Numer. Math. (2008) 108:487–500
DOI 10.1007/s00211-007-0121-y

Numerische Mathematik

A posteriori error estimates for mixed finite element approximations of elliptic problems

Mats G. Larson · Axel Målqvist

Received: 25 January 2005 / Revised: 25 September 2007 / Published online: 27 November 2007

Abstract We derive residual based a posteriori error estimates of the flux in L^2-norm for a general class of mixed methods for elliptic problems. The estimate is applicable to standard mixed methods such as the Raviart–Thomas–Nedelec and Brezzi–Douglas–Marini elements, as well as stabilized methods such as the Galerkin-Least squares method. The element residual in the estimate employs an elementwise computable postprocessed approximation of the displacement which gives optimal order.

Mathematics Subject Classification (2000) 65N30 · 65N15 · 65N12

1 Introduction

The model problem We consider the mixed formulation of the Poisson equation with Neumann boundary conditions:

$$\begin{cases} \boldsymbol{\sigma} - \nabla u = 0 & \text{in } \Omega, \\ -\nabla \cdot \boldsymbol{\sigma} = f & \text{in } \Omega, \\ \boldsymbol{n} \cdot \boldsymbol{\sigma} = 0 & \text{on } \Gamma, \end{cases} \tag{1.1}$$

where Ω is a polygonal domain in $\mathbf{R}^n$, $n = 2$ or 3 with boundary Γ. Assuming $\int_\Omega f\, dx = 0$, we get a well posed problem with a solution $u \in H^1(\Omega)/\mathbf{R}$ and

M. G. Larson (✉)
Department of Mathematics,
Umeå University, 90187 Umeå, Sweden
e-mail: mats.larson@math.umu.se

A. Målqvist
Division of Scientific Computing, Department of Information Technology,
Uppsala University, 75105 Uppsala, Sweden
e-mail: axel.malqvist@it.uu.se

$\boldsymbol{\sigma} \in \boldsymbol{V} = \{\boldsymbol{v} \in H(\text{div}, \Omega) : \boldsymbol{n} \cdot \boldsymbol{v} = 0 \text{ on } \Gamma\}$. See [9] for definitions of these function spaces.

Previous work Several works present a posteriori error estimates for mixed methods. In Carstensen [10] an error estimate in the $H(\text{div}, \Omega)$ norm of the flux is presented. The $H(\text{div}, \Omega)$ norm may be dominated by the div-part which is directly computable. When it comes to error estimates of the flux in L^2 norm of methods using richer spaces for the flux $\boldsymbol{\sigma}$ than the displacement u, such as Raviart–Thomas–Nedelec (RTN) elements, there are known difficulties. Braess and Verfürth presents a suboptimal estimate in [4]. The reason for the suboptimality is that the natural residual that arises from the first equation $\boldsymbol{\sigma} - \nabla u = 0$ in problem (1.1) may be large if the flux space is richer than the displacement space. In a recent paper Lovadina and Stenberg [12] derive an a posteriori error estimate of the L^2-norm of the flux for the RTN based methods which employs a particular postprocessed approximation U. The proof is based on a posteriori error analysis of an equivalent method which involves the postprocessed approximation U.

New contributions We derive a general a posteriori error estimate in the energy norm which is applicable to most mixed methods including the classical inf-sup stable elements, RTN elements and the Brezzi–Douglas–Marini (BDM) elements. Our estimate is closely related to the estimate presented by Lovadina and Stenberg [12], however, our proof is more general and also reveals the fact that one can use any piecewise polynomial approximation of the displacement when computing the residual. By a small adjustment of the argument we finally, derive an estimate for the stabilized mixed method of Masud and Hughes [13]. The same technique applies to other stabilized schemes, for instance the Galerkin least squares method.

Outline We start by presenting finite elements and the discrete version of equation (1.1) in Sect. 2 then we present the a posteriori error estimates in Sect. 3.

2 Weak formulation and the finite element method

Weak formulation We multiply the first equation in (1.1) by a test function $\boldsymbol{v} \in \boldsymbol{V}$ and integrate by parts. The second equation in (1.1) is multiplied by a test function $w \in W = L^2(\Omega)/\mathbf{R}$. The weak form reads: find $\boldsymbol{\sigma} \in \boldsymbol{V}$ and $u \in W$ such that,

$$\begin{cases} (\boldsymbol{\sigma}, \boldsymbol{v}) + (u, \nabla \cdot \boldsymbol{v}) = 0 & \text{for all } \boldsymbol{v} \in \boldsymbol{V}, \\ (-\nabla \cdot \boldsymbol{\sigma}, w) = (f, w) & \text{for all } w \in W. \end{cases} \tag{2.1}$$

Our aim is to derive a posteriori error estimates of finite element approximations $\{\boldsymbol{\Sigma}, U\}$ of the exact solution $\{\boldsymbol{\sigma}, u\}$ in the energy norm $\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0$, where $\|\cdot\|_0$ denotes the $L^2(\Omega)$ norm.

The mixed finite element method We let $\mathcal{K} = \{K\}$ be a partition of Ω into simplicial elements of diameter h_K and define the mesh function, $h(x) : \Omega \to \mathbf{R}^+$, by letting

$h(x) = h_K$ for $x \in K$. We assume that the elements are shape regular, i.e., there is a constant C such that $h_K/\rho_K \leq C$ for all $K \in \mathcal{K}$, where ρ_K is the diameter of the largest ball that can be inscribed in K. We seek an approximate solution in discrete spaces $\boldsymbol{V}_h \subset \boldsymbol{V}$ and $W_h \subset W$ defined on the partition $\mathcal{K}$. It is well known that for finite element methods based on the standard weak form (2.1) the discrete spaces must be chosen so that the inf-sup condition, see [9], is satisfied in order to guarantee a stable method. Only rather special constructions of the discrete spaces yield stable methods. In Sect. 3.4 we consider a stabilized mixed finite element method based on a modified weak formulation which can be based on standard continuous piecewise polynomials. We summarize some of the most well known choices of stable discrete spaces on triangles and tetrahedra for a given integer $k \geq 1$:

- *Raviart–Thomas–Nedelec* (RTN) elements, see [14,15],
  $\boldsymbol{V}_h = \{\boldsymbol{v} \in H(\text{div}, \Omega) : \boldsymbol{v}|_K \in [P_{k-1}(K)]^n \oplus \boldsymbol{x}\tilde{P}_{k-1}(K) \text{ for all } K \in \mathcal{K}\}$,
  $W_h = \{w \in L^2(\Omega)/\mathbf{R} : w|_K \in P_{k-1}(K) \text{ for all } K \in \mathcal{K}\}$.
- *Brezzi–Douglas–Marini* (BDM) elements, see [7,8],
  $\boldsymbol{V}_h = \{\boldsymbol{v} \in H(\text{div}, \Omega) : \boldsymbol{v}|_K \in [P_k(K)]^n \text{ for all } K \in \mathcal{K}\}$,
  $W_h = \{w \in L^2(\Omega)/\mathbf{R} : w|_K \in P_{k-1}(K) \text{ for all } K \in \mathcal{K}\}$.

Here $C(\Omega)$ denotes the space of continuous functions on Ω, $P_k(K)$ the space of polynomials of degree k on element K, and $\tilde{P}_k(K)$ the set of homogeneous polynomials of degree k. For a more complete account of these spaces and the inf-sup condition we refer to Brezzi–Fortin, [9]. Note, that our a posteriori error analysis does not use the inf-sup condition explicitly.

The mixed finite element method reads: find $\boldsymbol{\Sigma} \in \boldsymbol{V}_h$ and $U \in W_h$ such that:

$$\begin{cases} (\boldsymbol{\Sigma}, \mathbf{v}) + (U, \nabla \cdot \mathbf{v}) = 0 & \text{for all } \boldsymbol{v} \in \boldsymbol{V}_h, \\ (-\nabla \cdot \boldsymbol{\Sigma}, w) = (f, w) & \text{for all } w \in W_h. \end{cases} \tag{2.2}$$

3 A posteriori error estimates

3.1 Preliminaries

We use the following notation for the standard Sobolev norms, $\|\cdot\|_{s,\omega} = \|\cdot\|_{H^s(\omega)} = \|\cdot\|_{W_2^s(\omega)}$, see [1], and we let $(\cdot,\cdot)_\omega$ denote the $L^2(\omega)$ inner product. In the case $\omega = \Omega$ we simplify the notation and write $\|\cdot\|_{s,\Omega} = \|\cdot\|_s$ and $(\cdot,\cdot)_\Omega = (\cdot,\cdot)$. We shall also need suitable norms, see [18], on the element Sobolev spaces $H^1(K)$, $H^{1/2}(\partial K)$, and $H^{-1/2}(\partial K)$. For each element $K \in \mathcal{K}$ we define the following norms

$$|||v|||_{1,K}^2 = \|\nabla v\|_{0,K}^2 + h_K^{-2}\|v\|_{0,K}^2 \quad \text{for all } v \in H^1(K), \tag{3.1}$$

$$|||v|||_{1/2,\partial K}^2 = \inf_{\substack{\eta \in H^1(K) \\ \eta = v \text{ on } \partial K}} |||\eta|||_{1,K}^2 \quad \text{for all } v \in H^{1/2}(\partial K), \tag{3.2}$$

$$|||v|||_{-1/2,\partial K} = \sup_{w \in H^{1/2}(\partial K)} \frac{\langle v, w\rangle_{\partial K}}{|||w|||_{1/2,\partial K}} \quad \text{for all } v \in H^{-1/2}(\partial K), \tag{3.3}$$

where $\langle\cdot,\cdot\rangle_{\partial K}$ denotes the duality pairing between spaces $H^{-1/2}(\partial K)$ and $H^{1/2}(\partial K)$. With these definitions we clearly have the inequality

$$\langle v, w\rangle_{\partial K} \leq |||v|||_{-1/2,\partial K}|||w|||_{1/2,\partial K} \quad \text{for all } v \in H^{-1/2}(\partial K) \text{ and } w \in H^{1/2}(\partial K). \tag{3.4}$$

Furthermore, we have the following elementwise normal trace inequality.

Lemma 3.1 *The following trace inequality holds*

$$|||\boldsymbol{n}\cdot\boldsymbol{v}|||^2_{-1/2,\partial K} \leq C(\|\boldsymbol{v}\|^2_{0,K} + h_K^2\|\nabla\cdot\boldsymbol{v}\|^2_{0,K}), \tag{3.5}$$

for all $\boldsymbol{v} \in H(div, K)$ *with constant* C *independent of* h_K.

Proof To prove (3.5) we first derive the estimate on the reference element and then use a scaling argument to show uniformity in the size of the element h_K. Let $\widehat{K}$ be the reference element with boundary $\partial\widehat{K}$ and exterior unit normal $\widehat{\boldsymbol{n}}$. To prove the estimate on the reference element we employ Green's formula, see [11], to get the identity

$$\langle\widehat{\boldsymbol{n}}\cdot\widehat{\boldsymbol{v}}, \widehat{w}\rangle_{\partial\widehat{K}} = (\widehat{\boldsymbol{v}}, \nabla\widehat{w})_{\widehat{K}} + (\nabla\cdot\widehat{\boldsymbol{v}}, \widehat{w})_{\widehat{K}} \quad \text{for all } \widehat{\boldsymbol{v}} \in H(\text{div}, \widehat{K}) \quad \text{and} \quad \widehat{w} \in H^1(\widehat{K}). \tag{3.6}$$

Estimating the right hand side using Cauchy-Schwartz inequality we have

$$\langle\widehat{\boldsymbol{n}}\cdot\widehat{\boldsymbol{v}}, \widehat{w}\rangle_{\partial\widehat{K}} \leq (\|\widehat{\boldsymbol{v}}\|^2_{\widehat{K}} + h^2_{\widehat{K}}\|\nabla\cdot\widehat{\boldsymbol{v}}\|^2_{\widehat{K}})^{1/2}|||\widehat{w}|||_{\widehat{K}}, \tag{3.7}$$

where we divided and multiplied by a scaling factor of $h^{-2}_{\widehat{K}}$. Since (3.7) holds for all $\widehat{w} \in H^1(\widehat{K})$ we get, for any $\widehat{\mu} \in H^{1/2}(\partial\widehat{K})$,

$$\langle\widehat{\boldsymbol{n}}\cdot\widehat{\boldsymbol{v}}, \widehat{\mu}\rangle_{\partial\widehat{K}} \leq (\|\widehat{\boldsymbol{v}}\|^2_{\widehat{K}} + h^2_{\widehat{K}}\|\nabla\cdot\widehat{\boldsymbol{v}}\|^2_{\widehat{K}})^{1/2} \inf_{\substack{\widehat{w}\in H^1(\widehat{K})\\ \widehat{w}=\widehat{\mu} \text{ on } \partial\widehat{K}}} |||\widehat{w}|||_{\widehat{K}} \tag{3.8}$$

$$= (\|\widehat{\boldsymbol{v}}\|^2_{\widehat{K}} + h^2_{\widehat{K}}\|\nabla\cdot\widehat{\boldsymbol{v}}\|^2_{\widehat{K}})^{1/2}|||\widehat{\mu}|||_{1/2,\partial\widehat{K}}. \tag{3.9}$$

Dividing by $|||\widehat{\mu}|||_{1/2,\partial\widehat{K}}$ and taking the supremum over all $\widehat{\mu} \in H^{1/2}(\partial\widehat{K})$ we obtain the trace inequality (3.5) on the reference element $\widehat{K}$.

Given an element $K \in \mathcal{K}$, let $F : \mathbf{R}^n \to \mathbf{R}^n$ be an affine mapping of the form $F\widehat{\boldsymbol{x}} = \boldsymbol{b}+B\widehat{\boldsymbol{x}}$, where B is an invertible $n\times n$ matrix with positive determinant, $\boldsymbol{b} \in \mathbf{R}^n$, and $K = F\widehat{K}$. Furthermore, let $\|B\| = \sup_{\boldsymbol{x}\in\mathbf{R}^n} \|B\boldsymbol{x}\|/\|\boldsymbol{x}\|$ be the standard Euclidian matrix norm. We then have the following bounds $\|B\| \leq h_K/\rho_{\widehat{K}}$ and $\|B^{-1}\| \leq h_{\widehat{K}}/\rho_K$ and thus $\|B\|\,\|B^{-1}\| \leq C$ as a consequence of shape regularity.

Next we define a mapping $\mathcal{F} : H^1(\widehat{K}) \to H^1(K)$ by $v = \mathcal{F}\widehat{v} = \widehat{v}\circ F^{-1}$. We then have the estimate $|\widehat{v}|_{m,K} \leq C(\det B)^{-1/2}\|B\|^m|v|_{m,K}$, for $m = 0, 1$, see [9], and thus we conclude that

$$|||\,\widehat{v}\,|||_{1,\widehat{K}} \leq C(\det B)^{-1/2}\|B\|\;|||v|||_{1,K}. \tag{3.10}$$

Taking the appropriate infimum on both sides of (3.10) we obtain

$$|||\widehat{\mu}|||_{1/2,\partial\widehat{K}} \leq C(\det B)^{-1/2}\|B\|\;|||\mu|||_{1/2,\partial K}. \tag{3.11}$$

To transform vector valued functions we use the Piola transform $\mathcal{P} : H(\mathrm{div}, \widehat{K}) \to H(\mathrm{div}, K)$ defined by $\mathcal{P}\widehat{\boldsymbol{v}} = (\det B)^{-1} B\widehat{\boldsymbol{v}} \circ F^{-1}$. We then have the identity $\langle \boldsymbol{n} \cdot \boldsymbol{v}, w\rangle_{\partial K} = \langle \widehat{\boldsymbol{n}} \cdot \widehat{\boldsymbol{v}}, \widehat{w}\rangle_{\partial \widehat{K}}$ for $\boldsymbol{v} = \mathcal{P}\widehat{\boldsymbol{v}}$ and $w = \mathcal{F}\widehat{w}$. Furthermore, we have the estimates

$$\|\widehat{\boldsymbol{v}}\|_{\widehat{K}} \le (\det B)^{1/2} \|B^{-1}\| \, \|\boldsymbol{v}\|_K, \quad \|\nabla \cdot \widehat{\boldsymbol{v}}\|_{\widehat{K}} \le (\det B)^{1/2} \|\nabla \cdot \boldsymbol{v}\|_K, \tag{3.12}$$

see [9]. Using these results we obtain the following estimate

$$\begin{aligned} |||\boldsymbol{n} \cdot \boldsymbol{v}|||_{-1/2,\partial K} &= \sup_{\eta \in H^{1/2}(\partial K)} \frac{\langle \boldsymbol{n} \cdot \boldsymbol{v}, \eta\rangle_{\partial K}}{|||\eta|||_{1/2,\partial K}} && (3.13)\\ &\le C(\det B)^{-1/2} \|B\| \sup_{\widehat{\eta} \in H^{1/2}(\partial \widehat{K})} \frac{\langle \widehat{\boldsymbol{n}} \cdot \widehat{\boldsymbol{v}}, \widehat{\eta}\rangle_{\partial \widehat{K}}}{|||\widehat{\eta}|||_{1/2,\partial \widehat{K}}} && (3.14)\\ &\le C(\det B)^{-1/2} \|B\| |||\widehat{\boldsymbol{n}} \cdot \widehat{\boldsymbol{v}}|||_{-1/2,\partial \widehat{K}}. && (3.15) \end{aligned}$$

Finally, we can prove the desired estimate (3.5) on element K as follows

$$\begin{aligned} |||\boldsymbol{n} \cdot \boldsymbol{v}|||_{-1/2,\partial K} &\le C(\det B)^{-1/2} \|B\| |||\widehat{\boldsymbol{n}} \cdot \widehat{\boldsymbol{v}}|||_{-1/2,\partial \widehat{K}} && (3.16)\\ &\le C(\det B)^{-1/2} \|B\| (\|\widehat{\boldsymbol{v}}\|^2_{\widehat{K}} + h^2_{\widehat{K}} \|\nabla \cdot \widehat{\boldsymbol{v}}\|^2_{\widehat{K}})^{1/2} && (3.17)\\ &\le C(\det B)^{-1/2} \|B\| (\det B)^{1/2} \|B^{-1}\| \\ &\quad \times \left(\|\boldsymbol{v}\|^2_K + (h_{\widehat{K}} \|B\|)^2 \|\nabla \cdot \boldsymbol{v}\|^2_K\right)^{1/2} && (3.18)\\ &\le C(\|\boldsymbol{v}\|^2_K + h^2_K \|\nabla \cdot \boldsymbol{v}\|^2_K)^{1/2}, && (3.19) \end{aligned}$$

where we used: (3.15) in (3.16), the trace inequality (3.5) on the reference element $\widehat{K}$ in (3.17), the estimates for the Piola transform (3.12) in (3.18) together with the observation that $\|\nabla\cdot\boldsymbol{v}\| \le \|B^{-1}\| \, \|B\| \|\nabla\cdot\boldsymbol{v}\|$ since $1 = \|I\| \le \|B^{-1}\| \, \|B\|$, and finally in (3.19) we used the estimate $\|B\| \le h_K/\rho_{\widehat{K}}$ to conclude that $h_{\widehat{K}} \|B\| \le (h_{\widehat{K}}/\rho_{\widehat{K}}) h_K$. □

3.2 Estimate for standard mixed methods

Here we present a general a posteriori error estimate in the energy norm $\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0$ involving a piecewise polynomial function Q, which may be obtained by postprocessing U. The possibility to replace U by Q is important since it leads to a posteriori error estimates of optimal order. We are not interested in tracking the constants in the error estimates.

Theorem 3.1 *For arbitrary $Q \in \bigoplus_{K \in \mathcal{K}} P_l(K)$, with $l \ge 0$ and $f \in L^2(\Omega)$ it holds,*

$$\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|^2_0 \le C \sum_{K \in \mathcal{K}} \left(h^2_K \|f + \nabla \cdot \boldsymbol{\Sigma}\|^2_{0,K} + \|\boldsymbol{\Sigma} - \nabla Q\|^2_{0,K} + h^{-1}_K \| \, [Q] \, \|^2_{0,\partial K}\right), \tag{3.20}$$

where the jump denoted by $[\cdot]$ *is the difference in function value over a face in the mesh.*

Proof Starting with the left hand side we have

$$\begin{aligned}
\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 &= (\boldsymbol{\sigma} - \boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) && (3.21)\\
&= (\boldsymbol{\sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) && (3.22)\\
&= -(u, \nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma})) - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) && (3.23)\\
&= -(u - Q, \nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma})) + (Q, -\nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma})) - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) && (3.24)\\
&= (u - Q, f + \nabla \cdot \boldsymbol{\Sigma}) + \sum_{K \in \mathcal{K}} ((Q, -\nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma}))_K - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma})_K) && (3.25)\\
&= I + II. && (3.26)
\end{aligned}$$

We treat the two terms in equation (3.26) separately, beginning with I. From the second part of equation (2.2) we have the Galerkin orthogonality property $(f + \nabla \cdot \boldsymbol{\Sigma}, w) = 0$ for all $w \in W_h$. Let P_h denote the L^2 projection onto W_h. Using Galerkin orthogonality (2.2) to subtract the projection $P_h(u - Q) \in W_h$ of $(u - Q)$ followed by the projection error estimate $\|v - P_h v\|_{0,K} \le C h_K \|\nabla v\|_{0,K}$ we obtain

$$\begin{aligned}
I &\le |(f + \nabla \cdot \boldsymbol{\Sigma}, u - Q)| && (3.27)\\
&\le \|h(f + \nabla \cdot \boldsymbol{\Sigma})\|_0 \|h^{-1}(u - Q - P_h(u - Q))\|_0 && (3.28)\\
&\le C\|h(f + \nabla \cdot \boldsymbol{\Sigma})\|_0 \|\nabla(u - Q)\|_{0,\mathcal{K}} && (3.29)\\
&= C\|h(f + \nabla \cdot \boldsymbol{\Sigma})\|_0 \|\boldsymbol{\sigma} - \boldsymbol{\Sigma} + \boldsymbol{\Sigma} - \nabla Q\|_{0,\mathcal{K}} && (3.30)\\
&\le \frac{3C^2}{2}\|h(f + \nabla \cdot \boldsymbol{\Sigma})\|_0^2 + \frac{1}{4}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 + \frac{1}{2}\|\boldsymbol{\Sigma} - \nabla Q\|_{0,\mathcal{K}}^2. && (3.31)
\end{aligned}$$

Here, and below, $\|\cdot\|_{0,\mathcal{K}}$ denotes the broken L^2-norm defined by $\|v\|_{0,\mathcal{K}}^2 = \sum_{K \in \mathcal{K}} \|v\|_{0,K}^2$. We now turn to the second term II in equation (3.26) and start with integration by parts,

$$\begin{aligned}
II &= \sum_{K \in \mathcal{K}} ((Q, -\nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma}))_K - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma})_K) && (3.32)\\
&= \sum_{K \in \mathcal{K}} ((\nabla Q, \boldsymbol{\sigma} - \boldsymbol{\Sigma})_K - (Q, \boldsymbol{n} \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma}))_{\partial K} - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma})_K) && (3.33)\\
&= \sum_{K \in \mathcal{K}} (\nabla Q - \boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) - \langle Q, \boldsymbol{n} \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma})\rangle_{\partial K} && (3.34)\\
&\le \|\nabla Q - \boldsymbol{\Sigma}\|_{0,\mathcal{K}}^2 + \frac{1}{4}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 + \left|\sum_{K \in \mathcal{K}} \langle Q, \boldsymbol{n} \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma})\rangle_{\partial K}\right|. && (3.35)
\end{aligned}$$

Next we note that $\sum_{K\in\mathcal{K}}\langle v, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})\rangle_{\partial K} = \langle v, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})\rangle_{\partial\Omega} = 0$ for all $v\in H^1(\Omega)$, since $\boldsymbol{\sigma}-\boldsymbol{\Sigma}\in\boldsymbol{V}$. This identity follows from using Green's formula and decomposing the integral over Ω into a sum of integrals over the elements and then using Green's formula again. Thus we can subtract an arbitrary function $v\in H^1(\Omega)$ in the term $\sum_{K\in\mathcal{K}}\langle Q, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})\rangle_{\partial K} = \sum_{K\in\mathcal{K}}\langle Q-v, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})\rangle_{\partial K}$. We then have the estimate

$$II \le \|\nabla Q-\boldsymbol{\Sigma}\|_{0,\mathcal{K}}^2+\frac{1}{4}\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2+\left|\inf_{v\in H^1(\Omega)}\sum_{K\in\mathcal{K}}\langle Q-v, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})\rangle_{\partial K}\right|. \quad (3.36)$$

We now use inequality (3.4) followed by the trace inequality (3.5), to estimate the sum in equation (3.36) as follows

$$\left|\inf_{v\in H^1(\Omega)}\sum_{K\in\mathcal{K}}\langle Q-v, \boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})\rangle_{\partial K}\right| \quad (3.37)$$

$$\le \inf_{v\in H^1(\Omega)}\sum_{K\in\mathcal{K}} |||Q-v|||_{1/2,\partial K}|||\boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})|||_{-1/2,\partial K} \quad (3.38)$$

$$\le \inf_{v\in H^1(\Omega)}\left(\sum_{K\in\mathcal{K}}|||Q-v|||_{1/2,\partial K}^2\right)^{1/2}\left(\sum_{K\in\mathcal{K}}|||\boldsymbol{n}\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})|||_{-1/2,\partial K}^2\right)^{1/2} \quad (3.39)$$

$$\le C\inf_{v\in H^1(\Omega)}\left(\sum_{K\in\mathcal{K}}|||Q-v|||_{1/2,\partial K}^2\right)^{1/2}$$

$$\times\left(\sum_{K\in\mathcal{K}}\left(\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_{0,K}^2+h_K^2\|\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\Sigma})\|_{0,K}^2\right)\right)^{1/2} \quad (3.40)$$

$$\le C\inf_{v\in H^1(\Omega)}\left(\sum_{K\in\mathcal{K}}|||Q-v|||_{1/2,\partial K}^2\right)^{1/2}\left(\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2+\|h(f+\nabla\cdot\boldsymbol{\Sigma})\|_0^2\right)^{1/2} \quad (3.41)$$

$$\le \frac{3C^2}{2}\inf_{v\in H^1(\Omega)}\sum_{K\in\mathcal{K}}|||Q-v|||_{1/2,\partial K}^2+\frac{1}{4}\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2+\frac{1}{2}\|h(f+\nabla\cdot\boldsymbol{\Sigma})\|_0^2. \quad (3.42)$$

Together equation (3.36) and equations (3.37–3.42) give a bound of the second term, II, in equation (3.26),

$$II \le \|\nabla Q-\boldsymbol{\Sigma}\|_{0,\mathcal{K}}^2+\frac{1}{2}\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2+\frac{1}{2}\|h(f+\nabla\cdot\boldsymbol{\Sigma})\|_0^2$$
$$+\frac{3C^2}{2}\inf_{v\in H^1(\Omega)}\sum_{K\in\mathcal{K}}|||Q-v|||_{1/2,\partial K}^2. \quad (3.43)$$

We combine equation (3.31) and equation (3.43) to get,

$$
\begin{aligned}
I + II \leq \frac{3}{2}\|\nabla Q - \boldsymbol{\Sigma}\|_{0,\mathcal{K}}^2 + \frac{3}{4}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 + C\|h(f + \nabla \cdot \boldsymbol{\Sigma})\|_0^2 \\
+ \frac{3C^2}{2} \inf_{v \in H^1(\Omega)} \sum_{K \in \mathcal{K}} |||Q - v|||_{1/2,\partial K}^2.
\end{aligned} \tag{3.44}
$$

To estimate the last term on the right hand side in equation (3.44) we employ the technique of Lemma 4 in [3]. For completeness we include the details of the proof. We let $\mathcal{N}$ be the set of nodes in the mesh, $\{\phi\}_{i\in\mathcal{N}}$ be the lowest order Lagrange basis functions, $\omega_i = \operatorname{supp}(\phi_i)$, $CP_{l,i} = \{v \in C(\omega_i) : v|_K \in P_l(K)$, for all $K \in \mathcal{K}$ with $K \subset \omega_i\}$, i.e., continuous piecewise polynomials of degree l on ω_i, and $CP_l = \oplus_{i\in\mathcal{N}}\phi_i CP_{l,i} \subset H^1(\Omega)$.

Using that $CP_l \subset H^1(\Omega)$ followed by the inverse inequality $|||Q - v|||_{1/2,\partial K}^2 \leq Ch_K^{-1}\|Q - v\|_{0,\partial K}^2$, which holds since both v and Q are piecewise polynomials, we get

$$
\begin{aligned}
\inf_{v \in H^1(\Omega)} \sum_{K \in \mathcal{K}} |||Q - v|||_{1/2,\partial K}^2 &\leq \inf_{v \in CP_l} \sum_{K \in \mathcal{K}} |||Q - v|||_{1/2,\partial K}^2 \\
&\leq C \inf_{v \in CP_l} \sum_{K \in \mathcal{K}} h_K^{-1}\|Q - v\|_{0,\partial K}^2.
\end{aligned} \tag{3.45}
$$

We write $v = \sum_{i\in\mathcal{N}} \phi_i v_i \in CP_l$ and proceed with the estimate as follows

$$
\|Q - v\|_{0,\partial K}^2 = \sum_{i\in\mathcal{N}} (Q - v, \phi_i(Q - v_i))_{\partial K}, \tag{3.46}
$$

$$
\leq \sum_{i\in\mathcal{N}} \|\phi_i^{1/2}(Q - v)\|_{0,\partial K}\|\phi_i^{1/2}(Q - v_i)\|_{0,\partial K}, \tag{3.47}
$$

$$
\leq \|Q - v\|_{0,\partial K} \left(\sum_{i\in\mathcal{N}} \|\phi_i^{1/2}(Q - v_i)\|_{0,\partial K}^2 \right)^{1/2}, \tag{3.48}
$$

where we used that $\{\phi_i\}_{i\in\mathcal{N}}$ is a partition of unity. Dividing inequality (3.48) by $\|Q - v\|_{0,\partial K}$ and combining with estimate (3.45) we arrive at

$$
\inf_{v \in H^1(\Omega)} \sum_{K \in \mathcal{K}} |||Q - v|||_{1/2,\partial K}^2 \leq \sum_{i\in\mathcal{N}} \inf_{v_i \in CP_{l,i}} \sum_{K \in \mathcal{K}} h_K^{-1}\|\phi_i^{1/2}(Q - v_i)\|_{0,\partial K}^2. \tag{3.49}
$$

Next we employ the following inequality

$$
\inf_{v_i \in CP_{l,i}} \sum_{K \in \mathcal{K}} h_K^{-1}\|\phi_i^{1/2}(Q - v_i)\|_{0,\partial K}^2 \leq C \sum_{K \in \mathcal{K}} h_K^{-1}\|\phi_i^{1/2}\,[Q]\,\|_{0,\partial K}^2. \tag{3.50}
$$

We return to the proof of this inequality below. Together (3.49) and (3.50) give

$$\inf_{v\in H^1(\Omega)} \sum_{K\in\mathcal{K}} |||Q-v|||^2_{1/2,\partial K} \le C \sum_{K\in\mathcal{K}} \sum_{i\in\mathcal{N}} h_K^{-1} \|\phi_i^{1/2}\,[Q]\,\|^2_{0,\partial K} = C \sum_{K\in\mathcal{K}} h_K^{-1} \|\,[Q]\,\|^2_{0,\partial K}, \tag{3.51}$$

again we used that $\{\phi_i\}_{i\in\mathcal{N}}$ is a partition of unity.

Combining equation (3.43) and (3.51) we get

$$I+II \le \frac{3}{2}\|\nabla Q-\boldsymbol{\Sigma}\|^2_{0,\mathcal{K}}+\frac{3}{4}\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|^2_0+C\|h(f+\nabla\cdot\boldsymbol{\Sigma})\|^2_0+C\sum_{K\in\mathcal{K}} h_K^{-1}\|\,[Q]\,\|^2_{0,\partial K}. \tag{3.52}$$

Since $I + II = \|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2$ from equations (3.21–3.26) we just need to subtract $3/4\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2$ from both sides of equation (3.52) to prove the theorem.

It remains to prove inequality (3.50). We first note that it follows from shape regularity that there are constants c and C such that $c \le h_K/h_{K'} \le C$ for all elements $K, K' \in \mathcal{K}$ such that $K, K' \subset \omega_i$. We may therefore conclude that there is an h_i and constants c and C such that $ch_i \le h_K \le Ch_i$ for all $K \in \mathcal{K}$ with $K \subset \omega_i$. Next we observe that it also follows from shape regularity that there is only a finite number, say N_c, of possible element configurations in ω_i. Let $\widehat{\omega}_j$, $j = 1, \ldots, N_c$ be the corresponding reference configurations and let $j(i)$ denote the index of the reference configuration corresponding to patch ω_i. Let $F_{\omega_i} : \widehat{\omega}_{j(i)} \to \omega_i$ be a C^0-diffeomorphism such that $F_{\omega_i}|_{\widehat{K}}$ is affine. Let $\mathcal{E}$ denote the set of edges or faces in the mesh. For each $E \in \mathcal{E}$, $E \subset \partial K$, $K \subset \omega_i$ we let $F_{\omega_i,E} = F_{\omega_i}|_{\widehat{E}}$ with corresponding Jacobian $B_{\omega_i,E}$. Using shape regularity it follows that there are constants c and C such that $ch_i^{n-1} \le \det B_{\omega_i,E} \le Ch_i^{n-1}$ (where n is the spatial dimension).

With these preparations accomplished we begin with the estimate as follows

$$\sum_{K\in\mathcal{K}} h_K^{-1}\|\phi_i^{1/2}(Q-v_i)\|^2_{0,\partial K} = \sum_{K\in\mathcal{K}} h_K^{-1} \sum_{E\in\mathcal{E},E\subset\partial K} \|\phi_i^{1/2}(Q-v_i)\|^2_{0,E} \tag{3.53}$$

$$= \sum_{K\in\mathcal{K}} h_K^{-1} \sum_{E\in\mathcal{E},E\subset\partial K} \det B_{\omega_i,E}\|\widehat{\phi}_i^{1/2}(\widehat{Q}-\widehat{v}_i)\|^2_{0,\widehat{E}} \tag{3.54}$$

$$\le Ch_i^{n-2}\sum_{K\in\mathcal{K}} \sum_{E\in\mathcal{E},E\subset\partial K} \|\widehat{\phi}_i^{1/2}(\widehat{Q}-\widehat{v}_i)\|^2_{0,\widehat{E}}, \tag{3.55}$$

where we splitted the integral over the element boundaries to edge or face contributions in (3.53), mapped to the reference configuration in (3.54), and used the estimates $h_K^{-1} \le Ch_i^{-1}$ and $\det B_{\omega_i,E} \le Ch_i^{n-1}$ in (3.55).

Next we have the following inequality

$$\inf_{\widehat{v}_i \in \widehat{CP}_{l,i}} \sum_{K\in\mathcal{K}} \sum_{E\in\mathcal{E}, E\subset\partial K} \|\widehat{\phi}_i^{1/2}(\widehat{Q}-\widehat{v}_i)\|^2_{0,\widehat{E}} \le C \sum_{K\in\mathcal{K}} \sum_{E\in\mathcal{E}, E\subset\partial K} \|\widehat{\phi}_i^{1/2}[\widehat{Q}]\|^2_{0,\widehat{E}}, \tag{3.56}$$

for all $\widehat{Q} \in \bigoplus_{K\in\mathcal{K}, K\subset\omega_i} P_l(\widehat{K})$ and $i \in \mathcal{N}$. To prove (3.56) we observe that if the right hand side is zero then $\widehat{Q}$ is continuous on $\widehat{\omega}_i$ and thus we can take $\widehat{v}_i = \widehat{Q}$ and therefore the left hand side is also zero and then the inequality follows by using finite dimensionality of the discrete space $\bigoplus_{K\in\mathcal{K}, K\subset\omega_i} P_l(\widehat{K})$ together with the fact that there are only a finite number N_c of reference configurations.

Using (3.53–3.55) together with (3.56) and mapping back to the actual configuration we obtain

$$\sum_{K\in\mathcal{K}} h_K^{-1}\|\phi_i^{1/2}(Q-v_i)\|^2_{0,\partial K} \le Ch_i^{n-2} \sum_{K\in\mathcal{K}} \sum_{E\in\mathcal{E}, E\subset\partial K} \|\widehat{\phi}_i^{1/2}[\widehat{Q}]\|^2_{0,\widehat{E}} \tag{3.57}$$

$$\le Ch_i^{n-2} \sum_{K\in\mathcal{K}} \sum_{E\in\mathcal{E}, E\subset\partial K} (\det B_{\omega_i,E})^{-1}\|\phi_i^{1/2}[Q]\|^2_{0,E} \tag{3.58}$$

$$\le C \sum_{K\in\mathcal{K}} h_K^{-1} \sum_{E\in\mathcal{E}, E\subset\partial K} \|\phi_i^{1/2}[Q]\|^2_{0,E}, \tag{3.59}$$

where we finally used the estimates $(\det B_{\omega_i,E})^{-1} \le Ch_i^{-(n-1)}$ and $h_i \le Ch_K$ in (3.59). This estimate concludes the proof of (3.50). □

3.3 Estimate based on postprocessing

We now turn to the question of how to choose Q in Theorem 3.1. We know that choosing $Q = U$ results in a suboptimal estimate of the energy norm error, [4]. A natural idea is to choose Q to be a postprocessed version of U. There have been several works [5,8,12,16,17] following Arnold and Brezzi [2], published in the mid eighties, on postprocessing methods where information from the calculated flux $\mathbf{\Sigma}$ is used to compute an improved approximation of u.

We focus on the method considered in Lovadina and Stenberg [12] and show that Theorem 3.1 directly gives the estimate presented in [12]. We denote the postprocessed version of U by U^*. To define U^* we introduce some notations. For all $K \in \mathcal{K}$ we let $P_{h,K} : L^2(\Omega) \to W_{h,K}$ be the L^2 projection onto $W_{h,K}$, where $W_{h,K}$ is the restriction of W_h onto K. Furthermore, we let $W^*_{h,K}$ denote the following spaces: $W^*_{h,K} = P_k(K)$ for RTN elements and $W^*_{h,K} = P_{k+1}(K)$ for BDM elements.

Definition 3.1 (*Postprocessing method*) Find U^* such that $U^*|_K = U^*_K \in W^*_{h,K}$ where U^*_K is defined by

$$P_{h,K}U^*_K = U|_K, \tag{3.60}$$

and

$$(\nabla U^*, \nabla v)_K = (\mathbf{\Sigma}, \nabla v)_K \quad \text{for all } v \in (I - P_{h,K})W^*_{h,K}. \tag{3.61}$$

To compute U^* we need to solve one small problem for each element and thus the total cost is very low.

Corollary 3.1 *Given $f \in L^2(\Omega)$ it holds,*

$$\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 \le C \sum_{K \in \mathcal{K}} \left(h_K^2 \|f + \nabla \cdot \boldsymbol{\Sigma}\|_{0,K}^2 + \|\boldsymbol{\Sigma} - \nabla U^*\|_{0,K}^2 + h_K^{-1} \| \left[U^* \right] \|_{0,\partial K}^2 \right), \tag{3.62}$$

where U^* is taken from Definition 3.1.

Proof The proof follows directly from Theorem 3.1 with $Q = U^*$. □

Remark 3.1 In Corollary 2.8 on page 1667 in [12] the following a priori estimate of the error is presented for BDM and RTN elements,

$$\begin{aligned} \|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\| + \|u - U^*\|_{1,\mathcal{K}} &\le C h^{k+1} |u|_{k+2} \quad \text{for BDM}, \\ \|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\| + \|u - U^*\|_{1,\mathcal{K}} &\le C h^{k} |u|_{k+1} \quad \text{for RTN}, \end{aligned} \tag{3.63}$$

where $|\cdot|_k$ is the $H^k(\Omega)$ semi norm, see [1], and $\|\cdot\|_{1,\mathcal{K}}$ is the broken H^1 norm defined by $\|v\|_{1,\mathcal{K}}^2 = \sum_{K \in \mathcal{K}} \|v\|_{1,K}^2$. These estimates show that the postprocessed function U^* gives optimal order estimates. Further in Theorem 3.1. on the same page in [12] the critical term in the error estimate we present in Corollary 3.1, $\|\boldsymbol{\Sigma} - \nabla U^*\|_K^2$, is also proven to be of optimal order.

3.4 Estimate for stabilized methods

Here we extend our estimate to stabilized mixed methods, in particular, we consider the recent method presented in Masud and Hughes [13]. Stabilized methods are based on a modified weak formulation which yields a stable method for standard continuous piecewise polynomial approximations, e.g. piecewise linear functions for both displacement and flux.

The method reads: find $\boldsymbol{\Sigma} \in \boldsymbol{V}_h = \{\boldsymbol{v} \in [C(\Omega)]^n : \boldsymbol{v}|_K \in [P_k(K)]^n \text{ for all} K \in \mathcal{K}\}$ and $U \in W_h = \{v \in C(\Omega) : v|_K \in P_l(K) \text{ for all } K \in \mathcal{K}\}$, with $k, l \ge 1$ such that,

$$(-\nabla \cdot \boldsymbol{\Sigma}, w) + (\boldsymbol{\Sigma}, \boldsymbol{v}) + (U, \nabla \cdot \boldsymbol{v}) - \frac{1}{2}(\boldsymbol{\Sigma} - \nabla U, \boldsymbol{v} + \nabla w) = (f, w), \tag{3.64}$$

for all $\boldsymbol{v} \in \boldsymbol{V}_h$ and $w \in W_h$. Note, in particular, that in this method the order of polynomials in W_h may be higher than in $\boldsymbol{V}_h$ and thus in that case we do not expect post processing of the pressure to be necessary. Applying the same ideas as in Theorem 3.1 to this stabilized method we obtain the following a posteriori error estimate. The argument may be modified to cover other stabilized methods such as the Galerkin least squares method.

Proposition 3.1 *For the stabilized Galerkin method defined by (3.64) the following a posteriori error estimate holds,*

$$\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 \leq C \sum_{K \in \mathcal{K}} \left(h_K^2 \|f + \nabla \cdot \boldsymbol{\Sigma}\|_{0,K}^2 + \|\boldsymbol{\Sigma} - \nabla U\|_{0,K}^2 \right). \tag{3.65}$$

Proof Using the same arguments as in equations (3.21–3.25) in the proof of Theorem 3.1, we obtain the following error representation formula,

$$\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 = (u - Q, f + \nabla \cdot \boldsymbol{\Sigma}) + (Q, -\nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma})) - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma}) \tag{3.66}$$

$$= (u - Q, f + \nabla \cdot \boldsymbol{\Sigma}) + \sum_{K \in \mathcal{K}} (Q, -\nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Sigma}))_K - (\boldsymbol{\Sigma}, \boldsymbol{\sigma} - \boldsymbol{\Sigma})_K \tag{3.67}$$

$$= I + II. \tag{3.68}$$

Using the same technique as in the proof of Theorem 3.1, i.e. by combining equations (3.43) and (3.51), we can estimate the second term II as follows

$$II \leq \|\nabla Q - \boldsymbol{\Sigma}\|_{0,\mathcal{K}}^2 + \frac{1}{2}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 + \frac{1}{2}\|h(f + \nabla \cdot \boldsymbol{\Sigma})\|_0^2 + C \sum_{K \in \mathcal{K}} h_K^{-1} \|[Q]\|_{0,\partial K}^2. \tag{3.69}$$

We turn to the first term I in (3.68). We let $\pi_h : L^2(\Omega) \rightarrow W_h$ be the Scott–Zhang interpolant, see [6], and note, by letting $\boldsymbol{v} = 0$ in (3.64), that we have the Galerkin orthogonality property

$$(f + \nabla \cdot \boldsymbol{\Sigma}, w) = -\frac{1}{2}(\boldsymbol{\Sigma} - \nabla U, \nabla w), \tag{3.70}$$

for all $w \in W_h$. By choosing $w = \pi_h(u - Q)$ in equation (3.70) we get the following identity,

$$(f + \nabla \cdot \boldsymbol{\Sigma}, \pi_h(u - Q)) = -\frac{1}{2}(\boldsymbol{\Sigma} - \nabla U, \nabla \pi_h(u - Q)). \tag{3.71}$$

We then have,

$$I = (u - Q, f + \nabla \cdot \boldsymbol{\Sigma}) \tag{3.72}$$

$$= (u - Q - \pi_h(u - Q), f + \nabla \cdot \boldsymbol{\Sigma}) - \frac{1}{2}(\boldsymbol{\Sigma} - \nabla U, \nabla \pi_h(u - Q)), \tag{3.73}$$

using equation (3.71). In the first term we split the contributions over the elements and use the Cauchy–Schwarz inequality. In the second term we only use the Cauchy–Schwarz inequality,

$$I \leq \sum_{K \in \mathcal{K}} \|h^{-1}(u - Q - \pi_h(u - Q))\|_{0,K} \|h(f + \nabla \cdot \boldsymbol{\Sigma})\|_{0,K}$$
$$+ \frac{1}{2}\|\boldsymbol{\Sigma} - \nabla U\|_0 \|\nabla \pi_h(u - Q)\|_0. \tag{3.74}$$

By using a standard interpolation error estimate for the Scott-Zhang interpolant we get,

$$I \leq C \sum_{K\in\mathcal{K}} \|\nabla(u-Q)\|_{0,K}\|h(f+\nabla\cdot\boldsymbol{\Sigma})\|_{0,K}+\frac{1}{2}\|\boldsymbol{\Sigma}-\nabla U\|_0\|\nabla\pi_h(u-Q)\|_0. \quad (3.75)$$

We proceed by using the Cauchy-Schwarz inequality for sums and the identification $\boldsymbol{\sigma} = \nabla u$,

$$I \leq C\left(\sum_{K\in\mathcal{K}} \|\boldsymbol{\sigma} - \nabla Q\|_{0,K}^2\right)^{1/2}\left(\sum_{K\in\mathcal{K}} \|h(f + \nabla\cdot\boldsymbol{\Sigma})\|_{0,K}^2\right)^{1/2}$$
$$+\frac{1}{2}\|\boldsymbol{\Sigma} - \nabla U\|_0\|\nabla\pi_h(u - Q)\|_0. \quad (3.76)$$

Subtracting and adding $\boldsymbol{\Sigma}$ to the first term and using the inequality $ab \leq a^2/(2\epsilon) + \epsilon b^2/2$ we may choose ϵ such that,

$$I \leq \frac{1}{4}\|\boldsymbol{\Sigma} - \nabla Q\|_{0,\mathcal{K}}^2 + \frac{1}{4}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 + C\sum_{K\in\mathcal{K}} \|h(f + \nabla\cdot\boldsymbol{\Sigma})\|_{0,K}^2$$
$$+\frac{1}{2}\|\boldsymbol{\Sigma} - \nabla U\|_0\|\nabla\pi_h(u - Q)\|_0. \quad (3.77)$$

At this point we collect all terms from the estimates (3.77) of I and (3.69) of II,

$$I + II \leq C\sum_{K\in\mathcal{K}}\left(\|h(f + \nabla\cdot\boldsymbol{\Sigma})\|_{0,K}^2 + h_K^{-1}\|[Q]\|_{0,\partial K}^2\right) \quad (3.78)$$

$$+\frac{5}{4}\|\boldsymbol{\Sigma} - \nabla Q\|_{0,\mathcal{K}}^2 + \frac{3}{4}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 + +\frac{1}{2}\|\boldsymbol{\Sigma} - \nabla U\|_0\|\nabla\pi_h(u - Q)\|_0. \quad (3.79)$$

Next we choose $Q = U$ and observe that the jump terms vanish since U is continuous. We also take advantage of the fact that π_h is stable in H^1,

$$I + II \leq C\sum_{K\in\mathcal{K}} \|h(f + \nabla\cdot\boldsymbol{\Sigma})\|_{0,K}^2 \quad (3.80)$$

$$+\frac{5}{4}\|\boldsymbol{\Sigma} - \nabla U\|_{0,\mathcal{K}}^2 + \frac{3}{4}\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2 + C\|\boldsymbol{\Sigma} - \nabla U\|_0\|\boldsymbol{\sigma} - \nabla U\|_0. \quad (3.81)$$

Again we add and subtract $\boldsymbol{\Sigma}$ in the last term and use the same trick as above to get,

$$\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2 = I+II \leq C\sum_{K\in\mathcal{K}} \|h(f+\nabla\cdot\boldsymbol{\Sigma})\|_{0,K}^2+C\|\boldsymbol{\Sigma}-\nabla U\|_0^2+\frac{7}{8}\|\boldsymbol{\sigma}-\boldsymbol{\Sigma}\|_0^2, \quad (3.82)$$

and thus the proposition follows immediately by subtracting $7/8\|\boldsymbol{\sigma} - \boldsymbol{\Sigma}\|_0^2$ from both sides. □

Acknowledgments The authors wish to thank the referees for careful reading of the manuscript and helpful advice that resulted in significant improvements of this paper.

References

1. Adams, R.A.: Sobolev spaces. Pure and Applied Mathematics, vol. 65. Academic, New York (1975)
2. Arnold, D.N., Brezzi, F.: Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. RAIRO Modél. Math. Anal. Numér. **19**, 7–32 (1985)
3. Becker, R., Hansbo, P., Larson, M.G.: Energy norm a posteriori error estimation for discontinuous Galerkin methods. Comput. Methods Appl. Mech. Eng. **192**, 723–733 (2003)
4. Braess, D., Verfürt, R.: A posteriori error estimators for the Raviart–Thomas element. SIAM J. Numer. Anal. **33**, 2431–2444 (1996)
5. Bramble, J.H., Xu, J.: A local post-processing technique for improving the accuracy in mixed finite-element approximations. SIAM J. Numer. Anal. **26**, 1267–1275 (1989)
6. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods. Springer, Heidelberg (1994)
7. Brezzi, F., Douglas, J. Jr.., Durán, R., Fortin, M.: Mixed finite elements for second order elliptic problems in three variables. Numer. Math. **51**, 237–250 (1987)
8. Brezzi, F., Douglas, J. Jr.., Marini, D.L.: Two families of mixed finite elements for second order elliptic problems. Numer. Math. **47**, 217–235 (1985)
9. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer, Heidelberg (1991)
10. Carstensen, C.: A posteriori error estimate for the mixed finite element method. Math. Comp. **66**, 465–476 (1997)
11. Girault, V., Raviart, P.-A.: Finite Element Approximation of the Navier–Stokes Equation. Springer, Berlin (1979)
12. Lovadina, C., Stenberg, R.: Energy norm a posteriori error estimates for mixed finite element methods. Math. Comp. **75**, 1659–1674 (2006)
13. Masud, A., Hughes, T.J.R.: A stabilized finite element method for Darcy flow. Comput. Methods Appl. Mech. Eng. **191**, 4341–4370 (2002)
14. Nédélec, J.-C.: A new family of mixed finite elements in $\mathbf{R}^3$. Numer. Math. **50**, 57–81 (1986)
15. Raviart, P., Thomas, J.: A mixed finite element method for second order elliptic problems. In: Mathematical aspects of the finite element method. Lecture Notes in Math, vol. 606, pp. 292–315. Springer, Heidelberg (1977)
16. Stenberg, R.: Some new families of finite elements for the Stokes equations. Numer. Math. **56**, 827–838 (1990)
17. Stenberg, R.: Postprocessing schemes for some mixed finite elements. RAIRO Modél. Math. Anal. Numér. **25**, 151–167 (1991)
18. Roberts, J.E., Thomas, J.-M.: Mixed and hybrid methods. Ciarlet, P.G., Lions, J.L. (eds) Handbook of Numerical Analysis. Finite Element Methods (Part 1), vol. II, pp. 523–639. Elsevier, Amsterdam (1991)