Boiger and Carstensen *Advanced Modeling and Simulation in Engineering Sciences* 2013, **1**:5
http://www.amses-journal.com/content/1/1/5

Advanced Modeling and Simulation in Engineering Sciences
a SpringerOpen Journal

RESEARCH ARTICLE Open Access

A posteriori error analysis of stabilised FEM for degenerate convex minimisation problems under weak regularity assumptions

Wolfgang Boiger[1,2] and Carsten Carstensen[1,2*]

*Correspondence:
cc@math.hu-berlin.de
[1]Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099, Berlin, Germany
[2]Department of Computational Science and Engineering, Yonsei University, Unter den Linden 6, 120-749, Seoul, Korea

Abstract

Background: The discretisation of degenerate convex minimisation problems experiences numerical difficulties with a singular or nearly singular Hessian matrix.

Methods: Some discrete analog of the surface energy in microstrucures is added to the energy functional to define a stabilisation technique.

Results: This paper proves (a) strong convergence of the stress even without any smoothness assumption for a class of stabilised degenerate convex minimisation problems. Given the limitted a priori error control in those cases, the sharp a posteriori error control is of even higher relevance. This paper derives (b) guaranteed a posteriori error control via some equilibration technique which does not rely on the strict Galerkin orthogonality of the unperturbed problem. In the presence of L^2 control in the original minimisation problem, some realistic model scenario with piecewise smooth exact solution allows for strong convergence of the gradients plus refined a posteriori error estimates. This paper presents (c) an improved a posteriori error control in this interface problem and so narrows the efficiency reliability gap.

Conclusions: Numerical experiments illustrate the theoretical convergence rates for uniform and adaptive mesh-refinements and the improved a posteriori error control for four benchmark examples in the computational microstructures.

Keywords: Adaptive finite element method; Relaxation; Convexification; Calculus of variations; Degenerate convex problems; Energy reduction; Nonconvex minimisation; Partial differential equation; Stabilisation; Strong convergence; A posteriori error estimate; Reliability-efficiency gap; Euler-Lagrange equation; Guaranteed upper bound

Background

Infimising sequences of variational problems with non-quasiconvex energy densities, in general, develop finer and finer oscillations with no classical limit in Sobolev function spaces called microstructure [1-6]. Those oscillations cause difficulty to numerical methods because fine grids are necessary to resolve such oscillations which results in ineffective and tricky mesh-depending computations. Strong convergence of gradients of infimising sequences of the non-quasiconvex problem is impossible.

Relaxation techniques replace the nonconvex energy density by its (semi-)convex hull and lead to a macroscopic model. Since the convexified energy density obtained by this

method, in general, lacks strict convexity, numerical algorithms might encounter situations where the Hessian matrix is singular. For instance, the Newton minimisation algorithm fails on the convexified three-well problem of Subsection 'Three-well benchmark' below. Applications of relaxation techniques include models in computational microstructure [5-7], some optimal design problems [8,9], the nonlinear Laplacian [10] (where the Hessian can become arbitrarily ill-conditioned in spite of its strict convexity) and elastoplasticity [1].

Stabilisation techniques regularise the energy term by an additional positive semidefinite stabilisation function. The paper [11] discusses several choices of such stabilisation functions for P_1 conforming finite elements and quasiuniform meshes. It turns out that stabilisation can ensure strong convergence of the strain approximations under particular circumstances. A particular stabilisation in [12] leads to strong convergence even on unstructured grids but is still restricted to unrealistically smooth solutions. This paper studies the stabilisation technique of [12] and addresses the question of convergence (i) without extra regularity assumptions, (ii) in a realistic scenario called model interface problem, and (iii) establishes an a posteriori error control.

The stabilisation leads to improved condition numbers of the Hessian matrix and to reduced errors if the numerical solvers fail without stabilisation. Figure 1 shows the convergence of the discrete stress σ_ℓ of the three-well benchmark corresponding to the discrete minimisers of the energy $E_\ell(v_\ell) = E(v_\ell) + C/2\|\|v_\ell\|\|_\ell^2$. The errors are plotted for computations with uniform mesh refinements with various solver tolerances in the discrete minimisation procedure at a fixed triangulation and values of C, cf. Section 'Numerical experiments' for details on the MATLAB implementation. Without stabilisation, the convergence stagnates with a moderate tolerance of 10^{-5} which becomes visible as a "plateau" in Figure 1. The Newton solver even aborts prematurely due to the singular Hessian. In conclusion, stabilisation enables higher accuracies in numerical examples.

For $\beta \geqslant 0$ the convex energy functional assumes the form

$$E(v) := \int_\Omega \big(W(\mathrm{D}v(x)) + \beta|v(x) - g(x)|^2 - f(x)\cdot v(x)\big)\mathrm{d}x. \tag{1.1}$$

Assume that W is convex with quadratic growth so that there exist minimisers $u \in H^1_0(\Omega)$; below p-th order growth is included while $p = 2$ throughout this simplifying

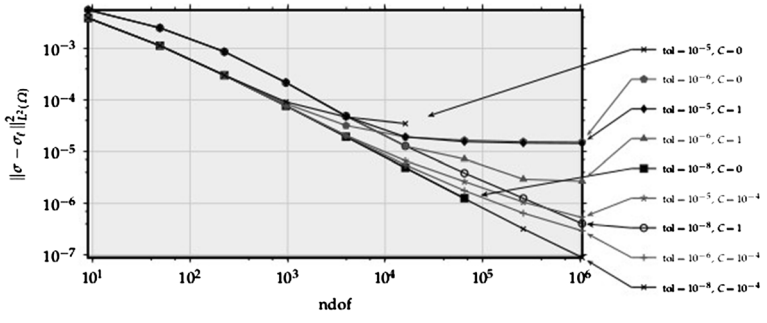

Figure 1 Impact of the stabilisation on the error. Error of the stabilised stress σ_ℓ with coefficients $C = 0, 10^{-4}, 1$ of the stabilisation and various tolerances tol $= 10^{-5}, 10^{-6}, 10^{-8}$ of the Newton solver.

introduction. Given a sequence of shape-regular triangulations $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ [13], let u_ℓ minimise the stabilised discrete energy

$$E_\ell(v_\ell) := E(v_\ell) + \frac{1}{2}|||v_\ell|||_\ell^2 \text{ with } |||v_\ell|||_\ell^2 := H_\ell^2 \sum_{F\in\mathcal{F}_\ell(\Omega)} h_F^{-1} || [\mathrm{D}v_\ell]_F ||_{L^2(F)}^2$$

amongst all conforming P_1 finite element functions v_ℓ on $\mathcal{T}_\ell$, where $[\mathrm{D}v_\ell]_F$ is the jump of the gradient $\mathrm{D}v_\ell$ along the interior side F, written $F \in \mathcal{F}_\ell(\Omega)$, and $H_\ell := \max_T h_T$ is the maximal diameter h_T of all simplices $T \in \mathcal{T}_\ell$.

Section 'Global convergence' verifies the strong convergence of the discrete solution u_ℓ and its stress $\sigma_\ell := \mathrm{D}W(\mathrm{D}u_\ell)$ to their respective continuous conterparts,

$$||\sigma - \sigma_\ell||_{L^2(\Omega)}^2 + \beta||u - u_\ell||_{L^2(\Omega)}^2 + |||u_\ell|||_\ell^2 \to 0 \text{ as } \ell \to \infty.$$

Section 'A posteriori error estimates' presents a novel application of [14-17] to nonlinear problems. For the L^2 projection Π_ℓ onto the space of piecewise P_0 functions, any Raviart-Thomas function $\tau_\ell \in RT_0(\mathcal{T}_\ell)$ satisfies

$$\begin{aligned}&||\sigma - \sigma_\ell||_{L^2(\Omega)}^2\\ &\quad\lesssim \big(||\sigma_\ell - \tau_\ell||_{L^2(\Omega)} + ||\Pi_\ell\Lambda_\ell + \mathrm{div}\ \tau_\ell||_{L^2(\Omega)} + \mathrm{osc}_{\ell,2}(\Lambda_\ell)\big)\, ||u - u_\ell||_{H^1(\Omega)}.\end{aligned}$$

This error bound holds for any discrete displacement u_ℓ that satisfies the boundary conditions; the point is that inexact solve is included — there is no Galerkin orthogonality required. The drawback is to minimise the expression on the right-hand side with respect to τ_ℓ in order to obtain a sharp error bound. This is a particular selection: degenerate convex minimisation problems do not allow for a control of $||u - u_\ell||_{H^1(\Omega)}$ and may even face multiple exact or discrete solutions while the discrete minimum of E_ℓ is unique. However, in some results of this paper, either W or the lower-order terms lead to some control over $||u - u_\ell||_{H^2(\Omega)}$ and the selection via stabilisation is correct.

Phase transition problems motivate the investigation of scenarios with a smooth solution u up to a one-dimensional interface $\Gamma \subset \overline{\Omega}$ [18]. Section 'Refined analysis for an interface model problem' proves that such problems allow even for strong convergence of the gradients for any unique solution u in $W^{1,\infty}(\Omega)\cap H^2(\Omega\setminus\Gamma)$ [19]. This result also leads to an improvement of the a posteriori error control of the discrete stresses and narrows the efficiency-reliability gap; the efficiency-reliability gap is the difference of the convergence rates of the guaranteed upper a posteriori error bound and the guaranteed lower a posteriori error bound.

Section 'Numerical experiments' complements the theoretical findings with numerical experiments to provide empirical evidence of the improved error control. The stabilisation technique competes in four benchmark examples, with and without known exact solution, for uniform and two different mesh-refining algorithms for the explicit residual-based error estimator of [7] and with an averaging-type error estimator of ([18], (1.11)).

Standard notation on Lebesgue and Sobolev spaces is employed throughout this paper and $a \lesssim b$ abbreviates $a \leqslant Cb$ with some generic constant $0 < C < \infty$ independent of crucial parameters (like the mesh-size on level ℓ); $a \approx b$ means $a \lesssim b \lesssim a$. Furthermore, $A : B$ abbreviates the matrix inner product that corresponds to the Frobenius norm.

Methods: Discretisation and Stabilisation

Based on the convergence results for unstructured grids, this paper will develop reliable error estimators for a class of stabilised convex minimisation problems described in the sequel. Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipshitz domain with polygonal boundary for $n = 2$ or 3. Given a continuous convex energy density $W : \mathbb{R}^{m\times n} \to \mathbb{R}$, $g,f \in L^2(\Omega;\mathbb{R}^m)$, $\beta \geqslant 0$, and $v \in W^{1,p}(\Omega;\mathbb{R}^m)$ with $2 \leqslant p < \infty$ and $m = 1,\dots,n$, the energy is given by (1.1).

Throughout this paper, the energy density $W \in \mathcal{C}^1(\mathbb{R}^{m\times n};\mathbb{R})$ satisfies (2.1)–(2.2) for parameters $1 < r \leqslant 2$, $0 \leqslant s < \infty$ and $s + r + p \leqslant rp$. The *two-sided growth condition* reads

$$|F|^p - 1 \lesssim W(F) \lesssim |F|^p + 1 \quad \text{for all } F \in \mathbb{R}^{m\times n}. \tag{2.1}$$

The *convexity control* assumption reads, for all $F_1, F_2 \in \mathbb{R}^{m\times n}$,

$$|\mathrm{D}\ W(F_1) - \mathrm{D}\ W(F_2)|^r \lesssim \big(1 + |F_1|^s + |F_2|^s\big)(\mathrm{D}\ W(F_1) - \mathrm{D}\ W(F_2)) : (F_1 - F_2). \tag{2.2}$$

The proof of Theorem 2 in [7] shows that (2.2) is crucial for the uniqueness of the stress tensor $\mathrm{D}W(\mathrm{D}u)$.

Given Dirichlet data $u_D \in W^{2,p}(\Omega;\mathbb{R}^m) \cap H^2(\partial\Omega;\mathbb{R}^m)$ for the set of admissible functions $\mathcal{A} := u_D + V := u_D + W_0^{1,p}(\Omega;\mathbb{R}^m)$, the continuous (convex) model problem reads

$$\text{minimise } E(v) \text{ within } v \in \mathcal{A}. \tag{2.3}$$

A finite element approximation of (2.3) is based on a family of regular triangulations $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ of the domain Ω into simplices in the sense of Ciarlet [13] (e.g., for $n = 2$, two non-disjoint triangles of $\mathcal{T}_\ell$ share either a common edge or a common node). The set of sides $\mathcal{F}_\ell$ consists of edges (for $n = 2$) or faces (for $n = 3$) of $\mathcal{T}_\ell$ and is split into the union of the sets of all interiour sides $\mathcal{F}_\ell(\Omega)$ and of all boundary sides $\mathcal{F}_\ell(\partial\Omega)$.

For latter reference, define the diameter $h_T := \mathrm{diam}T$ of a triangle (or tetrahedron) $T \in \mathcal{T}_\ell$ and the size $h_F := \mathrm{diam}F$ of a side $F \in \mathcal{F}_\ell$. The *mesh size function* $h_\ell : \Omega \to \mathbb{R}_{>0}$ is given by

$$h_\ell(x) := \begin{cases} h_T & \text{for } x \in \operatorname{int} T \in \mathcal{T}_\ell, \\ \min\{h_F : F \in \mathcal{F}_\ell \text{ and } x \in F\} & \text{otherwise.} \end{cases}$$

The global mesh size will be abbreviated by $H_\ell := \|h_\ell\|_{L^\infty(\Omega)}$. We presume the family $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ to be *shape-regular* so that $h_F \approx h_T$ for all $T \in \mathcal{T}_\ell$, $F \in \mathcal{F}_\ell$ and $F \subset T$.

The space of $\mathcal{T}_\ell$-piecewise polynomials of degree $\leqslant k \in \mathbb{N}_0$ is $P_k(\mathcal{T}_\ell)$. The nodal interpolation $I_\ell w \in P_1(\mathcal{T}_\ell) \cap C(\overline{\Omega})$ of $w \in C(\overline{\Omega})$ is given by $I_\ell w(z) = w(z)$ for all nodes z. Let furthermore $\Pi_\ell w$ be the L^2 projection of $w \in L^2(\Omega)$ onto $P_0(\mathcal{T}_\ell)$, and $\mathrm{osc}_{\ell,q}(w) := \|h_\ell(\mathrm{id} - \Pi_\ell)w\|_{L^q(\Omega)}$ be the oscillation of $w \in L^q(\Omega)$ for $2 \leqslant q \leqslant \infty$ with respect to the triangulation $\mathcal{T}_\ell$. The symbol id denotes the identity operator. Let $u_{D,\ell} = I_\ell u_D$, and

$$\mathcal{A}_\ell := u_{D,\ell} + V_\ell \quad \text{with } V_\ell := V \cap P_1(\mathcal{T}_\ell;\mathbb{R}^m) \cap C(\overline{\Omega}).$$

Given a function v on Ω which is possibly discontinuous along some side $F \in \mathcal{F}_\ell(\Omega)$ shared by the two elements $T_\pm$ such that there exist traces from either sides, the *jump* of v along F reads

$$[v]\,(x) = [v]_F\,(x) := \lim_{T_+ \ni y \to x} v(y) - \lim_{T_- \ni y \to x} v(y) \quad \text{for } x \in F.$$

The stabilisation of [12] will be used throughout this paper with $-1 < \gamma < \infty$ and

$$a_\ell(v,w) := \sum_{F \in \mathcal{F}_\ell(\Omega)} \frac{H_\ell^{1+\gamma}}{h_F} \int_F [\mathrm{D}v]_F : [\mathrm{D}w]_F \, \mathrm{d}s \quad \text{and} \quad |||v|||_\ell^2 := a_\ell(v,v). \tag{2.4}$$

The stabilised discrete problem reads

$$\text{minimise } E_\ell(v) := E(v) + \frac{1}{2} a_\ell(v,v) \text{ amongst } v \in \mathcal{A}_\ell. \tag{2.5}$$

Convergence of gradients with a guaranteed convergence rate is shown in [12] under unrealistically high regularity assumptions. A comprehensive collection of the results in [12] is summarised in the following theorem.

Theorem 2.1. *([12]) Let $u \in \mathcal{A} \cap H^{3/2+\varepsilon}(\Omega;\mathbb{R}^m)$ be some solution of (2.3) for some $\varepsilon > 0$; let p' and r' be the Hölder conjugate of p and r, $-1 < \gamma < 3$, and set*

$$\zeta := \min\left\{1+\gamma\,,\, r'\right\} \text{ for } \beta > 0 \text{ and } \zeta := \min\{1+\gamma\,,\,2\} \text{ for } \beta = 0.$$

Then the discrete solution $u_\ell \in \mathcal{A}_\ell$ of (2.5) and the continuous and discrete stress $\sigma := \mathrm{D}W(\mathrm{D}u) \in L^{p'}(\Omega;\mathbb{R}^{m\times n})$ and $\sigma_\ell := \mathrm{D}W(\mathrm{D}u_\ell) \in P_0(\mathcal{T}_\ell;\mathbb{R}^{m\times n})$ satisfy

$$\|\sigma - \sigma_\ell\|_{L^{p'}(\Omega)}^r + \|u - u_\ell\|_{L^2(\Omega)}^2 + |||u_\ell|||_\ell^2 + H_\ell^{(1+\gamma)/2} \|D(u-u_\ell)\|_{L^2(\Omega)}^2 \lesssim H_\ell^\zeta.$$

Proof. This combines Lemma 3.5 and 4.1–4.2 plus Theorem 3.8 and 4.4 in [12]. □

Global convergence

This section is devoted to the proof of a general convergence result *without* higher regularity assumptions. Let $u \in \mathcal{A}$ and $u_\ell \in \mathcal{A}_\ell$ solve the minimisation problem (2.3) and (2.5) and set $\sigma := \mathrm{D}W(\mathrm{D}u)$ and $\sigma_\ell := \mathrm{D}W(\mathrm{D}u_\ell)$. For the unstabilised approximation, the a priori error estimates of [7] plus a density argument prove convergence of

$$\|\sigma - \sigma_\ell\|_{L^{p'}(\Omega)}^r + \beta \|u - u_\ell\|_{L^2(\Omega)}^2 \to 0 \text{ as } H_\ell \to 0.$$

Note that $\beta = 0$ is permitted. Then, however, uniqueness of u and convergence of $\|u - u_\ell\|_{L^2(\Omega)}^2$ are guaranteed. The point in the following result is that the stabilised approximation converges as well as $|||u_\ell|||_\ell \to 0$ even for non-smooth or non-unique minimisers. Under special circumstances, uniqueness of u and the convergence $\|u - u_\ell\|_{L^2(\Omega)} \to 0$ can be shown even for $\beta = 0$, e.g., in Example 3.3.

Theorem 3.1. *(*Global Convergence*) Provided* $\lim_{\ell\to\infty} H_\ell = 0$ *it holds*

$$\|\sigma - \sigma_\ell\|_{L^{p'}(\Omega)}^r + \beta \|u - u_\ell\|_{L^2(\Omega)}^2 + |||u_\ell|||_\ell^2 \to 0 \text{ as } \ell \to \infty.$$

The proof is based on the following lemma.

Lemma 3.2. *The errors $\delta_\ell := \sigma - \sigma_\ell$ and $e_\ell := u - u_\ell$ satisfy, for all $v_\ell \in V_\ell$, that*

$$\|\delta_\ell\|_{L^{p'}(\Omega)}^r + \beta \|e_\ell\|_{L^2(\Omega)}^2 \lesssim |e_\ell - v_\ell|_{W^{1,p}(\Omega)}^{r'} + \beta \|e_\ell - v_\ell|_{L^2(\Omega)}^2 + a_\ell(u_\ell, v_\ell).$$

Proof. The minimisation problems (2.3) and (2.5) are equivalent to their respective Euler-Lagrange equations, namely for $v \in V$ and $v_\ell \in V_\ell$,

$$\int_\Omega \big(\sigma(x) : \mathrm{D}v(x) + 2\beta(u(x) - g(x)) \cdot v(x) - f(x) \cdot v(x)\big)\mathrm{d}x = 0; \tag{3.1}$$

$$\int_\Omega \big(\sigma_\ell(x) : \mathrm{D}v_\ell(x) + 2\beta(u_\ell(x) - g(x)) \cdot v_\ell(x) - f(x) \cdot v_\ell(x)\big)\mathrm{d}x + a_\ell(u_\ell, v_\ell) = 0. \tag{3.2}$$

Algebraic transformations of the difference of these two equations lead to

$$\int_\Omega \delta_\ell : \mathrm{D}e_\ell \mathrm{d}x + 2\beta\|e_\ell\|^2_{L^2(\Omega)} = \int_\Omega (\delta_\ell : \mathrm{D}(e_\ell - v_\ell) + 2\beta e_\ell \cdot (e_\ell - v_\ell))\mathrm{d}x + a_\ell(u_\ell, v_\ell).$$

It is shown in ([12], Lemma 3.5) that

$$\|\delta_\ell\|^r_{L^{p'}(\Omega)} \lesssim \int_\Omega \delta_\ell : \mathrm{D}e_\ell \mathrm{d}x. \tag{3.3}$$

Two Hölder inequalities on the right-hand side and absorbtions of $\|\delta_\ell\|_{L^{p'}(\Omega)}$ and $\|e_\ell\|_{L^2(\Omega)}$ eventually conclude the proof. Further details are dropped for brevity. □

Proof of Theorem 3.1. Given any positive ε, the density of smooth functions in $W^{1,p}_0(\Omega;\mathbb{R}^m)$ leads to some $v_\varepsilon \in \mathcal{D}(\Omega;\mathbb{R}^m)$ such that $\|u - u_D - v_\varepsilon\|_{W^{1,p}(\Omega)} \lesssim \varepsilon$. Hence $v_\ell := I_\ell(v_\varepsilon + u_D) - u_\ell \in V_\ell$ satisfies

$$e_\ell - v_\ell = (u - u_D - v_\varepsilon) + (\mathrm{id} - I_\ell)(v_\varepsilon + u_D).$$

Note that the nodal interpolation $I_\ell(v_\varepsilon + u_D)$ is well-defined since v_ε and u_D are assumed to be smooth. With ([12], Lemma 3.1–3.2) it follows that

$$\|(\mathrm{id} - I_\ell)(v_\varepsilon + u_D)\|_{W^{1,p}(\Omega)} \lesssim H_\ell \to 0 \text{ and}$$
$$|||I_\ell(v_\varepsilon + u_D)|||^2_\ell = |||(\mathrm{id} - I_\ell)(v_\varepsilon + u_D)|||^2_\ell \lesssim H_\ell^{1+\gamma} \to 0 \text{ as } \ell \to \infty.$$

Since $\|\cdot\|_{L^2(\Omega)} \lesssim \|\cdot\|_{W^{1,p}(\Omega)}$, this yields some $\ell_0 \in \mathbb{N}$ such that

$$|e_\ell - v_\ell|^{r'}_{W^{1,p}(\Omega)} + \beta\|e_\ell - v_\ell\|^2_{L^2}(\Omega) + |||I_\ell(v_\varepsilon + u_D)|||^2_\ell \lesssim \varepsilon \text{ for all } \ell \geqslant \ell_0.$$

A Cauchy inequality applied to the stabilisation norm proves

$$a_\ell(u_\ell, v_\ell) = -|||u_\ell|||^2_\ell + a_\ell(u_\ell, I_\ell(v_\varepsilon + u_D)) \leqslant -\frac{1}{2}|||u_\ell|||^2_\ell + \frac{1}{2}|||I_\ell(v_\varepsilon + u_D)|||^2_\ell.$$

Substitute $a_\ell(u_\ell, v_\ell)$ in Lemma 3.2 and add $\frac{1}{2}|||u_\ell|||^2_\ell$ on both sides. This leads to

$$\|\delta_\ell\|^r_{L^{p'}(\Omega)} + \beta\|e_\ell\|^2_{L^2(\Omega)} + |||u_\ell|||^2_\ell \lesssim \varepsilon \text{ for all } \ell \geqslant \ell_0.$$

□

Example 3.3. The two-well example from the computational benchmark [18] allows an estimate on $\|e_\ell\|_{L^2(\Omega)}$ even for $\beta = 0$. Let $n = 2$, let $F_1 := -F_2 := (3,2)/\sqrt{13}$, and let the energy density W be the convex hull of $F \mapsto |F - F_1|^2|F - F_2|^2$. That is

$$W(F) = \left(\max\left\{0, |F|^2 - 1\right\}\right)^2 + 4\left(|F|^2 - (3F(1) + 2F(2))^2/13\right). \tag{3.4}$$

Then ([11], Lemma 9.1) proves, for all $v_\ell \in V_\ell$, that

$$\|e_\ell\|^2_{L^2(\Omega)} \lesssim \int_\Omega \delta_\ell : \mathrm{D}e_\ell \mathrm{d}x + \|e_\ell - v_\ell\|^2_{H^1(\Omega)}.$$

Therefore, the arguments of Lemma 3.2 lead to

$$\|\delta_\ell\|^r_{L^{p'}(\Omega)} + \|e_\ell\|^2_{L^2(\Omega)} \lesssim |e_\ell - v_\ell|^{r'}_{W^{1,p}(\Omega)} + \|e_\ell - v_\ell\|^2_{H^1(\Omega)} + a_\ell(u_\ell, v_\ell).$$

This result can be used in the proof of Theorem 3.1 in order to obtain

$$\|\sigma - \sigma_\ell\|^r_{L^{p'}(\Omega)} + \|u - u_\ell\|^2_{L^2(\Omega)} + |||u_\ell|||^2_\ell \to 0 \text{ as } \ell \to \infty.$$

A posteriori error estimates

Beyond the a posteriori error analysis of [7], the additional stabilisation term in the discretisation of this paper causes an additional difficulty in that the Galerkin orthogonality does *not* hold for the natural residual. Inspired from novell developments in the a posteriori error control of elliptic PDEs motivated by inexact solve [14-17], this section presents some guaranteed upper error bound for the discretisation at hand for any approximation u_ℓ which does not necessarily satisfy (3.2) exactly. Thereby inexact solve is included.

Let $u \in \mathcal{A}$ solve (2.3) and let $u_\ell \in \mathcal{A}_\ell$ be arbitrary. It is *not* assumed that u_ℓ solves the discrete problem (2.5); the following theorem holds regardless of this. Recall the definitions of $\mathrm{osc}_{\ell,q}(\cdot)$ and Π_ℓ from Section 'Methods: Discretisation and Stabilisation' and given $\sigma := \mathrm{D}W(\mathrm{D}u)$ and $\sigma_\ell := \mathrm{D}W(\mathrm{D}u_\ell)$, abbreviate

$$\Lambda_\ell := -2\beta(u_\ell - g) + f, \quad e_\ell := u - u_\ell \quad \text{and} \quad \delta_\ell := \sigma - \sigma_\ell.$$

Theorem 4.1. *Given any $w_\ell \in W^{1,p}(\Omega;\mathbb{R}^m)$ with $w_\ell = u - u_\ell$ on the boundary $\partial\Omega$, and given any $\tau \in H(div,\Omega;\mathbb{R}^{m\times n})$, it holds, for all $2 \leqslant q \leqslant p$ and for some constant $\varkappa$ known from ([12], Lemma 3.5), that*

$$\begin{aligned}\varkappa/2\|\delta_\ell\|^r_{L^{p'}(\Omega)} + \beta\|e_\ell\|^2_{L^2(\Omega)} &\leqslant (r\varkappa/2)^{1-r'}/r'|w_\ell|^{r'}_{W^{1,p}(\Omega)} + \beta\|w_\ell\|^2_{L^2(\Omega)}\\ &+ \Big(\|\sigma_\ell - \tau\|_{L^{q'}(\Omega)} + \|\Pi_\ell\Lambda_\ell + div\tau\|_{L^{q'}(\Omega)} + osc_{\ell,q'}(\Lambda_\ell)\Big)\|e_\ell - w_\ell\|_{W^{1,q}(\Omega)}.\end{aligned}$$

The constant $\varkappa$ depends on problem-specific data such as $\|u\|_{W^{1,p}(\Omega)}$ and the size of the domain Ω. Refer to the proof of Lemma 3.5 in [12] for details.

Before the proofs conclude this section, some practical choice of τ in Theorem 4.1 is discussed as some Raviart-Thomas finite element functions in

$$RT_0(\mathcal{T}_\ell) := \big\{\tau_{\mathrm{RT}} \in P_1(\mathcal{T}_\ell)\cap H(\mathrm{div},\Omega) : \forall T \in \mathcal{T}_\ell\ \exists a,b,c \in \mathbb{R}\ \forall x \in T, \tau_{\mathrm{RT}}(x) = (a,b)+cx\big\}.$$

We suggest the computation (or an accurate approximation) of

$$\mu_\ell := \min_{\tau\in RT_0(\mathcal{T}_\ell)^m}\Big(\|\sigma_\ell - \tau\|_{L^{q'}(\Omega)} + \|\Pi_\ell\Lambda_\ell + \mathrm{div}\tau\|_{L^{q'}(\Omega)}\Big) \tag{4.1}$$

and emphasise that any upper bound is allowed in Theorem 4.1. This leads to

$$\begin{aligned}\varkappa/2\|\delta_\ell\|^r_{L^{p'}(\Omega)} + \beta\|e_\ell\|^2_{L^2(\Omega)} &\leqslant (r\varkappa/2)^{1-r'}/r'|w_\ell|^{r'}_{W^{1,p}(\Omega)} + \beta\|w_\ell\|^2_{L^2(\Omega)}\\ &+ \big(\mu_\ell + \mathrm{osc}_{\ell,q'}(\Lambda_\ell)\big)\|e_\ell - w_\ell\|_{W^{1,q}(\Omega)}.\end{aligned}$$

The algorithm of ([20], Prop. 4.1) computes some w_ℓ from $(\mathrm{id} - I_\ell)u_D$ with

$$\begin{aligned}\|w_\ell\|_{L^q(T)} &\approx h_T{}^{1/q}\|(\mathrm{id} - I_\ell)u_D\|_{L^q(\partial T\cap\partial\Omega)} \quad \text{and}\\ \|\mathrm{D}w_\ell\|_{L^q(T)} &\lesssim h_T^{1/q-1}\|(\mathrm{id} - I_\ell)u_D\|_{L^q(\partial T\cap\partial\Omega)} + h_T^{1/q}\|\partial(\mathrm{id} - I_\ell)u_D/\partial s\|_{L^q(\partial T\cap\partial\Omega)}.\end{aligned} \tag{4.2}$$

(The proof of the second assertion is analogous to that of ([20], Prop. 4.1) and the first is an immediate consequence of the design of w_ℓ). This and $\|e_\ell - w_\ell\|_{W^{1,q}(\Omega)} \lesssim 1$ for bounded u_ℓ (i.e. solely $\|u_\ell\|_{W^{1,p}(\Omega)} \lesssim 1$ is assumed) lead to the practical estimate μ_ℓ as a computable guaranteed upper bound of the left-hand side of Theorem 4.1. Since the minimisation of (4.1) is computationally intensive for $q \neq 2$, Section 'Numerical experiments' actually computes an approximation of μ_ℓ, based on $q = 2$.

The choice $\tau = \sigma$ in Theorem 4.1 shows that the right-hand side is in fact optimal up to oscillations. The reliability-efficiency gap of [18] is visible here in that we have no further estimate on $\|u_\ell\|_{W^{1,p}(\Omega)}$ [7,18]. However, additional smoothness assumptions on u may lead to refined estimates on the term $\|e_\ell - w_\ell\|_{W^{1,q}(\Omega)}$ (cf. Section 'Refined analysis for

an interface model problem'). The following result indicates that μ_ℓ is sharp in the sense that it converges with the correct convergence rate. This theorem employs the Fortin interpolation operator $I_{\mathrm{F},\ell}$ defined for $\tau \in H(\mathrm{div},\Omega) \cap L^t(\Omega;\mathbb{R}^n)$ with $t > 2$ by $I_{\mathrm{F},\ell}\tau \in RT_0(\mathcal{T}_\ell)$ and

$$\fint_F n_F \cdot (\mathrm{id} - I_{\mathrm{F},\ell})\tau \,\mathrm{d}s = 0 \quad \text{for all } F \in \mathcal{F}_\ell.$$

Here and in the following, n_F denotes a unit normal vector of the side F; the direction of n_F arbitrary, but fixed for a given side F. For the improved regularity of stress in the class of degenerate convex minimisation problems at hand, we refer to [3,21].

Theorem 4.2. *(*Efficiency*) If the exact stress σ is sufficiently regular such that its Fortin interpolant $\tau_\ell = I_{\mathrm{F},\ell}\sigma \in RT_0(\mathcal{T}_\ell;\mathbb{R}^{m\times n})$ is defined, it holds*

$$\begin{aligned}\|\sigma_\ell - \tau_\ell\|_{L^{q'}(\Omega)} + \|\Pi_\ell\Lambda_\ell + div\tau_\ell\|_{L^{q'}(\Omega)} \\ \lesssim \|\delta_\ell\|_{L^{q'}\Omega} + 2\beta\|e_\ell\|_{L^{q'}(\Omega)} + \|(id - I_{\mathrm{F},\ell})\sigma\|_{L^{q'}(\Omega)}.\end{aligned}$$

It is expected that $\|(\mathrm{id} - I_{\mathrm{F},\ell})\sigma\|_{L^{q'}(\Omega)} \lesssim H_\ell$. This is shown in ([22], Prop. 3.6) for $q' = 2$ and therefore also holds for $q' \leqslant 2$. Hence the right-hand side of the assertion of Theorem 4.2 converges with the (expected) optimal convergence rates.

Proof of Theorem 4.1. Let $\varkappa$ be the reciprocal of c_1 in ([12], Lemma 3.5), which is also the multiplicative constant hidden in (3.3). Recall Young's inequality, which reads $ab \leqslant a^r/r + b^{r'}/r'$ for $a, b > 0$. This, (3.3) and the continuous Euler-Lagrange equation (3.1) show, for $v = e_\ell - w_\ell \in V$, that

$$\begin{aligned}\varkappa\|\delta_\ell\|^r_{L^{p'}(\Omega)} + 2\beta\|e_\ell\|^2_{L^2(\Omega)} &\leqslant \int_\Omega (\delta_\ell : \mathrm{D}v + 2\beta e_\ell \cdot v)\mathrm{d}x \\ &\quad + \int_\Omega (\delta_\ell : \mathrm{D}w_\ell + 2\beta e_\ell \cdot w_\ell)\mathrm{d}x \\ &\leqslant -\int_\Omega (\sigma_\ell : \mathrm{D}v - \Lambda_\ell \cdot v)\mathrm{d}x \\ &\quad + \beta\|e_\ell\|^2_{L^2(\Omega)} + \beta\|w_\ell\|^2_{L^2(\Omega)} \\ &\quad + \varkappa/2\|\delta_\ell\|^r_{L^{p'}(\Omega)} + (r\varkappa/2)^{1-r'}/r'|w_\ell|^{r'}_{W^{1,p}(\Omega)}.\end{aligned}$$

Hence $\mathrm{Res}_\ell(v) := -\int_\Omega(\sigma_\ell : \mathrm{D}v - \Lambda_\ell \cdot v)\mathrm{d}x$ satisfies

$$\varkappa/2\|\delta_\ell\|^r_{L^{p'}(\Omega)} + \beta\|e_\ell\|^2_{L^2(\Omega)} \leqslant \mathrm{Res}_\ell(v) + (r\varkappa/2)^{1-r'}/r'|w_\ell|^{r'}_{W^{1,p}(\Omega)} + \beta\|w_\ell\|^2_{L^2(\Omega)}.$$

Let $C_{q'}$ denote the Poincaré constant of convex domains with respect to the $W^{1,q'}$ norm. The fundamental theorem of calculus on some one-dimensional arc shows that $C_\infty \leqslant 1$. The paper [23] proves $C_1 = 1/2$. Hence, operator-interpolation arguments [24,25] prove $C_{q'} \leqslant (1/2)^{1/q'} \leqslant 1$. The Poincaré inequality shows, for any $2 \leqslant q \leqslant p$, that

$$\begin{aligned}\int_\Omega (\mathrm{id} - \Pi_\ell)\Lambda_\ell \cdot v\mathrm{d}x &= \int_\Omega h_\ell(\mathrm{id} - \Pi_\ell)\Lambda_\ell \cdot \frac{1}{h_\ell}(\mathrm{id} - \Pi_\ell)v\mathrm{d}x \\ &\leqslant \|h_\ell(\mathrm{id} - \Pi_\ell)\Lambda_\ell\|_{L^{q'}(\Omega)}\|\mathrm{D}v\|_{L^q(\Omega)} = \mathrm{osc}_{\ell,q'}(\Lambda_\ell)\|\mathrm{D}v\|_{L^q(\Omega)}.\end{aligned}$$

For any $\tau \in H(\mathrm{div},\Omega;\mathbb{R}^{m\times n})$, the Hölder and Poincaré inequalities show

$$\begin{aligned}\mathrm{Res}_\ell(v) &= -\int_\Omega \big((\sigma_\ell - \tau) : \mathrm{D}v - (\Pi_\ell\Lambda_\ell + \mathrm{div}\tau)\cdot v - (\mathrm{id} - \Pi_\ell)\Lambda_\ell \cdot v\big)\mathrm{d}x \\ &\leqslant \Big(\|\sigma_\ell - \tau\|_{L^{q'}(\Omega)} + \|\Pi_\ell\Lambda_\ell + \mathrm{div}\tau\|_{L^{q'}(\Omega)} + \mathrm{osc}_{\ell,q'}(\Lambda_\ell)\Big)\|v\|_{W^{1,q}(\Omega)}.\end{aligned}$$

Proof of Theorem 4.2. The triangle inequality yields

$$\|\sigma_\ell - \tau_\ell\|_{L^{q'}(\Omega)} \leqslant \|(\mathrm{id} - I_{\mathrm{F},\ell})\sigma\|_{L^{q'}(\Omega)} + \|\delta_\ell\|_{L^{q'}\Omega}.$$

Since $f = 2\beta(u-g) - \mathrm{div}\sigma$, the commutative property $\mathrm{div} I_{\mathrm{F},\ell} = \Pi_\ell \mathrm{div}$ (cf. ([22], p. 129)) yields

$$\|\Pi_\ell \Lambda_\ell + \mathrm{div}\tau_\ell\|_{L^{q'}(\Omega)} = 2\beta\|\Pi_\ell e_\ell\|_{L^{q'}(\Omega)} \leqslant 2\beta\|e_\ell\|_{L^{q'}(\Omega)}.$$

Refined analysis for an interface model problem

This section is devoted for a model scenario from phase transition problems [18] with some solution u that is smooth outside some one-dimensional interface Γ. Suppose some (possibly non-unique) minimiser u of the continuous problem (2.3) satisfies $u \in W^{1,\infty}(\Omega;\mathbb{R}^m) \cap W^{2,p}(\Omega \setminus \Gamma;\mathbb{R}^m)$ for some finite union Γ of $(n-1)$ dimensional Lipschitz surfaces in $\overline{\Omega}$. Since Ω has a Lipschitz boundary, this implies Lipschitz continuity of u on Ω. We refer to [19] for sufficient conditions for $u \in W^{1,\infty}(\Omega;\mathbb{R}^m)$ and conclude that the remaining assumption $u \in W^{2,p}(\Omega \setminus \Gamma;\mathbb{R}^m)$ is the essential hypothesis expected in many interface problems. Let $u_\ell \in \mathcal{A}_\ell$ be the (unique) minimiser of the discrete stabilised problem (2.5). In the following, also $\Gamma = \emptyset$ is permitted to extend previous results [12] for highly regular minimisers.

The following theorem leads to a priori convergence rates for the interface model problem. Thereby it recovers the results of [12] for problems with piecewise smooth exact solution.

We will abbreviate the set of all triangles that are touched by Γ as $\mathcal{T}_\ell(\Gamma) := \{T \in \mathcal{T}_\ell : \mathrm{dist}(T,\Gamma) = 0\}$, its cardinality as $|\mathcal{T}_\ell(\Gamma)|$, its union as $\Omega_{\Gamma,\ell} := \mathrm{int}(\bigcup \mathcal{T}_\ell(\Gamma))$ with volume $|\Omega_{\Gamma,\ell}|$ and its complement as $\Omega^{\mathrm{C}}_{\Gamma,\ell} := \Omega \setminus \overline{\Omega_{\Gamma,\ell}}$.

Theorem 5.1. *Provided $\beta > 0$, it holds*

$$\|\delta_\ell\|^r_{L^{p'}(\Omega)} + \|e_\ell\|^2_{L^2(\Omega)} + |||u_\ell|||^2_\ell \lesssim H_\ell^{1+\gamma}|u|^2_{H^2(\Omega\setminus\Gamma)} + H_\ell^2|u|^2_{W^{1,\infty}(\Omega)} + H_\ell^{r/(r-1)}|u|^{r/(r-1)}_{W^{2,p}(\Omega^C_{\Gamma,\ell})} + H_\ell^{\gamma+n-1}|u|^2_{W^{1,\infty}(\Omega)}|\mathcal{T}_\ell(\Gamma)| + |u|^{r/(r-1)}_{W^{1,\infty}(\Omega)}|\Omega_{\Gamma,\ell}|^{r/((r-1)p)}.$$

Remark 5.2. In the case of uniform mesh refinements we may expect $|\mathcal{T}_\ell(\Gamma)| \approx H_\ell^{1-n}$ and $|\Omega_{\Gamma,\ell}| \approx H_\ell$ and Theorem 5.1 simplifies to

$$\|\delta_\ell\|^r_{L^{p'}(\Omega)} + \|e_\ell\|^2_{L^2(\Omega)} + |||u_\ell|||^2_\ell \lesssim H_\ell^{\min\{\gamma,2\}}|u|^2_{W^{1,\infty}(\Omega)} + H_\ell^{r/((r-1)p)}|u|^{r/(r-1)}_{W^{1,\infty}(\Omega)}.$$

Proof. With $w_\ell = (\mathrm{id} - I_\ell)e_\ell = (\mathrm{id} - I_\ell)u$, a Young inequality, (3.3) and ([12], Theorem 3.8) yield

$$\|\delta_\ell\|^r_{L^{p'}(\Omega)} + \|e_\ell\|^2_{L^2(\Omega)} + |||u_\ell|||^2_\ell \lesssim |w_\ell|^{r/(r-1)}_{W^{1,p}(\Omega)} + \|w_\ell\|^2_{L^2(\Omega)} + |||I_\ell u|||^2_\ell.$$

Theorem 4.4.4 in [25] shows $\|w_\ell\|_{L^2(\Omega)} \lesssim \|w_\ell\|_{L^\infty(\Omega)} \lesssim H_\ell|u|_{W^{1,\infty}(\Omega)}$ and

$$|w_\ell|^p_{W^{1,p}(\Omega)} = |w_\ell|^p_{W^{1,p}(\Omega_{\Gamma,\ell})} + |w_\ell|^p_{W^{1,p}(\Omega^C_{\Gamma,\ell})} \lesssim |u|^p_{W^{1,\infty}(\Omega_{\Gamma,\ell})}|\Omega_{\Gamma,\ell}| + H_\ell^p|u|^p_{W^{2,p}(\Omega^C_{\Gamma,\ell})}.$$

Let $\omega_F = \bigcup_{\substack{T\in\mathcal{T}_\ell \\ F\subset T}} T$ be the patch of a side $F \in \mathcal{F}_\ell$, and set $\mathcal{F}_\ell(\Gamma)$= $\{F \in \mathcal{F}_\ell(\Omega) : \omega_F \cap \Gamma \neq \emptyset\}$ and $\mathcal{F}_\ell^{\mathrm{C}}(\Gamma) = \mathcal{F}_\ell(\Omega) \setminus \mathcal{F}_\ell(\Gamma)$. Note that $[\mathrm{D}u]_F = 0$ for $F \in \mathcal{F}_\ell^{\mathrm{C}}(\Gamma)$. Then

$$|||I_\ell u|||_\ell^2 = H_\ell^{1+\gamma}\left(\sum_{F\in\mathcal{F}_\ell^{\mathrm{C}}(\Gamma)} h_F^{-1}\|[\mathrm{D}w_\ell]_F\|_{L^2(F)}^2 + \sum_{F\in\mathcal{F}_\ell(\Gamma)} h_F^{-1}\|[\mathrm{D}I_\ell u]_F\|_{L^2(F)}^2\right).$$

The first sum can be estimated as in the proof of ([12], Lemma 3.2), the second sum with

$$\|[\mathrm{D}I_\ell u]_F\|_{L^2(F)}^2 \lesssim h_F^{n-1}|I_\ell u|_{W^{1,\infty}(F)}^2 \lesssim h_F^{n-1}|u|_{W^{1,\infty}(F)}^2.$$

The observation $|\mathcal{F}_\ell(\Gamma)| \leqslant (n+1)|\mathcal{T}_\ell(\Gamma)|$ concludes the proof. □

Together with Theorem 5.1, the subsequent result implies strong convergence of the gradients in the model interface problem as $H_\ell \to 0$.

Theorem 5.3. *Under the aforementioned conditions on the (possibly non-unique) exact minimiser $u \in W^{1,\infty}(\Omega;\mathbb{R}^m) \cap W^{2,p}(\Omega\setminus\Gamma;\mathbb{R}^m)$, the error $e_\ell = u - u_\ell$ of the discrete solution $u_\ell \in \mathcal{A}_\ell$ of (2.5) satisfies*

$$\begin{aligned}\|De_\ell\|_{L^2(\Omega)} \lesssim& \|e_\ell\|_{L^2(\Omega)}^{1/3} + H_\ell^{5/6}\,\|\partial^2 u_D/\partial s^2\|_{L^2(\partial\Omega)}^{1/3} + H_\ell^{(1-\gamma)/2}|||u_\ell|||_\ell \\ &+ H_\ell^{-(1+\gamma)/4}|||u_\ell|||_\ell^{1/2}\Big(\|e_\ell\|_{L^2(\Omega)}^{1/2} + H_\ell^{5/4}\|\partial^2 u_D/\partial s^2\|_{L^2(\Omega)}^{1/2}\Big).\end{aligned}$$

Proof. The basic idea of gradient control is the generalisation of the interpolation estimate $H^1(\Omega) = [L^2(\Omega), H^2(\Omega)]_{1/2}$ for a reduced domain $\Omega\setminus\Gamma$; refer to [24,25] for a detailed analysis of interpolation spaces. Let w_ℓ be the boundary value interpolation of $(\mathrm{id} - I_\ell)u_D$ as described in ([20], Prop. 4.1), such that w_ℓ satisfies the inequalities in (4.2). A piecewise integration by parts shows, for $v := e_\ell - w_\ell \in W_0^{1,p}(\Omega;\mathbb{R}^m)$, that

$$\begin{aligned}\|\mathrm{D}e_\ell\|_{L^2(\Omega)}^2 &= \int_\Omega \mathrm{D}(u-u_\ell) : \mathrm{D}v\,\mathrm{d}x + \int_\Omega \mathrm{D}e_\ell : \mathrm{D}w_\ell\,\mathrm{d}x \\ &\leqslant \int_\Gamma v\cdot[\mathrm{D}u]_\Gamma\, n_\Gamma\,\mathrm{d}s - \int_{\Omega\setminus\Gamma} v\cdot\Delta u\,\mathrm{d}x - \sum_{F\in\mathcal{F}\ell(\Omega)}\int_F v\cdot[\mathrm{D}u_\ell]_F\, n_F\,\mathrm{d}s \\ &\quad + \|\mathrm{D}e_\ell\|_{L^2(\Omega)}\|\mathrm{D}w_\ell\|_{L^2(\Omega)},\end{aligned}$$

where n_Γ is a unit normal vector of the interface Γ. The Lipschitz continuity of u implies $|[\mathrm{D}u]_\Gamma n_\Gamma| \lesssim 1$. This and the trace inequality on Γ lead to

$$\int_\Gamma v\cdot[\mathrm{D}u]_\Gamma\, n_\Gamma\,\mathrm{d}s \lesssim \|v\|_{L^2(\Gamma)} \lesssim \|v\|_{L^2(\Omega)} + \|v\|_{L^2(\Omega)}^{1/2}\,\|\mathrm{D}v\|_{L^2(\Omega)}^{1/2}.$$

The case $\Gamma = \varnothing$ is contained in ([12], Theorem 4.4). The piecewise Laplacian of u is bounded in $L^2(\Omega)$ and so (with the generic constant $C := \|\Delta u\|_{L^2(\Omega\setminus\Gamma)}$ hidden in the notation $C \approx 1$)

$$\int_{\Omega\setminus\Gamma} v\cdot\Delta u\mathrm{d}x \lesssim \|v\|_{L^2(\Omega)}$$

The elementwise trace inequality ([25], Theorem 1.6.6, p. 39) for an n-dimensional simplex T and one of its sides F, and $f \in W^{1,q}(T;\mathbb{R}^m)$,$1 \leqslant q < \infty$, reads

$$\|f\|_{L^q(F)}^q \lesssim h_T^{-1}\|f\|_{L^q(T)}^q + \|f\|_{L^q(T)}^{q-1}\|\mathrm{D}f\|_{L^q(T)} \lesssim h_T^{-1}\|f\|_{L^q(T)}^q + h_T^{q-1}\|\mathrm{D}f\|_{L^q(T)}^q.$$

The term $\int_F v \cdot [\mathrm{D}u_\ell]_F \, n_F \, \mathrm{d}s$ and the stabilisation $|||u_\ell|||_\ell$ are already analysed in the *Estimate on C* in the proof of ([12], Theorem 4.4). This results in

$$\sum_{F\in\mathcal{F}_\ell(\Omega)} \int_F v \cdot [\mathrm{D}u_\ell]_F \, n_F \, \mathrm{d}s \lesssim |||u_\ell|||_\ell \left(H_\ell^{(1-\gamma)/2} \|\mathrm{D}v\|_{L^2(\Omega)} + H_\ell^{-(1+\gamma)/2} \|v\|_{L^2(\Omega)} \right) .$$

The preceding estimates plus the absorbtion of $\|\mathrm{D}e_\ell\|_{L^2(\Omega)}$ lead to

$$\begin{aligned} \|\mathrm{D}e_\ell\|_{L^2(\Omega)}^2 &\lesssim \|v\|_{L^2(\Omega)} + \|v\|_{L^2(\Omega)}^{1/2} \|\mathrm{D}v\|_{L^2(\Omega)}^{1/2} + \|\mathrm{D}w_\ell\|_{L^2(\Omega)}^2 \\ &\quad + |||u_\ell|||_\ell \left(H_\ell^{(1-\gamma)/2} \|\mathrm{D}v\|_{L^2(\Omega)} + H_\ell^{-(1+\gamma)/2} \|v\|_{L^2(\Omega)} \right) . \end{aligned}$$

The triangle inequality applied to $v = e_\ell - w_\ell$ and some careful elementary analysis to absorb $\|\mathrm{D}e_\ell\|_{L^2(\Omega)}^{1/2}$ eventually lead to

$$\begin{aligned} \|\mathrm{D}e_\ell\|_{L^2(\Omega)} &\lesssim \|e_\ell\|_{L^2(\Omega)}^{1/3} + \|w_\ell\|_{L^2(\Omega)}^{1/3} + |w_\ell|_{H^1(\Omega)} + H_\ell^{(1-\gamma)/2} |||u_\ell|||_\ell \\ &\quad + H_\ell^{-(1+\gamma)/4} |||u_\ell|||_\ell^{1/2} \big(\|e_\ell\|_{L^2(\Omega)} + \|w_\ell\|_{L^2(\Omega)} \big)^{1/2} . \end{aligned}$$

The inequalities (4.2), Poincaré and Friedrichs inequalities on sides $F \in \mathcal{F}_\ell(\partial\Omega)$ and removal of higher-order terms in H_ℓ conclude the proof. □

The following theorem is an improved a posteriori estimate based on Theorems 4.1 and 5.3.

Theorem 5.4. *Recall $u \in W^{1,\infty}(\Omega;\mathbb{R}^m) \cap W^{2,p}(\Omega \setminus \Gamma;\mathbb{R}^m)$, the definitions $e_\ell := u - u_\ell$ and $\delta_\ell := \sigma - \sigma_\ell$ for $\sigma := \mathrm{D}W(\mathrm{D}u)$ and $\sigma_\ell := \mathrm{D}W(\mathrm{D}u_\ell)$, and the definition of Λ_ℓ from Section 'A posteriori error estimates'. Set*

$$M(\tau) := \|\sigma_\ell - \tau\|_{L^2(\Omega)} + \|\Pi_\ell \Lambda_\ell + div\tau\|_{L^2(\Omega)} + osc_{\ell,2}(\Lambda_\ell) \quad \text{for all } \tau \in H(div,\Omega;\mathbb{R}^{m\times n}) .$$

Provided $\beta > 0$, it holds

$$\begin{aligned} \|\delta_\ell\|_{L^{p'}(\Omega)}^r + \|e_\ell\|_{L^2(\Omega)}^2 &\lesssim M(\tau)^{6/5} + H_\ell^{-(1+\gamma)/3} M(\tau)^{4/3} |||u_\ell|||_\ell^{2/3} \\ &\quad + M(\tau) \Big(H_\ell^{(1-\gamma)/2} |||u_\ell|||_\ell + H_\ell^{1-\gamma/4} |||u_\ell|||_\ell^{1/2} \Big) + H_\ell^{\min\{5,r'(1+1/p)\}} \end{aligned}$$

and

$$\begin{aligned} \|\mathrm{D}e_\ell\|_{L^2(\Omega)}^2 &\lesssim M(\tau)^{2/5} + H_\ell^{-(1+\gamma)/9} M(\tau)^{4/9} |||u_\ell|||_\ell^{2/9} + H_\ell^{\min\{5/3,r'(1+1/p)/3\}} \\ &\quad + M(\tau)^{1/3} \Big(H_\ell^{(1-\gamma)/2} |||u_\ell|||_\ell + H_\ell^{1-\gamma/4} |||u_\ell|||_\ell^{1/2} \Big)^{1/3} + H_\ell^{1-\gamma} |||u_\ell|||_\ell^2 \\ &\quad + H_\ell^{-(1+\gamma)/2} |||u_\ell|||_\ell \Big(M(\tau)^{6/5} + H_\ell^{-(1+\gamma)/3} M(\tau)^{4/3} |||u_\ell|||_\ell^{2/3} + H_\ell^{\min\{5,r'(1+1/p)\}} \Big)^{1/2} \\ &\quad + H_\ell^{-(1+\gamma)/2} |||u_\ell|||_\ell M(\tau)^{1/2} \Big(H_\ell^{(1-\gamma)/2} |||u_\ell|||_\ell + H_\ell^{1-\gamma/4} |||u_\ell|||_\ell^{1/2} \Big)^{1/2} \end{aligned}$$

The generic constants in Theorem 5.4 depend on problem-specific data such as the shapes of Ω and Γ as well as the generic constant $\varkappa$ of Theorem 4.1.

Theorem 5.5. *Theorem 5.4 holds verbatim in Example 3.3 and in the modified two-well problem of Subsection 'Modified two-well benchmark', where $\beta = 0$.*

Remark 5.6. The assertion of Theorem 5.4 holds for any discrete $u_\ell \in u_{D,\ell} + V_\ell$ which may approximate the discrete unique exact solution of (2.5). This allows the inexact SOLVE via an iterative procedure.

Proof of Theorem 5.4. Choose w_ℓ as in the proof of Theorem 5.3. Then Theorem 4.1 with $q = 2$ and (4.2) imply

$$\begin{aligned}\|\delta_\ell\|^r_{L^{p'}(\Omega)} + \|e_\ell\|^2_{L^2(\Omega)} &\lesssim M(\tau)\|e_\ell - w_\ell\|_{H^1(\Omega)} + |w_\ell|^{r'}_{W^{1,p}(\Omega)} + \|w_\ell\|^2_{L^2(\Omega)} \\ &\lesssim M(\tau)\Big(|e_\ell|_{H^1(\Omega)} + \|e_\ell\|_{L^2(\Omega)} + H_\ell^{3/2}\Big) + H_\ell^{\min\{5,r'(1+1/p)\}}.\end{aligned}$$

Theorem 5.3 provides an estimate of the semi-norm $|e_\ell|_{H^1(\Omega)}$. A Young inequality shows $H_\ell^{5/6}M(\tau) \lesssim H_\ell^5 + M(\tau)^6$. The absorbtion of $\|e_\ell\|_{L^2(\Omega)}$ then proves the first assertion. The second assertion is an immediate consequence of the first one, Theorem 5.3 and several algebraic transformations.

Numerical experiments

This section illustrates the theoretical estimates and their impact on the reliability-efficiency gap on 2D benchmarks in computational microstructures [18,26].

Numerical algorithms

The adaptive finite element method (AFEM) and algorithmic details on the implementation in MATLAB in the spirit of [27] concern the state-of-the-art AFEM loop

$$\text{SOLVE} \;\to\; \text{ESTIMATE} \;\to\; \text{MARK} \;\to\; \text{REFINE}$$

and are explained below together with some notation.

Solve

The stabilised discrete problem (2.5) is solved in a nested iteration on a given triangulation $\mathcal{T}_\ell$ with MATLAB's standard-minimiser `fminunc` with default tolerances. Gradient and Hessian of the discrete energy are available and therefore provided to `fminunc`. We set $\gamma = 1$ in the stabilisation term (2.4) in all our experiments. This is motivated by ([12], Theorem 4.4) which suggest that $\gamma = 1$ yields an optimal convergence rate. The discrete solution of the previous AFEM loop iteration serves as a start vector for `fminunc`; for the first iteration, the initial vector is zero everywhere up to the Dirichlet boundary nodes. Since the Galerkin orthogonality is *not* required in Theorem 4.1, the termination of an iterative realisation for SOLVE is *not* a sensitive issue. In the computational PDEs, it is a fundamental issue to involve inexact solve. In this paper, however, the numerical examples are run with the standard settings of MATLAB.

Estimate

The refinement indicator results from the error estimator of Theorem 4.1. As in the work of Repin [28], the computation of the minimiser $\tau \in RT_0(\mathcal{T}_\ell)^m$ of

$$\|\sigma_\ell - \tau\|_{L^2(\Omega)} + \|\Pi_\ell\Lambda_\ell + \mathrm{div}\tau\|_{L^2(\Omega)} \tag{6.1}$$

runs Algorithm 1 based on the formula

$$(a+b)^2 = \min_{s>0}\big((1+s)a^2 + (1+1/s)b^2\big) \quad \text{for } a, b > 0$$

The stopping criterion of Algorithm 1 monitors relative changes and avoids degenerate values of s. Undisplayed experiments have conviced us that a maxmium of three iterations and a stopping tolerance of $\varepsilon_{\mathrm{M}}^{0.8}$ (with the machine precision ε_{M}) yield satisfying results.

Algorithm 1 Approximate flux computation

Input: σ_ℓ, $\Pi_\ell \Lambda_\ell$
$s_1 = 1$;
for $k = 1, 2, 3$ **do**
  Compute minimiser τ_k of
  $M(s_k, \tau) = (1 + s_k)\|\sigma_\ell - \tau\|^2_{L^2(\Omega)} + (1 + 1/s_k)\|\Pi_\ell \Lambda_\ell + \mathrm{div}\tau\|^2_{L^2(\Omega)}$;
  if $D^2_\tau(s_k, \tau_k)$ *nearly singular (MATLAB "warning")* **then return** τ_k;
  $s_{k+1} = \|\Pi_\ell \Lambda_\ell + \mathrm{div}\tau_k\|_{L^2(\Omega)} / \|\sigma_\ell - \tau_k\|_{L^2(\Omega)}$;
  if $\max\left\{s_{k+1}, 1/s_{k+1}, \frac{|s_{k+1} - s_k|}{s_{k+1} + s_k}\right\} < \varepsilon_{\mathrm{M}}^{0.8}$ **then return** τ_k;
Output: approximate flux τ

The iteration is stopped whenever s, $1/s$ or the relative change of s drops below this tolerance. As an additional precaution, the iteration also stops if the linear system is deemed "nearly singular" by MATLAB. Our experiments convinced us that ignoring this warning causes a breakdown with NaNs. Note that if $q \neq 2$, we still minimise the L^2 sums in (6.1) to avoid the computational cost of a nonlinear solve. With the computed minimiser τ, Section 'A posteriori error estimates' yields the error estimator

$$\eta_{F,q'} := \|\sigma_\ell - \tau\|_{L^{q'}(\Omega)} + \|\Pi_\ell \Lambda_\ell + \mathrm{div}\tau\|_{L^{q'}(\Omega)} + \mathrm{osc}_{\ell,q'}(\Lambda_\ell)\,.$$

This will be compared with the well-established *residual based a posteriori error estimator* [7]

$$\eta_{R,q'} := \left(\sum_{T \in \mathcal{T}_\ell} h_T^{q'} \|\Lambda_\ell\|^{q'}_{L^{q'}(T)}\right)^{1/q'} + \left(\sum_{F \in \mathcal{F}_\ell(\Omega)} h_F \|[\sigma_\ell]_F \cdot n_F\|^{q'}_{L^{q'}(F)}\right)^{1/q'},$$

which is reliable for the original discretisation without stabilisation. Undisplayed experiments computed the *averaging error estimator* [18], which is founded on the same theoretical background as $\eta_{R,q'}$ and therefore yielded essentially the same convergence rates.

The error estimators in Theorem 5.4 read

$$\begin{aligned}\eta_{L,2} :=\ & \eta_{F,2}^{6/5} + H_\ell^{-(1+\gamma)/3} \eta_{F,2}^{4/3} |||u_\ell|||_\ell^{2/3} \\ & + \eta_{F,2}\left(H_\ell^{(1-\gamma)/2} |||u_\ell|||_\ell + H_\ell^{1-\gamma/4} |||u_\ell|||_\ell^{1/2}\right) + H_\ell^{\min\{5, r'(1+1/p)\}}\end{aligned}$$

$$\begin{aligned}\eta_{H,2} :=\ & \eta_{F,2}^{2/5} + H_\ell^{-(1+\gamma)/9} \eta_{F,2}^{4/9} |||u_\ell|||_\ell^{2/9} + H_\ell^{\min\{5/3, r'(1+1/p)/3\}} \\ & + \eta_{F,2}^{1/3}\left(H_\ell^{(1-\gamma)/2} |||u_\ell|||_\ell + H_\ell^{1-\gamma/4} |||u_\ell|||_\ell^{1/2}\right)^{1/3} + H_\ell^{1-\gamma} |||u_\ell|||_\ell^2 \\ & + H_\ell^{-(1+\gamma)/2} |||u_\ell|||_\ell \left(\eta_{F,2}^{6/5} + H_\ell^{-(1+\gamma)/3} \eta_{F,2}^{4/3} |||u_\ell|||_\ell^{2/3} + H_\ell^{\min\{5, r'(1+1/p)\}}\right)^{1/2} \\ & + H_\ell^{-(1+\gamma)/2} |||u_\ell|||_\ell \eta_{F,2}^{1/2}\left(H_\ell^{(1-\gamma)/2} |||u_\ell|||_\ell + H_\ell^{1-\gamma/4} |||u_\ell|||_\ell^{1/2}\right)^{1/2}.\end{aligned}$$

MARK

For any given $T \in \mathcal{T}_\ell$ with its set of faces $\mathcal{F}(T)$, $\partial T = \bigcup \mathcal{F}(T)$, and given τ from (6.1), set

$$\eta_F^{q'}(T) := \|\sigma_\ell - \tau\|_{L^{q'}(T)}^{q'} + \|\Pi_\ell \Lambda_\ell + \operatorname{div}\tau\|_{L^{q'}(T)}^{q'} + {h_T}^{q'} \|(\mathrm{id} - \Pi_\ell)\Lambda_\ell\|_{L^{q'}(T)}^{q'}.$$

$$\eta_R^{q'}(T) := |T|^{q'/n} \|\Lambda_\ell\|_{L^{q'}(T)}^{q'} + |T|^{1/n} \sum_{F \in \mathcal{F}_\ell(\Omega) \cap \mathcal{F}(T)} \|[\sigma_\ell]_F \cdot n_F\|_{L^{q'}(F)}^{q'}.$$

Let $\eta^{q'}(T)$ be one of the refinement indicators $\eta_F^{q'}(T)$ and $\eta_R^{q'}(T)$. Some greedy algorithm computes $\mathcal{M}_\ell \subset \mathcal{T}_\ell$ of (almost) minimal cardinality such that

$$\sum_{T \in \mathcal{M}_\ell} \eta^{q'}(T) \geqslant 1/2 \sum_{T \in \mathcal{T}_\ell} \eta^{q'}(T).$$

Refine

This step computes the smallest refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ with $\mathcal{M}_\ell \subset \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$ based on the red-green-blue refinement strategy as illustrated in Figure 2. This refinement involves some closure algorithm to avoid hanging nodes.

Two-well benchmark

The computational microstructure benchmark of ([18], Section 2) considers two wells with W from (3.4) in Example 3.3. The energy is given by (1.1) on the domain $\Omega = (0,1) \times (0,3/2) \subset \mathbb{R}^2$ with

$$g(x) := -3t^5/128 - t^3/3 \text{ and } u_D(x) := \begin{cases} g(x) & \text{for } t \leqslant 0, \\ t^3/24 + t & \text{for } t \geqslant 0 \end{cases}$$

for $t := (3(x_1 - 1) + 2x_2)/\sqrt{13}$; $p = q = 4$ and $f \equiv 0$. The unique minimiser u of $\min_{v \in \mathcal{A}} E(v)$ with $\mathcal{A} = u_D + W_0^{1,4}(\Omega)$ reads $u = u_D$ ([18], Theorem 2.1) and $\beta = 1$ allows for Theorems 5.1–5.4 to hold. An initial triangulation $\mathcal{T}_0$ is given by a criss triangulation of $(0,1) \times (0,3/2)$ with 12 congruent triangles and the two interior nodes $(1/2, 1/2)$ and $(1/2, 1)$. The adaptive algorithm of Subsection 'Numerical algorithms' computes a sequence of discrete solutions $(u_\ell)_\ell$ and stresses $(\sigma_\ell)_\ell$, as well as error estimators η_F and η_R with and without stabilisation for uniform and adaptive meshes and led to Figure 3 with overall observations of Section 'Conclusions'. The empirical convergence rates for uniform and R- as well as F-adapted mesh-refining are collected in Table 1. Note that the error estimator η_L performs better than η_F. This is evident from the table for uniform mesh refinements, but a closer look at Figure 3 reveals that even in the adaptive scenarios,

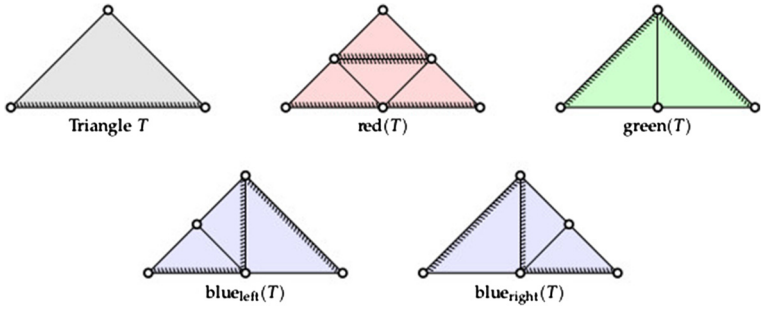

Figure 2 Possible refinements of a triangle *T*.

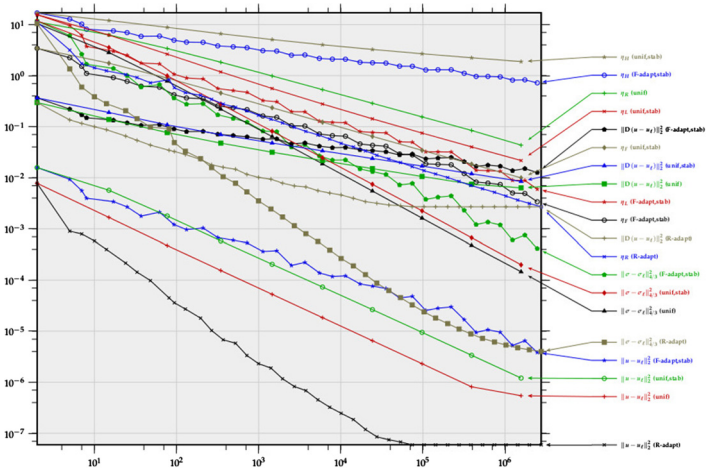

Figure 3 Convergence plot of the two-well benchmark. Errors and error estimators of the two-well benchmark of Subsection 'Two-well benchmark', plotted against the number of degrees of freedom.

η_L converges slightly faster than η_F. This is in accordance to the theory of Section 'Refined analysis for an interface model problem' where η_L is derived from η_F based on additional smoothness assumptions.

Modified two-well benchmark

This subsection concerns a modification of the previous problem with (3.4) and a linear right-hand side for $\beta = 0$ and $f(x) := -\mathrm{div}(\mathrm{D}W(\mathrm{D}u_D(x)))$ and unique solution $u = u_D$ as before. Note that Example 3.3 applies to this problem, and so the proof of Theorem 3.1 yields

$$\|\sigma - \sigma_\ell\|^r_{L^{p'}(\Omega)} + \|u - u_\ell\|^2_{L^2(\Omega)} + |||u_\ell|||^2_\ell \to 0 \ \text{ as } \ell \to \infty$$

and Theorems 5.1–5.4 hold as well. The algorithms of Subsection 'Numerical algorithms' ran with and without stabilisation for uniform and adaptive meshes with the same initial triangulation as in Subsection 'Two-well benchmark' and led to Figure 4 with overall observations of Section 'Conclusions'. The empirical convergence rates for uniform and R- as well as F-adapted mesh-refining are collected in Table 1 for completeness although they are almost identical with those observed in Subsection 'Two-well benchmark'.

Three-well benchmark

The energy density W of ([26], Example 5.9.3, p. 72) is the convex hull of $\min\{|F|^2, |F - (1,0)|^2, |F - (0,1)|^2\}$ with explicit form in ([26], Example 5.6.4, p. 58). Let furthermore $\Omega = (0,1)^2 \subset \mathbb{R}^2$ and $u_D(x_1, x_2) := a(x_1 - 1/4) + a(x_2 - 1/4)$ with $a(t) := t^3/6 + t/8$ for $t \leqslant 0$ and $a(t) := t^5/40 + t^3/8$ for $t \geqslant 0$. Then the energy is given by (1.1) with $\beta = 0$ and $f := -\mathrm{div}\mathrm{D}W(\mathrm{D}u_D)$. The exact solution $u = u_D$ satisfies the interface condition of Section 'Refined analysis for an interface model problem' and allows Theorem 5.3 to hold. Theorems 5.1 and 5.4 do not apply because $\beta = 0$. We use the grid of Figure 5 as initial triangulation to resolve discontinuities in ∇f.

The algorithms of Subsection 'Numerical algorithms' ran with and without stabilisation for uniform and adaptive meshes and led to Figure 6 with overall observations of

Table 1 Observed convergence rates in Figures 3, 4, 6 and 7 for uniform and adaptive mesh refining

Example of subsection		$\|\sigma - \sigma_\ell\|^2_{L^{p'}(\Omega)}$		$\|u - u_\ell\|^2_{L^2(\Omega)}$		η_R	η_F	η_L	$\|D(u - u_\ell)\|^2_{L^2(\Omega)}$		η_H
		unstab.	**stab.**	**unstab.**	**stab.**	**unstab.**	**stab.**	**stab.**	**unstab.**	**stab.**	**stab.**
'Two-well benchmark'	unif	5/3	5/3	3/2	7/5	4/5	4/5	1	3/5	1/2	1/3
	R-adapt	2	7/5	(5/3)	6/5	1	1	1	(2/3)	2/5	2/5
	F-adapt	2	4/3	(5/3)	6/5	1	1	1	(2/3)	2/5	2/5
'Modified two-well benchmark'	unif	5/3	5/3	3/2	7/5	4/5	4/5	1	3/5	1/2	1/3
	R-adapt	2	7/5	(5/3)	6/5	1	1	1	(2/3)	2/5	2/5
	F-adapt	11/5	4/3	(7/4)	6/5	1	1	1	(2/3)	2/5	2/5
'Three-well benchmark'	unif	(1)	3/2	—	7/5	1	4/5	1	—	1/2	2/5
	R-adapt	2	(1/4)	—	(1/4)	1	—	(1/3)	(1)	(1/5)	—
	F-adapt	9/5	1	—	4/5	1	3/5	4/5	—	1/3	1/3
'An optimal design example'	unif					4/5	4/5	6/5			2/5
	R-adapt					1	4/5	6/5			2/5
	F-adapt					1	4/5	1			2/5

Convergence rates are given as powers of the representative mesh-size $1/\sqrt{\text{ndof}}$ which is proportional to H_ℓ on uniform grids. Unavailable values are left blank, non-continuous rates are put in parantheses, inconclusive convergence behaviour is marked by "—".

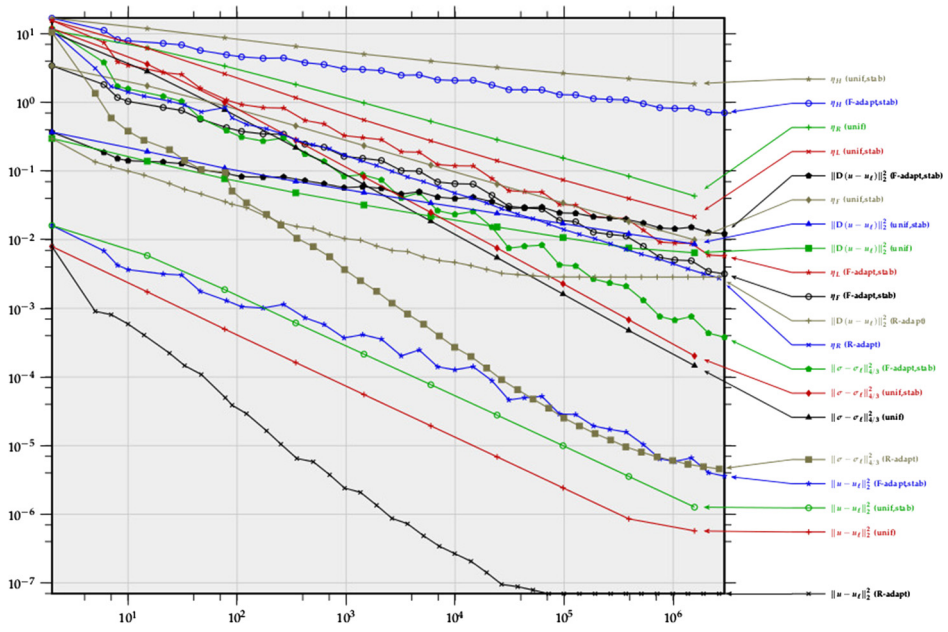

Figure 4 Convergence plot of the modified two-well benchmark. Errors and error estimators of the modified two-well benchmark of Subsection 'Modified two-well benchmark', plotted against the number of degrees of freedom.

Section 'Conclusions'. Beyond those general conclusions, this example demonstrates the difficulties with ill-conditioned Hessians. While the unstabilised method reaches 10^6 degrees of freedom without difficulty on uniform meshes, the adapted algorithms fail without stabilisation beyond 687 324 degrees of freedom (η_F-adaptive) and 33 169 degrees of freedom (η_R-adaptive). MATLAB's error message "Input to EIG must not contain NaN or Inf" indicates that a matrix operation returned non-finite numbers let `fminunc` break down. Undisplayed numerical experiments show condition numbers up to 10^{10} and

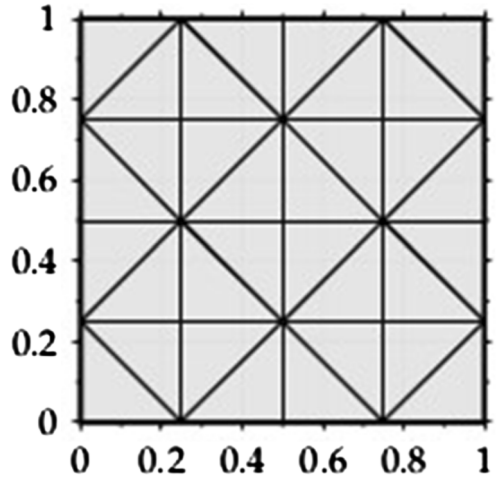

Figure 5 Initial grid for the three-well benchmark. Initial grid for the three-well benchmark of Subsection 'Three-well benchmark'.

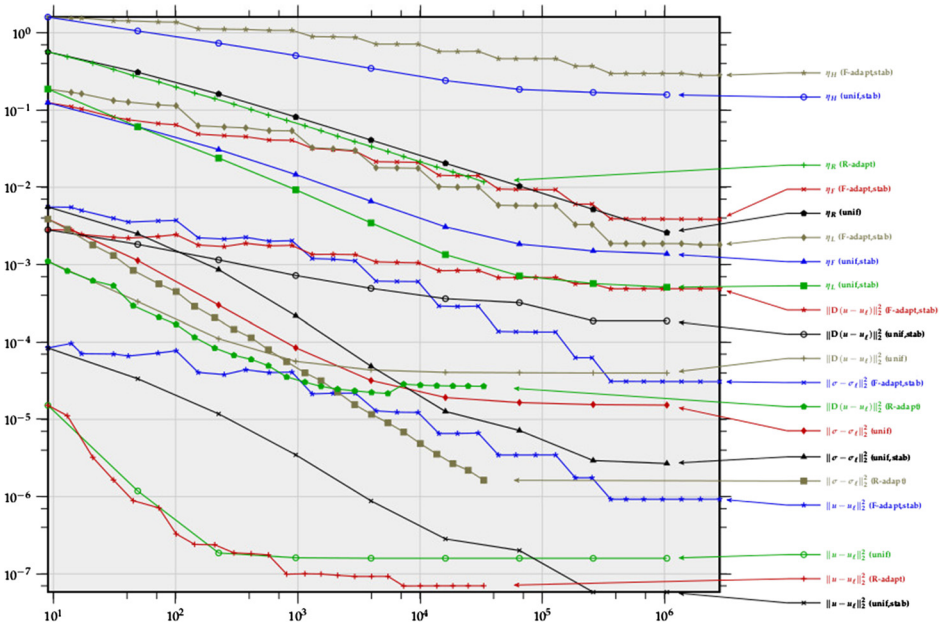

Figure 6 Convergence plot of the three-well benchmark. Errors and error estimators of the three-well benchmark of Subsection 'Three-well benchmark', plotted against the number of degrees of freedom.

beyond. The empirical convergence rates for uniform and R- as well as F-adapted mesh-refining are collected in Table 1. Moreover, Figure 1 in Section 'Background' reveals that stabilisation not only remedies ill-conditioned Hessians but thereby indeed allows for reduced errors in the discrete solution.

An optimal design example

The energy density of the topology optimisation problem of [3,8,29-33] reads

$$W(F) := \phi(|F|) \quad \text{for } F \in \mathbb{R}^2$$

$$\text{with } \phi(t) := \lambda/2 + \begin{cases} t^2 & \text{for } 0 \leqslant t \leqslant \sqrt{\lambda}, \\ 2\sqrt{\lambda}(t - \sqrt{\lambda}/2) & \text{for } \sqrt{\lambda} \leqslant t \leqslant 2\sqrt{\lambda}, \\ t^2/2 + \lambda & \text{for } t \geqslant 2\sqrt{\lambda}. \end{cases}$$

This leads to problem (2.3) with $\beta = 0, \lambda = 0.0084, u_D \equiv 0$ and $f \equiv 1$. Since regularity of the solutions is unclear, only the results of Sections 'Global convergence', 'A posteriori error estimates', 'Refined analysis for an interface model problem' and 'Numerical experiments' apply. As initial triangulation $\mathcal{T}_0$, we use the coarsest cross triangulation $\mathcal{T}_0 = \{\text{conv}\{(0,0),(1,0),(0,1)\}, \text{conv}\{(1,0),(0,1),(1,1)\}\}$ of $\Omega = (0,1)^2$.

The algorithms of Subsection 'Numerical algorithms' ran with and without stabilisation for uniform and adaptive meshes and led to Figure 7 with the overall observations of Section 'Conclusions'. The empirical convergence rates for uniform and R- as well as F-adapted mesh-refining are collected in Table 1. Undocumented experiments with a modified lower-order term f and known exact solution u led to the same convergence rates of the error estimators and confirm their accuracy.

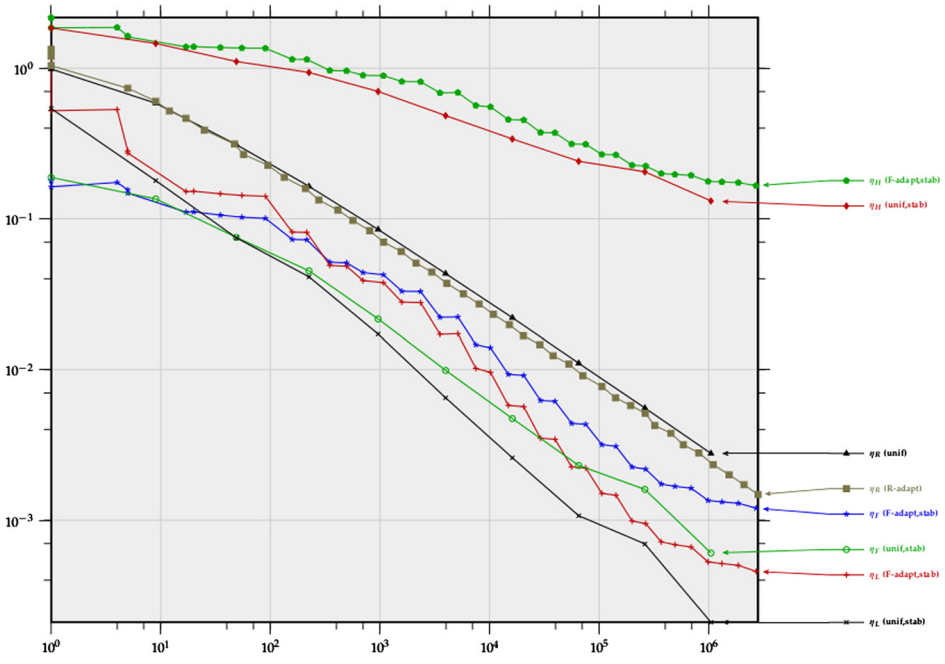

Figure 7 Convergence plot of the optimal design benchmark. Error estimators of the optimal design example of Subsection 'An optimal design example', plotted against the number of degrees of freedom.

Discussion of Empirical Convergence Rates

Global convergence without regularity assumptions

Theorem 3.1 asserts that $\|\sigma - \sigma_\ell\|_{L^{p'}(\Omega)}$, $\beta\|u - u_\ell\|_{L^2(\Omega)}$, and $|||u_\ell|||_\ell$ all tend to zero as $H_\ell \to 0$. The plain convergence result applies to all examples from Subsections 'Two-well benchmark', 'An optimal design example', 'Three-well benchmark', and 'An optimal design example' for the uniform mesh-refinements with $H_{\ell+1} = H_\ell/2$. The numerical experiments, however, show empirical convergence rates displayed in the first columns of Table 1. The adaptive algorithms do not reflect the condition $H_\ell \to 0$ explicitly and hence convergence is not guaranteed a priori. Undisplayed investigations show that indeed in the R-adapted version of the three-well example of Subsection 'Three-well benchmark', this condition $H_\ell \to 0$ does not appear to be true for more than 4 978 degrees of freedom. In all other experiments we observe convergence rates even for unstabilised discretisations.

Empirical convergence rates for interface model problems

Theorem 5.1 provides an a priori error estimate and an estimate of the stabilisation norm. It applies to the benchmark of Subsections 'Two-well benchmark', 'An optimal design example', 'Three-well benchmark', and 'An optimal design example' only, because of $\beta > 0$ and Example 3.3, and the smoothness conditions imposed upon u from Section 'Refined analysis for an interface model problem'. Recall the definitions of $\mathcal{T}_\ell(\Gamma)$, $\Omega_{\Gamma,\ell}$ and $\Omega^{\mathrm{C}}_{\Gamma,\ell}$ from Section 'Refined analysis for an interface model problem' and assume $\|u\|_{L^2(\Omega\setminus\Gamma)} \approx 1 \approx \|u\|_{W^{2,p}(\Omega^{\mathrm{C}}_{\Gamma,\ell})}$, $|\mathcal{T}_\ell(\Gamma)| \approx H_\ell^{-1}$ and $|\Omega_{\Gamma,\ell}| \approx H_\ell$ in this discussion. This leads to a convergence rate of $H_\ell^{2/p}$ for the right-hand side of Theorem 5.1. The observed convergence rates of $\|\sigma - \sigma_\ell\|_{L^{p'}(\Omega)}$ and $\|u - u_\ell\|_{L^2(\Omega)}$ for the stabilised

benchmark examples in Table 1 show convergence rates beyond those guaranteed in Theorem 5.1.

Theorem 5.3 implies, up to perturbations on the boundary,

$$\|\mathrm{D}(u-u_\ell)\|_{L^2(\Omega)} \lesssim \|u-u_\ell\|_{L^2(\Omega)}^{1/3} + |||u_\ell|||_\ell + H_\ell^{-1/2}|||u_\ell|||_\ell^{1/2}\|u-u_\ell\|_{L^2(\Omega)}^{1/2}.$$

Since the exact solutions of Subsections 'Two-well benchmark', 'An optimal design example', 'Three-well benchmark', and 'An optimal design example' are all smooth up to a one-dimensional interface line, Theorem 5.3 applies to these examples. The experiments shows that the right-hand side of Theorem 5.3 is dominated by $H_\ell^{-1/2}|||u_\ell|||_\ell^{1/2}\|u-u_\ell\|_{L^2(\Omega)}^{1/2}$ in all examples and that the inequality is satisfied.

Reliability without regularity assumptions

Up to boundary terms, Theorem 4.1 states

$$\|\sigma-\sigma_\ell\|_{L^{p'}(\Omega)}^2 + \beta\|u-u_\ell\|_{L^2(\Omega)}^2 \lesssim \eta_F\|u-u_\ell\|_{W^{1,p}(\Omega)}.$$

The convergence rates confirm this assertion for the general and rough estimate $\|u-u_\ell\|_{W^{1,p}(\Omega)} \lesssim 1$ in the sense that the rates for η_F are worse than or equal to those of $\|\sigma-\sigma_\ell\|_{L^{p'}(\Omega)}^2$ and $\|u-u_\ell\|_{L^2(\Omega)}^2$. In the numerical examples, $\|u-u_\ell\|_{H^1(\Omega)}$ is computed and displayed in Table 1 and the convergence rates of the product $\|u-u_\ell\|_{H^1(\Omega)}\eta_F$ can be compared with those of $\|\sigma-\sigma_\ell\|_{L^{p'}(\Omega)}^2 + \|u-u_\ell\|_{L^2(\Omega)}^2$. This comparison confirms the above a posteriori error estimate. In the examples with $p=2$ (of Subsections 'Two-well benchmark', 'An optimal design example', 'Three-well benchmark', and 'An optimal design example'), there holds even equality of the convergence rates which demonstrates the efficiency of the estimate of Theorem 4.1.

Efficiency without regularity assumptions

Up to oscillations and the (possibly) higher-order term $\|(\mathrm{id}-I_{\mathrm{F},\ell})\sigma\|_{L^{q'}(\Omega)}$, Theorem 4.2 states

$$\eta_F \lesssim \|\sigma-\sigma_\ell\|_{L^{p'}(\Omega)} + \beta\|u-u_\ell\|_{L^{p'}(\Omega)}.$$

The displayed convergence rates of Table 1 confirm this estimate.

Reliability of the refined a posteriori error control

Theorem 5.4 applies to the example of Subsection 'Two-well benchmark' and states

$$\|\sigma-\sigma_\ell\|_{L^{p'}(\Omega)}^2 + \|u-u_\ell\|_{L^2(\Omega)}^2 \lesssim \eta_L \quad\text{and}\quad \|\mathrm{D}(u-u_\ell)\|_{L^2(\Omega)}^2 \lesssim \eta_H.$$

Table 1 confirms this estimate and shows that the estimators η_L and η_H accurately predict the convergence rate of the errors, even with equality of the convergence rates in the case of adaptive mesh refinements in the examples of Subsections 'Two-well benchmark', 'An optimal design example', 'Three-well benchmark', and 'An optimal design example'.

All displayed convergence rates of η_L are better or at least equal to those of η_F. For instance, for uniform mesh-refining in Subsections 'Two-well benchmark', 'An optimal design example', 'Three-well benchmark', and 'An optimal design example', the error terms $\|\sigma-\sigma_\ell\|_{L^{p'}(\Omega)}^2 + \|u-u_\ell\|_{L^2(\Omega)}^2$ converge with the empirical convergence rate $7/5$ while the upper bound η_F does so with a reduced convergence rate $4/5$. The

refined error estimator η_L is a guaranteed upper bound (via Theorem 5.4) and converges with an empirical convergence rate 1.

Performance of the minimisation algorithm 1

In all numerical experiments of this paper, Algorithm 1 reaches the maximal number 3 of iterations. While this suggests that the optimal s is *not* found after three iterations, undisplayed experiments with higher iteration counts and hence higher computational efforts result solely in marginal improvements.

Conclusions

Effects of stabilisation

The empirical convergence rates of the error estimators η_F, η_R and the errors $\|u - u_\ell\|_{L^2(\Omega)}$ and $\|\sigma - \sigma_\ell\|_{L^{p'}(\Omega)}$ for uniform mesh-refinement with and without stabilisation coincide. This indicates that the choice $\gamma = 1$ leads to some significant perturbation but maintains the correct convergence rate at the same time. This is different for adaptive mesh refinement with less optimal convergence rates. Our conclusion is that an improved adaptive algorithm has to be developed with balance of local mesh-refinement and global stabilisation parameters in future research. The tested algorithm from Subsection 'Numerical algorithms' does neither reflect the effects of stabilisation nor that of inexact solve.

Another important aspect of the stabilisation is the regularisation of the Hessian in the step SOLVE of Subsection 'Numerical algorithms'. In the three-well problem of Subsection 'Three-well benchmark', the unstabilised adaptive algorithms fail.

Adaptive versus uniform mesh-refinement

The overall empirical convergence rates of the errors and estimators of the unstabilised computation for adaptive mesh-refinements are better than those for uniform mesh-refinements. This is in contrast to the stabilised computation, where the true errors $\|\sigma - \sigma_\ell\|_{L^{p'}(\Omega)}$ and $\|u - u_\ell\|_{L^2(\Omega)}$ behave better for uniform compared with the two adaptive mesh-refinments (with the exception in Subsection 'An optimal design example' where there is equality). It is observed that adaptivity does not necessarily improve the converegnce rates of the error $\|\sigma - \sigma_\ell\|_{L^{p'}(\Omega)}$ and $\|u - u_\ell\|_{L^2(\Omega)}$ in a stabilised computation. Surprisingly, the convergence of the gradient errors $\|\mathrm{D}(u - u_\ell)\|_{L^2(\Omega)}$ are slightly improved in the instabilised calculation by adaptive mesh-refinements. The adaptive mesh-refinement is expected to reduce the a posteriori error estimators in the first place: cf. [1,34] for the estimator reduction property. Indeed, the convergence rates of the a posteriori error estimators η_R, η_F, η_L, η_H are improved (or optimal) for adaptive mesh-refinements (except for the three-well example of Subsection 'Three-well benchmark').

Strong convergence of the gradients

The convergence of the gradient error of the stabilised problem surpasses the expectations of [12] in Subsection 'An optimal design example' but fails to do so in Subsections 'Two-well benchmark', 'An optimal design example', 'Three-well benchmark', and 'An optimal design example'. The improved error estimator η_H shows the same convergence rate as the error of the gradients in Subsections 'Two-well benchmark',

'An optimal design example', 'Three-well benchmark', and 'An optimal design example'. This holds for uniform and for adapted mesh refinements and suggests that η_H is in fact reliable and efficient for $\beta > 0$.

Guaranteed error control

The assertion on η_F in Theorem 4.1 is reflected in the numerical examples in that the stress approximations converge faster than η_F in all cases. This suggests that the estimate $\|u - u_\ell\|_{W^{1,p}(\Omega)} \lesssim 1$ is by far too pessimistic. In fact, the benchmark examples with known exact solution fulfil $\|\sigma - \sigma_\ell\|^2_{L^2(\Omega)} \lesssim \eta_F \|u - u_\ell\|_{H^1(\Omega)}$. Similar affirmative conclusions follow for Theorem 4.2 and 5.4.

Reliability-efficiency gap

In comparison with the residual-based error estimator of [7,18], the new a posteriori error estimators η_L and η_H of Theorem 5.4 lead to refined error control. The improvement is marginal for uniform meshes without stabilisation but significant for adaptive stabilised computations. η_L and η_H match the convergence of the errors and so narrow the reliability-efficiency gap.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally to all parts of this article. All authors read and approved the final manuscript.

Received: 29 July 2013 Accepted: 6 December 2013
Published: 29 January 2014

References

1. Carstensen C (2008) Convergence of an adaptive fem for a class of degenerate convex minimisation problems. IMA J Numer Anal 28(3): 423–439
2. Dacorogna B (2008) Direct methods in the calculus of variations, 2nd Ed. Applied Mathematical Sciences 78. Springer, Berlin. xii
3. Carstensen C, Müller S (2002) Local stress regularity in scalar non-convex variational problems. SIAM J Math Anal 34(2): 495–509
4. Chipot M (2000) Elements of Nonlinear Analysis. Birkhäuser Advanced Texts. Basel, Birkhäuser. vi
5. Müller S (1999) Variational models for microstructure and phase transisions. In: Hildebrandt S, et al. (eds) Calculus of variations and geometric evolution problems. Lectures given at the 2nd session of the, Centro Internazionale Matematico Estivo (CIME), Cetraro, Italy, June 15–22, 1996, Lect. Notes Math., 1713. Springer, Berlin, pp 85–210
6. Ball JM, James RD (1992) Proposed experimental tests for the theory of fine microstructures and the two-well problem. Phil Trans R Soc Lond A 338: 389–450
7. Carstensen C, Plecháč P (1997) Numerical solution of the scalar double-well problem allowing microstructure. Math Comp 66(219): 997–1026
8. Bartels S, Carstensen C (2007) A convergent adaptive finite element method for an optimal design problem. Numer Math 108: 359–385
9. Goodman J, Kohn RV, Reyna L (1986) Numerical study of a relaxed variational problem from optimal design. Comput Methods Appl Mech Eng 57: 107–127
10. Carstensen C, Klose R (2003) Guaranteed a posteriori finite element error control for the p-Laplace problem. SIAM J Sci Comput 25: 792–814
11. Bartels S, Carstensen C, Plecháč P, Prohl A (2004) Convergence for stabilisation of degenerate convex minimsation problems. IFB 6(2): 253–269
12. Boiger W, Carstensen C (2010) On the strong convergence of gradients in stabilised degenerate convex minimisation problems. SIAM J Numer Anal 47(6): 4569–4580
13. Ciarlet PG (2002) The finite element method for elliptic problems. Society for Industrial Mathematics, Philadelphia, PA, USA
14. El Alaoui L, Ern A, Vohralík M (2011) Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems. Comp Meth Appl Mech Eng 200(37–40): 2782–2795
15. Ern A, Nicaise S, Vohralík M (2007) An accurate *h*(div) flux reconstruction for discontinuous galerkin approximations of elliptic problems. C R, Math, Acad Sci Paris 345(12): 709–712
16. Luce R, Wohlmuth B (2004) A local a posteriori error estimator based on equilibrated fluxes. SIAM J Numer Anal 42(4): 1394–1414
17. Ainsworth M (2005) A synthesis of a posteriori error estimation techniques for conforming, non-conforming and discontinuous galerkin finite element methods. American Mathematical Society (AMS), Providence

18. Carstensen C, Jochimsen K (2003) Adaptive finite element methods for microstructures? Numerical experiments for a Two-well benchmark. Computing 71: 175–204
19. Chipot M, Evans LC (1986) Linearisation at infinity and Lipschitz estimates for certain problems in the calculus of variations. Proc Roy Soc Edinburgh Sect A 102(3–4): 291–303
20. Bartels S, Carstensen C, Dolzmann G (2004) Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error anylysis. Numer Math 99(1): 1–24
21. Knees D (2008) Global stress regularity of convex and some nonconvex variational problems. Ann Mat Pura Appl (4) 187(1): 157–184
22. Brezzi F, Fortin M (1991) Mixed and hybrid finite element methods. Springer series in computational mathematics. Springer-Verlag, New York
23. Acosta G, Durán RG (2004) An optimal Poincaré inequality in L^1 for convex domains. Proc Amer Math Soc 132(1): 195–202
24. Bergh J, Löfstrom J (1976) Interpolation spaces. Springer-Verlag, Berlin
25. Brenner SC, Scott LR (2002) The mathematical theory of finite element methods, 2nd Ed. Texts in Applied Mathematics. 15. Springer, Berlin. p361, xv
26. Bartels S (2001) Numerical analysis of some non-convex variational problems. Ph.D. thesis. Christian-Albrechts Universität zu Kiel, Kiel, Germany. [http://eldiss.uni-kiel.de/macau/receive/dissertation_diss_00000519]
27. Alberty J, Carstensen C, Funken SA (1999) Remarks around 50 lines of Matlab: short finite element implementation. Numer. Algorithms 20(2–3): 117–137
28. Repin SI, Sauter S, Smolianski A (2003) A posteriori error estimation for the dirichlet problem with account of the error in the approximation of boundary conditions. Computing 70(3): 205–233
29. Carstensen C, Günther D, Rabus H (2012) Mixed finite element method for a degenerate convex variational problem from topology optimization. SIAM J Math Anal 50(2): 522–543
30. Murat F, Tartar L (1985) Calcul des variations et homogénéisation. In: Bergman D, et al. (eds) Homogenization methods: theory and applications in physics. Collect Dir Études Rech Élec France, vol. 57. Éditions Eyrolles, Paris, France, pp 319–369
31. Kohn RV, Strang G (1986) Optimal design and relaxation of variational problems I–III. Comm Pure Appl Math 39(1–3): 113–137139182353377
32. Kawohl B, Stará J, Wittum G (1991) Analysis and numerical studies of a problem of shape design. Arch Rational Mech Anal 114(4): 349–363
33. Glowinski R, Lions J-L, Trémolières R (1981) Numerical analysis of variational inequalities. Studies in Mathematics and its Applications, vol. 8. North-Holland Publishing Co., Amsterdam. p 776
34. Cascón JM, Kreuzer C, Nochetto RH, Siebert KG (2008) Quasi-optimal convergence rate for an adaptive finite element method. SIAM J Numer Anal 46(5): 2524–2550

doi:10.1186/2213-7467-1-5
Cite this article as: Boiger and Carstensen: **A posteriori error analysis of stabilised FEM for degenerate convex minimisation problems under weak regularity assumptions.** *Advanced Modeling and Simulation in Engineering Sciences* 2013 **1**:5.