Comput. Methods Appl. Mech. Engrg. 253 (2013) 599–608

Contents lists available at SciVerse ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

An operator splitting algorithm for Tikhonov-regularized topology optimization

Cameron Talischi, Glaucio H. Paulino *

Department of Civil & Environmental Engineering, University of Illinois at Urbana-Champaign, USA

ARTICLE INFO

Article history:
Received 20 February 2012
Received in revised form 29 May 2012
Accepted 30 May 2012
Available online 20 June 2012

Keyword:
Topology optimization, Tikhonov regularization
Forward–backward splitting
Filtering

ABSTRACT

In this work, we investigate a Tikhonov-type regularization scheme to address the ill-posedness of the classical compliance minimization problem. We observe that a semi-implicit discretization of the gradient descent flow for minimization of the regularized objective function leads to a convolution of the original gradient descent step with the Green's function associated with the modified Helmholtz equation. The appearance of "filtering" in this update scheme is different from the current density and sensitivity filtering techniques in the literature. The next iterate is defined as the projection of this provisional density onto the space of admissible density functions. *For a particular choice of projection mapping, we show that the algorithm is identical to the well-known forward–backward splitting algorithm, an insight that can be further explored for topology optimization.* Also of interest is that with an appropriate choice of the projection parameter, nearly all intermediate densities are eliminated in the optimal solution using the common density material models. We show examples of near binary solutions even for large values of the regularization parameter.

1. Introduction

The lack of existence of solutions to the classical topology optimization problem is well-known (see, for example, [29,2] and references therein). The problem of minimizing compliance in structural optimization, for example, favors non-convergent minimizing sequences of shapes that exhibit progressively finer features. The commonly used density formulations, such as the popular Solid Isotropic Material with Penalization (SIMP) approach [6,42,41], wherein characteristic functions representing the shapes are replaced by density fields, continue to suffer from this pathology as the built-in penalization mechanism recovers solutions that are nearly binary in the optimal regime. A manifestation of this behavior in the finite element discretization of the problem is the dependence of solutions on the level of refinement of the spatial discretization. The problem of mesh dependency, just as the ill-posedness of the continuum problem, has led many researchers to devise formulations that are stable under mesh refinement. The search for a robust and yet mathematically consistent approach continues as evidenced by the growing number of publications on this issue [10,39,25,26,51,43,27,50].

We limit the following literature survey to density formulations but, as discussed above, the difficulty stems from a fundamental property of the original topology optimization problem and therefore similar measures are needed for other parameterizations of geometries, most notably the implicit functions methods (e.g., level sets). Placing a restriction on the perimeter of the admissible shapes is perhaps the oldest approach in the field. The set of admissible characteristic functions is restricted to a subset of functions of bounded variation with a prescribed upper bound on their total variation [3]. Existence of solutions follows from relative compactness of bounded sequences in BV in the L^1-topology and carries over to the corresponding density formulation [37]. Due to the difficulty of discretization of functions of bounded variation and robust linearization of the total variation functional [53], the perimeter formulation has perhaps fallen out of favor in the topology optimization community though it remains a significant point of reference. More recent approaches are based on the concept of filtering, which consists of implicitly imposing regularity on each admissible density function by means of convolution of an auxiliary field with a fixed and smooth filter. With such construction, all the admissible densities inherit the regularity of the filter, thereby ensuring compactness of the design space in the L^1-topology [10]. The filtering approach works well in practice since no explicit constraints on regularity of density functions are needed. Moreover the level of complexity of the final solutions (in fact all the admissible densities) is controlled directly by the regularity of the filtering kernel. *Herein lies the major drawback: the smoother the filtering kernel, the larger the amount of the intermediate densities since the transition between the extreme values of density over the domain cannot occur too rapidly*. Therefore, with more complexity control comes more "gray" regions and this, in some respect, undermines the basic premise of the density approach in that near characteristic functions are no longer recovered in the optimal regime (i.e., "0–1" or "black-and-white" designs are not obtained).

* Corresponding author.
E-mail addresses: paulino@illinois.edu, paulino@uiuc.edu (G.H. Paulino).

0045-7825/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cma.2012.05.024

We note that a similar issue arises in the slope constraint method of Petersson and Sigmund [38] where regularity is imposed explicitly by placing a constraint on the pointwise magnitude of the density gradient.

A recent trend [25,43,27,50] has focused on the so-called nonlinear filtering approaches. As pointed out in [49], the introduced nonlinearity usually amounts to a modification of the material interpolation model (e.g., SIMP) rather than a change in the filtering operation. For example, in the Heaviside filtering approach [25], both power law relations of SIMP—dependence of Young's modulus on ρ^p and volume on ρ—are augmented by the use of a smoothed Heaviside function. The additional parameter defining the sharpness of the Heaviside function controls the amount of gray that appears. To obtain good solutions, these parameters are often carefully increased throughout the course of the optimization algorithm.

It is no surprise that more fine-tuning is needed as one moves away from the simplicity of the original penalized density formulation. These schemes contain multiple penalty parameters in addition to the averaging effect of the underlying filter, which can adversely affect the quality of the reciprocal approximations of the objective function in the commonly used optimization algorithms, such as MMA [46], ultimately slowing down convergence rate or compromising the quality of the final solutions.

In this work, we examine the use of a simple Tikhonov-type regularization scheme for topology optimization. The admissible densities are defined as a subset of H^1 space with a uniform bound on their norm. In practice, this is achieved by appending the H^1 seminorm of the density function as a penalty term to the objective function. Existence of solutions follows from the compact embedding of H^1 in L^1. Such an approach has been previously studied by Borrvall in a review paper [8] where he examines penalty terms involving the L^p-norm of the density gradient (recovering the total variation regularization for $p = 1$ and the slope constraint method for $p = \infty$). This term also appears in phase field methods ([11,52,13,54,47,20]) as an interfacial energy term and is accompanied by a double-well potential penalizing intermediate densities. The two terms taken together with appropriately chosen coefficients (cf. Eq. (4.4)) serve as an approximation to the perimeter of the interface.

As we shall see in the present setting, filtering, in the form of inverse of the Helmholtz operator, naturally appears when the optimization iterations are obtained from a semi-implicit discretization of gradient flow associated with the regularized objective function. In contrast to the density and sensitivity filters, the effects of regularization term appear through smoothing of the gradient descent steps associated with the unregularized objective function. The next iterate is obtained from projection of the provisional density onto the space of admissible densities in order to enforce the 0–1 (void-solid) box constraint (and the pointwise move limit commonly introduced to stabilize the density evolution). We will show that, with a particular choice of projection map, this update scheme recovers the well-known forward–backward splitting algorithm [31,16,19,12]. This provides an alternative perspective on the proposed approach, which can be used to further investigate the theoretical and computational aspects of the algorithm leveraging the abundant literature of operator splitting and related methods. In particular, the uncoupled treatment of the Tikhonov term in the forward–backward method can be useful for more general (possibly nonsmooth) regularization approaches for topology optimization.

Finally, we note that the separation of filtering (i.e., the smoothening effect of regularization term) and projection operation in the general case offers some flexibility. In the extreme case of the L^2-projection, this algorithm, with the aid of SIMP penalization, eliminates nearly all intermediate densities regardless of the level of regularization and complexity of the final shapes. Numerically we have observed qualitatively good solutions obtained in moderate number of iterations without the need for continuation on the SIMP penalty parameter.

The remainder of the paper is organized as follows: the minimum compliance problem is stated in Section 2 and well-posedness of the regularized problem is discussed. The optimization algorithm and its relation to the forward–backward algorithm are discussed in Section 3. Several numerical results using the proposed framework are shown in Section 4. The paper is concluded with some remarks in Section 5.

Before concluding this section, we briefly describe the notation adopted in this paper. As usual, $L^p(\Omega)$ and $H^k(\Omega)$ denote the Lebesgue and Sobolev spaces defined over domain Ω. We also define the space $L^p(\Omega;K) := \{f \in L^p(\Omega) : f(\mathbf{x}) \in K \text{ for a.e. } \mathbf{x} \in \Omega\}$ given $K \subseteq \mathbb{R}$. Of particular interest are the inner product and norm associated with $L^2(\Omega)$, which are written as $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$, respectively. Similarly, the inner product, norm and semi-norm associated with $H^1(\Omega)$ are denoted by $\langle\cdot,\cdot\rangle_1$, $\|\cdot\|_1$ and $|\cdot|_1$, respectively. For $u,\ v \in H^1(\Omega)$, we shall also define the inner product $\langle u, v\rangle_\alpha := (1-\alpha)\langle u, v\rangle + \alpha\langle u, v\rangle_1$. Observe that for $0 < \alpha < 1$, the associated norm given by $\|\cdot\|_\alpha := \langle\cdot,\cdot\rangle_\alpha^{1/2}$ is equivalent to the standard norm $\|\cdot\|_1$.

2. Model problem and regularization

Let $\Omega \subseteq \mathbb{R}^d$ $(d = 2, 3)$ be a bounded set with a sufficiently smooth boundary representing the design domain for the problem and consider a nontrivial partition[1] of the boundary $\partial\Omega$ into disjoint sets Γ_D and Γ_N (cf. Fig. 2.1). In this paper, we focus our attention to the problem of compliance minimization where the objective function is given by

$$J(\rho) = \int_{\Gamma_N} \mathbf{u}_\rho \cdot \mathbf{t}\,ds + \lambda \int_\Omega \rho\,d\mathbf{x} \tag{2.1}$$

The displacement field $\mathbf{u}_\rho$ here solves the elasticity boundary value problem given by

$$\begin{aligned} &\operatorname{div}[\mathbf{C}_\rho : \boldsymbol{\varepsilon}(\mathbf{u})] = \mathbf{0} \quad \text{in } \Omega \\ &[\mathbf{C}_\rho : \boldsymbol{\varepsilon}(\mathbf{u})] \cdot \mathbf{n} = \mathbf{t} \quad \text{on } \Gamma_N \\ &\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma_D \end{aligned} \tag{2.2}$$

where $\boldsymbol{\varepsilon}(\mathbf{u}) = (\nabla\mathbf{u} + \nabla\mathbf{u}^T)/2$ is the linearized strain tensor, $\mathbf{t} \in L^2(\Gamma_N)^d$ denotes the prescribed traction vector, and $\mathbf{n}$ is the unit normal vector to $\partial\Omega$. The system of Eq. (2.2) and its solution depend on density function ρ through the elasticity tensor $\mathbf{C}_\rho$. In the SIMP model,

$$\mathbf{C}_\rho = [\epsilon + (1-\epsilon)\rho^p]\mathbf{C}^0 \tag{2.3}$$

where $p > 1$ is the penalty parameter, $\mathbf{C}^0$ is the elasticity tensor of the constituent material, and $0 < \epsilon \ll 1$ is a positive parameter ensuring well-posedness of the governing equations for every non-negative $\rho \in L^\infty(\Omega)$. In particular, given a measurable density function ρ taking values between zero and one, (2.2) admits a unique weak solution $\mathbf{u}_\rho$ in $H^1(\Omega)^d$. Since the second term in (2.1) represents a penalty on the volume of the material used, minimizing this objective amounts to finding the stiffest arrangement of $\mathbf{C}^0$ while using the least amount of material. The parameter $\lambda > 0$ determines the trade-off between the stiffness provided by the material and the amount of material that is used. Because of the monotonicity of these competing terms, it is expected that in the optimal regime, the density functions take extreme values of 0 and 1 throughout most of Ω provided that the penalty parameter p is sufficiently large. This fact was proven in [44,40,32] within

[1] That is, $\Gamma_D \cap \Gamma_N = \emptyset$, $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N$, and $|\Gamma_D| \neq 0$.

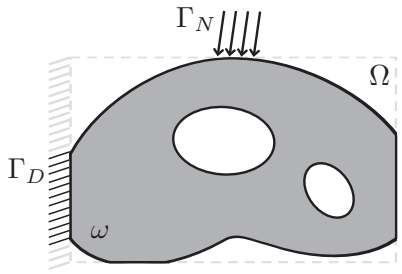

Fig. 2.1. Illustration of the prescribed boundary conditions defined on the design domain Ω. In a density formulation, each admissible shape $\omega \subseteq \Omega$ can be associated with some density function $\rho \in L^\infty(\Omega;[0,1])$.

the discrete setting where existence of solutions follows from the finite dimensionality of the problem.

We remark that frequently in the formulation of the minimum compliance problem, instead of using a penalty term as in (2.1), an explicit constraint of the form

$$\int_\Omega \rho \mathrm{d}\mathbf{x} \leqslant \overline{v}|\Omega|, \quad 0 \leqslant \overline{v} \leqslant 1 \tag{2.4}$$

is placed on the volume of the design. The two approaches are equivalent in the sense that for any prescribed volume fraction $\overline{v}$, there exists a penalty parameter λ such that the minimizer of (2.1) is also a solution to the problem with the explicit volume constraint. The converse is also true in that given $\lambda > 0$, we can define the equivalent $\overline{v}$ to be $|\Omega|^{-1}\int_\Omega \rho_\lambda^* \mathrm{d}\mathbf{x}$ where ρ_λ^* is a minimizer of $J(\rho)$. The drawback of the present formulation is that one needs to find a suitable value for the penalty parameter λ, which may not be immediately obvious. One the other hand, it has been our experience that for the compliance minimization problem, the penalty approach is more forgiving (the volume can exceed the final volume in the course of the algorithm) and leads to qualitatively better solutions.

As discussed in the introduction, to guarantee existence of solutions to the minimum compliance problem, the space of admissible densities must be restricted to a sufficiently regular subset of $L^\infty(\Omega;[0,1])$. One sufficient condition, investigated here, is to require each admissible ρ to belong to $H^1(\Omega)$ with $|\rho|_1 \leqslant M$ for some fixed positive constant M (see also, [7,8]). An alternative to adding this constraint is to modify the objective function as follows

$$\widetilde{J}(\rho) = J(\rho) + \frac{\beta}{2}|\rho|_1^2 \tag{2.5}$$

where β is a positive coefficient. The existence of minimizers for the regularized objective $\widetilde{J}$ can be established in a straight-forward manner. Any minimizing sequence ρ_n is bounded in $H^1(\Omega)$, since $\|\rho_n\| \leqslant |\Omega|$ for all n, and $\frac{\beta}{2}|\rho_n|_1^2$ remains finite. Therefore, there exists $\hat{\rho} \in H^1(\Omega)$ and a subsequence, again denoted by ρ_n, such that $\rho_n \rightharpoonup \hat{\rho}$ weakly in $H^1(\Omega)$ and by lower semicontinuity of norm under weak convergence, $|\hat{\rho}|_1 \leqslant \liminf_n |\rho_n|_1$. Moreover, by Rellich–Kondrachov theorem ([23]) and going to another subsequence, $\rho_n \to \hat{\rho}$ in $L^1(\Omega)$.[2] This in turn implies that the corresponding sequence of displacements, up to a subsequence, also converges, that is, $\mathbf{u}_{\rho_n} \to \mathbf{u}_{\hat{\rho}}$ in $H^1(\Omega)^d$ (see, for example, [9]). By continuity of compliance, we have $J(\rho_n) \to J(\hat{\rho})$ and so $\widetilde{J}(\hat{\rho}) \leqslant \liminf_n \widetilde{J}(\rho_n)$, which establishes the optimality of $\hat{\rho}$. Notice that $\hat{\rho}$ is indeed an admissible density function (satisfies the 0–1 box constraints) since on a subsequence, we have pointwise convergence of ρ_n to $\hat{\rho}$, which shows that $0 \leqslant \hat{\rho} \leqslant 1$ almost everywhere. We remark that existence of solutions to the regularized problem can be proven in a similar manner for general objective functions provided that the dependence on $(\rho, \mathbf{u}_\rho)$ is continuous in the strong topology of $L^1(\Omega) \times H^1(\Omega)^d$, which is the case for most problems of interest in engineering applications.

The Tikhonov-type regularization in (2.5) is commonly used for ill-posed inverse problems and we refer the reader to the abundant literature available (see, for example, [21] and references therein). We remark that from a theoretical perspective, regularization of densities in $H^1(\Omega)$ is more restrictive than regularization in $BV(\Omega)$ which is sufficient for guaranteeing existence of solutions for the compliance problem. However, from a practical point of view, $H^1(\Omega)$ is a simpler space to work with and, unlike total variation, the Tikhonov regularizer is smooth (differentiable) and has a quadratic form. Moreover, practical (engineering) considerations of complexity control in topology optimization (i.e., controlling feature size and orientation) can be accommodated here. A more general form of the regularization term is given by

$$R(\rho) = \frac{1}{2}\langle \nabla\rho, \kappa\nabla\rho\rangle = \frac{1}{2}\int_\Omega \nabla\rho(\mathbf{x})\cdot\kappa(\mathbf{x})\nabla\rho(\mathbf{x})\mathrm{d}\mathbf{x} \tag{2.6}$$

where κ belongs to $L^\infty(\Omega)^{d\times d}$ and for some $0 < k_1 < k_2$ satisfies $k_1 I_d \leqslant \kappa(\mathbf{x}) \leqslant k_2 I_d$ in the sense of quadratic forms for all $\mathbf{x} \in \Omega$. The existence of solutions can be shown in a similar manner since $R(\rho)$ is equivalent to the H^1 semi-norm. A suitable choice of κ can ensure a desired regularity of ρ in various parts of Ω, a fact illustrated later through a numerical example.

3. Optimization algorithm

In this section, we discuss the proposed optimization algorithm for solving the regularized topology optimization problem

$$\min_{\rho\in\mathcal{A}} \widetilde{J}(\rho) = J(\rho) + R(\rho) \tag{3.1}$$

where $J(\rho)$ is defined in (2.1), $R(\rho) = \frac{\beta}{2}|\rho|_1^2$, and the space of admissible density functions is given by $\mathcal{A} = \{\rho \in H^1(\Omega) : 0 \leqslant \rho \leqslant 1 \text{ a.e.}\}$. The *unconstrained* gradient flow corresponding to this minimization is given by

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\widetilde{J}'(\rho) = -\big[J'(\rho) + R'(\rho)\big] \tag{3.2}$$

where t is a pseudo-time variable that formally characterizes the evolution of the density function $\rho(t)$ along this descent flow. The gradient of compliance in the above expression can be calculated as ([7])

$$J'(\rho) = -(1-\epsilon)p\rho^{p-1}\boldsymbol{\varepsilon}(\mathbf{u}_\rho) : \mathbf{C}^0 : \boldsymbol{\varepsilon}(\mathbf{u}_\rho) + \lambda \tag{3.3}$$

The first term is a strain energy density field whose evaluation requires the solution $\mathbf{u}_\rho$ to (2.2). The gradient of regularization term is simply

$$R'(\rho) = -\beta\Delta\rho \tag{3.4}$$

(where Δ denotes the Lapacian) provided that $\partial\rho/\partial\mathbf{n} = 0$ on $\partial\Omega$.[3] An explicit discretization of (3.2) yields the usual gradient descent update, which due to the presence of the Laplacian term in $R'(\rho)$ requires small time increments, governed by the Courant–Friedrichs–Lewy condition, and subsequently a large number of iterations. Also, the resulting discrete dynamics may introduce

[2] In fact, by Vitali convergence theorem, this convergence holds in $L^q(\Omega)$ for any $1 \leqslant q < \infty$ since Ω has finite measure and $|\rho_n| \leqslant 1$ a.e. (see Theorem 2.24 and Proposition 2.27 in [24]).

[3] A remark is in order regarding the regularity of ρ and its boundary conditions implied by (3.4). We will essentially use the variational form of the gradient flow (3.2):

$$\langle \mathrm{d}\rho/\mathrm{d}t, \psi\rangle = \mathrm{d}\widetilde{J}(\rho)[\psi], \quad \forall\psi \in H^1(\Omega)$$

to evolve the density function (cf. Eq. (3.20)). Here $\mathrm{d}\widetilde{J}(\rho)[\psi]$ is the Gateaux derivative of $\widetilde{J}$ at ρ in the direction ψ. Observe $\mathrm{d}J(\rho)[\psi] = \langle J'(\rho), \psi\rangle$ where $J'(\rho)$ is defined in (3.3) and $\mathrm{d}R(\rho)[\psi] = \beta\langle\nabla\rho, \nabla\psi\rangle$. If additionally $\rho \in H^2(\Omega)$ and $\partial\rho/\partial\mathbf{n} = 0$ on $\partial\Omega$, then $\mathrm{d}R(\rho)[\psi] = \langle -\beta\Delta\rho, \psi\rangle = \langle R'(\rho), \psi\rangle$ but we do not need to place any additional constraints beyond $\rho \in H^1(\Omega)$.

additional regularization effects depending on the number and size of time increments, as discussed in [5], which in turn depend on the spatial discretization of ρ.

Due to the dependence of $J'(\rho)$ on the state equation, an implicit treatment of $J'(\rho)$ in (3.2) is not possible. Thus we consider a *semi-implicit* temporal discretization of the gradient flow equation, an approach also advocated by Bourdin and Chambolle [11] in their phase field method. Considering a fixed time increment τ and denoting by ρ_n the current density iterate at $t = n\tau$, the semi-implicit discretization of (3.2) takes the form

$$\frac{\rho^*_{n+1} - \rho_n}{\tau} = -J'(\rho_n) - R'(\rho^*_{n+1}) \tag{3.5}$$

where ρ^*_{n+1} is the interim or provisional iterate that may lie outside the admissible space $\mathcal{A}$ (i.e., violate the bounds on the admissible densities). We define the next iterate ρ_{n+1} to be the *projection* of ρ^*_{n+1} onto $\mathcal{A}$, that is,

$$\rho_{n+1} = \mathcal{P}_{\mathcal{A}}(\rho^*_{n+1}) \tag{3.6}$$

The algorithm consisting of (3.5) and (3.6) may be viewed as a semi-implicit form of the gradient projection method ([14]), which consists of the projection of an explicit update of (3.2). The proposed algorithm is similar to the two-step procedures for solution of the incompressible Navier–Stokes equations (cf. [33]). The standard optimality criteria (OC) algorithm also has the same two-step structure where the projection enforcing the box constraints constitutes the final step [7].

A minor modification to (3.6) allows us to accommodate a common approach for stabilizing the topology optimization algorithm, which consists of limiting the *point-wise* change in density in consecutive iterations. We can simply replace $\mathcal{A}$ in (3.6) by the subset

$$\mathcal{A}_n = \{\rho \in \mathcal{A} : \rho_n - m \leqslant \rho \leqslant \rho_n + m \text{ a.e.}\} \tag{3.7}$$

where $m \in (0, 1]$ is a prescribed *move limit*. Defining[4] $\rho^U_n = 1 \vee (\rho_n + m)$ and $\rho^L_n = 0 \wedge (\rho_n - m)$, we can write

$$\mathcal{A}_n = \left\{\rho \in H^1(\Omega) : \rho^L_n \leqslant \rho \leqslant \rho^U_n \text{a.e.}\right\} \tag{3.8}$$

The next iterate, accounting for this move limit constraint, is then

$$\rho_{n+1} = \mathcal{P}_{\mathcal{A}_n}(\rho^*_{n+1}) \tag{3.9}$$

Although reducing the time step τ can also increase the conservatism of the algorithm, based on our numerical experience so far, the move limit approach tends to perform better in practice and allows the use of a larger fixed step size τ. Of course, by setting $m = 1$, we get $\mathcal{A}_n = \mathcal{A}$ recovering the update (3.6).

The criterion for convergence of the algorithm is based on the change in the value of objective function $|\widetilde{J}(\rho_{n+1}) - \widetilde{J}(\rho_n)|/|\widetilde{J}(\rho_n)|$. Aside from the choice of the projection map $\mathcal{P}_{\mathcal{A}_n}$, the proposed update scheme contains two algorithmic parameters, namely τ and m, both of which are independent of β and the spatial discretization of ρ. More generally, in an extension of this algorithm, τ and m can be varied during the course of the algorithm to speed up convergence and/or improve stability.

3.1. Comparison with filtering methods

Substituting (3.4) into (3.5) and rearranging, we can see that the provisional iterate satisfies the modified Helmholtz equation

$$\rho^*_{n+1} - \beta\tau\Delta\rho^*_{n+1} = \rho_n - \tau J'(\rho_n) \tag{3.10}$$

with homogenous Neumann boundary conditions. It is interesting to note that the right-hand side term is the gradient descent update, with step size τ, for the *unregularized* objective function $J(\rho)$. Denoting by $G_{\beta\tau}$ the Green's function associated with Helmholtz operator in (3.10), we can write

$$\rho^*_{n+1}(\mathbf{x}) = \int_\Omega G_{\beta\tau}(\mathbf{x}, \mathbf{y})[\rho_n(\mathbf{y}) - \tau J'(\rho_n)(\mathbf{y})]\mathrm{d}\mathbf{y} \tag{3.11}$$

This shows that the candidate update ρ^*_{n+1} is obtained by filtering the "original" gradient descent update by the Gaussian kernel $G_{\beta\tau}$ whose support size depends on $\beta\tau$ (see Fig. 3.1). Although the Helmholtz equation has been previously used as a means to carry out filtering in topology optimization ([30,27]), the update expression (3.11) fundamentally differs from both density and sensitivity filtering methods as mentioned in the introduction.

It is in fact instructive to compare the proposed algorithm with the usual linear density filter, which essentially consists of the same two steps of smoothening and projection. Defining the filtering operator $\mathcal{F} = (1 - \beta\tau\Delta)^{-1}$, the update equation for the proposed algorithm can be written succinctly as

$$\rho_{n+1} = (\mathcal{P}_{\mathcal{A}_n} \circ \mathcal{F})[\rho_n - \tau J'(\rho_n)] \tag{3.12}$$

In the linear filtering method, the space of admissible densities is defined by

$$\overline{\mathcal{A}} = \{\mathcal{F}(\eta) : \eta \in L^\infty(\Omega; [0, 1])\} \tag{3.13}$$

The idea is that each density function automatically inherits its smoothness from the properties of $\mathcal{F}$ while the auxiliary functions $\eta \in L^\infty(\Omega; [0, 1])$ are updated in the optimization algorithm, and so the smoothness does not need to be enforced explicitly. If the gradient projection method is adopted[5] to do the optimization, the update expression for the auxiliary field is

$$\eta_{n+1} = \mathcal{P}_{\overline{\mathcal{A}}_n}[\eta_n - \tau J'_\eta(\eta_n)] \tag{3.14}$$

where J'_η denotes the gradient of J with respect to η and, analogously to (3.7), $\overline{\mathcal{A}}_n$ consists of $\eta \in L^\infty(\Omega; [0, 1])$ such that $|\eta - \eta_n| \leqslant m$ almost everywhere. The expression for the density update ρ_{n+1} is thus

$$\rho_{n+1} = \mathcal{F}(\eta_{n+1}) = (\mathcal{F} \circ \mathcal{P}_{\overline{\mathcal{A}}_n})\{\eta_n - \tau\mathcal{F}[J'(\rho_n)]\} \tag{3.15}$$

Comparing expressions (3.12) and (3.15), the most notable difference is the order of projection and filtering. In the density filter, by construction, the smoothness of the densities is dictated by properties of $\mathcal{F}$ (e.g., the value of $\beta\tau$) while in the proposed algorithm, the projection map also plays a role in the smoothness of the update. In the density filtering, η_n is typically binary in the optimal regime in the presence of SIMP penalization and yet, as discussed in the introduction, $\rho_n = \mathcal{F}(\eta_n)$ may contain large regions of intermediate densities depending on the smoothening effect of the filtering map $\mathcal{F}$. In the proposed algorithm, the projection map $\mathcal{P}_{\mathcal{A}_n}$ can be defined such that near binary densities are allowed. Of course, the two spaces $\mathcal{A}_n$ and $\overline{\mathcal{A}}_n$ have different structures and require different projection maps. The precise description of the projection operation is discussed next.

3.2. Projection map

We proceed to explore the possible definitions for the projection map. To this effect, consider projection with respect to the metric generated by $\|\cdot\|_\alpha$ defined in Section 1. For each $\psi \in H^1(\Omega)$, let

$$\mathcal{P}^\alpha_{\mathcal{B}}(\psi) := \operatorname*{argmin}_{\rho \in \mathcal{B}} \|\psi - \rho\|_\alpha \tag{3.16}$$

[4] Here $\wedge$ and $\vee$ denote the point-wise min/max operations.

[5] The optimality criteria (OC) is usually preferred to gradient (steepest) descent in structural optimization. We refer to [4] on the relationship between the two methods.

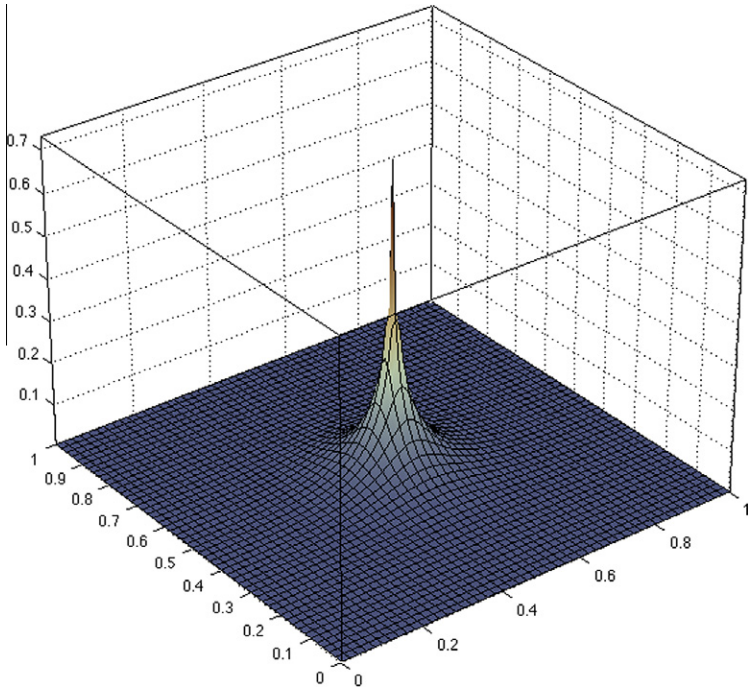

Fig. 3.1. Approximate Green's function computed numerically on a square domain Ω.

As noted before, $||\cdot||_\alpha$ defines an equivalent norm to the usual H^1 norm for $0 < \alpha < 1$. Provided that $\mathcal{B}$ is a closed convex subset of $H^1(\Omega)$, the projection $\mathcal{P}^\alpha_{\mathcal{B}}(\psi)$ exists and is unique for any $\psi \in H^1(\Omega)$. It is straightforward to show that $\mathcal{A}_n$ is closed and convex in $H^1(\Omega)$ and since $\rho^*_{n+1} \in H^1(\Omega)$ (cf. (3.10)), the update ρ_{n+1} is well-defined if we set $\mathcal{P}_{\mathcal{A}_n} = \mathcal{P}^\alpha_{\mathcal{A}_n}$ in (3.9). Note that even with the addition of an explicit volume constraint (cf. (2.4)), the space $\mathcal{A}_n$ remains closed and convex. However, in a more general setting and when dealing with nonconvex constraints, the projection operation may not be well-defined. Extending the present algorithm to such cases would require replacing the nonconvex constraints by suitable convex approximations, an approach we hope to pursue in the future.

The parameter α determines the smoothness of the projection map. Noting that

$$||\rho||^2_\alpha = \langle \rho, \rho\rangle + \alpha\langle \nabla\rho, \nabla\rho\rangle = ||\rho||^2 + \alpha|\rho|^2_1$$

we can see that the value of α determines the trade-off between minimizing the L^2 mismatch or matching the gradient values in the projection operation defined by $\mathcal{P}^\alpha_{\mathcal{B}}$. As shown in the next section, the choice $\alpha = \beta\tau$ has a particular significance for the proposed algorithm.

Also of significance is the case $\alpha = 0$ when $\mathcal{P}^\alpha_{\mathcal{B}}$ reduces to usual the L^2-projection. This is precisely the projection used in the density filtering algorithm (see Eq. (3.17) below), since the auxiliary functions in $\overline{\mathcal{A}}_n$ need not be differentiable. Even though $\mathcal{A}_n$ is *not* closed with respect to the L^2-norm, we nevertheless consider this case and define $\rho_{n+1} = \mathcal{P}^0_{\tilde{\mathcal{A}}_n}(\rho^*_{n+1})$ where $\tilde{\mathcal{A}}_n = \{\rho \in L^\infty(\Omega), \rho^L_n \leqslant \rho \leqslant \rho^U_n \text{a.e.}\}$. In fact, in the continuum setting, we can explicitly write

$$\rho_{n+1} = (\rho^*_{n+1} \wedge \rho^L_n) \vee \rho^U_n \tag{3.17}$$

Observe that the density ρ_{n+1} resulting from this projection need not lie in $H^1(\Omega)$ and so $R(\rho_{n+1}) = \frac{\beta}{2}|\rho_{n+1}|^2_1$ may not be defined. As such, the use of L^2 projection is inconsistent for solving the optimization problem (3.1). However, our numerical results show that it can produce noteworthy results (optimal densities that are nearly binary even for large β values). At any rate, the intention behind regularization for the topology optimization problem is controlling the complexity of final topologies. Restricting densities to $H^1(\Omega)$ is not necessary (or perhaps desirable) from a practical perspective. In fact, one can ignore the derivation and only focus on update Eq. (3.12) that features a particular use of filtering.

3.3. Relation to forward–backward splitting method

We next show that the proposed algorithm is related to the forward–backward splitting method (more broadly to the auxiliary problem principle ([18,17]) and cost approximation ([36]) methods) when $\mathcal{P}_{\mathcal{A}} = \mathcal{P}^{\beta\tau}_{\mathcal{A}_n}$, i.e., $\alpha = \beta\tau$ in the definition of the projection map. This connection allows us to place the algorithm on a more solid theoretical grounds and tap into the vast literature on these methods and explore the use of their many variations.

To this effect, we expand the update Eq. (3.9), which for the projection map defined in (3.16), can be equivalently written as

$$\rho_{n+1} = \operatorname*{argmin}_{\rho\in\mathcal{A}_n} ||\rho^*_{n+1} - \rho||^2_\alpha \tag{3.18}$$

Since adding or removing constant terms or multiplying by a scalar does not affect the minimizer, we have

$$\begin{aligned}
\rho_{n+1} &= \operatorname*{argmin}_{\rho\in\mathcal{A}_n} \langle\rho,\rho\rangle - 2\langle\rho^*_{n+1},\rho\rangle + \alpha\langle\nabla\rho - 2\nabla\rho^*_{n+1}, \nabla\rho\rangle \\
&= \operatorname*{argmin}_{\rho\in\mathcal{A}_n} \langle\rho,\rho\rangle - 2[\langle\rho_n - \tau J'(\rho_n),\rho\rangle \\
&\quad - \beta\tau\langle\nabla\rho^*_{n+1},\nabla\rho\rangle] + \alpha\langle\nabla\rho - 2\nabla\rho^*_{n+1},\nabla\rho\rangle \\
&= \operatorname*{argmin}_{\rho\in\mathcal{A}_n} ||\rho - [\rho_n - \tau J'(\rho_n)]||^2 + \alpha\langle\nabla\rho,\nabla\rho\rangle \\
&\quad + 2(\beta\tau - \alpha)\langle\nabla\rho^*_{n+1},\nabla\rho\rangle \\
&= \operatorname*{argmin}_{\rho\in\mathcal{A}_n} \frac{1}{2\tau}||\rho - [\rho_n - \tau J'(\rho_n)]||^2 + \frac{\alpha}{\beta\tau}R(\rho) \\
&\quad + \left(1 - \frac{\alpha}{\beta\tau}\right)\mathrm{d}R(\rho^*_{n+1})[\rho]
\end{aligned} \tag{3.19}$$

where $\mathrm{d}R(\rho^*_{n+1})[\rho]$ denotes the Gateaux derivative of R at ρ^*_{n+1} in the direction of ρ. Note that in the second equality above, we have used the fact that ρ^*_{n+1} solves the variational form of (3.10) given by

$$\langle \rho^*_{n+1}, \rho\rangle + \beta\tau\langle \nabla\rho^*_{n+1}, \nabla\rho\rangle = \langle \rho_n - \tau J'(\rho_n), \rho\rangle, \quad \forall \rho \in H^1(\Omega) \tag{3.20}$$

The first term in the minimization problem (3.19) measures the L^2 distance of ρ with the gradient descent step associated with J while the last two terms give an interpolation between $R(\rho)$ and its derivative at ρ^*_{n+1} as determined by projection parameter α. For $\alpha = \beta\tau$, this reduces to

$$\rho_{n+1} = \operatorname*{argmin}_{\rho\in\mathcal{A}_n} \frac{1}{2\tau}||\rho - [\rho_n - \tau J'(\rho_n)]||^2 + R(\rho) \tag{3.21}$$

which is precisely the iterations defined by the so-called *forward–backward splitting* procedure for minimization problem (3.1) ([16]). The intuition behind (3.21) is that the next iterate ρ_{n+1} is close to the gradient descent update on J, i.e., $\rho_n - \tau J'(\rho_n)$ while minimizing the complexity term $R(\rho)$. The forward–backward splitting allows for separate treatment of the constituent terms of $\widetilde{J}$, which is particularly useful when $R(\rho)$ is nonsmooth. Defining

$$K(\rho;\psi) = \frac{1}{2\tau}||\rho - [\psi - \tau J'(\psi)]||^2 + R(\rho)$$

it can be readily shown that if $\hat\rho$ minimizes $K(\rho;\hat\rho)$ in $\mathcal{A}$, then $\hat\rho$ is also a minimizer of $\widetilde{J}(\rho)$ ([18]). This illustrates the fact that if the sequence ρ_n produced by iterations (3.21) converges, then the limit is the solution to the minimization problem (3.1).

It is insightful to note that the forward-back iteration is equivalent to

$$\rho_{n+1} = \operatorname*{argmin}_{\rho\in\mathcal{A}_n} J(\rho_n) + \langle J'(\rho_n), \rho - \rho_n\rangle + \frac{1}{2\tau}||\rho - \rho_n||^2 + R(\rho) \tag{3.22}$$

We can see that, in the forward–backward subproblem, J is replaced by a local quadratic model whose curvature depends on $1/\tau$. The magnitude of τ affects how far ρ_{n+1} is from ρ_n. As noted before, the move limit constraint introduced in $\mathcal{A}_n$ also limits the change between ρ_n and ρ_{n+1} and so, for $\alpha = \beta\tau$ and fixed m, larger values of τ may accelerate convergence of the algorithm. Also, in light of (3.22), we can improve the performance of the algorithm by varying the step size parameter in the course of the algorithm. For example, we may choose step size τ_n in the nth iteration such that $\tau_n^{-1} I$ is a better approximation to the Hessian $J''(\rho_n)$.

4. Numerical investigations

In this section, we assess the performance of the proposed algorithm and present some numerical results for the compliance minimization problem. First we discuss some implementation aspects and efficiency considerations related to this algorithm. The numerical results are presented next with emphasis placed on the two extreme choices of the projection parameter, namely $\alpha = 0$ (the L^2-projection) and $\alpha = \beta\tau$ (the forward–backward algorithm).

4.1. Implementation

In this work, the density field is discretized by means of finite elements which allow for the use of unstructured grids necessary for representing arbitrary design domains Ω. Given an appropriate set of basis functions $\{N_i\}_{i=1}^M$, each admissible density function is written as $\rho = \sum_{i=1}^M N_i z_i$ where $0 \leqslant z_i \leqslant 1$ for all i and so it is characterized by the vector of nodal values $\mathbf{z} = [z_i]_{i=1}^M$. The Galerkin discretization of (3.10) yields the linear system

$$(\mathbf{M} + \beta\tau\mathbf{G})\mathbf{z}^*_{n+1} = \mathbf{M}\mathbf{z}_n - \tau\mathbf{F}_n \tag{4.1}$$

where $[\mathbf{M}]_{ij} = \langle N_i, N_j\rangle$ and $[\mathbf{G}]_{ij} = \langle \nabla N_i, \nabla N_j\rangle$ are the standard finite element matrices and $[\mathbf{F}_n]_i = \langle J'(\rho_n), N_i\rangle$. Note that the matrix $\mathbf{M} + \beta\tau\mathbf{G}$ does not change during the course of the algorithm (unless the mesh is changed) and thus can be factored once in the beginning of the algorithm. The cost of solving (4.1) is then negligible during the subsequent iterations. We note that for large-scale problems where factorization is not feasible and iterative solvers are necessary, one could use Krylov recycling (see, for example, [22,35]) to take advantage of the fact only the right-hand-side vector changes in the sequence of linear systems (4.1).

The discrete counterpart to the projection (3.18) is

$$\mathbf{z}_{n+1} = \operatorname*{argmin}_{\mathbf{z}_n^L \leqslant \mathbf{z} \leqslant \mathbf{z}_n^U} (\mathbf{z} - \mathbf{z}^*_{n+1})^T(\mathbf{M} + \alpha\mathbf{G})(\mathbf{z} - \mathbf{z}^*_{n+1}) \tag{4.2}$$

where $\mathbf{z}_n^L$ and $\mathbf{z}_n^U$ are defined in an obvious way. Note that minimization problem (4.2) is a sparse (strictly) convex quadratic program subject to simple bound constraints and can be solved efficiently using, for example, the active set method. However, we can again exploit the fact that Hessian $\mathbf{M} + \alpha\mathbf{G}$ is fixed in the course of the optimization. Using the Cholesky decomposition $\mathbf{M} + \alpha\mathbf{G} = \mathbf{R}^T\mathbf{R}$, we can write

$$\mathbf{z}_{n+1} = \operatorname*{argmin}_{\mathbf{z}_n^L \leqslant \mathbf{z} \leqslant \mathbf{z}_n^U} \left|\mathbf{R}\mathbf{z} - \mathbf{R}\mathbf{z}^*_{n+1}\right|^2 \tag{4.3}$$

where $|\cdot|$ is the standard Euclidean norm. Therefore, upon calculation of $\mathbf{R}$ once in the beginning of the algorithm, we only need to solve a bound constrained sparse least squares problem in each iteration, a simpler problem which can be solved efficiently, for example, by algorithms proposed in ([1]). In fact, with this approach, the dominant cost in each iteration of the topology optimization algorithm is computing the compliance sensitivities, i.e., vector $\mathbf{F}_n$, which requires the solution to the elasticity system (2.2). We remark that this is still the case if one wishes to enforce the volume constraint explicitly (cf. (2.4)). This requires including an additional linear constraint of the form $\mathbf{v}^T\mathbf{z} \leqslant \overline{v}|\Omega|$ where $[\mathbf{v}]_i = \int_\Omega N_i \mathrm{d}\mathbf{x}$ and the above-mentioned algorithms for solving sparse quadratic programs or least squares problems are capable of handling it.

For the sake of simplicity and following the common approach, we use the same finite element mesh describing the density field to solve the state equation. The concept of generalized isoparametric finite elements ([28]) is fitting as the density field (and consequently $\mathbf{C}_\rho$) and displacement fields are discretized on the same mesh. Within this framework, two accuracy considerations concerning (3.10) and (2.2) should guide the appropriate choice of finite element discretization (type of basis functions and level of mesh refinement). Though, we chose to use a fixed grid for the entire course of optimization for the results in this paper, these criteria can be used to devise an adaptive finite element strategy.

4.2. Results

We consider two benchmark compliance minimization problems, namely the MBB beam problem ([34]) and the bridge problem, whose domain geometry and prescribed loading and boundary conditions are shown in Fig. 4.1. The value of volume coefficient λ was set to $200|\Omega|^{-1}$ and $70|\Omega|^{-1}$ for these problems, respectively. For all the results, the constituent material $\mathbf{C}^0$ was assumed to be isotropic with unit Young's modulus and Poisson's ratio of $\nu = 0.3$ and the Ersatz stiffness was set to $\epsilon = 10^{-4}$. The SIMP penalty parameter, the move limit, and the step size were fixed at $p = 3$, $m = 0.02$ and $\tau = 0.75\lambda^{-1}$ throughout the course of

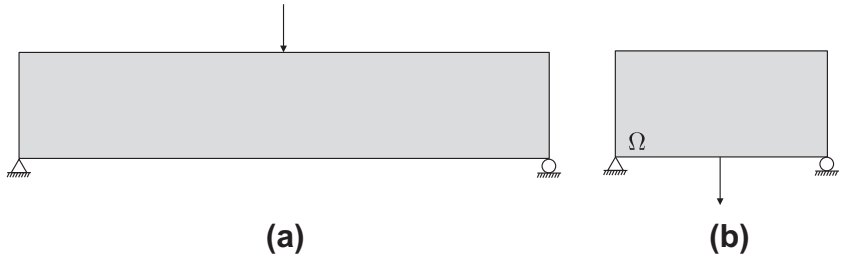

Fig. 4.1. Design domain and boundary conditions for (a) the MBB beam problem (the design domain has height $h = 1$ and width $w = 6$) and (b) the bridge problem (the design domain has height $h = 1$ and width $w = 2$). In both cases, the applied load has unit magnitude.

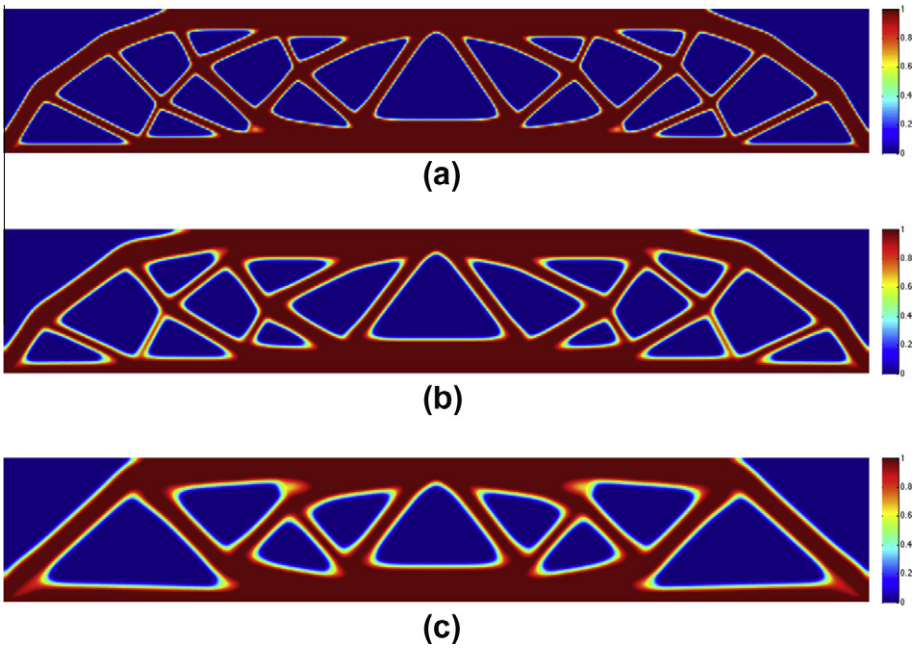

Fig. 4.2. Solutions to the MBB beam problem using the forward–backward algorithm, i.e, $\alpha = \beta\tau$, and complexity parameter (a) $\beta = 0.01$ (b) $\beta = 0.03$ and (c) $\beta = 0.06$.

optimization (i.e., no continuation was carried out) and the convergence tolerance of 10^{-6} was used. The initial guess in all cases was taken to be uniform density field with value of 0.5.

The first set of results explores the influence of β on the complexity of solutions to the MBB beam problem. The results were obtained using a uniform mesh of 360×60 standard bilinear quadrilateral elements for both density and displacement fields. Fig. 4.2 shows the solutions obtained using the forward–backward algorithm, i.e., $\alpha = \beta\tau$, for various values of β. As expected, increase in β results in smoother final solutions with fewer but large members. The same is true for the L^2-projection (with $\alpha = 0$) as shown in Fig. 4.3. Note, however, that the final densities are nearly binary with this type of projection even for large of values of β. In fact, only 10% of the nodal densities in the solution shown in Fig. 4.3(c) had values in the range (0.05,0.95) in contrast to the 25% for the solution in Fig. 4.2(c) using the forward–backward algorithm. For the sake of comparison, we have computed another measure of discreteness, proposed in [43], and the results are presented in Table 1.

Next we investigate the influence of mesh size on the final densities. The MBB beam problem was solved using finer grids consisting of 600×100 and 900×150 elements using the same complexity parameters as in the previous results. The final topologies were nearly identical for all the parameters (forward–backward and L^2-projection schemes) and due to the similarity with those obtained from coarser 360×60 mesh are not shown. In the case of L^2-projection, even though we evidently have convergence of the final densities under mesh refinement with respect to the L^p-norm, the regularization term $R(\rho)$ blows up (and convergence does not hold in H^1). For $\beta = 0.01$, the values of the regularization term for the final solutions were 7.56, 9.89 and 9.93 for the 360×60, 600×100 and 900×150 meshes, respectively. We note, however, that the perimeter (total variation of optimal density field) remained bounded under mesh refinement, which suggests a possible connection with perimeter constraint problem. For the case of $\beta = 0.01$, the total variation of the final solutions were 33.3, 32.9 and 32.7 for these meshes, respectively. Similarly, for $\beta = 0.05$, the total variation of the final solutions were 20.1, 19.7, and 19.5, respectively.

We intend to investigate this observation more in our future work but for now we note that the phase field approximation of the perimeter constraint problem is given by (see, for example [11,13])

$$\min_{\rho \in \mathcal{A}} \; J(\rho) + \gamma \left[\delta |\rho|_1^2 + \frac{1}{\delta} \int_\Omega W(\rho) \mathrm{d}\mathbf{x} \right] \tag{4.4}$$

where $\delta > 0$ and $W(\rho)$ is a strictly positive function that only vanishes at $\rho = 0$ and $\rho = 1$. Observe that the first term is identical to the present Tikhonov regularization term when $\gamma\delta = 2\beta$. Moreover, a binary density field (which is what our algorithm nearly produces

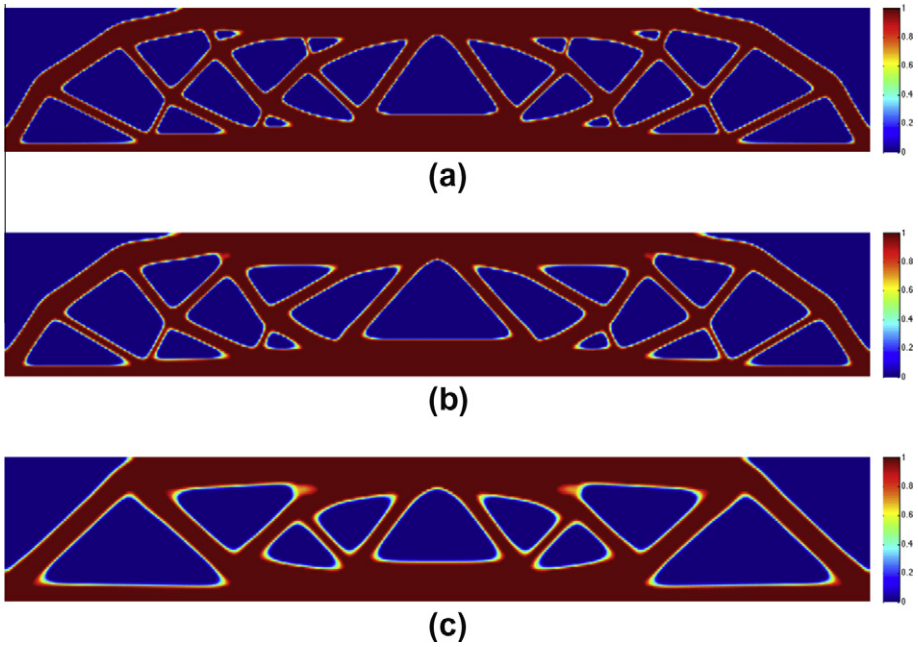

Fig. 4.3. Solutions to the MBB beam problem using the L^2 projection, i.e, $\alpha = 0$, and complexity parameter (a) $\beta = 0.01$ (b) $\beta = 0.02$ and (c) $\beta = 0.05$.

Table 1
Summary of the results for the MBB beam problem.

Projection type	$\alpha = \beta\tau$			$\alpha = 0$		
	0.01	0.03	0.06	0.01	0.02	0.05
$\int_{\Gamma_N} \mathbf{u}_\rho \cdot \mathbf{t} \mathrm{d}s$	101.91	101.32	100.72	95.92	95.16	94.57
$\lambda \int_\Omega \rho \mathrm{d}\mathbf{x}$	15.43	16.22	16.83	16.40	16.73	17.53
$\frac{\beta}{2}\int_\Omega \lvert\nabla\rho\rvert^2 \mathrm{d}\mathbf{x}$	5.706	8.626	8.927	7.563	11.11	11.29
$\frac{4}{\lvert\Omega\rvert}\int_\Omega \rho(1-\rho)\mathrm{d}\mathbf{x}$ (%)	7.62	13.7	14.7	3.22	5.00	8.65

Table 2
Summary of mesh refinement study for the bridge problem with $\beta = 0.02$ and L^2-projection.

Grid size	100×50	150×75	200×100	300×150
Number of iterations	91	105	94	93
$\int_{\Gamma_N} \mathbf{u}_\rho \cdot \mathbf{t} \mathrm{d}s$	20.35	20.84	21.22	21.81
$\lambda \int_\Omega \rho \mathrm{d}\mathbf{x}$	25.06	24.43	24.15	23.94
$\frac{\beta}{2}\int_\Omega \lvert\nabla\rho\rvert^2 \mathrm{d}\mathbf{x}$	3.340	3.828	4.214	4.544
$\int_\Omega \lvert\nabla\rho\rvert \mathrm{d}\mathbf{x}$	11.53	11.55	11.54	11.51

when $\alpha = 0$) minimizes the second term in the phase field approximation. This formal argument supports our numerical observations and highlights the fact that the underlying penalization mechanism in SIMP (combined with effects of nonsmooth L^2-projection) is an effective replacement for the penalization term in the phase field approximation of the perimeter constraint problem.

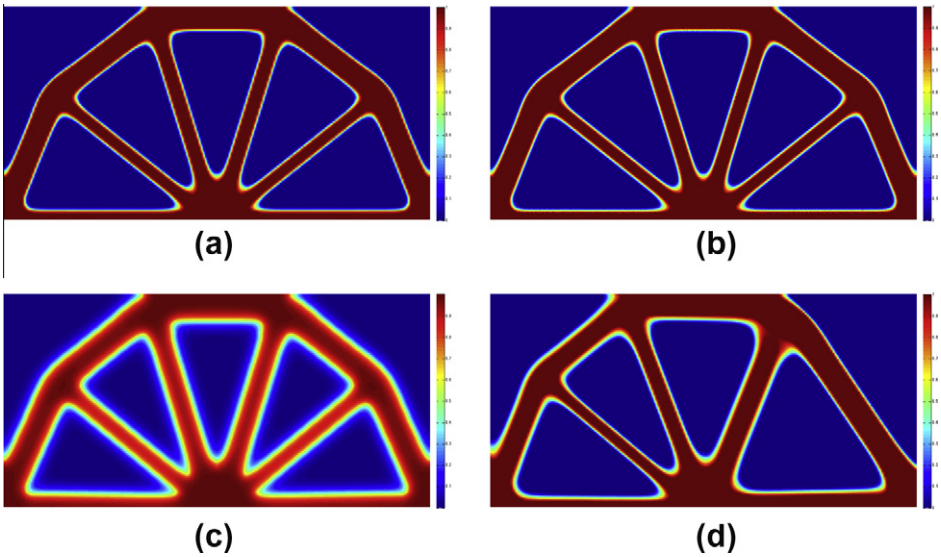

Fig. 4.4. Solutions to the the bridge problem using (a) L^2-projection, isotropic regularization and structured square mesh (b) L^2-projection, isotropic regularization and unstructured polygonal mesh (c) equivalent (same β and τ as the previous two cases) density filtering (d) L^2-projection scheme with anisotropic regularization term.

Fig. 5.1. The original image (left) with noise added (middle) and its reconstruction (left) using total variation minimization algorithm of the form (5.1). Images courtesy of [15].

The last set of numerical results, shown in Fig. 4.4, is for the bridge problem. The results were obtained for $\beta = 0.02$ and $\alpha = 0$ (i.e., the L^2-projection) using a structured square mesh and an unstructured mesh consisting of linear convex polygons (see [45,48] for the finite element formulation). Both meshes are made up of 20,000 elements. The final densities, shown in Fig. 4.4(a) and (b), are nearly identical despite the difference in the choice of the spatial discretization. It is interesting to note that even though the polygonal mesh used was not symmetric about the midspan axis, the final topology is symmetric. A summary of the mesh refinement study for the bridge is given in Table 2. Again we can see that the total variation of final solutions remains constant under mesh refinement while the regularization term grows.

For the sake of comparison, we also solved an equivalent problem (i.e., with the same parameter values for β and τ) using the density filtering method (cf. Eqs. (3.14) and (3.15)).[6] It can be seen from Fig. 4.4(c) that the final topology is the same but the density field is heavily smeared. In fact, only less than 48% of the nodal densities are in the range $[0, 0.05] \cup [0.95, 1]$ in contrast to 91% for the solution using the L^2-projection scheme. In terms of the discreteness measure of [43], the density filtering solution has a value of 33.1% while it is 7.04% for the L^2-projection result.

Lastly, we solved the bridge problem for the more general regularization term (2.6) with

$$\kappa(\mathbf{x}) = \begin{pmatrix} 0.05 & -0.03 \\ -0.03 & 0.05 \end{pmatrix}$$

taken to be constant over the entire domain. The eigenvectors of κ, namely $\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1,1)^T$ and $\mathbf{v}_2 = \frac{1}{\sqrt{2}}(1,-1)^T$, are rotated 45° from the horizontal axis and the corresponding eigenvalues are $\beta_1 = 0.02$ and $\beta_2 = 0.08$, respectively. We can see from the solution in Fig. 4.4(d) that the introduced anisotropy of the regularization term breaks the symmetry. The diagonal members on the right half, which are nearly perpendicular to $\mathbf{v}_2$, are penalized more and are thus collapsed into one member. This example illustrate the potentials of the proposed regularization scheme for control of local orientation and feature size.

5. Concluding remarks

In this work we proposed a simple but effective algorithm for solving the compliance minimization problem with Tikhonov-type regularization. As mentioned before, extensions of the present framework include more sophisticated choices of algorithmic parameters such as m and τ that can be adaptively changed to speed up or stabilize convergence as well as continuation on the complexity control regularization term (e.g., through the value of β) to avoid convergence to local minima. Also of interest is an adaptive finite element scheme based on the accuracy considerations pertaining to the representation of design (density) and response (displacement) fields. Additionally, a decoupling of the spatial discretization of these fields can conceivably lead to more computational savings provided that the appropriate and efficient infrastructure (e.g. data structure, transfer operators) is available.

We also find promising the use of operator splitting methods (such as the forward–backward algorithm) for nonsmooth regularization of the topology optimization. In particular, it would be interesting to explore such a decoupling approach for the perimeter constrained problem where the regularization term $R(\rho)$ is defined to be the total variation of the density field $\int_\Omega |\nabla\rho| d\mathbf{x}$. In this case, the forward–backward splitting leads to simpler subproblems of the form

$$\min_{\rho \in BV(\Omega;[0,1])} ||\rho - g_n||^2 + k \int_\Omega |\nabla\rho| d\mathbf{x} \tag{5.1}$$

for which efficient algorithms can be found in the image processing literature (e.g., [15,12]; also see Fig. 5.1 for an example of an image denoising problem). Owing to the simple structure of (5.1), there is no need to approximate total variation by a differentiable functional and therefore the effects of the regularizer can be captured with a high degree of fidelity.

We end with a general remark on the justification of restriction formulations for topology optimization such as the one presented here. Though they seemingly involve an arbitrary modification of the original problem (e.g., requiring the density fields to be uniformly smooth in the filtering method or belonging to a bounded subset of H^1 in the present work), such restrictions are more than a theoretical tool since they can be used to enforce manufacturing constraints on the admissible shapes that can be built for engineering applications or, in the case of inverse problems (e.g., obstacle identification), introduce *a priori* knowledge about the regularity of the unknown geometry. The ultimate test of the resulting algorithms, aside from usual criteria of robustness, feasibility and ease of implementation and computational cost, currently rests on "qualitative" inspection of the final topologies.

Acknowledgements

The authors acknowledge the support by the Department of Energy Computational Science Graduate Fellowship Program of the Office of Science and National Nuclear Security Administration in the Department of Energy under contract DE-FG02-97ER25308, and the National Science Foundation (NSF) through grant #1234243 (Civil, Mechanical and Manufacturing Innovation Division). The information presented in this paper is the sole opinion of the authors and does not necessarily reflect the views of the sponsoring agencies.

[6] Note that the smoothness of filter $\mathcal{F}$ depends on $\beta\tau$ for both density filtering and the proposed L^2-projection scheme. By contrast, the introduced complexity in the forward–backward algorithm only depends on β as evident from (3.22).

References

[1] M. Adlers, Sparse least squares problems with box constraints, Department of Mathematics, Linkoping University, Thesis, 1998.
[2] G. Allaire, Shape Optimization by the Homogenization Method, Springer, Berlin, Heidelberg, New York, 2001.
[3] L. Ambrosio, G. Buttazzo, An optimal design problem with perimeter penalization, Calc. Var. Partial Diff. Equat. 1 (1) (1993) 55–69.
[4] JS. Arora, Analysis of optimality criteria and gradient projection methods for optimal structural design, Comput. Methods Appl. Mech. Engrg. 23 (2) (1980) 185–213.
[5] UM. Ascher, H. Huang, K. Van den Doel, Artificial time integration, BIT Numer. Math. 47 (1) (2007) 3–25.
[6] MP. Bendsoe, Optimal design as material distribution probelm, Struct. Optimiz. 1 (1989) 193–202.
[7] MP. Bendsøe, O. Sigmund, Topology Optimization: Theory, Methods and Applications, Springer, 2003.
[8] T. Borrvall, Topology optimization of elastic continua using restriction, Arch. Comput. Methods E 8 (4) (2001) 251–285. Jun.
[9] T. Borrvall, J. Petersson, Topology optimization using regularized intermediate density control, Comput. Methods Appl. Mech. Engrg. 190 (37–38) (2001) 4911–4928.
[10] B. Bourdin, Filters in topology optimization, Int. J. Numer. Methods Engng 50 (9) (2001) 2143–2158.
[11] B. Bourdin, A. Chambolle, Design-dependent loads in topology optimization, ESAIM Control Optimiz. Calc. 9 (2) (2003) 19–48.
[12] K. Bredies, A forward–backward splitting algorithm for the minimization of non-smooth convex functionals in Banach space, Inverse Probl. 25 (2009) 015005.
[13] M. Burger, R. Stainko, Phase-field relaxation of topology optimization with local stress constraints, SIAM J Control Optimiz. 45 (4) (2006) 1447–1466.
[14] P.H. Calamai, J.J. Moré, Projected gradient methods for linearly constrained problems, Math. Program. 39 (1) (1987) 93–116.
[15] A. Chambolle, An algorithm for total variation minimization and applications, J. Math. Imaging Vis. 20 (1-2) (2004) 89–97.
[16] G.H.G. Chen, R.T. Rockafellar, Convergence rates in forward-backward splitting, SIAM J. Optimiz. 7 (2) (1997) 421–444.
[17] A. Cohen, R. Masson, Wavelet methods for second-order elliptic problems preconditioning and adaptivity, SIAM J. Sci. Comput. 21 (3) (2000) 1006–1026.
[18] G. Cohen, Optimization by decomposition and coordination: a unified approach, IEEE Trans. Automat. Control 23 (2) (1978) 222–232.
[19] P.L. Combettes, V.R. Wajs, Signal recovery by proximal forward-backward splitting, Multiscale Model. Simulat. 4 (4) (2006) 1168–1200.
[20] L. Dede, M.J. Borden, T.J.R. Hughes, Isogeometric analysis for topology optimization with a phase field model, ICES-Report 11-29, The Institute for Computational Engineering and Sciences, 2011.
[21] H.W. Engl, M. Hangke, A. Neubauer, Regularization of Inverse Problems, Kluwer Academic Publishers, 1996.
[22] J. Erhel, F. Guyomarc'h, An augmented conjugate gradient method for solving consecutive symmetric positive definite linear systems, SIAM J. Matrix. Anal. Appl. 21 (4) (2000) 1279–1299.
[23] L.C. Evans, Partial differential equations, American Mathematical Society Providence, Rhode Island, 1998.
[24] I. Fonseca, G. Leoni, Modern Methods in the Calculus of Variations: Lp Spaces, Springer, New York, 2007.
[25] J.K. Guest, J.H. Prevost, T. Belytschko, Achieving minimum length scale in topology optimization using nodal design variables and projection functions, Int. J. Numer. Meth. Engng. 61 (2) (2004) 238–254.
[26] X. Guo, Y.X. Gu, A new density-stiffness interpolation scheme for topology optimization of continuum structures, Engineering Computations 21 (1) (2004) 9–22.
[27] A. Kawamoto, T. Matsumori, S. Yamasaki, T. Nomura, T. Kondoh, S. Nishiwaki, Heaviside projection based topology optimization by a PDE-filtered scalar function, Struct. Multidisc Optimiz. 44 (1) (2010) 19–24.
[28] J.H. Kim, G.H. Paulino, Isoparametric graded finite elements for nonhomogeneous isotropic and orthotropic materials, J. Appl. Mech. T ASME 69 (4) (2002) 502–514.
[29] R.V. Kohn, G. Strang, Optimal design and relaxation of variational problems, Int. Commun. Pure Appl. Math. 39 (1) (1986) 113–137.
[30] B.S. Lazarov, O. Sigmund, Filters in topology optimization based on Helmholtz-type differential equations, Int. J. Numer. Methods Engrg. 86 (6) (2011) 765–781.
[31] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979) 964–979.
[32] J.M. Martinez, A note on the theoretical convergence properties of the SIMP method, Struct. Multidisc Optimiz. 29 (4) (2005) 319–323.
[33] M.L. Minion, Higher-order semi-implicit projection methods, in: Numerical Simulations of Incompressible Flows, 2003, pp.126–140.
[34] N. Olhoff, M.P. Bendsoe, J. Rasmussen, On CAD-integrated structural topology and design optimization, Comput. Methods Appl. Mech. Engrg. 89 (1–3) (1991) 259–279.
[35] M.L. Parks, E. de Sturler, G. Mackey, D.D. Johnson, S. Maiti, Recycling Krylov supaces for sequences of linear systems, SIAM J. Sci. Comput. 28 (5) (2006) 1651–1674.
[36] M. Patriksson, Cost approximation: a unified framework of descent algorithms for nonlinear programs, SIAM J Optimiz. 8 (2) (1998) 561–582.
[37] J. Petersson, Some convergence results in perimeter-controlled topology optimization, Comput. Methods Appl. Mech. Engrg. 171 (1–2) (1999) 123–140.
[38] J. Petersson, O. Sigmund, Slope constrained topology optimization, Int. J. Numer. Methods Engng. 41 (8) (1998) 1417–1434.
[39] T.A. Poulsen, A new scheme for imposing a minimum length scale in topology optimization, Int. J. Numer. Meth. Engng 57 (6) (2003) 741–760.
[40] A. Rietz, Sufficiency of a finite exponent in SIMP (power law) methods, Struct. Multidisc Optimiz. 21 (2001) 159–163.
[41] G.I.N. Rozvany, A critical review of established methods of structural topology optimization, Struct. Multidisc Optimiz. 37 (3) (2009) 217–237.
[42] G.I.N. Rozvany, M. Zhou, T. Birker, Generalized shape optimization without homogenization, Struct. Optimiz. 4 (3–4) (1992) 250–252.
[43] O. Sigmund, Morphology-based black and white filters for topology optimization, Struct. Multidisc Optimiz. 33 (4–5) (2007) 401–424.
[44] M. Stolpe, K. Svanberg, On the trajectories of penalization methods for topology optimization, Struct. Multidisc Optimiz. 21 (2) (2001) 128–139.
[45] N. Sukumar, A. Tabarraei, Conforming polygonal finite elements, Int. J. Numer. Methods Engng. 61 (12) (2004) 2045–2066, http://dx.doi.org/10.1002/nme.1141. Jan..
[46] K. Svanberg, The method of moving asymptotes–A new method for structural optimization, Int. J. Numer. Methods Engng 24 (2) (1987) 359–373.
[47] A. Takezawa, S. Nishiwaki, M. Kitamura, Shape and topology optimization based on the phase field method and sensitivity analysis, J. Comput. Phys. 229 (7) (2010) 2697–2718.
[48] C. Talischi, G.H. Paulino, A. Pereira, I.F.M. Menezes, Polygonal finite elements for topology optimization: a unifying paradigm, Int. J. Numer. Methods Engrg. 82 (6) (2010) 671–698.
[49] C. Talischi, G.H. Paulino, A. Pereira, I.F.M. Menezes, PolyTop: A Matlab implementation of a general topology optimization framework using unstructured polygonal finite element meshes, Struct. Multidisc Optimiz. 45 (2011) 329–357.
[50] F. Wang, B.S. Lazarov, O. Sigmund, On projection methods convergence robust formulations in topology optimization, Struct. Multidisc Optimiz. 43 (6) (2011) 767–784.
[51] M.Y. Wang, S.Y. Wang, Bilateral filtering for structural topology optimization, Int. J. Numer. Methods Engng. 63 (13) (2005) 1911–1938.
[52] M.Y. Wang, S. Zhou, Synthesis of shape and topology of multi-material structures with a phase-field method, J. Comput. Aided Mater. Des. 11 (2-3) (2004) 117–138.
[53] W.H. Zhang, P. Duysinx, Dual approach using a variant perimeter constraint efficient sub-iteration scheme for topology optimization, Comput. Struct. 81 (22-23) (2003) 2173–2181.
[54] S. Zhou, M.Y. Wang, Multimaterial structural topology optimization with a generalized cahn–hilliard model of multiphase transition, Struct. Multidisc Optimiz. 33 (2) (2007) 89–111.