

Convergent adaptive hybrid higher-order schemes for convex minimization

Carsten Carstensen, Ngoc Tien Tran

Angaben zur Veröffentlichung / Publication details:

Carstensen, Carsten, and Ngoc Tien Tran. 2022. "Convergent adaptive hybrid higher-order schemes for convex minimization." *Numerische Mathematik* 151 (2): 329–67. <https://doi.org/10.1007/s00211-022-01284-1>.



Convergent adaptive hybrid higher-order schemes for convex minimization

Carsten Carstensen¹ · Ngoc Tien Tran²

Received: 19 July 2021 / Revised: 20 February 2022 / Accepted: 27 February 2022 /

Published online: 19 April 2022

© The Author(s) 2022

Abstract

This paper proposes two convergent adaptive mesh-refining algorithms for the hybrid high-order method in convex minimization problems with two-sided p -growth. Examples include the p -Laplacian, an optimal design problem in topology optimization, and the convexified double-well problem. The hybrid high-order method utilizes a gradient reconstruction in the space of piecewise Raviart–Thomas finite element functions without stabilization on triangulations into simplices or in the space of piecewise polynomials with stabilization on polytopal meshes. The main results imply the convergence of the energy and, under further convexity properties, of the approximations of the primal resp. dual variable. Numerical experiments illustrate an efficient approximation of singular minimizers and improved convergence rates for higher polynomial degrees. Computer simulations provide striking numerical evidence that an adopted adaptive HHO algorithm can overcome the Lavrentiev gap phenomenon even with empirical higher convergence rates.

Keywords Convex minimization · Degenerate convex · Convexity control · Hybrid high-order · p -Laplacian · Optimal design problem · Double-well problem · A posteriori · Adaptive mesh-refining · Convergence · Lavrentiev gap

Mathematics Subject Classification 65N12 · 65N30 · 65Y20

✉ Carsten Carstensen
cc@math.hu-berlin.de

Ngoc Tien Tran
ngoc.tien.tran@uni-jena.de

¹ Department of mathematics, Humboldt-Universität zu Berlin, Berlin, Germany

² Department of mathematics, Friedrich-Schiller-Universität Jena, Jena, Germany

1 Introduction

Adaptive mesh-refining is vital in the computational sciences and engineering with optimal rates known in many linear problems [19, 21, 29, 54]. Besides eigenvalue problems [6, 23, 24, 33] much less is known for stationary nonlinear PDEs. The few positive results in the literature concern mainly conforming FEM with plain convergence results [4, 12, 22, 41, 56]. An important exception is the p -Laplacian in [5], where the notion of a quasi-norm enables two-sided error control. The next larger class of convex minimization problems from [27] emerged in the relaxation of non-convex minimization problems with enforced microstructures and this is in the focus of this paper. This class is characterized by a two-sided growth condition on a C^1 energy density W with an additional convexity control that enables a unique stress $DW(Du)$ independent of the multiple minimizers u on the continuous level. In fact, there is no further control of the convex closed set of minimizers in beyond a priori boundedness. This leads to the reliability-efficiency gap [14] in the a posteriori error control: If the mesh-size tends to zero, the known guaranteed lower and upper error bounds converge with a different convergence rate. In other words, the efficiency index tends to infinity. This dramatic loss of sharp error control does not prevent convergence of an adaptive algorithm in general, but it makes the analysis of plain convergence much harder and seemingly disables any proof of optimal rates.

The numerical experiments in [28, 38] motivate this paper on the adaptive HHO. The only nonconforming scheme known to converge for general convex minimization problems is [52] for the first-order Crouzeix–Raviart schemes and, according to the knowledge of the authors, there is no contribution for the convergence of an adaptive higher-order nonconforming scheme for nonlinear PDEs in the literature. In fact, this paper is the first one to guarantee plain convergence even for linear PDEs for the HHO schemes at all. The reason is a negative power of the mesh-size in the stabilization terms that is overcome in dG schemes for linear PDEs by over-penalization in [7] to be close to conforming approximations (and then enable arguments for optimal convergence rates) and recently by generalized Galerkin solutions in a limit space in [47] for plain convergence. One advantage of the HHO methodology is the absence of a stabilization parameter and, hence, this argument is not employed in this paper.

The main contributions of this paper are adaptive HHO methods with and without stabilization with guaranteed plain convergence for the class of convex minimization problems from [27]. Three types of results are available for those schemes.

- (a) If W is C^1 and convex with two-sided p -growth, then the minimal discrete energies converge to the exact minimal energy.
- (b) If furthermore W satisfies the convexity control in the class of degenerate convex minimization problems of [27], then the discrete stress approximations converge to the (unique) exact stress σ .
- (c) If W is even strongly convex, then the discrete approximations of the gradients converge to the gradient Du of the (unique) exact minimizer u .

The two-sided growth condition excludes problems that exhibit the Lavrentiev gap phenomenon [48] and so we only comment on the lowest-order schemes that overcome the Lavrentiev gap owing to the Jensen inequality.

Numerical experiments are carried out on simplicial meshes, but the design of the stabilized HHO method allows for polytopal meshes for a fairly flexible mesh-design, e.g., in 3D. The mesh-refinement of those schemes is less elaborated (e.g., in comparison with [55] on simplicial meshes) and remains as an important aspect for future research.

The remaining parts of this paper are organized as follows. Section 2 introduces the continuous minimization problem, the adaptive mesh-refining algorithm, and the main results of this paper. Section 3 reviews the discretization with the HHO methodology on simplicial triangulations without stabilization. Section 4 departs from discrete compactness, proves the plain convergence of an adaptive scheme, and concludes with an application to the Lavrentiev gap. Section 5 treats HHO methods on general polytopal meshes with stabilization and proves the results of Sect. 4.2. Numerical results for three model examples from Sect. 2.4 below are presented in Sect. 6 with conclusions drawn from the numerical experiments.

2 Mathematical setting and main results

This paper analyzes the convergence of an adaptive mesh-refining algorithm based on the hybrid high-order methodology [35, 37, 39] for convex minimization problems with a two-sided p -growth.

2.1 Continuous problem

Given a bounded polyhedral Lipschitz domain $\Omega \subset \mathbb{R}^n$ and $1 < p < \infty$, let $W \in C^1(\mathbb{M})$ with $\mathbb{M} := \mathbb{R}^{m \times n}$ satisfy

- (A1) (convexity) W is convex;
- (A2) (two-sided growth) $c_1|A|^p - c_2 \leq W(A) \leq c_3|A|^p + c_4$ for all $A \in \mathbb{M}$.

The constants $c_1, c_3 > 0$ and $c_2, c_4 \geq 0$ are universal in this paper and independent of the argument $A \in \mathbb{M}$; the same universality applies to c_5, c_6 in (2.5)–(2.6). Throughout this paper, the boundary $\partial\Omega$ of the domain Ω is divided into a compact Dirichlet part Γ_D with positive surface measure and a relatively open (and possibly empty) Neumann part $\Gamma_N = \partial\Omega \setminus \Gamma_D$. Given $f \in L^{p'}(\Omega; \mathbb{R}^m)$, $g \in L^{p'}(\Gamma_N; \mathbb{R}^m)$ with $1/p + 1/p' = 1$, and $u_D \in V := W^{1,p}(\Omega; \mathbb{R}^m)$, minimize the energy functional

$$E(v) := \int_{\Omega} (W(Dv) - f \cdot v) \, dx - \int_{\Gamma_N} g \cdot v \, ds \tag{2.1}$$

among admissible functions $v \in \mathcal{A} := u_D + V_D$ subject to the Dirichlet boundary condition $v|_{\Gamma_D} = u_D|_{\Gamma_D}$ and $V_D := \{v \in V : v|_{\Gamma_D} \equiv 0\}$.

2.2 Adaptive hybrid high-order method (AHHO)

The adaptive algorithm computes a sequence of discrete approximations of the minimal energy $\min E(\mathcal{A})$ in the affine space $\mathcal{A} = u_D + V_D$ of admissible functions in a

successive loop over the steps outlined below. The first version of the adaptive algorithm focuses on the newest-vertex-bisection (NVB) [55] and the first HHO method without stabilization on triangulations into simplices. It will be generalized to polytopal meshes in Sect. 5.

1. INPUT. The input is a regular initial triangulation \mathcal{T}_0 of Ω into simplices, a polynomial degree $k \geq 0$, a positive parameter $0 < \varepsilon \leq k + 1$, and a bulk parameter $0 < \theta < 1$.
2. SOLVE. Let \mathcal{T}_ℓ denote the triangulation associated to the level $\ell \in \mathbb{N}_0$ with the set of all sides \mathcal{F}_ℓ . The hybrid high-order method utilizes the discrete ansatz space $V(\mathcal{T}_\ell) := P_k(\mathcal{T}_\ell; \mathbb{R}^m) \times P_k(\mathcal{F}_\ell; \mathbb{R}^m)$ with a split of the discrete variables $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell})$ into a volume variable $v_{\mathcal{T}_\ell} \in P_k(\mathcal{T}_\ell; \mathbb{R}^m)$ and a skeleton variable $v_{\mathcal{F}_\ell} \in P_k(\mathcal{F}_\ell; \mathbb{R}^m)$ of polynomial degree at most $k \geq 0$ with respect to the simplices (\mathcal{T}_ℓ) and the sides (\mathcal{F}_ℓ) in the triangulation \mathcal{T}_ℓ . The proposed numerical scheme replaces Dv in (2.1) by a gradient reconstruction \mathcal{G}_ℓ in the space of piecewise Raviart–Thomas finite element functions $\Sigma(\mathcal{T}_\ell) = \text{RT}_k^{\text{pw}}(\mathcal{T}_\ell; \mathbb{M})$ for a shape-regular triangulation \mathcal{T}_ℓ of Ω into simplices. The details on the gradient reconstruction \mathcal{G}_ℓ are postponed to Sect. 3.4. The discrete problem computes a discrete minimizer u_ℓ of

$$E_\ell(v_\ell) := \int_\Omega (W(\mathcal{G}v_\ell) - f \cdot v_{\mathcal{T}_\ell}) \, dx - \int_{\Gamma_N} g \cdot v_{\mathcal{F}_\ell} \, ds \tag{2.2}$$

among $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) \in \mathcal{A}(\mathcal{T}_\ell)$ with the discrete analog $\mathcal{A}(\mathcal{T}_\ell)$ of \mathcal{A} so that $v_{\mathcal{F}_\ell}|_F = \Pi_F^k u_D$ for any Dirichlet side $F \in \mathcal{F}_\ell(\Gamma_D)$, where Π_F^k is the L^2 projection onto the polynomials $P_k(F)$ of degree at most k . Let $\sigma_\ell := \Pi_{\Sigma(\mathcal{T}_\ell)} DW(\mathcal{G}u_\ell) \in \Sigma(\mathcal{T}_\ell)$ be the L^2 projection of $DW(\mathcal{G}u_\ell)$ onto $\Sigma(\mathcal{T}_\ell)$. Further details on the hybrid high-order method follow in Sect. 3 below.

3. REFINEMENT INDICATORS. The computation of the refinement indicator η_ℓ utilizes an elliptic potential reconstruction $\mathcal{R}_\ell u_\ell \in P_{k+1}(\mathcal{T}_\ell; \mathbb{R}^m)$ of the discrete minimizer $u_\ell = (u_{\mathcal{T}_\ell}, u_{\mathcal{F}_\ell}) \in \mathcal{A}(\mathcal{T}_\ell)$ computed in SOLVE. The definition of $\mathcal{R}_\ell u_\ell$ follows in (3.4)–(3.5) below. Any interior side $F \in \mathcal{F}_\ell(\Omega)$ is shared by two simplices $T_+, T_- \in \mathcal{T}_\ell$ with $F = T_+ \cap T_-$. The jump $[\mathcal{R}_\ell u_\ell]_F$ along F is defined by $[\mathcal{R}_\ell u_\ell]_F := (\mathcal{R}_\ell u_\ell)|_{T_+} - (\mathcal{R}_\ell u_\ell)|_{T_-} \in P_{k+1}(F; \mathbb{R}^m)$. Given a positive parameter $0 < \varepsilon \leq k + 1$, compute the local refinement indicator

$$\begin{aligned} \eta_\ell^{(\varepsilon)}(T) := & |T|^{(\varepsilon p - p)/n} \|\Pi_T^k(\mathcal{R}_\ell u_\ell - u_T)\|_{L^p(T)}^p + |T|^{\varepsilon p'/n} \|\sigma_\ell - DW(\mathcal{G}u_\ell)\|_{L^{p'}(T)}^{p'} \\ & + |T|^{p'/n} \|(1 - \Pi_T^k)f\|_{L^{p'}(T)}^{p'} + |T|^{1/n} \sum_{F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Gamma_N)} \|(1 - \Pi_F^k)g\|_{L^{p'}(F)}^{p'} \\ & + |T|^{(\varepsilon p + 1 - p)/n} \left(\sum_{F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Gamma_D)} \|\mathcal{R}_\ell u_\ell - u_D\|_{L^p(F)}^p \right. \\ & \left. + \sum_{F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Omega)} \|[\mathcal{R}_\ell u_\ell]_F\|_{L^p(F)}^p + \sum_{F \in \mathcal{F}_\ell(T)} \|\Pi_F^k((\mathcal{R}_\ell u_\ell)|_T - u_F)\|_{L^p(F)}^p \right) \end{aligned} \tag{2.3}$$

for all $T \in \mathcal{T}_\ell$ of volume $|T|$ and sides $\mathcal{F}_\ell(T)$ with the abbreviation $u_T := u_{\mathcal{T}_\ell}|_T$ and $u_F := u_{\mathcal{F}_\ell}|_F$. Let $\eta_\ell^{(\varepsilon)} := \sum_{T \in \mathcal{T}_\ell} \eta_\ell^{(\varepsilon)}(T)$. The refinement indicator is motivated by the discrete compactness from Theorem 4.1 below. In fact, if $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)} = 0$, then there exists a $v \in \mathcal{A}$ such that, up to a subsequence, $\mathcal{G}_\ell u_\ell \rightharpoonup \nabla v$ weakly in $L^p(\Omega; \mathbb{M})$ and $u_{\mathcal{T}_\ell} \rightharpoonup v$ weakly in $L^p(\Omega; \mathbb{R}^m)$ as $\ell \rightarrow \infty$. It turns out that v is a minimizer of the continuous energy E from (2.1).

4. MARK and REFINER. Given a positive bulk parameter $0 < \theta < 1$, select a subset $\mathfrak{M}_\ell \subset \mathcal{T}_\ell$ of minimal cardinality such that

$$\theta \eta_\ell^{(\varepsilon)} \leq \eta_\ell^{(\varepsilon)}(\mathfrak{M}_\ell) := \sum_{T \in \mathfrak{M}_\ell} \eta_\ell^{(\varepsilon)}(T). \tag{2.4}$$

This marking strategy is known as Dörfler marking. The marked simplices are refined by the newest-vertex bisection [55] to define $\mathcal{T}_{\ell+1}$.

5. OUTPUT. The output is a sequence of shape-regular triangulations $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0}$, the corresponding discrete minimizers $(u_\ell)_{\ell \in \mathbb{N}_0}$, the discrete stresses $(\sigma_\ell)_{\ell \in \mathbb{N}_0}$, and the refinement indicators $(\eta_\ell^{(\varepsilon)})_{\ell \in \mathbb{N}_0}$. On each level $\ell \geq 0$, let $\mathcal{J}_\ell u_\ell \in V$ denote the conforming post-processing of u_ℓ from Lemma 3.4 below.

2.3 Main results

The main results establish the convergence of the sequence $(E_\ell(u_\ell))_{\ell \in \mathbb{N}_0}$ of minimal discrete energies computed by AHHO towards the exact minimal energy.

Theorem 2.1 (plain convergence) *Given the input \mathcal{T}_0 , $k \in \mathbb{N}_0$, $0 < \varepsilon \leq k + 1$, $0 < \theta < 1$, let $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0}$, $(u_\ell)_{\ell \in \mathbb{N}_0}$, and $(\sigma_\ell)_{\ell \in \mathbb{N}_0}$ be the output of the adaptive algorithm AHHO from Sect. 2.2. Assume that W satisfies (A1)–(A2), then (a)–(d) hold.*

- (a) $\lim_{\ell \rightarrow \infty} E_\ell(u_\ell) = \min E(\mathcal{A})$.
- (b) *The sequence of the post-processing $(\mathcal{J}_\ell u_\ell)_{\ell \in \mathbb{N}_0}$ is bounded in $V = W^{1,p}(\Omega; \mathbb{R}^n)$ and any weak accumulation point of $(\mathcal{J}_\ell u_\ell)_{\ell \in \mathbb{N}_0}$ in V minimizes E in \mathcal{A} .*
- (b) *Suppose there exists $c_5 > 0$ such that W satisfies, for all $A, B \in \mathbb{M}$,*

$$|A - B|^r \leq c_5(1 + |A|^s + |B|^s)(W(A) - W(B) - DW(B) : (A - B)) \tag{2.5}$$

with parameters r, s from Table 1. Then the minimizer u of E in \mathcal{A} is unique and $\lim_{\ell \rightarrow \infty} \mathcal{G}_\ell u_\ell = Du$ (strongly) in $L^p(\Omega; \mathbb{M})$ holds for the entire sequence.

- (d) *Suppose there exists $c_6 > 0$ such that W satisfies, for all $A, B \in \mathbb{M}$,*

$$|DW(A) - DW(B)|^{\tilde{r}} \leq c_6(1 + |A|^{\tilde{s}} + |B|^{\tilde{s}}) \times (W(A) - W(B) - DW(B) : (A - B)) \tag{2.6}$$

with parameters \tilde{r}, \tilde{s} from Table 1. Then the stress $\sigma := DW(Du) \in L^{p'}(\Omega; \mathbb{M})$ is unique (independent of the choice of a (possibly nonunique) minimizer u) and $\lim_{\ell \rightarrow \infty} DW(\mathcal{G}_\ell u_\ell) = \sigma$ (strongly) in $L^{p'}(\Omega; \mathbb{M})$ and $\sigma_\ell \rightharpoonup \sigma$ (weakly) in $L^{p'}(\Omega; \mathbb{M})$ hold for the entire sequence.

Table 1 Parameters $r, s, \tilde{r}, \tilde{s}$ in Theorem 2.1 and t, \tilde{t} in Sect. 4.2

| Case | r | s | t | \tilde{r} | \tilde{s} | \tilde{t} |
|---------------------|-----|---------|-------|-------------|-------------|-------------|
| $2 \leq p < \infty$ | p | 0 | 1 | 2 | $p - 2$ | $2/p'$ |
| $1 < p < 2$ | 2 | $2 - p$ | $2/p$ | p' | 0 | 1 |

A second focus is on the classical HHO method [34, 35, 37] on general polytopal meshes \mathcal{M}_ℓ with a stabilization $s_\ell(\bullet, \bullet)$ defined in (5.1) below. The convergence of AHHO for the stabilized HHO method on polytopal meshes is established under two assumptions (M1)–(M2). Further details on (M1)–(M2) and on the stabilized HHO method follow in Sect. 5.

Theorem 2.2 (plain convergence for stabilized HHO) *Given the input $\mathcal{M}_0, k \in \mathbb{N}_0, 0 < \varepsilon \leq \min\{k+1, (k+1)/(p-1)\}, 0 < \theta < 1$, let $(\mathcal{M}_\ell)_{\ell \in \mathbb{N}_0}, (u_\ell)_{\ell \in \mathbb{N}_0}$, and $(\sigma_\ell)_{\ell \in \mathbb{N}_0}$ be the output of the adaptive algorithm from Sect. 2.2. Suppose that (M1)–(M2) hold, then (a)–(d) from Theorem 2.1 hold verbatim and $\lim_{\ell \rightarrow \infty} s_\ell(u_\ell; u_\ell) = 0$.*

Notice that an additional restriction on the parameter ε is imposed in Theorem 2.2 to control the stabilization s_ℓ . The proofs of Theorems 2.1 and 2.2 are postponed to Sects. 4 and 5.

2.4 Examples

Theorem 2.1 applies to the following scalar examples with $m = 1$.

2.4.1 p -Laplace

The minimization of the energy $E : \mathcal{A} \rightarrow \mathbb{R}$ with the energy density $W \in C^1(\mathbb{R}^n)$,

$$W(a) := |a|^p/p \quad \text{for any } a \in \mathbb{R}^n \text{ with } 1 < p < \infty,$$

is related to the nonlinear PDE $-\operatorname{div} \sigma = f \in L^{p'}(\Omega)$ with $\sigma := \nabla W(\nabla u) = |\nabla u|^{p-2} \nabla u \in L^{p'}(\Omega; \mathbb{R}^n)$ subject to the boundary conditions $\sigma \nu = g$ on Γ_N and $u = u_D$ on Γ_D . The energy density W satisfies (A1)–(A2) and (2.5)–(2.6) [15, 44]. It is worth noticing that the convergence results of this paper are new even for a linear model problem with $p = 2$ for the two HHO algorithms.

2.4.2 Optimal design problem

The optimal design problem seeks the optimal distribution of two materials with fixed amounts to fill a given domain for maximal torsion stiffness [4, 46]. For fixed parameters $0 < \xi_1 < \xi_2$ and $0 < \mu_1 < \mu_2$ with $\xi_1 \mu_2 = \xi_2 \mu_1$, the energy density $W(a) := \psi(\xi), a \in \mathbb{R}^n, \xi := |a| \geq 0$ with

$$\psi(\xi) := \begin{cases} \mu_2 \xi^2 / 2 & \text{if } 0 \leq \xi \leq \xi_1, \\ \xi_1 \mu_2 (\xi - \xi_1 / 2) & \text{if } \xi_1 \leq \xi \leq \xi_2, \\ \mu_1 \xi^2 / 2 - \xi_1 \mu_2 (\xi_1 / 2 - \xi_2 / 2) & \text{if } \xi_2 \leq \xi \end{cases}$$

satisfies (A1)–(A2) and (2.6) [4, Prop. 4.2].

2.4.3 Relaxed two-well problem

Given distinct $F_1, F_2 \in \mathbb{R}^n$ in the two-well problem of [30], the convex envelope W of $|F - F_1|^2 |F - F_2|^2$ for $F \in \mathbb{R}^n$ reads

$$W(F) = \max\{0, |F - B|^2 - |A|^2\}^2 + 4(|A|^2 |F - B|^2 - (A \cdot (F - B))^2)$$

with $A = (F_2 - F_1)/2, B = (F_1 + F_2)/2$, and satisfies (A1)–(A2) and (2.6) [22, 27].

2.5 Notation

Standard notation for Sobolev and Lebesgue functions applies throughout this paper with the abbreviations $V := W^{1,p}(\Omega; \mathbb{R}^m) = W^{1,p}(\Omega)^m$ and $V_D := W_D^{1,p}(\Omega; \mathbb{R}^m) = \{v \in V : v|_{\Gamma_D} = 0\}$. In particular, $(\bullet, \bullet)_{L^2(\Omega)}$ denotes the scalar product of $L^2(\Omega)$ and $W^{p'}(\text{div}, \Omega; \mathbb{M}) := W^{p'}(\text{div}, \Omega)^m$ is the matrix-valued version of

$$W^{p'}(\text{div}, \Omega) := \{\tau \in L^{p'}(\Omega; \mathbb{R}^n) : \text{div} \tau \in L^{p'}(\Omega)\}. \tag{2.7}$$

For any $A, B \in \mathbb{M} := \mathbb{R}^{m \times n}$, $A : B$ denotes the Euclidean scalar product of A and B , which induces the Frobenius norm $|A| := (A : A)^{1/2}$ in \mathbb{M} . The context-dependent notation $|\bullet|$ denotes the length of a vector, the Frobenius norm of a matrix, the Lebesgue measure of a subset of \mathbb{R}^n , or the counting measure of a discrete set. For $1 < p < \infty, p' = p/(p - 1)$ denotes the Hölder conjugate of p with $1/p + 1/p' = 1$. The notation $A \lesssim B$ abbreviates $A \leq CB$ for a generic constant C independent of the mesh-size and $A \approx B$ abbreviates $A \lesssim B \lesssim A$.

3 Hybrid high-order method without stabilization

This section recalls the discrete ansatz space and reconstruction operators from the HHO methodology [35, 37, 39] for convenient reading.

3.1 Triangulation

A regular triangulation \mathcal{T}_ℓ of Ω in the sense of Ciarlet is a finite set of closed simplices T of positive volume $|T| > 0$ with boundary ∂T and outer unit normal ν_T such that $\cup_{T \in \mathcal{T}_\ell} T = \overline{\Omega}$ and two distinct simplices are either disjoint or share one common (lower-dimensional) subsimplex (vertex or edge in 2D and vertex, edge, or face in 3D).

Let $\mathcal{F}_\ell(T)$ denote the set of the $n + 1$ hyperfaces of T , called sides of T . Define the set of all sides $\mathcal{F}_\ell := \cup_{T \in \mathcal{T}_\ell} \mathcal{F}_\ell(T)$, the set of interior sides $\mathcal{F}_\ell(\Omega) := \mathcal{F}_\ell \setminus \{F \in \mathcal{F}_\ell : F \subset \partial\Omega\}$, the set of Dirichlet sides $\mathcal{F}_\ell(\Gamma_D) := \{F \in \mathcal{F}_\ell : F \subset \Gamma_D\}$, and the set of Neumann sides $\mathcal{F}_\ell(\Gamma_N) := \{F \in \mathcal{F}_\ell : F \subset \Gamma_N\}$ of \mathcal{T}_ℓ .

For any interior side $F \in \mathcal{F}_\ell(\Omega)$, there exist exactly two simplexes $T_+, T_- \in \mathcal{T}_\ell$ such that $\partial T_+ \cap \partial T_- = F$. The orientation of the outer normal unit $\nu_F = \nu_{T_+}|_F = -\nu_{T_-}|_F$ along F is fixed beforehand. Define the side patch $\omega_F := \text{int}(T_+ \cup T_-)$ of F . Let $[v]_F := (v|_{T_+})|_F - (v|_{T_-})|_F \in L^1(F)$ denote the jump of $v \in L^1(\omega_F)$ with $v \in W^{1,1}(T_+)$ and $v \in W^{1,1}(T_-)$ across F (with the abbreviations $W^{1,1}(T_+) := W^{1,1}(\text{int}(T_+))$ and $W^{1,1}(T_-) := W^{1,1}(\text{int}(T_-))$). For any boundary side $F \in \mathcal{F}_\ell(\partial\Omega) := \mathcal{F}_\ell \setminus \mathcal{F}_\ell(\Omega)$, there is a unique $T \in \mathcal{T}_\ell$ with $F \in \mathcal{F}_\ell(T)$. Then $\omega_F = \text{int}(T)$, $\nu_F := \nu_T$, and $[v]_F := (v|_T)|_F$. The differential operators div_{pw} and D_{pw} depend on the triangulation \mathcal{T}_ℓ and denote the piecewise application of div and D without explicit reference to \mathcal{T}_ℓ .

The shape regularity of a triangulation \mathcal{T} is the minimum $\min_{T \in \mathcal{T}} \varrho(T)$ of all ratios $\varrho(T) := r_i/r_c \leq 1$ of the maximal radius r_i of an inscribed ball and the minimal radius r_c of a circumscribed ball for a simplex $T \in \mathcal{T}$.

3.2 Discrete spaces

The discrete ansatz space of the HHO methods consists of piecewise polynomials on the triangulation \mathcal{T}_ℓ and on the skeleton $\partial\mathcal{T}_\ell := \cup \mathcal{F}_\ell$. For a simplex or a side $M \subset \mathbb{R}^n$ of diameter h_M , let $P_k(M)$ denote the space of polynomials of degree at most $k \in \mathbb{N}_0$ regarded as functions defined in M . The L^2 projection $\Pi_M^k v \in P_k(M)$ of $v \in L^1(M)$ is defined by $\Pi_M^k v \in P_k(M)$ with

$$\int_M \varphi_k(1 - \Pi_M^k)v \, dx = 0 \quad \text{for any } \varphi_k \in P_k(M).$$

The gradient reconstruction in $T \in \mathcal{T}_\ell$ maps in the space of Raviart–Thomas finite element functions

$$\text{RT}_k(T) := P_k(T; \mathbb{R}^n) + x P_k(T) \subset P_{k+1}(T; \mathbb{R}^n).$$

Let $P_k(\mathcal{T}_\ell)$, $P_k(\mathcal{F}_\ell)$, and $\text{RT}_k^{\text{pw}}(\mathcal{T}_\ell)$ denote the space of piecewise functions with respect to the mesh \mathcal{T}_ℓ or \mathcal{F}_ℓ and with restrictions to T or F in $P_k(T)$, $P_k(F)$, and $\text{RT}_k(T)$. The L^2 projections $\Pi_{\mathcal{T}_\ell}^k$ and $\Pi_{\mathcal{F}_\ell}^k$ onto the discrete spaces $P_k(\mathcal{T}_\ell)$ and $P_k(\mathcal{F}_\ell)$ are the global versions of Π_T^k and Π_F^k , e.g., $(\Pi_{\mathcal{T}_\ell}^k v)|_T := \Pi_T^k(v|_T)$ for $v \in L^1(\Omega)$. For vector-valued functions $v \in L^1(\Omega; \mathbb{R}^m)$, the L^2 projection $\Pi_{\mathcal{T}_\ell}^k$ onto $P_k(\mathcal{T}_\ell; \mathbb{R}^m) := P_k(\mathcal{T}_\ell)^m$ applies componentwise. This convention extends to the L^2 projections onto $P_k(M; \mathbb{R}^m) := P_k(M)^m$ and $P_k(\mathcal{F}_\ell; \mathbb{R}^m) := P_k(\mathcal{F}_\ell)^m$. The space of lowest-order Crouzeix–Raviart finite element functions reads

$$\text{CR}^1(\mathcal{T}_\ell) := \{v_{\text{CR}} \in P_1(\mathcal{T}_\ell) : v_{\text{CR}} \text{ is continuous at midpoints of } F \text{ for all } F \in \mathcal{F}_\ell(\Omega)\}. \tag{3.1}$$

Define the mesh-size function $h_\ell \in P_0(\mathcal{T}_\ell)$ with $h_\ell|_T \equiv |T|^{1/n}$ for all $T \in \mathcal{T}_\ell$, the (volume data) oscillation $\text{osc}(f, \mathcal{T}_\ell)^{p'} := \sum_{T \in \mathcal{T}_\ell} h_T \|(1 - \Pi_T^k) f\|_{L^{p'}(\Omega)}^{p'}$, and the (Neumann data) oscillation $\text{osc}_N(g, \mathcal{F}_\ell(\Gamma_N))^{p'} := \sum_{F \in \mathcal{F}_\ell(\Gamma_N)} h_F \|(1 - \Pi_F^k) g\|_{L^{p'}(F)}^{p'}$ with the diameter $h_F = \text{diam}(F)$ of $F \in \mathcal{F}_\ell$. (Notice that the shape regularity of \mathcal{T}_ℓ implies the equivalence $h_F \approx h_T \approx |T|^{1/n}$ for all $T \in \mathcal{T}_\ell, F \in \mathcal{F}_\ell(T)$.)

3.3 HHO ansatz space

For fixed $k \in \mathbb{N}_0$, let $V(\mathcal{T}_\ell) := P_k(\mathcal{T}_\ell; \mathbb{R}^m) \times P_k(\mathcal{F}_\ell; \mathbb{R}^m)$ denote the discrete ansatz space for V in HHO methods [35, 37]. The notation $v_\ell \in V(\mathcal{T}_\ell)$ means that $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) = ((v_T)_{T \in \mathcal{T}_\ell}, (v_F)_{F \in \mathcal{F}_\ell})$ for some $v_{\mathcal{T}_\ell} \in P_k(\mathcal{T}_\ell; \mathbb{R}^m)$ and $v_{\mathcal{F}_\ell} \in P_k(\mathcal{F}_\ell; \mathbb{R}^m)$ with the identification $v_T := v_{\mathcal{T}_\ell}|_T \in P_k(T; \mathbb{R}^m)$ and $v_F := v_{\mathcal{F}_\ell}|_F \in P_k(F; \mathbb{R}^m)$ for all $T \in \mathcal{T}_\ell, F \in \mathcal{F}_\ell$. The discrete space $V(\mathcal{T}_\ell)$ is endowed with the seminorm

$$\|v_\ell\|_\ell^p := \|D_{\text{pw}} v_{\mathcal{T}_\ell}\|_{L^p(\Omega)}^p + \sum_{T \in \mathcal{T}_\ell} \sum_{F \in \mathcal{F}_\ell(T)} h_F^{1-p} \|v_T - v_F\|_{L^p(F)}^p \tag{3.2}$$

for any $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) \in V(\mathcal{T}_\ell)$. The set $\mathcal{F}_\ell \setminus \mathcal{F}_\ell(\Gamma_D)$ of non-Dirichlet sides gives rise to the space $P_k(\mathcal{F}_\ell \setminus \mathcal{F}_\ell(\Gamma_D); \mathbb{R}^m)$ of piecewise polynomials $v_{\mathcal{F}_\ell} \in P_k(\mathcal{F}_\ell; \mathbb{R}^m)$ with the convention $v_{\mathcal{F}_\ell}|_F \equiv 0$ on $F \in \mathcal{F}_\ell(\Gamma_D)$ to model homogenous Dirichlet boundary conditions along the side $F \subset \Gamma_D$. The discrete linear space $V_D(\mathcal{T}_\ell) := P_k(\mathcal{T}_\ell; \mathbb{R}^m) \times P_k(\mathcal{F}_\ell \setminus \mathcal{F}_\ell(\Gamma_D); \mathbb{R}^m) \subset V(\mathcal{T}_\ell)$, equipped with the norm $\|\bullet\|_\ell$ from (3.2), is the discrete analogue to $V_D = W_D^{1,p}(\Omega; \mathbb{R}^m)$. The interpolation

$$I_\ell : V \rightarrow V(\mathcal{T}_\ell), v \mapsto (\Pi_{\mathcal{T}_\ell}^k v, \Pi_{\mathcal{F}_\ell}^k v) \tag{3.3}$$

gives rise to the discrete space $\mathcal{A}(\mathcal{T}_\ell) := I_\ell u_D + V_D(\mathcal{T}_\ell)$ of admissible functions.

3.4 Reconstruction operators

The reconstruction operators defined in this section link the two components of $v_\ell \in V(\mathcal{T}_\ell)$ and provide discrete approximations $\mathcal{R}_\ell v_\ell$ and $\mathcal{G}_\ell v_\ell$ of the displacement $v \in V$ and its derivative $Dv \in L^2(\Omega; \mathbb{M})$.

Potential reconstruction Given $T \in \mathcal{T}_\ell$ and $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) \in V(\mathcal{T}_\ell)$ with the convention $v_T = v_{\mathcal{T}_\ell}|_T$ and $v_F = v_{\mathcal{F}_\ell}|_F$ for all $F \in \mathcal{F}_\ell(T)$ from Sect. 3.3, the local potential reconstruction $\mathcal{R}_T v_\ell \in P_{k+1}(T; \mathbb{R}^m)$ satisfies

$$\begin{aligned} & \int_T D\mathcal{R}_T v_\ell : D\varphi_{k+1} \, dx \\ &= - \int_T \Delta\varphi_{k+1} \cdot v_T \, dx + \sum_{F \in \mathcal{F}(T)} \int_F v_F \cdot (D\varphi_{k+1} v_T)|_F \, ds \end{aligned} \tag{3.4}$$

for all $\varphi_{k+1} \in P_{k+1}(T; \mathbb{R}^m)$. The bilinear form $(D\bullet, D\bullet)_{L^2(T)}$ on the left-hand side of (3.4) defines a scalar product in the quotient space $P_{k+1}(T; \mathbb{R}^m)/\mathbb{R}^m$ and the right-hand side of (3.4) is a linear functional in $P_{k+1}(T; \mathbb{R}^m)/\mathbb{R}^m$. The Riesz representation $\mathcal{R}_T v_\ell \in P_{k+1}(T; \mathbb{R}^m)$ of this linear functional in $P_{k+1}(T; \mathbb{R}^m)/\mathbb{R}^m$ equipped with the energy scalar product is selected by

$$\int_T \mathcal{R}_T v_\ell \, dx = \int_T v_T \, dx. \tag{3.5}$$

The unique solution $\mathcal{R}_T v_\ell \in P_{k+1}(T; \mathbb{R}^m)$ to (3.4)–(3.5) gives rise to the potential reconstruction operator $\mathcal{R}_\ell : V(\mathcal{T}_\ell) \rightarrow P_{k+1}(\mathcal{T}_\ell; \mathbb{R}^m)$ with restriction $(\mathcal{R}_\ell v_\ell)|_T := \mathcal{R}_T v_\ell$ on each simplex $T \in \mathcal{T}_\ell$ for any $v_\ell \in V(\mathcal{T}_\ell)$.

Gradient reconstruction The gradient is reconstructed in the space $\Sigma(\mathcal{T}_\ell) = \text{RT}_k^{\text{pw}}(\mathcal{T}_\ell; \mathbb{M})$ of piecewise Raviart–Thomas finite element functions [1, 28]. Given $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) \in V(\mathcal{T}_\ell)$, its gradient reconstruction $\mathcal{G}_\ell v_\ell \in \Sigma(\mathcal{T}_\ell)$ solves

$$\int_\Omega \mathcal{G}_\ell v_\ell : \tau_\ell \, dx = - \int_\Omega v_{\mathcal{T}_\ell} \cdot \text{div}_{\text{pw}} \tau_\ell \, dx + \sum_{F \in \mathcal{F}_\ell} \int_F v_F \cdot [\tau_\ell v_F]_F \, ds \tag{3.6}$$

for all $\tau_\ell \in \Sigma(\mathcal{T}_\ell)$. In other words, $\mathcal{G}_\ell v_\ell$ is the Riesz representation of the linear functional on the right-hand side of (3.6) in the Hilbert space $\Sigma(\mathcal{T}_\ell)$ endowed with the L^2 scalar product. Since $D_{\text{pw}} P_{k+1}(\mathcal{T}_\ell; \mathbb{R}^m) \subset \Sigma(\mathcal{T}_\ell)$, it follows that $D_{\text{pw}} \mathcal{R}_\ell v_\ell$ is the L^2 projection of $\mathcal{G}_\ell v_\ell$ onto $D_{\text{pw}} P_{k+1}(\mathcal{T}_\ell; \mathbb{R}^m)$.

Lemma 3.1 (properties of \mathcal{G}) *Any $v \in V$ and $v_\ell \in V(\mathcal{T}_\ell)$ satisfy (a) $\|v_\ell\|_\ell \approx \|\mathcal{G}_\ell v_\ell\|_{L^p(\Omega)}$ and (b) $\Pi_{\Sigma(\mathcal{T}_\ell)} Dv = \mathcal{G}_\ell I_\ell v$. There exist positive constants C_{dF} and C_{dtr} that only depend on Ω , the shape regularity of \mathcal{T}_ℓ , k , and p such that (c) $\|v_{\mathcal{T}_\ell}\|_{L^p(\Omega)} \leq C_{\text{dF}} \|\mathcal{G}_\ell v_\ell\|_{L^p(\Omega)}$ and (d) $\|v_{\mathcal{F}_\ell}\|_{L^p(\Gamma_N)} \leq C_{\text{dtr}} \|\mathcal{G}_\ell v_\ell\|_{L^p(\Omega)}$ hold for all $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) \in V_D(\mathcal{T}_\ell)$.*

Proof The proofs of (a)–(b) are outlined in [1, 28]. The discrete Sobolev embedding $\|v_{\mathcal{T}_\ell}\|_{L^p(\Omega)} \lesssim \|v_\ell\|_\ell$ follows as in [11, 34, 36]. Theorem 4.4 in [11] and (c) lead to $\|v_{\mathcal{T}_\ell}\|_{L^p(\Gamma_N)} \lesssim \|v_\ell\|_\ell$. This and the triangle inequality $\|v_{\mathcal{F}_\ell}\|_{L^p(\Gamma_N)} \leq \|v_{\mathcal{T}_\ell}\|_{L^p(\Gamma_N)} + \|v_{\mathcal{T}_\ell} - v_{\mathcal{F}_\ell}\|_{L^p(\Gamma_N)}$ imply $\|v_{\mathcal{F}_\ell}\|_{L^p(\Gamma_N)} \lesssim \|v_\ell\|_\ell + \|v_{\mathcal{T}_\ell} - v_{\mathcal{F}_\ell}\|_{L^p(\Gamma_N)}$. The latter term is controlled by $\text{diam}(\Omega)^{1/p'} \|v_\ell\|_\ell$. This concludes the proof of (d). \square

3.5 Discrete problem

Lemma 3.1 implies the coercivity of E_ℓ in $\mathcal{A}(\mathcal{T}_\ell)$ with respect to the discrete seminorm $\|\mathcal{G}_\ell \bullet\|_{L^p(\Omega)}$ and the existence and the boundedness of discrete minimizers u_ℓ below.

Theorem 3.2 (Discrete minimizers) *The minimal discrete energy $\inf E_\ell(\mathcal{A}(\mathcal{T}_\ell))$ is attained. There exists a positive constant $C_1 > 0$ that depends only on $c_1, c_2, \Omega, \Gamma_D, u_D, f, g$, the shape regularity of \mathcal{T}_ℓ, k , and p with $\|\mathcal{G}_\ell u_\ell\|_{L^p(\Omega)} \leq C_1$ for all discrete minimizers $u_\ell \in \arg \min E_\ell(\mathcal{A}(\mathcal{T}_\ell))$. Any discrete stress $\sigma_\ell := \Pi_{\Sigma(\mathcal{T}_\ell)} \text{DW}(\mathcal{G}_\ell u_\ell) \in L^{p'}(\Omega; \mathbb{M})$ satisfies the discrete Euler–Lagrange equations*

$$\int_{\Omega} \sigma_\ell : \mathcal{G}_\ell v_\ell \, dx = \int_{\Omega} f \cdot v_{\mathcal{T}_\ell} \, dx + \int_{\Gamma_N} g \cdot v_{\mathcal{F}_\ell} \, ds \tag{3.7}$$

for all $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) \in V_D(\mathcal{T}_\ell)$. If W satisfies (2.5), then $u_\ell = \arg \min E_\ell(\mathcal{A}(\mathcal{T}_\ell))$ is unique. If W satisfies (2.6), then $\text{DW}(\mathcal{G}_\ell u_\ell) \in L^{p'}(\Omega; \mathbb{M})$ is unique (independent of the choice of a (possibly non-unique) discrete minimizer u_ℓ).

Proof The boundedness $\inf E_\ell(\mathcal{A}(\mathcal{T}_\ell)) > -\infty$ of E_ℓ in $\mathcal{A}(\mathcal{T}_\ell)$ follows from the lower p -growth of W , the discrete Friedrichs, and the discrete trace inequality from Lemma 3.1, cf., e.g., [27, 28, 32]. The direct method in the calculus of variations [32] implies the existence of discrete minimizers $u_\ell \in \arg \min E_\ell(\mathcal{A}(\mathcal{T}_\ell))$. The bound $\|\mathcal{G}_\ell u_\ell\|_{L^p(\Omega)} \leq C_1$ is a consequence of the coercivity of E_ℓ in $\mathcal{A}(\mathcal{T}_\ell)$ with respect to $\|\bullet\|_\ell$ as in [28]. If W satisfies (2.5), then W is strictly convex and the discrete minimizer $u_\ell \in \arg \min E_\ell(\mathcal{A}(\mathcal{T}_\ell))$ is unique. If W satisfies (2.6), then the uniqueness of $\text{DW}(\mathcal{G}_\ell u_\ell)$ follows as in [16, 27, 28]. \square

Remark 3.3 ($H(\text{div})$ conformity) The discrete Euler–Lagrange equations (3.7) imply the continuity of the normal jumps $[\sigma_\ell v_F]_F$ of $\sigma_\ell \in \Pi_{\Sigma(\mathcal{T}_\ell)} \text{DW}(\mathcal{G}_\ell u_\ell)$ along all interior side $F \in \mathcal{F}_\ell(\Omega)$ [28, Theorem 3.2]. In other words, $\sigma_\ell \in \Sigma(\mathcal{T}_\ell) \cap W^{p'}(\text{div}, \Omega; \mathbb{M})$ with $\text{div} \sigma_\ell = -\Pi_{\mathcal{T}_\ell}^k f$ and $\sigma_\ell v_F = \Pi_F^k g$ for all $F \in \mathcal{F}_\ell(\Gamma_N)$.

3.6 Conforming companion

The companion operator $\mathcal{J}_\ell : V(\mathcal{T}_\ell) \rightarrow V$ is a right-inverse of the interpolation $\mathcal{I}_\ell : V \rightarrow V(\mathcal{T}_\ell)$ in spirit of [18, 23, 26, 42]. In particular, \mathcal{J}_ℓ preserves the moments

$$\Pi_{\mathcal{T}_\ell}^k \mathcal{J}_\ell v_\ell = v_{\mathcal{T}_\ell} \quad \text{and} \quad \Pi_{\mathcal{F}_\ell}^k \mathcal{J}_\ell v_\ell = v_{\mathcal{F}_\ell} \quad \text{for any } v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) \in V(\mathcal{T}_\ell). \tag{3.8}$$

An explicit construction of $\mathcal{J}_\ell v_\ell$ on simplicial meshes is presented in [42, Section 4.3] for simplicial triangulations with the following properties.

Lemma 3.4 (right-inverse) *There exists a linear operator $\mathcal{J}_\ell : V(\mathcal{T}_\ell) \rightarrow V$ with (3.8) such that any $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) \in V(\mathcal{T}_\ell)$ satisfies, for all $T \in \mathcal{T}_\ell$,*

$$\begin{aligned} \|\mathcal{G}_\ell v_\ell - \text{D}\mathcal{J}_\ell v_\ell\|_{L^p(T)}^p &\lesssim \sum_{E \in \mathcal{F}_\ell(\Omega), E \cap T \neq \emptyset} h_E^{1-p} \|[\mathcal{R}_\ell v_\ell]_E\|_{L^p(E)}^p \\ &+ \sum_{F \in \mathcal{F}_\ell(T)} h_F^{1-p} \|\Pi_F^k((\mathcal{R}_\ell v_\ell)_T - v_F)\|_{L^p(F)}^p + h_T^{-p} \|\Pi_T^k(\mathcal{R}_\ell v_\ell - v_T)\|_{L^p(T)}^p. \end{aligned} \tag{3.9}$$

In particular, \mathcal{J}_ℓ is stable in the sense that $\|\mathbf{D}\mathcal{J}_\ell v_\ell\|_{L^p(\Omega)} \leq \Lambda_0 \|v_\ell\|_\ell$ holds with the constant Λ_0 that exclusively depends on k, p , and the shape regularity of \mathcal{T}_ℓ .

Proof For $p = 2$, the right-hand side of (3.9) is an upper bound for $\|\mathbf{D}(\mathcal{R}_\ell v_\ell - \mathcal{J}_\ell v_\ell)\|_{L^2(T)}$, cf. [42, Proof of Proposition 4.7], and scaling arguments confirm this for $1 < p < \infty$. The L^p stability of the L^2 projection [34, Lemma 3.2] and the orthogonality $\mathcal{G}_\ell v_\ell - \mathbf{D}\mathcal{J}_\ell v_\ell \perp \text{RT}_k(T; \mathbb{M})$ in $L^2(T; \mathbb{M})$ imply $\|\mathcal{G}_\ell v_\ell - \mathbf{D}\mathcal{J}_\ell v_\ell\|_{L^p(T)} \lesssim \|\mathbf{D}(\mathcal{R}_\ell v_\ell - \mathcal{J}_\ell v_\ell)\|_{L^p(T)}$. This proves (3.9). The right-hand side of (3.9) can be bounded by

$$\begin{aligned} & \sum_{E \in \mathcal{F}_\ell(\Omega), E \cap T \neq \emptyset} h_E^{1-p} \|[\mathcal{R}_\ell v_\ell]_E\|_{L^p(E)}^p + \sum_{F \in \mathcal{F}_\ell(T)} h_F^{1-p} \|\Pi_F^k((\mathcal{R}_\ell v_\ell)_T - v_F)\|_{L^p(F)}^p \\ & + h_T^{-p} \|\Pi_T^k(\mathcal{R}_\ell v_\ell - v_T)\|_{L^p(T)}^p \lesssim \sum_{K \in \mathcal{T}_\ell, K \cap T \neq \emptyset} \sum_{E \in \mathcal{F}_\ell(K)} h_E^{1-p} \|v_K - v_F\|_{L^p(E)}^p \end{aligned} \tag{3.10}$$

with a hidden constant that only depends on the shape regularity of \mathcal{T}_ℓ, k , and p [42, Proof of Proposition 4.7]. The sum of this over all simplices $T \in \mathcal{T}_\ell$, a triangle inequality, and the shape regularity of \mathcal{T}_ℓ imply $\|\mathcal{G}_\ell v_\ell - \mathbf{D}\mathcal{J}_\ell v_\ell\|_{L^p(\Omega)} \lesssim \|v_\ell\|_\ell$. This, a reverse triangle inequality, and the norm equivalence from Lemma 3.1a conclude the stability $\|\mathbf{D}\mathcal{J}_\ell v_\ell\|_{L^p(\Omega)} \lesssim \|v_\ell\|_\ell$. \square

4 Proof of Theorem 2.1

This section is devoted to the proof of the convergence results in Theorem 2.1.

4.1 Discrete compactness

The proof of Theorem 2.1 departs from a discrete compactness in spirit of [11, 34, 36] and generalizes [34]. Recall the mesh-size function $h_\ell \in P_0(\mathcal{T}_\ell)$ from Sect. 3.2 with $h_\ell|_T = |T|^{1/n}$ for $T \in \mathcal{T}_\ell$ and the seminorm $\|\bullet\|_\ell$ in $V(\mathcal{T}_\ell)$ from (3.2).

Theorem 4.1 (discrete compactness) *Given a uniformly shape-regular sequence $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0}$ of triangulations and $(v_\ell)_{\ell \in \mathbb{N}_0}$ with $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) \in \mathcal{A}(\mathcal{T}_\ell)$ for all $\ell \in \mathbb{N}_0$. Suppose that the sequence $(\|v_\ell\|_\ell)_{\ell \in \mathbb{N}_0}$ is bounded and suppose that $\lim_{\ell \rightarrow \infty} \mu_\ell(v_\ell) = 0$ with*

$$\mu_\ell(v_\ell) := \|h_\ell^{k+1}(\mathcal{G}_\ell v_\ell - \mathbf{D}\mathcal{J}_\ell v_\ell)\|_{L^p(\Omega)}^p + \sum_{F \in \mathcal{F}_\ell(\Gamma_D)} h_F^{kp+1} \|\mathcal{J}_\ell v_\ell - u_D\|_{L^p(F)}^p. \tag{4.1}$$

Then there exist a subsequence $(v_{\ell_j})_{j \in \mathbb{N}_0}$ of $(v_\ell)_{\ell \in \mathbb{N}_0}$ and a weak limit $v \in \mathcal{A}$ such that $\mathcal{J}_{\ell_j} v_{\ell_j} \rightharpoonup v$ weakly in V and $\mathcal{G}_{\ell_j} v_{\ell_j} \rightharpoonup \mathbf{D}v$ weakly in $L^p(\Omega; \mathbb{M})$ as $j \rightarrow \infty$.

Proof The first part of the proof proves the uniform boundedness

$$\|\mathcal{J}_\ell v_\ell\|_{W^{1,p}(\Omega)} \lesssim \|v_\ell\|_\ell + \|u_D\|_{W^{1,p}(\Omega)} \lesssim 1 \tag{4.2}$$

of the sequence $\mathcal{J}_\ell v_\ell$ in $W^{1,p}(\Omega; \mathbb{R}^m)$. Since $\|D\mathcal{J}_\ell v_\ell\|_{L^p(\Omega)} \lesssim \|v_\ell\|_\ell$ from the stability of \mathcal{J}_ℓ in Lemma 3.4, it remains to show $\|\mathcal{J}_\ell v_\ell\|_{L^p(\Omega)} \lesssim \|v_\ell\|_\ell + \|u_D\|_{W^{1,p}(\Omega)}$ to obtain (4.2). The triangle inequality implies

$$\|\mathcal{J}_\ell v_\ell\|_{L^p(\Omega)} \leq \|\mathcal{J}_\ell v_\ell - v_{\mathcal{T}_\ell}\|_{L^p(\Omega)} + \|v_{\mathcal{T}_\ell} - \Pi_{\mathcal{T}_\ell}^k u_D\|_{L^p(\Omega)} + \|\Pi_{\mathcal{T}_\ell}^k u_D\|_{L^p(\Omega)}. \tag{4.3}$$

The right-inverse \mathcal{J}_ℓ of the interpolation I_ℓ from Lemma 3.4 satisfies the L^2 orthogonality $\mathcal{J}_\ell v_\ell - v_{\mathcal{T}_\ell} \perp P_k(\mathcal{T}_\ell; \mathbb{R}^m)$. This, a piecewise application of the Poincaré inequality, a triangle inequality, and $\|h_\ell\|_{L^\infty(\Omega)} \leq \text{diam}(\Omega)$ lead to

$$\begin{aligned} \|\mathcal{J}_\ell v_\ell - v_{\mathcal{T}_\ell}\|_{L^p(\Omega)} &\lesssim \|h_\ell D_{pw}(\mathcal{J}_\ell v_\ell - v_{\mathcal{T}_\ell})\|_{L^p(\Omega)} \\ &\lesssim \|D\mathcal{J}_\ell v_\ell\|_{L^p(\Omega)} + \|D_{pw} v_{\mathcal{T}_\ell}\|_{L^p(\Omega)}. \end{aligned} \tag{4.4}$$

Since $v_\ell - I_\ell u_D \in V_D(\mathcal{T}_\ell)$, the Sobolev embedding from Lemma 3.1c and a triangle inequality show $\|v_{\mathcal{T}_\ell} - \Pi_{\mathcal{T}_\ell}^k u_D\|_{L^p(\Omega)} \lesssim \|\mathcal{G}_\ell(v_\ell - I_\ell u_D)\|_{L^p(\Omega)} \leq \|\mathcal{G}_\ell v_\ell\|_{L^p(\Omega)} + \|\mathcal{G}_\ell I_\ell u_D\|_{L^p(\Omega)}$. This, the equivalence $\|\mathcal{G}_\ell v_\ell\|_{L^p(\Omega)} \approx \|v_\ell\|_\ell$ from Lemma 3.1a, the commutativity $\mathcal{G}_\ell I_\ell u_D = \Pi_{\Sigma(\mathcal{T}_\ell)} D u_D$ from Lemma 3.1b, and the L^p stability of the L^2 projection $\Pi_{\Sigma(\mathcal{T}_\ell)}$ [34, Lemma 3.2] provide

$$\|v_{\mathcal{T}_\ell} - \Pi_{\mathcal{T}_\ell}^k u_D\|_{L^p(\Omega)} \lesssim \|v_\ell\|_\ell + \|D u_D\|_{L^p(\Omega)}. \tag{4.5}$$

Lemma 3.4 and the definition of the discrete norm $\|v_\ell\|_\ell$ in (3.2) prove that the right-hand side of (4.4) is controlled by $\|v_\ell\|_\ell$. Hence, the combination of (4.3)–(4.5) concludes (4.2).

The Banach–Alaoglu theorem [10, Theorem 3.18] ensures the existence of a (not relabelled) subsequence of $(\mathcal{J}_\ell v_\ell)_{\ell \in \mathbb{N}_0}$ and a weak limit $v \in V$ such that $\mathcal{J}_\ell v_\ell \rightharpoonup v$ weakly in V as $\ell \rightarrow \infty$. Lemma 3.1a assures that the sequence $(\mathcal{G}_\ell v_\ell)_{\ell \in \mathbb{N}_0}$ is uniformly bounded in $L^p(\Omega; \mathbb{M})$. Hence there exist a (not relabelled) subsequence of $(v_\ell)_{\ell \in \mathbb{N}_0}$ and its weak limit $G \in L^p(\Omega; \mathbb{M})$ such that $\mathcal{G}_\ell v_\ell \rightharpoonup G$ weakly in $L^p(\Omega; \mathbb{M})$ as $\ell \rightarrow \infty$. The second part of the proof verifies $Dv = G$ in Ω and $v = u_D$ on Γ_D (and so $v \in \mathcal{A}$). Since $I_\ell \mathcal{J}_\ell v_\ell = v_\ell$, the commutativity $\mathcal{G}_\ell v_\ell = \Pi_{\Sigma(\mathcal{T}_\ell)} D \mathcal{J}_\ell v_\ell$ from Lemma 3.1b proves the L^2 orthogonality $\mathcal{G}_\ell v_\ell - D \mathcal{J}_\ell v_\ell \perp \Sigma(\mathcal{T}_\ell)$. This and an integration by parts verify, for all $\Phi \in C^\infty(\overline{\Omega}; \mathbb{M})$ with $\Phi \equiv 0$ on Γ_N , that

$$\begin{aligned} \int_\Omega \mathcal{G}_\ell v_\ell : \Phi \, dx &= \int_\Omega (\mathcal{G}_\ell v_\ell - D \mathcal{J}_\ell v_\ell) : \Phi \, dx + \int_\Omega D \mathcal{J}_\ell v_\ell : \Phi \, dx \\ &= \int_\Omega (\mathcal{G}_\ell v_\ell - D \mathcal{J}_\ell v_\ell) : (1 - \Pi_{\Sigma(\mathcal{T}_\ell)}) \Phi \, dx - \int_\Omega \mathcal{J}_\ell v_\ell \cdot \text{div} \Phi \, dx \\ &\quad + \int_{\Gamma_D} (\mathcal{J}_\ell v_\ell - u_D) \cdot \Phi \nu \, ds + \int_{\Gamma_D} u_D \cdot \Phi \nu \, ds. \end{aligned} \tag{4.6}$$

The approximation property of piecewise polynomials, also known under the name Bramble–Hilbert lemma [9, Lemma 4.3.8], leads to

$$\|h_\ell^{-(k+1)}(1 - \Pi_{\Sigma(\mathcal{T}_\ell)})\Phi\|_{L^{p'}(\Omega)} \lesssim |\Phi|_{W^{k+1,p'}(\Omega)}. \tag{4.7}$$

This and a Hölder inequality imply

$$\begin{aligned} & \int_{\Omega} (\mathcal{G}_\ell v_\ell - D\mathcal{J}_\ell v_\ell) : (1 - \Pi_{\Sigma(\mathcal{T}_\ell)})\Phi \, dx \\ & \lesssim \|h_\ell^{k+1}(\mathcal{G}_\ell v_\ell - D\mathcal{J}_\ell v_\ell)\|_{L^p(\Omega)} |\Phi|_{W^{k+1,p'}(\Omega)}. \end{aligned} \tag{4.8}$$

The L^2 orthogonality $(\mathcal{J}_\ell u_\ell - u_D)|_F \perp P_k(F; \mathbb{R}^m)$ for each side $F \in \mathcal{F}_\ell(\Gamma_D)$ on the Dirichlet boundary, a piecewise application of the trace inequality, and (4.7) imply

$$\int_{\Gamma_D} (\mathcal{J}_\ell v_\ell - u_D) \cdot \Phi \nu \, ds \lesssim \left(\sum_{F \in \mathcal{F}_\ell(\Gamma_D)} h_F^{k p + 1} \|\mathcal{J}_\ell v_\ell - u_D\|_{L^p(F)}^p \right)^{1/p} |\Phi|_{W^{k+1,p'}(\Omega)}. \tag{4.9}$$

The right-hand sides of (4.8)–(4.9) vanish in the limit as $\ell \rightarrow \infty$ by assumption (4.1). This, (4.6), $\mathcal{G}_\ell v_\ell \rightarrow G$ in $L^p(\Omega; \mathbb{M})$, and $\mathcal{J}_\ell v_\ell \rightarrow v$ in V prove

$$\int_{\Omega} (G : \Phi + v \cdot \operatorname{div}\Phi) \, dx - \int_{\Gamma_D} u_D \cdot \Phi \nu \, ds = 0$$

for all $\Phi \in C^\infty(\bar{\Omega}; \mathbb{M})$ with $\Phi \equiv 0$ on Γ_N . This implies $Dv = G$ a.e. in Ω with $v = u_D$ on Γ_D and concludes the proof. □

Since $\mathcal{J}_\ell v_\ell$ cannot attain the exact value u_D on Γ_D in general, a (Dirichlet boundary data) oscillation arises in (4.1), but is controlled by the contributions of $\eta_\ell^{(\varepsilon)}$.

Lemma 4.2 (Dirichlet boundary data oscillation) *Given $F \in \mathcal{F}_\ell(\Gamma_D)$, let $T \in \mathcal{T}_\ell$ be the unique simplex with $F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Gamma_D)$. Then it holds, for all $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) \in V(\mathcal{T}_\ell)$, that*

$$\begin{aligned} \|\mathcal{J}_\ell v_\ell - u_D\|_{L^p(F)}^p & \lesssim \sum_{E \in \mathcal{F}_\ell(\Omega), E \cap F \neq \emptyset} \|[\mathcal{R}_\ell v_\ell]_E\|_{L^p(E)}^p \\ & + \|\Pi_F^k(\mathcal{R}_\ell v_\ell - v_F)\|_{L^p(F)}^p + \|\mathcal{R}_\ell v_\ell - u_D\|_{L^p(F)}^p. \end{aligned}$$

Proof The proof of Lemma 4.2 utilizes standard averaging and bubble-function techniques, cf., e.g, [20, 42, 57, 58]; further details are therefore omitted. □

4.2 Plain convergence

Before the remaining parts of this subsection prove Theorem 2.1, the following lemma establishes the reduction of the mesh-size function h_ℓ with $h_\ell|_T \equiv |T|^{1/n}$ for all $T \in \mathcal{T}_\ell$.

Lemma 4.3 (Mesh-size reduction) *Given the output $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0}$ of AHHO from Sect. 2.2, let $\Omega_\ell := \text{int}(\cup(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}))$ for all level $\ell \in \mathbb{N}_0$. Then it holds*

$$\lim_{\ell \rightarrow \infty} \|h_\ell\|_{L^\infty(\Omega_\ell)} = 0. \tag{4.10}$$

Proof The proof is omitted for two reasons. First this is known from [50, Lemma 9] and second it is a particular case of Lemma 5.2 below. \square

Proof of Theorem 2.1 This proof is motivated by [4, 13, 22, 49, 50, 52] and is divided into five steps.

Step 1 establishes $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}) = 0$. Since no suitable residual-based a posteriori control is available in the general setting (A1)–(A2), standard arguments, e.g., reliability, efficiency, or estimator reduction [22, 41, 49] fail. The proof of Step 1 rather relies on a positive power of the mesh-size that arises from the smoothness of test functions in Theorem 4.1. This is done in [52] for a similar setting and in [11, 34, 36] for uniform mesh-refinements. Let $\mu_\ell^{(\varepsilon)}(T)$ abbreviate some contributions of $\eta_\ell^{(\varepsilon)}(T)$ from (2.3) related to $\mu_\ell(u_\ell)$ in (4.1), namely

$$\begin{aligned} \mu_\ell^{(\varepsilon)}(T) &:= |T|^{(\varepsilon p - p)/n} \|\Pi_T^k(\mathcal{R}_\ell u_\ell - u_T)\|_{L^p(T)}^p \\ &+ |T|^{(\varepsilon p + 1 - p)/n} \left(\sum_{F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Gamma_D)} \|\mathcal{R}_\ell u_\ell - u_D\|_{L^p(F)}^p \right. \\ &+ \left. \sum_{F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Omega)} \|[\mathcal{R}_\ell u_\ell]_F\|_{L^p(F)}^p + \sum_{F \in \mathcal{F}_\ell(T)} \|\Pi_F^k((\mathcal{R}_\ell u_\ell)|_T - u_F)\|_{L^p(F)}^p \right). \end{aligned} \tag{4.11}$$

Denote $\mu_\ell^{(\varepsilon)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}) := \sum_{T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}} \mu_\ell^{(\varepsilon)}(T)$. Given any $T \in \mathcal{T}_\ell$ and $F \in \mathcal{F}_\ell(\Gamma_N)$, the $L^{p'}$ stability of the L^2 projection $\Pi_{\text{RT}^k(T; \mathbb{M})}$ resp. Π_T^k or Π_F^k [34, Lemma 3.2] implies $\|\sigma_\ell - \text{DW}(\mathcal{G}_\ell u_\ell)\|_{L^{p'}(T)} \lesssim \|\text{DW}(\mathcal{G}_\ell u_\ell)\|_{L^{p'}(T)}$ resp. $\|(1 - \Pi_T^k)f\|_{L^{p'}(T)} \lesssim \|f\|_{L^{p'}(T)}$ or $\|(1 - \Pi_F^k)g\|_{L^{p'}(F)} \lesssim \|g\|_{L^{p'}(F)}$. Since $\mu_\ell^{(\varepsilon)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}) \leq \|h_\ell\|_{L^\infty(\Omega_\ell)}^{\varepsilon p} \mu_\ell^{(0)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1})$ with $\|h_\ell\|_{L^\infty(\Omega_\ell)} = \sup_{T \in \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}} |T|^{1/n}$ and $\Omega_\ell = \text{int}(\cup(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}))$ from Lemma 4.3, this leads to

$$\begin{aligned} \eta_\ell^{(\varepsilon)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}) &\lesssim \|h_\ell\|_{L^\infty(\Omega_\ell)}^{\varepsilon p} \mu_\ell^{(0)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}) + \|h_\ell\|_{L^\infty(\Omega_\ell)}^{\varepsilon p'} \|\text{DW}(\mathcal{G}_\ell u_\ell)\|_{L^{p'}(\Omega)}^{p'} \\ &+ \|h_\ell\|_{L^\infty(\Omega_\ell)}^{p'} \|f\|_{L^{p'}(\Omega)}^{p'} + \|h_\ell\|_{L^\infty(\Omega_\ell)} \|g\|_{L^{p'}(\Gamma_N)}^{p'}. \end{aligned} \tag{4.12}$$

The two-sided growth $|A|^p - 1 \lesssim W(A) \lesssim |A|^p + 1$ implies $|\text{DW}(A)|^{p'} \lesssim |A|^p + 1$ [28, Lemma 2.1] and so Theorem 3.2 provides $\|\text{DW}(\mathcal{G}_\ell u_\ell)\|_{L^{p'}(\Omega)}^{p'} \lesssim \|\mathcal{G}_\ell u_\ell\|_{L^p(\Omega)}^p +$

$|\Omega| \lesssim 1$. This and (4.10) prove

$$\begin{aligned} \lim_{\ell \rightarrow \infty} & \left(\|h_\ell\|_{L^\infty(\Omega_\ell)}^{\varepsilon p'} \|DW(\mathcal{G}u_\ell)\|_{L^{p'}(\Omega)}^{p'} \right. \\ & \left. + \|h_\ell\|_{L^\infty(\Omega_\ell)}^{p'} \|f\|_{L^{p'}(\Omega)}^{p'} + \|h_\ell\|_{L^\infty(\Omega_\ell)} \|g\|_{L^{p'}(\Gamma_N)}^{p'} \right) = 0, \end{aligned} \tag{4.13}$$

whence, in order to obtain $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}) = 0$, it suffices to prove that $\mu_\ell^{(0)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1})$ is uniformly bounded. The estimate (3.10) provides control over all but only one contribution of $\mu_\ell^{(0)}(T)$ in (4.11); that is $\|h_F^{-1/p'}(\mathcal{R}_\ell u_\ell - u_D)\|_{L^p(F)}$ for any $F \in \mathcal{F}_\ell(\Gamma_D)$. Triangle inequalities and $u_F = \Pi_F^k u_D$ imply

$$\begin{aligned} \|h_F^{-1/p'}(\mathcal{R}_\ell u_\ell - u_D)\|_{L^p(F)} & \leq \|h_F^{-1/p'}(1 - \Pi_F^k)\mathcal{R}_\ell u_\ell\|_{L^p(F)} \\ & + \|h_F^{-1/p'}\Pi_F^k(\mathcal{R}_\ell u_\ell - u_F)\|_{L^p(F)} + \|h_F^{-1/p'}(1 - \Pi_F^k)u_D\|_{L^p(F)}. \end{aligned} \tag{4.14}$$

Given any $F \in \mathcal{F}_\ell(\Gamma_D)$, let $T \in \mathcal{T}_\ell$ be the unique simplex with $F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Gamma_D)$. The L^p stability of the L^2 projection Π_F^k [34, Lemma 3.2] and a trace inequality show $\|h_F^{-1/p'}(1 - \Pi_F^k)u_D\|_{L^p(F)} \lesssim \|Du_D\|_{L^p(T)}$ and $\|h_F^{-1/p'}(1 - \Pi_F^k)(\mathcal{R}_\ell u_\ell)|_F\|_{L^p(F)} \lesssim \|D\mathcal{R}_\ell u_\ell\|_{L^p(T)}$. Recall that $D\mathcal{R}_\ell u_\ell$ is the L^2 projection of $\mathcal{G}_\ell u_\ell$ onto $D_{pw}P_{k+1}(T)$, whence the L^p stability of L^2 projections [34, Lemma 3.2] proves $\|D\mathcal{R}_\ell u_\ell\|_{L^p(T)} \lesssim \|\mathcal{G}_\ell u_\ell\|_{L^p(T)}$. Hence, the right-hand side of (4.14) is controlled by $\|\mathcal{G}_\ell u_\ell\|_{L^p(T)} + \|Du_D\|_{L^p(T)} + \|h_F^{-1/p'}\Pi_F^k(\mathcal{R}_\ell u_\ell - u_F)\|_{L^p(F)}$. This, (3.10), and $\|u_\ell\|_\ell \approx \|\mathcal{G}_\ell u_\ell\|_{L^p(\Omega)} \leq C_1$ from Lemma 3.1a and Theorem 3.2 lead to

$$\mu_\ell^{(0)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}) \leq \mu_\ell^{(0)} := \sum_{T \in \mathcal{T}_\ell} \mu_\ell^{(0)}(T) \lesssim \|u_\ell\|_\ell^p + \|Du_D\|_{L^p(\Omega)}^p \lesssim 1.$$

Hence, the combination of (4.12)–(4.13) with $\lim_{\ell \rightarrow \infty} \|h_\ell\|_{L^\infty(\Omega_\ell)} = 0$ in (4.10) confirms $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}) = 0$.

Step 2 establishes $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)} = 0$. Recall the set \mathfrak{M}_ℓ of marked simplices on level $\ell \in \mathbb{N}_0$ from Sect. 2.2. Since all simplices in $\mathfrak{M}_\ell \subset \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$ are refined and the Dörfler marking enforces $\eta_\ell^{(\varepsilon)} \leq \theta^{-1} \eta_\ell^{(\varepsilon)}(\mathfrak{M}_\ell) \leq \theta^{-1} \eta_\ell^{(\varepsilon)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1})$ in (2.4), the convergence $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)}(\mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}) = 0$ in Step 1 implies $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)} = 0$.

Step 3 provides the lower energy bound (LEB)

$$\begin{aligned} \text{LEB}_\ell := E_\ell(u_\ell) & + \int_\Omega (1 - \Pi_{\Sigma(\mathcal{T}_\ell)})DW(\mathcal{G}_\ell u_\ell) : Du \, dx \\ & - C_3(\text{osc}(f, \mathcal{T}_\ell) + \text{osc}_N(g, \mathcal{F}_\ell(\Gamma_N))) \leq E(u). \end{aligned} \tag{4.15}$$

The convexity of $W \in C^1(\mathbb{M})$ implies $DW(\mathcal{G}_\ell u_\ell) : (Du - \mathcal{G}_\ell u_\ell) \leq W(Du) - W(\mathcal{G}_\ell u_\ell)$ a.e. in Ω . The integral of this inequality with $\sigma_\ell := \Pi_{\Sigma(\mathcal{T}_\ell)} DW(\mathcal{G}_\ell u_\ell)$ reads

$$0 \leq \int_{\Omega} (W(Du) - W(\mathcal{G}_\ell u_\ell) - (1 - \Pi_{\Sigma(\mathcal{T}_\ell)})DW(\mathcal{G}_\ell u_\ell) : Du) \, dx - \int_{\Omega} \sigma_\ell : (Du - \mathcal{G}_\ell u_\ell) \, dx. \tag{4.16}$$

The commutativity $\Pi_{\Sigma(\mathcal{T}_\ell)} Du = \mathcal{G}_\ell \mathbf{I}_\ell u$ from Lemma 3.1b and the discrete Euler-Lagrange equations (3.7) lead to

$$\int_{\Omega} \sigma_\ell : (Du - \mathcal{G}_\ell u_\ell) \, dx = \int_{\Omega} f \cdot (\Pi_{\mathcal{T}_\ell}^k u - u_{\mathcal{T}_\ell}) \, dx + \int_{\Gamma_N} g \cdot (\Pi_{\mathcal{F}_\ell}^k u - u_{\mathcal{F}_\ell}) \, dx.$$

The substitution of this in (4.16), the definition of E in (2.1), and the definition of E_h in (2.2) result in

$$0 \leq E(u) - E_\ell(u_\ell) - \int_{\Omega} (1 - \Pi_{\Sigma(\mathcal{T}_\ell)})DW(\mathcal{G}_\ell u_\ell) : Du \, dx + \int_{\Omega} (u - \Pi_{\mathcal{T}_\ell}^k u) \cdot (f - \Pi_{\mathcal{T}_\ell}^k f) \, ds + \int_{\Gamma_N} (u - \Pi_{\mathcal{F}_\ell}^k u) \cdot (g - \Pi_{\mathcal{F}_\ell}^k g) \, ds. \tag{4.17}$$

The final two integrals on the right-hand side of (4.17) give rise to the data oscillations $\text{osc}(f, \mathcal{T}_\ell)$ and $\text{osc}_N(g, \mathcal{F}_\ell(\Gamma_N))$ defined in Sect. 3.2. In fact, a Hölder inequality and a piecewise application of the Poincaré inequality show

$$\int_{\Omega} (u - \Pi_{\mathcal{T}_\ell}^k u) \cdot (f - \Pi_{\mathcal{T}_\ell}^k f) \, dx \lesssim \|Du\|_{L^p(\Omega)} \text{osc}(f, \mathcal{T}_\ell).$$

A trace inequality and the Bramble–Hilbert lemma [9, Lemma 4.3.8] lead, for all $F \in \mathcal{F}_\ell(\Gamma_N)$ and the unique $T \in \mathcal{T}_\ell$ with $F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Gamma_N)$, to $\|h_F^{-1/p'}(u - \Pi_F^k u)\|_{L^p(F)} \lesssim \|Du\|_{L^p(T)}$. Consequently,

$$\int_{\Gamma_N} (u - \Pi_{\mathcal{F}_\ell}^k u) \cdot (g - \Pi_{\mathcal{F}_\ell}^k g) \, ds \lesssim \|Du\|_{L^p(\Omega)} \text{osc}_N(g, \mathcal{F}_\ell(\Gamma_N)).$$

The lower p -growth $c_1|A|^p - c_2 \leq W(A)$ for all $A \in \mathbb{M}$ implies the coercivity of E in the seminorm $\|D \bullet\|_{L^p(\Omega)}$ and so the bound $\|Du\|_{L^p(\Omega)} \leq C_2$ with a positive constant C_2 , that exclusively depends on $c_1, c_2, \Omega, \Gamma_D, f, g$, and u_D , cf., e.g. [32, Theorem 4.1]. Thus there exists a positive constant C_3 independent of the mesh-size with

$$\int_{\Omega} (u - \Pi_{\mathcal{T}_\ell}^k u) \cdot (f - \Pi_{\mathcal{T}_\ell}^k f) \, ds + \int_{\Gamma_N} (u - \Pi_{\mathcal{F}_\ell}^k u) \cdot (g - \Pi_{\mathcal{F}_\ell}^k g) \, ds \leq C_3 (\text{osc}(f, \mathcal{T}_\ell) + \text{osc}_N(g, \mathcal{F}_\ell(\Gamma_N))). \tag{4.18}$$

The combination of this with (4.17) concludes the proof of (4.15).

Step 4 establishes $\lim_{\ell \rightarrow \infty} E_\ell(u_\ell) = E(u)$. On the one hand, the discrete compactness from Theorem 4.1 and the weak lower semicontinuity of the energy functional imply $E(u) \leq \liminf_{\ell \rightarrow \infty} \text{LEB}_\ell$. On the other hand, $\text{LEB}_\ell \leq E(u)$ from (4.15). This proves $\lim_{\ell \rightarrow \infty} E_\ell(u_\ell) = E(u)$ as follows. Given any $\Phi \in C^\infty(\Omega; \mathbb{M})$, the definition $\sigma_\ell := \Pi_{\Sigma(\mathcal{T}_\ell)} DW(\mathcal{G}_\ell u_\ell)$, a Hölder inequality, and (4.7) show

$$\left| \int_\Omega (\sigma_\ell - DW(\mathcal{G}_\ell u_\ell)) : \Phi \, dx \right| = \left| \int_\Omega (\sigma_\ell - DW(\mathcal{G}_\ell u_\ell)) : (1 - \Pi_{\Sigma(\mathcal{T}_\ell)}) \Phi \, dx \right| \lesssim \|h_\ell^{k+1} (\sigma_\ell - DW(\mathcal{G}_\ell u_\ell))\|_{L^{p'}(\Omega)} |\Phi|_{W^{k+1,p}(\Omega)}. \tag{4.19}$$

Since $\|h_\ell^{k+1} (\sigma_\ell - DW(\mathcal{G}_\ell u_\ell))\|_{L^{p'}(\Omega)}^{p'} \leq \eta_\ell^{(k+1)} \lesssim \eta_\ell^{(\varepsilon)} \rightarrow 0$ as $\ell \rightarrow \infty$ from Step 2, the right-hand side of (4.19) vanishes in the limit as $\ell \rightarrow \infty$. This, the density of $C^\infty(\Omega; \mathbb{M})$ in $L^p(\Omega; \mathbb{M})$, and the uniform boundedness of the sequence $(\sigma_\ell - DW(\mathcal{G}_\ell u_\ell))_{\ell \in \mathbb{N}_0}$ in $L^{p'}(\Omega; \mathbb{M})$ prove $\sigma_\ell - DW(\mathcal{G}_\ell u_\ell) \rightarrow 0$ (weakly) in $L^{p'}(\Omega; \mathbb{M})$ as $\ell \rightarrow \infty$. In particular,

$$\lim_{\ell \rightarrow \infty} \int_\Omega (\sigma_\ell - DW(\mathcal{G}_\ell u_\ell)) : Du \, dx = 0. \tag{4.20}$$

Recall $\mu_\ell(u_\ell)$ from (4.1) and $\mu_\ell^{(\varepsilon)}$ from (4.11). The combination of (3.9) with the bound of the Dirichlet data oscillation from Lemma 4.2 and the equivalence $h_F \approx h_T \approx |T|^{1/n}$ for all $T \in \mathcal{T}_\ell, F \in \mathcal{F}_\ell(T)$ from the shape regularity of \mathcal{T}_ℓ result in

$$\mu_\ell(u_\ell) \lesssim \mu_\ell^{(k+1)} \leq \text{diam}(\Omega)^{(k+1-\varepsilon)p} \mu_\ell^{(\varepsilon)} \lesssim \eta_\ell^{(\varepsilon)}.$$

This and $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)} = 0$ from Step 2 imply $\lim_{\ell \rightarrow \infty} \mu_\ell(u_\ell) = 0$. Since $\|u_\ell\|_\ell \approx \|\mathcal{G}_\ell u_\ell\|_{L^p(\Omega)} \leq C_1$ from Lemma 3.1a and Theorem 3.2, the discrete compactness from Theorem 4.1 leads to a (not relabelled) subsequence of $(u_\ell)_{\ell \in \mathbb{N}_0}$ and a weak limit $v \in \mathcal{A}$ such that $\mathcal{J}_\ell u_\ell \rightarrow v$ weakly in V and $\mathcal{G}_\ell u_\ell \rightarrow Dv$ weakly in $L^p(\Omega; \mathbb{M})$ as $\ell \rightarrow \infty$. The boundedness of the linear trace operator $\gamma : V \rightarrow L^p(\partial\Omega; \mathbb{R}^m)$ [10, Chapter 9] implies $(\mathcal{J}_\ell u_\ell)|_{\partial\Omega} \rightarrow v|_{\partial\Omega}$ (weakly) in $L^p(\partial\Omega; \mathbb{R}^m)$. Hence

$$\lim_{\ell \rightarrow \infty} \int_{\Gamma_N} g \cdot \mathcal{J}_\ell u_\ell \, ds = \int_{\Gamma_N} g \cdot v \, ds.$$

This, $\mathcal{J}_\ell u_\ell \rightarrow v$ (weakly) in $V, \mathcal{G}_\ell u_\ell \rightarrow Dv$ (weakly) in $L^p(\Omega; \mathbb{M})$, the sequential weak lower semicontinuity of the functional $\int_\Omega W(\bullet) \, dx$ in $L^p(\Omega; \mathbb{M})$, and (3.8) verify

$$\begin{aligned} E(v) &\leq \liminf_{\ell \rightarrow \infty} \left(\int_\Omega (W(\mathcal{G}_\ell u_\ell) - f \cdot \mathcal{J}_\ell u_\ell) \, dx - \int_{\Gamma_N} g \cdot \mathcal{J}_\ell u_\ell \, ds \right) \\ &= \liminf_{\ell \rightarrow \infty} \left(E_\ell(u_\ell) - \int_\Omega \mathcal{J}_\ell u_\ell \cdot (1 - \Pi_{\mathcal{T}_\ell}^k) f \, dx - \int_{\Gamma_N} \mathcal{J}_\ell u_\ell \cdot (1 - \Pi_{\mathcal{F}_\ell}^k) g \, ds \right). \end{aligned} \tag{4.21}$$

As in (4.18), a piecewise application of the Poincaré inequality, the trace inequality, the approximation property of polynomials, and the uniform bound $\|D\mathcal{J}_\ell u_\ell\|_{L^p(\Omega)} \lesssim 1$ from (4.2) confirm

$$\begin{aligned} & \left| \int_{\Omega} \mathcal{J}_\ell u_\ell \cdot (1 - \Pi_{\mathcal{T}_\ell}^k) f \, dx \right| + \left| \int_{\Gamma_N} \mathcal{J}_\ell u_\ell \cdot (1 - \Pi_{\mathcal{F}_\ell}^k) g \, ds \right| \\ & \lesssim \text{osc}(f, \mathcal{T}_\ell) + \text{osc}_N(g, \mathcal{F}_\ell(\Gamma_N)). \end{aligned} \tag{4.22}$$

Since $\text{osc}(f, \mathcal{T}_\ell)^{p'} + \text{osc}_N(g, \mathcal{F}_\ell(\Gamma_N))^{p'} \lesssim \eta_\ell^{(\varepsilon)}$ and $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)} = 0$ from Step 2, the LEB from (4.15) and (4.20)–(4.22) lead to

$$E(u) \leq E(v) \leq \liminf_{\ell \rightarrow \infty} E_\ell(u_\ell) = \liminf_{\ell \rightarrow \infty} \text{LEB}_\ell \leq E(u).$$

Hence $\lim_{\ell \rightarrow \infty} E_\ell(u_\ell) = \lim_{\ell \rightarrow \infty} \text{LEB}_\ell = E(u)$ for a (not relabelled) subsequence of $(u_\ell)_{\ell \in \mathbb{N}_0}$. Since the above arguments from Step 4 apply to all subsequences of $(u_\ell)_{\ell \in \mathbb{N}_0}$ and the limit $E(u)$ is unique, this holds for the entire sequence.

Step 5 is the finish of the proof. Suppose that W satisfies (2.5). Then the arguments from [27, 28] show, for all $\varrho, \xi \in L^p(\Omega; \mathbb{M})$ and r, t from Table 1, that

$$\begin{aligned} \|\varrho - \xi\|_{L^p(\Omega)}^r & \leq 3c_5(|\Omega| + \|\varrho\|_{L^p(\Omega)}^p + \|\xi\|_{L^p(\Omega)}^p)^{t/t'} \\ & \quad \times \int_{\Omega} (W(\varrho) - W(\xi) - DW(\xi) : (\varrho - \xi)) \, dx. \end{aligned} \tag{4.23}$$

The choice $\varrho := Du$ and $\xi := \mathcal{G}_\ell u_\ell$ in (4.23) and the bounds $\|Du\|_{L^p(\Omega)} \leq C_2$ and $\|\mathcal{G}_\ell u_\ell\|_{L^p(\Omega)} \leq C_1$ lead, with the constant $C_4 := 3c_5(|\Omega| + C_1^p + C_2^p)^{t/t'}$, to

$$\begin{aligned} & C_4^{-1} \|Du - \mathcal{G}_\ell u_\ell\|_{L^p(\Omega)}^r \\ & \leq \int_{\Omega} (W(Du) - W(\mathcal{G}_\ell u_\ell) - DW(\mathcal{G}_\ell u_\ell) : (Du - \mathcal{G}_\ell u_\ell)) \, dx. \end{aligned} \tag{4.24}$$

The right-hand side of (4.24) coincides with the right-hand side of (4.16). The latter is bounded by the right-hand side of (4.15) in Step 3. This implies

$$C_4^{-1} \|Du - \mathcal{G}_\ell u_\ell\|_{L^p(\Omega)}^r \leq E(u) - \text{LEB}_\ell. \tag{4.25}$$

Step 4 proves that $E(u) - \text{LEB}_\ell$ vanishes in the limit as $\ell \rightarrow \infty$. Thus,

$$\lim_{\ell \rightarrow \infty} \mathcal{G}_\ell u_\ell = Du \text{ (strongly) in } L^p(\Omega; \mathbb{M}).$$

If W satisfies (2.6), then [28, Lemma 4.2] implies, for all $\varrho, \xi \in L^p(\Omega; \mathbb{M})$ and \tilde{r}, \tilde{t} from Table 1, that

$$\begin{aligned} \|DW(\varrho) - DW(\xi)\|_{L^p(\Omega)}^{\tilde{r}} &\leq 3c_6(|\Omega| + \|\varrho\|_{L^p(\Omega)}^p + \|\xi\|_{L^p(\Omega)}^p)^{\tilde{t}/\tilde{r}} \\ &\times \int_{\Omega} (W(\varrho) - W(\xi) - DW(\xi) : (\varrho - \xi)) \, dx, \end{aligned} \tag{4.26}$$

whence the left-hand side of (4.25) can be replaced by $C_5^{-1} \|\sigma - DW(\mathcal{G}_\ell u_\ell)\|_{L^{p'}(\Omega)}^{\tilde{r}}$ with $C_5 := 3c_6(|\Omega| + C_1^p + C_2^p)^{\tilde{t}/\tilde{r}}$. This and Step 4 conclude the proof of

$$\lim_{\ell \rightarrow \infty} DW(\mathcal{G}_\ell u_\ell) = \sigma \text{ (strongly) in } L^{p'}(\Omega; \mathbb{M}).$$

Remark 4.4 (necessity of $\varepsilon > 0$) The counter example in [52, Subsection 3.4] shows that the restriction $\varepsilon > 0$ is necessary. Indeed, for $k = 0$, the data $W, \Omega, \Gamma_D, \Gamma_N, f, g, u_D$, and $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}_0}$ from [52, Subsection 3.4], there exists a sequence of discrete minimizers $(u_\ell)_{\ell \in \mathbb{N}_0}$ such that $\mathcal{J}_\ell u_\ell \rightarrow 0$ weakly in V and $\mathcal{G}_\ell u_\ell \rightarrow 0$ weakly in $L^p(\Omega; \mathbb{M})$ as $\ell \rightarrow \infty$, but $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)} \neq 0$.

4.3 The Lavrentiev gap

A particular challenge in the computational calculus of variations is the Lavrentiev phenomenon $\inf E(\mathcal{A}) < \inf E(\mathcal{A} \cap W^{1,\infty}(\Omega))$ [48]. Its presence is equivalent to the failure of standard conforming FEMs [17, Theorem 2.1] in the sense that a wrong minimal energy is approximated. As a remedy, the nonconforming Crouzeix–Raviart FEM in [3, 51, 52] can overcome the Lavrentiev gap under fairly general assumptions on W : Throughout the remaining parts of this section, let $W \in C^1(\mathbb{M})$ be convex with the one-sided lower growth

$$c_1|A|^p - c_2 \leq W(A) \text{ for all } A \in \mathbb{M} \text{ and some } 1 < p < \infty.$$

(A two-sided growth of W excludes a Lavrentiev gap.) Since there is no upper growth of W , the dual variable $\sigma := DW(Du)$ is not guaranteed to be in $L^{p'}(\Omega; \mathbb{M})$. This denies an access to the Euler–Lagrange equations and, therefore, the convergence analysis of [51, 52] solely relies on the Jensen inequality. For $k = 0$, HHO methods can overcome the Lavrentiev gap because the Crouzeix–Raviart FEM can.

Lemma 4.5 (Lower-energy bound for $k = 0$) *Let $k = 0$. There exists a positive constant C_6 such that, for all level $\ell \in \mathbb{N}_0$,*

$$\min E_\ell(\mathcal{A}(\mathcal{T}_\ell)) - C_6(\|h_\ell f\|_{L^{p'}(\Omega)} - \text{osc}_N(g, \mathcal{F}_\ell(\Gamma_N))) \leq \min E(\mathcal{A}).$$

Proof Recall the discrete space $\text{CR}^1(\mathcal{T}_\ell; \mathbb{R}^m)$ of Crouzeix–Raviart finite element functions from (3.1). Define

$$\text{CR}_D^1(\mathcal{T}_\ell; \mathbb{R}^m) := \{v_{\text{CR}} \in \text{CR}^1(\mathcal{T}_\ell; \mathbb{R}^m) : v_{\text{CR}}(\text{mid}(F)) = 0 \text{ for all } F \in \mathcal{F}_\ell(\Gamma_D)\}$$

and the nonconforming interpolation $I_{CR} : V \rightarrow CR^1(\mathcal{T}_\ell; \mathbb{R}^m)$ [31] with

$$I_{CR}v(\text{mid}(F)) := \int_F v \, ds / |F| \quad \text{for all } F \in \mathcal{F}_\ell, v \in V.$$

The discrete CR-FEM minimizes the non-conforming energy

$$E_{NC}(v_{CR}) := \int_\Omega (W(D_{pw}v_{CR}) - \Pi_{\mathcal{T}_\ell}^0 f \cdot v_{CR}) \, dx - \int_{\Gamma_N} \Pi_{\mathcal{F}_\ell}^0 g \cdot v_{CR} \, ds$$

among $v_{CR} \in \mathcal{A}_{NC} := I_{NC}u_D + CR_D^1(\mathcal{T}_\ell; \mathbb{R}^m)$. A straight-forward modification of the proof of [52, Lemma 4] shows, for a positive constant $C_6 > 0$, that

$$\min E_{NC}(\mathcal{A}_{NC}) - C_6(\|h_\ell f\|_{L^{p'}(\Omega)} - \text{osc}_N(g, \mathcal{F}_\ell(\Gamma_N))) \leq \min E(\mathcal{A}). \quad (4.27)$$

Notice that I_{CR} does *not* provide the L^2 orthogonality $I_{CR}v - v \perp P_0(\mathcal{T}_\ell; \mathbb{R}^m)$ in $L^2(\Omega; \mathbb{R}^m)$, but $(I_{CR}v - v)|_F \perp P_0(F; \mathbb{R}^m)$ in $L^2(F; \mathbb{R}^m)$ for all $F \in \mathcal{F}_\ell(\Gamma_N)$ and $v \in V$. Hence the Neumann boundary data oscillations $\text{osc}_N(g, \mathcal{F}_\ell(\Gamma_N))$ arise in (4.27), but $\|h_\ell f\|_{L^{p'}(\Omega)}$ cannot be replaced by $\text{osc}(f, \mathcal{T}_\ell)$. For any $v_{CR} \in \mathcal{A}_{NC}$, $v_\ell := (\Pi_{\mathcal{T}_\ell}^0 v_{CR}, \Pi_{\mathcal{F}_\ell}^0 v_{CR}) \in \mathcal{A}(\mathcal{T}_\ell)$ satisfies $\mathcal{G}_\ell v_\ell = D_{pw}v_{CR}$ and hence, $\min E_\ell(\mathcal{A}(\mathcal{T}_\ell)) \leq \min E_{NC}(\mathcal{A}_{NC})$. This and (4.27) conclude the proof. \square

The discrete compactness from Theorem 4.1, the LEB in (4.27), and straightforward modifications of the proof of Theorem 2.1 lead to $\lim_{\ell \rightarrow \infty} E_\ell(u_\ell) = \min E(\mathcal{A})$ for the output $(u_\ell)_{\ell \in \mathbb{N}_0}$ of the adaptive algorithm in Sect. 2.2 with the refinement indicator, for all $T \in \mathcal{T}_\ell$,

$$\eta_\ell^{(\varepsilon)}(T) := \mu_\ell^{(\varepsilon)}(T) + |T|^{p'/n} \|f\|_{L^{p'}(T)}^{p'} + |T|^{1/n} \sum_{F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Gamma_N)} \|(1 - \Pi_F^0)g\|_{L^{p'}(F)}^{p'}.$$

For $k \geq 1$, the consistency error $\sigma_\ell - DW(\mathcal{G}_\ell u_\ell)$ arises in (4.15), but is not guaranteed to be bounded in $L^{p'}(\Omega; \mathbb{M})$ in the limit as $\ell \rightarrow \infty$ in general. Thus, in the absence of further conditions, the convergence $\lim_{\ell \rightarrow \infty} E_\ell(u_\ell) = \min E(\mathcal{A})$ cannot be proven for $k \geq 1$ with this methodology.

5 Stabilized HHO method on polytopal meshes

The classical HHO methodology [35, 37] allows even polytopal partitions of the domain Ω . The assumption (M1) on the mesh follows the works [35, 37, 40].

5.1 Polytopal meshes

Let \mathcal{M}_ℓ be a finite collection of closed polytopes of positive volume with overlap of volume measure zero that cover $\overline{\Omega} = \cup_{K \in \mathcal{M}_\ell} K$. A side S of the mesh \mathcal{M}_ℓ is the (in

general disconnected) closed subset of a hyperplane $H_S \subset \Omega$ with positive $((n - 1)$ -dimensional) surface measure such that either (a) there exist $K_1, K_2 \in \mathcal{M}_\ell$ with $S = \partial K_1 \cap \partial K_2 \cap H_S$ (interior side) or (b) there exists $K \in \mathcal{M}_\ell$ with $S = \partial K \cap \partial \Omega \cap H_S$ (boundary side). Let Σ_ℓ denote the set of all sides of \mathcal{M}_ℓ and adapt the notation $\Sigma_\ell(K)$, $\Sigma_\ell(\Omega)$, $\Sigma_\ell(\Gamma_D)$, and $\Sigma_\ell(\Gamma_N)$ from Sect. 3.1. The convergence results of this section are established under the assumptions (M1)–(M2) below.

- (M1) Assume that there exists a universal constant $\varrho > 0$ such that, for all level $\ell \in \mathbb{N}_0$, \mathcal{M}_ℓ admits a shape-regular simplicial subtriangulation \mathcal{T}_ℓ with the shape regularity $\geq \varrho$ defined in Sect. 3.1 and, for each simplex $T \in \mathcal{T}_\ell$, there exists a unique cell $K \in \mathcal{M}_\ell$ with $T \subseteq K$ and $\varrho h_K \leq h_T$.
- (M2) Assume the existence of a universal constant $0 < \gamma < 1$ such that $|\widehat{K}| \leq \gamma|K|$ holds for all $K \in \mathcal{M}_\ell \setminus \mathcal{M}_{\ell+1}$, $\widehat{K} \in \mathcal{M}_{\ell+1}$ with $\widehat{K} \subset K$ and level $\ell \in \mathbb{N}_0$, i.e., the volume measure of all children \widehat{K} of a refined cell K is at most $\gamma|K|$.

The assumption (M1) is typical for the error analysis of HHO methods on polytopal meshes, cf., e.g., [34, 35, 37, 40, 42]. The assumption (M2) holds for the newest-vertex bisection on simplicial triangulations with $\gamma = 1/2$.

Remark 5.1 (equivalence of side lengths) The assumption (M1) ensures that $h_S \approx h_K \approx |K|^{1/n}$ holds for all $K \in \mathcal{M}_\ell$ and $S \in \Sigma_\ell(K)$ with equivalence constants that exclusively depend on the universal constant ϱ in (M1) [40, Lemma 1.42].

Lemma 5.2 (mesh-size reduction) *Suppose that the sequence $(\mathcal{M}_\ell)_{\ell \in \mathbb{N}_0}$ satisfies (M2), then the mesh-size function $h_\ell \in P_0(\mathcal{M}_\ell)$ with $h_\ell|_K := |K|^{1/n}$ for all $K \in \mathcal{M}_\ell$ satisfies $\lim_{\ell \rightarrow \infty} \|h_\ell\|_{L^\infty(\Omega_\ell)} = 0$ for $\Omega_\ell := \text{int}(\cup(\mathcal{M}_\ell \setminus \mathcal{M}_{\ell+1}))$.*

Proof Given any $j \in \mathbb{N}_0$ and $\alpha_j := \gamma^j|\Omega|$, define the set $\mathcal{M}(j) \subset \cup_{\ell \in \mathbb{N}_0} \mathcal{M}_\ell$ of all polytopes K with volume measure $\alpha_{j+1} < |K| \leq \alpha_j$. Since the volume measure of any refined polytope is at least reduced by the factor γ , the polytopes of $\mathcal{M}(j)$ are not children of each other and so $|K \cap T| = 0$ holds for any two distinct polytopes $K, T \in \mathcal{M}(j)$. This implies that the cardinality $|\mathcal{M}(j)|$ of $\mathcal{M}(j)$ satisfies $|\mathcal{M}(j)| < \gamma^{-(j+1)}$. For any level $\ell \in \mathbb{N}_0$, select some $K_\ell \in \mathcal{M}_\ell \setminus \mathcal{M}_{\ell+1}$ with $|K_\ell| = \|h_\ell\|_{L^\infty(\Omega_\ell)}^n$. Since $K_\ell \notin \mathcal{M}_j$ for all $j > \ell$, the polytopes K_0, K_1, K_2, \dots are pairwise distinct. Given $N \in \mathbb{N}_0$, the number $|\{\ell \in \mathbb{N}_0 : |K_\ell| > \alpha_{N+1}\}|$ of all indices $\ell \in \mathbb{N}_0$ with $|K_\ell| > \alpha_{N+1}$ is bounded by $|\mathcal{M}(0)| + |\mathcal{M}(1)| + \dots + |\mathcal{M}(N)| \leq (\gamma^{-(N+1)} - 1)/(1 - \gamma)$. Hence there exists a maximal index L such that $\|h_\ell\|_{L^\infty(\Omega_\ell)}^n = |K_\ell| \leq \alpha_{N+1}$ for all $\ell \geq L$. Notice that Lemma 4.3 follows for simplicial triangulations with $\gamma = 1/2$. \square

5.2 Stabilization

The classical HHO method [1, 34] utilizes a gradient reconstruction $\mathcal{G}_\ell : V(\mathcal{M}_\ell) \rightarrow \Sigma(\mathcal{M}_\ell)$ in the space $\Sigma(\mathcal{M}_\ell) := P_k(\mathcal{M}_\ell; \mathbb{M})$ of matrix-valued piecewise polynomials of total degree at most k . The discrete seminorm $\|\bullet\|_\ell$ of $V(\mathcal{M}_\ell)$ and the operators $I_\ell, \mathcal{R}_\ell, \mathcal{G}_\ell$ of this section are defined by the formulas (3.2)–(3.6) in Sect. 3.3 with adapted notation, i.e., \mathcal{T}_ℓ (resp. \mathcal{F}_ℓ) is replaced by \mathcal{M}_ℓ (resp. Σ_ℓ).

Remark 5.3 (need of stabilization) The kernel of the gradient reconstruction \mathcal{G}_ℓ restricted to $V_D(\mathcal{M}_\ell)$ is not trivial. For instance, any $v_\ell = (v_{\mathcal{M}_\ell}, 0) \in V_D(\mathcal{M}_\ell)$

with $v_{\mathcal{M}_\ell} \in P_k(\mathcal{M}_\ell; \mathbb{R}^m)$ and $v_{\mathcal{M}_\ell} \perp P_{k-1}(\mathcal{M}_\ell; \mathbb{R}^m)$ (with the convention $P_{-1}(\mathcal{M}_0; \mathbb{R}^m) := \{0\}$) satisfies $\mathcal{G}_\ell v_\ell = 0$ and the norm equivalence in Lemma 3.1a fails. On simplicial meshes, a gradient reconstruction in any discrete space $\Sigma(\mathcal{M}_\ell)$ with $\text{RT}_k(\mathcal{M}_\ell; \mathbb{M}) \subset \Sigma(\mathcal{M}_\ell)$ is stable, but the commutativity from Lemma 3.1b may fail if $\Sigma(\mathcal{M}_\ell)$ is too large, e.g., $\Sigma(\mathcal{M}_\ell) = P_{k+1}(\mathcal{M}_\ell; \mathbb{M})$ [1].

The stabilization function $s_\ell : V(\mathcal{M}_\ell) \times V(\mathcal{M}_\ell) \rightarrow \mathbb{R}$ in the HHO methodology is defined, for any $u_\ell, v_\ell = (v_{\mathcal{M}_\ell}, v_{\Sigma_\ell}) \in V(\mathcal{M}_\ell)$ and any side $S \in \Sigma_\ell(K)$ of $K \in \mathcal{M}_\ell$ with diameter $h_S = \text{diam}(S)$, by $s_\ell(u_\ell; v_\ell) := \sum_{K \in \mathcal{M}_\ell} s_K(u_\ell; v_\ell)$ and

$$s_{K,S} v_\ell := \Pi_S^k(v_S - v_K - (1 - \Pi_K^k)(\mathcal{R}_\ell v_\ell)|_K) \in P_k(S; \mathbb{R}^m),$$

$$s_K(u_\ell; v_\ell) := \sum_{S \in \Sigma_\ell(K)} h_S^{1-p} \int_S |S_{K,S} u_\ell|^{p-2} S_{K,S} u_\ell \cdot S_{K,S} v_\ell \, ds. \tag{5.1}$$

Notice that $s_\ell(\bullet; \bullet)$ is linear in the second component, but *not* in the first unless $p = 2$. The relevant properties of $s_\ell(\bullet; \bullet)$ are summarized below.

Lemma 5.4 (stabilization) *Any $u_\ell, v_\ell = (v_{\mathcal{M}_\ell}, v_{\Sigma_\ell}) \in V(\mathcal{M}_\ell), v \in V$, and $K \in \mathcal{M}_\ell$ satisfy (a)–(e) with parameters p, r, s, t from Table 1.*

- (a) $\|v_\ell\|_\ell^p \approx \|\mathcal{G}_\ell v_\ell\|_{L^p(\Omega)}^p + s_\ell(v_\ell; v_\ell)$.
- (b) $s_K(\mathbb{I}_\ell v; \mathbb{I}_\ell v)^{1/p} \lesssim \min_{\varphi_h \in P_{k+1}(K; \mathbb{R}^m)} \|\mathbb{D}(v - \varphi_h)\|_{L^p(K)}$. In particular, if $v \in W^{k+2,p}(K; \mathbb{R}^m)$, then $s_K(\mathbb{I}_\ell v; \mathbb{I}_\ell v)^{1/p} \lesssim h_K^{k+1} |v|_{W^{k+2,p}(K)}$.
- (c)

$$s_K(v_\ell; v_\ell) \lesssim h_K^{-p} \|\Pi_K^k(\mathcal{R}_\ell v_\ell - v_K)\|_{L^p(K)}^p + \sum_{S \in \Sigma_\ell(K)} h_S^{1-p} \|\Pi_S^k((\mathcal{R}_\ell v_\ell)_K - v_S)\|_{L^p(S)}^p.$$

- (d) $s_K(u_\ell; v_\ell) \leq s_K(u_\ell; u_\ell)^{1/p'} s_K(v_\ell; v_\ell)^{1/p}$.
- (e)

$$\sum_{K \in \mathcal{M}_\ell} \sum_{S \in \Sigma_\ell(K)} \|h_S^{-1/p'} S_{K,S}(u_\ell - v_\ell)\|_{L^p(S)}^r$$

$$\lesssim (1 + s_\ell(u_\ell; u_\ell) + s_\ell(v_\ell; v_\ell))^{t/t'} (s_\ell(v_\ell; v_\ell)/p - s_\ell(u_\ell; v_\ell) + s_\ell(u_\ell; u_\ell)/p').$$

Proof The norm equivalence in (a) is established in [37, Lemma 4] for $p = 2$ and extended to $1 \leq p < \infty$ in [34, Lemma 5.2]; the approximation property (b) is [42, Lemma 3.2]. The upper bound (c) follows immediately from a triangle and a discrete trace inequality. The proof of (d) concerns $K \in \mathcal{M}_\ell$ and $S \in \Sigma_\ell(K)$. A

Hölder inequality with the exponents p, p' and $1 - p + 1/p' = (1 - p)/p'$ show

$$\begin{aligned} & h_S^{1-p} \int_S |\mathcal{S}_{K,Su_\ell}|^{p-2} \mathcal{S}_{K,Su_\ell} \cdot \mathcal{S}_{K,Sv_\ell} \, ds \\ & \leq \|h_S^{(1-p)/p'} |\mathcal{S}_{K,Su_\ell}|^{p-2} \mathcal{S}_{K,Su_\ell}\|_{L^{p'}(S)} \|h_S^{-1/p'} \mathcal{S}_{K,Sv_\ell}\|_{L^p(S)} \\ & = \|h_S^{-1/p'} \mathcal{S}_{K,Su_\ell}\|_{L^p(S)}^{p/p'} \|h_S^{-1/p'} \mathcal{S}_{K,Sv_\ell}\|_{L^p(S)}. \end{aligned}$$

The sum of this over all $S \in \Sigma_\ell(K)$ and a Cauchy inequality prove (d). The proof of (e) departs from the function $W(a) := |a|^p/p$ for $a \in \mathbb{R}^m$ with the convexity control (2.5). The integral of (2.5) over the side S leads to (4.23) for all $\varrho, \xi \in L^p(S; \mathbb{R}^m)$ and Ω (resp. \mathbb{M}) replaced by S (resp. \mathbb{R}^m). The choice $\varrho := h_S^{-1/p'} \mathcal{S}_{K,Sv_\ell}$ and $\xi := h_S^{-1/p'} \mathcal{S}_{K,Su_\ell}$ in (4.23) leads to

$$\begin{aligned} & (3c_5)^{-1} (|S| + \|h_S^{-1/p'} \mathcal{S}_{K,Su_\ell}\|_{L^p(S)}^p + \|h_S^{-1/p'} \mathcal{S}_{K,Sv_\ell}\|_{L^p(S)}^p)^{-t/t'} \\ & \quad \times \|h_S^{-1/p'} \mathcal{S}_{K,S}(u_\ell - v_\ell)\|_{L^p(S)}^r \leq \|h_S^{-1/p'} \mathcal{S}_{K,Sv_\ell}\|_{L^p(S)}^p / p \\ & \quad - \int_S h_S^{1-p} |\mathcal{S}_{K,Su_\ell}|^{p-2} \mathcal{S}_{K,Su_\ell} \cdot \mathcal{S}_{K,Sv_\ell} \, ds + \|h_S^{-1/p'} \mathcal{S}_{K,Su_\ell}\|_{L^p(S)}^p / p'. \end{aligned} \tag{5.2}$$

The sum of this over all $S \in \Sigma_\ell(K)$ and $K \in \mathcal{M}_\ell$ concludes the proof of (e). □

5.3 Stabilized HHO method on a polytopal mesh

The discrete problem minimizes the discrete energy

$$E_\ell(v_\ell) := \int_\Omega (W(\mathcal{G}_\ell v_\ell) - f \cdot v_{\mathcal{M}_\ell}) \, dx - \int_{\Gamma_N} g \cdot v_{\Sigma_\ell} \, ds + s_\ell(v_\ell; v_\ell) / p \tag{5.3}$$

among $v_\ell = (v_{\mathcal{M}_\ell}, v_{\Sigma_\ell}) \in \mathcal{A}(\mathcal{M}_\ell)$.

Theorem 5.5 (discrete minimizers) *The minimal discrete energy $\inf E_\ell(\mathcal{A}(\mathcal{M}_\ell))$ is attained. There exists a positive constant $C_7 > 0$ that merely depends on $c_1, c_2, \Omega, \Gamma_D, u_D, f, g, \varrho$ in (M1), k , and p with $\|\mathcal{G}_\ell u_\ell\|_{L^p(\Omega)}^p + s_\ell(u_\ell; u_\ell) \leq C_7^p$ for any discrete minimizer $u_\ell \in \arg \min E_\ell(\mathcal{A}(\mathcal{M}_\ell))$. Any discrete stress $\sigma_\ell := \Pi_{\mathcal{M}_\ell}^k \text{DW}(\mathcal{G}_\ell u_\ell)$ satisfies the discrete Euler–Lagrange equations*

$$\int_\Omega \sigma_\ell : \mathcal{G}_\ell v_\ell \, dx = \int_\Omega f \cdot v_{\mathcal{M}_\ell} \, dx + \int_{\Gamma_N} g \cdot v_{\Sigma_\ell} \, ds - s_\ell(u_\ell; v_\ell) \tag{5.4}$$

for all $v_\ell = (v_{\mathcal{M}_\ell}, v_{\Sigma_\ell}) \in V_D(\mathcal{M}_\ell)$. If W satisfies (2.5), then $u_\ell = \arg \min E_\ell(\mathcal{A}(\mathcal{M}_\ell))$ is unique. If W satisfies (2.6), then $\text{DW}(\mathcal{G}_\ell u_\ell) \in L^{p'}(\Omega; \mathbb{M})$ is unique (independent of the choice of a (possibly non-unique) discrete minimizer u_ℓ).

Proof The proof follows that of Theorem 3.2. The norm equivalence in Lemma 5.4a and the lower growth of W lead to the coercivity of E_ℓ in $\mathcal{A}(\mathcal{M}_\ell)$ with respect to

the seminorm $\|\bullet\|_\ell^p \approx \|\mathcal{G}_\ell \bullet\|_{L^p(\Omega)}^p + s_\ell(\bullet; \bullet)$ from Lemma 5.4a. This implies the existence of discrete minimizers and the bound $\|\mathcal{G}_\ell u_\ell\|_{L^p(\Omega)}^p + s_\ell(u_\ell; u_\ell) \leq C_7^p$ for all $u_\ell \in \arg \min E_\ell(\mathcal{A}(\mathcal{M}_\ell))$. If W satisfies (2.5), then the strict convexity of W and of s_ℓ in Lemma 5.4c leads to the uniqueness of $u_\ell = \arg \min E_\ell(\mathcal{A}(\mathcal{M}_\ell))$. If W satisfies (2.6), then the uniqueness of $DW(\mathcal{G}_\ell u_\ell)$ follows as in [16, 27, 28]. \square

The following lemma extends Lemma 3.4 to polytopal meshes.

Lemma 5.6 *There exists a linear operator $\mathcal{J}_\ell : V(\mathcal{M}_\ell) \rightarrow V$ such that any $v_\ell = (v_{\mathcal{M}_\ell}, v_{\Sigma_\ell}) \in V(\mathcal{M}_\ell)$ satisfies*

$$\Pi_{\mathcal{M}_\ell}^k \mathcal{J}_\ell v_\ell = v_{\mathcal{M}_\ell} \quad \text{and} \quad \Pi_{\Sigma_\ell}^k \mathcal{J}_\ell v_\ell = v_{\Sigma_\ell} \tag{5.5}$$

and, for any $K \in \mathcal{M}_\ell$, the estimate

$$\begin{aligned} \|\mathcal{G}_\ell v_\ell - D\mathcal{J}_\ell v_\ell\|_{L^p(K)}^p &\lesssim \sum_{S \in \Sigma_\ell(\Omega), S \cap K \neq \emptyset} h_S^{1-p} \|[\mathcal{R}_\ell v_\ell]_S\|_{L^p(S)}^p \\ &+ \sum_{S \in \Sigma_\ell(K)} h_S^{1-p} \|(\mathcal{R}_\ell v_\ell)|_K - v_S\|_{L^p(S)}^p + h_K^{-p} \|\mathcal{R}_\ell v_\ell - v_K\|_{L^p(K)}^p. \end{aligned} \tag{5.6}$$

In particular, \mathcal{J}_ℓ is stable in the sense that $\|D\mathcal{J}_\ell v_\ell\|_{L^p(\Omega)} \leq \Lambda_1 \|v_\ell\|_\ell$ holds with the constant Λ_1 that exclusively depends on k, p , and q in (M1).

Proof The construction of the conforming operator \mathcal{J}_ℓ on polytopal meshes in [42, Section 5] utilizes averaging and bubble-function techniques on the subtriangulation \mathcal{T}_ℓ and give rise an upper bound of $\|\mathcal{G}_\ell v_\ell - D\mathcal{J}_\ell v_\ell\|_{L^p(K)}$, namely

$$\begin{aligned} &\sum_{S \in \Sigma_\ell(\Omega), S \cap K \neq \emptyset} h_S^{1-p} \|[\mathcal{R}_\ell v_\ell]_S\|_{L^p(S)}^p + \sum_{T \in \mathcal{T}_\ell, T \subset K} h_T^{-p} \|\Pi_T^k(\mathcal{R}_\ell v_\ell - v_K)\|_{L^p(T)}^p \\ &+ \sum_{S \in \Sigma_\ell(K)} \sum_{F \in \mathcal{F}_\ell, F \subset S} h_F^{1-p} \|\Pi_F((\mathcal{R}_\ell v_\ell)|_K - v_S)\|_{L^p(F)}^p. \end{aligned} \tag{5.7}$$

Since the L^2 projection Π_T^k (resp. Π_F^k) is stable in $L^p(T; \mathbb{R}^m)$ (resp. $L^p(F; \mathbb{R}^m)$) [34, Lemma 3.2], it can be omitted in (5.7). This, the equivalence $h_T \approx h_K$ for all $K \in \mathcal{M}_\ell, T \in \mathcal{T}_\ell$ with $T \subset K$ from (M1), and $h_F \approx h_S$ for all $S \in \Sigma_\ell, F \in \mathcal{F}_\ell$ with $F \subset S$ from (M1) and Remark 5.1 show (5.6). This implies the stability $\|D\mathcal{J}_\ell v_\ell\|_{L^p(\Omega)} \lesssim \|v_\ell\|_\ell$, cf., e.g., [42, Subsection 4.3] for more details. Notice that the computation of the right-hand side of (5.6) does *not* require explicit information on the subtriangulation \mathcal{T}_ℓ . \square

Remark 5.7 (discrete compactness) The discrete compactness from Theorem 4.1 holds verbatim with \mathcal{T}_ℓ (resp. \mathcal{F}_ℓ) replaced by \mathcal{M}_ℓ (resp. Σ_ℓ). Notice that \mathcal{G}_ℓ from Sect. 5.2 and \mathcal{J}_ℓ from Lemma 5.6 in this section are different objects. With adapted notation, all arguments from the proof of Theorem 4.1 apply verbatim. Indeed, the commutativity $\Pi_{\Sigma(\mathcal{M}_\ell)} Dv = \mathcal{G}_\ell I_\ell v$ for all $v \in V$ from Lemma 3.1b remains valid [1, 34]. This and

(5.5) imply the L^2 orthogonality $\mathcal{G}_\ell v_\ell - \text{DW}(\mathcal{G}_\ell u_\ell) \perp \Sigma(\mathcal{M}_\ell)$ for any $v_\ell \in V(\mathcal{M}_\ell)$. This is the key argument in the proof of Theorem 4.1 and provides a positive power of the mesh-size in (4.1).

5.4 Proof of Theorem 2.2

Given any $K \in \mathcal{M}_\ell$, Lemma 5.6 motivates the refinement indicator

$$\eta_\ell^{(\varepsilon)}(K) := \mu_\ell^{(\varepsilon)}(K) + |K|^{\varepsilon p'/n} \|\sigma_\ell - \text{DW}(\mathcal{G}_\ell u_\ell)\|_{L^{p'}(K)}^{p'} + |K|^{p'/n} \|(1 - \Pi_K^k) f\|_{L^{p'}(K)}^{p'} + |K|^{1/n} \sum_{S \in \Sigma_\ell(K) \cap \Sigma_\ell(\Gamma_N)} \|(1 - \Pi_S^k) g\|_{L^{p'}(S)}^{p'}$$

with

$$\begin{aligned} \mu_\ell^{(\varepsilon)}(K) := & |K|^{(\varepsilon p - p)/n} \|\mathcal{R}_\ell u_\ell - u_K\|_{L^p(K)}^p \\ & + |K|^{(\varepsilon p + 1 - p)/n} \left(\sum_{S \in \Sigma_\ell(K) \cap \Sigma_\ell(\Gamma_D)} \|\mathcal{R}_\ell u_\ell - u_D\|_{L^p(S)}^p \right. \\ & \left. + \sum_{S \in \Sigma_\ell(K) \cap \Sigma_\ell(\Omega)} \|[\mathcal{R}_\ell u_\ell]_S\|_{L^p(S)}^p + \sum_{S \in \Sigma_\ell(K)} \|(\mathcal{R}_\ell u_\ell)|_K - u_S\|_{L^p(S)}^p \right). \end{aligned}$$

The remaining parts of this section are devoted to the proof of the convergence results in Theorem 2.2.

Proof of Theorem 2.2 The proof follows that of Theorem 2.1.

Step 1 establishes $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)} = 0$. The key argument from Step 1 of the proof of Theorem 2.1 is the positive power of the mesh size in $\eta_\ell^{(\varepsilon)}$ in the sense that

$$\begin{aligned} \eta_\ell^{(\varepsilon)}(\mathcal{M}_\ell \setminus \mathcal{M}_{\ell+1}) & \lesssim \|h_\ell\|_{L^\infty(\Omega_\ell)}^{\varepsilon p} (\|u_\ell\|_\ell^p + \|\text{DU}_D\|_{L^p(\Omega)}^p) \\ & \quad + \|h_\ell\|_{L^\infty(\Omega_\ell)}^{\varepsilon p'} \|\text{DW}(\mathcal{G}_\ell u_\ell)\|_{L^{p'}(\Omega)}^{p'} \\ & \quad + \|h_\ell\|_{L^\infty(\Omega_\ell)}^{p'} \|f\|_{L^{p'}(\Omega)}^{p'} + \|h_\ell\|_{L^\infty(\Omega_\ell)} \|g\|_{L^{p'}(\Gamma_N)}^{p'}. \end{aligned}$$

Hence $\lim_{\ell \rightarrow \infty} \|h_\ell\|_{L^\infty(\Omega_\ell)} = 0$ from Lemma 5.2 implies $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)}(\mathcal{M}_\ell \setminus \mathcal{M}_{\ell+1}) = 0$. This and the Dörfler marking in (2.4) conclude $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)} = 0$.

Step 2 provides a LEB with the extra stabilization term $s_\ell(u_\ell; \text{I}_\ell u)$, namely

$$\begin{aligned} \text{LEB}_\ell := & E_\ell(u_\ell) + \int_\Omega (1 - \Pi_{\mathcal{M}_\ell}^k) \text{DW}(\mathcal{G}_\ell u_\ell) : \text{Du} \, dx - s_\ell(u_\ell; \text{I}_\ell u) \\ & - C_3 (\text{osc}(f, \mathcal{M}_\ell) + \text{osc}_N(g, \Sigma_\ell(\Gamma_N))) \leq E(u) - s_\ell(u_\ell; u_\ell)/p' \leq E(u). \end{aligned} \tag{5.8}$$

The commutativity $\Pi_{\Sigma(\mathcal{M}_\ell)} Du = \mathcal{G}_\ell I_\ell u$ from Lemma 3.1b and the discrete Euler–Lagrange equations (5.4) show that

$$\int_{\Omega} \sigma_\ell : (Du - \mathcal{G}_\ell u_\ell) \, dx = \int_{\Omega} f \cdot (\Pi_{\mathcal{M}_\ell}^k u - u_{\mathcal{M}_\ell}) \, dx + \int_{\Gamma_N} g \cdot (\Pi_{\Sigma_\ell}^k u - u_{\Sigma_\ell}) \, dx - s_\ell(u_\ell; I_\ell u - u_\ell). \tag{5.9}$$

This, (4.16), and (4.18) (with adapted notation) conclude the proof of (5.8).

Step 3 establishes $\lim_{\ell \rightarrow \infty} E_\ell(u_\ell) = E(u)$. Notice from (5.6) that $\eta_\ell^{(\varepsilon)}$ is an upper bound for $\mu_\ell(u_\ell)$ in (4.1). Hence the discrete compactness (from Remark 5.7) implies the existence of a (not relabelled) subsequence of $(u_\ell)_{\ell \in \mathbb{N}_0}$ and a weak limit $v \in \mathcal{A}$ such that $\mathcal{J}_\ell u_\ell \rightharpoonup v$ weakly in V and $\mathcal{G}_\ell u_\ell \rightharpoonup Dv$ weakly in $L^p(\Omega; \mathbb{M})$ as $\ell \rightarrow \infty$. The only difference between the LEB in (5.8) and that in (4.15) for simplicial meshes is the additional term $s_\ell(u_\ell; Iu)$ in this proof.

Lemma 5.8 (convergence of $s_\ell(u_\ell; Iu)$) *Given a sequence $(u_\ell)_{\ell \in \mathbb{N}_0}$ with $u_\ell \in V(\mathcal{M}_\ell)$ for all $\ell \in \mathbb{N}_0$, suppose that $s_\ell(u_\ell; u_\ell) \leq C_8$ for a universal constant C_8 independent of the level ℓ and $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)} = 0$ with $\varepsilon \leq \min\{k + 1, (k + 1)/(p - 1)\}$. Then*

$$\lim_{\ell \rightarrow \infty} s_\ell(u_\ell; Iu) = 0. \tag{5.10}$$

Proof of Lemma 5.8 The proof of (5.10) first establishes this for smooth functions. Given any $\varphi \in C^\infty(\overline{\Omega}; \mathbb{R}^m)$, the Hölder inequality from Lemma 5.4d, $h_K \approx |K|^{1/n}$ from Remark 5.1, and the interpolation error from Lemma 5.4b prove

$$\begin{aligned} |s_\ell(u_\ell; I_\ell \varphi)| &\leq \sum_{K \in \mathcal{M}_\ell} s_K(u_\ell; u_\ell)^{1/p'} s_K(I_\ell \varphi; I_\ell \varphi)^{1/p} \\ &\lesssim \left(\sum_{K \in \mathcal{M}_\ell} |K|^{(k+1)p'/n} s_K(u_\ell; u_\ell) \right)^{1/p'} |\varphi|_{W^{k+2,p}(\Omega)}. \end{aligned} \tag{5.11}$$

Lemma 5.4c implies that $\eta_\ell^{(\varepsilon)}$ controls the stabilization in the sense that

$$\sum_{K \in \mathcal{M}_\ell} |K|^{(k+1)p'/n} s_K(u_\ell; u_\ell) \lesssim \|h_\ell\|_{L^\infty(\Omega)}^{((k+1)p' - \varepsilon p)/n} \eta_\ell^{(\varepsilon)}. \tag{5.12}$$

The restriction $\varepsilon \leq (k+1)/(p-1)$ provides $(k+1)p' - \varepsilon p > 0$. Hence $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)} = 0$ implies that the right-hand side of (5.12) vanishes in the limit as $\ell \rightarrow \infty$. This and (5.11)–(5.12) lead to $\lim_{\ell \rightarrow \infty} s_\ell(u_\ell; I_\ell \varphi) = 0$ for all $\varphi \in C^\infty(\overline{\Omega}; \mathbb{R}^m)$. Given any $\delta > 0$, let $\varphi \in C^\infty(\overline{\Omega}; \mathbb{R}^m)$ such that $\|D(u - \varphi)\|_{L^p(\Omega)} \leq \delta$. The interpolation error from Lemma 5.4b proves $s_\ell(I_\ell(u - \varphi); I_\ell(u - \varphi)) \leq C_9 L^p \|D(u - \varphi)\|_{L^p(\Omega)}^p \leq C_9^p \delta^p$ with a universal constant $C_9 > 0$. The convergence $\lim_{\ell \rightarrow \infty} s_\ell(u_\ell; I_\ell \varphi) = 0$ implies the existence of $N \in \mathbb{N}_0$ with $|s_\ell(u_\ell; I_\ell \varphi)| \leq \delta$ for all $\ell \geq N$. This, a triangle

inequality, a Hölder inequality, and the bound $s_\ell(u_\ell; u_\ell) \leq C_8$ (by assumption) verify

$$|s_\ell(u_\ell; \mathbf{I}_\ell u)| \leq |s_\ell(u_\ell; \mathbf{I}_\ell \varphi)| + |s_\ell(u_\ell; \mathbf{I}_\ell(u - \varphi))| \leq |s_\ell(u_\ell; \mathbf{I}_\ell \varphi)| + s_\ell(u_\ell; u_\ell)^{1/p'} s_\ell(\mathbf{I}_\ell(u - \varphi); \mathbf{I}_\ell(u - \varphi))^{1/p} \leq (1 + C_8^{1/p'} C_9) \delta.$$

This concludes the proof of $\lim_{\ell \rightarrow \infty} s_\ell(u_\ell; \mathbf{I}u) = 0$ in (5.10). □

We return to proof of Theorem 2.2 and recall $s_\ell(u_\ell; u_\ell) \leq C_7^p$ from Theorem 5.5 and $\lim_{\ell \rightarrow \infty} \eta_\ell^{(\varepsilon)} = 0$ from Step 1. Hence Lemma 5.8 applies and (5.10) follows. With this additional argument (5.10) and the remaining conclusions, that lead to (4.21) in the proof of Theorem 2.1, $E(u) \leq E(v) \leq \liminf_{\ell \rightarrow \infty} \text{LEB}_\ell \leq E(u)$ follows for the weak limit v . This implies $\lim_{\ell \rightarrow \infty} E_\ell(u_\ell) = \lim_{\ell \rightarrow \infty} \text{LEB}_\ell = E(u)$. Since $s_\ell(u_\ell; u_\ell)/p' \leq E(u) - \text{LEB}_\ell$ from (5.8), $s_\ell(u_\ell; u_\ell)$ vanishes in the limit as $\ell \rightarrow \infty$. If W satisfies (2.5), then the choice $\varrho := Du$ and $\xi := \mathcal{G}_\ell u_\ell$ in (4.23), (5.9) and the data oscillations from (4.18) imply

$$C_{10}^{-1} \|Du - \mathcal{G}_\ell u_\ell\|_{L^p(\Omega)}^r + s_\ell(u_\ell; u_\ell)/p' \leq E(u) - \text{LEB}_\ell \tag{5.13}$$

with the constant $C_{10} := 3c_5(|\Omega| + C_2^p + C_7^p)^{t/t'}$ and r, t from Table 1. This shows $\lim_{\ell \rightarrow \infty} \mathcal{G}_\ell u_\ell = Du$ (strongly) in $L^p(\Omega; \mathbb{M})$. If W satisfies (2.6), then (4.26) holds and $C_{10}^{-1} \|Du - \mathcal{G}_\ell u_\ell\|_{L^p(\Omega)}^r$ on the left-hand side of (5.13) can be replaced by $C_{11}^{-1} \|\sigma - DW(\mathcal{G}_\ell u_\ell)\|_{L^{p'}(\Omega)}^{\tilde{r}}$ with $C_{11} := 3c_6(|\Omega| + C_2^p + C_7^p)^{\tilde{t}/\tilde{t}'}$ and \tilde{r}, \tilde{t} from Table 1. Hence $\lim_{\ell \rightarrow \infty} DW(\mathcal{G}_\ell u_\ell) = \sigma$ (strongly) in $L^{p'}(\Omega; \mathbb{M})$. □

6 Numerical examples

Some remarks on the implementation precede the numerical benchmarks for the three examples of Sect. 2.4 and the experiments in the Foss–Hrusa–Mizel example with the Lavrentiev gap in Sect. 6.5.

6.1 Implementation

The realization in MATLAB follows that of [28, Subsubsection 5.1.1] with the parameters `FunctionTolerance = OptimalityTolerance = StepTolerance = 10-15` and `MaxIterations = Inf` for improved accuracy.

The class of minimization problems at hand allows, in general, for multiple exact and discrete solutions. The numerical experiments select one (of those) by the approximation in `fminunc` with the initial value computed as follows. On the coarse initial triangulations \mathcal{T}_0 , the initial value $v_0 = (v_{\mathcal{T}_0}, v_{\mathcal{F}_0}) \in V(\mathcal{T}_0)$ is defined by $v_{\mathcal{T}_0} \equiv 1$, $v_{\mathcal{F}_0}|_F \equiv 1$ on any $F \in \mathcal{F}_0(\Omega)$, and $v_{\mathcal{F}_0}|_F = \Pi_F^k u_D$ for all $F \in \mathcal{F}_0(\Gamma_D)$. On each refinement $\mathcal{T}_{\ell+1}$ of some triangulation \mathcal{T}_ℓ , the initial approximation is defined by a prolongation of the output u_ℓ of the call `fminunc` on the coarse triangulation \mathcal{T}_ℓ . The prolongation maps u_ℓ onto $v_{\ell+1} := \mathbf{I}_{\ell+1} \mathcal{J}_\ell u_\ell \in V(\mathcal{T}_{\ell+1})$.

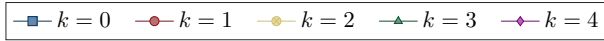


Fig. 1 Polynomial degrees $k = 0, \dots, 4$ in the numerical benchmarks of Sect. 6

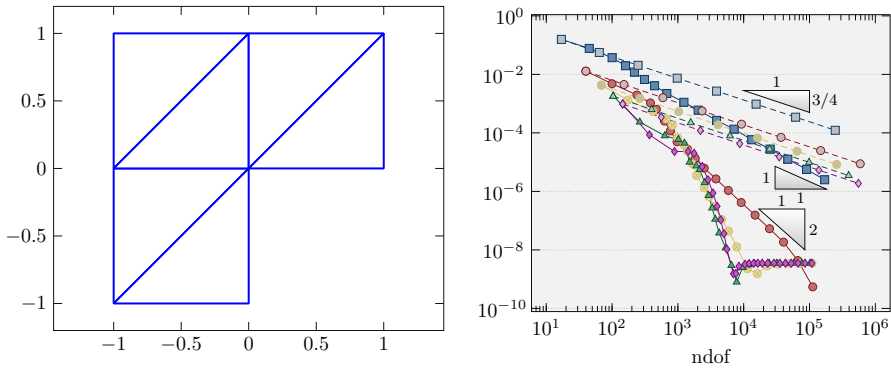


Fig. 2 Initial triangulation \mathcal{T}_0 (left) of the L-shaped domain and convergence history plot (right) of $|E(u) - E_\ell(u_\ell)|$ with k from Fig. 1 on uniform (dashed line) and adaptive (solid line) triangulations for the p -Laplace benchmark in Sect. 6.2

The numerical integration of polynomials is exact with the quadrature formula in [45]: For non-polynomial functions such as $W(\mathcal{G}_\ell v_\ell)$ with $v_\ell \in V(\mathcal{T}_\ell)$, the number of chosen quadrature points allows for exact integration of polynomials of order $p(k + 1)$ with the growth p of W and the polynomial order k of the discretization; the same quadrature formula also applies to the integration of the dual energy density W^* in (6.2). The implementation is based on the in-house AFEM software package in MATLAB [2, 8]. Adaptive computations are carried out with $\theta = 0.5, \varepsilon = (k + 1)/100$, and the polynomial degrees k from Fig. 1. Undisplayed computer experiments suggest only marginal influence of the choice of ε on the convergence rates of the errors.

The uniform or adaptive mesh-refinement leads to convergence history plots of the energy error $|E(u) - E_\ell(u_\ell)|$ or the stress error $\|\sigma - \nabla W(\mathcal{G}_\ell u_\ell)\|_{L^{p'}(\Omega)}^2$ plotted against the number of degrees of freedom (ndof) in Figs. 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11. (Recall the scaling $\text{ndof} \propto h_{\max}^{-2}$ in 2D for uniform mesh refinements with maximal mesh size h_{\max} in a log-log plot.) In the numerical experiments without a priori knowledge of u , the reference value displayed for $\min E(\mathcal{A})$ stems from an Aitken extrapolation of the numerical results for a sequence of uniformly refined triangulations.

6.2 The p -Laplace equation

The third numerical benchmark from [15, Section 6] for the p -Laplace problem in Sect. 2.4.1 considers $p = 4$, the right-hand side

$$f(r, \varphi) := 343/2048 r^{-11/8} \sin(7\varphi/8),$$

on the L-shaped domain $\Omega := (-1, 1)^2 \setminus ([0, 1) \times (-1, 0])$ with the initial triangulation \mathcal{T}_0 displayed in Fig. 2a, the Dirichlet boundary data $u_D(r, \varphi) := r^{7/8} \sin(7\varphi/8)$

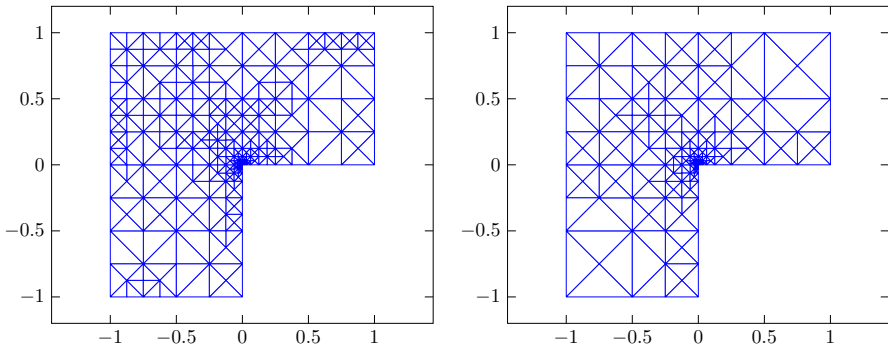


Fig. 3 Adaptive triangulations of the L-shaped domain into 492 triangles (1238 dofs) for $k = 0$ (left) and 490 triangles (7824 dofs) for $k = 3$ (right) for the p -Laplace benchmark in Sect. 6.2

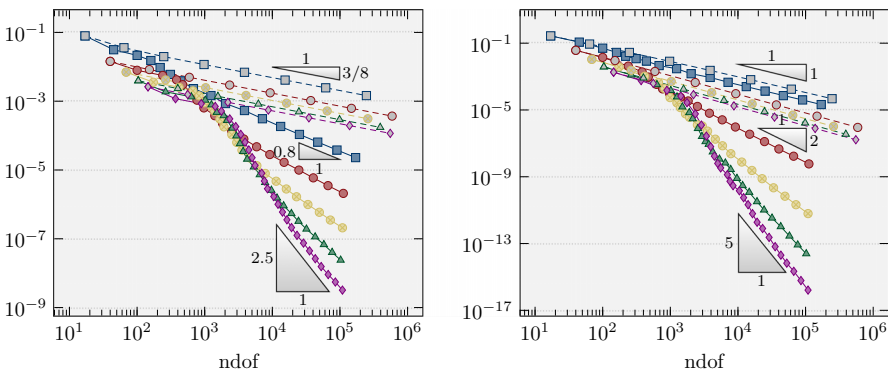


Fig. 4 Convergence history plot of $\|\nabla u - \mathcal{G}_\ell u_\ell\|_{L^4(\Omega)}^2$ (left) and $\|\sigma - \nabla W(\mathcal{G}_\ell u_\ell)\|_{L^{4/3}(\Omega)}^2$ (right) with k from Fig. 1 on uniform (dashed line) and adaptive (solid line) triangulations for the p -Laplace benchmark in Sect. 6.2

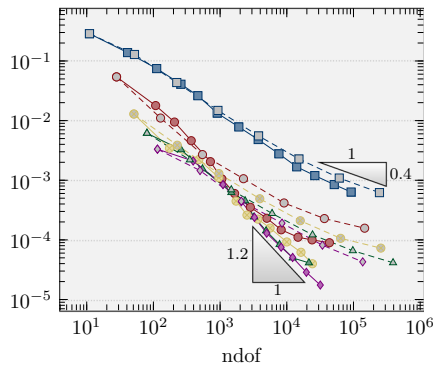
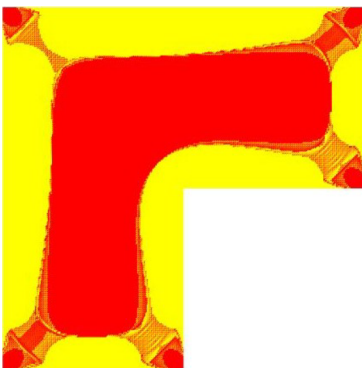


Fig. 5 Material distribution of the L-shaped domain (left) and convergence history plot (right) of RHS_ℓ in (6.1) with k from Fig. 1 on uniform (dashed line) and adaptive (solid line) triangulations for the optimal design problem in Sect. 6.3

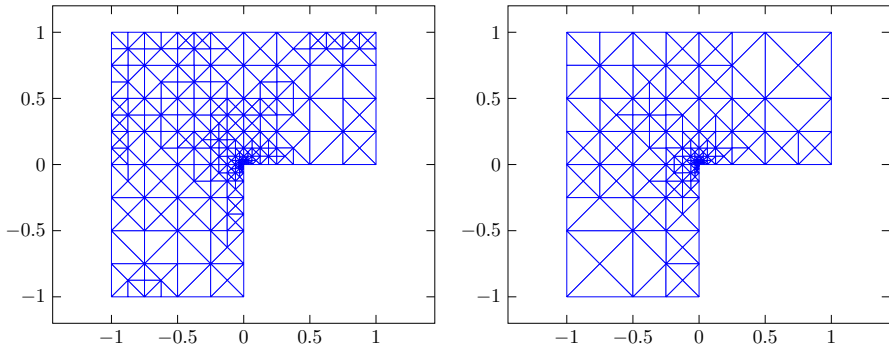


Fig. 6 Adaptive triangulation of the L-shaped domain into 1510 triangles (3721 dofs) for $k = 0$ (left) and 1351 triangles (21,450 dofs) for $k = 3$ (right) for the optimal design problem in Sect. 6.3

$\Gamma_D := (0 \times [-1, 0]) \cup ([0, 1] \times 0)$, and the Neumann boundary data

$$g(r, \varphi) := 343/512 r^{-3/8} (-\sin(\varphi/8), \cos(\varphi/8)) \cdot \nu$$

in polar coordinates with the outer normal unit vector ν on $\Gamma_N := \partial\Omega \setminus \Gamma_D$. The minimal energy $\min E(\mathcal{A}) = -1.4423089582447$ is attained at the unique minimizer

$$u(r, \varphi) := r^{7/8} \sin(7\varphi/8).$$

Since u is singular at the origin, reduced convergence rates are expected for uniform mesh-refining. Figure 2b displays the suboptimal convergence rates 0.75 for the energy error $|E(u) - E_\ell(u_\ell)|$ and all polynomial degrees $k = 0, \dots, 4$. The adaptive mesh-refining algorithm refines towards the origin as depicted in Fig. 3 and we observed a stronger local refinement for larger polynomial degree k . Since W satisfies (2.5)–(2.6), the interest is on the displacement error $\|\nabla u - \mathcal{G}_\ell u_\ell\|_{L^4(\Omega)}$ and the stress error $\|\sigma - \nabla W(\mathcal{G}_\ell u_\ell)\|_{L^{4/3}(\Omega)}$. On uniformly refined meshes, $\|\nabla u - \mathcal{G}_\ell u_\ell\|_{L^4(\Omega)}^2$ converges with the suboptimal convergence rate 0.375 and adaptive computation improves the convergence rate to 0.8 for $k = 0$ and 2.5 for $k = 4$ as depicted in Fig. 4a, b displays the convergence rate 1 for the stress error $\|\sigma - \nabla W(\mathcal{G}_\ell u_\ell)\|_{L^{4/3}(\Omega)}^2$ on uniform triangulations for all $k = 0, \dots, 4$. This is optimal for $k = 0$, but *not* for $k \geq 1$. The adaptive mesh-refining algorithm recovers the optimal convergence rates $k + 1$ for $k \geq 1$.

6.3 The optimal design problem

Consider W from Sect. 2.4.2 for $\mu_1 = 1, \mu_2 = 2, \xi_1 = \sqrt{2\lambda\mu_1/\mu_2}$, and $\xi_2 = \mu_2\xi_1/\mu_1$ with the fixed parameter $\lambda = 0.0145$ on the L-shaped domain $\Omega := (-1, 1)^2 \setminus ([0, 1] \times (-1, 0])$ from [4, Figure 1.1]. Let $f \equiv 1$ in Ω and $u_D \equiv 0$ on $\Gamma_D = \partial\Omega$ with the reference value $\min E(\mathcal{A}) = -0.0745512$.

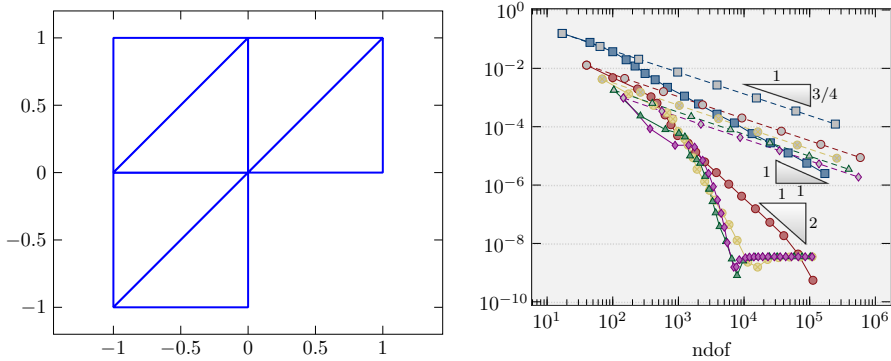


Fig. 7 Initial triangulation (left) of the rectangular domain Ω and convergence history plot (right) of $|E(u) - E_\ell(u_\ell)|$ with k from Fig. 1 on uniform (dashed line) and adaptive (solid line) triangulations for the two-well benchmark in Sect. 6.4

The material distribution in Fig. 5a consists of two homogenous phases, an interior (red) and a boundary (yellow) layer, and a transition layer, also called microstructure zone with a fine mixture of the two materials [4, 16, 25, 28]. The approximated volume fractions $\Lambda(|\Pi_{\mathcal{T}_\ell}^0 \mathcal{G}_\ell u_\ell|)$ for a discrete minimizer u_ℓ with $\Lambda(\xi) = 0$ if $0 \leq \xi \leq \xi_1$, $\Lambda(\xi) = (\xi - \xi_1)/(\xi_2 - \xi_1)$ if $\xi_1 \leq \xi \leq \xi_2$, and $\Lambda(\xi) = 1$ if $\xi \geq \xi_2$, define the colour map of the fraction plot of Fig. 5. Since W satisfies (2.6), Theorem 2.1 implies the convergence of $|E(u) - E_\ell(u_\ell)|$ and $\|\sigma - \nabla W(\mathcal{G}_\ell u_\ell)\|_{L^2(\Omega)}$. Since the exact solution is unknown, the numerical experiment computes RHS_ℓ in

$$\begin{aligned} & \|\sigma - \nabla W(\mathcal{G}_\ell u_\ell)\|_{L^2(\Omega)}^2 + |E(u) - E_\ell(u_\ell)| \\ & \lesssim \text{RHS}_\ell := E_\ell(u_\ell) - E^*(\sigma_\ell) + \text{osc}(f, \mathcal{T}_\ell) + \|\mathcal{G}_\ell u_\ell - \nabla \mathcal{J}_\ell u_\ell\|_{L^2(\Omega)}^2 \end{aligned} \tag{6.1}$$

from [28, Theorem 4.6] with the convex conjugate $W^* \in C(\mathbb{M})$ [53, Corollary 12.2.2] and the dual energy

$$E^*(\sigma_\ell) := - \int_{\Omega} W^*(\sigma_\ell) \, dx. \tag{6.2}$$

Figure 5b displays the suboptimal convergence rate 0.4 for RHS_ℓ on uniform triangulations. The adaptive algorithm refines towards the reentrant corner and the boundaries of the microstructure zone as displayed in Fig. 6. This improves the convergence rates up to 1.2 for $k = 4$. Undisplayed computer experiments show significant improvement for the convergence rates of RHS_ℓ for examples with small microstructure zones in agreement with the related empirical observations in [25].

6.4 The relaxed two-well benchmark

Let $\Omega := (0, 1) \times (0, 3/2)$ with pure Dirichlet boundary $\Gamma_D := \partial\Omega$. The computational benchmark from [15] considers the two distinct wells $F_1 = -(3, 2)/\sqrt{13} = -F_2$

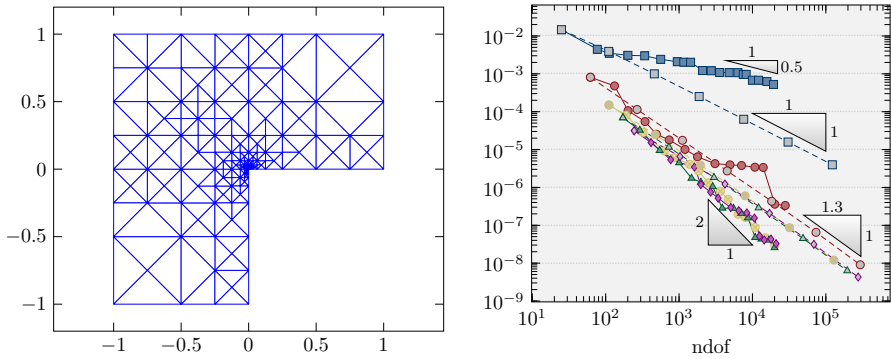


Fig. 8 Adaptive triangulation (left) of the rectangular domain Ω into 1192 triangles (7104 dofs) for $k = 1$ and convergence history plot (right) of $\|u - u_{\mathcal{T}_\ell}\|_{L^2(\Omega)}^2$ with k from Fig. 1 on uniform (dashed line) and adaptive (solid line) triangulations for the two-well benchmark in Sect. 6.4

in the definition of W from Sect. 2.4.3 and introduces an additional quadratic term $\|\zeta - v\|_{L^2(\Omega)}^2$ in the energy

$$E(v) := \int_{\Omega} (W(\nabla v) - fv) \, dx + \|\zeta - v\|_{L^2(\Omega)}^2 / 2$$

for all $v \in \mathcal{A} := u_D + W_0^{1,4}(\Omega)$ with $f(x, y) := -3\varrho^5/128 - \varrho^3/3$, $\zeta(x, y) := \varrho^3/24 + \varrho$,

$$u(x, y) := u_D(x, y) := \begin{cases} f(x, y) & \text{if } -1/2 \leq \varrho \leq 0, \\ \zeta(x, y) & \text{if } 0 \leq \varrho \leq 1/2 \end{cases}$$

at $(x, y) \in \mathbb{R}^2$ and $\varrho := (3(x - 1) + 2y)/\sqrt{13}$. Since E is strictly convex in \mathcal{A} , the minimal energy $\min E(\mathcal{A}) = E(u) = 0.1078147674$ is attained at the unique minimizer u . The discrete minimizer $u_\ell = (u_{\mathcal{T}_\ell}, u_{\mathcal{F}_\ell})$ of the discrete energy

$$E_\ell(v_\ell) := \int_{\Omega} (W(\mathcal{G}_\ell v_\ell) - fv_{\mathcal{T}_\ell}) \, dx + \|\zeta - v_{\mathcal{T}_\ell}\|_{L^2(\Omega)}^2 / 2$$

among $v_\ell = (v_{\mathcal{T}_\ell}, v_{\mathcal{F}_\ell}) \in \mathcal{A}(\mathcal{T}_\ell)$ is unique in the volume component $u_{\mathcal{T}_\ell}$ only. The convergence analysis can be extended to the situation at hand with the refinement indicator $\tilde{\eta}_\ell^{(\varepsilon)}(T) := \eta_\ell^{(\varepsilon)}(T) + |T| \|(1 - \Pi_T^k)\zeta\|_{L^2(T)}^2$ and leads to $\lim_{\ell \rightarrow \infty} E_\ell(u_\ell) = E(u)$, $\lim_{\ell \rightarrow \infty} \nabla W(\mathcal{G}_\ell u_\ell) = \sigma$ (strongly) in $L^{4/3}(\Omega; \mathbb{R}^2)$, and $\lim_{\ell \rightarrow \infty} u_{\mathcal{T}_\ell} = u$ (strongly) in $L^4(\Omega)$.

The exact solution u is piecewise smooth and the derivative ∇u jumps across the interface $\Gamma = \text{conv}\{(1, 0), (0, 3/2)\}$. For an aligned initial triangulation, where Γ coincides with the sides of the triangulation, the numerical results from [28] display optimal convergence rates $k + 1$ for $|E(u) - E_\ell(u_\ell)|$, $\|\sigma - \sigma_\ell\|_{L^{4/3}(\Omega)}^2$, $\|u - u_{\mathcal{T}_\ell}\|_{L^2(\Omega)}^2$, and $\|\nabla u - \mathcal{G}_\ell u_\ell\|_{L^4(\Omega)}^2$ on uniformly refined meshes. Since a priori information on

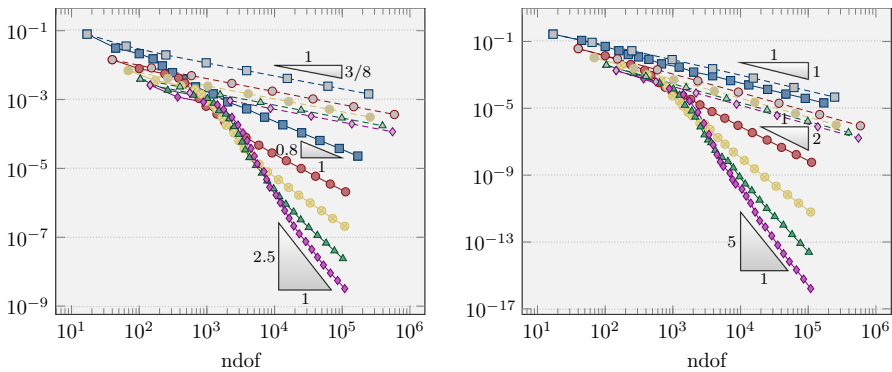


Fig. 9 Convergence history plot of $\|\nabla u - \mathcal{G}_\ell u_\ell\|_{L^4(\Omega)}^2$ (left) and $\|\sigma - \nabla W(\mathcal{G}_\ell u_\ell)\|_{L^{4/3}(\Omega)}^2$ (right) with k from Fig. 1 on uniform (dashed line) and adaptive (right) triangulations for the two-well benchmark in Sect. 6.4

u is not available in general, this numerical benchmark considers the non-aligned initial triangulation \mathcal{T}_0 in Fig. 7a, where Γ cannot be resolved exactly (even not with adaptively refined triangulations of \mathcal{T}_0). In this case, Carstensen and Jochimsen [14] predicted

$$\|(1 - \Pi_{\mathcal{T}_\ell}^0)u\|_{L^4(\Omega)} + \|(1 - \Pi_{\mathcal{T}_\ell}^0)\sigma\|_{L^{4/3}(\Omega)} \lesssim H_\ell, \|(1 - \Pi_{\mathcal{T}_\ell}^0)\nabla u\|_{L^4(\Omega)} \lesssim H_\ell^{1/4}$$

for $H_\ell := \|h_\ell\|_{L^\infty(\Omega)}$. These expected (optimal) convergence rates on uniform meshes are indeed observed empirically for the lowest-order HHO scheme. Figures 7b, 8b and 9 display the convergence rate 1, 1, 1/4, and 1 for $|E(u) - E_\ell(u_\ell)|$, $\|u - u_{\mathcal{T}_\ell}\|_{L^2(\Omega)}^2$, $\|\nabla u - \mathcal{G}_\ell u_\ell\|_{L^4(\Omega)}^2$, and $\|\sigma - \nabla W(\mathcal{G}_\ell u_\ell)\|_{L^{4/3}(\Omega)}^2$, respectively. This improves the convergence rate 3/4 of the stress error from the lowest-order Courant FEM in [14]. The adaptive algorithm generates adaptive meshes with a strong local mesh-refinement near the interface Γ and improve the convergence rate of $|E(u) - E_\ell(u_\ell)|$ to 2.2 in Fig. 7b, of $\|u - u_{\mathcal{T}_\ell}\|_{L^2(\Omega)}^2$ to 2 in Fig. 8b, and of $\|\sigma - \nabla W(\mathcal{G}_\ell u_\ell)\|_{L^{4/3}(\Omega)}^2$ to 2.5 in Fig. 9b for polynomial degrees $k \geq 2$. For $k = 1$, adaptive mesh refinements only leads to marginal improvements. Since optimal convergence rates are obtained for $\|u - u_{\mathcal{T}_\ell}\|_{L^2(\Omega)}$ and $\|\sigma - \nabla W(\mathcal{G}_\ell u_\ell)\|_{L^{4/3}(\Omega)}$ with $k = 0$ on uniform meshes, there is not much gain from adaptive computation.

6.5 Modified Foss–Hrusa–Mizel benchmark

The final example considers a modified Foss–Hrusa–Mizel [43] benchmark in [52], extended to the domain $\Omega := (-1, 1) \times (0, 1)$ with $\Gamma_1 := [-1, 0] \times \{0\}$, $\Gamma_2 := [0, 1] \times \{0\}$, $\Gamma_3 := \{x = (x_1, x_2) \in \partial\Omega : x_1 = -1 \text{ or } x_1 = 1 \text{ or } x_2 = 1\}$, and the initial triangulation \mathcal{T}_0 of Fig. 10a. Define the energy density $W(A) := (|A|^2 - 2 \det A)^4 +$

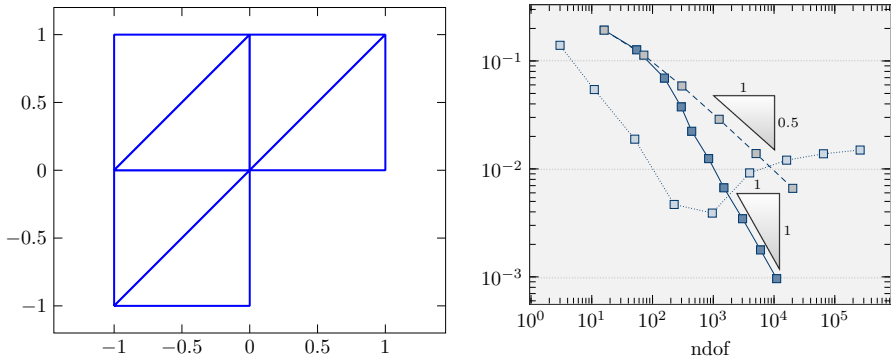


Fig. 10 Initial triangulation T_0 (left) of Ω and empirical verification of the Lavrentiev gap (right) for the modified Foss–Hrusa–Mizel benchmark in Sect. 6.5: convergence history plot of $|E(u) - E_\ell(u_\ell)|$ for the Courant FEM (dotted line) and the lowest-order HHO method on uniform (dashed line) and adaptive (solid line) triangulations

$|A|^2/2$ for all $A \in \mathbb{M} := \mathbb{R}^{2 \times 2}$, the set

$$\mathcal{A} := \{v = (v_1, v_2) \in W^{1,2}(\Omega; \mathbb{R}^2) : v_1 \equiv 0 \text{ on } \Gamma_1, v_2 \equiv 0 \text{ on } \Gamma_2, v = u_D \text{ on } \Gamma_3\}$$

of admissible functions in $W^{1,2}(\Omega; \mathbb{R}^2)$ with $u_D := (\cos(\varphi/2), \sin(\varphi/2))$ in polar coordinates, and the vanishing right-hand side $f \equiv 0$. The minimal energy $E(u) = \min E(\mathcal{A}) = 0.88137023556$ of

$$E(v) := \int_{\Omega} W(Dv) \, dx \text{ among } v \in \mathcal{A}$$

is attained at $u := r^{1/2}(\cos(\varphi/2), \sin(\varphi/2))$ in polar coordinates. The energy density $W \in C^1(\mathbb{M})$ is convex and satisfies the lower growth $W(A) \geq |A|^2/2$ of order $p = 2$, but *no* upper growth of order 2.

The application of the discrete compactness to this model example with free boundary requires the modified refinement indicator

$$\begin{aligned} \eta_\ell^{(\epsilon)}(T) := & |T|^{\epsilon-1} \|\Pi_T^k((\mathcal{R}_\ell u_\ell)|_T - u_T)\|_{L^2(T)}^2 + |T|^{\epsilon-1/2} \\ & \times \left(\sum_{F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Gamma_1)} \|(\mathcal{R}_\ell u_\ell)|_F \cdot e_1\|_{L^2(F)}^2 + \sum_{F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Gamma_2)} \|(\mathcal{R}_\ell u_\ell)|_F \cdot e_2\|_{L^2(F)}^2 \right. \\ & + \sum_{F \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Gamma_3)} \|(\mathcal{R}_\ell u_\ell)|_F - u_D\|_{L^2(F)}^2 + \sum_{E \in \mathcal{F}_\ell(T) \cap \mathcal{F}_\ell(\Omega)} \|[\mathcal{R}_\ell u_\ell]_E\|_{L^2(E)}^2 \\ & \left. + \sum_{F \in \mathcal{F}_\ell(T)} \|\Pi_F((\mathcal{R}_\ell u_\ell)_T - u_F)\|_{L^2(F)}^2 \right) \end{aligned}$$

with the j -th canonical unit vector $e_j \in \mathbb{R}^2$. Since the presence of the Lavrentiev gap is equivalent to the failure of conforming FEMs [17, Theorem 2.1], the lowest-order HHO can be utilized to detect the Lavrentiev gap, cf. Sect. 4.3. Figure 10b provides

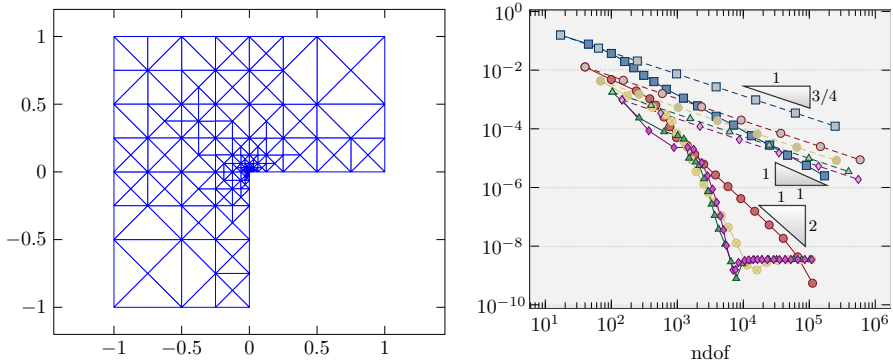


Fig. 11 Adaptive triangulation of Ω (left) into 614 triangles (7336 dofs) for $k = 1$ and convergence history plot (right) of $|E(u) - E_\ell(u_\ell)|$ with k from Fig. 1 on uniform (dashed line) and adaptive (solid line) triangulations for the modified Foss–Hrusa–Mizel benchmark in Sect. 6.5

empirical evidence that there is a Lavrentiev gap: $|E(u) - E_\ell(u_\ell)|$ converges with the suboptimal convergence rate 0.5 on uniformly refined meshes, but the Courant FEM seems to approximate a wrong energy. The adaptive mesh-refining algorithm refines towards the origin as depicted in Fig. 11a. It is outlined in Sect. 4.3 that a convergence proof of AHHO for minimization problems with the Lavrentiev gap is impossible with the known mathematical methodology for $k \geq 1$. It comes as a welcome surprise that optimal convergence rates $k + 1$ are obtained for any polynomial degrees k on adaptively refined meshes in Fig. 11b.

6.6 Conclusions

The numerical results from Sect. 6 confirm the theoretical findings in Theorem 2.1. In particular, the convergence of the energy $\lim_{\ell \rightarrow \infty} \min E_\ell(\mathcal{A}(\mathcal{T}_\ell)) = \min E(\mathcal{A})$ is observed in all examples. The introduced adaptive mesh-refining algorithm of Sect. 2.2 provides efficient approximations of singular solutions and even leads to improved empirical convergence rates. The choice of the parameter ε only has marginal influence on the convergence rates and convergence is observed for $\varepsilon = 0$ in undisplayed computer experiments. Better convergence rates are obtained for larger polynomial degrees k . The computer experiments provide empirical evidence that the HHO method can overcome the Lavrentiev gap for any polynomial degree k .

Acknowledgements This work has been supported by the Deutsche Forschungsgemeinschaft (DFG) in the Priority Program 1748 Reliable simulation techniques in solid mechanics: Development of non-standard discretization methods, mechanical and mathematical analysis under the project CA 151/22. The second author is also supported by the ERC through the Starting Grant DAFNE, agreement ID 891734.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence,

and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Abbas, M., Ern, A., Pignet, N.: Hybrid high-order methods for finite deformations of hyperelastic materials. *Comput. Mech.* **62**(4), 909–928 (2018)
2. Albery, J., Carstensen, C., Funken, S.A.: Remarks around 50 lines of Matlab: short finite element implementation. *Numer. Algorithms* **20**(2–3), 117–137 (1999)
3. Balci, A.K., Ortner, C., Storn, J.: Crouzeix–Raviart finite element method for non-autonomous variational problems with Lavrentiev gap. [arXiv:2106.06837](https://arxiv.org/abs/2106.06837) (2021)
4. Bartels, S., Carstensen, C.: A convergent adaptive finite element method for an optimal design problem. *Numer. Math.* **108**(3), 359–385 (2008)
5. Belenki, L., Diening, L., Kreuzer, C.: Optimality of an adaptive finite element method for the p-Laplacian equation. *IMA J. Numer. Anal.* **32**(2), 484–510 (2012)
6. Boffi, D., et al.: Optimal convergence of adaptive FEM for eigenvalue clusters in mixed form. *Math. Comput.* **86**(307), 2213–2237 (2017)
7. Bonito, A., Nochetto, R.H.: Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method. *SIAM J. Numer. Anal.* **48**(2), 734–771 (2010)
8. Brenner, S.C., Carstensen, C.: Finite element methods. In: Deconinck, H., Ricchiuto, M. (eds.) *Encyclopedia of Computational Mechanics*, 2nd edn., pp. 1–47. American Cancer Society, Atlanta (2017)
9. Brenner, S.C., Scott, L.R.: *The Mathematical Theory of Finite Element Methods*. Texts in Applied Mathematics, vol. 15. Springer, New York (2008)
10. Brezis, H.: *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York (2011)
11. Buffa, A., Ortner, C.: Compact embeddings of broken Sobolev spaces and applications. *IMA J. Numer. Anal.* **29**(4), 827–855 (2009)
12. Carstensen, C., Dolzmann, G.: Convergence of adaptive finite element methods for a nonconvex double-well minimization problem. *Math. Comput.* **84**, 2111–2135 (2015)
13. Carstensen, C., Dolzmann, G.: Convergence of adaptive finite element methods for a nonconvex double-well minimization problem. *Math. Comput.* **84**(295), 2111–2135 (2015)
14. Carstensen, C., Jochimsen, K.: Adaptive finite element methods for microstructures? Numerical experiments for a 2-well benchmark. *Computing* **71**(2), 175–204 (2003)
15. Carstensen, C., Klose, R.: A posteriori finite element error control for the p-Laplace problem. *SIAM J. Sci. Comput.* **25**, 792–814 (2003)
16. Carstensen, C., Liu, D.J.: Nonconforming FEMs for an optimal design problem. *SIAM J. Numer. Anal.* **53**(2), 874–894 (2015)
17. Carstensen, C., Ortner, C.: Analysis of a class of penalty methods for computing singular minimizers. *Comput. Methods Appl. Math.* **10**(2), 137–163 (2010)
18. Carstensen, C., Puttkammer, S.: How to prove the discrete reliability for nonconforming finite element methods. *J. Comput. Math.* **38**(1), 142–175 (2020)
19. Carstensen, C., Rabus, H.: Axioms of adaptivity with separate marking for data resolution. *SIAM J. Numer. Anal.* **55**(6), 2644–2665 (2017)
20. Carstensen, C., et al.: A review of unified a posteriori finite element error control. *Numer. Math. Theory Methods Appl.* **5**(4), 509–558 (2012)
21. Carstensen, C., et al.: Axioms of adaptivity. *Comput. Math. Appl.* **67**(6), 1195–1253 (2014)
22. Carstensen, C.: Convergence of an adaptive FEM for a class of degenerate convex minimization problems. *IMA J. Numer. Anal.* **28**(3), 423–439 (2008)
23. Carstensen, C., Gallistl, D., Schedensack, M.: Adaptive nonconforming Crouzeix–Raviart FEM for eigenvalue problems. *Math. Comput.* **84**(293), 1061–1087 (2015)
24. Carstensen, C., Gedicke, J.: An adaptive finite element eigenvalue solver of asymptotic quasi-optimal computational complexity. *SIAM J. Numer. Anal.* **50**(3), 1029–1057 (2012)

25. Carstensen, C., Günther, D., Rabus, H.: Mixed finite element method for a degenerate convex variational problem from topology optimization. *SIAM J. Numer. Anal.* **50**(2), 522–543 (2012)
26. Carstensen, C., Nataraj, N.: A priori and a posteriori error analysis of the Crouzeix–Raviart and Morley FEM with original and modified right-hand sides. *Comput. Methods Appl. Math.* **21**(2), 289–315 (2021)
27. Carstensen, C., Plecháč, P.: Numerical solution of the scalar double-well problem allowing microstructure. *Math. Comput.* **66**(219), 997–1026 (1997)
28. Carstensen, C., Tran, T.: Unstabilized hybrid high-order method for a class of degenerate convex minimization problems. *SIAM J. Numer. Anal.* **59**(3), 1348–1373 (2021)
29. Cascon, J.M., et al.: Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.* **46**(5), 2524–2550 (2008)
30. Chipot, M., Collins, C.: Numerical approximations in variational problems with potential wells. *SIAM J. Numer. Anal.* **29**(4), 1002–1019 (1992)
31. Crouzeix, M., Raviart, P.-A.: Conforming and nonconforming finite element methods for solving the stationary Stokes equations I. *RAIRO Sér. Rouge* **7**(R–3), 33–75 (1973)
32. Dacorogna, B.: *Direct Methods in the Calculus of Variations*. Applied Mathematical Sciences, vol. 78, 2nd edn. Springer, New York (2008)
33. Dai, X., Xu, J., Zhou, A.: Convergence and optimal complexity of adaptive finite element eigenvalue computations. *Numer. Math.* **110**(3), 313–355 (2008)
34. Di Pietro, D.A., Droniou, J.: A hybrid high-order method for Leray–Lions elliptic equations on general meshes. *Math. Comput.* **86**(307), 2159–2191 (2017)
35. Di Pietro, D.A., Ern, A.: A hybrid high-order locking-free method for linear elasticity on general meshes. *Comput. Methods Appl. Mech. Eng.* **283**, 1–21 (2015)
36. Di Pietro, D.A., Ern, A.: Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible Navier–Stokes equations. *Math. Comput.* **79**(271), 1303–1330 (2010)
37. Di Pietro, D.A., Ern, A., Lemaire, S.: An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators. *Comput. Methods Appl. Math.* **14**(4), 461–472 (2014)
38. Di Pietro, D.A., Specogna, R.: An a posteriori-driven adaptive mixed high-order method with application to electrostatics. *J. Comput. Phys.* **326**, 35–55 (2016)
39. Di Pietro, D. A., Droniou, J.: The hybrid high-order method for polytopal meshes. In: *MS& A. Modeling, Simulation and Applications. Design, Analysis, and Applications*, vol. 19, p. 525. Springer, Cham (2020)
40. Di Pietro, D.A., Ern, A.: *Mathematical aspects of discontinuous Galerkin methods*, vol. 69. *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg (2012)
41. Diening, L., Kreuzer, C.: Linear convergence of an adaptive finite element method for the p-Laplacian equation. *SIAM J. Numer. Anal.* **46**(2), 614–638 (2008)
42. Ern, A., Zanotti, P.: A quasi-optimal variant of the hybrid high-order method for elliptic partial differential equations with H-1 loads. *IMA J. Numer. Anal.* **40**(4), 2163–2188 (2020)
43. Foss, M., Hrusa, W.J., Mizel, V.J.: The Lavrentiev gap phenomenon in nonlinear elasticity. *Arch. Ration. Mech. Anal.* **167**(4), 337–365 (2003)
44. Glowinski, R., Marrocco, A.: Sur l’approximation, par éléments finis d’ordre un, et la résolution, par pénalisation-dualité, d’une classe de problèmes de Dirichlet non linéaires. *RAIRO Sér. Rouge Anal. Numér.* **9**(R–2), 41–76 (1975)
45. Hammer, P.C., Marlowe, O.J., Stroud, A.H.: Numerical integration over simplexes and cones. *Math. Tables Aids Comput.* **10**, 130–137 (1956)
46. Kohn, R.V., Strang, G.: Optimal design and relaxation of variational problems. I. *Commun. Pure Appl. Math.* **39**(1), 113–137 (1986)
47. Kreuzer, C., Georgoulis, E.H.: Corrigendum to “Convergence of adaptive, discontinuous Galerkin methods”. *Math. Comput.* **90**(328), 637–640 (2021)
48. Lavrentieff, M.: Sur quelques problèmes du calcul des variations. *Ann. Mat. Pura Appl.* **4**(1), 7–28 (1927)
49. Morin, P., Siebert, K.G., Veeger, A.: A basic convergence result for conforming adaptive finite elements. *Math. Models Methods Appl. Sci.* **18**(5), 707–737 (2008)
50. Nochetto, R.H., Veeger, A.: “Primer of Adaptive Finite Element Methods”. *Multiscale and Adaptive-ity: Modeling, Numerics and Applications. Lecture Notes in Mathematics*, vol. 2040, pp. 125–225. Springer, Heidelberg (20120)

51. Ortner, C.: Nonconforming finite-element discretization of convex variational problems. *IMA J. Numer. Anal.* **31**(3), 847–864 (2011)
52. Ortner, C., Praetorius, D.: On the convergence of adaptive nonconforming finite element methods for a class of convex variational problems. *SIAM J. Numer. Anal.* **49**(1), 346–367 (2011)
53. Rockafellar, R.T.: *Convex Analysis* Princeton Mathematical Series, No. 28. Princeton University Press, Princeton (1970)
54. Stevenson, R.: Optimality of a standard adaptive finite element method. *Found. Comput. Math.* **7**(2), 245–269 (2007)
55. Stevenson, R.: The completion of locally refined simplicial partitions created by bisection. *Math. Comput.* **77**(261), 227–241 (2008)
56. Veese, A.: Convergent adaptive finite elements for the nonlinear Laplacian. *Numer. Math.* **92**(4), 743–770 (2002)
57. Veese, A., Zanotti, P.: Quasi-optimal nonconforming methods for symmetric elliptic problems. II-Overconsistency and classical nonconforming elements. *SIAM J. Numer. Anal.* **57**(1), 266–292 (2019)
58. Verfürth, R.: *A Posteriori Error Estimation Techniques for Finite Element Methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford (2013)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.