

Numerical analysis of the SIMP model for the topology optimization problem of minimizing compliance in linear elasticity

Ioannis P. A. Papadopoulos

submitted: September 29, 2025

Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: ioannis.papadopoulos@wias-berlin.de

No. 3219
Berlin 2025



2020 Mathematics Subject Classification. 65K10, 65N30, 74F99, 90C26.

Key words and phrases. Topology optimization, linear elasticity, nonconvex variational problem, multiple local minimizers, finite element method.

This work was completed with the support of the EPSRC grant EP/T022132/1 Spectral element methods for fractional differential equations, with applications in applied analysis and medical imaging, the Leverhulme Trust Research Project Grant RPG-2019-144 Constructive approximation theory on and inside algebraic curves and surfaces and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – The Berlin Mathematics Research Center MATH + (EXC-2046/1, project ID: 390685689).

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: <http://www.wias-berlin.de/>

Numerical analysis of the SIMP model for the topology optimization problem of minimizing compliance in linear elasticity

Ioannis P. A. Papadopoulos

Abstract

We study the finite element approximation of the solid isotropic material with penalization method (SIMP) for the topology optimization problem of minimizing the compliance of a linearly elastic structure. To ensure the existence of a local minimizer to the infinite-dimensional problem, we consider two popular regularization methods: $W^{1,p}$ -type penalty methods and density filtering. Previous results prove weak(-*) convergence in the space of the material distribution to a local minimizer of the infinite-dimensional problem. Notably, convergence was not guaranteed to *all* the isolated local minimizers. In this work, we show that, for every isolated local or global minimizer, there exists a sequence of finite element local minimizers that strongly converges to the minimizer in the appropriate space. As a by-product, this ensures that there exists a sequence of unfiltered discretized material distributions that does not exhibit checkerboarding.

1 Introduction

Topology optimization is an important mathematical technique extensively used in the initial stages of engineering design. The goal is to find the optimal design of a structure or device that minimizes an objective functional. In this work, we focus on the numerical analysis of the topology optimization problem of minimizing the compliance of linearly elastic materials. This is perhaps the most studied topology optimization problem, with roots that go back to the original works by Bendsøe and Kikuchi [9]. The topology optimization of elastic structures and its extensions have been used in numerous applications, for instance airplane wing design [1], cantilevers, and beams [8].

We consider the density formulation via the solid isotropic material with penalization method (SIMP), one of the most popular models for such problems. Here the optimal layout of elastic material is encoded by a function, ρ , known as the material distribution. Regions where $\rho = 1$ a.e. indicate the presence of elastic material, whereas where $\rho = 0$ a.e. are areas where material is absent, as modelled by a material with negligible stiffness. The density approach results in an infinite-dimensional and nonconvex optimization problem with PDE, box, and inequality constraints. In order to obtain a well-posed problem, the infinite-dimensional model is often coupled with a *regularization method* [8]. Irrespective of the choice of the regularization method, even with fixed model parameters, there may exist multiple (local) minimizers to the same density formulated problem, e.g. [35, Sec. 4.5 & 4.6].

Due to the nonlinear nature of the optimization problem, closed-form expressions for the local minimizers are difficult to find. In practice, the full optimization problem is discretized along with the elastic displacement \mathbf{u} and the material distribution ρ . In a discretize-then-optimize approach, the objective functional as well as the PDE, box, and inequality constraints are also discretized and one utilizes a finite-dimensional optimization algorithm to discover discrete local minima. Otherwise, in an optimize-then-discretize approach, one derives the first-order optimality conditions of the infinite-dimensional

optimization problem and discretizes those, utilizing solvers to find discrete solutions. In either case, the finite element (FE) method is the most common choice for discretizing the SIMP optimization problem. Despite its abundant use, proofs that local FE minimizers (rigorously defined in Definition 3.1) converge to the local minimizers of the original infinite-dimensional problem, in a suitably strong sense, is often missing in the literature. Different classes of regularization methods require different assumptions and techniques for their analysis. Another difficulty is caused by the nonconvexity of the problem (in the general case). Literature that does prove convergence typically shows that a subsequence of the local FE minimizers converge to a local infinite-dimensional minimizer of the original problem. However, since the nonconvexity provides multiple candidates for the limits of these sequences, there is an inherent ambiguity in these statements. In particular, these arguments do not guarantee that *every* isolated minimizer can be arbitrarily closely approximated by a sequence of local FE minimizers.

In the context of the topology optimization of Stokes flow [11], the issue of approximating multiple minimizers with the FE method was recently resolved [35, 36, 33, 34]. It was shown that, for every isolated local or global minimizer of the problem, there exists a sequence of FE solutions to the first-order optimality conditions that strongly converges to the infinite-dimensional minimizer as the mesh size tends to zero. We endeavour to prove a similar result here for any conforming FE spaces with an approximation property. The different physics (Stokes vs. linear elasticity) and the additional analysis of the regularization method are the two major obstacles that prevent the analysis in [36, 33] from directly translating to the SIMP model. To reiterate, it is the discretization of the SIMP model for elasticity that is analyzed in this paper.

We consider two different regularization methods in this work: $W^{1,p}$ -type penalty methods and density filtering (both described in Section 2). Recall that ρ and \mathbf{u} denote the material distribution and displacement of the elastic structure, respectively. Let h denote the mesh size. Our goal is to show that for every isolated local or global infinite-dimensional minimizer of the SIMP model, there exists a sequence of local FE minimizers that converges to the infinite-dimensional minimizer. In particular we show that:

- With a $W^{1,p}$ -type penalty one has that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$, $\rho_h \rightarrow \rho$ strongly in $W^{1,p}(\Omega)$, and $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$.
- Instead if one uses a (nonlinear) density filter, then we show that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$ and $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$. Moreover, we prove that $\tilde{\rho}_h \rightarrow \tilde{\rho}$ strongly in $W^{1,q}(\Omega)$, provided $\tilde{\rho}_h \in W^{1,q}(\Omega)$, where $\tilde{\rho}$ and $\tilde{\rho}_h$ denote the filtered material distribution and the discretized filtered material distribution, respectively.

All the stated convergence results are novel, particularly for nonlinear density filters [47, 26].

The main results of this work are stated in Theorems 3.1–3.3. The SIMP model, $W^{1,p}$ -type penalty methods, and density filtering are introduced in Section 2, Section 2.1, and Section 2.2, respectively. We discuss reformulating the generalized linear elasticity PDE constraint in variational form in Section 2.3, detail existence theorems in Section 2.4, and define isolated minimizers in Section 2.5. We provide a literature review of previous FE results in Section 2.6 and discuss the implications of the convergence results in Section 2.7. We establish the FE discretization in Section 3 where we discuss the assumptions of the discretization. In Section 3.4 we give useful auxiliary results known in the literature. Their proofs are provided in Section A for completeness. In Section 3.5, we focus on $W^{1,p}$ -type penalty methods and prove Theorem 3.1. We consider density filtering in Section 3.6 and prove Theorems 3.2 and 3.3. Finally, in Section 4, we give our conclusions and detail potential future extensions.

2 The SIMP model

Consider a bounded and open Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$. Let $W^{s,p}(\Omega)$, $s \in (0, \infty)$, $p \in [1, \infty]$, and $L^q(\Omega)$, $q \in [1, \infty]$, denote the standard Sobolev and Lebesgue spaces, respectively [2]. Define $H^s(\Omega) := W^{s,2}(\Omega)$. Let $\Gamma \subseteq \partial\Omega$ be a subset of the boundary with nonzero Hausdorff measure $\mathcal{H}^{d-1}(\Gamma) > 0$. Then we define:

$$H_\Gamma^1(\Omega)^d := \{\mathbf{v} \in H^1(\Omega)^d : \mathbf{v}|_\Gamma = \mathbf{0}\}, \quad (2.1)$$

where $|_\Gamma$ is the standard boundary trace operator $|_\Gamma : W^{1,p}(\Omega) \rightarrow W^{1-\frac{1}{p},p}(\partial\Omega)$ [24].

The topology optimization problem is to find $\mathbf{u} : \Omega \rightarrow \mathbb{R}^d$, $\mathbf{u} \in H^1(\Omega)^d$ and $\rho : \Omega \rightarrow \mathbb{R}$, $\rho \in L^\infty(\Omega)$, that minimize

$$J(\mathbf{u}, \rho) := l(\mathbf{u}) + \mathcal{R}(\rho) \quad (2.2)$$

where

$$l(\mathbf{u}) := \int_\Omega \mathbf{f} \cdot \mathbf{u} \, dx + \int_{\Gamma_N^t} g^t \mathbf{u}_t \, ds + \int_{\Gamma_N^n} g^n \mathbf{u}_n \, ds, \quad (2.3)$$

subject to the generalized linear elasticity PDE constraint, cf. [3, Sec. 2] and [7],

$$\begin{cases} -\operatorname{div}(\mathbf{S}) = \mathbf{f} & \text{in } \Omega, \\ \mathbf{S} = k(F(\rho)) [2\mu \mathbf{D}(\mathbf{u}) + \lambda \operatorname{div}(\mathbf{u}) \mathbf{I}] & \text{in } \Omega, \\ \mathbf{u}_t = 0 & \text{on } \Gamma_D^t, \\ \mathbf{u}_n = 0 & \text{on } \Gamma_D^n, \\ (\mathbf{S} \hat{\mathbf{n}})_t = g^t & \text{on } \Gamma_N^t, \\ (\mathbf{S} \hat{\mathbf{n}})_n = g^n & \text{on } \Gamma_N^n, \end{cases} \quad (2.4)$$

as well as the following box and inequality constraints on $\rho \in L^\infty(\Omega)$:

$$0 \leq \rho \leq 1 \text{ a.e. in } \Omega \quad \text{and} \quad \int_\Omega F(\rho) \, dx \leq \gamma |\Omega|. \quad (2.5)$$

The state \mathbf{u} denotes the elastic displacement of the structure and \mathbf{S} denotes the stress tensor. The body force $\mathbf{f} \in L^2(\Omega)^d$ and traction forces $g^t \in L^2(\Gamma_N^t)$ and $g^n \in L^2(\Gamma_N^n)$ are known. Here $\hat{\mathbf{n}}$ denotes the unit outward normal, $\mathbf{v}_n := \mathbf{v}|_{\partial\Omega} \cdot \hat{\mathbf{n}}$, and $\mathbf{v}_t := \mathbf{v}|_{\partial\Omega} - \mathbf{v}_n \hat{\mathbf{n}}$, whenever \mathbf{v}_n and \mathbf{v}_t are well-defined. $\Gamma_N^n, \Gamma_D^n \subset \partial\Omega$ (respectively $\Gamma_N^t, \Gamma_D^t \subset \partial\Omega$) are known disjoint boundaries on $\partial\Omega$ such that $\Gamma_N^n \cup \Gamma_D^n = \partial\Omega$ (respectively $\Gamma_N^t \cup \Gamma_D^t = \partial\Omega$). μ and λ are the Lamé coefficients, $\operatorname{tr}(\cdot)$ is the matrix-trace operator, \mathbf{I} is the $d \times d$ identity matrix, and

$$\mathbf{D}(\mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^\top), \quad k(\rho) = \epsilon_{\text{SIMP}} + (1 - \epsilon_{\text{SIMP}})\rho^{p_s},$$

where $0 < \epsilon_{\text{SIMP}} \ll 1$ and $p_s \geq 1$. A typical “magic” choice resulting in crisp material distributions is $p_s = 3$ [42, Sec. 3]. The constant $\gamma \in (0, 1)$ is the fraction of the total volume of the domain that the material distribution can occupy. Here, $|\cdot|$ denotes the Lebesgue measure of a set. $\mathcal{R} : W^{1,p}(\Omega) \rightarrow \mathbb{R}$ is a functional that models a $W^{1,p}$ -type penalty and $F : \{\eta \in L^\infty(\Omega) : 0 \leq \eta \leq 1 \text{ a.e.}\} \rightarrow W^{1,\infty}(\Omega)$ is a function that encodes density filtering.

Remark 2.1. We have provided the linear elasticity PDE constraint in (2.4) with very general boundary conditions. One could instead reduce the boundary conditions to two boundaries Γ_D and Γ_N , $\Gamma_D \cup \Gamma_N = \partial\Omega$, such that

$$\begin{cases} \mathbf{u}|_{\Gamma_D} = \mathbf{0} & \text{on } \Gamma_D, \\ (\mathbf{S}\hat{\mathbf{n}})|_{\Gamma_N} = \mathbf{g} & \text{on } \Gamma_N, \end{cases} \quad (2.6)$$

for some $\mathbf{g} \in L^2(\Gamma_N)^d$. All the subsequent results in this work still hold. However, we note that these simpler boundary conditions are insufficient to describe the physics for the MBB beam problem in Example 2.1 and many other common SIMP models.

For a physical interpretation of the SIMP model we refer to Bendsøe and Sigmund [8, Ch. 1]. When $\rho \approx 1$, then $k(\rho) \approx 1$ indicating the presence of material, whereas where $\rho \approx 0$, then $k(\rho) \approx \epsilon_{\text{SIMP}}$ indicating void. The optimal material distribution that minimizes the compliance is $\rho = 1$ a.e. in Ω . However, due to the volume constraint restriction and the limitation $\rho \leq 1$ a.e., regions where $\rho < 1$ a.e. become necessary. The SIMP model penalizes intermediate regions where $0 < \rho < 1$ a.e. because intermediate regions have a stiffness proportional to $k(F(\rho))$ for a volume that is proportional to $F(\rho)$. This makes such regions inefficient at minimizing compliance under a volume constraint. Thus, a local minimizer will favour larger regions where either $\rho = 0$ or $\rho = 1$ a.e., while significantly diminishing regions where $0 < \rho < 1$ a.e. as $p_s \rightarrow \infty$. Hence, by raising ρ to the power of p_s , values of ρ between 0 and 1 are penalized.

The optimization problem (2.2)–(2.5) is normally stated with $\mathcal{R}(\rho) = 0$ and $F(\rho) = \rho$. With these choices (2.2)–(2.5) is ill-posed in general, i.e. it does not have minimizers in the continuous setting [10]. In particular any attempt at a proof of existence, via a direct method in the calculus of variations [18], fails. Naive attempts at finding minimizers often yield checkerboard patterns of ρ (as discussed in Section 2.7.2). Checkerboarding can be avoided, without modifications to the discretized optimization problem induced by (2.2)–(2.5), by particular choices of FE spaces for \mathbf{u} and ρ . However, the discretized material distribution will still be mesh dependent i.e. as the mesh is refined, the beams of the discretized material distribution will become ever thinner, leading to the homogenized limit. There are several schemes employed by the topology optimization community to obtain physically reasonable local minimizers known as regularization methods [8]. An excellent review is given by Sigmund and Petersson [43] as well as Borrvall [10]. We consider two regularization methods in this work: $W^{1,p}$ -type penalty methods and density filtering.

2.1 $W^{1,p}$ -type penalty methods

We use the functional \mathcal{R} to model $W^{1,p}$ -type penalties of ρ which forces one to seek a local minimizer with the regularity $\rho \in L^\infty(\Omega) \cap W^{1,p}(\Omega)$.

Definition 2.1 ($W^{1,p}$ -type penalty methods). For $p \in (1, \infty)$, we say that the regularization method is a $W^{1,p}$ -type penalty method if $\mathcal{R}(\rho) = \frac{\epsilon}{p} \|\nabla \rho\|_{L^p(\Omega)}^p + m(\rho)$, where $m(\rho)$ is continuous in the strong topology of $L^p(\Omega)$ and models lower order terms not involving $\nabla \rho$. For instance $m(\rho) = \int_\Omega \rho(1 - \rho) dx$ (if $p \geq 2$). We refer to \mathcal{R} as a $W^{1,p}$ -type penalty functional.

We define the material distribution space \mathcal{H}_γ^p as

$$\mathcal{H}_\gamma^p := \{\rho \in L^\infty(\Omega) \cap W^{1,p}(\Omega) : 0 \leq \rho \leq 1 \text{ a.e., } \int_\Omega \rho dx \leq \gamma|\Omega|\}. \quad (2.7)$$

Examples of choices for \mathcal{R} include, for $\epsilon > 0, \beta > 0$,

$$\mathcal{R}(\rho) = \frac{\epsilon}{p} \|\nabla \rho\|_{L^p(\Omega)}^p, \quad p \in (1, \infty), \quad (W^{1,p}\text{-penalty}), \quad (2.8)$$

$$\mathcal{R}(\rho) = \frac{\beta\epsilon}{2} \|\nabla \rho\|_{L^2(\Omega)}^2 + \frac{\beta}{2\epsilon} \int_{\Omega} \rho(1 - \rho) \, dx \quad (\text{Ginzburg–Landau penalty}). \quad (2.9)$$

If $p = 1$ or $p = \infty$, one recovers total variation regularization [38] and a Lipschitz continuity penalty method [40], respectively, which are beyond the scope of this work.

Often the $W^{1,p}$ -type penalty is added as an additional constraint rather than at the level of the objective functional. For instance, $\mathcal{R}(\rho) = 0$ but one adds the constraint $\|\nabla \rho\|_{L^p(\Omega)} \leq M < \infty$ for a user-chosen $M > 0$. Although the implementation differs, the analysis is similar and we expect many of the results discussed in this work to continue to hold.

Example 2.1 (MBB beam). *To illustrate the SIMP model with a $W^{1,p}$ -type penalty, we provide an example of an MBB beam. Consider the domain $\Omega = (0, 3) \times (0, 1)$. Consider the boundaries:*

$$\Gamma_{D_1}^n = \{x = 0\}, \quad \Gamma_{D_2}^n = \{y = 0, 2.9 \leq x \leq 3\}, \quad \Gamma_{N_1}^n = \{y = 1, 0 \leq x \leq 0.1\}. \quad (2.10)$$

The Dirichlet and Neumann boundaries are

$$\Gamma_D^n = \Gamma_{D_1}^n \cup \Gamma_{D_2}^n, \quad \Gamma_D^t = \emptyset, \quad \Gamma_N^n = \partial\Omega \setminus \Gamma_D^n, \quad \Gamma_N^t = \partial\Omega. \quad (2.11)$$

The body and traction forces are $\mathbf{f} = \mathbf{0}$, $g^t = 0$, and

$$g^n(x, y) = \begin{cases} -10 & \text{if } (x, y) \in \Gamma_{N_1}^n, \\ 0 & \text{otherwise.} \end{cases} \quad (2.12)$$

These conditions describe a half-beam that is fixed horizontally on the y -axis and fixed vertically at its bottom right corner on the x -axis. There is a boundary force pushing vertically downwards at the top left corner, which represents the middle of the beam when the half-beam is mirrored. This is visualised in Figure 1.

Consider the volume fraction $\gamma = 0.535$, the Lamé coefficients $\mu = 75.38$ and $\lambda = 64.62$, the SIMP parameters, $\epsilon_{\text{SIMP}} = 10^{-5}$ and $p_s = 3$. If one uses a Ginzburg–Landau penalty, it is possible to find two local minima as depicted in Figure 2.

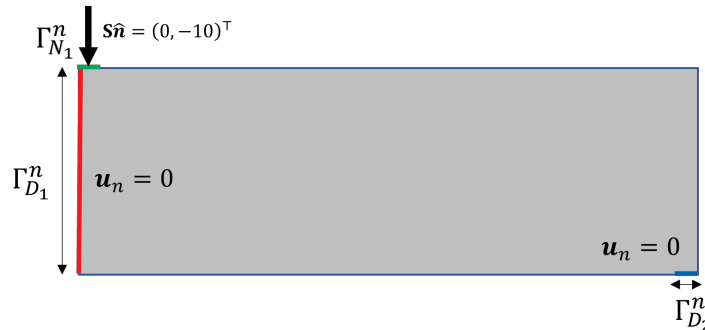


Figure 1: Setup of the MBB beam. The remaining unlabeled boundary conditions are $(\mathbf{S}\hat{\mathbf{n}})_t = 0$ on $\Gamma_{D_1}^n \cup \Gamma_{D_2}^n$ and $\mathbf{S}\hat{\mathbf{n}} = (0, 0)^\top$ on $\partial\Omega \setminus (\Gamma_{D_1}^n \cup \Gamma_{D_2}^n \cup \Gamma_{N_1}^n)$.

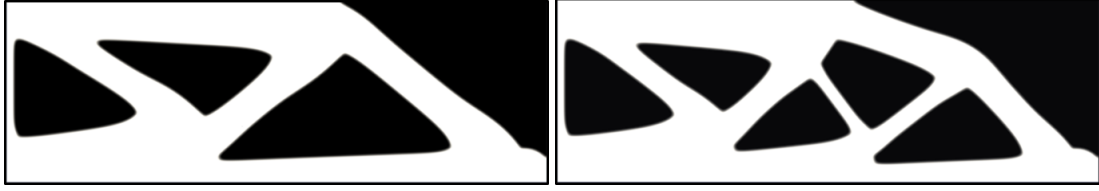


Figure 2: The material distribution of two (local) minima of the MBB beam problem [35, Sec. 4.6]. In white regions $\rho = 1$ whereas in black regions $\rho = 0$. The Ginzburg–Landau parameters, as defined in (2.9), are $\epsilon = 1.90 \times 10^{-2}$, $\beta = 9 \times 10^{-3}$. Moreover, $\gamma = 0.535$, $\epsilon_{\text{SIMP}} = 10^{-5}$, $p_s = 3$, and the Lamé coefficients are $\mu = 75.38$ and $\lambda = 64.62$.

2.2 Density filtering

The function $F(\cdot)$ models a *density filtering* of the problem [13, 26, 47]. Filtering averages ρ in small neighbourhoods at every point. The first application in the context of topology optimization of compliance is attributed to Bruns and Tortorelli [16]; existence and FE convergence was first shown by Bourdin [13]. We define the unfiltered material distribution space \mathcal{H}_γ^F as

$$\mathcal{H}_\gamma^F := \{\rho \in L^\infty(\Omega) : 0 \leq \rho \leq 1 \text{ a.e.}, \int_\Omega F(\rho) \, dx \leq \gamma |\Omega|\} \quad (2.13)$$

and the filtered material distribution space $\tilde{\mathcal{H}}_\gamma^F$

$$\tilde{\mathcal{H}}_\gamma^F := \{\tilde{\eta} \in W^{1,\infty}(\Omega) : \exists \eta \in \mathcal{H}_\gamma^F \text{ s.t. } \tilde{\eta} = F(\eta)\}. \quad (2.14)$$

In the context of density filtering, the material distribution, prior to filtering, $\rho \in \mathcal{H}_\gamma^F$, is known as the *unfiltered* material distribution and $\tilde{\rho} = F(\rho) \in \tilde{\mathcal{H}}_\gamma^F$ is known as the *filtered* material distribution.

We make the following assumptions on the density filters:

- (F1) For any $\eta \in \mathcal{H}_\gamma^F$, then $\|F(\eta)\|_{W^{1,\infty}(\Omega)} \leq C < \infty$ with C independent of η ;
- (F2) For any $\eta \in \mathcal{H}_\gamma^F$, we have $0 \leq F(\eta) \leq 1$ a.e. in Ω ;
- (F3) Suppose there exists an L^s -bounded sequence, $s \in [1, \infty]$, such that $\eta_j \rightharpoonup \eta$ weakly in $L^s(\Omega)$ ($\eta_j \xrightarrow{*} \eta$ weakly-* in $L^\infty(\Omega)$ if $s = \infty$). Then there exists a subsequence (up to relabelling) such that $F(\eta_j) \rightarrow F(\eta)$ strongly in $L^s(\Omega)$.

To prove convergence of the filtered material distribution in $W^{1,s}(\Omega)$, for some $s \in (1, \infty)$, we require the following assumption:

- (F4) The space of filtered material distributions, $\tilde{\mathcal{H}}_\gamma^F$, is a norm-closed and convex subset of $W^{1,q}(\Omega)$ for some $q \in (1, \infty)$.

One class of density filters are known as linear density filters.

Definition 2.2. A density filter F is a linear density filter if the material distribution ρ is convolved with a function that is not dependent on ρ , i.e.

$$F(\rho)(x) = \int_\Omega f(x-y)\rho(y) \, dy, \quad (2.15)$$

where $f \in W^{1,\infty}(\mathbb{R}^d)$, $f \geq 0$ a.e. and $\|f\|_{L^1(\mathbb{R}^d)} = 1$.

The following proposition is a known result and we give a proof in Section A for completeness.

Proposition 2.1. *Suppose that F is a linear density filter. Then, (F1)–(F4) are satisfied.*

We do not analyze *sensitivity filtering* in this work [41]. Sensitivity filtering is cheap and effective and, under suitable conditions, may be understood as the identification of the Fréchet derivative of the objective functional with a gradient in a suitably chosen Hilbert space [5, Sec. 5.2], see also [21, 30] for other studies. However, it does not fit under our framework and thus it is not studied here.

2.3 Variational form

For a FE discretization of the linear elasticity PDE constraint, we must recast (2.4) in variational form. The most common choice, in the topology optimization literature, is the primal formulation and all known FE convergence results for the SIMP model only deal with this case. The primal formulation substitutes the explicit formula for the stress tensor \mathbf{S} into the first equation of (2.4). Thus the two PDE constraints reduce to one PDE constraint in terms of \mathbf{u} and ρ . This PDE constraint is then multiplied by a test function \mathbf{v} and an integration by parts is performed. The advantage of the primal formulation lies in its simplicity and cost-effectiveness, attributed to the reduced computational load from not discretizing the stress tensor, thus saving degrees of freedom.

By choosing not to eliminate the stress tensor, one may form a mixed formulation whereby one discretizes and solves for the stress tensor and elastic displacement independently. There are multiple mixed formulations found in the literature and we refer the interested reader to [14, Ch. VI.3]. In this work we solely consider a primal formulation and leave the analysis of different mixed formulations for future work.

Let $H_D^1(\Omega)^d := \{\mathbf{v} \in H^1(\Omega)^d : \mathbf{v}_t|_{\Gamma_D^t} = 0, \mathbf{v}_n|_{\Gamma_D^n} = 0\}$ and denote the space of rigid body motions by RM^d [7, Sec. 2.2]:

$$\text{RM}^d := \{\mathbf{R}\mathbf{x} + \mathbf{a} : \mathbf{a} \in \mathbb{R}^d, \mathbf{R} \in \text{skew}_d\}, \quad (2.16)$$

where skew_d denotes the vector space of constant skew-symmetric $d \times d$ matrices. For instance, $\text{RM}^2 = \text{span}\{[1, 0]^\top, [0, 1]^\top, [-y, x]^\top\}$. Define the space of displacements by \mathbf{V}_D where

$$\mathbf{V}_D := \{\mathbf{v} \in H_D^1(\Omega)^d : (\nabla \mathbf{v}, \mathbf{R})_{L^2(\Omega)} = 0 \forall \mathbf{R} \in \nabla(\text{RM}^d \cap H_D^1(\Omega)^d)\}, \quad (2.17)$$

where, given two matrices of functions $\mathbf{A}, \mathbf{B} \in L^2(\Omega)^{d \times d}$, then $(\mathbf{A}, \mathbf{B})_{L^2(\Omega)}^2 := \sum_{i,j=1}^d (\mathbf{A}_{ij}, \mathbf{B}_{ij})_{L^2(\Omega)}^2$. For instance, given the boundary conditions in (2.6) we have that $H_D^1(\Omega)^d \cap \text{RM}^d = \{\mathbf{0}\}$ and the definition of \mathbf{V}_D in (2.17) reduces to $\mathbf{V}_D = H_D^1(\Omega)^d$. The additional condition $(\nabla \mathbf{v}, \mathbf{R})_{L^2(\Omega)} = 0$ plays a role if the boundary conditions of the elastic problem are not sufficient to uniquely determine the solution.

Recall the definition of $l(\mathbf{v})$ in (2.3) and let \mathcal{H}_γ denote

$$\mathcal{H}_\gamma := \begin{cases} \mathcal{H}_\gamma^p & \text{if } \mathcal{R}(\rho) \neq 0 \text{ and } F(\rho) = \rho, \\ \mathcal{H}_\gamma^F & \text{if } \mathcal{R}(\rho) = 0 \text{ and } F(\rho) \neq \rho, \\ \mathcal{H}_\gamma^p \cap \mathcal{H}_\gamma^F & \text{if } \mathcal{R}(\rho) \neq 0 \text{ and } F(\rho) \neq \rho. \end{cases} \quad (2.18)$$

In the variational form, the primal formulation of the optimization problem reduces to,

$$\min_{(\mathbf{u}, \rho) \in \mathbf{V}_D \times \mathcal{H}_\gamma} J(\mathbf{u}, \rho) \quad \text{subject to} \quad a(\mathbf{u}, \mathbf{v}; \rho) = l(\mathbf{v}) \text{ for all } \mathbf{v} \in \mathbf{V}_D, \quad (\text{SIMP})$$

where

$$a(\mathbf{u}, \mathbf{v}; \rho) := \int_{\Omega} k(F(\rho)) \mathbf{E} \mathbf{u} : \mathbf{E} \mathbf{v} \, dx, \quad (2.19)$$

$$\mathbf{E} \mathbf{u} : \mathbf{E} \mathbf{v} := 2\mu \mathbf{D}(\mathbf{u}) : \mathbf{D}(\mathbf{v}) + \lambda \operatorname{div}(\mathbf{u}) \operatorname{div}(\mathbf{v}). \quad (2.20)$$

Note that ρ only appears in the optimization problem (2.2)–(2.5) in the box and volume constraints as well as the $W^{1,p}$ -type penalty term \mathcal{R} . Moreover, by assumption (F2), $0 \leq \rho \leq 1$ a.e. implies that $0 \leq F(\rho) \leq 1$. If $\mathcal{R}(\rho) = 0$ (no $W^{1,p}$ -type penalty, but a density filter is applied) then (SIMP) may be recast as

$$\min_{(\mathbf{u}, \tilde{\rho}) \in \mathbf{V}_D \times \tilde{\mathcal{H}}_{\gamma}^F} \tilde{J}(\mathbf{u}, \tilde{\rho}) \quad \text{subject to} \quad \tilde{a}(\mathbf{u}, \mathbf{v}; \tilde{\rho}) = l(\mathbf{v}) \text{ for all } \mathbf{v} \in \mathbf{V}_D, \quad (\text{F-SIMP})$$

where $\tilde{J}(\mathbf{u}, \tilde{\rho}) = J(\mathbf{u}, \rho)$ and $\tilde{a}(\mathbf{u}, \mathbf{v}; \tilde{\rho}) = a(\mathbf{u}, \mathbf{v}; \rho)$.

2.4 Existence of local minimizers

In this subsection we provide existence theorems for (SIMP).

Proposition 2.2 (Existence & uniqueness with a fixed ρ). *Suppose that $\rho \in \mathcal{H}_{\gamma}$ is fixed, $l(\mathbf{v}) = 0$ for all $\mathbf{v} \in H_D^1(\Omega)^d \cap \operatorname{RM}^d$, and (F2) holds. Then, there exists a unique $\mathbf{u} \in \mathbf{V}_D$ that satisfies the PDE constraint in (SIMP).*

Proof. The result is standard and follows as a direct consequence of Korn's inequality with tangential or normal boundary conditions [7, Th. 11] and the Lax–Milgram Theorem [20, Ch. 6.2, Th. 1]. See also [15, Ch. 11]. \square

An essential result utilized in this work is the existence of a local minimizer with either a $W^{1,p}$ -penalty method or a density filter.

Theorem 2.1 (Existence of a local minimizer to (SIMP)). *Suppose that the conditions of Proposition 2.2 hold. Consider the two cases:*

(Case 1). *If a $W^{1,p}$ -type penalty method is used, then there exists a local minimizer $(\mathbf{u}, \rho) \in \mathbf{V}_D \times \mathcal{H}_{\gamma}^p$ of (SIMP).*

(Case 2). *If a density filter is used and (F1)–(F3) hold, then there exists a local minimizer $(\mathbf{u}, \rho) \in \mathbf{V}_D \times \mathcal{H}_{\gamma}^F$ of (SIMP).*

Proof. The proof of (Case 1) may be found in [8, Ch. 5.2.2] and the proof of (Case 2) is found in [13, Sec. 3]. \square

2.5 Isolated local minimizers

A key assumption for the FE convergence results in this work is that the local and global minimizers in the limit are *isolated*. Consider an element $u \in U$ where U is a Banach space. We denote the closed ball of radius r by $B_{r,U}(u)$:

$$B_{r,U}(u) := \{v \in U : \|u - v\|_U \leq r\}. \quad (2.21)$$

This is extended to tuples of functions as follows: consider $u_i \in U_i$, for $i = 1, \dots, n$, where U_i are Banach spaces. We define $B_{r, U_1 \times \dots \times U_n}(u_1, \dots, u_n)$ as

$$\begin{aligned} & B_{r, U_1 \times \dots \times U_n}(u_1, \dots, u_n) \\ & := \{(v_1, \dots, v_n) \in U_1 \times \dots \times U_n : \sum_{i=1}^n \|v_i - u_i\|_{U_i} \leq r\}. \end{aligned} \quad (2.22)$$

Definition 2.3 (Isolated local minimizer of (SIMP)). *We say that a local minimizer $(\mathbf{u}, \rho) \in \mathbf{V}_D \times \mathcal{H}_\gamma$ of (SIMP) is isolated if there exists an $r > 0$ such that (\mathbf{u}, ρ) is the unique local minimizer in $B_{r, H^1(\Omega) \times Y}(\mathbf{u}, \rho) \cap (\mathbf{V}_D \times \mathcal{H}_\gamma)$. Here*

$$Y = \begin{cases} W^{1,p}(\Omega) & \text{if a } W^{1,p}\text{-type penalty is used,} \\ L^2(\Omega) & \text{otherwise.} \end{cases} \quad (2.23)$$

We define the radius of the basin of attraction as the largest value r such that (\mathbf{u}, ρ) is the unique minimizer. We note that

$$\begin{aligned} & (\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_\gamma \cap B_{r/2, Y}(\rho)) \\ & \subset B_{r, H^1(\Omega) \times Y}(\mathbf{u}, \rho) \cap (\mathbf{V}_D \times \mathcal{H}_\gamma) \end{aligned}$$

and hence (\mathbf{u}, ρ) is also the unique minimizer in $(\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_\gamma \cap B_{r/2, Y}(\rho))$.

Remark 2.2. We make the assumption that $\rho \in \mathcal{H}_\gamma$ is isolated with respect to the Y -norm (as opposed to the L^∞ -norm). This is a stronger isolation assumption as discussed in [36, Rem. 7]. To our knowledge, the validity of the Y -isolation in SIMP optimization problems, regularized by a $W^{1,p}$ -type penalty method or density filtering, is an open problem. However, qualitatively and numerically, it appears to hold in many practical problems e.g. in Example 2.1. We make this stronger isolation assumption as continuous functions are not dense in $L^\infty(\Omega)$, but are dense in Y . This has implications in the assumptions for the FE spaces (in particular (D2)).

Remark 2.3 (Violating the isolation assumption). The results in Theorems 3.1–3.3 heavily rely on the isolation assumption in order to deduce the limit of a FE minimizing sequence. If the isolation assumption was lifted, the results, at least as derived in this work, would no longer be valid. Compliance problems, in full generality, can support very interesting solution landscapes, e.g. an infinite number of global minimizers, no classical minimizers or local minimizers that are not necessarily isolated [4, Ch. 4.1.5]. A study of such local minimizers is beyond the scope of this work.

2.6 Literature review

The focus of this subsection is to conduct a literature review and contextualize our aims. We define a local FE minimizer in Definition 3.1. Our goal is to answer the following open problems in the numerical analysis:

- (P1) With the exception of cases where a unique global minimizer exists, prior results do not explicitly identify the specific local minimizer towards which the sequence of local FE minimizers converges. This ambiguity arises due to the nonconvex nature of the problem, which results in multiple potential candidates for the limits. We show that for every isolated local minimizer, there exists a sequence of local FE minimizers that strongly converges to it;

- (P2) ($W^{1,p}$ -type penalty). The strongest results currently found in the literature show that $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$, for any $s \in [1, \infty)$. We prove the stronger result: $\rho_h \rightarrow \rho$ strongly in $W^{1,p}(\Omega)$;
- (P3) (Density filter). Previous results showed that the unfiltered material distribution $\rho_h \xrightarrow{*} \rho$ weakly- $*$ in $L^\infty(\Omega)$. We improve this result to $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$, for any $s \in [1, \infty)$;
- (P4) (Density filter). It is known that the discretized filtered material distribution $F_h(\rho_h) \rightarrow F(\rho)$ uniformly on Ω (F and F_h are defined in Section 2 and Section 3, respectively). We prove that there exists a sequence such that $F_h(\rho_h) \rightarrow F(\rho)$ strongly in $W^{1,q}(\Omega)$ for any $q \in [1, \infty)$.

In both Theorems 3.1 and 3.2 we tackle the issue of multiple local minimizers (P1). In Theorems 3.1, 3.2, and 3.3 we prove (P2), (P3), and (P4), respectively.

Below we give a (perhaps non-exhaustive) list of the known FE convergence results. We emphasize that in all of the following results convergence is *not* guaranteed to all isolated local minimizers (unless there only exists one local minimizer).

SIMP coupled with a $W^{1,p}$ -type penalty method.

- Petersson and Sigmund analyzed the primal formulation in the context of slope constraints [40], which have a similar flavour to a $W^{1,\infty}$ -type penalty [10, Sec. 6.3]. Here the admissible set of material distributions is reduced to Lipschitz continuous functions and a pointwise bound is placed on $\partial_{x_i}\rho$ for $i = 1, 2$. They showed that a $(CG_1)^2$ (continuous piecewise linear) discretization for \mathbf{u} and a DG_0 (piecewise constant) discretization for ρ converges to a local minimum of the problem. In particular $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^2$, $\rho_h \rightarrow \rho$ uniformly in Ω and, therefore, $\rho_h \rightarrow \rho$ strongly in $L^\infty(\Omega)$.
- Petersson [38] investigated a regularization in the flavour of a $W^{1,1}(\Omega)$ -penalty together with penalization terms such as $\epsilon^{-1} \int_\Omega \rho(1 - \rho) \, dx$. He considered a $(CG_1)^2$ discretization for \mathbf{u} and a DG_0 discretization for ρ . He showed that a sequence of local FE minimizers of the discretized (SIMP), (\mathbf{u}_h, ρ_h) , converges to a local minimizer (\mathbf{u}, ρ) such that $(\mathbf{u}_h, \rho_h) \rightarrow (\mathbf{u}, \rho)$ strongly in $H^1(\Omega)^d \times L^s(\Omega)$, for any $s \in [1, \infty)$.
- Talischi and Paulino studied a $W^{1,2}$ -penalty [45]. They used a $(CG_k)^2$ and (CG_k) , $k \geq 1$, for the discretization of \mathbf{u} and ρ , respectively. However, they allow for ρ to be projected onto the space DG_0 in the PDE constraint. They showed that there exists a sequence of local FE minimizers, (\mathbf{u}_h, ρ_h) converging to a local minimizer (\mathbf{u}, ρ) such that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^2$ and $\rho_h \rightharpoonup \rho$ weakly in $H^1(\Omega)$.

SIMP with density filtering.

- In the context of linear density filtering, Bourdin showed that, for a $(CG_1)^2$ discretization for \mathbf{u} and a DG_0 discretization for ρ , the sequence of minimizers (\mathbf{u}_h, ρ_h) of the discretized (SIMP) converges to a local minimizer (\mathbf{u}, ρ) such that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$, $F_h(\rho_h) \rightarrow F(\rho)$ uniformly on Ω , and $\rho_h \rightharpoonup \rho$ weakly in $L^s(\Omega)$ for any $s \in [1, \infty)$ [13].

Other results.

- One of the first FE convergence results for the topology optimization of a linearly elastic material was shown by Petersson and Haslinger [39]. They showed that a sequence of local FE minimizers, (\mathbf{u}_h, ρ_h) , of a problem similar to the discretization of (SIMP), converge to a local minimizer (\mathbf{u}, ρ) . In particular $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$ and $\rho_h \xrightarrow{*} \rho$ weakly- $*$ in $L^\infty(\Omega)$.

- The problem reduces to the problem of the optimal variable thickness of sheets when $p_s = 1$. Petersson analyzed such problems [37] and showed that, under a regularity and biaxiality assumption on the optimal stress field, the minimizer ρ is unique. Moreover, for any conforming discretization of \mathbf{u} and a DG_0 discretization for ρ , the minimizing sequence of unique FE minimizers converges to the unique minimizer (\mathbf{u}, ρ) such that $(\mathbf{u}_h, \rho_h) \rightarrow (\mathbf{u}, \rho)$ strongly in $H^1(\Omega)^2 \times L^s(\Omega)$ for any $s \in [1, \infty)$.
- Borrvall and Petersson [12] considered a problem with no filtering, $F(\rho) = \rho$, but with a regularization term $\mathcal{R}(\rho)$ reminiscent of a density filter. For example $\mathcal{R}(\rho) = g \circ \int_{\Omega} f(x - y)\rho(y) \, dy$ where $g(\tilde{\rho}) = \int_{\Omega} (\tilde{\rho} - \underline{\rho})(1 - \tilde{\rho}) \, dx$, $\underline{\rho} > 0$, and \circ denotes the composition operator. They showed that, for any conforming discretization of \mathbf{u} and a DG_0 discretization for ρ , there exists a sequence of local FE minimizers ρ_h such that $\rho_h \xrightarrow{*} \rho$ weakly-* in $L^\infty(\Omega)$ where ρ is a local minimizer of the problem. This is subsequently strengthened to $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega_{\rho,b})$, for any $s \in [1, \infty)$, where $\Omega_{\rho,b} := \{\mathbf{x} \in \Omega : \rho(\mathbf{x}) = 0 \text{ or } \rho(\mathbf{x}) = 1\}$, if such a set $\Omega_{\rho,b}$ exists and is measurable.
- Greifstein and Stingl [25] consider a material distribution that is partially density filtered and partially regularized with slope constraints (as defined in [40]). They show that there exists a sequence of FE minimizers that converge to a local minimizer. In particular $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$, the unfiltered material distribution converges weakly-* in $L^\infty(\Omega)$, whereas the filtered and slope constrained material distribution both converge uniformly. For more details on how the hybrid formulation is designed and how \mathbf{u} depends on the material distribution, we refer the reader to [25].

Although the above list might not be exhaustive, all works follow very similar arguments as Petersson and Sigmund [40] or Bourdin [13] for penalty and density filtering techniques, respectively.

2.7 Implications

2.7.1 Multiple local minima

The optimization problem is nonconvex and there may exist multiple local minima to the same optimization problem. In [35, Sec. 4.5 & 4.6], there are examples of cantilevers and MBB beams (see Example 2.1), coupled with a Ginzburg–Landau penalty method, where two local minima are found for each example. All previous results only showed that a sequence of FE minimizers converge to a local minimizer of the infinite-dimensional problem. Our novel results guarantee strong convergence to all the isolated local minimizers. This is of critical importance in practical implementations.

2.7.2 Checkerboard

We use the term *checkerboarding* to describe the phenomenon where the discretized material distribution wildly oscillates in value between neighbouring cells in the mesh and this behaviour persists as the mesh size tends to zero [43, Sec. 2]. Terminating the mesh refinement at any point is problematic as, without careful post-processing, the discretized material distribution suggests designs which are not manufacturable. This may occur in the discretized material distribution of the SIMP model when a low-order discretization is used without a regularization method. The causes of checkerboarding are due to artificial stiffness granted by low-order finite elements [19] which in turn violate a discrete

inf-sup condition of the discretized problem [27]. Suitable regularization methods simultaneously guarantee the existence of a local minimizer to the SIMP model and prevent checkerboard patterns from persisting as $h \rightarrow 0$.

Previous results in literature proved that a density filter ensures the existence of a sequence of unfiltered discretized material distributions such that $\rho_h \xrightarrow{*} \rho$ weakly-* in $L^\infty(\Omega)$ where ρ is an unfiltered local material distribution minimizer of the infinite-dimensional problem. However, weak-* convergence is not sufficient to prevent a checkerboard pattern from forming and persisting. For instance, consider the one-dimensional checkerboard sequence $\phi_n \in L^\infty(0, 1)$, $n \in \mathbb{N}$,

$$\phi_n(x) = \begin{cases} 1 & \text{if } j/n < x < (j + 1/2)/n, \\ -1 & \text{if } (j + 1/2)/n < x < (j + 1)/n, \end{cases} \quad (2.24)$$

where $j \in \{0, 1, \dots, n-1\}$. Then $\phi_n \xrightarrow{*} 0$ weakly-* in $L^\infty(\Omega)$ [20, Sec. 8.7]. However $\phi_n \not\rightarrow 0$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty]$ since $\|\phi_n\|_{L^s(\Omega)} = 1$ for all $n \in \mathbb{N}$. Thus, although weak-* convergence is not sufficient, by proving that there exists a subsequence (up to relabelling) where $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$, for any $s \in [1, \infty)$, we guarantee that a checkerboard cannot persist as $h \rightarrow 0$ in both two and three dimensions [43, Sec. 2]. Although this fact was qualitatively observed in two dimensions [13, Fig. 2 & 3], it was not proven for the general class of density filters and FE methods we consider here. The unfiltered material distribution has sharper boundaries than the filtered material distribution and may be useful for inferring the eventual manufactured design.

3 Finite element approximation

For the remainder of this work we assume that $\Omega \subset \mathbb{R}^d$ is a polygonal domain in two dimensions or a polyhedral Lipschitz domain in three dimensions. Moreover, we assume that there exists a (local or global) isolated minimizer of (SIMP) and fix one such *local or global isolated minimizer* $(\mathbf{u}, \rho) \in \mathbf{V}_D \times \mathcal{H}_\gamma$ of (SIMP).

Consider a family of quasi-uniform, non-degenerate, and simplicial meshes (\mathcal{T}_h) as $h \rightarrow 0$ [15, Def. 4.4.13] such that $\cup_{K \in \mathcal{T}_h} K = \bar{\Omega}$ for every h . We choose conforming FE discretizations $\mathbf{V}_h \subset H^1(\Omega)^d$, $\mathcal{H}_{\gamma,h}^p \subset \mathcal{H}_\gamma^p$, $\mathcal{H}_{\gamma,h}^F \subset \mathcal{H}_\gamma^F$. To clarify, a FE discretization is conforming if the discretized space is a subspace of the infinite-dimensional space that it is discretizing.

In general, it will not be possible to represent the body and traction forces \mathbf{f} , g^n , g^t exactly in the displacement FE space. Hence, for each h , we instead consider \mathbf{f}_h , g_h^n , g_h^t (which can be represented). We define the discretized functional:

$$l_h(\mathbf{u}_h) := \int_{\Omega} \mathbf{f}_h \cdot \mathbf{u}_h \, dx + \int_{\Gamma_N^t} g_h^t(\mathbf{u}_h)_t \, ds + \int_{\Gamma_N^n} g_h^n(\mathbf{u}_h)_n \, ds. \quad (3.1)$$

We henceforth assume that:

$$(D1) \quad \mathbf{f}_h \rightarrow \mathbf{f} \text{ strongly in } L^2(\Omega)^d, \quad g_h^n \rightarrow g^n \text{ strongly in } L^2(\Gamma_N^n), \text{ and } g_h^t \rightarrow g^t \text{ strongly in } L^2(\Gamma_N^t).$$

Moreover,

$$(D2) \quad \text{The FE spaces are dense in their respective function spaces, i.e. for any } \mathbf{v} \in H^1(\Omega)^d, \eta \in \mathcal{H}_\gamma^p,$$

and $\theta \in \mathcal{H}_\gamma^F$,

$$\begin{aligned} \lim_{h \rightarrow 0} \inf_{\mathbf{w}_h \in \mathbf{V}_h} \|\mathbf{v} - \mathbf{w}_h\|_{H^1(\Omega)} &= \lim_{h \rightarrow 0} \inf_{\zeta_h \in \mathcal{H}_{\gamma,h}^F} \|\eta - \zeta_h\|_{W^{1,p}(\Omega)} \\ &= \lim_{h \rightarrow 0} \inf_{\zeta_h \in \mathcal{H}_{\gamma,h}^F} \|\theta - \zeta_h\|_{L^2(\Omega)} = 0. \end{aligned}$$

(D3) $l(\mathbf{v}) = l_h(\mathbf{v}) = 0$ for all $\mathbf{v} \in H_D^1(\Omega)^d \cap \text{RM}^d$.

We denote the restriction of \mathbf{V}_h to the boundary conditions as:

$$\mathbf{V}_{D,h} := \mathbf{V}_h \cap \mathbf{V}_D. \quad (3.2)$$

3.1 FE assumptions on the density filter

Let $\mathcal{P}_k(\mathcal{T}_h) \subset W^{1,\infty}(\Omega)$, $k \geq 0$ denote the space of continuous piecewise polynomials of degree k which are polynomials when restricted to each cell $K \in \mathcal{T}_h$. In general, the filtered discretized material distribution $F(\rho_h) \notin \mathcal{P}_k(\mathcal{T}_h)$. Hence, we are required to project $F(\rho_h)$ into the space $\mathcal{P}_k(\mathcal{T}_h)$. We denote the projection of $F(\rho_h)$ onto $\mathcal{P}_k(\mathcal{T}_h)$ as $F_h(\rho_h) = \Pi_h F(\rho_h) \in \mathcal{P}_k(\mathcal{T}_h)$. Here Π_h denotes the projection operator cf. [15, Def. 2.3.9]. We assume that:

- (F5) For any $\eta \in \mathcal{H}_\gamma^F$, then $\|F_h(\eta)\|_{W^{s,q}(\Omega)} \leq C \|F(\eta)\|_{W^{s,q}(\Omega)}$, $s \in \{0, 1\}$, $q \in [1, \infty]$, with C independent of h ;
- (F6) The projection is a linear operator, i.e. for any $u, v \in L^s(\Omega)$, $s \in [1, \infty]$, $a, b \in \mathbb{R}$, $\Pi_h(au + bv) = a\Pi_h u + b\Pi_h v$;
- (F7) For any $\eta \in \mathcal{H}_\gamma^F$, $\|F(\eta) - F_h(\eta)\|_{L^s(\Omega)} \rightarrow 0$ for any $s \in [1, \infty]$.

The interpolant operator, as defined in [15, Def. 3.3.9], would satisfy the assumptions (F5)–(F7), cf. [15, Th. 4.4.20] and [15, Prop. 3.3.4]. Recall the definition of $\tilde{\mathcal{H}}_\gamma^F$ in (2.14). We define the FE space $\tilde{\mathcal{H}}_{\gamma,h}^F$ as

$$\tilde{\mathcal{H}}_{\gamma,h}^F := \{\tilde{\eta}_h \in \mathcal{H}_{\gamma,h}^F : \exists \eta_h \in \mathcal{H}_{\gamma,h}^F \text{ such that } \tilde{\eta}_h = F_h(\eta_h)\}. \quad (3.3)$$

We make the following density assumption:

(D4) The FE space $\tilde{\mathcal{H}}_{\gamma,h}^F$ is dense in $\tilde{\mathcal{H}}_\gamma^F$, i.e. for any $\tilde{\eta} \in \tilde{\mathcal{H}}_\gamma^F$, for some $q \in (1, \infty)$,

$$\lim_{h \rightarrow 0} \inf_{\tilde{\zeta}_h \in \tilde{\mathcal{H}}_{\gamma,h}^F} \|\tilde{\eta} - \tilde{\zeta}_h\|_{W^{1,q}(\Omega)} = 0. \quad (3.4)$$

3.2 FE discretized optimization problem

The discretized compliance is denoted by J_h ,

$$J_h(\mathbf{u}_h, \rho_h) := l_h(\mathbf{u}_h) + \mathcal{R}(\rho_h), \quad (3.5)$$

where l_h is defined in (3.1). We define the bilinear form

$$a_h(\mathbf{u}_h, \mathbf{v}_h; \rho_h) := \int_{\Omega} k(F_h(\rho_h)) \mathbf{E} \mathbf{u}_h : \mathbf{E} \mathbf{v}_h \, dx. \quad (3.6)$$

Definition 3.1. Consider the FE spaces $\mathbf{V}_{D,h} \subset H^1(\Omega)^d$ and $\mathcal{H}_{\gamma,h} \subset \mathcal{H}_\gamma$. The FE discretized optimization problem is to find $(\mathbf{u}_h, \rho_h) \in \mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}$ that satisfies,

$$\begin{cases} \min_{(\mathbf{w}_h, \eta_h) \in \mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}} J_h(\mathbf{w}_h, \eta_h) \\ \text{subject to } a_h(\mathbf{w}_h, \mathbf{v}_h; \eta_h) = l_h(\mathbf{v}_h) \text{ for all } \mathbf{v}_h \in \mathbf{V}_{D,h}, \end{cases} \quad (\text{SIMP}_h) \quad (3.7)$$

where J_h , l_h , and $a_h(\cdot, \cdot; \cdot)$ are defined in (3.5), (3.1), and (3.6), respectively. A local minimizer of (SIMP_h) is called a local FE minimizer.

3.3 The main results

The following three theorems are the core aims of this work. In both Theorems 3.1 and 3.2 we tackle the open issue of multiple local minimizers (P1). We prove that every isolated minimizer is arbitrarily closely approximated by the FE method. If a $W^{1,p}$ -type penalty is used, in Theorem 3.1, we show that there exists a sequence of local FE minimizers satisfying $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$ and $\rho_h \rightarrow \rho$ strongly in $W^{1,p}(\Omega)$. This resolves the open problem (P2). If one chooses a density filter as a regularization method, then in Theorem 3.2, we show that $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$, resolving the open problem (P3). Finally, we also prove that $F_h(\rho_h) \rightarrow F(\rho)$ strongly in $W^{1,q}(\Omega)$ as a result of Theorem 3.3. This observation resolves (P4).

Theorem 3.1 (First main theorem: $W^{1,p}$ -type penalty methods). Suppose that \mathcal{R} is a $W^{1,p}$ -type penalty functional for some $p \in (1, \infty)$ and $F(\rho) = \rho$. Consider the FE spaces $\mathbf{V}_h \subset H^1(\Omega)^d$ and $\mathcal{H}_{\gamma,h}^p \subset \mathcal{H}_\gamma^p$. Suppose that (D1)–(D3) hold. Fix an isolated local minimizer $(\mathbf{u}, \rho) \in \mathbf{V}_D \times \mathcal{H}_\gamma^p$ of (SIMP).

Then, as $h \rightarrow 0$, there exists a sequence of local FE minimizers $(\mathbf{u}_h, \rho_h) \in \mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}^p$ of (SIMP_h) such that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$ and $\rho_h \rightarrow \rho$ strongly in $W^{1,p}(\Omega)$.

Theorem 3.2 (Second main theorem: density filters). Suppose that a density filter is used that satisfies (F1)–(F7) and $\mathcal{R}(\rho) = 0$. Consider the FE spaces $\mathbf{V}_h \subset H^1(\Omega)^d$ and $\mathcal{H}_{\gamma,h}^F \subset \mathcal{H}_\gamma^F$. Suppose that (D1)–(D3) hold. Fix an isolated local minimizer $(\mathbf{u}, \rho) \in \mathbf{V}_D \times \mathcal{H}_\gamma^F$ of (SIMP).

Then, as $h \rightarrow 0$, there exists a sequence of local FE minimizers $(\mathbf{u}_h, \rho_h) \in \mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}^F$ of (SIMP_h) such that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$ and $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$.

Recall that in the case of pure filtering (SIMP) is equivalent to (F-SIMP). Hence, if $(\mathbf{u}, \tilde{\rho})$ is the minimizer of (F-SIMP), then $(\mathbf{u}, \tilde{\rho}) = (\mathbf{u}, F(\rho))$ where (\mathbf{u}, ρ) is the equivalent minimizer of (SIMP).

Theorem 3.3 (Third main theorem: density filters). Suppose that the conditions of Theorem 3.2 hold, the space of filtered material distributions is $W^{1,q}$ -conforming, $\tilde{\mathcal{H}}_{\gamma,h}^F \subset (\tilde{\mathcal{H}}_\gamma^F \cap W^{1,q}(\Omega))$, for some $q \in (1, \infty)$, and the density assumption of (D4) is valid.

Assume that the minimizer $(\mathbf{u}, \tilde{\rho}) = (\mathbf{u}, F(\rho)) \in \mathbf{V}_D \times \tilde{\mathcal{H}}_\gamma^F$ is an isolated minimizer of (F-SIMP).

Then, there exists a subsequence (up to relabelling) such that

$$\mathbf{u}_h \rightarrow \mathbf{u} \text{ strongly in } H^1(\Omega)^d, \quad (3.7)$$

$$F_h(\rho_h) \rightarrow F(\rho) \text{ strongly in } W^{1,q}(\Omega). \quad (3.8)$$

3.4 Auxiliary results

The following lemmas are useful results and are used later in the convergence analysis. Their proofs are provided in Section A for completeness.

Lemma 3.1. *Consider a sequence $(\eta_j) \subset L^\infty(\Omega)$, $0 \leq \eta_j \leq 1$ a.e. such that $\eta_j \rightarrow \eta$ strongly in $L^q(\Omega)$ for some $q \in [1, \infty)$. Then $\eta_j \rightarrow \eta$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$.*

Lemma 3.2. *Suppose that (F1)–(F7) hold. Consider any sequence $(\eta_h) \subset \mathcal{H}_{\gamma,h}^F$ such that $\eta_h \rightharpoonup \eta \in \mathcal{H}_\gamma^F$ weakly in $L^q(\Omega)$ ($\eta_h \xrightarrow{*} \eta$ weakly-* in $L^\infty(\Omega)$ if $q = \infty$) for $q \in [1, \infty]$. Then $F_h(\eta_h) \rightarrow F(\eta)$ strongly in $L^q(\Omega)$.*

Lemma 3.3. *Suppose that (D1)–(D3) hold. Consider the following two cases:*

(Case 1: $W^{1,p}$ -type penalty). *Suppose that \mathcal{R} is a $W^{1,p}$ -type penalty functional, and $F(\rho) = \rho$. Consider any sequence $(\hat{\mathbf{u}}_h, \hat{\rho}_h) \in \mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}^p$ satisfying the PDE constraint in (SIMP)_h such that $\hat{\mathbf{u}}_h \rightharpoonup \hat{\mathbf{u}} \in \mathbf{V}_D$ weakly in $H^1(\Omega)^d$. Moreover, suppose that $\hat{\rho}_h \rightarrow \hat{\rho} \in \mathcal{H}_\gamma^p$ strongly in $L^q(\Omega)$ for some $q \in [2, \infty)$. Then the limit $(\hat{\mathbf{u}}, \hat{\rho})$ satisfies the PDE constraint in (SIMP).*

(Case 2: density filtering). *Suppose that a density filter is used that satisfies (F1)–(F7), and $\mathcal{R}(\rho) = 0$. Consider any sequence $(\hat{\mathbf{u}}_h, \hat{\rho}_h) \in \mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}^F$ satisfying the PDE constraint in (SIMP)_h such that $\hat{\mathbf{u}}_h \rightharpoonup \hat{\mathbf{u}} \in \mathbf{V}_D$ weakly in $H^1(\Omega)^d$. Moreover, suppose that $\hat{\rho}_h \rightharpoonup \hat{\rho}$ weakly in $L^q(\Omega)$ for some $q \in [2, \infty)$ ($\hat{\rho}_h \xrightarrow{*} \hat{\rho}$ weakly-* in $L^\infty(\Omega)$ if $q = \infty$). Then the limit $(\hat{\mathbf{u}}, \hat{\rho})$ satisfies the PDE constraint in (SIMP).*

Results similar to Lemma 3.3 are found in the literature e.g. [38, Lem. 2.1].

Lemma 3.4 (Existence of a strongly converging sequence $\hat{\mathbf{u}}_h$ in $H^1(\Omega)^d$). *Suppose that (D1)–(D3) hold. Consider a pair $(\mathbf{u}, \rho) \in \mathbf{V}_D \times \mathcal{H}_\gamma$ that satisfies the PDE constraint in (SIMP). Consider the following two cases:*

(Case 1). *Suppose that $F(\rho) = \rho$. Consider any sequence such that $\hat{\rho}_h \rightarrow \rho$ strongly in $L^q(\Omega)$, for some $q \in [2, \infty]$, and the corresponding (unique) sequence of elastic displacements such that $(\hat{\mathbf{u}}_h, \hat{\rho}_h)$ satisfies the PDE constraint in (SIMP)_h. Then, $\hat{\mathbf{u}}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$.*

(Case 2). *Suppose that a density filter is used that satisfies (F1)–(F7). Consider any sequence such that $F_h(\hat{\rho}_h) \rightarrow F(\rho)$ strongly in $L^q(\Omega)$, for some $q \in [2, \infty]$, and the corresponding (unique) sequence of elastic displacements such that $(\hat{\mathbf{u}}_h, \hat{\rho}_h)$ satisfies the PDE constraint in (SIMP)_h. Then, $\hat{\mathbf{u}}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$.*

3.5 FE analysis: $W^{1,p}$ -type penalty methods

In this subsection we assume that a $W^{1,p}$ -type penalty is used, with functional \mathcal{R} , as defined in Definition 2.1 and no density filter is used, i.e. $F(\eta) = F_h(\eta) = \eta$. Fix a local isolated minimizer $(\mathbf{u}, \rho) \in \mathbf{V}_D \times \mathcal{H}_\gamma^p$ of (SIMP). Consider the finite-dimensional optimization problem: find $(\mathbf{u}_h, \rho_h) \in \mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}^p$ that minimizes

$$\begin{cases} \min_{(\mathbf{v}_h, \eta_h) \in (\mathbf{V}_{D,h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_{\gamma,h}^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho))} J_h(\mathbf{v}_h, \eta_h), \\ \text{subject to } a_h(\mathbf{u}_h, \mathbf{v}_h; \rho_h) = l_h(\mathbf{v}_h) \text{ for all } \mathbf{v}_h \in \mathbf{V}_{D,h}. \end{cases} \quad (\text{I-SIMP}_h^p)$$

If there exists a pair $(\mathbf{u}_h, \rho_h) \in (\mathbf{V}_{D,h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_{\gamma,h}^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho))$ such that $J_h(\mathbf{u}_h, \rho_h) \leq J_h(\mathbf{v}_h, \eta_h)$ for all $(\mathbf{v}_h, \eta_h) \in (\mathbf{V}_{D,h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_{\gamma,h}^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho))$, then we call (\mathbf{u}_h, ρ_h) a *global FE minimizer*.

In Proposition 3.1, we prove the existence of finite-dimensional global minimizers to the discretized problem (I-SIMP_h^p). Then, in Proposition 3.2, we prove weak convergence of the discretized global minimizers to the isolated infinite-dimensional local minimizer as $h \rightarrow 0$. Direct consequences of the weak convergence are derived in Corollaries 3.1 and 3.2. We derive a strong convergence result in Proposition 3.3 and conclude this subsection by proving the result of Theorem 3.1.

Proposition 3.1. *Suppose that the conditions of Theorem 3.1 hold. Then a global FE minimizer $(\mathbf{u}_h, \rho_h) \in \mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}^p$ of (I-SIMP_h^p) exists.*

Proof. The set $(\mathbf{V}_{D,h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_{\gamma,h}^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho))$ is a finite-dimensional, closed, and bounded set. Moreover, for sufficiently small h it is non-empty. Therefore, it is sequentially compact by the Heine–Borel theorem [22, Th. 11.18]. We say that $(\mathbf{v}, \eta) \in \mathcal{S}$ if (\mathbf{v}, η) satisfies the PDE constraint in (I-SIMP_h^p). By Korn’s inequality [7, Th. 11], and the assumption that $\epsilon_{\text{SIMP}} > 0$, then the bilinear form $a_h(\cdot, \cdot; \eta_h)$ is coercive for any $\eta_h \in \mathcal{H}_{\gamma,h}^p$. Moreover, the bilinear form is bounded, (D3) asserts that $l_h(\mathbf{v}) = 0$ for all $\mathbf{v} \in H^1(\Omega)^d \cap \text{RM}^d$, and therefore, by the Lax–Milgram theorem [20, Ch. 6.2, Th. 1], for every $\eta_h \in \mathcal{H}_{\gamma,h}^p$, there exists a unique $\mathbf{v}_h \in \mathbf{V}_{D,h}$ such that $(\mathbf{v}_h, \eta_h) \in \mathcal{S}$. Moreover, $\|\mathbf{v}_h\|_{H^1(\Omega)}$ is bounded thanks to the boundedness of $\|\mathbf{f}_h\|_{L^2(\Omega)}$, $\|g_h^n\|_{L^2(\Gamma_N^n)}$, and $\|g_h^t\|_{L^2(\Gamma_N^t)}$. Hence, $\mathcal{S} \cap (\mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}^p)$ is finite-dimensional, closed, and bounded and thus sequentially compact by the Heine–Borel theorem. As the spaces are Hausdorff, the intersection of two compact sets is compact and, hence, $\mathcal{S} \cap ((\mathbf{V}_{D,h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_{\gamma,h}^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho)))$ is sequentially compact.

By assumption the functional J_h is continuous in the strong topology of the finite-dimensional space $\mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}^p$. Hence J_h obtains its infimum in $\mathcal{S} \cap ((\mathbf{V}_{D,h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_{\gamma,h}^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho)))$ and therefore, a global FE minimizer (\mathbf{u}_h, ρ_h) of (I-SIMP_h^p) exists. \square

The following proposition is the first step in tackling the open problem (P1).

Proposition 3.2 (Weak \times weak convergence of (\mathbf{u}_h, ρ_h) in $H^1(\Omega)^d \times W^{1,p}(\Omega)$). *Suppose that the conditions of Theorem 3.1 hold. Then there exists a subsequence of the global FE minimizers of (I-SIMP_h^p), that satisfies*

$$\mathbf{u}_h \rightharpoonup \mathbf{u} \text{ weakly in } H^1(\Omega)^d, \quad (3.9)$$

$$\rho_h \rightharpoonup \rho \text{ weakly in } W^{1,p}(\Omega). \quad (3.10)$$

Proof. By a corollary of Kakutani’s Theorem [23, Th. A.65], if a Banach space is reflexive then every norm-closed, bounded and convex subset of the Banach space is weakly compact and thus, by the Eberlein–Šmulian theorem [23, Th. A.62], sequentially weakly compact. It can be checked that $\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})$ and $\mathcal{H}_\gamma^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho)$ are norm-closed, bounded and convex subsets of the reflexive Banach spaces $H^1(\Omega)^d$ and $W^{1,p}(\Omega)$, respectively. Therefore, $\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})$ is weakly sequentially compact in $H^1(\Omega)^d$ and $\mathcal{H}_\gamma^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho)$ is weakly sequentially compact in $W^{1,p}(\Omega)$.

Hence we extract subsequences (up to relabelling), (\mathbf{u}_h) and (ρ_h) of the sequence generated by the global FE minimizers of (I-SIMP_h^p) such that

$$\mathbf{u}_h \rightharpoonup \mathbf{u}_0 \in \mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u}) \text{ weakly in } H^1(\Omega)^d, \quad (3.11)$$

$$\rho_h \rightharpoonup \rho_0 \in \mathcal{H}_\gamma^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho) \text{ weakly in } W^{1,p}(\Omega). \quad (3.12)$$

In the remainder of the proof, we show that the weak limit (\mathbf{u}_0, ρ_0) is in fact the isolated minimizer (\mathbf{u}, ρ) . By the Rellich–Kondrachov theorem, there exists a subsequence (up to relabelling) that satisfies

$$\rho_h \rightarrow \rho_0 \text{ strongly in } L^1(\Omega) \quad (3.13)$$

and thus $\rho_h \rightarrow \rho_0$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$ by Lemma 3.1. Hence, we satisfy the requirement of Lemma 3.3 and deduce that (\mathbf{u}_0, ρ_0) satisfies the PDE constraint in (SIMP).

The next step is to prove that (\mathbf{u}_0, ρ_0) is a minimizing pair. By assumption (D2), there exists a sequence of FE functions $\hat{\rho}_h \in \mathcal{H}_{\gamma,h}^p$ that strongly converges to ρ in $W^{1,p}(\Omega)$. Moreover, by Lemma 3.4, there exists a sequence $(\hat{\mathbf{u}}_h)$ such that $(\hat{\mathbf{u}}_h, \hat{\rho}_h)$ satisfies the PDE constraint in (I-SIMP $_h^p$) and $\hat{\mathbf{u}}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$. Thus

$$|J_h(\hat{\mathbf{u}}_h, \hat{\rho}_h) - J(\mathbf{u}, \rho)| \leq |l_h(\hat{\mathbf{u}}_h) - l(\mathbf{u})| + |\mathcal{R}(\rho) - \mathcal{R}(\hat{\rho}_h)|. \quad (3.14)$$

By (D1), we have strong convergence of \mathbf{f}_h in $L^2(\Omega)$, g_h^n in $L^2(\Gamma_N^n)$, and g_h^t in $L^2(\Gamma_N^t)$. Thus together with the strong convergence of $\hat{\rho}_h$ and $\hat{\mathbf{u}}_h$ we see that

$$J_h(\hat{\mathbf{u}}_h, \hat{\rho}_h) \rightarrow J(\mathbf{u}, \rho) \text{ as } h \rightarrow 0.$$

Furthermore, for sufficiently small $h > 0$,

$$(\hat{\mathbf{u}}_h, \hat{\rho}_h) \in (\mathbf{V}_{D,h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_{\gamma,h}^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho)).$$

Therefore,

$$J_h(\mathbf{u}_h, \rho_h) \leq J_h(\hat{\mathbf{u}}_h, \hat{\rho}_h). \quad (3.15)$$

By taking the limit as $h \rightarrow 0$ and utilizing the strong convergence of $\hat{\mathbf{u}}_h$ and $\hat{\rho}_h$ to \mathbf{u} and ρ , respectively, we see that

$$\limsup_{h \rightarrow 0} J_h(\mathbf{u}_h, \rho_h) \leq J(\mathbf{u}, \rho). \quad (3.16)$$

Recall that $J(\mathbf{u}, \rho) = l(\mathbf{u}) + \mathcal{R}(\rho)$. Then by the linearity of $l(\mathbf{u})$ and Definition 2.1, J is weakly lower semicontinuous on $H^1(\Omega)^d \times W^{1,p}(\Omega)$. Therefore,

$$J(\mathbf{u}_0, \rho_0) \leq \liminf_{h \rightarrow 0} J_h(\mathbf{u}_h, \rho_h). \quad (3.17)$$

(\mathbf{u}, ρ) is the unique local minimizer of (SIMP) in $(\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_{\gamma}^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho))$. Thus $J(\mathbf{u}, \rho) \leq J(\mathbf{u}_0, \rho_0)$. Hence, from (3.16) and (3.17), it follows that

$$J(\mathbf{u}_0, \rho_0) = J(\mathbf{u}, \rho). \quad (3.18)$$

Since (\mathbf{u}, ρ) is the unique local minimizer in the spaces we consider and (\mathbf{u}_0, ρ_0) satisfies the PDE constraint in (SIMP), we identify the limits \mathbf{u}_0 and ρ_0 as \mathbf{u} and ρ , respectively, and state that $\mathbf{u}_h \rightharpoonup \mathbf{u}$ weakly in $H^1(\Omega)^d$ and $\rho_h \rightharpoonup \rho$ weakly in $W^{1,p}(\Omega)$. \square

Corollary 3.1. *Suppose that the conditions of Theorem 3.1 hold. Then there exists a subsequence of the global FE minimizers (\mathbf{u}_h, ρ_h) of (I-SIMP $_h^p$) (up to relabelling) such that $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$.*

Proof. By Proposition 3.2, $\rho_h \rightharpoonup \rho$ weakly in $W^{1,p}(\Omega)$. Thus, for $p \geq 1$, the Rellich–Kondrachov theorem implies that there exists a subsequence such that $\rho_h \rightarrow \rho$ strongly in $L^1(\Omega)$. The result follows from Lemma 3.1. \square

Corollary 3.2 (Strong convergence of \mathbf{u}_h in $H^1(\Omega)^d$). *Suppose that the conditions of Theorem 3.1 hold. There exists a subsequence of the global FE minimizers (\mathbf{u}_h, ρ_h) of (I-SIMP $_h^p$) such that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$.*

Proof. By Corollary 3.1, we have that $\rho_h \rightarrow \rho$ strongly in $L^2(\Omega)$. Hence, the result follows as a direct consequence of (Case 1) in Lemma 3.4. \square

We now focus on proving the stronger results of Theorem 3.1. We show that there exists a subsequence (up to relabelling) such that $\rho_h \rightarrow \rho$ strongly in $W^{1,p}(\Omega)$.

Proposition 3.3 (Strong convergence of ρ_h in $W^{1,p}(\Omega)$, $p \in (1, \infty)$). *Suppose that the conditions of Theorem 3.1 hold. Then, there exists a subsequence (up to relabelling) of global FE minimizers of (I-SIMP $_h^p$), that satisfies*

$$\rho_h \rightarrow \rho \text{ strongly in } W^{1,p}(\Omega). \quad (3.19)$$

Proof. By (D2), there exists a sequence of functions $(\hat{\rho}_h)$ such that $\hat{\rho}_h \rightarrow \rho$ strongly in $W^{1,p}(\Omega)$. Moreover, by (Case 1) in Lemma 3.4 the corresponding sequence of elastic displacements $(\hat{\mathbf{u}}_h)$ such that the pair $(\hat{\mathbf{u}}_h, \hat{\rho}_h)$ satisfies the PDE constraint in (SIMP $_h$) also satisfies $\hat{\mathbf{u}}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$. Thanks to the strong convergence, for sufficiently small h , we have that $(\hat{\mathbf{u}}_h, \hat{\rho}_h) \in (\mathbf{V}_{D,h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_{\gamma,h}^p \cap B_{r/2, W^{1,p}(\Omega)}(\rho))$. Hence

$$J(\mathbf{u}, \rho) \leq J_h(\mathbf{u}_h, \rho_h) \leq J_h(\hat{\mathbf{u}}_h, \hat{\rho}_h). \quad (3.20)$$

In turn this implies that

$$\begin{aligned} & \frac{\epsilon}{p} \left| \|\nabla \rho\|_{L^p(\Omega)}^p - \|\nabla \rho_h\|_{L^p(\Omega)}^p \right| \\ & \leq |J_h(\hat{\mathbf{u}}_h, \hat{\rho}_h) - J(\mathbf{u}, \rho)| + |l(\mathbf{u}) - l_h(\mathbf{u}_h)| + |m(\rho) - m(\rho_h)|. \end{aligned} \quad (3.21)$$

Label the three terms on the right-hand side of (3.21) by (I), (II), and (III), respectively. Thanks to the strong convergence of $(\hat{\mathbf{u}}_h, \hat{\rho}_h)$ in $H^1(\Omega)^d \times W^{1,p}(\Omega)$ then (I) $\rightarrow 0$. Similarly due to (D1) and Corollary 3.2, we have that (II) $\rightarrow 0$. Finally since, by assumption, $m(\cdot)$ is continuous in $L^p(\Omega)$, Corollary 3.1 implies that (III) $\rightarrow 0$. Therefore, $\|\nabla \rho_h\|_{L^p(\Omega)}^p \rightarrow \|\nabla \rho\|_{L^p(\Omega)}^p$ and consequently

$$\|\nabla \rho_h\|_{L^p(\Omega)} \rightarrow \|\nabla \rho\|_{L^p(\Omega)}. \quad (3.22)$$

Whenever $p \in (1, \infty)$, $L^p(\Omega)$ is a uniformly convex Banach space. Hence, by the fact that $\rho_h \rightharpoonup \rho$ weakly in $W^{1,p}(\Omega)$ (as shown in Proposition 3.2) and (3.22) holds, the Radon–Riesz property [23, Th. A.70] implies that $\nabla \rho_h \rightarrow \nabla \rho$ strongly in $L^p(\Omega)$. Consequently $\rho_h \rightarrow \rho$ strongly in $W^{1,p}(\Omega)$. \square

Corollary 3.3. *Suppose that the conditions of Theorem 3.1 hold. Then for sufficiently small h , there exists a subsequence of global FE minimizers (\mathbf{u}_h, ρ_h) of (I-SIMP $_h^p$) that are also local FE minimizers of (SIMP $_h$).*

Proof. Corollary 3.2 and Proposition 3.3 imply that there exists a subsequence (\mathbf{u}_h, ρ_h) such that $(\mathbf{u}_h, \rho_h) \rightarrow (\mathbf{u}, \rho)$ strongly in $H^1(\Omega)^d \times W^{1,p}(\Omega)$. By the definition of strong convergence, there exists an \bar{h} such that for all $h \leq \bar{h}$, $\|\mathbf{u} - \mathbf{u}_h\|_{H^1(\Omega)} + \|\rho - \rho_h\|_{W^{1,p}(\Omega)} \leq r/4$. Thus the basin of attraction constraint is not active and the subsequence of global FE minimizers of (I-SIMP_h^p) are also local FE minimizers of (SIMP_h). \square

We now have the sufficient ingredients to prove Theorem 3.1.

Proof of Theorem 3.1. By Proposition 3.1, there exists a sequence of global FE minimizers (\mathbf{u}_h, ρ_h) of (I-SIMP_h^p) that satisfies $(\mathbf{u}_h, \rho_h) \rightharpoonup (\mathbf{u}, \rho)$ weakly in $H^1(\Omega)^d \times W^{1,p}(\Omega)$. In Corollary 3.2, it was shown that there exists a subsequence (up to relabelling) such that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$. Then, in Proposition 3.3, we demonstrate that there exists a subsequence (up to relabelling) such that $\rho_h \rightarrow \rho$ strongly in $W^{1,p}(\Omega)$. Finally, in Corollary 3.3, we conclude that a subsequence of the global FE minimizers are also local FE minimizers of (SIMP_h). \square

3.6 FE analysis: density filtering

In this subsection we assume a density filter is applied, with a mapping F that satisfies (F1)–(F7), and that no penalty method is used, i.e. $\mathcal{R} \equiv 0$. Fix an isolated local minimizer $(\mathbf{u}, \rho) \in \mathbf{V}_D \times \mathcal{H}_\gamma^F$ of (SIMP). Consider the finite-dimensional optimization problem: find $(\mathbf{u}_h, \rho_h) \in \mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}^F$ that minimizes

$$\begin{cases} \min_{(\mathbf{v}_h, \eta_h) \in (\mathbf{V}_{D,h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_{\gamma,h}^F \cap B_{r/2, L^2(\Omega)}(\rho))} J_h(\mathbf{v}_h, \eta_h), \\ \text{subject to } a_h(\mathbf{u}_h, \mathbf{v}_h; \rho_h) = l_h(\mathbf{v}_h) \text{ for any } \mathbf{v}_h \in \mathbf{V}_{D,h}. \end{cases} \quad (\text{I-SIMP}_h^F)$$

In Proposition 3.4, we prove the existence of a finite-dimensional global minimizer to the discretized problem (I-SIMP_h^F). Then, in Proposition 3.5, we prove weak convergence of the discretized minimizers to the isolated infinite-dimensional local minimizer as $h \rightarrow 0$. Direct consequences of the weak convergence are derived in Corollaries 3.4, 3.5, and 3.6. We derive a strong convergence result in Proposition 3.6 for unfiltered material distributions which allows us to deduce the result of Theorem 3.2. We conclude this subsection by proving the result of Theorem 3.3.

Proposition 3.4. *Suppose that the conditions of Theorem 3.2 hold. Then a global FE minimizer $(\mathbf{u}_h, \rho_h) \in \mathbf{V}_{D,h} \times \mathcal{H}_{\gamma,h}^F$ of (I-SIMP_h^F) exists.*

Proof. The proof of this result follows the proof of Proposition 3.1 with some small modifications, up to replacing the spaces \mathcal{H}_γ^p and $\mathcal{H}_{\gamma,h}^p$ with \mathcal{H}_γ^F and $\mathcal{H}_{\gamma,h}^F$, respectively, and the optimization problem (I-SIMP_h^p) with (I-SIMP_h^F). \square

Proposition 3.5 (Weak \times weak convergence of (\mathbf{u}_h, ρ_h) in $H^1(\Omega)^d \times L^2(\Omega)$). *Suppose that the conditions of Theorem 3.2 hold. Then there exist subsequences (up to relabelling) of the global FE minimizers of (I-SIMP_h^F), that satisfy*

$$\mathbf{u}_h \rightharpoonup \mathbf{u} \text{ weakly in } H^1(\Omega)^d, \quad (3.23)$$

$$\rho_h \rightharpoonup \rho \text{ weakly in } L^2(\Omega). \quad (3.24)$$

Proof. The proof of this result follows the proof of Proposition 3.2 with some small modifications, up to replacing the spaces \mathcal{H}_γ^p , $\mathcal{H}_{\gamma,h}^p$, and $W^{1,p}(\Omega)$ with \mathcal{H}_γ^F , $\mathcal{H}_{\gamma,h}^F$, and $L^2(\Omega)$, respectively, and the optimization problem (I-SIMP $_h^p$) with (I-SIMP $_h^F$). Moreover, we do not need to invoke the Rellich–Kondrachov theorem to deduce the strongly converging subsequence in (3.13) as, in this framework, the conditions of Lemma 3.3 are already satisfied. \square

Corollary 3.4. *Suppose that the conditions of Theorem 3.2 hold. Then there exists a subsequence (up to relabelling) of the global FE minimizers (ρ_h) of (I-SIMP $_h^F$) such that $\rho_h \xrightarrow{*} \rho$ weakly-* in $L^\infty(\Omega)$ and consequently $\rho_h \rightharpoonup \rho$ weakly in $L^s(\Omega)$ for all $s \in [1, \infty)$.*

Proof. By the Banach–Alaoglu theorem, the closed unit ball of the dual space of a normed vector space is compact in the weak-* topology and, therefore, the closed unit ball of $L^\infty(\Omega)$ is weak-* compact. Hence we find a subsequence (up to relabelling) such that $\rho_h \xrightarrow{*} \rho_0 \in \mathcal{H}_\gamma^F \cap \{\eta : \|\rho - \eta\|_{L^\infty(\Omega)} \leq r/2\}$ weakly-* in $L^\infty(\Omega)$. By the uniqueness of the weak limit, we identify $\rho_0 = \rho$ a.e. in Ω and, thus, we deduce that $\rho_h \xrightarrow{*} \rho$ weakly-* in $L^\infty(\Omega)$, i.e. $\int_\Omega \rho_h \eta \, dx \rightarrow \int_\Omega \rho \eta \, dx$ for any $\eta \in L^1(\Omega)$. Since $L^s(\Omega) \subset L^1(\Omega)$, for any $s > 1$, we note that $\int_\Omega \rho_h \eta \, dx \rightarrow \int_\Omega \rho \eta \, dx$ for any $\eta \in L^s(\Omega)$, $s \geq 1$. Hence, by definition, $\rho_h \rightharpoonup \rho$ weakly in $L^s(\Omega)$ for all $s \in [1, \infty)$. \square

Corollary 3.5. *Suppose that the conditions of Theorem 3.2 hold. Let (ρ_h) denote the sequence of global FE minimizers to (I-SIMP $_h^F$) such that $\rho_h \xrightarrow{*} \rho$ weakly-* in $L^\infty(\Omega)$. Then $F_h(\rho_h) \rightarrow F(\rho)$ strongly in $L^\infty(\Omega)$.*

Proof. The result is a direct consequence of Lemma 3.2. \square

Corollary 3.6 (Strong convergence of \mathbf{u}_h in $H^1(\Omega)^d$). *Suppose that the conditions of Theorem 3.2 hold. Then there exists a subsequence (up to relabelling) of the global FE minimizers (\mathbf{u}_h, ρ_h) of (I-SIMP $_h^F$) such that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$.*

Proof. By Corollary 3.5, we have a subsequence such that $F_h(\rho_h) \rightarrow F(\rho)$ strongly in $L^\infty(\Omega)$. Hence, the result follows as a direct consequence of (Case 2) in Lemma 3.4. \square

In the remainder of this subsection we tackle the open problems (P3) and (P4). In particular, we prove the stronger results of Theorem 3.2 and Theorem 3.3 Recall that (\mathbf{u}, ρ) is the fixed local minimizer of (SIMP).

We first prove the following two lemmas.

Lemma 3.5. *Suppose that the conditions of Theorem 3.2 hold. Then a local minimizer exists to the following ϵ -perturbed optimization problem:*

$$\begin{cases} \min_{\mathbf{u}_\epsilon, \rho_\epsilon} J_\epsilon(\mathbf{u}_\epsilon, \rho_\epsilon) := J(\mathbf{u}_\epsilon, \rho_\epsilon) + \frac{\epsilon}{2} \|\rho_\epsilon\|_{L^2(\Omega)}^2, \\ \text{where } (\mathbf{u}_\epsilon, \rho_\epsilon) \in (\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_\gamma^F \cap B_{r/2, L^2(\Omega)}(\rho)) \\ \text{subject to } a(\mathbf{u}_\epsilon, \mathbf{v}; \rho_\epsilon) = l(\mathbf{v}) \text{ for all } \mathbf{v} \in \mathbf{V}_D. \end{cases} \quad (\text{I-SIMP}_\epsilon^F)$$

Proof. Consider a minimizing sequence $(\mathbf{u}_n, \rho_n) \in (\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_\gamma^F \cap B_{r/2, L^2(\Omega)}(\rho))$. Since $\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})$ and $\mathcal{H}_\gamma^F \cap B_{r/2, L^2(\Omega)}(\rho)$ are norm-closed, bounded, and convex subsets of $H^1(\Omega)^d$ and $L^2(\Omega)$, respectively, then there exists a limit $(\mathbf{u}_\epsilon, \rho_\epsilon) \in (\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_\gamma^F \cap B_{r/2, L^2(\Omega)}(\rho))$ such that

$$\mathbf{u}_n \rightharpoonup \mathbf{u}_\epsilon \text{ weakly in } H^1(\Omega)^d, \quad (3.25)$$

$$\rho_n \rightharpoonup \rho_\epsilon \text{ weakly in } L^2(\Omega). \quad (3.26)$$

By Lemma 3.3, $(\mathbf{u}_\epsilon, \rho_\epsilon)$ satisfies the PDE constraint in (I-SIMP $_\epsilon^F$). Moreover, $J(\mathbf{u}, \rho)$ is bounded and weak \times weak lower semicontinuous on $H^1(\Omega)^d \times L^2(\Omega)$. Thus $J_\epsilon(\mathbf{u}_\epsilon, \rho_\epsilon) \leq \liminf_{n \rightarrow \infty} J_\epsilon(\mathbf{u}_n, \rho_n)$. Hence, a local minimizer exists. \square

Lemma 3.6. *Suppose that the conditions of Theorem 3.2 hold. Consider a sequence of local minimizers $(\mathbf{u}_\epsilon, \rho_\epsilon)$ of (I-SIMP $_\epsilon^F$). Then, there exists a subsequence (up to relabelling) such that $\mathbf{u}_\epsilon \rightharpoonup \mathbf{u}$ weakly in $H^1(\Omega)^d$ and $\rho_\epsilon \rightharpoonup \rho$ weakly in $L^2(\Omega)$ as $\epsilon \rightarrow 0$.*

Proof. Since $\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})$ and $\mathcal{H}_\gamma^F \cap B_{r/2, L^2(\Omega)}(\rho)$ are norm-closed, bounded, and convex subsets of $H^1(\Omega)^d$ and $L^2(\Omega)$, respectively, then there exists a limit $(\mathbf{u}_0, \rho_0) \in (\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_\gamma^F \cap B_{r/2, L^2(\Omega)}(\rho))$ such that

$$\mathbf{u}_\epsilon \rightharpoonup \mathbf{u}_0 \text{ weakly in } H^1(\Omega)^d, \quad (3.27)$$

$$\rho_\epsilon \rightharpoonup \rho_0 \text{ weakly in } L^2(\Omega). \quad (3.28)$$

By Lemma 3.3, (\mathbf{u}_0, ρ_0) satisfies the PDE constraint in (SIMP).

Consider any sequence of functions $\rho_n \in \mathcal{H}_\gamma^F$ such that $\rho_n \rightarrow \rho$ strongly in $L^2(\Omega)$. Then the corresponding sequence of elastic displacements $\mathbf{u}_n \in \mathbf{V}_D$ such that the pair (\mathbf{u}_n, ρ_n) satisfies the PDE constraint in (SIMP) also satisfies $\mathbf{u}_n \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$. For any $\epsilon > 0$ and sufficiently large $n \in \mathbb{N}$,

$$J_\epsilon(\mathbf{u}_\epsilon, \rho_\epsilon) \leq J_\epsilon(\mathbf{u}_n, \rho_n). \quad (3.29)$$

J_ϵ is weak \times weak lower semicontinuous on $H^1(\Omega)^d \times L^2(\Omega)$. Thus by taking the limits $\epsilon \rightarrow 0$ and $n \rightarrow \infty$ and using the weak convergence of $(\mathbf{u}_\epsilon, \rho_\epsilon)$ and the strong convergence of (\mathbf{u}_n, ρ_n) we see that

$$J(\mathbf{u}_0, \rho_0) \leq \liminf_{\epsilon \rightarrow 0} J_\epsilon(\mathbf{u}_\epsilon, \rho_\epsilon) \leq J(\mathbf{u}, \rho). \quad (3.30)$$

Since (\mathbf{u}, ρ) is the unique minimizer of (SIMP) in $(\mathbf{V}_D \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_\gamma^F \cap B_{r/2, L^2(\Omega)}(\rho))$, we identify $\mathbf{u}_0 = \mathbf{u}$ and $\rho_0 = \rho$ a.e. in Ω . \square

The following proposition is concerned with the open problem (P3). We show that if a density filter is used, then there exists a sequence of the *unfiltered* discretized material distributions that converges strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$.

Proposition 3.6 (Strong convergence of ρ_h in $L^s(\Omega)$, $s \in [1, \infty)$). *Suppose that the conditions of Theorem 3.2 hold. Then there exists a subsequence (up to relabelling) of global FE minimizers of (I-SIMP $_h^F$), that satisfy, for any $s \in [1, \infty)$,*

$$\rho_h \rightarrow \rho \text{ strongly in } L^s(\Omega). \quad (3.31)$$

In Figure 3, we provide the structure of the proof. The first step is to prove the two downward limits, as $\epsilon \rightarrow 0$, of minimizers of the ϵ -perturbed problem (I-SIMP $_\epsilon^F$) to minimizers of the problems (SIMP) and (I-SIMP $_h^F$). This is achieved via weak convergence, the convergence of the supremum value of the norms, and the Radon–Riesz property [23, Th. A.70]. Next, the top arrow, as $h \rightarrow 0$, in the ϵ -perturbed problem is proven using the properties of the minimization problem and the Radon–Riesz property. Then, since $\rho_{\epsilon, h}$, ρ_ϵ , ρ_h , and ρ are all bounded in $L^s(\Omega)$ uniformly in h and ϵ , we invoke the Osgood theorem [49, Ch. 4, Sec. 11, Th. 2'] for exchanging limits together with the Radon–Riesz property to conclude the result.

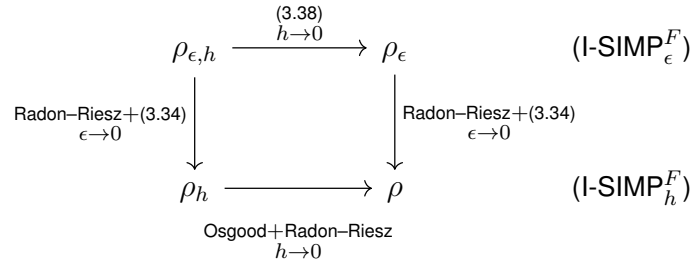


Figure 3: A summary of the proof of Proposition 3.6.

Proof of Proposition 3.6. Consider the global minimizer $(\mathbf{u}_{\epsilon}, \rho_{\epsilon})$ of $(\text{I-SIMP}_{\epsilon}^F)$. If there is more than one global minimizer, then fix one. After discretizing $(\text{I-SIMP}_{\epsilon}^F)$, then for each $\epsilon > 0$, by using the same arguments as in the proof of Proposition 3.5, there exists a sequence such that $\mathbf{u}_{\epsilon, h_{\epsilon}} \rightarrow \mathbf{u}_{\epsilon}$ strongly in $H^1(\Omega)^d$ and $\rho_{\epsilon, h_{\epsilon}} \xrightarrow{*} \rho_{\epsilon}$ weakly-* in $L^{\infty}(\Omega)$ as $h_{\epsilon} \rightarrow 0$ where $(\mathbf{u}_{\epsilon, h_{\epsilon}}, \rho_{\epsilon, h_{\epsilon}})$ is the (possibly non-unique) global minimizer of the FE discretized optimization problem $(\text{I-SIMP}_{\epsilon}^F)$ with mesh size h_{ϵ} .

By Lemma 3.6, there exists a sequence (up to relabelling) $(\mathbf{u}_{\epsilon}, \rho_{\epsilon})$ such that $\rho_{\epsilon} \rightharpoonup \rho$ weakly in $L^2(\Omega)$ and $\mathbf{u}_{\epsilon} \rightharpoonup \mathbf{u}$ weakly in $H^1(\Omega)^d$ as $\epsilon \rightarrow 0$. We now strengthen the convergence of material distribution to strong convergence in $L^2(\Omega)$ as $\epsilon \rightarrow 0$. Suppose that for any $\epsilon > 0$ under a certain threshold,

$$\|\rho\|_{L^2(\Omega)} < \|\rho_{\epsilon}\|_{L^2(\Omega)}. \quad (3.32)$$

This would imply that

$$J(\mathbf{u}, \rho) + \frac{\epsilon}{2} \|\rho\|_{L^2(\Omega)}^2 < J(\mathbf{u}_{\epsilon}, \rho_{\epsilon}) + \frac{\epsilon}{2} \|\rho_{\epsilon}\|_{L^2(\Omega)}^2. \quad (3.33)$$

However, this would be a contradiction with the definition of $(\mathbf{u}_{\epsilon}, \rho_{\epsilon})$. Thus, for the sequence such that $\rho_{\epsilon} \rightharpoonup \rho$ in $L^2(\Omega)$,

$$\limsup_{\epsilon \rightarrow 0} \|\rho_{\epsilon}\|_{L^2(\Omega)} \leq \|\rho\|_{L^2(\Omega)}. \quad (3.34)$$

Since $L^2(\Omega)$ is a uniformly convex Banach space, then by a generalized result of the Radon–Riesz property [23, Th. A.70], we have that $\rho_{\epsilon} \rightarrow \rho$ strongly in $L^2(\Omega)$. A similar argument gives, for each $h > 0$, $\rho_{\epsilon, h} \rightarrow \rho_h$ strongly in $L^2(\Omega)$, as $\epsilon_h \rightarrow 0$.

The next step is to show that, for each $\epsilon > 0$, there exists a subsequence such that $\|\rho_{\epsilon, h_{\epsilon}}\|_{L^2(\Omega)} \rightarrow \|\rho_{\epsilon}\|_{L^2(\Omega)}$ as $h_{\epsilon} \rightarrow 0$. The deduction of the result is similar to the proof of Proposition 3.3. By (D2), there exists a sequence of functions $(\hat{\rho}_{\epsilon, h_{\epsilon}})$ such that $\hat{\rho}_{\epsilon, h_{\epsilon}} \rightarrow \rho_{\epsilon}$ strongly in $L^2(\Omega)$. Moreover, by (Case 2) in Lemma 3.4 the corresponding sequence of elastic displacements $(\hat{\mathbf{u}}_{\epsilon, h_{\epsilon}})$ such that the pair $(\hat{\mathbf{u}}_{\epsilon, h_{\epsilon}}, \hat{\rho}_{\epsilon, h_{\epsilon}})$ satisfies the PDE constraint in $(\text{I-SIMP}_{\epsilon}^F)$ also satisfies $\hat{\mathbf{u}}_{\epsilon, h_{\epsilon}} \rightarrow \mathbf{u}_{\epsilon}$ strongly in $H^1(\Omega)^d$. Thanks to the strong convergence, for sufficiently small h_{ϵ} , we have that $(\hat{\mathbf{u}}_{\epsilon, h_{\epsilon}}, \hat{\rho}_{\epsilon, h_{\epsilon}}) \in (\mathbf{V}_{D, h} \cap B_{r/2, H^1(\Omega)}(\mathbf{u})) \times (\mathcal{H}_{\gamma, h}^F \cap B_{r/2, L^2(\Omega)}(\rho))$. Hence

$$J_{\epsilon}(\mathbf{u}_{\epsilon}, \rho_{\epsilon}) \leq J_{\epsilon, h_{\epsilon}}(\mathbf{u}_{\epsilon, h_{\epsilon}}, \rho_{\epsilon, h_{\epsilon}}) \leq J_{\epsilon, h_{\epsilon}}(\hat{\mathbf{u}}_{\epsilon, h_{\epsilon}}, \hat{\rho}_{\epsilon, h_{\epsilon}}), \quad (3.35)$$

where $J_{\epsilon, h}(\mathbf{v}, \eta) = l_h(\mathbf{v}) + \frac{\epsilon}{2} \|\eta\|_{L^2(\Omega)}^2$. In turn this implies that

$$\begin{aligned} & \frac{\epsilon}{2} \left| \|\rho_{\epsilon}\|_{L^2(\Omega)}^2 - \|\rho_{\epsilon, h_{\epsilon}}\|_{L^2(\Omega)}^2 \right| \\ & \leq |J_{\epsilon, h_{\epsilon}}(\hat{\mathbf{u}}_{\epsilon, h_{\epsilon}}, \hat{\rho}_{\epsilon, h_{\epsilon}}) - J_{\epsilon}(\mathbf{u}_{\epsilon}, \rho_{\epsilon})| + |l(\mathbf{u}_{\epsilon}) - l_h(\mathbf{u}_{\epsilon, h_{\epsilon}})|. \end{aligned} \quad (3.36)$$

Label the two terms on the right-hand side of (3.36) by (I) and (II), respectively. Thanks to the strong convergence of $(\hat{\mathbf{u}}_{\epsilon, h_\epsilon}, \hat{\rho}_{\epsilon, h_\epsilon})$ in $H^1(\Omega)^d \times L^2(\Omega)$ to $(\mathbf{u}_\epsilon, \rho_\epsilon)$ then (I) $\rightarrow 0$. Similarly due to (D1) and Corollary 3.6, we have that (II) $\rightarrow 0$. Therefore, $\|\rho_{\epsilon, h_\epsilon}\|_{L^2(\Omega)}^2 \rightarrow \|\rho_\epsilon\|_{L^2(\Omega)}^2$ and consequently $\|\rho_{\epsilon, h_\epsilon}\|_{L^2(\Omega)} \rightarrow \|\rho_\epsilon\|_{L^2(\Omega)}$.

We have shown that $\lim_{\epsilon_{h_\epsilon} \rightarrow 0} \|\rho_{\epsilon_{h_\epsilon}, h_\epsilon}\|_{L^2(\Omega)} = \|\rho_h\|_{L^2(\Omega)}$ for each h and $\lim_{h_\epsilon \rightarrow 0} \|\rho_{\epsilon, h_\epsilon}\|_{L^2(\Omega)} = \|\rho_\epsilon\|_{L^2(\Omega)}$ for each ϵ . Thus after a diagonalization argument, we may extract a subsequence (h_j, ϵ_j) , $j \in \mathbb{N}$ such that $\lim_{\epsilon_j, h_j \rightarrow 0} \|\rho_{\epsilon_j, h_j}\|_{L^2(\Omega)} = L$ for some $L \geq 0$. Dropping the subsequence subscript j , then an application of the Osgood Theorem [49, Ch. 4, Sec. 11, Th. 2'] for exchanging limits reveals that the limits $\lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|\rho_{\epsilon, h}\|_{L^2(\Omega)}$ and $\lim_{\epsilon \rightarrow 0} \lim_{h \rightarrow 0} \|\rho_{\epsilon, h}\|_{L^2(\Omega)}$ exist and

$$L = \lim_{\epsilon \rightarrow 0} \lim_{h \rightarrow 0} \|\rho_{\epsilon, h}\|_{L^2(\Omega)} = \lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|\rho_{\epsilon, h}\|_{L^2(\Omega)}. \quad (3.37)$$

By first taking the inner limit and then the outer limit in the second term and only the inner limit in the third term of (3.37), we find that

$$\|\rho\|_{L^2(\Omega)} = \lim_{h \rightarrow 0} \|\rho_h\|_{L^2(\Omega)}. \quad (3.38)$$

Since $L^2(\Omega)$ is a uniformly convex Banach space, $\rho_h \rightharpoonup \rho$ weakly in $L^2(\Omega)$ and (3.38) holds, then by the Radon–Riesz property [23, Th. A.70], we have that $\rho_h \rightarrow \rho$ strongly in $L^2(\Omega)$. Lemma 3.1 implies that $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$. \square

Corollary 3.7. *Suppose that the conditions in Theorem 3.2 hold. Then, for sufficiently small h , there exists a subsequence of global FE minimizers (\mathbf{u}_h, ρ_h) of (I-SIMP $_h^F$) that are also local FE minimizers of (SIMP $_h$).*

Proof. Corollary 3.6 and Proposition 3.6 imply that there exists a subsequence (\mathbf{u}_h, ρ_h) such that $(\mathbf{u}_h, \rho_h) \rightarrow (\mathbf{u}, \rho)$ strongly in $H^1(\Omega)^d \times L^2(\Omega)$. By the definition of strong convergence, there exists an \bar{h} such that for all $h \leq \bar{h}$, $\|\mathbf{u} - \mathbf{u}_h\|_{H^1(\Omega)} + \|\rho - \rho_h\|_{L^2(\Omega)} \leq r/4$. Thus the basin of attraction constraint is not active and the sequence of global FE minimizers of (I-SIMP $_h^F$) are also local FE minimizers of (SIMP $_h$). \square

We now have the sufficient ingredients to prove Theorem 3.2.

Proof of Theorem 3.2. By Proposition 3.4 and Corollary 3.4, there exists a sequence of global FE minimizers (\mathbf{u}_h, ρ_h) minimizing (I-SIMP $_h^F$) that satisfy $\mathbf{u}_h \rightharpoonup \mathbf{u}$ weakly in $H^1(\Omega)^d$ and $\rho_h \overset{*}{\rightharpoonup} \rho$ weakly- $*$ in $L^\infty(\Omega)$. In Corollary 3.6, we show that there exists a subsequence (up to relabelling) such that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)$. Then, in Proposition 3.6, we show that there exists a subsequence (up to relabelling) such that $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$. Finally, in Corollary 3.7, we show that a subsequence of the global FE minimizers are also local FE minimizers of (SIMP $_h$). \square

The final result we consider is the proof of Theorem 3.3 concerning the strong convergence of the filtered material distributions in $W^{1,q}(\Omega)$, $q \in [1, \infty)$. The proof of this result follows a similar pattern to the proof of Proposition 3.6. Essentially, the goal is to prove strong convergence of an ϵ -perturbed problem, then by proving convergence in the value of the norms, weak convergence, and utilizing the Radon–Riesz property, we deduce strong convergence.

Proof of Theorem 3.3. Recall that $(\mathbf{u}, F(\rho))$ is an isolated local minimizer of (F-SIMP). Let $\tilde{\rho} := F(\rho)$ and $\tilde{\rho}_h := F_h(\rho_h)$. Theorem 3.2 implies the existence of a sequence of local FE minimizers of

(SIMP_h) such that $\mathbf{u}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$ and $\rho_h \rightarrow \rho$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$. Subsequently by Lemma 3.2, $\tilde{\rho}_h \rightarrow \tilde{\rho}$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$.

Let \tilde{r} denote the radius of the basin of attraction. Denote the global minimizers of the following optimization problem by $(\mathbf{u}_\epsilon, \tilde{\rho}_\epsilon)$:

$$\begin{cases} \min_{\mathbf{v}, \tilde{\eta}} \tilde{J}(\mathbf{v}, \tilde{\eta}) + \frac{\epsilon}{q} \|\nabla \tilde{\eta}\|_{L^q(\Omega)}^q, \\ \text{where } (\mathbf{v}, \tilde{\eta}) \in (\mathbf{V}_D \cap B_{\tilde{r}/2, H^1(\Omega)}(\mathbf{u})) \times (\tilde{\mathcal{H}}_\gamma^F \cap B_{\tilde{r}/2, W^{1,q}(\Omega)}(\tilde{\rho})) \\ \text{subject to } \tilde{a}(\mathbf{u}_\epsilon, \mathbf{v}; \tilde{\rho}_\epsilon) = l(\mathbf{v}) \text{ for all } \mathbf{v} \in \mathbf{V}_D. \end{cases} \quad (\tilde{\text{I-SIMP}}_\epsilon^F)$$

The existence of a local minimizer follows by a calculus of variations argument similar to the proof of Lemma 3.5. Consider the sequence of filtered material distribution global minimizers $\tilde{\rho}_\epsilon$ as $\epsilon \rightarrow 0$. Since $\tilde{\mathcal{H}}_\gamma^F \cap B_{\tilde{r}/2, W^{1,q}(\Omega)}(\tilde{\rho})$ is a norm-closed, bounded and convex subset of the reflexive Banach space $W^{1,q}(\Omega)$, then there exists a subsequence (not relabeled) such that, as $\epsilon \rightarrow 0$,

$$\tilde{\rho}_\epsilon \rightharpoonup \tilde{\rho}_0 \in \tilde{\mathcal{H}}_\gamma^F \cap B_{\tilde{r}/2, W^{1,q}(\Omega)}(\tilde{\rho}) \text{ weakly in } W^{1,q}(\Omega). \quad (3.39)$$

Since $(\mathbf{u}, \tilde{\rho})$ is the unique minimizer of J in $(\mathbf{V}_D \cap B_{\tilde{r}/2, H^1(\Omega)}(\mathbf{u})) \times (\tilde{\mathcal{H}}_\gamma^F \cap B_{\tilde{r}/2, W^{1,q}(\Omega)}(\tilde{\rho}))$, then using the same arguments as in the proof of Lemma 3.6, we may identify $\tilde{\rho}_0 = \tilde{\rho}$ a.e.

We now wish to strengthen the convergence to strong convergence in $W^{1,q}(\Omega)$. Suppose that for any $\epsilon > 0$ under a certain threshold,

$$\|\nabla \tilde{\rho}\|_{L^q(\Omega)} < \|\nabla \tilde{\rho}_\epsilon\|_{L^q(\Omega)}. \quad (3.40)$$

This would imply that

$$\tilde{J}(\mathbf{u}, \tilde{\rho}) + \frac{\epsilon}{q} \|\nabla \tilde{\rho}\|_{L^q(\Omega)}^p < \tilde{J}(\mathbf{u}_\epsilon, \tilde{\rho}_\epsilon) + \frac{\epsilon}{q} \|\nabla \tilde{\rho}_\epsilon\|_{L^q(\Omega)}^p, \quad (3.41)$$

as $\tilde{J}(\mathbf{u}_\epsilon, \tilde{\rho}_\epsilon) \geq \tilde{J}(\mathbf{u}, \tilde{\rho})$. However, this would be a contradiction with the definition of $(\mathbf{u}_\epsilon, \tilde{\rho}_\epsilon)$. Thus for the sequence such that $\tilde{\rho}_\epsilon \rightharpoonup \tilde{\rho}$ in $W^{1,q}(\Omega)$ and thus $\tilde{J}_\epsilon(\mathbf{u}_\epsilon, \tilde{\rho}_\epsilon) \rightarrow \tilde{J}(\mathbf{u}, \tilde{\rho})$,

$$\limsup_{\epsilon \rightarrow 0} \|\nabla \tilde{\rho}_\epsilon\|_{L^q(\Omega)} \leq \|\nabla \tilde{\rho}\|_{L^q(\Omega)}. \quad (3.42)$$

Since $L^q(\Omega)$ is a uniformly convex Banach space for any $q \in (1, \infty)$, then by a generalized result of the Radon–Riesz property [23, Th. A.70], we have that $\nabla \tilde{\rho}_\epsilon \rightarrow \nabla \tilde{\rho}$ strongly in $L^q(\Omega)$. By an application of the Rellich–Kondrachov theorem, (F2), and Lemma 3.1 we deduce that $\tilde{\rho}_\epsilon \rightarrow \tilde{\rho}$ strongly in $W^{1,q}(\Omega)$. For each $h > 0$, a similar argument implies the existence of a sequence $\epsilon_h \rightarrow 0$ that satisfies $\tilde{\rho}_{\epsilon_h, h} \rightarrow \tilde{\rho}_h$ strongly in $W^{1,q}(\Omega)$.

With only small changes to the proof of Proposition 3.3 we see that, for each $\epsilon > 0$, there exists a subsequence $h_\epsilon \rightarrow 0$ such that $\|\tilde{\rho}_{\epsilon, h_\epsilon}\|_{W^{1,q}(\Omega)} \rightarrow \|\tilde{\rho}_\epsilon\|_{W^{1,q}(\Omega)}$.

We have shown that $\lim_{\epsilon_{h_\epsilon} \rightarrow 0} \|\tilde{\rho}_{\epsilon_{h_\epsilon}, h_\epsilon}\|_{W^{1,q}(\Omega)} = \|\tilde{\rho}_h\|_{W^{1,q}(\Omega)}$ for each h and $\lim_{h_\epsilon \rightarrow 0} \|\rho_{\epsilon, h_\epsilon}\|_{W^{1,q}(\Omega)} = \|\rho_\epsilon\|_{W^{1,q}(\Omega)}$ for each ϵ in their respective sequences. Thus after a diagonalization argument, we may extract a subsequence (h_j, ϵ_j) , $j \in \mathbb{N}$ such that $\lim_{\epsilon_j, h_j \rightarrow 0} \|\rho_{\epsilon_j, h_j}\|_{W^{1,q}(\Omega)} = L$ for some $L \geq 0$. Dropping the subsequence subscript j , then an application of the Osgood Theorem [49, Ch. 4, Sec. 11, Th. 2'] for exchanging limits reveals that the limits $\lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|\rho_{\epsilon, h}\|_{W^{1,q}(\Omega)}$ and $\lim_{\epsilon \rightarrow 0} \lim_{h \rightarrow 0} \|\rho_{\epsilon, h}\|_{W^{1,q}(\Omega)}$ exist and

$$L = \lim_{\epsilon \rightarrow 0} \lim_{h \rightarrow 0} \|\tilde{\rho}_{\epsilon, h}\|_{W^{1,q}(\Omega)} = \lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|\tilde{\rho}_{\epsilon, h}\|_{W^{1,q}(\Omega)}. \quad (3.43)$$

Thus by first taking the inner limit and then the outer limit in the second term and only the inner limit in the third term of (3.43), we find that

$$\|\tilde{\rho}\|_{W^{1,q}(\Omega)} = \lim_{h \rightarrow 0} \|\tilde{\rho}_h\|_{W^{1,q}(\Omega)}. \quad (3.44)$$

Since $W^{1,q}(\Omega)$ is a uniformly convex Banach space for any $q \in (1, \infty)$, $\tilde{\rho}_h \rightharpoonup \tilde{\rho}$ weakly in $W^{1,q}(\Omega)$ and (3.44) holds, then by the Radon–Riesz property [23, Th. A.70], we have that $\tilde{\rho}_h \rightarrow \tilde{\rho}$ strongly in $W^{1,q}(\Omega)$. Thus $F_h(\rho_h) \rightarrow F(\rho)$ strongly in $W^{1,q}(\Omega)$. \square

4 Conclusions and future directions

In this work we studied the convergence of a conforming FE discretization of the SIMP model for the linear elasticity compliance topology optimization problem. To ensure existence, we considered two types of regularization methods: $W^{1,p}$ -type penalty methods and density filtering. The nonconvexity of the optimization problem was handled by fixing any isolated local or global minimizer and introducing a modified optimization problem with the chosen isolated minimizer as its unique global minimizer. We then showed that there exists a sequence of discretized local FE minimizers that converges to the infinite-dimensional minimizer in the appropriate norms. The elastic displacement and material distribution converge strongly in $H^1(\Omega)^d$ and $L^p(\Omega)$, for any $p \in [1, \infty)$, respectively, for either a $W^{1,p}$ -type penalty or a density filter. If a $W^{1,p}$ -type penalty is used, we further deduced that there exists a sequence such that $\rho_h \rightarrow \rho$ strongly in $W^{1,p}(\Omega)$. In contrast, if a density filter is used then the filtered material distribution converges strongly in $W^{1,q}(\Omega)$, $q \in [1, \infty)$, provided the discretization for the filtered material distribution is $W^{1,q}(\Omega)$ -conforming. Thanks to the strong convergence, a subsequence of the local FE minimizers of the modified optimization problem are also minimizers of the original optimization problem.

We now outline some future directions.

Rates of convergence

To the best of the author's knowledge, there are currently no results that prove the rate of convergence of a sequence of local FE minimizers to an infinite-dimensional local minimizer. Numerical studies compare the convergence of local FE minimizers to a heavily refined local FE minimizer, which serves as a reference solution. Ultimately relying solely on this approach may yield misleading outcomes. It is of practical importance to know the reduction in the error when reducing the mesh size. We hope that the techniques used to prove the strong convergence results in this work will contribute significantly in this endeavour.

Adaptive mesh refinement

Topology optimization is a field that significantly benefits from adaptive mesh refinement. Intuitively, refining in regions where $0 < \rho \leq 1$ a.e. reduces the error more dramatically than in areas where $\rho = 0$ a.e. These strategies enable one to achieve smaller errors with fewer degrees of freedom, thereby cutting down the overall computational cost. Notable progress has been made in this direction cf. [48, 32, 44, 17, 46]. However, none of these cited work prove that their proposed mesh adaptivity strategies are guaranteed to converge to an infinite-dimensional local minimizer in the limit. There is a possibility that cells in the mesh, where the error is nonzero, may not meet the refinement criteria and

consequently may never be refined, hindering the convergence of the local FE minimizer in the limit. We hope that by utilizing the framework introduced in this work, one may leverage the results in [31, 29, 28] to devise mesh adaptive algorithms that provably converge in the limit.

Locking

A common criticism of the primal formulation of classical linear elasticity (without the topology optimization component) is that as the material becomes incompressible ($\lambda \rightarrow \infty$), the operator norm becomes unbounded. This can result in *locking*, the phenomenon where the approximation has sub-optimal convergence, or even appear to diverge, for large ranges of the mesh size h (before the asymptotic regime kicks in). In the primal formulation for linear elasticity, it is often possible to prove an error estimate of the form, for $\mathbf{u} \in H^s(\Omega)^d$ [3, Eq. (2.9a)]:

$$\|\mathbf{u} - \mathbf{u}_h\|_{H^1(\Omega)} \leq C(\lambda) h^{\min(k, s-1)} \|\mathbf{u}\|_{H^s(\Omega)}, \quad (4.1)$$

where k denotes the polynomial degree of the FE discretization of \mathbf{u} . Moreover $C(\lambda) \rightarrow \infty$ as $\lambda \rightarrow \infty$, which corresponds to the incompressible limit. Asymptotically, the error should reduce at rate of $h^{\min(k, s-1)}$. However, if $C(\lambda)$ is extremely large, the constant will dominate the error resulting in locking. Locking can be alleviated by using a high-order element for \mathbf{u} , e.g. quartics or higher [3] or by using alternatives to the primal formulation such as the pressure and symmetric stress formulations [14, Ch. VI.3].

The appearance of locking is particularly troublesome when considering the SIMP model. The primal formulation with low-order elements is used in almost all implementations of the SIMP model, e.g. [6]. Moreover, since there are almost no known formulae for local minimizers of SIMP models, there are few comprehensive studies on FE convergence. We conjecture that locking occurs near the incompressible limit in SIMP models and, consequently, the computed discretized minimizers have a high error even for small mesh sizes.

Rigorously describing locking is challenging and the results of [3] and [14, Ch. VI.3] do not (immediately) extend to the SIMP model. Thus any proof of non-locking formulations was beyond the scope of this work. Nevertheless, the new convergence results presented in this study aim to lay some groundwork for addressing the locking issue in future research. A comprehensive mathematical understanding of locking in the SIMP model would have significant implications for future implementations.

A Supplementary proofs

Proof of Proposition 2.1. Since $\rho \in \mathcal{H}_\gamma^F$ and thus $f, \rho \geq 0$ a.e. then $F(\rho)(x) \geq 0$ for all $x \in \Omega$. Moreover,

$$|F(\rho)(x)| = \left| \int_{\Omega} f(x-y)\rho(y) dy \right| \leq \|f\|_{L^1(\mathbb{R}^d)} \|\rho\|_{L^\infty(\Omega)} \leq 1. \quad (\text{A.1})$$

Thus (F2) is satisfied. Moreover, (A.1) implies that

$$\|F(\rho)\|_{L^\infty(\Omega)} = \left\| \int_{\Omega} f(x-y)\rho(y) dy \right\|_{L^\infty(\Omega)} \leq \|f\|_{L^\infty(\mathbb{R}^d)} \leq C < \infty. \quad (\text{A.2})$$

It can be shown that $h(x, y) := f(x - y)\rho(y)$ satisfies the condition for the Leibniz integral rule and, therefore,

$$\|\nabla F(\rho)\|_{L^\infty(\Omega)} = \left\| \int_{\Omega} \nabla_x [f(x - y)]\rho(y) dy \right\|_{L^\infty(\Omega)} \leq \|\nabla f\|_{L^\infty(\mathbb{R}^d)} \leq C < \infty. \quad (\text{A.3})$$

This implies that $F(\rho) \in W^{1,\infty}(\Omega)$. Hence, (F1) is satisfied.

Consider a weakly-* converging sequence $\rho_n \xrightarrow{*} \rho \in \mathcal{H}_\gamma^F$ weakly-* in $L^\infty(\Omega)$. This defines a bounded sequence of Lipschitz functions $F(\rho_n)$. Thus, by the Arzelà–Ascoli theorem [2, Th. 1.33], there exists a limit $F_0 \in L^\infty(\Omega)$ such that a subsequence (up to relabelling) satisfies $F(\rho_n) \rightarrow F_0$ uniformly in Ω . Hence, $F(\rho_n) \rightarrow F_0$ strongly in $L^\infty(\Omega)$. Moreover,

$$F(\rho_n)(x) = \int_{\Omega} f(x - y)\rho_n(y) dy \rightarrow \int_{\Omega} f(x - y)\rho(y) dy = F(\rho)(x). \quad (\text{A.4})$$

Hence $F(\rho_n) \rightarrow F(\rho)$ pointwise. Thus, by the uniqueness of limits, $F_0 = F(\rho)$ and $F(\rho_n) \rightarrow F(\rho)$ strongly in $L^\infty(\Omega)$. A similar argument reveals that if $\rho_n \rightharpoonup \rho$ weakly in $L^s(\Omega)$, $s \in [1, \infty)$, then $F(\rho_n) \rightarrow F(\rho)$ strongly in $L^s(\Omega)$. Thus (F3) is satisfied.

Note that $F(\rho) + t(F(\rho) - F(\eta)) = F(\rho + t(\eta - \rho)) \in \tilde{\mathcal{H}}_\gamma^F$ since $\rho + t(\eta - \rho) \in \mathcal{H}_\gamma^F$. Hence $\tilde{\mathcal{H}}_\gamma^F$ is a convex space. Consider any sequence such that $F(\rho_n) \rightarrow F_0$ strongly in $W^{1,s}(\Omega)$ for any $s \in [1, \infty]$. This defines a bounded sequence (ρ_n) in $L^\infty(\Omega)$. Hence, since \mathcal{H}_γ^F is a norm-closed, bounded, and convex subset of $L^\infty(\Omega)$, then by the Banach–Alaoglu theorem, there exists an $\rho_0 \in \mathcal{H}_\gamma^F$ such that a subsequence (up to relabelling) satisfies $\rho_n \xrightarrow{*} \rho_0$ weakly-* in $L^\infty(\Omega)$. As previously shown, this implies that $F(\rho_n) \rightarrow F(\rho_0)$ strongly in $L^\infty(\Omega)$. Hence, $F_0 = F(\rho_0)$ and, therefore, $F_0 \in \tilde{\mathcal{H}}_\gamma^F$. Hence, $\tilde{\mathcal{H}}_\gamma^F$ is a norm-closed, bounded, and convex subset of $W^{1,\infty}(\Omega)$. A similar argument reveals that $\tilde{\mathcal{H}}_\gamma^F$ is a norm-closed, bounded, and convex subset of $W^{1,s}(\Omega)$ for any $s \in [1, \infty]$. Therefore, (F4) is satisfied. \square

Proof of Lemma 3.1. By Hölder's inequality $\|\eta - \eta_j\|_{L^1(\Omega)} \leq |\Omega|^{1/p} \|\eta - \eta_j\|_{L^p(\Omega)}$. Thus $\eta_j \rightarrow \eta$ strongly in $L^1(\Omega)$. Therefore, for any $s \in (1, \infty)$,

$$\int_{\Omega} |\eta - \eta_j|^s dx = \int_{\Omega} |\eta - \eta_j|^{s-1} |\eta - \eta_j| dx \leq 1^{s-1} \|\eta - \eta_j\|_{L^1(\Omega)}, \quad (\text{A.5})$$

which implies that $\eta_j \rightarrow \eta$ strongly in $L^s(\Omega)$ for any $s \in [1, \infty)$. \square

Proof of Lemma 3.2. Fix an $\eta \in L^q(\Omega)$ such that $\eta_h \rightharpoonup \eta \in \mathcal{H}_\gamma^F$ weakly in $L^q(\Omega)$ ($\eta_h \xrightarrow{*} \eta$ weakly-* in $L^\infty(\Omega)$ if $q = \infty$) for $q \in [1, \infty]$. Then,

$$\|F(\eta) - F_h(\eta_h)\|_{L^q(\Omega)} \leq \|F(\eta) - F_h(\eta)\|_{L^q(\Omega)} + \|F_h(\eta) - F_h(\eta_h)\|_{L^q(\Omega)}. \quad (\text{A.6})$$

The first term on the right-hand side of (A.6) tends to zero thanks to assumption (F7). For the second term we note that, by assumptions (F6) and (F5),

$$\|F_h(\eta) - F_h(\eta_h)\|_{L^q(\Omega)} \leq C \|F(\eta) - F(\eta_h)\|_{L^q(\Omega)} \quad (\text{A.7})$$

for a constant $C < \infty$ independent of h . The right-hand side of (A.7) tends to zero thanks to (F3). Hence $\eta_h \rightharpoonup \eta$ weakly in $L^q(\Omega)$ ($\eta_h \xrightarrow{*} \eta$ weakly-* in $L^\infty(\Omega)$ if $q = \infty$) implies that $F_h(\eta_h) \rightarrow F(\eta)$ strongly in $L^q(\Omega)$. \square

Proof of Lemma 3.3. We prove the result for (Case 1) and note that (Case 2) follows with only a couple of small modifications. For any $\mathbf{v} \in \mathbf{V}_D$ and $\mathbf{v}_h \in \mathbf{V}_{D,h}$, consider

$$\begin{aligned} a_h(\hat{\mathbf{u}}_h, \mathbf{v}; \hat{\rho}_h) &= a_h(\hat{\mathbf{u}}_h, \mathbf{v}_h; \hat{\rho}_h) + a_h(\hat{\mathbf{u}}_h, \mathbf{v} - \mathbf{v}_h; \hat{\rho}_h) \\ &= l_h(\mathbf{v}_h) + a_h(\hat{\mathbf{u}}_h, \mathbf{v} - \mathbf{v}_h; \hat{\rho}_h). \end{aligned} \quad (\text{A.8})$$

By (D2) we may choose \mathbf{v}_h such that $\mathbf{v}_h \rightarrow \mathbf{v}$ strongly in $H^1(\Omega)^d$. Thus

$$l_h(\mathbf{v}_h) \rightarrow l(\mathbf{v}). \quad (\text{A.9})$$

due to the linearity of $l(\cdot)$, (D1), and (D2). Moreover,

$$|a_h(\hat{\mathbf{u}}_h, \mathbf{v} - \mathbf{v}_h; \hat{\rho}_h)| \leq \|k(\hat{\rho}_h)\|_{L^\infty(\Omega)} \|\mathbf{E}\hat{\mathbf{u}}_h\|_{L^2(\Omega)} \|\mathbf{E}(\mathbf{v} - \mathbf{v}_h)\|_{L^2(\Omega)} \rightarrow 0, \quad (\text{A.10})$$

thanks to the boundedness of $k(\hat{\rho}_h)$, $\mathbf{E}\hat{\mathbf{u}}_h$ and (D2). The boundedness of $\mathbf{E}\hat{\mathbf{u}}_h$ follows from the a priori estimates in the PDE of (SIMP_h). Hence,

$$a_h(\hat{\mathbf{u}}_h, \mathbf{v}; \hat{\rho}_h) \rightarrow l(\mathbf{v}). \quad (\text{A.11})$$

We note that

$$\begin{aligned} &|a_h(\hat{\mathbf{u}}_h, \mathbf{v}; \hat{\rho}_h) - a(\hat{\mathbf{u}}, \mathbf{v}; \hat{\rho})| \\ &\leq |a(\hat{\mathbf{u}}_h - \hat{\mathbf{u}}, \mathbf{v}; \hat{\rho})| + |a_h(\hat{\mathbf{u}}_h, \mathbf{v}; \hat{\rho}_h) - a(\hat{\mathbf{u}}_h, \mathbf{v}; \hat{\rho})|. \end{aligned} \quad (\text{A.12})$$

The first term on the right-hand side tends to zero thanks to weak convergence of $\hat{\mathbf{u}}_h$ to $\hat{\mathbf{u}}$. For the second term we see that

$$|a_h(\hat{\mathbf{u}}_h, \mathbf{v}; \hat{\rho}_h) - a(\hat{\mathbf{u}}_h, \mathbf{v}; \hat{\rho})| \leq \|(k(\hat{\rho}_h) - k(\hat{\rho}))\mathbf{E}\mathbf{v}\|_{L^2(\Omega)} \|\mathbf{E}\hat{\mathbf{u}}_h\|_{L^2(\Omega)}. \quad (\text{A.13})$$

By assumption $\hat{\rho}_h \rightharpoonup \hat{\rho}$ in $W^{1,q}(\Omega)$ for some $q \in [2, \infty)$ and, by the Rellich–Kondrachov theorem and Lemma 3.1, $\hat{\rho}_h \rightarrow \hat{\rho}$ strongly in $L^2(\Omega)$. Thus by utilizing Lemma 3.1, for any smooth test function $\mathbf{v} \in C_c^\infty(\Omega)^d$,

$$a_h(\hat{\mathbf{u}}_h, \mathbf{v}; \hat{\rho}_h) \rightarrow a(\hat{\mathbf{u}}, \mathbf{v}; \hat{\rho}). \quad (\text{A.14})$$

(A.14) follows for any $\mathbf{v} \in \mathbf{V}_D$ by the density of smooth and compactly supported functions in \mathbf{V}_D . Hence, (A.11) and (A.14) imply that, for any $\mathbf{v} \in \mathbf{V}_D$,

$$a(\hat{\mathbf{u}}, \mathbf{v}; \hat{\rho}) = l(\mathbf{v}). \quad (\text{A.15})$$

□

Proof of Lemma 3.4. We prove the result for (Case 1) and note that (Case 2) follows with only a couple of small modifications. The first step is to show that $\hat{\mathbf{u}}_h \rightharpoonup \mathbf{u}$ in $H^1(\Omega)^d$. The sequence generated by $\hat{\rho}_h$ defines a bounded sequence of $\hat{\mathbf{u}}_h$ as $h \rightarrow 0$. Since \mathbf{V}_D is a reflexive space, then there exists a limit $\hat{\mathbf{u}} \in \mathbf{V}_D$ and a subsequence (up to relabelling) such that $\hat{\mathbf{u}}_h \rightharpoonup \hat{\mathbf{u}} \in \mathbf{V}_D$. We must now identify $\hat{\mathbf{u}}$ with \mathbf{u} .

The requirements of Lemma 3.3 are satisfied and thus, for any $\mathbf{v} \in \mathbf{V}_D$, $a(\hat{\mathbf{u}}, \mathbf{v}; \rho) = l(\mathbf{v})$. Due to the uniqueness of limits and the uniqueness of solutions of the linear elasticity problem with a fixed ρ (Proposition 2.2), we deduce that $\hat{\mathbf{u}} = \mathbf{u}$ a.e. and $\hat{\mathbf{u}}_h \rightharpoonup \mathbf{u}$ in $H^1(\Omega)^d$.

We now strengthen the convergence to strong convergence. By Korn's inequality [7, Th. 11], and the assumption that $\epsilon_{\text{SIMP}} > 0$, there exists a $c > 0$ such that

$$\begin{aligned} c\|\mathbf{u} - \hat{\mathbf{u}}_h\|_{H^1(\Omega)}^2 &\leq a_h(\mathbf{u} - \hat{\mathbf{u}}_h, \mathbf{u} - \hat{\mathbf{u}}_h; \hat{\rho}_h) \\ &= a_h(\mathbf{u} - \hat{\mathbf{u}}_h, \mathbf{u}; \hat{\rho}_h) + a_h(\hat{\mathbf{u}}_h, \hat{\mathbf{u}}_h; \hat{\rho}_h) - a_h(\mathbf{u}, \hat{\mathbf{u}}_h; \hat{\rho}_h) \\ &= a_h(\mathbf{u} - \hat{\mathbf{u}}_h, \mathbf{u}; \hat{\rho}_h) + l_h(\hat{\mathbf{u}}_h) - a_h(\mathbf{u}, \hat{\mathbf{u}}_h; \hat{\rho}_h). \end{aligned} \quad (\text{A.16})$$

The first term on the rightmost-hand side of (A.16) is bounded above as follows

$$\begin{aligned} a_h(\mathbf{u} - \hat{\mathbf{u}}_h, \mathbf{u}; \hat{\rho}_h) \\ \leq a(\mathbf{u} - \hat{\mathbf{u}}_h, \mathbf{u}; \rho) + \|k(\rho) - k(\rho_h)\|_{L^2(\Omega)} \|\mathbf{E}(\mathbf{u} - \hat{\mathbf{u}}_h)\|_{L^2(\Omega)} \|\mathbf{E}\mathbf{u}\|_{L^2(\Omega)}. \end{aligned} \quad (\text{A.17})$$

The first term tends to zero thanks to $\hat{\mathbf{u}}_h \rightharpoonup \mathbf{u}$ weakly in $H^1(\Omega)^d$. The second term converges to zero since $\rho_h \rightarrow \rho$ strongly in $L^q(\Omega)$ for some $q \in [2, \infty]$ (and thus for any $q \in [1, \infty)$ by Lemma 3.1).

The second term on the rightmost-hand side of (A.16) tends to $l(\mathbf{u})$ thanks to (D1) and the weak convergence of $\hat{\mathbf{u}}_h$ in $H^1(\Omega)^d$. By the same arguments as (A.8)–(A.10), we note that $a_h(\mathbf{u}, \hat{\mathbf{u}}_h; \hat{\rho}_h) \rightarrow l(\mathbf{u})$. Hence, we conclude that $\hat{\mathbf{u}}_h \rightarrow \mathbf{u}$ strongly in $H^1(\Omega)^d$. \square