Sparse basis pursuit for compliance minimization in the vanishing volume ratio limit

Evgrafov, Anton ; Sigmund, Ole

Published in:
Zeitschrift fuer Angewandte Mathematik und Mechanik

Link to article, DOI:
10.1002/zamm.202000008

Publication date:
2020

Document Version
Peer reviewed version

Received 08 January 2020; | Revised XXX; | Accepted XXX

DOI: xxx/xxxx

ORIGINAL PAPER

Sparse basis pursuit for compliance minimization in the vanishing volume ratio limit

Anton Evgrafov* | Ole Sigmund

[1]Department of Mechanical Engineering, Technical University of Denmark, DK–2800 Kgs. Lyngby, Denmark

Correspondence
*Anton Evgrafov, Department of Mechanical Engineering, Technical University of Denmark, DK–2800 Kgs. Lyngby, Denmark.
Email: aaev@mek.dtu.dk

Funding Information
This research was supported by the Villum Fonden, Villum Investigator Project InnoTop.

ABSTRACT

We consider a non-smooth convex variational problem appearing as a formal limit of compliance minimization in the vanishing volume ratio limit. The problem has a classical basis pursuit form, and several successful algorithms have been utilized to solve problems of this class in other application contexts. We discuss the well-posedness and regularity of solutions to these problems, possible solution algorithms, and their discretizations as relevant in this mechanical engineering context. We then test the algorithms on a few benchmark problems with available analytical solutions. We find that whereas many algorithms are successful in estimating the optimal objective value to the problem to a high accuracy, the same cannot be said about finding the optimal solutions themselves. In particular, in some examples the algorithms struggle to properly identify the areas where the solutions should vanish entirely. We also discuss an example where the found optimal solutions are not sparse even though sparse(r) solutions exist.

KEYWORDS:
compliance minimization, convex optimization, non-smooth optimization, splitting algorithms, sparse solutions

1 | INTRODUCTION

Non-smooth variational problems based on minimizing sparsity-encouraging [1] norms, or sparse basis pursuit [2], is a well established technique for robustly recovering sparse solutions in many applied scientific disciplines. We refer the interested reader to "LASSO" (least absolute shrinkage and selection operator) methods in statistics [3,4], robust denoising of images [5], or sparse filtering in compressed sensing [6]. The classical compliance minimization problem in the vanishing volume ratio limit and single-load case can also be, at least formally, stated as a minimization of an L^1-type norm of solutions to an underdetermined force equilibrium system [7,8,9,10,11], naturally lending itself to an applicaton of the rich theory and methodology developed specifically for this class of variational problems.

In this work we pursue precisely this line of research. Our goal is to construct an efficient numerical method, which produces approximate, but reasonably accurate solutions to the compliance minimization problem in the zero volume ratio limit on discretizations of modest sizes. The motivation for this particular set of requirements comes from the fact that we envision the utilization of this approach as a first step in an algorithm following the ideas of [12,13], where coarse provisional computations are post-processed to obtain final high-resolution results. Indeed, in the low volume ratio limit one is forced to work with extremely soft ersatz materials occupying largerly void areas of the design domain, thereby having to repeatedly solve PDEs with nearly singular (more than usual even for topology optimization problems) stiffness tensor, something which we can avoid with the

current approach. Last, but not the least, we hope that the convexity and the well studied structure of the basis pursuit problems could be advantageously utilized for algorithmic purposes.

We begin by recalling the precise optimization problem statement in Section 2, where we also illustrate that one cannot expect unique or regular solutions, even for arbitrarily smooth input data. To put the problem in a more computationally accessible Hilbert (in particular, Lebesgue and Sobolev) functional space framework, in Section 3 we propose two versions of Tikhonov regularizations of the problem, akin to displacement and stress (equilibrium) formulations of linear elasticity problems, each admitting a unique optimal solution. We then review the classical efficient solution algorithms applicable to problems of this class in Section 4, making sure to relate them to the underlying mechanical problem, and comment on a somewhat delicate question of selecting the appropriate spaces of piecewise polynomials for the finite element discretization in Section 5. In Section 6 we study the behaviour of these algorithms on a few benchmark examples, for which either analytical or high quality numerical solutions are available. We conclude our work by summarizing our observations in Section 7. Some of the technical material is presented in the appendices.

2 | PROBLEM STATEMENT

Let Ω be a bounded open domain in $\mathbb{R}^n$, $n \in \{2, 3\}$ with Lipschitz boundary $\Gamma = \partial\Omega$ with the outwards facing unit normal $\hat{n}$. We split Γ into two open sets Γ_t and Γ_u ($\Gamma \subset \mathrm{cl}(\Gamma_t \cup \Gamma_u)$, $\Gamma_t \cap \Gamma_u = \emptyset$) representing areas where traction and homogeneous displacement boundary conditions are going to be enforced. Given a volumetric load $f : \Omega \to \mathbb{R}^n$ and a boundary (traction) load $g : \Gamma_t \to \mathbb{R}^n$, the compliance minimization problem consists of finding the infimum of the complementary elastic energy $0.5 \int_\Omega \sigma(x) : \mathcal{S}(x) : \sigma(x)\,\mathrm{d}x$ with respect to the stresses σ and the elastic material properties encoded in the spatially heterogeneous compliance tensor $\mathcal{S}$, subject to the force equilibrium constraint and a restriction on the volume of the utilized material expressed as an integral quantity of $\mathcal{S}^{-1}$. This problem has been a subject of intensive study for many years owing to its practical significance, and homogenization theory provides us with a full characterization of optimal material distributions, see for example[14,15,16,17]. In particular, in a single-load case, the optimal elastic tensor is a limit, in the sense of homogenized material properties, of the ranked laminates with layer directions aligned with the directions of principal stress (eigenvectors of the stress tensor). Even more specifically, in the formal vanishing material volume limit the quantity $\sigma(x) : \mathcal{S}(x) : \sigma(x)$ reduces to $\|\lambda(\sigma(x))\|_1 =: \|\sigma(x)\|_{S^1}$, where $\lambda : \mathbb{S}^n \to \mathbb{R}^n$ denotes the spectrum of a symmetric matrix, that is, to the sum of the absolute values of principal stresses. Incidentally, this quantity also indicates how much composite material should be utilized at a given point, with laminae directions encoded in the eigenvectors of $\sigma(x)$ as indicated previously.

In this fashion, the material distribution variables are naturally eliminated from the optimization problem statement, which can therefore be written in term of stresses only:

$$\underset{\sigma \in \tilde{\Sigma}_{f,g}}{\text{minimize}} \; \|\sigma\|_{S^1}, \tag{1}$$

where $\|\sigma\|_{S^1}$ is an integral quantity accumulating $\|\sigma(x)\|_{S^1}$ over $\mathrm{cl}\,\Omega$, and $\tilde{\Sigma}_{f,g}$ is the set of stresses satisfying the force equilibrium constraint. We define both the objective function and the feasible set of (1) in APPENDIX A: in such a way that the problem is guaranteed to possess globally optimal solutions. To motivate such a somewhat technical definition of these quantities, let us first consider the following example.

Example 1 (Lack of solution regularity for (1))**.** Let $n = 2$, $\Omega =]0, 1[^n$–a convex polygon, $f(x_1, x_2) \equiv 0$, and suppose that the boundary load $g(x_1, x_2) = [\psi(x_1), 0]^{\mathrm{T}}$, where ψ is an arbitrary continuos or even smooth function supported on $]-\infty, 1[$. Mechanically speaking, the horizontal line load g is best carried by a horizontal bar on the boundary. The stresses in such a design, which also solve (1), to are then statically determined and can be easily computed from the fact that their divergence equals to the negative of the applied loads:

$$\sigma(x_1, x_2) = \begin{cases} \begin{bmatrix} \int_{x_1}^1 \psi(\xi)\,\mathrm{d}\xi & 0 \\ 0 & 0 \end{bmatrix}, & (x_1, x_2) \in \Gamma_t, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

see Figure 1 (a). Note that the stresses are distributed only along the boundary (more precisely, they are measures supported on the lines and not ordinary functions belonging to Lebesgue or Sobolev spaces).

The purpose of this example is two-fold. Firstly, regardless of how "nice" the problem's data (Ω, f, g) is, its solutions are not even guaranteed to be ordinary functions: they may have concentrations on lower-dimensional sets, such as for example

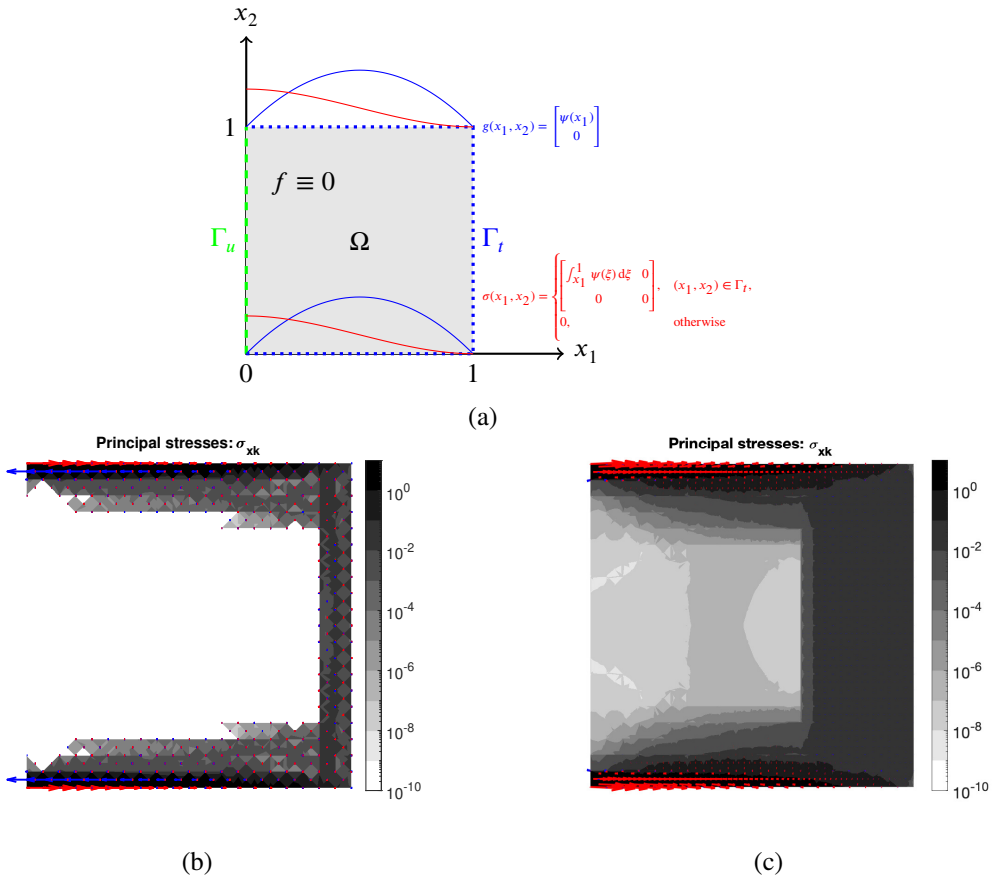


FIGURE 1 (a) Drawing for the example 1. (b) Stresses computed by discretizing (1) using displacement-based FEM. (c) Stresses computed by discretizing (1) using equilibrium-based FEM. In (b), (c), and throughout the manuscript we vizualize $\|\sigma(x)\|_{S^1}$ with background colour intensity, see colourbar for the reference. Red and blue arrows represent respectively the tensile and compressive principal stress directions. The length of the arrow represents the corresponding principal stress magnitude, scaled to fit the figure.

δ-functions, and other oddities. Thus we cannot expect any solution regularity, and it is not very likely that it is possible to accurately approximate such irregular solutions using, for example, the usual piecewise-polynomial based finite element method.

Secondly, we cannot expect stresses solving (1) to satisfy traction boundary conditions in the usual sense; indeed in the present case we have $0 = \sigma \cdot \hat{n} \neq g$.[1] Therefore, the set of equilibrium-satisfying stresses $\tilde{\Sigma}_{f,g}$ participating in the definition of problem (1) has to be defined in a non-standard fashion, see APPENDIX A: for details. Indeed, applying optimization algorithms (to be discussed later in this document), which operate assuming the standard traction boundary conditions prescribed either as the essential or natural boundary conditions, as in equilibrium or displacement-based formulations of elasticity equations, results in the computed stress distributions shown in Figure 1 (b) and (c), where we used $\psi(x_1) = 1 - x_1$. In this case, the optimal $\sigma_{11}(x_1) = x_1^2/2 - x_1 + 1/2$ corresponding to the optimal value of the objective function $2\int_0^1 |\sigma_{11}(x_1)|\, dx_1 = 1/3$. Note that "bars" become finite-width members after the discretization, whose thickness may converge to zero with mesh refinement. Note also the appearence of a vertical member needed to equilibrate the applied traction boundary conditions manifesting themselves in the usual manner as shear loads. Shown stress distributions correspond to the objective value of approximately 0.3342 (displacement-based FEM) and 0.3469 (equilibrium-based FEM).

[1] By testing equation (A2) with $\phi \in C_c^1(\Omega; \mathbb{R}^n)$ it is easy to check that $\text{supp}(\sigma) \cap \Omega = \emptyset$. However, if we insist on satisfying the equation $\sigma \cdot \hat{n} = g$ on the boundary and set either σ_{12} or σ_{22} to be non-zero, then (A2) cannot be satisfied as easily checked using integration by parts on the horizontal lines $\Gamma_t \cap \text{supp}\, g$.

Example 1 shows that it is necessary to seek solutions to (1) in a space significantly larger than the usual function spaces. In fact, it is quite natural to seek stresses satisfying problem (1) in the space $\mathcal{M}(\mathrm{cl}\,\Omega;\mathbb{S}^n)$ of finite signed symmetric matrix-valued measures supported on $\mathrm{cl}\,\Omega$. Whereas we leave the details of the problem statement of (1) for APPENDIX A:, it is worth mentioning that it has been extensively studied in the literature (see for example [7,8,9,10,11]) owing to its intimate relation with a certain classical optimal design problem, namely that of Michell trusses [18]. In particular, [11] establishes that traction-force only version of this problem in dimension $n = 2$ is indeed the Γ-limit of compliance minimization problems in the vanishing volume ratio limit, precisely as conjectured by Kohn and Allaire [9] and as verbally outlined previously in the introduction.

Before we proceed, we would like to bring to the reader's attention the lack of strong convexity of the objective function in (1). This has several implications: for example one cannot in general estimate the distance to optimal solutions using the distance of the objective function value from the optimal value, which in turn has consequences for possibility of estimating the speed of convergence of optimization algorithms. However, the most obvious effect is that in general we cannot expect the problem to have a unique solution. The following example provides a specific instance of (1) with multiple solutions.

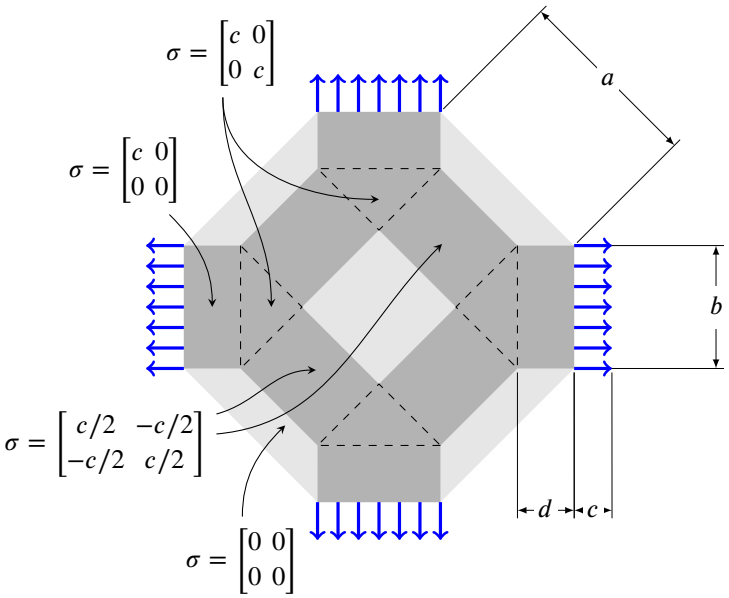


FIGURE 2 Optimization problem setup and a family of optimal solutions discussed in Example 2.

Example 2 (Solution multiplicity)**.** Consider the problem sketched in Figure 2 , where $\Omega \subset \mathbb{R}^n$, $n = 2$ is a convex octagon as shown in the figure determined by two parameters $a \geq 0$ and $b > 0$, $\Gamma_u = \emptyset, \Gamma_t = \partial\Omega$, and $f \equiv 0$. At four parts of the boundary we apply uniformly distributed traction loads with magnitude $c > 0$, with load locations and directions shown in the figure. In this case problem (1) has infinitely many optimal solutions corresponding to the objective value $\|\sigma\|_{S^1} = 2\sqrt{2}abc + 2b^2c$. A single-parametric family of possible optimal solutions, determined by a parameter $d \in [0, a/\sqrt{2}]$, with the corresponding values of piecewise-constant optimal stresses is shown in Figure 2 as well (values of stresses in other subdomains can be easily determined from the symmetry considerations). These are not the only optimal solutions to this problem; indeed owing to the convexity of (1) any convex combination of the presented optimal solutions is yet again a solution to (1).

3 | TIKHONOV REGULARIZATION

Since our goal is to discuss practical algorithms for approximately solving (1), it will be easier to do so if we work in a Hilbert space setting, which does not apply directly to (1). Henceforth we will make an assumption $f \in L^2(\Omega;\mathbb{R}^n)$ and $g \in H^{-1/2}(\Gamma_t;\mathbb{R}^n)$. Let $\Sigma = \{\, \sigma \in L^2(\Omega;\mathbb{S}^n) \,\}$. For non-negative weights $\alpha, \beta \in L^\infty(\Omega)$ we define the regularized objective

function $J_{\alpha,\beta} : \Sigma \to \mathbb{R}$ by

$$J_{\alpha,\beta}(\sigma) := \int_\Omega \alpha(x)\|\sigma(x)\|_{S^1}\,\mathrm{d}x + \frac{1}{2}\int_\Omega \beta(x)\|\sigma(x)\|_F^2\,\mathrm{d}x, \tag{3}$$

where $\|\cdot\|_F = \|\cdot\|_{S^2}$ is the Frobenius norm of the matrix. We will formulate the optimization algorithms for our problem having in mind the setting $\alpha \equiv 1$ and $\beta \approx 0$, and we will strive to find algorithms whose steps remain well-defined even when $\beta \equiv 0$.[2] To complete the definition of the regularization of (1) it remains to define the set of stresses $\sigma \in \Sigma$, which satisfy the force equilibrium condition. Adapting directly definition (A2) in APPENDIX A: to the situation at hand we arrive at the following definition:

$$\begin{aligned}
\Sigma_{f,g}^w &= \{\, \sigma \in \Sigma \mid \int_\Omega \sigma(x) : \nabla u(x)\,\mathrm{d}x = \ell(u), \forall u \in U_0 \,\}, \quad \text{where} \\
\ell(u) &= \int_\Omega f(x)\cdot u(x)\,\mathrm{d}x + \int_{\Gamma_t} g(x)\cdot u(x)\,\mathrm{d}x, \quad \text{and} \\
U_0 &= \{\, u \in H^1(\Omega;\mathbb{R}^n) \mid u = 0 \text{ on } \Gamma_u \,\}.
\end{aligned} \tag{4}$$

Thus, the force equilibrium constraint is enforced only weakly, as in the displacement-based formulations of linearized elasticity. An alternative is to take inspiraiton in "equilibrium"-based formulations of elasticity and use the following definition:

$$\Sigma_{f,g}^s = \{\, \sigma \in \Sigma \cap H(\mathrm{div},\Omega;\mathbb{R}^{n\times n}) \mid -\mathrm{div}\,\sigma = f,\ \text{in } \Omega,\ \sigma\cdot\hat{n} = g,\ \text{on } \Gamma_t \,\}. \tag{5}$$

Regardless of which definition is used, we can state the following result.

Proposition 1. *Assume that $\beta \geq \underline{\beta} > 0$, and that $\Sigma_{f,g}^\diamond \neq \emptyset$, where $\diamond \in \{w, s\}$. Then the optimization problem*

$$\underset{\sigma\in\Sigma_{f,g}^\diamond}{\text{minimize}}\ J_{\alpha,\beta}(\sigma), \tag{6}$$

admits a unique optimal solution.

Proof. In both cases we deal with a minimization of a convex function over a non-empty closed affine manifold. If $\diamond = w$ the considered objective is coercive and strongly convex over Σ. If $\diamond = s$, then the objective is only strongly convex and coercive over $\Sigma_{f,g}^s$ given the fact that $\mathrm{div}\,\sigma$ *is constant over $\Sigma_{f,g}^s$. In either case, the existence follows from the generalized Weierstrass' theorem[19] Section 7.3, and the uniqueness from the strong convexity.* □

Example 3 (Example 2 revisited)**.** Let us recall Example 2, see Figure 2 . Let $\alpha \equiv 1$ and β be an arbitrary positive constant. Note that all members of the infinite single-parametric solution family to (1) sketched in Figure 2 have the same value of the scaled L^2-regularization term, which equals to $\|\sigma\|^2_{L^2(\Omega;\mathbb{S}^n)} = 2\sqrt{2}abc^2 + 2b^2c^2$. However, this is in no way contradicting the uniqueness of solutions to problem (6). Indeed, owing to convexity of the problem (1), any convex combination of the solutions from the sketched single-parametric family is also a solution to the non-regularized problem (1) with a smaller value of $\|\sigma\|^2_{L^2(\Omega;\mathbb{S}^n)}$. Owing to the strong convexity of the regularization term, there is a unique solution attaining the smallest value of $\|\sigma\|^2_{L^2(\Omega;\mathbb{S}^n)}$, and this is the solution we can hope to approximate when $\beta \to 0$ in (6).

4 | POSSIBLE SOLUTION APPROACHES FOR (1)

We start this section by recalling that the problem (1) and its regularization (6) are particular instances of the classical basis pursuit problem, see for example[2]. In the present situation, we try to find the expansion of the applied loads f and g in terms of σ, which is optimal with respect to $\|\cdot\|_{S^1}$ or its regularization $J_{\alpha,\beta}(\cdot)$. As often is the case for such problems, the norm we utilize is equivalent with L^1-norm, and as such is expected to be sparsity-enforcing[1].

There are at least two difficulties associated with problems (1) and (6). One is the non-smoothness of the objective function related to the non-smoothness of the Schatten matrix norm $\|\cdot\|_{S^1}$. The second issue, confined to (6) with $\diamond = s$ is associated with the usual difficulties in constructing piecewise-polynomial subspaces of $H(\mathrm{div},\Omega;\mathbb{S}^n) = \Sigma \cap H(\mathrm{div},\Omega;\mathbb{R}^{n\times n})$ on general meshes, see for example[20,21,22,23,24].

[2]Note that the setting $\alpha \equiv 0$, $\beta \equiv 1$ can be used for testing the convergence of the proposed discretization of the problem with respect to the mesh refinement, as this setting corresponds to the linear elasticity problem.

There are two classes of standard approaches to solving basis pursuit problems, both involving introduction of auxiliary variables. One is to reformulate the problem as a linear conic programming problem, thereby making it (or rather its discretization) amenable to the direct application of interior point (IP) optimization algorithms. The second class of algorithms casts the original problem into solving a sequence of easily solvable smooth and non-smooth convex optimization problems. A comprehensive introduction to such algorithms, including split and linearized Bregman-type iterations and their connection to a variety of alternating direction method of multipliers (ADMM) and similar algorithms is provided in [25]. We will utilize the primal-dual algorithm of Chambolle and Pock [5] as a convenient framework for describing splitting type algorithms. In what follows we provide a brief description of some of the algorithms from each class as applied to the problem at hand.

4.1 | Reformulation as a linear conic problem

To reformulate (6) as a linear conic problem we split a variable $\sigma \in \Sigma$ into positively and negatively semi-definite parts, that is, $\sigma = \sigma_+ - \sigma_-$, where

$$\sigma_+, \sigma_- \in \Sigma_+ := \{\, \tau \in \Sigma \mid \tau(x) \in \mathbb{S}^n_+, \text{for almost all } x \in \Omega \,\},$$

and $\mathbb{S}^n_+ \subset \mathbb{S}^n$ is a set of positively semidefinite symmetric matrices. With this notation the problem (6) can be equivalently stated as follows:

$$\begin{aligned} \underset{(\sigma,\sigma_+,\sigma_-,z)}{\text{minimize}} \quad & \int_\Omega \alpha \operatorname{tr}(\sigma_+ + \sigma_-) + z, \\ \text{subject to} \quad & \sigma \in \Sigma^\diamond_{f,g}, \\ & \sigma_+ - \sigma_- = \sigma, \\ & \|\beta^{1/2}\sigma\|^2_{L^2(\Omega;\mathbb{S}^n)} \le 2z, \\ & \sigma_+, \sigma_- \in \Sigma_+, \\ & z \in \mathbb{R}. \end{aligned} \tag{7}$$

Thus the continuum problem involves only linear equality constraints (force equilibrium, stress splitting), a "global" rotated second order conic constraint for (z, σ) resulting from the quadratic Tikhonov regularization term [26], and a direct product of pointwise semidefinite conic constraints for $\sigma_\pm$. After an appropriate discretization, (7) is directly solvable by standard methods for linear conic problems, most notably primal-dual interior point solvers such as SeDuMi [27].

4.2 | Primal-dual algorithm for (6) with $\diamond = w$

We refer the reader to the algorithmic description in APPENDIX B: and the original paper [5]. To put our problem (6) in the required form, we let $X = \Sigma$, $Y = U_0$, $G = J_{\alpha,\beta}$. Let us define $\hat{u} \in U_0$ and $K : X \to Y$ through the following variational problems, each equivalent to two decoupled scalar elliptic (depending on which inner product on U_0 is used, possibly Laplace) equations:

$$\begin{aligned} (\hat{u}, \tilde{u})_{U_0} &= \ell(\tilde{u}), && \forall \tilde{u} \in U_0, \\ (K\sigma, \tilde{u})_{U_0} &= (\sigma, \nabla\tilde{u})_{L^2(\Omega;\mathbb{S}^n)}, && \forall \tilde{u} \in U_0. \end{aligned} \tag{8}$$

We then set

$$F(u) = I_{\hat{u}}(u) = \begin{cases} 0, & \text{if } u = \hat{u}, \\ +\infty, & \text{otherwise}, \end{cases} \tag{9}$$

to be the indicator function of the set $\{\hat{u}\}$. In this notation the problem (6) with $\diamond = w$ is equivalent to minimizing $F(K\sigma)+G(\sigma)$ over $\sigma \in \Sigma$. After some routine derivations the remaining components of the algorithm (B5) applied to the problem at hand can be identified as follows:

$$K^* u = \nabla_s u,$$
$$\operatorname{prox}_{\rho F^*}(u) = u - \rho\hat{u},$$

where ∇_s is the symmetric gradient (linearized strain) operator, and finally $\sigma_{\text{out}} = \operatorname{prox}_{\rho G}(\sigma_{\text{in}})$ can be defined pointwise:

$$
\begin{aligned}
\sigma_{\text{out}}(x) &= Q(x)\tilde{\Lambda}(x)Q^{\mathrm{T}}(x), \quad \text{where} \\
\frac{1}{1+\rho\beta(x)}\sigma_{\text{in}}(x) &= Q(x)\Lambda(x)Q^{\mathrm{T}}(x), \quad \text{is an eigenvalue decomposition, and} \\
\tilde{\Lambda}_{jj}(x) &= \begin{cases} \Lambda_{jj}(x) - \rho\alpha(x)(1+\rho\beta(x))^{-1}, & \text{if } \Lambda_{jj}(x) \geq \rho\alpha(x)(1+\rho\beta(x))^{-1}, \\ \Lambda_{jj}(x) + \rho\alpha(x)(1+\rho\beta(x))^{-1}, & \text{if } \Lambda_{jj}(x) \leq -\rho\alpha(x)(1+\rho\beta(x))^{-1}, \\ 0, & \text{otherwise,} \end{cases}
\end{aligned}
\tag{10}
$$

is an application of soft thresholding to Λ.

4.3 | Primal-dual algorithm for (6) with $\diamond = s$

Again we refer the reader to the algorithmic description in APPENDIX B: and the original paper[5]. To put our problem (6) in the desired form, we have several options. Perhaps the simplest is to proceed as follows: let $X = \Sigma \cap H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n})$, $Y = \Sigma$, $F = J_{\alpha,\beta}$, $G = I_{\Sigma^s_{f,g}}$, and finally $K = i$, that is the embedding (identity) operator. With these identifications the problem (6) with $\diamond = s$ is equivalent to minimizing $F(K\sigma) + G(\sigma)$ over $\sigma \in \Sigma \cap H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n})$. The remaining components of the algorithm are evaluated as follows. The primal step $\sigma_{\text{out}} = \operatorname{prox}_{\rho G}(\sigma_{\text{in}} - \rho K^*\tau)$ is determined as the solution of the following variational problem: find $(\sigma_{\text{out}}, u) \in \Sigma \cap H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}) \times L^2(\Omega; \mathbb{R}^n)$, such that $\sigma \cdot \hat{n}|_{\Gamma_t} = g$ and

$$
\begin{aligned}
(\sigma_{\text{out}}, \tilde{\sigma})_{H(\operatorname{div},\Omega;\mathbb{R}^{n\times n})} - (\operatorname{div}\tilde{\sigma}, u)_{L^2(\Omega;\mathbb{R}^n)} &= (\sigma_{\text{in}}, \tilde{\sigma})_{H(\operatorname{div},\Omega;\mathbb{R}^{n\times n})} - \rho(\tau, \tilde{\sigma})_{L^2(\Omega;\mathbb{R}^{n\times n})}, \\
-(\operatorname{div}\sigma, \tilde{u})_{L^2(\Omega;\mathbb{R}^n)} &= (f, \tilde{u}),
\end{aligned}
\tag{11}
$$

$\forall(\tilde{\sigma}, \tilde{u}) \in \Sigma \cap H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}) \times L^2(\Omega; \mathbb{R}^n)$, such that $\tilde{\sigma} \cdot \hat{n}|_{\Gamma_t} = 0$. The dual step $\sigma_{\text{out}} = \operatorname{prox}_{\rho F^*}(\sigma_{\text{in}}) = \operatorname{prox}_{\rho J^*_{\alpha,\beta}}(\sigma_{\text{in}})$ can be determined using (10) and Morreau's identity $\operatorname{prox}_{\rho J^*_{\alpha,\beta}}(\sigma_{\text{in}}) = \sigma_{\text{in}} - \rho \operatorname{prox}_{\rho^{-1} J_{\alpha,\beta}}(\rho^{-1}\sigma_{\text{in}})$.

Remark 1. Note that $\operatorname{div}\sigma$ is constant over $\Sigma^s_{f,g}$, which means that $\|\cdot\|_{L^2(\Omega;\mathbb{S}^n)} = \|\cdot\|_{H(\operatorname{div},\Omega;\mathbb{R}^{n\times n})}$ on $\Sigma^s_{0,0}$. Let us split $\sigma = \hat{\sigma} + \Delta\sigma$, where $\hat{\sigma} \in \Sigma^s_{f,g}$ is arbitrary but fixed for the duration of the algorithm. One can consider the following variation of the algorithm: $X = (\Sigma^s_{0,0}, \|\cdot\|_{L^2(\Omega;\mathbb{S}^n)})$, $Y = L^2(\Omega; \mathbb{S}^n)$, $K = i$, $F(\Delta\sigma) = J_{\alpha,\beta}(\hat{\sigma} + \Delta\sigma)$, $G(\Delta\sigma) = 0$. After working through the routine derivations we arrive at a fixed point iteration, which in terms of the original variable σ is obtained by the algorithm we have stated above by replacing both $H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n})$-inner products in (11) with $L^2(\Omega; \mathbb{S}^n)$-inner products. With this modification, the left hand side of (11) is precisely the equations of elasticity written in the mixed form corresponding to S being the identity tensor.

With the discretizations we have employed we do not observe significant difference in algorithmic performance between the two variations of the algorithms.

Remark 2. As we have already mentioned, constructing piecewise-polynomial subspaces of $\Sigma \cap H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n})$ is generally speaking a non-trivial task. To deal with this issue we can consider the following modification of the algorithm. We put $X = H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n})$, $Y = L^2(\Omega; \mathbb{R}^{n\times n})$, $\Sigma^{\times}_{f,g} = \{\, \sigma \in H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n}) \mid -\operatorname{div}\sigma = f, \text{ in } \Omega, \sigma \cdot \hat{n} = g, \text{ on } \Gamma_t \,\}$, $G = I_{\Sigma^{\times}_{f,g}}$, $K = i$, and finally $F = J_{\alpha,\beta} + I_{L^2(\Omega;\mathbb{S}^n)}$.[3] In this situation most derivations remain the same, except (10) is applied to $L^2(\Omega; \mathbb{R}^{n\times n})$-projection of σ_{in} onto $L^2(\Omega; \mathbb{S}^n)$, that is, to $0.5(\sigma_{\text{in}} + \sigma^{\mathrm{T}}_{\text{in}})$. Of course the advantage of this formulation is that (11) only requires $H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n})$-conforming elments. Additionally, this tensorial problem splits into two independent vector-problems (mixed formulation of a Laplace-like equation). Remark 1 applies to the non-symmetric stress formulation as well.

4.4 | Discussion

Needless to say, the algorithms mentioned in this section are not the only way of solving problem (6). In particular, it is worth mentioning that one could apply subgradient or spectral bundle algorithms[28,29] directly to the non-smooth problem, or come up with different splitting strategies.

[3] Note that strictly speaking $J_{\alpha,\beta}$ is undefined for non-symmetric matrices, but this is of little consequence as F is infinite for such arguments because of the second term.

As most first order splitting methods result in a fixed point iteration, standard fixed point acceleration tools may be utilized, such as for example Newton's method [30] or Andersson acceleration [31,32], just to name a few. A different approach to accelerating splitting methods is to equip them with a linesearch mechanism, see [33], thereby potentially allowing the methods to take longer steps when beneficial.

5 | DISCRETIZATION

One can see that solution approaches outlined in the previous section involve a range of variational (sub-)problems posed over a variety of vector spaces, affine subspaces thereof, or cones therein. Some of these sets are significantly easier to construct conforming piecewise-polynomial approximations to than others.

Owing to the expected low regularity of the solutions, see Example 1, we generally focus on the low order methods.

5.1 | Splitting methods

Essentially all the splitting methods we consider involve two steps: solving an elasticity (or Laplace)-like PDE in the mixed or primal variable form, followed by pointwise thresholding operations.

Splitting methods involving $H(\mathrm{div}, \Omega; \mathbb{S}^n)$-conforming stresses, as discussed in Subsection 4.3, have the strongest requirements to the approximating piecewise-polynomial spaces, to comply with which many approaches employ higher order polynomials to achieve stability [24,34]. In this paper we focus on utilizing the piecewise-linear compound triangle element [20].

For variations of the methods involving stresses conforming only to $H(\mathrm{div}, \Omega; \mathbb{R}^{n\times n})$, as discussed in Remark 2, we employ the lowest order BDM elements [35].

Finally, for the $H^1(\Omega; \mathbb{R}^n)$-conforming method discussed in Subsection 4.2 we utilize second order Lagrange elements.

For the pointwise-evaluated prox-operator only conformity to $L^2(\Omega; \mathbb{S}^n)$ is required. It is most convenient to utilize nodal discontinuous Galerkin bases based on Gauss quadrature points within the element for discretizing this part of the algorithm.

5.2 | Linear conic optimization

The interesting additional challenge in the approach outlined in Section 4.1, when compared with others, is related to choosing a suitable discrete space for the positive and negative definite parts of the stresses $\sigma_\pm \in \Sigma_+$.

Indeed, without striving for full generality, in the procedures we have in mind we represent the trial functions $\sigma_{\pm,h}(x) = \sum_{j=1}^{M} \sigma_{\pm,j}\phi_j(x)$, where $\sigma_{\pm,j} \in \mathbb{S}^n$ are (tensorial) global degrees of freedom, and $\phi_j : \Omega \to \mathbb{R}$ form a global basis or frame in the trial space. We then let the optimization algorithm, which is applied to the discretized problem, work with variables $\sigma_{\pm,i} \in \mathbb{S}^n$, $i = 1, \dots, M$. A typical interior point algorithm would guarantee, that $\sigma_{\pm,i} \in \mathbb{S}^n_+$, $i = 1, \dots, M$. It is then natural to seek methods, which under such assumptions guarantee that $\sigma_{\pm,h} \in \Sigma_+$.

Such a requirement is satisfied, for example, by tensor products of the following elements: DP0/DQ0 (piecewise constants on triangles/quads); P1 (piecewise linear shape functions on triangles, where the tensor value inside the triangle is represented as a convex combination of nodal degrees of freedom); affinely transformed Q1 elements (value at each interior point can be written as a convex combination of values at points on opposing edges, and those in turn can be written as convex combinations of nodal degrees of freedom).

For approximating symmetric $H(\mathrm{div}, \Omega; \mathbb{S}^n)$-conforming stresses we can choose conforming elements mentioned in subsection 5.1. We have tested our implementation with a piecewise-linear compound triangle element [20] and a piecewise-linear non-conforming quadrilateral on uniform grids [36].

6 | NUMERICAL EXAMPLES

In addition to the problem described in (2), we consider a few more benchmark examples. In all of them we use a rectangular design domain $\Omega =]0, L[\times]0, H[$, where L, H are positive parameters. We also put $\Gamma_t = \{ (x_1, x_2) \in \partial\Omega \mid x_1 > 0 \}$, $f \equiv 0$. The presented analytical optimal solutions assume $\alpha = 1$, and $\beta = 0$. Unstructured (and some structured) simplicial meshes are prepared with Gmsh [37] in such a way that the applied traction boundary conditions can be exactly represented by the underlying piecewise polynomial subspaces. The discretized optimization algorithms are implemented in Matlab [38].

In all experiments we keep $\alpha = 1$ and $\beta = 0$. (We have also experimented with letting $\beta = 10^{-10}$, which allows us to utilize the prox-parameter selection strategy for the strongly convex case, as described in [5]. In our experience, the difference between these two settings appears to be relatively insignificant.)

We use the standard L^2, $H(\mathrm{div})$ and H^1 norms in our implementations, which leads to $\|K\| \leq 1$ in all discussed versions of Chambolle and Pock's algorithm. We initialize $\rho_x \approx \rho_y \approx 1$, or more precisely we put one of the prox-parameters to 0.9 to satisfy the strict inequality $\rho_x \rho_y \|K\|^2 < 1$.

We stop the splitting algorithms when $\|x_{k+1} - x_k\|_X < 10^{-6}$ and $\|y_{k+1} - y_k\|_Y < 10^{-6}$, while SeDuMi is allowed to run until it runs into numerical problems.

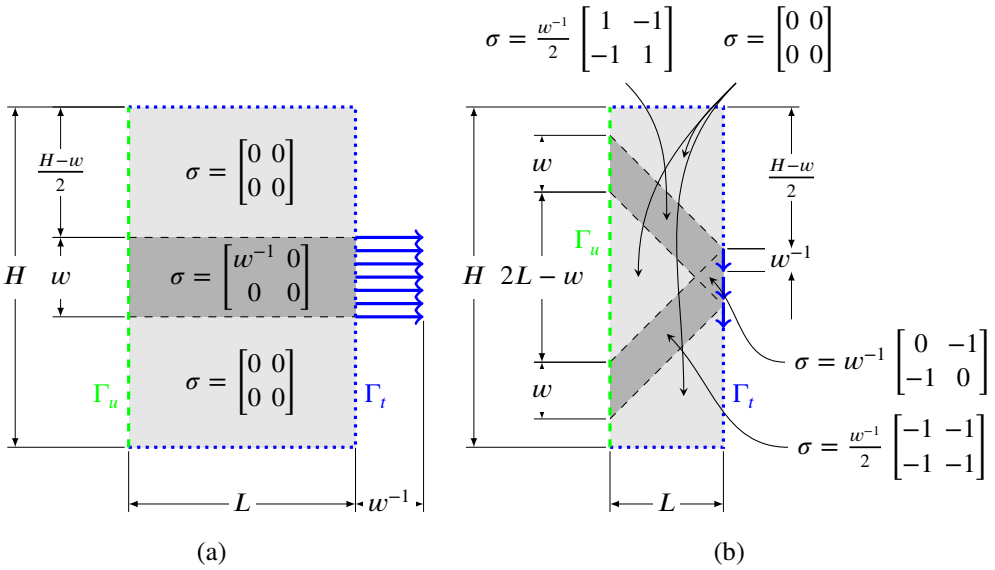


FIGURE 3 (a) 1-bar example; (b) 2-bar/cantilever example. In both cases, position and direction of the non-zero traction forces is shown with blue arrows. As one can intuitively expect, the optimal stresses are only non-zero within areas outlined by black dashed lines.

6.1 | 1-bar example

This problem together with its unique optimal solution is sketched in Figure 3 (a). We consider the case $L = H = 1$ and $w = 0.2$.

6.2 | 2-bar example

This problem together with its unique optimal solution is sketched in Figure 3 (b), and corresponds to the choice $2L + w < H$. In our case we put $L = 0.5$, $H = 1.5$, $w = 0.2$.

6.3 | Short cantilever

The problem setup is similar to the previous case, see Figure 3 (b), but now we set $L = 2.0$, $H = 1.0$, $w = 0.2$. In this case, the analytical solution is only available for a point (measure) load case, see for example [39]. However, the problem has been studied numerically in [40].

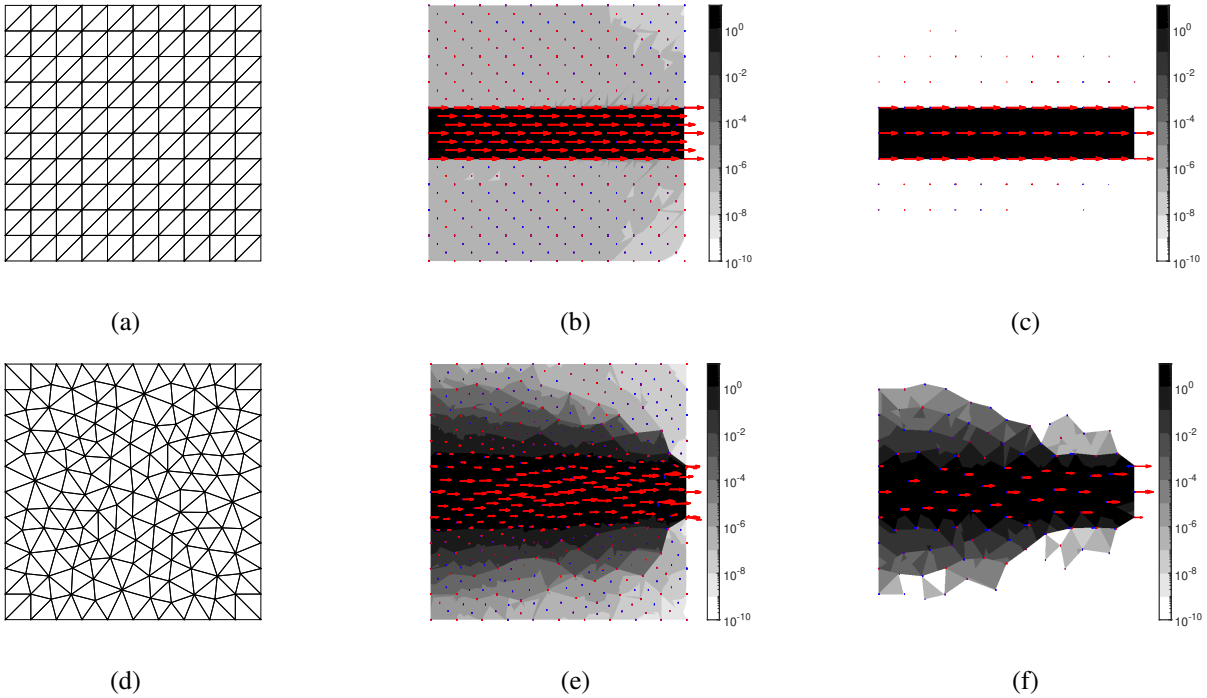


FIGURE 4 Influence of the mesh on the solution: 1-bar example. From top to bottom: analytical solution-conforming mesh; unstructured mesh. From left to right: mesh; solution found using $\Sigma^s_{f,g}$-based splitting algorithm; solution found using $\Sigma^w_{f,g}$-based splitting algorithm. Also see Table 1 .

Id	# iterations	error, objective	L^2-error, stresses
(b)	5528	$3.3 \cdot 10^{-7}$	$4.2 \cdot 10^{-6}$
(c)	8432	$2.0 \cdot 10^{-9}$	$2.9 \cdot 10^{-3}$
(e)	95467	$1.4 \cdot 10^{-2}$	$3.4 \cdot 10^{-1}$
(f)	38426	$9.1 \cdot 10^{-9}$	$3.1 \cdot 10^{-1}$

TABLE 1 Influence of the mesh on the solution: 1-bar example. Id column refers to the labels in Figure 4 . All errors are relative.

6.4 | Influence of the mesh alignment

Generally speaking, mesh turns out to play a huge role in the quality/accuracy of the computed solutions. We illustrate our point by solving two benchmark problems, 1-bar and 2-bar examples, on coarse meshes, which are either aligned or misaligned relative to the known analytical piecewise-constant stresses. The results are summarized in Figures 4 and 5 , and accompanying Tables 1 and 2 . Within statistics and image processing community it is a well known fact that objective functions of the type we consider may lead to "staircaising effect", that is, where solution follows the mesh for small values of the stresses. One cure for this effect has been proposed within the context of statistics by Huber[41]. In essense, it amounts to smoothing the L^1-type norm by replacing it with a quadratic L^2-type term in the vicinity of zero. Such a replacement, however, implies that zero-stress "void" areas, whose identification is nearly the most important part of the solution, would not be identified correctly any longer. We believe that within the topology optimization context a different approach to this issue, such as for example directional mesh adaptivity, would be more appropriate.

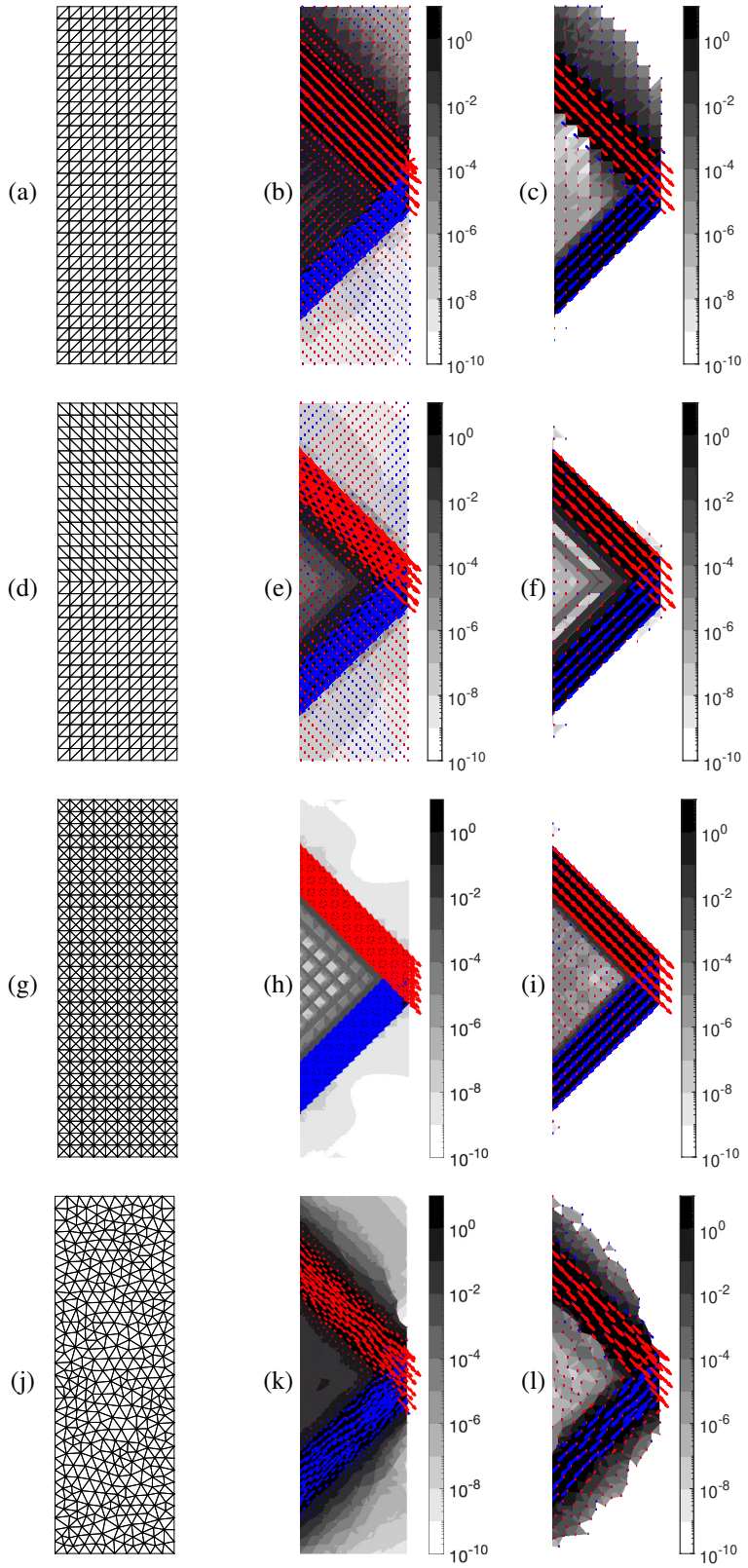


FIGURE 5 Influence of the mesh on the solution: 2-bar example. From top to bottom: three structured and one unstructured mesh; only the mesh in row three is fully solution-conforming. From left to right: mesh; solution found using $\Sigma^{s}_{f,g}$-based splitting algorithm; solution found using $\Sigma^{w}_{f,g}$-based splitting algorithm. Also see Table 2 .

Id	# iterations	error, objective	L^2-error, stresses
(b)	40574	$9.5 \cdot 10^{-3}$	$3.9 \cdot 10^{-1}$
(c)	54587	$2.4 \cdot 10^{-7}$	$2.0 \cdot 10^{-1}$
(e)	16482	$5.0 \cdot 10^{-4}$	$2.0 \cdot 10^{-1}$
(f)	20463	$1.5 \cdot 10^{-9}$	$1.4 \cdot 10^{-1}$
(h)	21866	$1.1 \cdot 10^{-8}$	$3.9 \cdot 10^{-3}$
(i)	41100	$5.7 \cdot 10^{-9}$	$1.8 \cdot 10^{-2}$
(k)	72666	$3.5 \cdot 10^{-3}$	$3.2 \cdot 10^{-1}$
(l)	35386	$1.7 \cdot 10^{-7}$	$2.6 \cdot 10^{-1}$

TABLE 2 Influence of the mesh on the solution: 2-bar example. Id column refers to the labels in Figure 5 . All errors are relative.

6.5 | Influence of the choice $\diamond \in \{w, s\}$

This dichotomy is relatively simple. In the stress space $\Sigma^w_{f,g}$ the equilibrium constraint is enforced only weakly, similarly to the original problem (1). When compared to $\Sigma^s_{f,g}$, this stress space is larger, and is more amenable to construction of conforming piecewise-polynomial subspaces, as we mentioned previously. Empirically we observe that the choice $\diamond = w$ often results in stress fields with more contrast, which is advantageous for the considered application, and also in more accurate objective values. At the same time we should also mention that in several cases we obtain significantly more accurate approximation of stresses with respect to $L^2(\Omega)$-error with the choice $\diamond = s$, which is consistent with the general wisdom that equilibrium-based FEM discretizations produce more accurate approximations of stresses than displacement-based FEM discretizations, all other things being equal.

6.6 | Influence of the numerical quadrature

Owing to the point-wise definition of the non-linear soft thresholding operator, we evaluate it at Gauss quadrature integration points, see Section 5. Equivalently, the number of quadrature points determines the piecewise-polynomial subspace for one of the components of the solution within Chambolle–Pock's splitting algorithm applied to our problem. Additionally, since the soft thresholding generally speaking converts piecewise polynomials to non-polynomials (with the exception of piecewise-constant polynomials), we cannot in general expect to integrate the result of such a prox-operator exaclty. Thus the number of quadrature points should be determined from the stablity and accuracy considerations on benchmark examples.

We should mention that examples in the previous subsections have been run using a 3-point Gauss quadrature on each mesh triangle (sub-triangle in the case of a composite triangle element).

To explore the issue of the influence of the numeric quadrature on the solution, we reconsider a 2-bar example discretized with a regular grid of 10×30 squares as shown in Figure 5 (a) and (g). The results are summarized in Table 3 . Note that in a particularly attractive choice of 1-point quadrature, corresponding to the piecewise-constant stresses in Σ, converges very fast in the case of the weak enforcement of equilibrium constraint to a wrong solution shown in Figure 6 with more than a 100% relative error in stresses!

Based on these computations we conclude that utilizing a 3-point Gauss quadrature, or piecewise-linear stresses on each (sub-)triangle provides an adequate choice for the selected PDE discretization.

6.7 | Influence of the mesh size

To study the behaviour of the algorithms with respect to successive mesh refinement (h-refinement only) we apply the interior point algorithm as well as the splitting algorithm of Chambolle and Pock to 1-bar and 2-bar examples discretised on a solution-conforming structured grids, see Figures 4 (a) and 5 (g). We can draw some meaningful conclusions even from these limited experiments, where it is important to remember that the analytical solution can be represented exactly on each discretization level, and therefore the error (at least, in $\Sigma^s_{f,g}$) can be reduced to nearly zero, see Table 4 where we solve the same examples using Chambolle and Pock's splitting algorithm stopped when $\|x_{k+1} - x_k\|_X < 10^{-12}$ and $\|y_{k+1} - y_k\|_Y < 10^{-12}$. With this in mind, it is clear that the problem becomes more difficult to solve numerically on the refined meshes. Most notably SeDuMi runs

Id	# q.points	# iterations	error, objective	L^2-error, stresses
	1	33742	$1.8 \cdot 10^{-3}$	$3.0 \cdot 10^{-1}$
(b)	3	40574	$9.5 \cdot 10^{-3}$	$3.9 \cdot 10^{-1}$
	6	51044	$1.5 \cdot 10^{-2}$	$4.0 \cdot 10^{-1}$
	12	50188	$7.5 \cdot 10^{-3}$	$4.0 \cdot 10^{-1}$
	37	54427	$1.0 \cdot 10^{-2}$	$4.0 \cdot 10^{-1}$
Id	# q.points	# iterations	error, objective	L^2-error, stresses
	1	19535	$4.2 \cdot 10^{-1}$	$1.1 \cdot 10^{0}$
(c)	3	54587	$2.4 \cdot 10^{-7}$	$2.0 \cdot 10^{-1}$
	6	74507	$2.4 \cdot 10^{-7}$	$2.3 \cdot 10^{-1}$
	12	76526	$2.2 \cdot 10^{-7}$	$2.2 \cdot 10^{-1}$
	37	101878	$1.1 \cdot 10^{-7}$	$2.4 \cdot 10^{-1}$
Id	# q.points	# iterations	error, objective	L^2-error, stresses
	1	51255	$3.2 \cdot 10^{-8}$	$1.1 \cdot 10^{-2}$
(h)	3	21866	$1.1 \cdot 10^{-8}$	$3.9 \cdot 10^{-3}$
	6	21336	$5.9 \cdot 10^{-6}$	$3.9 \cdot 10^{-3}$
	12	21096	$7.9 \cdot 10^{-7}$	$3.9 \cdot 10^{-3}$
	37	20783	$2.6 \cdot 10^{-8}$	$3.8 \cdot 10^{-3}$
Id	# q.points	# iterations	error, objective	L^2-error, stresses
	1	5180	$5.8 \cdot 10^{-9}$	$4.5 \cdot 10^{-4}$
(i)	3	41100	$5.7 \cdot 10^{-9}$	$1.8 \cdot 10^{-2}$
	6	79462	$3.0 \cdot 10^{-8}$	$4.4 \cdot 10^{-2}$
	12	103907	$7.8 \cdot 10^{-8}$	$7.2 \cdot 10^{-2}$
	37	128385	$1.4 \cdot 10^{-7}$	$1.1 \cdot 10^{-1}$

TABLE 3 Influence of the numerical quadrature on the algorithm. Id column refers to Figure 5 . All errors are relative

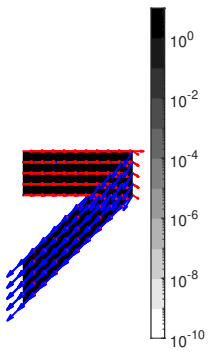


FIGURE 6 Computed solution corresponding to 1-point Gauss quadrature and weak enforcement of force equilibrium, 2-bar example.

into numerical problems earlier and earlier in terms of reducing the error. Chambolle and Pock's splitting algorithms, which we stop based on the progress they make with their primal and dual steps, are also stopped earlier because their convergence rate becomes slower with mesh refinement.

Example	Grid size	# iterations	error, objective	L^2-error, stresses
1-bar, $\Sigma^s_{f,g}$	5×5	3779	$2.8 \cdot 10^{-13}$	$3.3 \cdot 10^{-13}$
2-bar, $\Sigma^s_{f,g}$	10×30	$> 10^7$	$O(10^{-12})$	$O(10^{-9})$
1-bar, $\Sigma^w_{f,g}$	5×5	16216	$2.0 \cdot 10^{-14}$	$3.9 \cdot 10^{-11}$
2-bar, $\Sigma^w_{f,g}$	10×30	244906	$2.5 \cdot 10^{-16}$	$1.8 \cdot 10^{-2}$

TABLE 4 Chambolle and Pock's algorithm run to "full" numerical accuracy on small meshes for 1-bar and 2-bar examples, see Figures 4 (a) and 5 (g). In the second row of the table the algorithm reduces the objective function error to the order of 10^{-12}, at which point both primal and dual steps are of the order 10^{-11}. It is likely that at this point the round-off errors play a significant role in the accuracy of calculation of prox-operators.

	Grid size	# iterations	error, objective	L^2-error, stresses
(a)	5×5	718	$4.3 \cdot 10^{-7}$	$1.4 \cdot 10^{-6}$
	10×10	5528	$3.3 \cdot 10^{-7}$	$4.2 \cdot 10^{-6}$
	20×20	18461	$1.3 \cdot 10^{-7}$	$3.1 \cdot 10^{-6}$
	40×40	80114	$1.6 \cdot 10^{-7}$	$6.0 \cdot 10^{-4}$
	80×80	146017	$1.7 \cdot 10^{-7}$	$1.4 \cdot 10^{-3}$
	Grid size	# iterations	error, objective	L^2-error, stresses
(b)	5×5	5484	$2.4 \cdot 10^{-8}$	$5.4 \cdot 10^{-5}$
	10×10	8432	$2.0 \cdot 10^{-9}$	$2.9 \cdot 10^{-3}$
	20×20	18408	$5.1 \cdot 10^{-9}$	$4.5 \cdot 10^{-3}$
	40×40	53945	$2.6 \cdot 10^{-9}$	$3.7 \cdot 10^{-3}$
	80×80	113628	$7.2 \cdot 10^{-9}$	$9.0 \cdot 10^{-3}$
	Grid size	# iterations	error, objective	L^2-error, stresses
(c)	5×5	10	$2.6 \cdot 10^{-5}$	$3.1 \cdot 10^{-5}$
	10×10	11	$3.5 \cdot 10^{-5}$	$6.3 \cdot 10^{-5}$
	20×20	16	$8.6 \cdot 10^{-6}$	$1.8 \cdot 10^{-4}$
	40×40	18	$5.4 \cdot 10^{-5}$	$4.2 \cdot 10^{-3}$
	80×80	21	$7.3 \cdot 10^{-5}$	$1.6 \cdot 10^{-2}$

TABLE 5 Influence of the mesh size on the solution: 1-bar example. (a) Chambolle and Pock's algorithm in $\Sigma^s_{f,g}$; (b) Chambolle and Pock's algorithm in $\Sigma^w_{f,g}$; (c) Linear conic solver (SeDuMi) in $\Sigma^s_{f,g}$. All errors are relative.

6.8 | Influence of proximal step parameters

Naturally, varying the proximal step parameters in Chambolle and Pock's algorithm (or, more generally, selecting one of infinitely many equivalent possibilities of equipping the associated Hilbert spaces X and Y with inner products) has implications for the rate of convergence of the non-linear fixed point iteration constituting the algorithm. To provide an illustration of this issue we apply the algorithm to our two benchmarks while varying ρ_x and keeping $\rho_y = 0.9/\rho_x$. The results are summarized in Figure 7 . We only note that because we stop the algorithm based on the residual in the fixed point iteration scheme, smaller number of iterations does not necessarily imply smaller distance to the optimal solution.

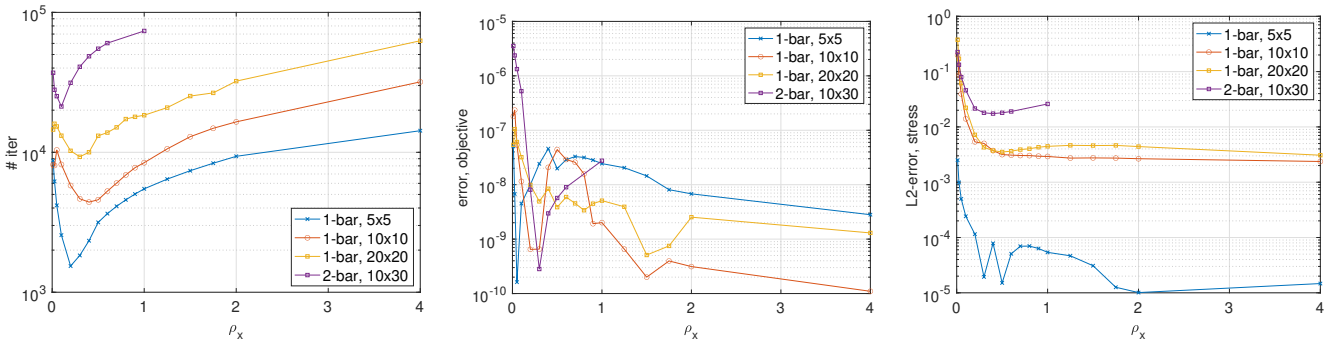
6.9 | Other benchmarks

In this section we apply the algorithms we have developed to other benchmark problems we have previously mentioned. We limit ourselves to the algorithm of Chambolle and Pock with weak enforcement of stress equilibrium, as this algorithm has the cheapest iterations and accuracy sufficient for illustrative purposes.

For the short cantilever problem on a structured grid of 320×160 squares split into four triangles we obtain the solution shown in Figure 8 , top. The algorithm stops after 2297748 (!!!) iterations with objective function of 7.030286. To put things into perspective, after 10000 iterations the value of the objective function is 7.033975, that is, within 0.05% of the final value.

	Grid size	# iterations	error, objective	L^2-error, stresses
(a)	10×30	41100	$1.1 \cdot 10^{-8}$	$3.9 \cdot 10^{-3}$
	20×60	46837	$2.8 \cdot 10^{-8}$	$9.2 \cdot 10^{-3}$
	40×120	74700	$7.0 \cdot 10^{-8}$	$2.8 \cdot 10^{-2}$
	Grid size	# iterations	error, objective	L^2-error, stresses
(b)	10×30	41100	$5.7 \cdot 10^{-9}$	$1.8 \cdot 10^{-2}$
	20×60	73666	$2.8 \cdot 10^{-8}$	$2.6 \cdot 10^{-2}$
	40×120	61919	$4.5 \cdot 10^{-8}$	$5.2 \cdot 10^{-2}$
	Grid size	# iterations	error, objective	L^2-error, stresses
(c)	10×30	20	$5.8 \cdot 10^{-5}$	$1.8 \cdot 10^{-2}$
	20×60	25	$1.2 \cdot 10^{-4}$	$3.8 \cdot 10^{-2}$
	40×120	32	$1.5 \cdot 10^{-4}$	$6.1 \cdot 10^{-2}$

TABLE 6 Influence of the mesh size on the solution: 2-bar example. (a) Chambolle and Pock's algorithm in $\Sigma^s_{f,g}$; (b) Chambolle and Pock's algorithm in $\Sigma^w_{f,g}$; (c) Linear conic solver (SeDuMi) in $\Sigma^s_{f,g}$. All errors are relative.



(a) (b) (c)

FIGURE 7 Influence of the steplength parameter on the algorithmic performance. (a) Number of iterations; (b) Relative error, objective function; (c) Relative L^2-error, stresses.

Note the presence of small stresses in the triangle adjacent to the left (fixed) boundary, where the solution should be identically zero, something we have already observed with the two-bar benchmark (see Figure 5).

For the "octagon" example (Example 2), we put $a = 0.8\sqrt{2}$, $b = 0.4$, and $c = 1$, which corresponds to the optimal objective of 1.6. We discretize the domain using an unstructured siplicial mesh with 4652 elements, and obtain the solution shown in Figure 8 , bottom. The algorithm stops after 1775 iterations with relative error in the objective value of $1.3 \cdot 10^{-9}$. Note that the obtained solution is not sparse, even though sparse(r) globally optimal solutions exist as described in Example 2.

7 | CONCLUSIONS

Despite its convexity and deceptively simple formulation (at least, as it appears in the engineering literature, see for example [16] p. 185), problem (1) seems to be quite difficult to solve. Indeed, many of the algorithms we have tested, including the classical ADMM/Douglas–Rachford iteration based optimization algorithms, are able to approximate the objective value of the problem rather accurately. Unfortunately, owing to the lack of strong convexity in the problem, this does not lead to accurate approximations of optimal solutions (stresses). Furthermore, the computed solutions seem to be rather sensitive to discretization/numerical details, as discussed in Section 6. In particular, we found that stresses tend to align themselves with the mesh, and "diffuse" when the orientations of the elements do not conform well to the orientation of the optimal stresses — or indeed in some cases even if the mesh is aligned with the solution, see for example Figure 5 (h) and (i). Therefore the strategy of working on regular

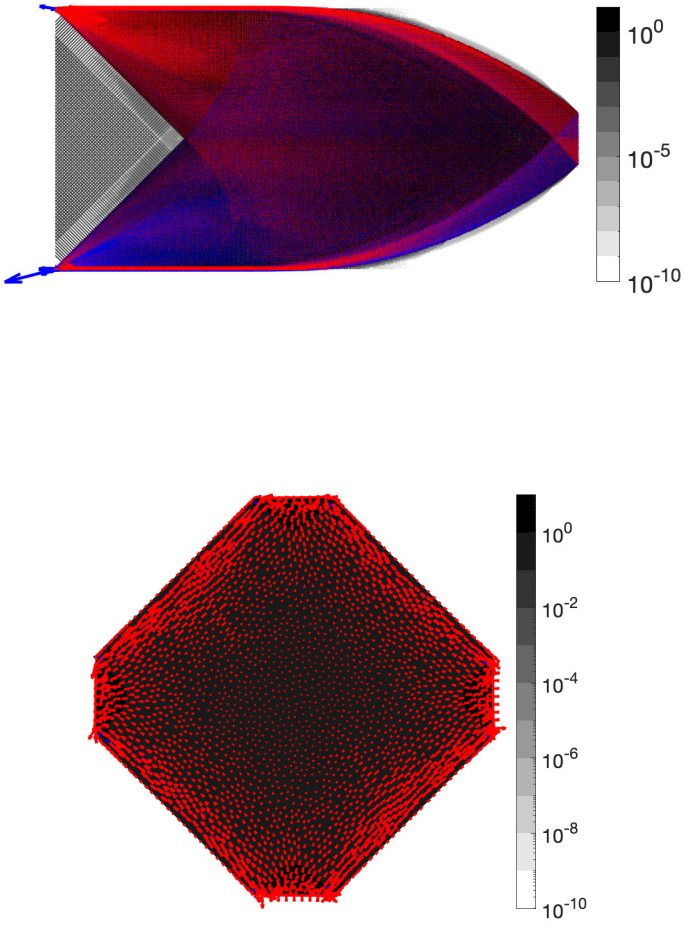


FIGURE 8 Numerical solutions to other benchmarks. Top: short cantilever. Bottom: "octagon", see Example 2.

structured grids, which has been exceedingly successfully utilized for solving topology optimization problems for many years may not be the most appropriate for problems in this class.

Even within the considered class of problems some instances appear to be considerably more difficult to solve than others. For example, the considered "1-bar" problem seems to be rather easy (even if it becomes more difficult with mesh refinement): we were very successful in accelerating the algorithms based on a fixed-point iteration formulation using standard approaches of Levenberg–Marquardt (damped semismooth Newton) or even Anderson. The same strategies fail rather miserably when applied to the "2-bar" example, particularly with the strong enforcement of force equilibrium constraint.

In general, it appears that the objective function that we utilize, which translates to compliance after the design variables are reintroduced into the problem formulation, is extremely flat. Indeed, nearly-optimal objective values are often easily attained by stress (and therefore also design) fields, which are very far from analytical optimum, and are often not "sparse enough", in the sense that they include large areas of relatively small non-zero stresses. Perhaps such solutions are suitable in applications where accuracy is not of ultimate importance (possibly certain applications of image processing), but their relevance in the mechanical engineering context is questionable.

We have also demonstrated with a simple example 1 the inherent lack of regularity of solutions to (1), thereby further emphasizing the non-triviality of the question of constructing numerical methods for this problem. Last, but most certainly not the least, we have discussed a specific example 3 for which problem (1) admits infinitely many solutions, both sparse and not

sparse, thereby showing that its objective function is not necessarily sparsity enforcing in the topology optimization context. This example alone is casting a strong shadow of a doubt onto the practical viability of methods based on (1).

References

[1] David L Donoho, *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences* **2006** *59*(6), 797.

[2] Stephen Boyd, Lieven Vandenberghe, *Convex optimization*, Cambridge university press, **2004**.

[3] Fadil Santosa, William W Symes, *SIAM Journal on Scientific and Statistical Computing* **1986** *7*(4), 1307.

[4] Robert Tibshirani, *Journal of the Royal Statistical Society: Series B (Methodological)* **1996** *58*(1), 267.

[5] Antonin Chambolle, Thomas Pock, *Journal of mathematical imaging and vision* **2011** *40*(1), 120.

[6] M Davenport, *IEEE Signal Processing Society Online Tutorial Library* **2013**.

[7] Gilbert Strang, Robert V Kohn, *Computer Methods in Applied Mechanics and Engineering* **1983** *36*(2), 207.

[8] MP Bendsøe, RB Haber, *Structural optimization* **1993** *6*(4), 263.

[9] G Allaire, RV Kohn, *European Journal of Mechanics A-solids* **1993** *12*(6), 839.

[10] Guy Bouchitté, Wilfrid Gangbo, Pierre Seppecher, *Mathematical Models and Methods in Applied Sciences* **2008** *18*(09), 1571.

[11] Heiner Olbermann, *Calculus of Variations and Partial Differential Equations* **2017** *56*(6), 166, doi:10.1007/s00526-017-1266-x.

[12] Olivier Pantz, Karim Trabelsi, *SIAM Journal on Control and Optimization* **2008** *47*(3), 1380.

[13] Jeroen P Groen, Ole Sigmund, *International Journal for Numerical Methods in Engineering* **2018** *113*(8), 1148.

[14] Andrej Cherkaev, Robert Kohn (Eds:), *Topics in the mathematical modelling of composite materials*, Birkhäuser Boston Inc., Boston, MA, **1997**.

[15] Andrej Cherkaev, *Variational Methods for Structural Optimization*, Springer-Verlag, Berlin / Heidelberg / New York / etc., **2000**.

[16] MP Bendøse, Ole Sigmund, *Topology Optimization: Theory, Methods and Applications*, Springer, **2003**.

[17] Grégoire Allaire, *Shape optimization by the homogenization method*, *Vol. 146*, Springer Science & Business Media, **2012**.

[18] A. G. M. Michell, *Philosophical Magazine* **1904** *8*(43-48), 589.

[19] Andrew J Kurdila, Michael Zabarankin, *Convex functional analysis*, Springer Science & Business Media, **2006**.

[20] C. Johnson, B. Mercier, *Numerische Mathematik* **1978** *30*(1), 103, doi:10.1007/BF01403910.

[21] Douglas Arnold, Ragnar Winther, *Numerische Mathematik* **2002** *92*(3), 401.

[22] Douglas Arnold, Gerard Awanou, Ragnar Winther, *Mathematics of Computation* **2008** *77*(263), 1229.

[23] Leszek Demkowicz, Jay Gopalakrishnan, *Numerical Methods for Partial Differential Equations* **2011** *27*(1), 70.

[24] Johnny Guzmán, Michael Neilan, *Numerische Mathematik* **2014** *126*(1), 153.

[25] Ernie Esser, *CAM report* **2009**, *9*, 31.

[26] Miguel Sousa Lobo, Lieven Vandenberghe, Stephen Boyd, Hervé Lebret, *Linear Algebra and its Applications* **1998** *284*(1), 193 , International Linear Algebra Society (ILAS) Symposium on Fast Algorithms for Control, Signals and Image Processing, doi:https://doi.org/10.1016/S0024-3795(98)10032-0.

[27] Jos F. Sturm, *Optimization Methods and Software* **1999** *11*(1-4), 625, doi:10.1080/10556789908805766.

[28] Amir Beck, *First-Order Methods in Optimization*, *Vol. 25*, SIAM, **2017**.

[29] Christoph Helmberg, Franz Rendl, *SIAM Journal on Optimization* **2000** *10*(3), 673.

[30] Alnur Ali, Eric Wong, J Zico Kolter, in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, JMLR. org, **2017**, pp. 70–79.

[31] Donald G Anderson, *Journal of the ACM (JACM)* **1965** *12*(4), 547.

[32] Anqi Fu, Junzi Zhang, Stephen Boyd, *arXiv preprint arXiv:1908.11482* **2019**.

[33] Pontus Giselsson, Mattias Fält, Stephen Boyd, in *2016 IEEE 55th Conference on Decision and Control (CDC)*, IEEE, **2016**, pp. 1015–1022.

[34] Jun Hu, Hongying Man, Shangyou Zhang, *Journal of Scientific Computing* **2014** *58*(2), 367, doi:10.1007/s10915-013-9736-6.

[35] Franco Brezzi, Jim Douglas, L. D. Marini, *Numerische Mathematik* **1985** *47*(2), 217, doi:10.1007/BF01389710.

[36] Jun Hu, Hongying Man, Jianye Wang, Shangyou Zhang, *Computers & Mathematics with Applications* **2016** *71*(7), 1317 , doi:https://doi.org/10.1016/j.camwa.2016.01.023.

[37] Christophe Geuzaine, Jean-François Remacle, *International journal for numerical methods in engineering* **2009** *79*(11), 1309.

[38] *MATLAB*, **2019**, The MathWorks, Natick, MA, USA.

[39] GIN Rozvany, *Structural optimization* **1998** *15*(1), 42.

[40] Ole Sigmund, Niels Aage, Erik Andreassen, *Structural and Multidisciplinary Optimization* **2016** *54*(2), 361.

[41] Peter J Huber, et al., *The Annals of Statistics* **1973** *1*(5), 799.

How cite this article: A. Evgrafov, and O. Sigmund (2020), Sparse basis pursuit for compliance minimization in the vanishing volume ratio limit, *Z Angew Math Mech.*, *2020;*.

APPENDIX A: FORMAL PROBLEM STATEMENT OF (1)

The material in this appendix is presented for the reader's convenience to keep this paper self-contained.

A.1 Objective function

Let $\lambda : \mathbb{S}^n \to \mathbb{R}^n$ be the function mapping a symmetric matrix to its eigenvalues in non-descenting order. Let further $\|\cdot\|_{S^p} = \|\lambda(\cdot)\|_p$, where $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}$ is the usual p-th vector norm, $1 \le p \le \infty$.[4] $\|\cdot\|_{S^p}$ is a convex function on $\mathbb{S}^n$, see for example[28] Theorem 7.17. Let us recall the duality relation

$$\|v\|_p = \sup\{\, v \cdot w = v^{\mathrm{T}} w \mid w \in \mathbb{R}^n, \|w\|_q \le 1 \,\},$$

where $p^{-1} + q^{-1} = 1$, and Fan's theorem $s : t = \operatorname{tr}(s^{\mathrm{t}}\tau) \le \lambda(s) \cdot \lambda(t)$, $\forall\, s, t \in \mathbb{S}^n$, with equality when the matrices are simultaneously diagonalizable, see for example[28] Theorem 7.14. As their consequence we have a similar duality relationship for Schatten norms:

$$\|s\|_{S^p} = \sup\{\, s : t \mid t \in \mathbb{S}^n, \|t\|_{S^q} \le 1 \,\},$$

where $p^{-1} + q^{-1} = 1$.

Let now $\sigma \in \mathcal{M}(\mathbb{R}^n; \mathbb{S}^n)$ be a finite symmetric matrix valued measure with compact support. We put

$$\|\sigma\|_{S^1} = \sup \left\{ \int_{\mathbb{R}^n} \phi(x) : d\sigma(x) \mid \phi \in C^0(\mathbb{R}^n; \mathbb{S}^n), \|\phi(x)\|_{S^\infty} \le 1 \right\}. \tag{A1}$$

Directly from this definition it follows that $\|\cdot\|_{S^1} : \mathcal{M}(\mathbb{R}^n; \mathbb{S}^n) \to \mathbb{R}$ is non-negative, convex, weakly* lower semicontinuous, and coercive (in fact, degree 1 homogeneous).

A.2 Force equilibrium constraint

Let $F \in \mathcal{M}(\operatorname{cl}\Omega; \mathbb{R}^n)$ be a vector-valued finite measure defined on $\operatorname{cl}\Omega$, representing a given load. In the situations we are interested in $F = f\mathcal{H}^n \llcorner \Omega + g\mathcal{H}^{n-1} \llcorner \Gamma_t$, where $\mathcal{H}^n$ and $\mathcal{H}^{n-1}$ are n and $n-1$ dimensional Hausdorff measures, $\llcorner$ denotes the restriction of a measure to a set[5], and $f \in L^1(\Omega; \mathbb{R}^n)$ and $g \in L^1(\Omega; \Gamma_t)$ are volumetric and boundary loads. We define

$$\tilde{\Sigma}_F = \left\{ \sigma \in \mathcal{M}(\operatorname{cl}\Omega; \mathbb{S}^n) \mid \int_{\operatorname{cl}\Omega} \nabla\phi(x) : d\sigma(x) = \int_{\operatorname{cl}\Omega} \phi(x) \cdot dF(x),\ \forall \phi \in C^1(\operatorname{cl}\Omega; \mathbb{R}^n), \phi(x) = 0,\ \text{on } \Gamma_u \right\}, \tag{A2}$$

which is a natural weak enforcement of the requirement $-\operatorname{div}\sigma = F$. Directly from the definition it follows that $\tilde{\Sigma}_F$ is a closed affine manifold in $\mathcal{M}(\operatorname{cl}\Omega; \mathbb{S}^n)$, and consequently also closed in weak* topology of $\mathcal{M}(\operatorname{cl}\Omega; \mathbb{S}^n)$.

A.3 Existence of solutions to (1)

As a direct consequence of the definitions (A1) and (A2) we are in a position to apply the generalized Weierstrass' theorem, see for example[19] Section 7.3.

Proposition 2. *Assume that $\tilde{\Sigma}_F \neq \emptyset$. Then the optimization problem* (1) *admits at least one globally optimal solution.*

Proof. As we already mentioned, we are minimizing a weakly lower semicontinuous and coercive function over a closed affine manifold, which is non-empty owing to the assumptions. Therefore the conclusion follows directly from the generalized Weierstrass' theorem.* □

[4] For $1 \le p < \infty$ we have $\|v\|_p^p = \sum_{i=1}^n |v_i|^p$, while for $p = \infty$ we have $\|v\|_\infty = \max_{1 \le i \le n} |v_i|$.
[5] For a measure μ and measurable sets A and B we have, by definition, $\mu \llcorner A(B) = \mu(A \cap B)$.

APPENDIX B: PRIMAL-DUAL ALGORITHM OF CHAMBOLLE AND POCK

Let us recall that for a proper convex and lower semi-continunous function $h : X \to \mathbb{R} \cup \{+\infty\}$ on a Hilbert space X one defines its Fenchel conjugate to be $h^*(x) := \sup_{\tilde{x}\in X}[(x,\tilde{x})_X - h(\tilde{x})]$, and a proximal operator by $\text{prox}_h(x) := \arg\min_{\tilde{x}\in X}[h(\tilde{x}) + 0.5\|x-\tilde{x}\|_X^2]$.

Let now $F : Y \to \mathbb{R}\cup\{+\infty\}, G : X \to \mathbb{R}\cup\{+\infty\}$ be two proper convex lower semi-continuous functions over Hilbert spaces X and Y with simply computable prox-operators, and $K : X \to Y$ be a bounded linear operator. For a convex minimization problem

$$\underset{x\in X}{\text{minimize}}[F(Kx) + G(x)], \tag{B3}$$

with an equivalent saddle-point reformulation

$$\min_{x\in X}\max_{y\in Y}[(y,Kx)_Y - F^*(y) + G(x)], \tag{B4}$$

the first order primal-dual algorithm of Chambolle and Pock[5] amounts to the following fixed-point iteration:

$$\begin{aligned} y^{(k+1)} &= \text{prox}_{\rho_y F^*}(y^{(k)} + \rho_y K\bar{x}^{(k)}),\\ x^{(k+1)} &= \text{prox}_{\rho_x G}(x^{(k)} - \rho_x K^* y^{(k+1)}),\\ \bar{x}^{(k+1)} &= x^{(k+1)} + \theta(x^{(k+1)} - x^{(k)}), \end{aligned} \tag{B5}$$

where $\theta \in [0,1]$, $\rho_x > 0$, $\rho_y > 0$ are parameters such that $\rho_x\rho_y\|K\|^2 < 1$, and $K^* : Y \to X$ is the Hilbert space adjoint of K defined by the variational statement $(K^*y, x)_X = (y, Kx)_Y$, $\forall(x,y) \in X \times Y$. The algorithm can be related both to the classical Dunford–Rachford splitting and ADMM, see[5] for details.