame structure composed of two Euler-Bernoulli frame elements, see Fig. 2a. Two loads are applied, each of them acting as a separate load case, and weighted equally by $\boldsymbol{\omega} = \mathbf{1}$. Both these frame elements possess the Young modulus $E = 1$, and their overall volume is bounded by $\bar{V} = 0.816597322$ from above³. Accordingly with Fig. 2b, the elements $i = \{1, 2\}$ have rectangular cross-sections with areas $a_i = 0.3h_i$. Then, $I_i = \frac{1}{40}h_i^3$, which implies that $c_{II} = 0$ and $c_{III} = 25/27$ in Eq. (12).

The feasible domain of the optimization problem shown in Fig. 2c reveals that there are three global optima of the objective function value 7.738, corresponding to the following cases: (i) $a_1^* = \bar{V}/\sqrt{2}$ and $a_2^* = 0$, (ii) $a_1^* = 0$ and

² For rank computation we considered the eigenvalues with the absolute value smaller than 10^{-8} to be singular.

³ Fewer digits may prevent the solver from reaching all three global optima. Although an analytical formula for this specific \bar{V} can be derived, we omit it for the sake of brevity.

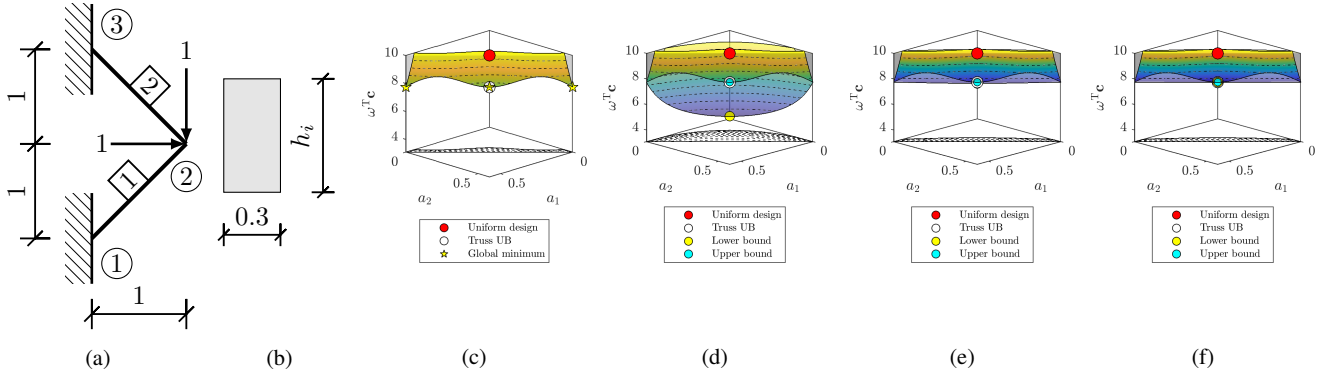


Fig. 2: Frame structure composed of two elements: (a) boundary conditions, (b) the cross-section parametrization, and the sub-level set $\omega^T \mathbf{c} \leq 10$ of the (c) feasible space and of the (d) second, (e) third, and (f) fourth outer approximations with the associated lower- and upper-bounds. Variables a_1 and a_2 stand for the cross-section areas of the two elements and $\omega^T \mathbf{c}$ denotes the corresponding weighted compliance of the two load cases (assuming the moments of inertia $I_i = 25/27a_i^3$, $i \in \{1, 2\}$), and h_i is the cross-section height.

$a_2^* = \bar{V}/\sqrt{2}$, and (iii) $a_1^* = a_2^* = \bar{V}\sqrt{2}/4$. All these solutions are extracted in the fourth relaxation of the moment-sum-of-squares hierarchy (Fig. 2f), which converged based on the rank condition (8) with the rank equal to $s = 3$ and also based on Theorem 3, $\varepsilon = 6 \times 10^{-9}$. Notice that in all the relaxations, the upper bounds recovered by Theorem 2 are global minima, Figs. 2d-2f.

Because all local minima are also global, all tested optimization algorithms converge to the optimal objective function value, see Table 1. Among these algorithms, OC and MMA exhibited the best performance in terms of computational time.

4.2 Irreducible positive optimality gap

Let us now modify the optimization problem described in the preceding section by fixing the volume bound to

method	a_1	a_2	LB	UB	time [s]
OC	0.289	0.289	-	7.738	0.009
MMA	0.289	0.289	-	7.738	0.011
fmincon	0.289	0.289	-	7.738	0.113
NSDP	0.289	0.289	-	7.738	0.409
PO ⁽²⁾ , Th. 3	0.289	0.289	5.065	7.738	0.023
PO ⁽³⁾ , Th. 3	0.289	0.289	7.647	7.738	0.071
PO ⁽⁴⁾ , Th. 3	0.289	0.289	7.738	7.738	0.438

Table 1: Different optimization methods applied to the first optimization problem. LB denotes lower bound, UB abbreviates feasible upper bounds, and PO stands for polynomial optimization. The entries a_i denote cross-section areas of the i -th element, or the areas constructed from the first-order moments in the case of PO.

some $\bar{V} \in (0.816597322, 2.73603242)$. Whilst the boundary points of this open interval match the cases when three global optima occur, the interval interior removes $a_1 = a_2 = \bar{V}\sqrt{2}/4$ from the set of globally optimal solutions. In what follows, we set \bar{V} to the center of the interval.

Solving this modified optimization problem with the moment-sum-of-squares hierarchy produces the sequence of lower- and upper-bounds shown in Fig. 3. Although the hierarchy exhibited a finite convergence based on the rank condition (8) with $s = 2$, the corresponding optimality gap remains strictly positive ($\varepsilon = 0.521$) and cannot be reduced in the subsequent relaxations. Clearly, all outer convex approximations must contain the convex combination of their limit points. Hence, if the limit points denote the global optima, also their convex combinations attain the globally optimal objective function value. Therefore, they are also optimal

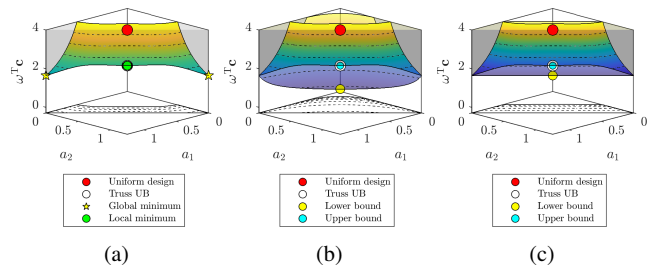


Fig. 3: Frame structure possessing a non-zero optimality gap. The sub-level set $\omega^T \mathbf{c} \leq 4$ of the (a) feasible space and of the (b) second, and (c) third outer approximations with the associated lower- and upper-bounds. Variables a_1 and a_2 stand for the cross-section areas of the two elements and $\omega^T \mathbf{c}$ denotes the corresponding weighted compliance of the two load cases (assuming the moments of inertia $I_i = 25/27a_i^3$, $i \in \{1, 2\}$).

method	a_1	a_2	LB	UB	time [s]
OC	0.628	0.628	-	2.161	0.003
MMA	0.628	0.628	-	2.161	0.004
fmincon	0.628	0.628	-	2.161	0.049
NSDP	0.628	0.628	-	2.161	0.200
PO ⁽²⁾ , Th. 3	0.628	0.628	0.936	2.161	0.022
PO ⁽³⁾ , Th. 3	0.628	0.628	1.640	2.161	0.063
PO ⁽³⁾ , Eq. (8)	1.256	0.000	1.640	1.640	0.063
	0.000	1.256	1.640	1.640	0.063

Table 2: Different optimization methods applied to the second optimization problem. LB denotes lower bound, UB abbreviates feasible upper bounds, and PO stands for polynomial optimization. The entries a_i denote cross-sectional areas of the i -th element, or the areas constructed from the first-order moments in the case of PO.

for the associated relaxation, but may lack feasibility to the original problem. Depending on the optimization algorithm and its settings, one can either reach a lower bound that is actually feasible for the original problem (as was the case in Section 4.1), i.e., a zero optimality gap, or a positive optimality gap that cannot be further reduced, which is the case here.

For this particular problem, all local optimization techniques, using their default starting points and settings, missed the global optima, see Table 2. In fact, they approached the feasible upper-bound that was provided by Theorem 2.

4.3 Frame structure with self-weight

For the previous examples, $f(\mathbf{a}) = \mathbf{f}$ was constant, so that the optimum designs utilized the entire available volume \bar{V} .

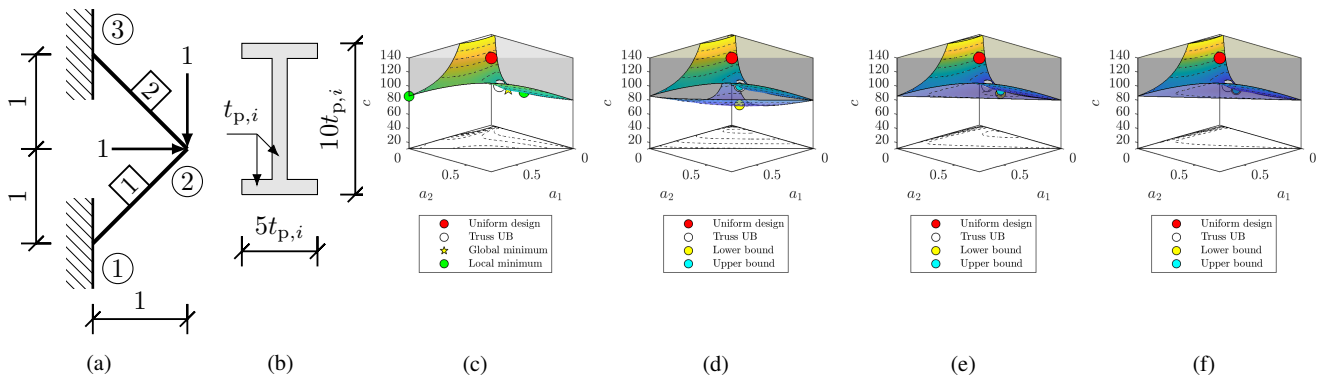


Fig. 4: Frame structure with self-weight: (a) boundary conditions, (b) the cross-section parametrization, and the sub-level set $c \leq 140$ of the (c) feasible space and of the (d) first, (e) second, and (f) third outer approximations with the associated lower- and upper-bounds. Variables a_1 and a_2 stand for the cross-section areas of the two elements and c denotes the corresponding compliance (assuming the moments of inertia $I_i = 41/54a_i^2$, $i \in \{1, 2\}$), and $t_{p,i}$ stands for the flanges and web thickness.

method	a_1	a_2	LB	UB	time [s]
OC	0.022	0.166	-	70.442	1.129
MMA	0.022	0.166	-	70.442	0.935
NSDP	0.707	0.000	-	85.846	1.448
PO ⁽¹⁾ , Th. 3	0.050	0.119	48.246	74.171	0.006
PO ⁽²⁾ , Th. 3	0.034	0.220	68.328	71.594	0.015
PO ⁽³⁾ , Th. 3	0.022	0.166	70.442	70.442	0.058

Table 3: Different optimization methods applied to the optimization problem with self-weight. LB denotes lower bound, UB abbreviates feasible upper bounds, and PO stands for polynomial optimization. The entries a_i denote cross-sectional areas of the i -th element, or the areas constructed from the first-order moments in the case of PO.

In these cases, the volume inequality constraint could have been changed to equality, and, therefore, one design variable eliminated. However, such a procedure cannot be applied when design-dependent loads are present.

To visualize this, let our third illustration be the single-load-case frame structure in Fig. 4a, composed of two frame elements with $E = 1$ with I-shaped cross-sections, Fig. 4b, parametrized by the thickness $t_{p,i}$. The overall volume is bounded from above by $\bar{V} = 1$. The self-weight applies in the vertical direction and is parametrized by the material density $\rho = 10$. For the considered cross-sections, we have $a_i = 18t_{p,i}^2$ and $I_i = 246t_{p,i}^4$. Hence, $c_{II} = 41/54$ and $c_{III} = 0$ in Eq. (12).

The feasible domain of this optimization problem, Fig. 4c, reveals three local optima, and one of them is the global solution. Computation of the optimum by the moment-sum-of-squares hierarchy required three relaxations, Figs. 4d–4f, which converged based on both the rank condition (8) with $s = 1$ and on Theorem 3 with $\varepsilon = -7 \times 10^{-8}$; the slightly negative value of ε is due to

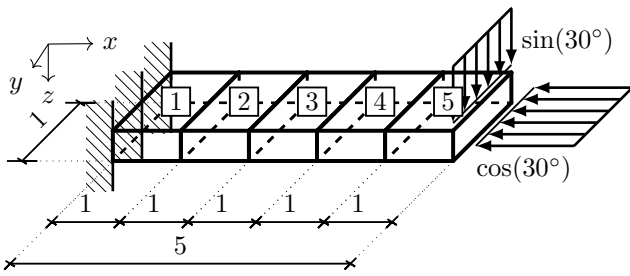


Fig. 5: Boundary conditions of the cantilever beam design problem.

the numerical accuracy of the optimizer. Also notice that the upper-bounds based on Theorem 2 are of very high qualities, see Table 3.

Using local optimization techniques, only OC and MMA were able to arrive at the global optimum, see Table 3. Among other formulations, NSDP approached the worst local optimum and `fmincon` failed even to find a feasible solution.

4.4 Different element types on a cantilever beam

A certain generality of the developed approach is illustrated on a cantilever beam/plate design problem, Fig. 5. The dimensions of the cantilever are 5 in length and 1 in width, and the thicknesses of 5 finite elements are to be found in the optimization. The beam is made of a linear-elastic material

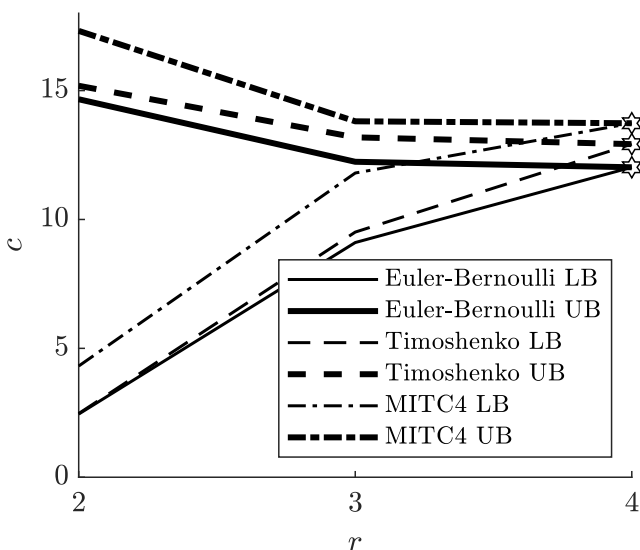


Fig. 6: Convergence of the moment-sum-of-squares hierarchy for the cantilever problem with three finite element types. Variable c denotes compliance, r stands for the relation degree, and LB with UB abbreviate lower- and upper-bound.

	Euler-Bernoulli	Timoshenko	MITC4
a_1^*	2.775	2.724	2.754
a_2^*	2.454	2.414	2.462
a_3^*	2.086	2.060	2.091
a_4^*	1.639	1.643	1.651
a_5^*	1.047	1.159	1.041
c^*	12.025	12.922	13.734
time [s]	65.363	58.633	967.086
Th. 3, ε	-2×10^{-9}	-9×10^{-10}	-2×10^{-9}
Eq. 8, s	1	1	1

Table 4: Globally optimal thicknesses a_1^*, \dots, a_5^* and compliances c^* for the cantilever problem for three element types: Euler-Bernoulli and Timoshenko frame elements, and the MITC4 shell element. Variables ε and s denote the optimality gap in Theorem 3 and the rank of the moment matrices according to (8), respectively.

with the Young modulus $E = 1$ and Poisson ratio $\nu = 0.25$. This structure is subjected to a tip distributed load of magnitude 1 induced under 30° angle with respect to the midline/midsurface. We optimize the frame/shell thicknesses t_i (of rectangular cross-sections) while satisfying $\bar{V} = 10$. The shear correction factor is set to $5/6$ where appropriate.

In what follows, we compare the optimization results of the cantilever problem for three finite element types: Euler-Bernoulli and Timoshenko frame elements, and the quadrilateral Mixed Interpolation Tensorial Component (MITC4) shell element (Bathe and Dvorkin, 1986). For both of the frame element types, we have $c_{II} = 0$ and $c_{III} = I_i(a)/a_i^3 = 1/12$ in Eq (12), whereas $c_{II} = 0$ and $c_{III} = 1$ for the MITC4 element.

The moment-sum-of-squares hierarchy required three steps (degree-four relaxation) to converge in all three cases, Fig. 6, and approached very similar optimal thicknesses, Table 4. As expected, the lowest compliance is provided by the Euler-Bernoulli frame elements, which neglect the shear effects. We account for these effects in the Timoshenko frame elements, increasing so the value of optimal compliance. Another generalization occurs when using the MITC4 shell elements, which not only consider the effects of shear, but further incorporate effects induced by bending about z axis, recall Fig. 5. Therefore, the optimal compliance associated with the MITC4 elements is the highest. These results thus reveal the importance of using an appropriate finite element type for particular problem, as neglecting a physical phenomenon may result in a suboptimal design. Moreover, this influence on optimal minimum-energy designs can be rigorously studied by the proposed approach.

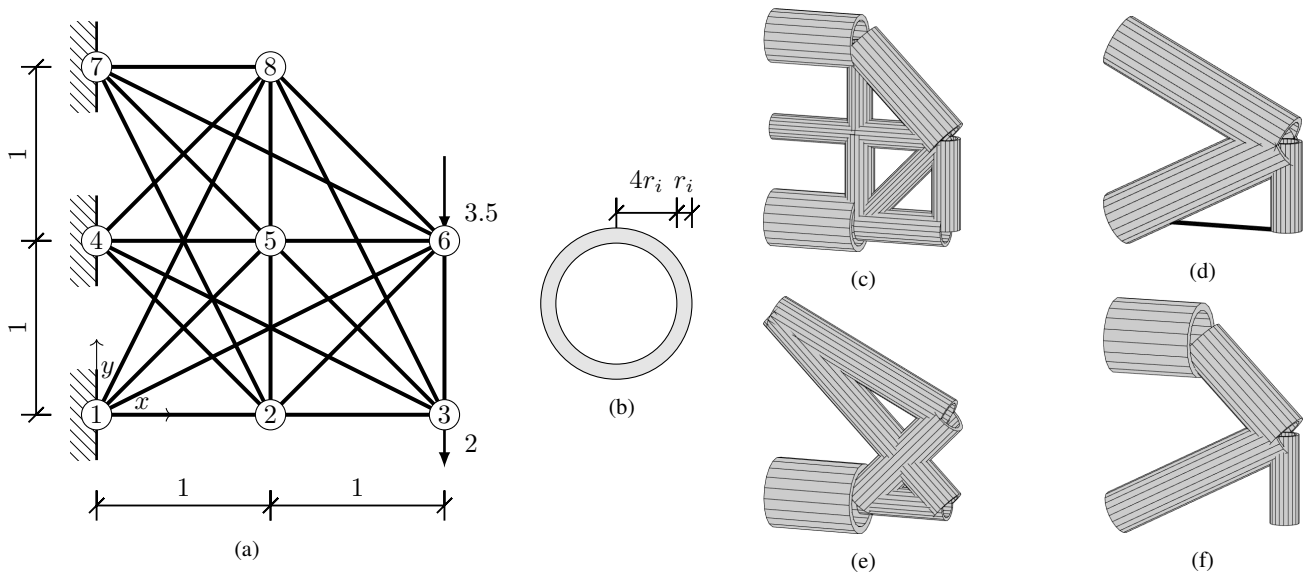


Fig. 7: (a) Ground structure of the 22-elements frame optimization problem, (b) cross-section parameterized by r_i , and optimized designs of compliances: (c) $c = 3276.3$ obtained by $\text{PO}^{(1)}$, (d) $c^* = 1668.6$ resulting from $\text{PO}^{(2)}$ and OC, (e) $c = 1697.7$ reached by MMA and fmincon , and (f) $c = 1741.1$ optimized by NSDP.

4.5 22-elements frame structure

Our final example investigates topology optimization of a 22-element frame structure shown in Fig. 7a. Two loads are applied at nodes 3 and 6 in a single load case. In addition, we set $E = 1$ and $\bar{V} = 0.5$. All the structural elements possess a thin-walled circular cross-section with the radius $5r_i$ and the wall thickness r_i , Fig. 7b. Hence, $a_i = 9\pi r_i^2$ and $I_i = 46.125\pi r_i^4$, so that $c_{\text{II}} = 46.125/(81\pi)$ and $c_{\text{III}} = 0$.

The moment-sum-of-squares hierarchy requires two relaxations to achieve a guaranteed global optimum, both based on Theorem 3 with $\varepsilon = 2 \times 10^{-5}$ and on Eq. (8) with $s = 1$. However, even the second relaxation is fairly computationally expensive (see Table 5), prohibiting solution of higher relaxations of similarly-sized problems on standard hardware.

method	LB	UB	time [s]
OC	-	1668.585	2.454
MMA	-	1697.749	13.816
fmincon	-	1697.665	0.650
NSDP	-	1741.062	4.406
$\text{PO}^{(1)}$, Th. 3, Eq. (8)	1062.105	3276.294	0.103
$\text{PO}^{(2)}$, Th. 3, Eq. (8)	1668.584	1668.584	1492.842

Table 5: Different optimization methods applied to the 22-frame structure design problem. LB denotes lower bounds, UB abbreviates feasible upper bounds, and PO stands for polynomial optimization.

Evaluation of the local optimization algorithms revealed that only OC converged to the global optimum ($c^* = 1668.6$ shown in Fig. 7d). The remaining optimization approaches reached local optima of comparable performance but considerably different topologies: MMA and fmincon converged to the design shown in Fig. 7e with $c = 1697.7$, and NSDP reached the design in Fig. 7f with $c = 1741.1$.

5 Conclusions

Our contribution has addressed the fundamental question in the structural design: how to find globally-optimal minimum-compliant bending-resistant structures in discrete topology optimization with continuous design variables. For the cases of frame and shell structures, multiple loading conditions and design-dependent loads, we have formulated this optimization problem as a (non-linear) semidefinite program constrained by a polynomial matrix inequality. The feasible space of this optimization problem forms a semialgebraic set; hence, powerful results on polynomial optimization—the moment-sum-of-squares hierarchy—facilitate computation of the global solutions.

This hierarchy generates a sequence of tightening outer convex approximations in the space of moments of the probability measures, so that the first-order moments converge monotonically to the convex hull of the original problem. Therefore, a non-decreasing sequence of lower bounds is established. Using the first-order moments only, we have shown that a sequence of feasible upper bounds can be obtained by a simple correction. Consequently, because lower

and upper bounds are available in each relaxation, the upper-bound design quality can be assessed, establishing a sufficient condition of global ε -optimality. This condition is very simple to check and complements the traditional rank-based certificate of global optimality, e.g., (Henrion and Lasserre, 2006).

Our condition fails to be necessary because the first-order moments are not unique when considered optimization problem possesses multiple global optima, potentially leaving a strictly positive optimality gap. For the case of the unique global optimum, we have shown that the hierarchy eventually attains a zero optimality gap as the relaxation number approaches infinity. We note here that the possibility of the global minima multiplicity can almost be avoided in practice when the symmetry of structure and boundary conditions are exploited.

These theoretical results have been illustrated on five problems, which indicate the merits and weaknesses of the presented strategy. First, all of our test problems exhibited a rapid convergence of the hierarchy, allowing for extraction of all global solutions based on the flat extension theorem of Curto and Fialkow (1996). However, the computational complexity is currently fairly high, also when compared to investigated local optimization techniques, leaving the ability to compute proven global optima for small-scale optimization problems only. Yet, even for middle-scale problems, the hierarchy still provides a sequence of upper bounds of reasonable qualities, and, especially, the certificate of their ε -optimality.

In the future, we plan to extend these results in multiple directions. First, we believe that Theorem 4 can be strengthened to certify a zero optimality gap when convergence of the hierarchy in a finite number of steps occurs. Second, the computational demands could be decreased by exploiting the structural sparsity via clique-based chordal decomposition in the spirit of (Kim et al., 2010; Kočvara, 2020). Another research directions may explore eigenvalue constraints and optimization (Achtziger and Kočvara, 2008; Tyburec et al., 2019), the minimum-weight setting, or, eventually, investigate performance of the hierarchy for different topology optimization formulations.

Acknowledgements We thank Edita Dvořáková for providing us with her implementation of the MITC4 shell elements (Dvořáková, 2015).

Marek Tyburec, Jan Zeman, and Martin Kružík acknowledge the support of the Czech Science Foundation project No. 19-26143X.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Replication of results Source codes are available at (Tyburec et al., 2020b).

A Relation to truss topology optimization

The problem formulation (21) has already been known in the context of truss topology optimization (Vandenberghe and Boyd, 1996), for which the constraints (21b) reduce to linear matrix inequalities (LMI). Consequently, the feasible set is convex, allowing for an efficient solution of (21) by interior point methods, for example.

A natural question then arises: What happens when the rotational degrees of freedom are neglected, solving truss topology optimization problem instead of the frame one? To this goal, however, we must first satisfy the rather restrictive assumption that the truss ground structure is capable of carrying the loads $\mathbf{f}_j(\mathbf{a})$, i.e.,

$$\mathbf{f}_j(\mathbf{a}) \in \text{Im}(\mathbf{K}_{t,j}(\mathbf{a})), \forall j \in \{1 \dots n_{1c}\}, \quad (43)$$

where $\mathbf{K}_{t,j}(\mathbf{a}) = \mathbf{K}_{j,0} + \sum_{i=1}^{n_e} \mathbf{K}_{j,i}^{(1)} a_i$. Suppose now that \mathbf{a}_t^* are optimal cross-sections obtained from a solution to (21) with the terms $\mathbf{K}_{j,i}^{(2)}$ and $\mathbf{K}_{j,i}^{(3)}$ neglected, and $\omega^T \mathbf{c}_t^*$ is the associated optimal objective function value, which can be computed from \mathbf{a}_t^* as

$$\omega^T \mathbf{c}_t^* = \sum_{j=1}^{n_{1c}} \left(\omega_j [\mathbf{f}_j(\mathbf{a}_t^*)]^T [\mathbf{K}_{t,j}(\mathbf{a}_t^*)]^\dagger \mathbf{f}_j(\mathbf{a}_t^*) \right). \quad (44)$$

When the optimal cross-sections of a truss structure, \mathbf{a}_t^* , are reused in a frame structure, the resulting objective function value changes to

$$\omega^T \mathbf{c}_f = \sum_{j=1}^{n_{1c}} \left(\omega_j [\mathbf{f}_j(\mathbf{a}_t^*)]^T [\mathbf{K}_{t,j}(\mathbf{a}_t^*) + \mathbf{K}_{b,j}(\mathbf{a}_t^*)]^\dagger \mathbf{f}_j(\mathbf{a}_t^*) \right) \quad (45)$$

with $\mathbf{K}_{b,j}(\mathbf{a}) = \sum_{i=1}^{n_e} \left(c_{II} \mathbf{K}_{j,i}^{(2)} a_{t,i}^2 + c_{III} \mathbf{K}_{j,i}^{(3)} a_{t,i}^3 \right)$.

To state a relation between $\omega^T \mathbf{c}_t^*$ and $\omega^T \mathbf{c}_f$ we recall a useful lemma:

Lemma 3 (Kovanic, 1979) *Let $\mathbf{A} \in \mathbb{S}^n$ and $\mathbf{B} \in \mathbb{R}^{n \times q}$. Then,*

$$\begin{aligned} (\mathbf{A} + \mathbf{B}\mathbf{B}^T)^\dagger &= \mathbf{A}^\dagger \\ &- \mathbf{A}^\dagger \mathbf{B} (\mathbf{I} + \mathbf{B}^T \mathbf{A}^\dagger \mathbf{B})^{-1} \mathbf{B}^T \mathbf{A}^\dagger + \left(\mathbf{B}_\perp^\dagger \right)^T \mathbf{B}_\perp^\dagger \end{aligned} \quad (46)$$

with $\mathbf{B}_\perp = (\mathbf{I} - \mathbf{A}\mathbf{A}^\dagger) \mathbf{B}$.

Using this lemma, we prove that $\omega^T \mathbf{c}_t$ provides an upper bound for $\omega^T \mathbf{c}_f$.

Lemma 4 *Suppose that $\omega^T \mathbf{c}^*$ is the optimal objective function value of the frame structure design problem (21) and (43) holds. Then, $\omega^T \mathbf{c}^* \leq \omega^T \mathbf{c}_f \leq \omega^T \mathbf{c}_t^*$.*

Proof. Because of $\mathbf{f}_j(\mathbf{a}_t^*) \in \text{Im}(\mathbf{K}_{t,j}(\mathbf{a}_t^*))$, we clearly have $\mathbf{f}_j(\mathbf{a}_t^*) \in \text{Im}(\mathbf{K}_{t,j}(\mathbf{a}_t^*) + \mathbf{K}_{b,j}(\mathbf{a}_t^*))$. Therefore, \mathbf{a}_t^* is a feasible solution to the frame structure design problem (21) and the associated objective function is bounded from below by the global optimum $\omega^T \mathbf{c}^*$. Hence, $\omega^T \mathbf{c}^* \leq \omega^T \mathbf{c}_f$.

For the other inequality, we express (45) using Lemma 3. To this goal, let $\mathbf{K}_{b,j}(\mathbf{a}_t^*) = \mathbf{B}_j \mathbf{B}_j^T$, where \mathbf{B}_j is a real matrix because $\mathbf{K}_{b,j}(\mathbf{a}_t^*) \succeq 0$ by definition. Then, (45) can be written as

$$\omega^T \mathbf{c}_f = \omega^T \mathbf{c}_t^* - \omega^T \mathbf{c}_a + \omega^T \mathbf{c}_b, \quad (47)$$

where

$$c_{a,j} = [\mathbf{f}_j(\mathbf{a}_t^*)]^T \mathbf{A}_j^\dagger \mathbf{B}_j \left(\mathbf{I} + \mathbf{B}_j^T \mathbf{A}_j^\dagger \mathbf{B}_j \right)^{-1} \mathbf{B}_j^T \mathbf{A}_j^\dagger \mathbf{f}_j(\mathbf{a}_t^*), \quad (48a)$$

$$c_{b,j} = [\mathbf{f}_j(\mathbf{a}_t^*)]^T \left(\mathbf{B}_{\perp,j}^\dagger \right)^T \mathbf{B}_{\perp,j}^\dagger \mathbf{f}_j(\mathbf{a}_t^*), \quad (48b)$$

with $\mathbf{A}_j = \mathbf{K}_{t,j}(\mathbf{a}_t^*)$. Clearly, Eq. (48a) is non-negative. For (48b), $\left(\mathbf{B}_{\perp,j}^\dagger \right)^T \mathbf{B}_{\perp,j}^\dagger \in \text{Ker}(\mathbf{A}_j)$, so that $\omega^T \mathbf{c}_b$ vanishes. Hence, $\omega^T \mathbf{c}_f = \omega^T \mathbf{c}_t^* - \omega^T \mathbf{c}_a \leq \omega^T \mathbf{c}_t^*$. \square

Thus, when (43) holds true, the truss topology optimization produces an upper bound to the optimal objective of the frame structure topology optimization problem.

References

- Achtziger, W. and Kočvara, M. Structural topology optimization with eigenvalues. *SIAM J. Optim.*, 18(4):1129–1164, 2008. doi: 10.1137/060651446.
- Achtziger, W., Bendsøe, M., Ben-Tal, A., and Zowe, J. Equivalent displacement based formulations for maximum strength truss topology design. *Impact Comput. Sci. Eng.*, 4(4):315–345, 1992. doi: 10.1016/0899-8248(92)90005-s.
- An, H. and Huang, H. Topology and sizing optimization for frame structures with a two-level approximation method. *AIAA J.*, 55(3): 1044–1057, 2017. doi: 10.2514/1.j055020.
- Anjos, M. F. and Lasserre, J. B., editors. *Handbook on semidefinite, conic and polynomial optimization*. Springer US, 2012. doi: 10.1007/978-1-4614-0769-0.
- Anjos, M. F., Emine, Y., Lodi, A., and Sun, Z. A note on the Lasserre hierarchy for different formulations of the maximum independent set problem. Technical report, Polytechnique Montreal, 2020. URL http://www.optimization-online.org/DB_HTML/2020/03/7692.html.
- Bathe, K.-J. and Dvorkin, E. N. A formulation of general shell elements—the use of mixed interpolation of tensorial components. *Int. J. Numer. Methods Eng.*, 22(3):697–722, 1986. doi: 10.1002/nme.1620220312.
- Ben-Tal, A. and Nemirovski, A. *Lectures on modern convex optimization*. Society for Industrial and Applied Mathematics, 2001. doi: 10.1137/1.9780898718829.
- Bendsøe, M., Ben-Tal, A., and Haftka, R. New displacement-based methods for optimal truss topology design. In *32nd Structures, Structural Dynamics, and Materials Conference*. American Institute of Aeronautics and Astronautics, 1991. doi: 10.2514/6.1991-1215.
- Bendsøe, M. P. and Sigmund, O. *Topology optimization: theory, methods, and applications*. Springer Berlin Heidelberg, 2004. doi: 10.1007/978-3-662-05086-6.
- Chan, C.-M., Grierson, D. E., and Sherbourne, A. N. Automatic optimal design of tall steel building frameworks. *J. Struct. Eng.*, 121(5): 838–847, 1995. doi: 10.1061/(asce)0733-9445(1995)121:5(838).
- Curto, R. E. and Fialkow, L. A. Solution of the truncated complex moment problem for flat data. *Mem. Am. Math. Soc.*, 119(568), 1996. doi: 10.1090/memo/0568.
- Dorn, W. S., Gomory, R. E., and Greenberg, H. J. Automatic design of optimal structures. *J. Mec.*, 3(1):25–52, 1964.
- Dvořáková, E. Finite elements for analysis of plates and shells. Master's thesis, Czech Technical University in Prague, Prague, Czech Republic, 2015. URL https://mech.fsv.cvut.cz/wiki/images/f/f3/DP_Dvorakova.pdf.
- Fiala, J., Kočvara, M., and Stingl, M. PENLAB: A MATLAB solver for nonlinear semidefinite optimization, 2013. arXiv:1311.5240.
- Fredricson, H. Topology optimization of frame structures—joint penalty and material selection. *Struct. Multidiscip. Optim.*, 30(3): 193–200, 2005. doi: 10.1007/s00158-005-0515-3.
- Fredricson, H., Johansen, T., Klarbring, A., and Petersson, J. Topology optimization of frame structures with flexible joints. *Struct. Multidiscip. Optim.*, 25(3):199–214, 2003. doi: 10.1007/s00158-003-0281-z.
- Gallier, J. *Geometric methods and applications*. Springer New York, 2011. doi: 10.1007/978-1-4419-9961-0.
- Henrion, D. and Lasserre, J.-B. Convergent relaxations of polynomial matrix inequalities and static output feedback. *IEEE Trans. Autom. Control*, 51(2):192–202, 2006. doi: 10.1109/tac.2005.863494.
- Kanno, Y. Mixed-integer second-order cone programming for global optimization of compliance of frame structure with discrete design variables. *Struct. Multidiscip. Optim.*, 54(2):301–316, 2016. doi: 10.1007/s00158-016-1406-5.
- Khan, M. R. Optimality criterion techniques applied to frames having general cross-sectional relationships. *AIAA J.*, 22(5):669–676, 1984. doi: 10.2514/3.8653.
- Kim, S., Kojima, M., Mevissen, M., and Yamashita, M. Exploiting sparsity in linear and nonlinear matrix inequalities via positive semidefinite matrix completion. *Math. Program.*, 129(1):33–68, 2010. doi: 10.1007/s10107-010-0402-6.
- Kočvara, M. Decomposition of arrow type positive semidefinite matrices with application to topology optimization. *Math. Program.*, 2020. doi: 10.1007/s10107-020-01526-w.
- Kovanic, P. On the pseudoinverse of a sum of symmetric matrices with applications to estimation. *Kybernetika*, 15(5):341–348, 1979.
- Lasserre, J. B. *An introduction to polynomial and semi-algebraic optimization*. Cambridge University Press, 2015. doi: 10.1017/cbo9781107447226.
- Lobo, M. S., Vandenberghe, L., Boyd, S., and Lebret, H. Applications of second-order cone programming. *Linear Alg. Appl.*, 284(1-3): 193–228, 1998. doi: 10.1016/s0024-3795(98)10032-0.
- Löffberg, J. YALMIP: a toolbox for modeling and optimization in MATLAB. In *In Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.
- Michell, A. The limits of economy of material in frame-structures. *Lond. Edinb. Phil. Mag.*, 8(47):589–597, 1904. doi: 10.1080/14786440409463229.
- MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 9.0.*, 2019. URL <http://docs.mosek.com/9.0/toolbox/index.html>.
- Rozvany, G. I. N. *Structural design via optimality criteria*. Springer Netherlands, 1989. doi: 10.1007/978-94-009-1161-1.
- Saka, M. Optimum design of rigidly jointed frames. *Comput. Struct.*, 11(5):411–419, 1980. doi: 10.1016/0045-7949(80)90107-8.
- Svanberg, K. The method of moving asymptotes—a new method for structural optimization. *Int. J. Numer. Methods Eng.*, 24(2):359–373, 1987. doi: 10.1002/nme.1620240207.
- Tyburec, M., Zeman, J., Novák, J., Lepš, M., Plachý, T., and Poul, R. Designing modular 3D printed reinforcement of wound composite hollow beams with semidefinite programming. *Mater. Des.*, 183: 108131, 2019. doi: 10.1016/j.matdes.2019.108131.
- Tyburec, M., Zeman, J., Kružík, M., and Henrion, D. On optimum design of frame structures. *Acta Polytech. CTU Proc.*, 26:117–125, 2020a. doi: 10.14311/app.2020.26.0117.
- Tyburec, M., Zeman, J., Kružík, M., and Henrion, D. Source codes for preprint Global optimality in minimum compliance topology optimization of frames and shells by moment-sum-of-squares hierarchy, 2020b. doi: 10.5281/zenodo.4048828.
- Tyburec, M., Zeman, J., Doškář, M., Kružík, M., and Lepš, M. Modular-topology optimization with Wang tilings: an application to truss structures. *Struct. Multidiscip. Optim.*, forthcoming. doi: 10.1007/s00158-020-02744-8.
- Vandenberghe, L. and Boyd, S. Semidefinite programming. *SIAM Rev.*, 38(1):49–95, 1996. doi: 10.1137/1038003.
- Wang, Q. and Arora, J. S. Alternative formulations for structural optimization: an evaluation using frames. *J. Struct. Eng.*, 132(12):1880–1889, 2006. doi: 10.1061/(asce)0733-9445(2006)132:12(1880).
- Yamada, S. and Kanno, Y. Relaxation approach to topology optimization of frame structure under frequency constraint. *Struct. Multidiscip. Optim.*, 53(4):731–744, 2015. doi: 10.1007/s00158-015-1353-6.