Numer. Math. (2017) 136:1097–1115
DOI 10.1007/s00211-017-0866-x

Numerische Mathematik

Convergence of natural adaptive least squares finite element methods

Carsten Carstensen[1] · **Eun-Jae Park**[2] · **Philipp Bringmann**[1]

Received: 23 December 2013 / Revised: 30 March 2016 / Published online: 23 February 2017

Abstract The first-order div least squares finite element methods provide inherent a posteriori error estimator by the elementwise evaluation of the functional. In this paper we prove Q-linear convergence of the associated adaptive mesh-refining strategy for a sufficiently fine initial mesh with some sufficiently large bulk parameter for piecewise constant right-hand sides in a Poisson model problem. The proof relies on some modification of known supercloseness results to non-convex polygonal domains plus the flux representation formula. The analysis is carried out for the lowest-order case in two-dimensions for the simplicity of the presentation.

Mathematics Subject Classification 65N12 · 65N15 · 65N30 · 65N50 · 65Y20

1 Introduction

The mathematical theory of the adaptive finite element method (AFEM) has been developed significantly over the past decade. In particular, the adaptive mesh-refining

This research was supported by the Deutsche Forschungsgemeinschaft in the Priority Program 1748 "Reliable simulation techniques in solid mechanics. Development of non- standard discretization methods, mechanical and mathematical analysis" under the project "Foundation and application of generalized mixed FEM towards nonlinear problems in solid mechanics" (CA 151/22-1). This research was supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology NRF 2011-0030934 and NRF-2015R1A5A1009350.

✉ Eun-Jae Park
ejpark@yonsei.ac.kr

1 Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

2 Department of Computational Science and Engineering, Yonsei University, Seoul 03722, Korea

method has been understood to converge with optimal convergence rates with respect to the concept of a nonlinear approximations class [4,19,21,28]. Although optimal convergence rates are often observed in many numerical experiments [1,3,27], even the plain convergence is not understood for the adaptive least squares finite element method (ALSFEM).

This paper analyses the convergence of natural adaptive mesh-refining first-order div least squares finite element methods. The adaptive scheme monitors the local contributions of the least squares residual and converges for a sufficiently fine initial mesh with some large bulk parameter for piecewise constant right-hand sides.

The reliable and efficient error control of the first-order div least squares finite element method (LSFEM) [6] for a Poisson model problem (PMP) with the homogeneous Dirichlet boundary condition is immediately available by the least squares functional

$$LS(f; \mathbf{p}_\ell, u_\ell) := \|f + \operatorname{div} \mathbf{p}_\ell\|^2 + \|\mathbf{p}_\ell - \nabla u_\ell\|^2$$

with the L^2 norm $\|\cdot\| := \|\cdot\|_{L^2(\Omega)}$ evaluated for the discrete approximations $(\mathbf{p}_\ell, u_\ell)$ in some Raviart–Thomas and Courant finite element subspaces of the Sobolev spaces $H(\operatorname{div}; \Omega) \times H_0^1(\Omega)$ with respect to a triangulation $\mathcal{T}_\ell$ of the polygonal domain Ω.

It is expected that the elementwise evaluation of the least squares functional leads to an effective ALSFEM [1,3,27]. One difficulty in the convergence analysis of those schemes is the question whether the least squares residual is indeed strictly reduced provided the mesh is refined [1].

The main contribution of this paper is a first convergence proof for this natural strategy with the contraction property of the modified least squares functional

$$\widehat{LS}_\ell := LS(f; \mathbf{p}_\ell, u_\ell) + \Lambda_1 \|(\bullet - \operatorname{mid}(\mathcal{T}_\ell)) \operatorname{div} \mathbf{p}_\ell\|^2 \tag{1.1}$$

for some appropriate constant $0 < \Lambda_1 < \infty$ and the additional divergence term with the piecewise affine pre-factor $\bullet - \operatorname{mid}(\mathcal{T}_\ell) \in P_1(\mathcal{T}_\ell; \mathbb{R}^2)$ as a weight equal to $x - \operatorname{mid}(T)$ at $x \in T \in \mathcal{T}_\ell$ in the triangle T with centre of inertia $\operatorname{mid}(T)$. Saturation holds in the sense that there exists some $0 < \varrho_1 < 1$ with

$$\widehat{LS}_{\ell+1} \le \varrho_1 \widehat{LS}_\ell \quad \text{for all } \ell = 0, 1, 2, \ldots \tag{1.2}$$

Let Π_ℓ denote the L^2 orthogonal projection onto the piecewise constants $P_0(\mathcal{T}_\ell)$ (or any power like $P_0(\mathcal{T}_\ell; \mathbb{R}^2)$) to illustrate the difference of LS_ℓ and $\widehat{LS}_\ell$ in the sequel. The lowest-order Raviart–Thomas finite element functions and piecewise orthogonal splits guarantee that

$$||f - \Pi_\ell f||^2 + ||\Pi_\ell f + \operatorname{div} \mathbf{p}_\ell||^2 + \lambda ||(1 - \Pi_\ell)\mathbf{p}_\ell||^2 + ||\Pi_\ell \mathbf{p}_\ell - \nabla u_\ell||^2$$

equals $LS_\ell := LS(f; \mathbf{p}_\ell, u_\ell)$ for $\lambda := 1$ (and $\widehat{LS}_\ell$ for $\lambda := 1 + 4\Lambda_1$). In other words, LS_ℓ and $\widehat{LS}_\ell$ differ solely in the weights of the preceding four contributions. In this sense, (1.2) may be seen as a saturation for the (slightly modified) least squares functional.

The affirmative mathematical analysis is performed under the following three restrictions (i)–(iii) to ensure (1.2). (i) The initial triangulation $\mathcal{T}_0$ is required to be sufficiently fine; (ii) an ad-hoc version of adaptive mesh-refining does *not* resolve the data properly and then (1.2) holds exclusively for a piecewise constant source f; (iii) the sequence of shape-regular triangulations is generated with some bulk parameter $0 < \Theta_1 < 1$ sufficiently close to 1.

Some remarks are in order regarding those restrictions. The fineness condition (i) on the initial mesh was first used in [21]. The following counter-example illustrates the severe difficulty (ii). It gives the warning that, in general, overall refinement does *not* lead to strict reduction of the least squares functional. Suppose that $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}}$ is a sequence of uniform mesh-refinements (e.g. with bisec5 depicted in Fig. 1) and pick a natural number, say, $k = 100$ and a source term f in $L^2(\Omega)$ as right-hand side in the PMP with piecewise integral mean zero with respect to the mesh $\mathcal{T}_k$, written $f_k \equiv 0$, but with a non-zero integral mean $f_{k+1} \not\equiv 0$ with respect to the mesh $\mathcal{T}_{k+1}$. Then, the respective discrete solutions $(\mathbf{p}_\ell, u_\ell)$ of the LSFEM vanish for level $\ell = 0, \ldots, k$ while $(\mathbf{p}_{k+1}, u_{k+1})$ does not. Consequently, the sequence $LS_\ell := LS(f; \mathbf{p}_\ell, u_\ell)$ of the minimal least squares functionals satisfies

$$LS_{k+1} < LS_k = LS_{k-1} = \cdots = LS_0 = ||f||^2. \tag{1.3}$$

It is clear that the sequence of the least squares functionals is monotone decreasing, but the convergence may *not* be strict. An example of this type can be constructed for standard Galerkin methods as well. But those methods are accompanied by a residual-based error estimate $\|h_\ell f\|$ with a mesh-size h_ℓ in front of the right-hand side $f \in L^2(\Omega)$, which is reduced. The difference is that, here, some refinement on the level $\ell < k = 100$ does neither reduce the error nor the aforementioned equivalent error measures.

The condition (iii) on the bulk parameter $0 < \Theta_1 < 1$ sufficiently close to 1 contradicts the discrete reliability in the sense of Stevenson [28] which is key to the proof of optimal convergence rates: All known optimality results follow [28] and require the bulk parameter to be sufficiently small! As a consequence, the authors propose an alternative error analysis with explicit residual-based error estimates in [15] which

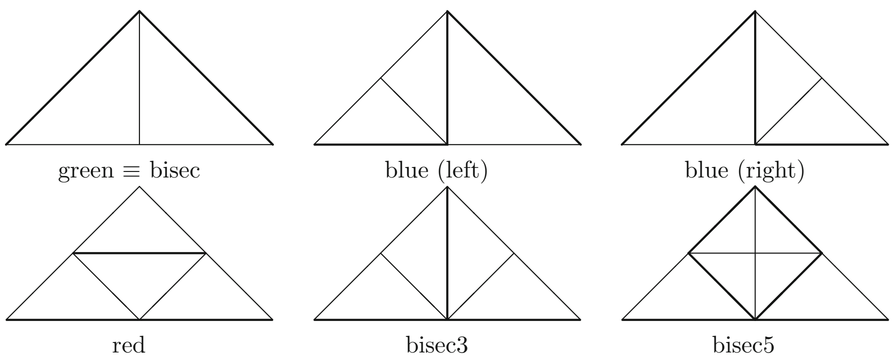

Fig. 1 Possible refinements of a triangle

then allows the arguments of [12,28] to guarantee optimal convergence rates under the extra assumption of exact solve. Nevertheless, the natural adaptive mesh-refining algorithm of the present paper is employed in practice without further understanding. A similar situation is encountered in the context of saturation of conforming finite element methods and hierarchical error estimators in [22].

The remaining parts of this paper are organised as follows. Section 2 introduces the PMP and its least squares discretisation with Raviart–Thomas-type flux approximations. The analysis is exploited for polygonal domains in two space dimensions while the generalisation to three dimensions is incremental. The proof of the saturation property (1.2) relies on some generalisation of the supercloseness results of [8] to the non-smooth case of non-convex polygonal domains in Sect. 3. The reduction factor in saturation depends on the maximal mesh-size to some power $0 < s < \pi/\omega$ from reduced elliptic regularity of the PMP of Sect. 2 in the non-convex polygonal domain Ω with maximal interior angle ω.

Section 4 presents some natural ALSFEM with marking based on the elementwise contributions of the least squares functional and proves the saturation property (1.2) for large bulk parameter Θ and fine initial meshes $\mathcal{T}_0$. Section 5 presents numerical experiments for the investigation of the choice of the bulk parameter Θ.

Standard notation on Lebesgue and Sobolev spaces and norms is employed throughout this paper: $\|\cdot\|$ denotes the L^2 norm and $|||\cdot|||$ denotes the H^1 seminorm over the entire domain Ω, while $|||\cdot|||_{NC} := \|\nabla_{NC}\cdot\|$ is some piecewise version thereof. Finally, $a \lesssim b$ denotes $a \le c\, b$ with some generic constant c which may depend on the domain and the initial coarse mesh $\mathcal{T}_0$ but which is independent of the level ℓ or the mesh-size $H_\ell = \max\{h_T : T \in \mathcal{T}_\ell\}$ which is the maximal piecewise mesh-size $h_\ell \in L^\infty(\Omega)$ defined by $h_\ell|_T := h_T = |T|^{1/2}$ for the area $|T|$ of a triangle $T \in \mathcal{T}_\ell$. Similarly, $a \approx b$ abbreviates $a \lesssim b \lesssim a$.

The measure $|\cdot|$ is context-sensitive and refers to the number of elements of some finite set (e.g. the number $|\mathcal{T}|$ of triangles in a triangulation $\mathcal{T}$) or the length $|E|$ of an edge E or the area $|T|$ of some domain T and not just the modulus of a real number or the Euclidean length of a vector.

It is expected that the results can be generalized to higher-order FEM in 3D as well despite the severe difficulties that nonconforming FEMs are not available in 3D for all polynomial degrees.

2 Poisson model problem (PMP) and its least squares discretisation

Given $f \in L^2(\Omega)$ on a simply-connected bounded polygonal Lipschitz domain $\Omega \subset \mathbb{R}^2$, the PMP seeks some function $u \in C_0(\overline{\Omega}) \cap H^2_{\mathrm{loc}}(\Omega)$ with

$$-\Delta u = f \text{ in } \Omega \quad \text{and} \quad u = 0 \text{ on } \partial\Omega. \tag{2.1}$$

The least squares methods consider the equivalent first-order system

$$-\operatorname{div} \mathbf{p} = f \quad \text{and} \quad \mathbf{p} - \nabla u = \mathbf{0} \quad \text{in } \Omega. \tag{2.2}$$

The weak form involves the L^2 inner product $(\cdot,\cdot)_{L^2(\Omega)}$ and its L^2 norm $\|\cdot\|$ over Ω. Standard notation is employed for the Sobolev space $H^1(\Omega)$ with seminorm $|||\cdot|||$ and $V := H^1_0(\Omega)$. The Hilbert space

$$H(\mathrm{div};\Omega) = \{\mathbf{q} \in L^2(\Omega;\mathbb{R}^2) : \mathrm{div}\,\mathbf{q} \in L^2(\Omega)\}$$

consists of all L^2 vector functions $\mathbf{q} = (q_1, q_2)$ with weak divergence $\mathrm{div}\,\mathbf{q} := \partial_{x_1} q_1 + \partial_{x_2} q_2$ in $L^2(\Omega)$ and associated norm $\|\cdot\|_{H(\mathrm{div})}$ [7,10,11,23]. The least squares method solves system (2.2) by minimising the residual functional, for $(\mathbf{q}, v) \in H(\mathrm{div};\Omega) \times V$,

$$LS(f;\mathbf{q},v) := \|f + \mathrm{div}\,\mathbf{q}\|^2 + \|\mathbf{q} - \nabla v\|^2. \tag{2.3}$$

The associated Euler–Lagrange equations lead to the equivalent weak problem: Seek $(\mathbf{p}, u) \in H(\mathrm{div};\Omega) \times V$ such that, for all $(\mathbf{q}, v) \in H(\mathrm{div};\Omega) \times V$,

$$\int_\Omega (f + \mathrm{div}\,\mathbf{p})\,\mathrm{div}\,\mathbf{q}\,dx + \int_\Omega (\mathbf{p} - \nabla u)\cdot(\mathbf{q} - \nabla v)\,dx = 0. \tag{2.4}$$

The well-established equivalence of the norm in $H(\mathrm{div};\Omega) \times V$ with the least squares functional

$$LS(0;\mathbf{q},v) \approx \|\mathbf{q}\|^2_{H(\mathrm{div})} + |||v|||^2 \quad \text{for all } (\mathbf{q}, v) \in H(\mathrm{div};\Omega) \times V \tag{2.5}$$

leads to the unique existence of a minimiser of $LS(f;\cdot)$ and weak solution $(\mathbf{p}, u) \in H(\mathrm{div};\Omega) \times V$ [6]. Moreover, the conforming discretisation leads to a quasi-optimal convergence.

The prototype example for a discretisation is the lowest-order Raviart–Thomas function space $RT_0(\mathcal{T})$ based on a regular triangulation $\mathcal{T}$ of Ω in closed triangles in the sense of Ciarlet [10,20], i.e., $\cup\mathcal{T} = \overline{\Omega}$ and any two distinct triangles in $\mathcal{T}$ are either disjoint or share exactly one vertex or one common edge. Given any regular triangulation $\mathcal{T}$, let

$$\begin{aligned} V(\mathcal{T}) &:= P_1(\mathcal{T}) \cap V,\\ RT_0(\mathcal{T}) &:= \{\mathbf{q} \in P_1(\mathcal{T};\mathbb{R}^2) \cap H(\mathrm{div};\Omega) : \forall T \in \mathcal{T}, \exists a_T, b_T, c_T \in \mathbb{R},\\ &\qquad \forall x \in T,\ \mathbf{q}(x) = (a_T, b_T)^\top + c_T\, x\}. \end{aligned}$$

There exists a unique minimiser $(\mathbf{p}_{LS}, u_{LS})$ of $LS(f;\cdot)$ in $RT_0(\mathcal{T}) \times V(\mathcal{T})$ and this is characterised as the weak solution of the discrete analog (2.6) of (2.4). In other words, the LSFEM solution $(\mathbf{p}_{LS}, u_{LS}) \in RT_0(\mathcal{T}) \times V(\mathcal{T}) \subset H(\mathrm{div};\Omega) \times V$ satisfies, for all $(\mathbf{q}_{RT}, v_C) \in RT_0(\mathcal{T}) \times V(\mathcal{T}) \subset H(\mathrm{div};\Omega) \times V$, that

$$\int_\Omega (f + \mathrm{div}\,\mathbf{p}_{LS})\,\mathrm{div}\,\mathbf{q}_{RT}\,dx + \int_\Omega (\mathbf{p}_{LS} - \nabla u_{LS})\cdot(\mathbf{q}_{RT} - \nabla v_C)\,dx = 0. \tag{2.6}$$

The Céa lemma leads to the best approximation property

$$\|\mathbf{p}-\mathbf{p}_{LS}\|_{H(\mathrm{div})}+|||u-u_{LS}|||\lesssim \min_{\mathbf{q}_{RT}\in RT_0(\mathcal{T})}\|\mathbf{p}-\mathbf{q}_{RT}\|_{H(\mathrm{div})}+\min_{v_C\in V(\mathcal{T})}|||u-v_C|||.$$

Provided the exact solution u belongs to $H^2(\Omega)$ (e.g. for a convex domain Ω), standard approximation results lead to linear convergence in the maximal mesh-size. However, in case of reduced elliptic regularity (e.g. for a non-convex domain Ω), appropriate mesh-refining strategies are required to avoid suboptimal convergence rates for less regular problems.

This section concludes with some representation result which is frequently employed throughout this paper. Denote by Π_0 the L^2 orthogonal projection onto the piecewise constants $P_0(\mathcal{T};\mathbb{R}^m)$ for $m=1,2$ with respect to the present triangulation $\mathcal{T}$. Let $CR_0^1(\mathcal{T})$ denote the functions in $P_1(\mathcal{T})$ which are continuous at the midpoints of all interior edges $\mathcal{E}(\Omega)$ and vanish at the midpoints of all boundary edges $\mathcal{E}(\partial\Omega)$. Let ∇_{NC} denote the piecewise action of the gradient.

Proposition 2.1 *Any Raviart–Thomas function* $\mathbf{q}_{RT}\in RT_0(\mathcal{T})$ *reads*

$$\mathbf{q}_{RT}=\Pi_0\mathbf{q}_{RT}+(\bullet-\mathrm{mid}(\mathcal{T}))\frac{\mathrm{div}\,\mathbf{q}_{RT}}{2}\quad \textit{a.e. in } \Omega \tag{2.7}$$

(*where* $\bullet-\mathrm{mid}(\mathcal{T})$ *abbreviates* $x-\mathrm{mid}(T)$ *at any* $x\in T\in\mathcal{T}$ *with centre of inertia* $\mathrm{mid}(T)$) *and satisfies, for unique* $v_{CR}\in CR_0^1(\mathcal{T})$ *and* $w_C\in V(\mathcal{T})/\mathbb{R}$, *that*

$$\Pi_0\mathbf{q}_{RT}=\nabla_{NC}v_{CR}+\mathrm{Curl}\,w_C. \tag{2.8}$$

Therein, $v_{CR}\in CR_0^1(\mathcal{T})$ *is the Crouzeix–Raviart solution of the PMP with right-hand side* $-\mathrm{div}\,\mathbf{q}_{RT}\in L^2(\Omega)$*, i.e.,* v_{CR} *solves the nonconforming finite element problem, hereafter referred to as NCFEM,*

$$\int_\Omega \nabla_{NC}v_{CR}\cdot\nabla_{NC}w_{CR}\,dx=-\int_\Omega w_{CR}\,\mathrm{div}\,\mathbf{q}_{RT}\,dx\quad \textit{for all } w_{CR}\in CR_0^1(\mathcal{T}). \tag{2.9}$$

Moreover, for any discrete solution $\mathbf{q}_{RT}$ *of a mixed finite element problem or any LSFEM solution* $\mathbf{q}_{RT}:=\mathbf{p}_{LS}$ *of* (2.6)*,* $w_C\equiv 0$ *holds in* (2.8)*. In other words, those particular Raviart–Thomas fluxes are* L^2 *orthogonal onto* $\mathrm{Curl}(V(\mathcal{T}))$.

Proof The identities (2.7)–(2.9) are proven in [24] but essentially known since [2]. The formula (2.7) follows from the very definition of the Raviart–Thomas functions. The formula (2.8) is a discrete Helmholtz decomposition for simply-connected domains of any piecewise constant vector field.

The proof of the L^2 orthogonality follows from the observation that any function in $\mathrm{Curl}(V(\mathcal{T}))$ is a divergence-free Raviart–Thomas function; the converse holds as well for the simply-connected domain. This plus the discrete equation with such a test function leads to the asserted L^2 orthogonality. □

3 Supercloseness results

This section is devoted to the proof that the divergence term in the least squares functional at the discrete minimiser is much smaller than the dominating flux term. Although the proof below is different from that in [8] and based on L^2 error control for Crouzeix–Raviart nonconforming FEM, we believe that it is known. Since the following result seems unavailable in the literature for non-convex domains, some direct proof is given below for convenient reading.

Let H denote the maximal mesh-size in the current regular triangulation $\mathcal{T}$ and let $1/2 < s < \pi/\omega$ for the maximal interior angle ω of the non-convex polygonal domain Ω. Note the regularity index s attains the value 1 for convex domains.

Theorem 3.1 *The LSFEM solution satisfies*

$$\|\Pi_0 f + \operatorname{div} \mathbf{p}_{LS}\| \lesssim H^s \|\mathbf{p}_{LS} - \nabla u_{LS}\|. \tag{3.1}$$

Before the remaining part of this section is devoted to the proof of Theorem 3.1, various supercloseness results are deduced from it. Recall that $(\mathbf{p}_{LS}, u_{LS})$ denotes the least squares solution and $\mathbf{p}_{RT}$ denotes the lowest-order Raviart–Thomas mixed FEM approximation of the PMP [7,10,11,23], i.e., there exists $(\mathbf{p}_{RT}, u_{RT}) \in RT_0(\mathcal{T}) \times P_0(\mathcal{T})$ with

$$\int_\Omega \mathbf{p}_{RT} \cdot \mathbf{q}_{RT}\, dx + \int_\Omega u_{RT} \operatorname{div} \mathbf{q}_{RT}\, dx = 0 \quad \text{for all } \mathbf{q}_{RT} \in RT_0(\mathcal{T}),$$
$$\int_T (f + \operatorname{div} \mathbf{p}_{RT})\, dx = 0 \quad \text{for all } T \in \mathcal{T}.$$

Moreover, let u_{CR} (resp. $\widehat{u}_{CR}$) denote the NCFEM approximation of the PMP with right-hand side f (resp. $-\operatorname{div} \mathbf{p}_{LS}$). Proposition 2.1 leads to some $\widehat{u}_{CR} \in CR_0^1(\mathcal{T})$ with $\Pi_0 \mathbf{p}_{LS} = \nabla_{NC} \widehat{u}_{CR}$. Let $u_C \in V(\mathcal{T})$ denote the Courant finite element solution of the PMP with right-hand side $f \in L^2(\Omega)$ with the oscillation term

$$\operatorname{osc}^2(f, \mathcal{T}) := \sum_{T \in \mathcal{T}} |T|\, ||f - \Pi_0 f||^2_{L^2(T)}.$$

Recall that $|||\cdot||| \equiv ||\nabla \cdot||$ denotes the H^1 seminorm and define its discrete version $|||\cdot|||_{NC} := ||\nabla_{NC} \cdot||$ with respect to the underlying regular triangulation $\mathcal{T}$.

Corollary 3.2 *The aforementioned approximations satisfy*

$$\|\mathbf{p}_{LS} - \mathbf{p}_{RT}\|^2 + |||\widehat{u}_{CR} - u_{CR}|||^2_{NC} + |||u_{LS} - u_C|||^2$$
$$\lesssim H^{2s}\, LS(\Pi_0 f; \mathbf{p}_{LS}, u_{LS}) + \operatorname{osc}^2(f, \mathcal{T}).$$

Proof All the three terms on the left-hand side are controlled by $\|\Pi_0 f + \operatorname{div} \mathbf{p}_{LS}\|$ and then the corollary follows from (3.1). Since the supercloseness of the nonconforming Crouzeix–Raviart FEM seems to be new, the proof below focusses on the estimation of $|||\widehat{u}_{CR} - u_{CR}|||_{NC}$.

Let $\widehat{u}_{CR}$ (resp. u_{CR} and $\tilde{u}_{CR}$) solve the NCFEM for the PMP with right-hand side $-\operatorname{div}\mathbf{p}_{LS}$ (resp. f and $\Pi_0 f$). Proposition 2.1 and the admissible test function $\widehat{u}_{CR} - \tilde{u}_{CR}$ in NCFEM lead to

$$|||\widehat{u}_{CR} - \tilde{u}_{CR}|||_{NC}^2 = \int_\Omega (\Pi_0 f + \operatorname{div}\mathbf{p}_{LS})(\tilde{u}_{CR} - \widehat{u}_{CR})\,dx.$$

The discrete Friedrichs inequality [10] for functions in $CR_0^1(\mathcal{T})$, i.e.,

$$\|v_{CR}\| \le c_{dF}|||v_{CR}|||_{NC} \quad \text{for all } v_{CR} \in CR_0^1(\mathcal{T}),$$

results in

$$|||\widehat{u}_{CR} - \tilde{u}_{CR}|||_{NC} \le c_{dF}\|\Pi_0 f + \operatorname{div}\mathbf{p}_{LS}\|. \tag{3.2}$$

The solution u_{CR} of NCFEM with right-hand side f satisfies

$$|||u_{CR} - \tilde{u}_{CR}|||_{NC}^2 = \int_\Omega (f - \Pi_0 f)(u_{CR} - \tilde{u}_{CR})\,dx.$$

Since $\int_T (f - \Pi_0 f)\,dx = 0$ for any $T \in \mathcal{T}$, this equals

$$\int_\Omega (f - \Pi_0 f)(1 - \Pi_0)(u_{CR} - \tilde{u}_{CR})\,dx.$$

The piecewise Poincaré inequality (with $h_{\mathcal{T}}|_T := |T|^{1/2}$ for any triangle $T \in \mathcal{T}$) shows

$$\begin{aligned}\int_\Omega (f - \Pi_0 f)(1 - \Pi_0)(u_{CR} - \tilde{u}_{CR})\,dx &\lesssim \|h_{\mathcal{T}}(f - \Pi_0 f)\|\;|||u_{CR} - \tilde{u}_{CR}|||_{NC}\\ &= \operatorname{osc}(f, \mathcal{T})|||u_{CR} - \tilde{u}_{CR}|||_{NC}.\end{aligned}$$

Consequently,

$$|||u_{CR} - \tilde{u}_{CR}|||_{NC} \lesssim \operatorname{osc}(f, \mathcal{T}). \tag{3.3}$$

The triangle inequality and the estimates (3.2)–(3.3) show that $|||\widehat{u}_{CR} - u_{CR}|||_{NC}^2$ is controlled by the right-hand side in the corollary. □

The main tool in the proof of (3.1) is the following superior convergence of the Crouzeix–Raviart errors in $L^2(\Omega)$ (when compared with the nonconforming energy norm) which is standard [7,10] for H^2 regular problems when Ω is convex. Recall that H denotes the maximal mesh-size and s is the index of elliptic regularity.

Lemma 3.3 (L^2 error estimate for NCFEM) *Any $v_{CR} \in CR_0^1(\mathcal{T})$ with*

$$\int_\Omega \nabla_{NC} v_{CR} \cdot \nabla w_C\,dx = 0 \quad \textit{for all } w_C \in V(\mathcal{T})$$

satisfies

$$\|v_{CR}\| \lesssim H^s |||v_{CR}|||_{NC}.$$

Proof Let $z \in H^{1+s}(\Omega) \cap H_0^1(\Omega)$ be the solution of the PMP with right-hand side $v_{CR} = -\Delta z$. In case $0 < s < 1$, the textbook analysis is not applicable and, hence, arguments from a medius analysis are exploited. Given $v_{CR} \in CR_0^1(\mathcal{T})$, one defines a conforming approximation by the averaging of the possible values

$$v_1(z) := v_{CR}^*(z) := \lim_{\delta\to 0} \int_{B(z,\delta)} v_{CR}\, dx/|B(z,\delta)|$$

of the (possibly) discontinuous v_{CR} at any interior node $z \in \mathcal{N}(\Omega)$ (v_{CR}^* denotes the precise representation of the Lebesgue function v_{CR}). Linear interpolation of those values defines $v_1 \in P_1(\mathcal{T}) \cap C_0(\Omega)$.
In the second step, one defines $v_2 \in P_2(\mathcal{T}) \cap C_0(\Omega)$ which equals v_1 at all nodes $\mathcal{N}$ and satisfies

$$\int_E v_{CR}\, ds = \int_E v_2\, ds \quad \text{for all } E \in \mathcal{E}(\Omega).$$

In the third step, one adds the cubic bubble functions to v_2 such that the resulting function $v_3 \in P_3(\mathcal{T}) \cap C_0(\Omega)$ equals v_2 along the edges and satisfies

$$\int_T v_{CR}\, dx = \int_T v_3\, dx \quad \text{for all } T \in \mathcal{T}.$$

Therefore, an integration by parts shows

$$\int_T \nabla v_{CR}\, dx = \int_T \nabla v_3\, dx \quad \text{for all } T \in \mathcal{T}.$$

Altogether,

$$\begin{aligned}\|v_{CR}\|^2 &= \int_\Omega (v_{CR} - v_3) v_{CR}\, dx + \int_\Omega v_3 v_{CR}\, dx \\ &= \int_\Omega (v_{CR} - v_3)(1 - \Pi_0) v_{CR}\, dx - \int_\Omega v_3 \Delta z\, dx.\end{aligned}$$

Piecewise Poincaré inequalities lead to

$$\int_\Omega (v_{CR} - v_3)(1 - \Pi_0) v_{CR}\, dx \le H^2 |||v_{CR}|||_{NC}\, |||v_{CR} - v_3|||_{NC}.$$

The design of the dual solution z leads to

$$-\int_\Omega v_3 \Delta z\, dx = \int_\Omega \nabla v_3 \cdot \nabla z\, dx = \int_\Omega \nabla_{NC} v_{CR} \cdot \nabla z\, dx + \int_\Omega \nabla_{NC}(v_3 - v_{CR}) \cdot \nabla z\, dx.$$

Since v_{CR} is perpendicular to the conforming nodal interpolation $I_C z \in V(\mathcal{T})$ and since $\int_T \nabla(v_3 - v_{CR})\,dx = 0$ for all $T \in \mathcal{T}$, the last expression equals

$$\int_\Omega \nabla_{NC} v_{CR} \cdot \nabla(z - I_C z)\,dx + \int_\Omega \nabla_{NC}(v_3 - v_{CR}) \cdot \nabla(z - I_C z)\,dx.$$

The reduced elliptic regularity of the PMP plus standard finite element interpolation estimates on polygonal domains bound the previous terms with some $C(\Omega, s) \lesssim 1$ from above by

$$C(\Omega, s) H^s \|v_{CR}\| (|||v_{CR}|||_{NC} + |||v_3 - v_{CR}|||_{NC}).$$

The approximation and stability properties of v_1 has been studied in a former work on preconditioners for nonconforming FEM [9] (called enrichment therein). This and standard arguments also prove stability in the sense that

$$|||v_3|||_{NC} \lesssim |||v_{CR}|||_{NC}.$$

The combination of the above estimates concludes the proof. □

Proof of Theorem 3.1. Given the piecewise constant $\Pi_0 f + \operatorname{div} \mathbf{p}_{LS} \in P_0(\mathcal{T})$ and the inf-sup condition (also called LBB condition) for the lowest-order Raviart–Thomas functions, there exists some $\mathbf{q}_{RT} \in RT_0(\mathcal{T})$ with

$$-\operatorname{div} \mathbf{q}_{RT} = -\Pi_0 f - \operatorname{div} \mathbf{p}_{LS} \quad \text{and} \quad \|\mathbf{q}_{RT}\|_{H(\operatorname{div},\Omega)} \lesssim \|\Pi_0 f + \operatorname{div} \mathbf{p}_{LS}\|.$$

Amongst all possible $\mathbf{q}_{RT} \in RT_0(\mathcal{T})$ with prescribed divergence, the mixed finite element solution minimises the L^2 norm $\|\mathbf{q}_{RT}\|$ of the flux and, hence, is orthogonal onto $\operatorname{Curl}(V(\mathcal{T}))$. Hence, Proposition 2.1 shows that we may and will assume that

$$\mathbf{q}_{RT} = \nabla_{NC} v_{CR} + \frac{\operatorname{div} \mathbf{q}_{RT}}{2}(\bullet - \operatorname{mid}(\mathcal{T})) \quad \text{a.e. in } \Omega$$

with the Crouzeix–Raviart solution v_{CR} of the PMP with right-hand side $-\operatorname{div} \mathbf{q}_{RT}$. Recall the analog identity for $\mathbf{p}_{LS}$ with $\Pi_0 \mathbf{p}_{LS} = \nabla_{NC} \widehat{u}_{CR}$. The LSFEM leads to

$$\|\Pi_0 f + \operatorname{div} \mathbf{p}_{LS}\|^2 = \int_\Omega (\Pi_0 f + \operatorname{div} \mathbf{p}_{LS})\, \operatorname{div} \mathbf{q}_{RT}\,dx = \int_\Omega (\nabla u_{LS} - \mathbf{p}_{LS}) \cdot \mathbf{q}_{RT}\,dx.$$

The aforementioned identities for $\mathbf{q}_{RT}$ and $\mathbf{p}_{LS}$ show that the above term equals

$$-\frac{1}{4}\int_\Omega \operatorname{div} \mathbf{q}_{RT}\, \operatorname{div} \mathbf{p}_{LS} |x - \operatorname{mid}(\mathcal{T})|^2\,dx + \int_\Omega \nabla_{NC}(u_{LS} - \widehat{u}_{CR}) \cdot \nabla_{NC} v_{CR}\,dx.$$

The first term is controlled by some $H||\operatorname{div} \mathbf{q}_{RT}||\, ||h_\mathcal{T} \operatorname{div} \mathbf{p}_{LS}||$ in terms of the local mesh-size $h_\mathcal{T} \in P_0(\mathcal{T})$. An inverse inequality proves that

$$||h_\mathcal{T} \operatorname{div}_{NC}(\mathbf{p}_{LS} - \nabla u_{LS})|| \lesssim ||\mathbf{p}_{LS} - \nabla u_{LS}||.$$

This results in

$$-\frac{1}{4}\int_\Omega \operatorname{div}\mathbf{q}_{RT}\operatorname{div}\mathbf{p}_{LS}|x-\operatorname{mid}(\mathcal{T})|^2\,dx \lesssim H\,||\mathbf{p}_{LS}-\nabla u_{LS}||\,||\operatorname{div}\mathbf{q}_{RT}||.$$

The second term is recast with the observation that, given any $v_C \in V(\mathcal{T})$ with $\nabla v_C \in P_0(\mathcal{T};\mathbb{R}^2)$, it follows

$$\int_\Omega \nabla_{NC}(u_{LS}-\widehat{u}_{CR})\cdot\nabla v_C\,dx = \int_\Omega (\mathbf{p}_{LS}-\Pi_0\mathbf{p}_{LS})\cdot\nabla v_C\,dx = 0.$$

Hence, Lemma 3.3 implies

$$||\widehat{u}_{CR}-u_{LS}|| \lesssim H^s|||\widehat{u}_{CR}-u_{LS}|||_{NC}.$$

This is applied at the last step after $\Pi_0\mathbf{q}_{RT} = \nabla_{NC}v_{CR}$ followed by an integration by parts, namely

$$\begin{aligned}\int_\Omega \nabla_{NC}(u_{LS}-\widehat{u}_{CR})\cdot\nabla_{NC}v_{CR}\,dx &= \int_\Omega \mathbf{q}_{RT}\cdot\nabla_{NC}(u_{LS}-\widehat{u}_{CR})\,dx\\ &= \int_\Omega (\widehat{u}_{CR}-u_{LS})\operatorname{div}\mathbf{q}_{RT}\,dx\\ &\lesssim H^s|||\widehat{u}_{CR}-u_{LS}|||_{NC}||\operatorname{div}\mathbf{q}_{RT}||.\end{aligned}$$

Notice that $|||\widehat{u}_{CR}-u_{LS}|||_{NC} = ||\Pi_0\mathbf{p}_{LS}-\nabla u_{LS}|| \le ||\mathbf{p}_{LS}-\nabla u_{LS}||$. The combination of the respective upper bounds for the first and second term yields (3.1). □

4 Saturation for large bulk parameter

This section is devoted to the proof of the existence of constants $0 < \Lambda_1 < \infty$ and $0 < \varrho_1 < 1$ with (1.1)–(1.2) for a uniform mesh-refining or some particular adaptive mesh-refining strategy. The point of departure for the specification of the latter, is the discussion of a general marking with the localisation of the least squares residual into

$$\begin{aligned}\mu_\ell^2(T) := ||f-\Pi_\ell f||^2_{L^2(T)} + ||\Pi_\ell f+\operatorname{div}\mathbf{p}_\ell||^2_{L^2(T)}\\ + ||(1-\Pi_\ell)\mathbf{p}_\ell||^2_{L^2(T)} + ||\Pi_\ell\mathbf{p}_\ell-\nabla u_\ell||^2_{L^2(T)}\end{aligned}$$

for the LSFEM solution $(\mathbf{p}_\ell, u_\ell)$ (and the L^2 orthogonal projection Π_ℓ onto the piecewise constants) with respect to the regular triangulation $\mathcal{T}_\ell$ and a triangle $T \in \mathcal{T}_\ell$. For any subset $\mathcal{M}_\ell \subset \mathcal{T}_\ell$ of triangles, its contribution to the least squares functional $LS(f;\mathbf{p}_\ell, u_\ell)$ is abbreviated as

$$\mu_\ell^2(\mathcal{M}_\ell) := \sum_{T\in\mathcal{M}_\ell}\mu_\ell^2(T) \quad\text{and so}\quad \mu_\ell^2(\mathcal{T}_\ell) \equiv LS(f;\mathbf{p}_\ell, u_\ell).$$

Given any bulk parameter $0 < \Theta < 1$, the step MARK of an adaptive refinement selects some subset (e.g., of almost minimal cardinality) $\mathcal{M}_\ell \subset \mathcal{T}_\ell$ with

$$\Theta \mu_\ell^2(\mathcal{T}_\ell) \le \mu_\ell^2(\mathcal{M}_\ell). \tag{4.1}$$

This is equivalent to

$$\mu_\ell^2(\mathcal{T}_\ell \backslash \mathcal{M}_\ell) \le (1-\Theta)\mu_\ell^2(\mathcal{T}_\ell) = (1-\Theta)\, LS(f; \mathbf{p}_\ell, u_\ell).$$

The highly-oscillatory data example from the introduction with (1.3) has to be excluded to guarantee saturation (1.2). The underlying assumption throughout this section will be that the data resolution error $\|f - f_\ell\|$ is small compared to $LS(f_\ell; \mathbf{p}_\ell, u_\ell)$ for $f_\ell := \Pi_\ell f$. At least the refined triangulation $\mathcal{T}_{\ell+1}$ shall resolve the data and then the last condition implies

$$\sum_{T \in \mathcal{T}_\ell \backslash \mathcal{M}_\ell} \|\mathbf{p}_\ell - \nabla u_\ell\|_{L^2(T)}^2 \le (1-\Theta)\, LS(f_{\ell+1}; \mathbf{p}_\ell, u_\ell). \tag{4.2}$$

All the triangulations in this paper are defined by a sequence of one-level refinements starting with the initial triangulation $\mathcal{T}_0$ of Ω into triangles. A one-level refinement consists of markings in the newest-vertex bisection as depicted in Fig. 1 to generate a shape-regular refinement.

The following two refinement conditions (R1)–(R2) are imposed on the regular triangulations $\mathcal{T}_\ell$ and $\mathcal{T}_{\ell+1}$ of Ω with their respective LSFEM solutions $(\mathbf{p}_\ell, u_\ell)$ and $(\mathbf{p}_{\ell+1}, u_{\ell+1})$ for saturation.

- (R1) The LSFEM solution $(\mathbf{p}_\ell, u_\ell)$ satisfies (4.2) with $\Theta_2 \le \Theta \le 1$.
- (R2) The regular triangulation $\mathcal{T}_{\ell+1}$ is a one-level refinement of $\mathcal{T}_\ell$ such that any triangle in $\mathcal{M}_\ell$ is red-refined.

Theorem 4.1 *Provided the initial regular triangulation $\mathcal{T}_0$ is sufficiently fine, there exist constants $0 < \varrho_2 < 1$, $0 < \Lambda_2 < \infty$, and $0 < \Theta_2 < 1$ such that* (R1)–(R2) *imply*

$$\begin{aligned} &LS(f_{\ell+1}; \mathbf{p}_{\ell+1}, u_{\ell+1}) + \Lambda_2 \|(1-\Pi_{\ell+1})\mathbf{p}_{\ell+1}\|^2 \\ &\quad \le \varrho_2 \left(LS(f_{\ell+1}; \mathbf{p}_\ell, u_\ell) + \Lambda_2 \|(1-\Pi_\ell)\mathbf{p}_\ell\|^2 \right). \end{aligned}$$

Some comments are in order on the parameters before the proof of Theorem 4.1.

Remark 4.1 Theorem 4.1 implies saturation (1.2) under the aforementioned assumption $f = f_{\ell+1} := \Pi_{\ell+1} f$. Since the refinement rules (R1)–(R2) do not provide the resolution of the data, additional algorithms are required to guarantee this assumption, e.g., the data approximation algorithm in the separate marking of [15,18].

Remark 4.2 The crucial point is that $\Theta_2 < 1$ may be large (i.e., close to one) and so is Θ with (R1). Some closer investigations on the parameters at the very end of this section reveal that $\varrho_2 < 1$ implies $2/3 \le \Theta_2$. This is crucial and seems to expel the proof of optimal convergence rates with arguments from [12,18,28].

Remark 4.3 Some closer investigations on the parameters at the very end of this section reveal that $0 < \Lambda_2 < \infty$ can be arbitrarily small. However, $0 < \Lambda_2 << 1$ implies that $\varrho_2, \Theta_2 < 1$ are very close to one and the initial triangulation $\mathcal{T}_0$ is very fine.

Remark 4.4 The condition (R2) on the red-refinement can be relaxed. However, Lemma 4.3 requires that all edges of any triangle in $\mathcal{M}_\ell$ are bisected.

The proof of Theorem 4.1 is split into four ingredients. The first of those is based on the stability of the mixed FEM plus elementary algebra.

Lemma 4.2 *It holds*

$$\|f_\ell + \operatorname{div}\mathbf{p}_\ell\|^2 + 2\|f_{\ell+1} - f_\ell + \operatorname{div}(\mathbf{p}_{\ell+1} - \mathbf{p}_\ell)\|^2 - 2\|f_{\ell+1} + \operatorname{div}\mathbf{p}_{\ell+1}\|^2$$
$$\lesssim \|\mathbf{p}_{\ell+1} - \mathbf{p}_\ell - \nabla(u_{\ell+1} - u_\ell)\|^2.$$

Proof The inf-sup condition from the Proof of Theorem 3.1 leads to $\mathbf{q}_\ell \in RT_0(\mathcal{T}_\ell)$ with $f_\ell + \operatorname{div}\mathbf{p}_\ell = \operatorname{div}\mathbf{q}_\ell$ and some stability constant $C_{stab} \approx 1$ with

$$\|\mathbf{q}_\ell\|^2 \leq C_{stab}\|\operatorname{div}\mathbf{q}_\ell\|^2 = C_{stab}\|f_\ell + \operatorname{div}\mathbf{p}_\ell\|^2.$$

The LSFEM on the level ℓ shows

$$\|f_\ell + \operatorname{div}\mathbf{p}_\ell\|^2 = \int_\Omega (\nabla u_\ell - \mathbf{p}_\ell)\cdot\mathbf{q}_\ell\,dx.$$

The LSFEM on the level $\ell+1$ with test function $\mathbf{q}_\ell$ shows that the last term equals

$$\int_\Omega (\mathbf{p}_{\ell+1} - \mathbf{p}_\ell - \nabla(u_{\ell+1} - u_\ell))\cdot\mathbf{q}_\ell\,dx + \int_\Omega (f_{\ell+1} + \operatorname{div}\mathbf{p}_{\ell+1})\operatorname{div}\mathbf{q}_\ell\,dx.$$

Since $\operatorname{div}\mathbf{q}_\ell = f_\ell + \operatorname{div}\mathbf{p}_\ell$, the binomial formula shows that the second summand in the last term equals one half times

$$\|f_{\ell+1} + \operatorname{div}\mathbf{p}_{\ell+1}\|^2 + \|f_\ell + \operatorname{div}\mathbf{p}_\ell\|^2 - \|f_{\ell+1} - f_\ell + \operatorname{div}(\mathbf{p}_{\ell+1} - \mathbf{p}_\ell)\|^2.$$

The combination of the aforementioned identities shows

$$\begin{aligned}&\|f_\ell + \operatorname{div}\mathbf{p}_\ell\|^2 + \|f_{\ell+1} - f_\ell + \operatorname{div}(\mathbf{p}_{\ell+1} - \mathbf{p}_\ell)\|^2 - \|f_{\ell+1} + \operatorname{div}\mathbf{p}_{\ell+1}\|^2\\ &\quad = 2\int_\Omega (\mathbf{p}_{\ell+1} - \mathbf{p}_\ell - \nabla(u_{\ell+1} - u_\ell))\cdot\mathbf{q}_\ell\,dx\\ &\quad \leq \|\mathbf{q}_\ell\|^2/(2C_{stab}) + 2C_{stab}\|\mathbf{p}_{\ell+1} - \mathbf{p}_\ell - \nabla(u_{\ell+1} - u_\ell)\|^2.\end{aligned}$$

The combination of $\|\mathbf{q}_\ell\|^2/C_{stab} \leq \|f_\ell + \operatorname{div}\mathbf{p}_\ell\|^2$ with the previous inequality concludes the proof. □

The second ingredient exploits arguments from a discrete efficiency analysis of adaptive mixed FEM.

Lemma 4.3 *It holds*

$$\|\Pi_\ell \mathbf{p}_\ell - \nabla u_\ell\|^2 \lesssim \|\mathbf{p}_{\ell+1} - \mathbf{p}_\ell - \nabla(u_{\ell+1} - u_\ell)\|^2 + (1-\Theta) LS(f_{\ell+1}; \mathbf{p}_\ell, u_\ell).$$

Proof Recall the representation $\Pi_\ell \mathbf{p}_\ell = \nabla_{NC} \widehat{u}_\ell$ from Proposition 2.1. Then $v_{CR} := \widehat{u}_\ell - u_\ell \in CR_0^1(\mathcal{T}_\ell)$ satisfies

$$\|\Pi_\ell \mathbf{p}_\ell - \nabla u_\ell\|^2 = \int_\Omega (\mathbf{p}_\ell - \nabla u_\ell) \cdot \nabla_{NC} v_{CR}\, dx.$$

The nonconforming v_{CR} is first approximated similar to the proof of Lemma 3.3 by some $v_\ell \in V(\mathcal{T}_\ell)$ with $v_{CR}^*(z) = v_\ell(z) = v_{\ell+1}(z)$ for any interior node $z \in \mathcal{N}_\ell(\Omega)$. Whenever an interior edge $E \in \mathcal{E}_\ell$ of length $|E|$ is refined, written $E \in \mathcal{E}_\ell(\Omega)\backslash\mathcal{E}_{\ell+1}$, its midpoint $\mathrm{mid}(E) \in \mathcal{N}_{\ell+1}(\Omega)$ is an interior node in the refined triangulation $\mathcal{T}_{\ell+1}$ and has some conforming nodal basis function $\varphi_E \in V(\mathcal{T}_{\ell+1})$. Then

$$v_{\ell+1} := v_\ell + \sum_{E \in \mathcal{E}_\ell(\Omega)\backslash\mathcal{E}_{\ell+1}} 2 \int_E (v_{CR} - v_\ell) ds/|E|\, \varphi_E \in V(\mathcal{T}_{\ell+1})$$

satisfies, for all $E \in \mathcal{E}_\ell(\Omega)\backslash\mathcal{E}_{\ell+1}$, that

$$\int_E v_{\ell+1}\, ds = \int_E v_{CR}\, ds.$$

Since any $T \in \mathcal{M}_\ell$ is red-refined by (R2), all its edges are bisected and the previous identity leads (via an integration by parts) to

$$\int_T (\mathbf{p}_\ell - \nabla u_\ell) \cdot \nabla(v_{CR} - v_{\ell+1})\, dx = 0 \quad \text{for all } T \in \mathcal{M}_\ell.$$

Let $\Omega' := \cup(\mathcal{T}_\ell \backslash \mathcal{M}_\ell)$ abbreviate that part of the domain which is not covered by the marked triangles. Then,

$$\|\Pi_\ell \mathbf{p}_\ell - \nabla u_\ell\|^2 = \int_\Omega (\mathbf{p}_\ell - \nabla u_\ell) \cdot \nabla v_{\ell+1}\, dx + \int_{\Omega'} (\mathbf{p}_\ell - \nabla u_\ell) \cdot \nabla_{NC}(v_{CR} - v_{\ell+1})\, dx.$$

The test function $v_{\ell+1} \in V(\mathcal{T}_{\ell+1})$ in LSFEM on the level $\ell+1$ satisfies

$$\int_\Omega (\mathbf{p}_{\ell+1} - \nabla u_{\ell+1}) \cdot \nabla v_{\ell+1}\, dx = 0.$$

Therefore,

$$\begin{aligned}\|\Pi_\ell \mathbf{p}_\ell - \nabla u_\ell\|^2 \le &-\int_\Omega (\mathbf{p}_{\ell+1} - \mathbf{p}_\ell - \nabla(u_{\ell+1} - u_\ell)) \cdot \nabla v_{\ell+1}\, dx \\ &+ ||\mathbf{p}_\ell - \nabla u_\ell||_{L^2(\Omega')}\, |||v_{CR} - v_{\ell+1}|||_{NC}.\end{aligned}$$

The stability of the discrete approximation operators [9] reads

$$|||v_{\ell+1}||| \lesssim |||v_{CR}|||_{NC} = ||\Pi_\ell \mathbf{p}_\ell - \nabla u_\ell||.$$

Observe that (R1) and (4.2) imply

$$||\mathbf{p}_\ell - \nabla u_\ell||^2_{L^2(\Omega')} \le (1-\Theta) LS(f_{\ell+1}; \mathbf{p}_\ell, u_\ell).$$

The combination of the previous three displayed formulas concludes the proof. □

The third ingredient is the well-established Galerkin orthogonality for LSFEM.

Lemma 4.4 *It holds*

$$LS(0; \mathbf{p}_{\ell+1} - \mathbf{p}_\ell, u_{\ell+1} - u_\ell) = LS(f; \mathbf{p}_\ell, u_\ell) - LS(f; \mathbf{p}_{\ell+1}, u_{\ell+1}).$$

Proof The proof is straightforward with elementary algebra and the Galerkin orthogonality of the LSFEM. □

The last ingredient is the explicit reduction for $\|(1-\Pi_\ell)\mathbf{p}_\ell\|$.

Lemma 4.5 *The refinement conditions* (R1)–(R2) *with* (4.2) *imply*

$$\|(1-\Pi_{\ell+1})\mathbf{p}_\ell\|^2 \le 1/4\, \|(1-\Pi_\ell)\mathbf{p}_\ell\|^2 + 3(1-\Theta)/4\, LS(f_{\ell+1}; \mathbf{p}_\ell, u_\ell).$$

Proof Any $T \in \mathcal{M}_\ell$ with vertices P_1, P_2, P_3 and opposite edges E_1, E_2, E_3 of lengths $|E_1|, |E_2|, |E_3|$ satisfies

$$|| \bullet - \mathrm{mid}(T)||^2_{L^2(T)} = |T|(|E_1|^2 + |E_2|^2 + |E_3|^2)/36. \tag{4.3}$$

(The proof of (4.3) is by direct calculations and hence omitted.) The red-refinement $\mathcal{T}_{\ell+1}(T)$ of T consist of the four congruent subtriangles T_1, T_2, T_3, T_4 enumerated such that T_4 is the triangle in the centre and the subtriangle T_j has the vertex P_j for any $j = 1, 2, 3$. The four contributions of the four subtriangles are equal to each other and can also be calculated with (4.3). This results in

$$\int_{T_j} | \bullet - \mathrm{mid}(T_j)|^2\, dx = |T_4|(|E_1|^2 + |E_2|^2 + |E_3|^2)/144 \quad \text{for } j = 1, \dots, 4.$$

The comparison with (4.3) proves that the identity mapping $\bullet$ and its piecewise constant integral means (which interpolate at the centres of inertia) satisfy

$$\|(1-\Pi_{\ell+1}) \bullet \|^2_{L^2(T)} = 1/4\, \|(1-\Pi_\ell) \bullet \|^2_{L^2(T)}.$$

The multiplication with the constant $|\operatorname{div} \mathbf{p}_\ell|_T|^2$ proves

$$\|(1-\Pi_{\ell+1})\mathbf{p}_\ell\|^2_{L^2(T)} = 1/4\, \|(1-\Pi_\ell)\mathbf{p}_\ell\|^2_{L^2(T)}.$$

This verifies even equality in the key estimate

$$\|(1-\Pi_{\ell+1})\mathbf{p}_\ell\|^2_{L^2(T)} \le \varrho_3\, \|(1-\Pi_\ell)\mathbf{p}_\ell\|^2_{L^2(T)} \quad \text{for all } T\in\mathcal{M}_\ell$$

with $\varrho_3 = 1/4$. Any triangle which is bisected (also called green refined), shows the previous inequality with $\varrho_3 = 1/2$ while $\varrho = 1$ for the unrefined; further details are omitted.

Recall from the proof of Lemma 4.3 that $\Omega' := \cup(\mathcal{T}_\ell\backslash\mathcal{M}_\ell)$ abbreviates that part of the domain which is not covered by the marked triangles. The sum of all triangles leads to

$$\begin{aligned}\|(1-\Pi_{\ell+1})\mathbf{p}_\ell\|^2 &\le 1/4\|(1-\Pi_\ell)\mathbf{p}_\ell\|^2_{L^2(\Omega\backslash\Omega')} + \|(1-\Pi_\ell)\mathbf{p}_\ell\|^2_{L^2(\Omega')}\\ &= 1/4\|(1-\Pi_\ell)\mathbf{p}_\ell\|^2 + 3/4\|(1-\Pi_\ell)\mathbf{p}_\ell\|^2_{L^2(\Omega')}.\end{aligned}$$

This and (4.2) conclude the proof. □

Proof of Theorem 4.1. Recall (2.3) and set $LS'_\ell := LS(f_\ell;\mathbf{p}_\ell,u_\ell)$, $\widetilde{LS}_\ell := LS(f_{\ell+1};\mathbf{p}_\ell,u_\ell) = LS'_\ell + ||f_{\ell+1}-f_\ell||^2$, and $\widetilde{LS}_{\ell+1} := LS(f_{\ell+1};\,\mathbf{p}_{\ell+1},u_{\ell+1})$. This and Lemmas 4.2–4.3 show, with some $C\approx 1$, that

$$\begin{aligned}LS'_\ell - \|(1-\Pi_\ell)\mathbf{p}_\ell\|^2 &= \|f_\ell + \operatorname{div}\mathbf{p}_\ell\|^2 + \|\Pi_\ell\mathbf{p}_\ell - \nabla u_\ell\|^2\\ &\le 2\|f_{\ell+1}+\operatorname{div}\mathbf{p}_{\ell+1}\|^2 - 2\|f_{\ell+1}-f_\ell+\operatorname{div}(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|^2\\ &\quad + C\|\mathbf{p}_{\ell+1}-\mathbf{p}_\ell-\nabla(u_{\ell+1}-u_\ell)\|^2 + C(1-\Theta)\widetilde{LS}_\ell.\end{aligned}$$

The multiplication by any δ with $0<\delta<\min\{1/2,1/C\}$ and Lemma 4.4 lead to

$$\begin{aligned}\widetilde{LS}_{\ell+1} &+ \|\operatorname{div}(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|^2 + 2\delta\|f_{\ell+1}-f_\ell+\operatorname{div}(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|^2\\ &\le (1-\delta)\widetilde{LS}_\ell + \delta||f_{\ell+1}-f_\ell||^2 + 2\delta\|f_{\ell+1}+\operatorname{div}\mathbf{p}_{\ell+1}\|^2\\ &\quad + \delta\|(1-\Pi_\ell)\mathbf{p}_\ell\|^2 + (1-\Theta)\widetilde{LS}_\ell.\end{aligned}\tag{4.4}$$

The further analysis uses the following list of arguments (a)–(c) for the estimation of three terms on the right-hand side in the preceding inequality.

(a) The Young inequality

$$\|f_{\ell+1}-f_\ell\|^2 \le 2\|f_{\ell+1}-f_\ell+\operatorname{div}(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|^2 + 2\|\operatorname{div}(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|^2.$$

(b) The supercloseness result of Theorem 3.1 for $\mathcal{T}\equiv\mathcal{T}_{\ell+1}$ reads

$$\|f_{\ell+1}+\operatorname{div}\mathbf{p}_{\ell+1}\|^2 \le \varepsilon\,\widetilde{LS}_{\ell+1}$$

with some $\varepsilon\approx H^s$ which tends to zero as the maximal mesh-size of $\mathcal{T}_\ell$ tends to zero; $0<\varepsilon<1/2$ will be chosen sufficiently small via the condition that the initial triangulation $\mathcal{T}_0$ is sufficiently fine.

(c) The Young inequality shows, for any $0 < \lambda < \infty$

$$\|(1-\Pi_{\ell+1})\mathbf{p}_{\ell+1}\|^2 \leq (1+\lambda)\|(1-\Pi_{\ell+1})\mathbf{p}_\ell\|^2 + (1+1/\lambda)\|(1-\Pi_{\ell+1})(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|^2.$$

Lemma 4.5 implies some reduction formula for $\|(1-\Pi_{\ell+1})\mathbf{p}_\ell\|$. To control the second term $\|(1-\Pi_{\ell+1})(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|$, observe that the integrand $\|(1-\Pi_{\ell+1})(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|^2$ is equal to the constant divergence of $\mathbf{p}_{\ell+1}-\mathbf{p}_\ell$ in T' times $|x-\mathrm{mid}(T')|^2 \leq h_{\ell+1}^2|_{T'} := |T'|$ of $T' \in \mathcal{T}_{\ell+1}$ at any x in some triangle T' in $\mathcal{T}_{\ell+1}$. This proves

$$\|(1-\Pi_{\ell+1})(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|^2 \leq \|h_{\ell+1}\,\mathrm{div}(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|^2. \tag{4.5}$$

The combination of all arguments of (c) results in

$$\begin{aligned}
&\|(1-\Pi_{\ell+1})\mathbf{p}_{\ell+1}\|^2\\
&\quad\leq (1+\lambda)/4\|(1-\Pi_\ell)\mathbf{p}_\ell\|^2 + 3(1+\lambda)(1-\Theta)/4\,\widetilde{LS}_\ell\\
&\qquad+(1+1/\lambda)\;H^2\|\,\mathrm{div}(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|^2.
\end{aligned}$$

Multiply the respective inequalities (a), (b), and (c) by δ, 2δ, and some factor $0 < \Lambda < \infty$ and add them to (4.4). The result is equivalent to

$$\begin{aligned}
&(1-2\delta\varepsilon)\widetilde{LS}_{\ell+1} + \Lambda\|(1-\Pi_{\ell+1})\mathbf{p}_{\ell+1}\|^2\\
&\quad\leq (2-\Theta-\delta+3\Lambda(1+\lambda)(1-\Theta)/4)\;\widetilde{LS}_\ell\\
&\qquad+(\delta+\Lambda(1+\lambda)/4)\;\|(1-\Pi_\ell)\mathbf{p}_\ell\|^2\\
&\qquad+\Big(h^2\Lambda(1+1/\lambda)+2\delta-1\Big)\;\|\,\mathrm{div}(\mathbf{p}_{\ell+1}-\mathbf{p}_\ell)\|^2.
\end{aligned}$$

This inequality is divided by $1-2\delta\varepsilon$ and then proves the assertion with $\Lambda_2 := \Lambda/\,(1-2\delta\varepsilon)$ and $\varrho_2 := \varrho/\,(1-2\delta\varepsilon)$ for

$$\varrho := \max\{1-\delta+(1-\Theta)\,(1+3\Lambda(1+\lambda)/4)\,,\,\delta/\Lambda+(1+\lambda)/4\}$$

provided that

$$H^2\Lambda(1+1/\lambda)+2\delta \leq 1.$$

The latter condition as well as $\varrho_3 < 1$ follow for sufficiently fine meshes (as h and ε become small) once the parameters $0 < \lambda, \Lambda < \infty$ and $0 < \delta < \min\{1/2, 1/C\}$ are fixed with $\varrho < 1$. The crucial condition $\delta < 1/2$ and $\varrho < 1$ is feasible for large Θ_2; further details on the parameter choice are omitted for brevity. □

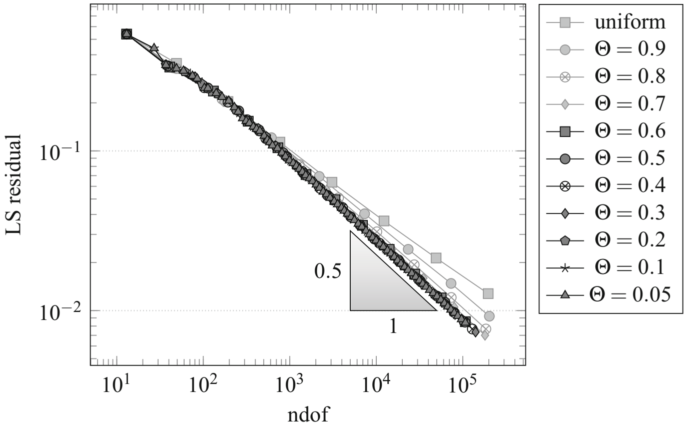

Fig. 2 Convergence history plot of the least squares functional $LS(f; p_\ell, u_\ell)$ in Sect. 5

5 Numerical experiments on L-shaped domain

Let $\Omega := (-1, 1)^2 \backslash [0, 1]^2$ be the L-shaped domain and let $f \equiv 1$. The Fig. 2 shows convergence of the natural adaptive LSFEM for a wide range of bulk parameters $0 < \Theta \leq 1$. This indicates that the restrictions on the parameter in condition (R1) does not seem to be very sharp. However, they are crucial for the analysis at hand.

Acknowledgements The authors thank Eunjung Lee of the Department of Computational Science and Engineering of Yonsei University in Seoul, Korea, and Gerhard Starke from the Faculty of Mathematics of the University of Duisburg-Essen, Germany, for valuable discussions.

References

1. Adler, J.H., Manteuffel, T.A., McCormick, S.F., Nolting, J.W., Ruge, J.W., Tang, L.: Efficiency based adaptive local refinement for first-order system least-squares formulations. SIAM J. Sci. Comput. **33**(1), 1–24 (2011)
2. Arnold, D., Brezzi, F.: Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. RAIRO Model. Math. Anal. Numer. **19**(1), 7–32 (1985)
3. Berndt, M., Manteuffel, T.A., McCormick, S.F.: Local error estimates and adaptive refinement for first-order system least squares (FOSLS). Electron. Trans. Numer. Anal. **6**, 35–43 (1997). ((electronic) Special issue on multilevel methods (Copper Mountain, CO, 1997))
4. Binev, P., Dahmen, W., DeVore, R.: Adaptive finite element methods with convergence rates. Numer. Math. **97**(2), 219–268 (2004)
5. Binev, P., DeVore, R.: Fast computation in adaptive tree approximation. Numer. Math. **97**(2), 193–217 (2004)
6. Bochev, P.B., Gunzburger, M.D.: Least-squares Finite Element Methods, volume 166 of Applied Mathematical Sciences. Springer, New York (2009)
7. Braess, D.: Finite Elements. Cambridge University Press, Cambridge (2007). third edition
8. Brandts, J., Chen, Y., Yang, J.: A note on least-squares mixed finite elements in relation to standard and mixed finite elements. IMA J. Numer. Anal. **26**(4), 779–789 (2006)
9. Brenner, S.C.: Two-level additive Schwarz preconditioners for nonconforming finite element methods. Math. Comp. **65**(215), 897–921 (1996)
10. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods. Texts in Applied Mathematics, vol. 15, 3rd edn. Springer, New York (2008)

11. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods, volume 15 of Springer Series in Computational Mathematics. Springer, New York (1991)
12. Carstensen, C., Feischl, M., Page, M., Praetorius, D.: Axioms of adaptivity. Comput. Math. Appl. **67**(6), 1195–1253 (2014)
13. Carstensen, C., Hoppe, R.W.H.: Error reduction and convergence for an adaptive mixed finite element method. Math. Comp. **75**, 1033–1042 (2006)
14. Carstensen, C., Hoppe, R.W.H., Eigel, M., Löbhard, C.: A review of unified a posteriori finite element error control. Numer. Math. Theory Methods Appl. **5**, 509–558 (2012)
15. Carstensen, C., Park, E.-J.: Convergence and optimality of adaptive least squares finite element methods. SIAM J. Numer. Anal. **53**(1), 43–62 (2015)
16. Carstensen, C., Peterseim, D., Schedensack, M.: Comparison results of finite element methods for the Poisson model problem. SIAM J. Numer. Anal. **50**(6), 2803–2823 (2012)
17. Carstensen, C., Rabus, H.: An optimal adaptive mixed finite element method. Math. Comp. **80**(274), 649–667 (2011)
18. Carstensen, C., Rabus, H.: Axioms of adaptivity for separate marking. (2016). arXiv:1606.02165 [math.NA] (Submitted)
19. Cascon, J.M., Kreuzer, Ch., Nochetto, R.H., Siebert, K.G.: Quasi-optimal convergence rate for an adaptive finite element method. SIAM J. Numer. Anal. **46**, 2524–2550 (2008)
20. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. Classics in Applied Mathematics, vol. 40. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2002)
21. Dörfler, W.: A convergent adaptive algorithm for Poisson's equation. SIAM J. Numer. Anal. **33**(3), 1106–1124 (1996)
22. Ferraz-Leite, C., Ortner, S., Praetorius, D.: Convergence of simple adaptive Galerkin schemes based on $h - h/2$ error estimators. Numer. Math. **116**(2), 291–316 (2010)
23. Girault, V., Raviart, P.-A.: Finite Element Methods for Navier–Stokes Equations, volume 5 of Springer Series in Computational Mathematics. Springer, Berlin (1986)
24. Marini, L.D.: An inexpensive method for the evaluation of the solution of the lowest order Raviart–Thomas mixed method. SIAM J. Numer. Anal. **22**(3), 493–496 (1985)
25. Nochetto, R.H., Siebert, K.G., Veeser, A.: Theory of Adaptive Finite Element Methods: An Introduction Multiscale Nonlinear and Adaptive Approximation. Springer, Berlin (2009). (electronic)
26. Scott, L.R., Zhang, S.: Finite element interpolation of nonsmooth functions satisfying boundary conditions. Math. Comput. **54**(190), 483–493 (1990)
27. Starke, G.: An adaptive least-squares mixed finite element method for elasto-plasticity. SIAM J. Numer. Anal. **45**(1), 371–388 (2007)
28. Stevenson, R.: Optimality of a standard adaptive finite element method. Found. Comput. Math. **7**(2), 245–269 (2007)
29. Verfürth, R.: A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques. Wiley and Teubner, New York (1996)