*IMA Journal of Numerical Analysis* (2014) **34**, 1–27
doi:10.1093/imanum/drs048
Advance Access publication on May 3, 2013

# Quasi-optimal *a priori* estimates for fluxes in mixed finite element methods and an application to the Stokes–Darcy coupling

J. M. Melenk
*Institut für Analysis und Scientific Computing, TU Wien, Wiedner Hauptstrasse 8-10, A-1040 Wien, Austria*
melenk@tuwien.ac.at

and

H. Rezaijafari and B. Wohlmuth*
*M2 Zentrum Mathematik, Technische Universität München, Boltzmannstraße 3, D-85748 Garching, Germany*
*Corresponding author: rezaijafari@ma.tum.de wohlmuth@ma.tum.de

[Received on 2 February 2012; revised on 28 September 2012]

We show improved *a priori* convergence results in the $L^2$ norm on interfaces for the approximation of the normal component of the flux in mixed finite element methods. Compared with standard estimates for this problem class, additional factors of $\sqrt{h}|\log h|$ for the lowest-order case and of $\sqrt{h}$ in the higher-order case in the *a priori* bound for the flux variable are obtained. An important role in the analysis play new error estimates in strips of width $\mathcal{O}(h)$ and the use of anisotropic and weighted norms. Numerical examples including an application to the Stokes–Darcy coupling illustrate our theoretical results.

*Keywords*: anisotropic norms; local FEM error analysis; mixed finite elements; saddle point problem; Stokes–Darcy coupling; weighted norms.

## 1. Introduction

An important goal of many simulations in applications is to obtain accurate and reliable values for the normal flux across certain interfaces or the boundary of the domain. As an example, we mention that the treatment of complex problems in physics or engineering quite often requires the use of a variety of models in different parts of the computational domain, which in turn are coupled through the normal flux across common interfaces. On the level of numerical methods, this entails a need to understand and quantify the discretization error in the normal flux at interfaces. In the present paper, we study this question, taking the Poisson problem in mixed form as our model problem. Our setting is motivated by more complex problems in porous media applications such as the well-known Stokes–Darcy coupling problem. There, discretizations that are (locally) conservative are of a particular interest, and one such class is mixed finite element methods (FEMs). An attractive feature of mixed FEM is that, in contrast to the popular, well-established finite volume schemes, methods of arbitrary order are available.

In numerical methods that are based on a primal–dual formulation, the normal flux at an interface can be extracted directly from the flux variable. The errors in the primal and dual variables are linked to each other, and the standard saddle point theory as described, for example, in Nicolaides (1982) or Brezzi & Fortin (1991) leads to *a priori* estimates for the flux variable in the $L^2$ norm on an interface which are at most of order $l - \frac{1}{2}$, where $l$ is the order of the flux error in the $L^2$ norm on the domain.

However, the *best approximation* error for the normal flux in mixed FEMs is typically better by a factor $\sqrt{h}$. It is this gap in the *a priori* analysis that the present paper removes (up to a logarithmic factor in the lowest-order case). We mention at this point that this improved estimate is fairly easily achievable if optimal-order estimates in $L^\infty$ are available; however, this requires significantly more regularity than the present analysis.

In view of the technical nature of the paper, in Section 2 we formulate our model problem and state the main result which yields quasi-optimal *a priori* error estimates for the normal flux. The remainder of the paper is devoted to the proofs of the *a priori* bound and to numerical results. In Section 3, we introduce a suitable anisotropic norm and a dual problem with right-hand sides that are supported in a strip of width $\mathcal{O}(h)$ near the interface. Section 3 also discusses the regularity properties of the solutions of these dual problems. Section 4 quantifies the approximation properties of the Fortin operator in these anisotropic norms. In Section 5, the convergence analysis for the dual problems is given, and the proof of the main result, Theorem 2.3, is presented. Finally, in Section 6, we provide numerical results including the application to a Stokes–Darcy coupling.

## 2. Problem formulation and main results

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ be a convex and bounded polyhedral domain and let $f \in L^2(\Omega)$. We consider the model problem

$$-\Delta u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega$$

in its saddle point formulation based on $H(\text{div}; \Omega)$ and $L^2(\Omega)$, where

$$H(\text{div}; \Omega) := \{\tau \in L^2(\Omega)^d, \ \text{div}\, \tau \in L^2(\Omega)\}.$$

We state the saddle point formulation: find $(\sigma, u) \in H(\text{div}; \Omega) \times L^2(\Omega)$ such that

$$a(\sigma, \tilde{\sigma}) + b(\tilde{\sigma}, u) = 0, \qquad \tilde{\sigma} \in H(\text{div}; \Omega), \tag{2.1a}$$

$$b(\sigma, \tilde{u}) = -(f, \tilde{u})_0, \qquad \tilde{u} \in L^2(\Omega), \tag{2.1b}$$

where the bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ are given, for $\tau, \tilde{\tau} \in H(\text{div}; \Omega)$ and $v \in L^2(\Omega)$, by

$$a(\tau, \tilde{\tau}) := \int_\Omega \tau \cdot \tilde{\tau}\, \mathrm{d}x, \quad b(\tau, v) := \int_\Omega \text{div}\, \tau v\, \mathrm{d}x.$$

The saddle point formulation is well posed (Brezzi & Fortin, 1991, Section IV.1.2). We note that in contrast to the primal weak formulation, the homogeneous Dirichlet boundary conditions do not enter into the definition of the spaces.

For integer $k \in \mathbb{N}_0$, Sobolev norms on $\Omega$ are denoted by $\|\cdot\|_k$; the seminorm $k \geqslant 1$ is denoted by $|\cdot|_k$. For $s \notin \mathbb{N}_0$ the Aronstein–Slobodeckij characterization for norms and seminorms is employed. A second, lower index, for example, $\|\cdot\|_{s;\omega}$ or $|\cdot|_{s;\omega}$, indicates that the norm or seminorm is not considered on $\Omega$ but on $\omega$, which will typically be an element or an edge or a face. We will also work with the Besov spaces $B^s_{2,q}(\Omega)$, which are defined as interpolation spaces using the 'real method' (see Tartar, 2007, Chapter 22 and Triebel, 1995, Section 1.3,4.4 for details): for positive $s \notin \mathbb{N}$ and $q \in [1, \infty]$, we set

$$B^s_{2,q}(\Omega) := (H^{\lfloor s\rfloor}(\Omega), H^{\lceil s\rceil}(\Omega))_{s-\lfloor s\rfloor,q}.$$

To assist readers more familiar with Sobolev spaces, we mention that, for each $\varepsilon > 0$ and noninteger $s$, we have the continuous embedding $H^{s+\varepsilon}(\Omega) \subset B^s_{2,1}(\Omega) \subset H^s(\Omega)$.

### 2.1 *Discretization*

For simplicity of notation, we restrict ourselves to a family of quasi-uniform simplicial meshes $\mathcal{T}_h$ and use standard mixed finite elements. We write $\mathcal{E}_h$ for the set of edges in two dimensions and for the set of faces in three dimensions. We consider uniformly inf–sup-stable pairings $V_h^k \times M_h^k \subset H(\mathrm{div};\Omega) \times L^2(\Omega)$ where $V_h^k$ is either a Raviart–Thomas (RT) or a Brezzi–Douglas–Marini (BDM) finite element space. For details, we refer the reader to Brezzi (1989, 2003), Brezzi & Fortin (1991) and Wriggers & Carstensen (2009) and the references therein and to the original contributions Raviart & Thomas (1977), Nédélec (1980), Arnold & Brezzi (1985) and Brezzi *et al.* (1985, 1986). More precisely, we set

$$\mathrm{RT}_h^k := \{\tau \in H(\mathrm{div};\Omega),\ \tau|_T \in \mathrm{RT}_k(T), T \in \mathcal{T}_h\}, \quad \mathrm{RT}_k(T) := (P_k(T))^d + P_k(T)\mathbf{x},$$

$$\mathrm{BDM}_h^k := \{\tau \in H(\mathrm{div};\Omega),\ \tau|_T \in \mathrm{BDM}_k(T), T \in \mathcal{T}_h\}, \quad \mathrm{BDM}_k(T) := (P_k(T))^d,$$

where $k \in \mathbb{N}_0$ in the case of RT elements and $k \in \mathbb{N}$ for BDM elements. The local spaces on the element $T$ are denoted by $V_k(T)$. We recall that $\mathrm{RT}_h^k \subset \mathrm{BDM}_h^{k+1} \subset \mathrm{RT}_h^{k+1}$, $k \in \mathbb{N}_0$. Quite often, lowest-order finite element spaces are used. The popular choice $\mathrm{RT}_h^0$ has exactly one degree of freedom per edge/face $e \in \mathcal{E}_h$, whereas $\mathrm{BDM}_h^1$ has two/three degrees of freedom per edge/face.

For the approximation in $L^2(\Omega)$, we use piecewise polynomials

$$P_h^k := \{v \in L^2(\Omega),\ v|_T \in P_k(T), T \in \mathcal{T}_h\}.$$

It is well known that the pairings $(V_h^k, M_h^k) := (\mathrm{RT}_h^k, P_h^k)$, $k \in \mathbb{N}_0$ and $(V_h^k, M_h^k) := (\mathrm{BDM}_h^k, P_h^{k-1})$, $k \in \mathbb{N}$ are uniformly inf–sup stable (Brezzi & Fortin, 1991, Section IV.1.2). As can be easily seen, mixed finite elements satisfy the inverse estimate

$$\|\tau_h \mathbf{n}\|_{0;e} \leqslant \frac{C}{\sqrt{h}} \|\tau_h\|_{0;T}, \quad \tau_h \in V_k(T), \quad T \in \mathcal{T}_h, e \in \mathcal{E}_h \text{ with } e \subset \partial T. \tag{2.2}$$

We note that all our constants $0 < c, C < \infty$ are generic constants and do not depend on the mesh size but possibly depend on the order $k$.

Of crucial importance for our analysis will be the so-called Fortin operator $I_h^k$ (see, for example, Brezzi & Fortin, 1991, Section III.3.3), which maps a dense subset of $H(\mathrm{div};\Omega)$ onto $V_h^k$. Analogously to the nodal Lagrange interpolation operator for standard conforming elements, $I_h^k \tau|_T \in V_k(T)$ is uniquely defined by $\tau$ restricted to $T$. Introducing $\Pi_h^k$ as the elementwise defined $L^2$ projection onto $P_h^k$, we have

$$\mathrm{div}\, I_h^k \tau = \Pi_h^k \,\mathrm{div}\, \tau \quad \text{if } V_h^k = \mathrm{RT}_h^k, \tag{2.3a}$$

$$\mathrm{div}\, I_h^k \tau = \Pi_h^{k-1} \,\mathrm{div}\, \tau \quad \text{if } V_h^k = \mathrm{BDM}_h^k. \tag{2.3b}$$

For simplicity of notation, we abbreviate

$$\Pi_h^* := \begin{cases} \Pi_h^k & \text{if } V_h^k = \mathrm{RT}_h^k, \\ \Pi_h^{k-1} & \text{if } V_h^k = \mathrm{BDM}_h^k. \end{cases} \tag{2.4}$$

Hence, $\operatorname{div} \circ I_h^k = \Pi_h^* \circ \operatorname{div}$. Furthermore, $I_h^k$ commutes with $\pi_h^k$, i.e.,

$$(I_h^k \tau)\mathbf{n}|_e = \pi_h^k(\tau \mathbf{n})|_e, \tag{2.5}$$

where $\pi_h^k$ is the $L^2$ projection onto $\prod_{e \in \mathcal{E}_h} P_k(e)$. To each $e \in \mathcal{E}_h$, we associate a unit normal $\mathbf{n}$. If $e \subset \partial\Omega$, then $\mathbf{n}$ is given by the outer unit normal, otherwise the orientation is arbitrary but fixed. Moreover, the Fortin operator has the following local best approximation properties (Brezzi & Fortin, 1991, Proposition 3.6, Section III.3.3):

$$\|\tau - I_h^k \tau\|_{j;T} \leqslant C h^{s+1-j} |\tau|_{s+1;T}, \quad \tau \in (H^{s+1}(T))^d, \;\; 0 \leqslant s \leqslant k, \;\; j \in \{0, 1\}, \tag{2.6a}$$

$$\|(\tau - I_h^k \tau)\mathbf{n}\|_{0;e} \leqslant C h^{s+1} |\tau \mathbf{n}|_{s+1;e}, \quad \tau \in H^{s+1}(e), \;\; -\tfrac{1}{2} \leqslant s \leqslant k. \tag{2.6b}$$

It is obvious that (2.6b) results directly from (2.5). We remark that (2.6) holds for both choices of $V_h^k$, whereas for estimates in the $L^2$ norm of the divergence, we have to consider the two families separately; due to (2.3), we have

$$\|\operatorname{div} \tau - \operatorname{div} I_h^k \tau\|_{0;T} \leqslant C h^{s+1} |\operatorname{div} \tau|_{s+1;T}, \quad \operatorname{div} \tau \in H^{s+1}(T), \; \begin{cases} 0 \leqslant s \leqslant k & \text{if } V_h^k = \mathrm{RT}_h^k, \\ 0 \leqslant s \leqslant k-1 & \text{if } V_h^k = \mathrm{BDM}_h^k. \end{cases}$$

By $(\sigma_h, u_h) \in V_h^k \times M_h^k$, we denote the finite element solution of the mixed formulation, i.e., $(\sigma_h, u_h)$ satisfies (2.1) if the test spaces are restricted to $V_h^k$ and $M_h^k$. Moreover, $(\sigma_h, u_h) \in V_h^k \times M_h^k$ is uniquely characterized by the following Galerkin orthogonalities:

$$a(\sigma - \sigma_h, \tilde{\sigma}_h) + b(\tilde{\sigma}_h, u - u_h) = 0, \quad \tilde{\sigma}_h \in V_h^k, \tag{2.7a}$$

$$b(\sigma - \sigma_h, \tilde{u}_h) = 0, \quad \tilde{u}_h \in M_h^k. \tag{2.7b}$$

With $\Pi_h^*$ defined in (2.4), we remark that (2.7b) implies the relationship

$$\operatorname{div} \sigma_h = \Pi_h^* \operatorname{div} \sigma. \tag{2.8}$$

## 2.2 *Main result*

We start with a result from Li *et al.* (2010, Lemma 2.1) that will be an important ingredient for our recovering an additional factor $\sqrt{h}$ for the $L^2$ error of the flux on $\Gamma$ when compared with standard estimates. This result allows us to control the $L^2$ norm of a function in a small tubular neighbourhood of a $(d-1)$-dimensional manifold. For sufficiently smooth functions, this $L^2$ norm scales with the volume of this tubular neighbourhood.

LEMMA 2.1 Let $\gamma$ be a finite union of $(d-1)$-dimensional manifolds such that $\Omega$ is decomposed into a finite number of Lipschitz domains. For $\alpha_{\max} \geqslant \alpha \geqslant 1$ and $h > 0$, define tubular neighbourhoods of $\gamma$ by

$$S_{\alpha h,\gamma} := \{x \in \Omega \mid \operatorname{dist}(x,\gamma) < \alpha h\}. \tag{2.9}$$

Then there exists a constant $C > 0$, independent of $h$ and $\alpha$ but depending on the Lipschitz character of $\gamma$ and $\alpha_{\max}$, such that

$$\|z\|_{0;S_{\alpha h,\gamma}} \leqslant C\sqrt{h}\|z\|_{B^{1/2}_{2,1}(\Omega)}, \quad z \in B^{1/2}_{2,1}(\Omega), \tag{2.10}$$

$$\|z\|_{(B^{1/2}_{2,1})'} \leqslant C\sqrt{h}\|z\|_0, \quad z \in L^2(\Omega), \quad \operatorname{supp} z \subset S_{\alpha h,\gamma}. \tag{2.11}$$

*Proof.* The estimate (2.10) is taken directly from Li *et al.* (2010, Lemma 2.1). The bound (2.11) is obtained from (2.10) by a simple duality argument, which can be found in Melenk & Wohlmuth (2012, proof of Lemma 5.2). □

Let $\Gamma$ be a finite union of $(d-1)$-dimensional manifolds such that $\Omega$ is decomposed into finitely many Lipschitz domains by $\Gamma$. We stress that, while $\Omega$ is assumed to be convex, the subdomains need not be convex. We assume, furthermore, that the mesh $\mathcal{T}_h$ resolves $\Gamma$. Hence, $\Gamma$ can be written as the union of $\mathcal{O}(h^{1-d})$ edges/faces in $\mathcal{E}_h$, i.e., $\bar{\Gamma} := \bigcup_{e \in \mathcal{E}_\Gamma \subset \mathcal{E}_h} \bar{e}$. Using the definition (2.9), we set

$$S_{\alpha h} := S_{\alpha h,\Gamma}, \quad 1 \leqslant \alpha \leqslant \alpha_{\max} < \infty. \tag{2.12}$$

LEMMA 2.2 Let $(\sigma, u) \in H(\operatorname{div};\Omega) \times L^2(\Omega)$ be the solution of (2.1) and let $(\sigma_h, u_h) \in V^k_h \times M^k_h$ be its finite element approximation determined by (2.7). If $\sigma \in (B^{k+3/2}_{2,1}(\Omega))^d$, then the $L^2$-norm error of the flux on the interface $\Gamma$ can be bounded by

$$\|(\sigma - \sigma_h)\mathbf{n}\|_{0;\Gamma} \leqslant C\left(h^{k+1}\|\sigma\|_{B^{k+3/2}_{2,1}} + \frac{1}{\sqrt{h}}\|\sigma - \sigma_h\|_{0;S_h}\right),$$

where $\|\cdot\|_{B^{k+3/2}_{2,1}}$ stands for the Besov space $(B^{k+3/2}_{2,1}(\Omega))^d$ norm.

*Proof.* Starting with the triangle inequality and using (2.2) and (2.5), we obtain the upper bound

$$\|(\sigma - \sigma_h)\mathbf{n}\|_{0;\Gamma} \leqslant C\left(\|\sigma\mathbf{n} - \pi^k_h(\sigma\mathbf{n})\|_{0;\Gamma} + \frac{1}{\sqrt{h}}\|\sigma - I^k_h\sigma\|_{0;S_h} + \frac{1}{\sqrt{h}}\|\sigma - \sigma_h\|_{0;S_h}\right).$$

The first two terms on the right-hand side yield, due to the best approximation property of $\pi^k_h$ and the local character of $I^k_h$, order $h^{k+1}$ estimates, provided that the solution is sufficiently smooth. More precisely, for the second term, we can apply (2.6a) in combination with (2.10).

The first term can be bounded using (2.6b) with $s = k$ and the fact that the trace operator is a bounded linear operator $B^{k+3/2}_{2,1}(\Omega) \to H^{k+1}(\Gamma)$ (Triebel, 1995, Theorem 2.9.3). Here, the norm of $H^{k+1}(\Gamma)$ is understood to be taken edge/facewise. □

Lemma 2.2 shows that there is some hope to recover an extra factor of $\sqrt{h}$ in the *a priori* estimates for the normal flux at the interface. We point out that this factor can be trivially found if the regularity permits optimal-order $L^\infty$ estimates. We refer to Wang (1989) for $L^\infty$ estimates for mixed finite elements and note that these estimates require rather strong regularity assumptions. In the following theorem, which is the principal result of the paper, this assumption is considerably relaxed.

THEOREM 2.3 Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$ be a convex polygon/polyhedron. Fix an interface $\Gamma$ satisfying the above assumptions. Let $(\sigma, u) \in H(\mathrm{div};\Omega) \times L^2(\Omega)$ be the solution of the model problem (2.1) and let $(\sigma_h, u_h) \in V_h^k \times M_h^k$ be its finite element approximation, which satisfies (2.7). If $\sigma \in (B_{2,1}^{k+3/2}(\Omega))^d$, then the $L^2$-norm error of the flux on the interface $\Gamma$ can be bounded by

$$\|(\sigma - \sigma_h)\mathbf{n}\|_{0;\Gamma} \leqslant Ch^{k+1}\|\sigma\|_{B_{2,1}^{k+3/2}} \begin{cases} |\log h| & \text{for the lowest-order case,} \\ 1 & \text{else.} \end{cases} \tag{2.13}$$

Here, the lowest-order case is $k=0$ for RT elements and $k=1$ for BDM elements.

## 3. Dual problems and their regularity

The analysis of the $L^2$ error on the strip $S_h$, defined by (2.12), is based on a dual problem and closely related to the Aubin–Nitsche trick. However, we have to use suitable anisotropic norms and study the dual problem with right-hand sides supported by $S_h$.

### 3.1 *Dual problem formulation*

We denote by $(\lambda, w) \in H(\mathrm{div};\Omega) \times L^2(\Omega)$ the solution of the dual problem

$$a(\lambda, \tilde{\sigma}) + b(\tilde{\sigma}, w) = (\chi(\sigma - \sigma_h), \tilde{\sigma})_0, \quad \tilde{\sigma} \in H(\mathrm{div};\Omega), \tag{3.1a}$$

$$b(\lambda, \tilde{w}) = 0, \quad \tilde{w} \in L^2(\Omega), \tag{3.1b}$$

where $0 \leqslant \chi \leqslant 1$ is a smooth cut-off function that is equal to 1 in $S_h$ and vanishes on $\Omega \setminus S_{\kappa h}$ with $\kappa$ sufficiently large but independent of the mesh size. We will also assume

$$\|\nabla\chi/\sqrt{\chi}\|_{L^\infty} \leqslant Ch^{-1}. \tag{3.2}$$

The mixed finite element approximation to (3.1) is denoted by $(\lambda_h, w_h) \in V_h^k \times M_h^k$ and satisfies the Galerkin orthogonalities

$$a(\lambda - \lambda_h, \tilde{\sigma}_h) + b(\tilde{\sigma}_h, w - w_h) = 0, \quad \tilde{\sigma}_h \in V_h^k, \tag{3.3a}$$

$$b(\lambda - \lambda_h, \tilde{u}_h) = 0, \quad \tilde{u}_h \in M_h^k. \tag{3.3b}$$

It is well known that a higher-order *a priori* estimate can be obtained for the pressure; namely, using the convexity of $\Omega$, one can show (see, for example, Brezzi & Fortin, 1991, outset of Section V.3)

$$\|w_h - \Pi_h^* w\|_0 \leqslant Ch\|\lambda - \lambda_h\|_0. \tag{3.4}$$

For further developments, it will be useful to note that for sufficiently regular $w$, we have

$$-\Delta w = \mathrm{div}(\chi(\sigma - \sigma_h)) \quad \text{in } \Omega, \quad w = 0 \quad \text{on } \partial\Omega \tag{3.5}$$

and correspondingly $\lambda = \chi(\sigma - \sigma_h) + \nabla w$.

### 3.2 *Regularity*

Our *a priori* analysis is based on regularity results for the solution $(\lambda, w)$ of (3.1). Let us study this problem in more generality by considering, for $g \in (L^2(\Omega))^d$, the problem of finding $(\lambda_g, w_g) \in H(\mathrm{div};\Omega) \times L^2(\Omega)$ such that

$$a(\lambda_g, \tilde{\sigma}) + b(\tilde{\sigma}, w_g) = (g, \tilde{\sigma})_0, \quad \tilde{\sigma} \in H(\mathrm{div};\Omega), \tag{3.6a}$$

$$b(\lambda_g, \tilde{w}) = 0, \qquad \tilde{w} \in L^2(\Omega). \tag{3.6b}$$

Let us denote by $T^M = (T^M_\lambda, T^M_w)$ the solution operator $g \mapsto (\lambda_g, w_g)$ for (3.6), i.e., $\lambda_g = T^M_\lambda g$ and $w_g = T^M_w g$. Then, the following two technical lemmas give us suitable regularity and stability results for the $w$ component $T^M_w g$.

LEMMA 3.1 Let $\Omega$ be a bounded Lipschitz domain. Then $T^M_w$ is a bounded linear operator with the following mapping properties:

(i) $T^M_w : (H(\mathrm{div};\Omega))' \to L^2(\Omega)$;

(ii) $T^M_w : (L^2(\Omega))^d \to H^1_0(\Omega)$;

(iii) if $\Omega$ is convex, then $T^M_w : H(\mathrm{div};\Omega) \to H^2(\Omega) \cap H^1_0(\Omega)$;

(iv) if $\Omega$ is convex, then $T^M_w : ((B^{1/2}_{2,1}(\Omega))^d)' \to B^{1/2}_{2,\infty}(\Omega)$.

*Proof.* Statement (i) follows from the well-posedness of the saddle point problem (3.6). To see (ii), let $\hat{w}_g \in H^1_0(\Omega)$ satisfy

$$(\nabla \hat{w}_g, \nabla \varphi)_0 = -(g, \nabla \varphi)_0, \quad \varphi \in H^1_0(\Omega) \tag{3.7}$$

and set $\hat{\lambda}_g := g + \nabla \hat{w}_g$. Then, we find $\mathrm{div}\,\hat{\lambda}_g = 0$ and thus $(\hat{\lambda}_g, \hat{w}_g) \in H(\mathrm{div};\Omega) \times L^2(\Omega)$. Moreover, $(\hat{\lambda}_g, \hat{w}_g)$ satisfies (3.6a) and (3.6b). Since the solution of (3.6) is unique, we conclude $w_g = \hat{w}_g$; thus (ii) is valid.

For $g \in H(\mathrm{div};\Omega)$, an integration by parts shows that $w_g$ not only solves (3.7) but also

$$(\nabla w_g, \nabla \varphi)_0 = (\mathrm{div}\, g, \varphi)_0, \quad \varphi \in H^1_0(\Omega).$$

The standard shift theorem for convex domains then gives $w_g \in H^2(\Omega)$, and thus (iii) holds.

Finally, we show (iv). The proof exploits an equivalence of the weak and the very weak formulation of Poisson problems in convex domains. We consider the variational problem: find $y \in L^2(\Omega)$ such that

$$B(y, \varphi) := (y, \Delta\varphi)_0 = \langle g, \nabla\varphi \rangle_{((H^1(\Omega))^d)' \times (H^1(\Omega))^d}, \quad \varphi \in H^2(\Omega) \cap H^1_0(\Omega), \tag{3.8}$$

where $\langle \cdot, \rangle_{(((H^1(\Omega)))^d)' \times (H^1(\Omega))^d}$ stands for the duality pairing between $((H^1(\Omega))^d)'$ and $(H^1(\Omega)^d)$ (with pivot space $(L^2(\Omega))^d$). By the convexity of $\Omega$, the bilinear form $B$ satisfies an inf–sup condition, and thus the solution operator $T^D_{vw} : (H^1(\Omega)^d)' \to L^2(\Omega)$ given by $g \mapsto y$ is bounded and linear. Selecting $\varphi \in H^2(\Omega) \cap H^1_0(\Omega)$ in (3.7) and integrating by parts shows that $w_g$ also solves (3.8). By uniqueness, we thus get that the solution $y = T^D_{vw} g = w_g \in H^1_0(\Omega)$ if $g \in (L^2(\Omega))^d$. Having that $T^D_{vw} : (L^2(\Omega))^d \to$

$H_0^1(\Omega) \subset H^1(\Omega)$ is bounded and linear, we can apply an interpolation argument to find that

$$T_{vw}^D : ((H^1(\Omega))^d)', ((L^2(\Omega))^d)_{1/2,\infty} \to (L^2(\Omega), H^1(\Omega))_{1/2,\infty} = B^{1/2}_{2,\infty}(\Omega)$$

is bounded. Finally, we recall that $T_{vw}^D = T_w^M$ on $(L^2(\Omega))^d$ and note that (see, for example (Triebel, 1995, Theorem 1.11.2), or (Tartar, 2007, Lemma 41.3)), we have

$$((H^1(\Omega))^d)', ((L^2(\Omega))^d)_{1/2,\infty} = ((H^1(\Omega))^d, (L^2(\Omega))^d)_{1/2,1})' = (B^{1/2}_{2,1}(\Omega)^d)'. \qquad \square$$

REMARK 3.2 The assumption of the convexity of $\Omega$ in Lemma 3.1(iv) can be weakened: it suffices that $\Omega$ admits a shift theorem by more than $1/2$; see Lemma A3, for details.

Next, we provide stability results in weighted Sobolev norms. As weight, we introduce the regularized distance $\delta$ from $\Gamma$, namely

$$\delta(x) := h + \mathrm{dist}(x, \Gamma). \tag{3.9}$$

LEMMA 3.3 Let the bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$ be a polygon ($d = 2$) or a polyhedron ($d = 3$). Fix $c_0 \geqslant 1$ and $\beta \in \mathbb{R}$. Then there exists $c_2 > c_1 > c_0$ independent of $h$ such that the following is true: if $y \in B^{1/2}_{2,\infty}(\Omega)$ solves $-\Delta y = 0$ in $\Omega \setminus S_{c_0 h}$, then

$$\|\sqrt{\delta}\nabla y\|_{0;\Omega\setminus S_{c_2 h}} \leqslant C\sqrt{|\log h|}\|y\|_{B^{1/2}_{2,\infty}(\Omega)}, \tag{3.10a}$$

$$\|\sqrt{\delta}\nabla^2 y\|_{0;\Omega\setminus S_{c_2 h}} \leqslant C\sqrt{|\log h|}\|y\|_{B^{3/2}_{2,\infty}(\Omega)} \quad \text{if } y \in B^{3/2}_{2,\infty}(\Omega), \tag{3.10b}$$

$$\|\delta^\beta \nabla^3 y\|_{0;\Omega\setminus S_{c_2 h}} \leqslant C\|\delta^{\beta-1}\nabla^2 y\|_{0;\Omega\setminus S_{c_1 h}}. \tag{3.10c}$$

*Proof.* The upper bound (3.10b) is given by Melenk & Wohlmuth (2012, proof of Lemma 5.4, Equation (5.4)), and the proof of (3.10a) follows by the same type of arguments. The estimate (3.10c) expresses interior regularity for harmonic functions and is also taken from Melenk & Wohlmuth (2012, Lemma 5.4). $\square$

We are now in a position to apply Lemmas 3.1 and 3.3 to the dual problem (3.1), i.e., we set $g = \chi(\sigma - \sigma_h)$ in (3.6).

LEMMA 3.4 Let $\Omega$ be convex and $(\lambda, w) \in H(\mathrm{div};\Omega) \times L^2(\Omega)$ be the solution of (3.1). Then $w \in H^2(\Omega) \cap H_0^1(\Omega)$, and it satisfies

$$\|\sqrt{\delta}\nabla w\|_0 \leqslant C\sqrt{h}\sqrt{|\log h|}\|\chi(\sigma - \sigma_h)\|_0, \tag{3.11a}$$

$$\|\sqrt{\delta}\nabla^2 w\|_0 \leqslant C\sqrt{h}\sqrt{|\log h|}\|\,\mathrm{div}(\chi(\sigma - \sigma_h))\|_0. \tag{3.11b}$$

*Proof.* We note that the support properties of $\chi$ imply that $w$ is harmonic in $\Omega \setminus S_{\kappa h}$, and thus we are in the setting of Lemma 3.3 with $c_0$ equal to $\kappa$, where the constant $\kappa$ is the constant appearing in the definition of $\chi$. We start with the bound (3.11a) and decompose the domain $\Omega$ into $\Omega \setminus S_{c_2 h}$ and $S_{c_2 h}$.

Then Lemma 3.1(iv) in combination with (3.10a) and (2.11) yields

$$\|\sqrt{\delta}\nabla w\|_{0;\Omega\setminus S_{c_2h}} \leqslant C\sqrt{|\log h|}\|\chi(\sigma-\sigma_h)\|_{((B^{1/2}_{2,1}(\Omega))^d)'} \leqslant C\sqrt{|\log h|}\sqrt{h}\|\chi(\sigma-\sigma_h)\|_0.$$

Recalling that $\delta$ is bounded by $Ch$ on $S_{c_2h}$, we have the trivial bound $\|\sqrt{\delta}\nabla w\|_{0;S_{c_2h}} \leqslant C\sqrt{h}\|\nabla w\|_{0;S_{c_2h}}$. Then Lemma 3.1(ii) implies $\|\nabla w\|_0 \leqslant C\|\chi(\sigma-\sigma_h)\|_0$ and thus $\|\sqrt{\delta}\nabla w\|_{0;S_{c_2h}} \leqslant C\sqrt{h}\|\chi(\sigma-\sigma_h)\|_0$.

To prove (3.11b), we proceed as in the proof of (3.11a). Starting with the $L^2$ norm on $S_{c_2h}$, we get, in view of Lemma 3.1(iii),

$$\|\sqrt{\delta}\nabla^2 w\|_{0;S_{c_2h}} \leqslant C\sqrt{h}\|\nabla^2 w\|_0 \leqslant C\sqrt{h}\|\operatorname{div}(\chi(\sigma-\sigma_h))\|_0.$$

We note that Melenk & Wohlmuth (2012, Lemma 5.2) states

$$\|w\|_{B^{3/2}_{2,\infty}(\Omega)} \leqslant C\|\operatorname{div}(\chi(\sigma-\sigma_h))\|_{(B^{1/2}_{2,1}(\Omega))'}. \tag{3.12}$$

To bound the weighted norm on $\Omega\setminus S_{c_2h}$, we use (3.10b), (3.12) and (2.11) to obtain

$$\|\sqrt{\delta}\nabla^2 w\|_{0;\Omega\setminus S_{c_2h}} \leqslant C\sqrt{|\log h|}\|w\|_{B^{3/2}_{2,\infty}(\Omega)} \leqslant C\sqrt{|\log h|}\sqrt{h}\|\operatorname{div}(\chi(\sigma-\sigma_h))\|_0. \qquad \square$$

## 4. Approximation in anisotropic norms

In this section, we introduce anisotropic norms and reconsider the approximation properties of the Fortin operator $I_h^k$ and the $L^2$ projection $\Pi_h^*$ with respect to these norms.

### 4.1 *Anisotropic norms*

The definition of our anisotropic norms is based on the idea of viewing a $d$-dimensional domain as the product of a one-dimensional interval and a $(d-1)$-dimensional manifold. The anisotropic norm arises from treating these two directions differently. As will become clear below, this point of view is closely related to certain weighted Sobolev spaces with weight given by the distance from $\Gamma$.

For $\tau \geqslant 0$, let us introduce the $(d-1)$-dimensional manifold $\gamma_\tau$ by

$$\gamma_\tau := \{x\in\Omega,\ \operatorname{dist}(x,\Gamma)=\tau\}, \qquad \tau\geqslant 0.$$

We set $D := \operatorname{diam}\Omega$ and note that, for $\tau \geqslant D$, we have $\gamma_\tau = \emptyset$. We place ourselves in the setting of the Fubini–Tonelli formula for integration over $\Omega$ and assume the existence of a measure $\mathrm{d}\mu^\tau$ such that

$$\int_\Omega w\,\mathrm{d}x = \int_{\tau=0}^{D}\int_{\gamma_\tau} w\,\mathrm{d}\mu^\tau\,\mathrm{d}\tau.$$

We point out that this is done for simplicity of exposition—in the general case, we can use a localization technique and fitted coordinate systems as in Melenk & Wohlmuth (2012). A second simplifying assumption is that the Lipschitz character of $\gamma_\tau$ may be assumed to be independent of $\tau\in[0,D]$.

Next, we introduce our anisotropic norms,

$$\|v\|_{L(p;2)} := \|v\|_{L^p((0,D);L^2(\gamma_\tau))} := \left(\int_{\tau=0}^{D}\left(\int_{\gamma_\tau} v^2\,\mathrm{d}\mu^\tau\right)^{p/2}\mathrm{d}\tau\right)^{1/p}, \quad 1 \leqslant p < \infty, \tag{4.1a}$$

$$\|v\|_{L(\infty;2)} := \|v\|_{L^\infty((0,D);L^2(\gamma_\tau))} := \sup_{\tau\in(0,D)}\left(\int_{\gamma_\tau} v^2\,\mathrm{d}\mu^\tau\right)^{1/2}, \tag{4.1b}$$

and observe that, for $p = 2$, we recover by Fubini–Tonelli the standard $L^2(\Omega)$ norm. As a consequence of the one-dimensional Hölder inequality, we find

$$\left|\int_\Omega v\tilde{v}\,\mathrm{d}x\right| \leqslant \|v\|_{L(p;2)}\|\tilde{v}\|_{L(q;2)}, \quad \frac{1}{p}+\frac{1}{q}=1. \tag{4.2}$$

### 4.2 *Approximation in anisotropic norms*

In this section, we reconsider the Fortin operator and its approximation properties with respect to our newly defined anisotropic norms. The definition (4.1b) of the $L(\infty;2)$ norm shows that we have to consider the $L^2(\gamma_\tau)$ norm in more detail. As a preliminary step, we introduce the set

$$\mathcal{T}_\tau := \{T \in \mathcal{T}_h\colon \gamma_\tau \cap \bar{T} \neq \emptyset\} \tag{4.3}$$

and observe that $\bigcup_{T\in\mathcal{T}_\tau}\bar{T} \subset S_{h,\gamma_\tau}$.

Lemma 4.1 For $\sigma \in (B^{k+3/2}_{2,1}(\Omega))^d$, we have

$$\|\sigma - I_h^k\sigma\|_{L(\infty;2)} \leqslant Ch^{k+1}\|\sigma\|_{B^{k+3/2}_{2,1}}, \tag{4.4a}$$

$$\|\operatorname{div}\sigma - \Pi_h^*\operatorname{div}\sigma\|_{L(\infty;2)} \leqslant Ch^k\|\operatorname{div}\sigma\|_{B^{k+1/2}_{2,1}}. \tag{4.4b}$$

*Proof.* Recalling (4.1b), we see that we have to bound the $L^2(\gamma_\tau)$ norm. A scaling argument allows us to bound the $L^2(\gamma_\tau\cap\bar{T})$ norm in terms of a combination of weighted $L^2(T)$ and $H^1(T)$ norms. Owing to (2.6a), we then obtain

$$\begin{aligned}\|\sigma - I_h^k\sigma\|^2_{0;\gamma_\tau} &= \sum_{T\in\mathcal{T}_\tau}\|\sigma - I_h^k\sigma\|^2_{0;\gamma_\tau\cap\bar{T}} \leqslant C\sum_{T\in\mathcal{T}_\tau}\left(\frac{1}{h}\|\sigma - I_h^k\sigma\|^2_{0;T} + h\|\nabla(\sigma - I_h^k\sigma)\|^2_{0;T}\right)\\ &\leqslant C\sum_{T\in\mathcal{T}_\tau}h^{2k+1}|\sigma|^2_{k+1;T} \leqslant Ch^{2k+1}|\sigma|^2_{k+1;S_{h,\gamma_\tau}}.\end{aligned}$$

The definition (4.3) guarantees that an additional factor of $h$ can be recovered using Lemma 2.1; that is, $|\sigma|^2_{k+1;S_{h,\gamma_\tau}} \leqslant Ch\|\sigma\|^2_{B^{k+3/2}_{2,1}}$ and thus (4.1b) yields (4.4a).

Now, we focus on (4.4b) and proceed as before:

$$
\begin{aligned}
\| \operatorname{div} \sigma - \Pi_h^* \operatorname{div} \sigma \|_{0;\gamma_\tau}^2 &= \sum_{T \in \mathcal{T}_\tau} \| \operatorname{div} \sigma - \Pi_h^* \operatorname{div} \sigma \|_{0;\gamma_\tau \cap \bar{T}}^2 \\
&\leqslant C \sum_{T \in \mathcal{T}_\tau} \left( \frac{1}{h} \| \operatorname{div} \sigma - \Pi_h^* \operatorname{div} \sigma \|_{0;T}^2 + h \| \nabla (\operatorname{div} \sigma - \Pi_h^* \operatorname{div} \sigma) \|_{0;T}^2 \right) \\
&\leqslant C \sum_{T \in \mathcal{T}_\tau} h^{2k-1} | \operatorname{div} \sigma |_{k;T}^2 \leqslant C h^{2k-1} | \operatorname{div} \sigma |_{k;S_{h,\gamma_\tau}}^2 \leqslant C h^{2k} \| \operatorname{div} \sigma \|_{B_{2,1}^{k+1/2}}^2 . \qquad \square
\end{aligned}
$$

REMARK 4.2 We note that Lemma 4.1 is not sharp in the case where $V_h^k = \mathrm{RT}_h^k$. Then $k$ on the right-hand side of (4.4b) can be replaced by $k+1$, provided that the solution is regular enough. However, this sharper result does not significantly improve the global estimate for the normal flux on the interface and is thus not stated.

## 5. Proof of the main result, Theorem 2.3

In this section, we provide the proof of Theorem 2.3. To start with, in Section 5.1 we consider a local $L^2$ estimate for the error $\sigma - \sigma_h$. In Section 5.3, we focus on *a priori* bounds for the error in the flux of the dual problem. Finally in Section 5.4, the main result (2.13) is established.

### 5.1 *Local $L^2$ estimates*

The Aubin–Nitsche trick in combination with the Hölder-type inequality (4.2) allows us to bound $\| \sqrt{\chi} (\sigma - \sigma_h) \|_0$.

LEMMA 5.1 Let $(\sigma, u) \in H(\operatorname{div}; \Omega) \times L^2(\Omega)$ be the solution of (2.1) and $(\sigma_h, u_h) \in V_h^k \times M_h^k$ be its finite element approximation. Let $(\lambda, w) \in H(\operatorname{div}; \Omega) \times L^2(\Omega)$ be the solution of the dual problem (3.1) and $(\lambda_h, w_h) \in V_h^k \times M_h^k$ be its finite element approximation. Then, for $\sigma \in B_{2,1}^{k+3/2}(\Omega)$, we have

$$
\| \sqrt{\chi} (\sigma - \sigma_h) \|_0^2 \leqslant C h^k ( h \| \lambda - \lambda_h \|_{L(1;2)} + \| w - \Pi_h^* w \|_{L(1;2)} ) \| \sigma \|_{B_{2,1}^{k+3/2}} . \tag{5.1}
$$

*Proof.* A crucial observation for the proof is that $\operatorname{div} \lambda_h = 0$. This follows from the fact that $\operatorname{div} V_h^k = M_h^k$ and (3.1b), (3.3b). Moreover, we recall that, by (2.1b), we have $b(\sigma - \sigma_h, v_h) = 0$ for all $v_h \in M_h^k$. The symmetry of the bilinear form $a(\cdot,\cdot)$ yields, in combination with (2.1a), that $a(\tau_h, \sigma - \sigma_h) = 0$ for all $\tau_h \in V_h^k$ with $\operatorname{div} \tau_h = 0$. Using the definition of the dual solution and exploiting the Galerkin orthogonality (2.7a), we find

$$
\begin{aligned}
\| \sqrt{\chi} (\sigma - \sigma_h) \|_0^2 &= (\chi (\sigma - \sigma_h), \sigma - \sigma_h)_0 = a(\lambda, \sigma - \sigma_h) + b(\sigma - \sigma_h, w) \\
&= a(\lambda - \lambda_h, \sigma - \sigma_h) + b(\sigma - \sigma_h, w - w_h).
\end{aligned}
$$

Now, using the Galerkin orthogonality (3.3a), we can replace the finite element solution $\sigma_h$ by the Fortin interpolation of $\sigma$ and $w_h$ by the $L^2$ projection of $w$:

$$
\begin{aligned}
\| \sqrt{\chi} (\sigma - \sigma_h) \|_0^2 &= a(\lambda - \lambda_h, \sigma - I_h^k \sigma) + b(\sigma - I_h^k \sigma, w - w_h) \\
&= a(\lambda - \lambda_h, \sigma - I_h^k \sigma) + b(\sigma - I_h^k \sigma, w - \Pi_h^* w).
\end{aligned}
$$

The Hölder-type inequality (4.2) for our anisotropic norms yields

$$\|\sqrt{\chi}(\sigma - \sigma_h)\|_0^2 \leqslant \|\lambda - \lambda_h\|_{L(1;2)}\|\sigma - I_h^k\sigma\|_{L(\infty;2)} \\ + \|w - \Pi_h^* w\|_{L(1;2)}\|\operatorname{div}\sigma - \Pi_h^*(\operatorname{div}\sigma)\|_{L(\infty;2)}.$$

To obtain a bound for $\|\sqrt{\chi}(\sigma - \sigma_h)\|_0^2$ in terms of the mesh size, we have to control the terms on the right-hand side. The two terms associated with the solution $\sigma$ and the Fortin interpolant $I_h^k\sigma$ are covered by Lemma 4.1. For $\sigma \in B_{2,1}^{k+3/2}(\Omega)$, we have $\|\operatorname{div}\sigma\|_{B_{2,1}^{k+1/2}} \leqslant \|\sigma\|_{B_{2,1}^{k+3/2}}$ and thus (4.4a) in combination with (4.4b) gives (5.1). □

### 5.2 *Estimating* $\|w - \Pi_h^* w\|_{L(1;2)}$

Lemma 5.1 shows that we need to estimate $\|\lambda - \lambda_h\|_{L(1;2)}$ and $\|w - \Pi_h^* w\|_{L(1;2)}$. The case of $\|w - \Pi_h^* w\|_{L(1;2)}$ is significantly simpler and covered in the present subsection. Several ingredients of its proof will reappear in the treatment of $\|\lambda - \lambda_h\|_{L(1;2)}$.

First, we relate our anisotropic $L(1;2)$ norm, defined by (4.1a), to a weighted $L^2$ norm: it can easily be shown by decomposing the interval $(0, D)$ into the two subintervals $(0, \alpha h)$ and $(\alpha h, D)$ that

$$\begin{aligned} \|v\|_{L(1;2)} &= \int_{\tau=0}^{\alpha h}\left(\int_{\gamma_\tau} v^2\,\mathrm{d}\mu^\tau\right)^{1/2}\mathrm{d}\tau + \int_{\tau=\alpha h}^{D}\left(\int_{\gamma_\tau} v^2\,\mathrm{d}\mu^\tau\right)^{1/2}\mathrm{d}\tau \\ &\leqslant \sqrt{\alpha h}\|v\|_{0;S_{\alpha h}} + \left(\int_{\tau=\alpha h}^{D}\tau^{-1}\,\mathrm{d}\tau\int_{\tau=\alpha h}^{D}\int_{\gamma_\tau}\tau v^2\,\mathrm{d}x\,\mathrm{d}\tau\right)^{1/2} \\ &\leqslant \sqrt{\alpha h}\|v\|_{0;S_{\alpha h}} + C|\log\alpha h|^{1/2}\|\sqrt{\delta}v\|_{0;\Omega\setminus S_{\alpha h}}. \end{aligned} \tag{5.2}$$

Similarly, for arbitrary $\varepsilon > 0$, one can show that

$$\|v\|_{L(1;2)} \leqslant \sqrt{\alpha h}\|v\|_{0;S_{\alpha h}} + C_\varepsilon h^{-\varepsilon}\|\delta^{1/2+\varepsilon}v\|_{0;\Omega\setminus S_{\alpha h}}. \tag{5.3}$$

We will use the bound (5.2) to handle the lowest-order case and (5.3) for higher-order elements.

With these observations in hand, we formulate the following result.

LEMMA 5.2 (i) For the lowest-order cases (i.e., $k = 0$ for RT elements and $k = 1$ for BDM elements) there holds

$$\|w - \Pi_h^* w\|_{L(1;2)} \leqslant Ch^{3/2}|\log h|\|\chi(\sigma - \sigma_h)\|_0. \tag{5.4}$$

(ii) For higher-order elements, there holds

$$\|w - \Pi_h^* w\|_{L(1;2)} \leqslant Ch^{3/2}\|\chi(\sigma - \sigma_h)\|_0.$$

*Proof.* The case of lowest-order elements is proved by combining (5.2) with (3.11a) to obtain

$$\|w - \Pi_h^* w\|_{L(1;2)}^2 \leqslant Ch\|\nabla w\|_{0;S_{2h}}^2 + Ch^2|\log h|\|\sqrt{\delta}\nabla w\|_0^2 \leqslant Ch^3|\log h|^2\|\chi(\sigma - \sigma)\|_0^2.$$

The case of higher-order elements relies on (5.3). Let $c \geqslant 1$ be such that $w$ is harmonic on $\Omega \setminus S_{ch}$. For $\varepsilon > 0$ sufficiently small but fixed, we get from (5.3), in view of the fact that the image space of $\Pi_h^*$

contains (discontinuous) piecewise linear functions,

$$\begin{aligned}\|w - \Pi_h^* w\|_{L(1;2)}^2 &\leqslant Ch\|w - \Pi_h^* w\|_{0;S_{2ch}}^2 + Ch^{-2\varepsilon}\|\delta^{1/2+\varepsilon}(w - \Pi_h^* w)\|_{0;\Omega\setminus S_{2ch}}^2 \\ &\leqslant Ch\|w - \Pi_h^* w\|_{0;S_{2ch}}^2 + Ch^{-2\varepsilon}h^4\|\delta^{1/2+\varepsilon}\nabla^2 w\|_{0;\Omega\setminus S_{(2c-1)h}}^2 \\ &\leqslant Ch\|w - \Pi_h^* w\|_{0;S_{2ch}}^2 + Ch^3\|\nabla w\|_0^2 \leqslant Ch^3\|\nabla w\|_0^2.\end{aligned}$$

In the last step, we have exploited that $w$ is harmonic on $\Omega \setminus S_{ch}$ and thus, in view of (3.10c), we obtain $\|\delta^{1/2+\varepsilon}\nabla^2 w\|_{0;\Omega\setminus S_{(2c-1)h}} \leqslant C\|\delta^{-1/2+\varepsilon}\nabla w\|_0$. Then, the stability assertion of Lemma 3.1 (ii) allows us to conclude the argument. □

### 5.3 *A priori bounds on the error in the dual flux*

The estimate for $\sigma - \sigma_h$ in Lemma 5.1 involves anisotropic norms of the FEM error $\lambda - \lambda_h$ and the approximation error $w - \Pi_h^* w$ for the solution $(\lambda, w)$ of the dual problem (3.1). In this subsection, we focus on the error in the flux variable $\lambda$ and use the regularity assertions for $w$ given in Lemma 3.4.

We start with some preliminary technical results which play an important role in the bound for the flux error. As is standard for localized estimates in finite element approximation, we have to use a 'superapproximation' property formulated in Lemma 5.3. We refer to, for example, Wahlbin (1995, Chapter 2.3) for its use in a Poisson-type problem and to Gastaldi & Nochetto (1989) for its application in mixed FEMs.

LEMMA 5.3 ('Superapproximation') Fix $T \in \mathcal{T}_h$. Let $z \in W^{1;\infty}(T)$. Then there exists a constant $C > 0$ depending only on the shape regularity of $T$ and $k$ such that

$$\|z\tau_h - I_h^k(z\tau_h)\|_{0;T} \leqslant Ch\|\nabla z\|_{W^{0;\infty}(T)}\|\tau_h\|_{0;T}, \quad \tau_h \in V_k(T), \tag{5.5a}$$

$$\|I_h^k(z\tau_h)\|_{0;T} \leqslant C(\|z\|_{W^{0;\infty}(T)} + h\|\nabla z\|_{W^{0;\infty}(T)})\|\tau_h\|_{0;T}, \quad \tau_h \in V_k(T). \tag{5.5b}$$

*Proof.* We start with the stability bound (5.5b). Since the Fortin operator is not $H(\mathrm{div}; T)$ stable, we use the triangle inequality, the approximation property (2.6a) and an inverse estimate for polynomials to find

$$\begin{aligned}\|I_h^k(z\tau_h)\|_{0;T} &\leqslant \|z\tau_h - I_h^k(z\tau_h)\|_{0;T} + \|z\tau_h\|_{0;T} \leqslant C(h|z\tau_h|_{1;T} + \|z\tau_h\|_{0;T}) \\ &\leqslant C(h\|\nabla z\|_{W^{0;\infty}(T)}\|\tau_h\|_{0;T} + \|z\|_{W^{0;\infty}(T)}\|\tau_h\|_{0;T}).\end{aligned}$$

Equation (5.5a) can easily be shown using (5.5b). Recalling that $\Pi_h^0 z|_T \tau_h \in V_k(T)$ and that the definition of $I_h^k(z\tau_h)$ only involves values of $z\tau_h$ restricted to $T$, we obtain

$$\|z\tau_h - I_h^k(z\tau_h)\|_{0;T} = \|(z - \Pi_h^0 z)\tau_h - I_h^k((z - \Pi_h^0 z)\tau_h)\|_{0;T} \leqslant Ch\|\nabla z\|_{W^{0;\infty}(T)}\|\tau_h\|_{0;T},$$

where we observed $\|z - \Pi_h^0 z\|_{W^{0;\infty}(T)} \leqslant Ch\|\nabla z\|_{W^{0;\infty}(T)}$ and applied (5.5a) with $z - \Pi_h^0 z$ taking the role of $z$ there. □

The proof of the following Lemma 5.4 requires the introduction of some notation. Recall the definition of $\delta$ in (3.9). For $x \in \Omega$ we select two balls $\tilde{B}_{\delta;x}^i$, $i \in \{1, 2\}$, centred at $x$ with radii $\kappa_i\delta(x)$, where $0 < \kappa_1 < \kappa_2$ are suitably chosen independent of the mesh size so that certain covering arguments can be carried out below. We set $B_{\delta;x}^i := \tilde{B}_{\delta;x}^i \cap \Omega$, select $\chi_x \in W^{1;\infty}(\mathbb{R}^d)$, with supp $\chi_x \subset \tilde{B}_{\delta;x}^2$ and $\chi_x \equiv 1$

on $B^1_{\delta;x}$, and require $\|\nabla\chi_x\|_{W^{0;\infty}} \leqslant C\delta(x)^{-1}$. Then we obtain, with the aid of the local superapproximation property (5.5a) and the bound on the gradient of $\chi_x$, the estimate

$$\|\chi_x\tau_h - I_h^k(\chi_x\tau_h)\|_0 \leqslant C\frac{h}{\delta(x)}\|\tau_h\|_{0;B^2_{\delta;x}}, \quad \tau_h \in V_h^k. \tag{5.6}$$

LEMMA 5.4 Let $(\sigma, u) \in H(\mathrm{div};\Omega) \times L^2(\Omega)$ be the solution of (2.1) and $(\sigma_h, u_h) \in V_h^k \times M_h^k$ be its finite element approximation. Let $(\lambda, w) \in H(\mathrm{div};\Omega) \times L^2(\Omega)$ be the solution of the dual problem (3.1) and $(\lambda_h, w_h) \in V_h^k \times M_h^k$ be its finite element approximation. Then, we have the bounds

$$\|\lambda - \lambda_h\|^2_{L(1;2)} \leqslant Ch|\log h|\|\chi(\sigma - \sigma_h)\|_0^2 + h^3|\log h|^2\|\operatorname{div}(\chi(\sigma - \sigma_h))\|_0^2 \quad \text{if } k = 0, \tag{5.7a}$$

$$\|\lambda - \lambda_h\|^2_{L(1;2)} \leqslant Ch\|\chi(\sigma - \sigma_h)\|_0^2 + h^3\|\operatorname{div}(\chi(\sigma - \sigma_h))\|_0^2 \quad \text{if } k \geqslant 1. \tag{5.7b}$$

*Proof.* We start with the proof of (5.7a). For $c \geqslant 1$ sufficiently large, we use (5.2) to write

$$\|\lambda - \lambda_h\|^2_{L(1;2)} \leqslant Ch\|\lambda - \lambda_h\|^2_{0;S_{2ch}} + C|\log h|\|\sqrt{\delta}(\lambda - \lambda_h)\|^2_{0;\Omega\setminus S_{2ch}} \tag{5.8}$$

and note that we have to bound the two terms on the right-hand side separately. Recalling $\operatorname{div}\lambda = \operatorname{div}\lambda_h = 0$, we easily get, from (3.1) and (3.3), the stability estimate

$$\sqrt{h}\|\lambda - \lambda_h\|_{0;S_{ch}} \leqslant \sqrt{h}\|\lambda - \lambda_h\|_0 \leqslant \sqrt{h}\|\chi(\sigma - \sigma_h)\|_0; \tag{5.9}$$

hence, we have a stronger estimate for the first term on the right-hand side of (5.8) than required.

For the treatment of $\|\sqrt{\delta}(\lambda - \lambda_h)\|_{0;\Omega\setminus S_{2ch}}$, we assume that $c \geqslant 1$ is so large that $w$ is harmonic on $\Omega \setminus S_{ch}$. We fix $x \in \Omega$. We start by considering the $L^2$ norm of $\lambda - \lambda_h$ restricted to $B^1_{\delta;x}$. Using the superapproximation property (5.6), the Galerkin orthogonality (3.3a), and the fact that $\delta(x) \geqslant h$, we find

$$\begin{aligned}
\|\lambda - \lambda_h\|^2_{0;B^1_{\delta;x}} &\leqslant (\chi_x(\lambda - \lambda_h), \lambda - \lambda_h)_0 = (\chi_x(\lambda - I_h^k\lambda), \lambda - \lambda_h)_0 + (\chi_x(I_h^k\lambda - \lambda_h), \lambda - \lambda_h)_0 \\
&\leqslant C\|\lambda - \lambda_h\|_{0;B^2_{\delta;x}}\left(\|\lambda - I_h^k\lambda\|_{0;B^2_{\delta;x}} + \frac{h}{\delta(x)}\|\lambda_h - I_h^k\lambda\|_{0;B^2_{\delta;x}}\right) \\
&\quad + b(I_h^k(\chi_x(I_h^k\lambda - \lambda_h)), w_h - w) \\
&\leqslant C\|\lambda - \lambda_h\|_{0;B^2_{\delta;x}}\left(\|\lambda - I_h^k\lambda\|_{0;B^2_{\delta;x}} + \frac{h}{\delta(x)}\|\lambda_h - \lambda\|_{0;B^2_{\delta;x}}\right) \\
&\quad + b(I_h^k(\chi_x(I_h^k\lambda - \lambda_h)), w_h - \Pi_h^* w).
\end{aligned} \tag{5.10}$$

To estimate the contribution from the bilinear form $b(\cdot,\cdot)$, we use the properties (2.3a), (2.3b), the bound for $\nabla\chi_x$, the product rule, and the fact that $I_h^k\lambda$ and $\lambda_h$ are divergence-free, to obtain

$$\begin{aligned}
b(I_h^k(\chi_x(I_h^k\lambda - \lambda_h)), w_h - \Pi_h^* w) &= b(\chi_x(I_h^k\lambda - \lambda_h), w_h - \Pi_h^* w) \\
&\leqslant \|\operatorname{div}(\chi_x(I_h^k\lambda - \lambda_h))\|_0\|w_h - \Pi_h^* w\|_{0;B^2_{\delta;x}} \\
&\leqslant \frac{C}{\delta(x)}\|I_h^k\lambda - \lambda_h\|_{0;B^2_{\delta;x}}\|w_h - \Pi_h^* w\|_{0;B^2_{\delta;x}}.
\end{aligned} \tag{5.11}$$

Next, we consider a countable, locally finite covering of $\Omega \setminus S_{2ch}$ by balls $\{B^1_{\delta;x_i}\}_{x_i}$ such that the associated covering $\{B^2_{\delta;x_i}\}_{x_i}$ is locally finite; for details of the construction we refer to Melenk & Wohlmuth (2012, Appendix). We assume, furthermore, that the sets $B^2_{\delta;x_i}$ are contained in $\Omega \setminus S_{ch}$. We note that, for each $y \in B^2_{\delta;x_i}$, we have the equivalence of $\delta(y)$ and $\delta(x_i)$.

Applying Young's inequality for $\epsilon > 0$, using (5.9–5.11), we get, by summation and the fact that the coverings are locally finite,

$$\begin{aligned}
\|\sqrt{\delta}(\lambda-\lambda_h)\|_0^2 &\leqslant C\sum_{x_i}\delta(x_i)\|\lambda-\lambda_h\|^2_{0;B^1_{\delta;x_i}} + \|\sqrt{\delta}(\lambda-\lambda_h)\|^2_{0;S_{2ch}} \\
&\leqslant C\sum_{x_i}(\delta(x_i)\|\lambda - I_h^k\lambda\|_{0;B^2_{\delta;x_i}}\|\lambda-\lambda_h\|_{0;B^2_{\delta;x_i}} + h\|\lambda_h-\lambda\|^2_{0;B^2_{\delta;x_i}}) \\
&\quad + C\sum_{x_i}\frac{1}{\sqrt{h}}\|w_h - \Pi_h^* w\|_{0;B^2_{\delta;x_i}}\sqrt{h}\|I_h^k\lambda-\lambda_h\|_{0;B^2_{\delta;x_i}} + h\|(\lambda-\lambda_h)\|^2_{0;S_{2ch}} \\
&\leqslant C\left(\left(\frac{1}{\epsilon}+1\right)\|\sqrt{\delta}(\lambda - I_h^k\lambda)\|^2_{0;\Omega\setminus S_{ch}} + \epsilon\|\sqrt{\delta}(\lambda-\lambda_h)\|^2_{0;\Omega\setminus S_{ch}}\right. \\
&\quad \left. + h\|\lambda_h-\lambda\|_0^2 + \frac{1}{h}\|w_h-\Pi_h^* w\|_0^2\right). \qquad (5.12)
\end{aligned}$$

Selecting $\varepsilon > 0$ sufficiently small, but fixed, to absorb the term $\varepsilon\|\sqrt{\delta}(\lambda-\lambda_h)\|^2_{0;\Omega\setminus S_{ch}}$ of the right-hand side in the left-hand side, we obtain from (2.6a) with $s=0$ and the observation $\lambda = -\nabla w$ for some suitable $c' < c$,

$$\begin{aligned}
\|\sqrt{\delta}(\lambda-\lambda_h)\|_0^2 &\leqslant C\left(h^2\|\sqrt{\delta}\nabla\lambda\|^2_{0;\Omega\setminus S_{c'h}} + h\|\lambda_h-\lambda\|_0^2 + \frac{1}{h}\|w_h-\Pi_h^* w\|_0^2\right) \\
&\leqslant C(h^2\|\sqrt{\delta}\nabla^2 w\|^2_{0;\Omega\setminus S_{c'h}} + h\|\chi(\sigma-\sigma_h)\|_0^2).
\end{aligned}$$

In the last step, we have first used the superapproximation property (3.4) which bounds $\|w_h - \Pi_h^* w\|_{0;\Omega}$, and second (5.9). To bound $\|\sqrt{\delta}\nabla^2 w\|_0$, we use (3.11b) and note that this term introduces another log factor in the upper bound.

The proof of (5.7b) is quite similar. The main difference is that we do not use the weighted $L^2$ norm of $\nabla^2 w$ given by (3.11a) but the standard $L^2$ estimate given in Lemma 3.1(iii). For small $\alpha > 0$ we compute, similarly as in (5.12),

$$\begin{aligned}
\|\delta^{1/2+\alpha}(\lambda-\lambda_h)\|_0^2 &\leqslant C\sum_{x_i}\delta(x_i)^{1+2\alpha}\|\lambda-\lambda_h\|^2_{0;B^1_{\delta;x_i}} + \|\delta^{1/2+\alpha}(\lambda-\lambda_h)\|^2_{0;S_{2ch}} \\
&\leqslant C\sum_{x_i}\left(\delta(x_i)^{1+2\alpha}\|\lambda - I_h^k\lambda\|_{0;B^2_{\delta;x_i}}\|\lambda-\lambda_h\|_{0;B^2_{\delta;x_i}} + h\frac{1}{\delta(x_i)}\|\delta^{1/2+\alpha}(\lambda_h-\lambda)\|^2_{0;B^2_{\delta;x_i}}\right) \\
&\quad + C\sum_{x_i}\delta(x_i)^{-1/2+\alpha}\|w_h-\Pi_h^* w\|_{0;B^2_{\delta;x_i}}\delta(x_i)^{1/2+\alpha}\|I_h^k\lambda-\lambda_h\|_{0;B^2_{\delta;x_i}} + h^{1+2\alpha}\|(\lambda-\lambda_h)\|^2_{0;S_{2ch}}
\end{aligned}$$

$$\leqslant C\left(\left(\frac{1}{\epsilon}+1\right)\|\delta^{1/2+\alpha}(\lambda-I_h^k\lambda)\|^2_{0;\Omega\setminus S_{ch}}+\epsilon\|\delta^{1/2+\alpha}(\lambda-\lambda_h)\|^2_{0;\Omega\setminus S_{ch}}\right.$$
$$\left.+\sup_{x\in\Omega\setminus S_{2ch}}\frac{h}{\delta(x)}\|\delta^{1/2+\alpha}(\lambda_h-\lambda)\|_0^2+\frac{1}{\varepsilon}h^{-1+2\alpha}\|w_h-\Pi_h^* w\|_0^2+\varepsilon\|\delta^{1/2+\alpha}(I_h^k\lambda-\lambda_h)\|_0^2\right).$$

The constant $C>0$ is independent of $c$, and $\varepsilon>0$. Selecting now $\varepsilon$ sufficiently small and fixing $c\geqslant 1$ sufficiently large so as to make $\sup_{x\in\Omega\setminus S_{ch}} h/\delta(x)$ sufficiently small, we arrive at

$$\|\delta^{1/2+\alpha}(\lambda-\lambda_h)\|_0^2\leqslant C\|\delta^{1/2+\alpha}(\lambda-I_h^k\lambda)\|^2_{0;\Omega\setminus S_{ch}}+h^{-1+2\alpha}\|w_h-\Pi_h^* w\|_0^2. \tag{5.13}$$

Selecting $\varepsilon>0$ sufficiently small, we get, from (2.6a) with $s=0$ and the observation $\lambda=-\nabla w$ for some $c'<c$,

$$\|\delta^{1/2+\alpha}(\lambda-\lambda_h)\|_0^2\leqslant C\left(h^2\|\delta^{1/2+\alpha}(\lambda-I_h^k\lambda)\|^2_{0;\Omega\setminus S_{c'h}}+h\|\lambda_h-\lambda\|_0^2+\frac{1}{h}\|w_h-\Pi_h^* w\|_0^2\right)$$
$$\leqslant C(h^2\|\delta^{1/2+\alpha}\nabla^2 w\|^2_{0;\Omega\setminus S_{c'h}}+h\|\chi(\sigma-\sigma_h)\|_0^2).$$

Since we assume $k\geqslant 1$ and have $\lambda=-\nabla w$, we get, with the approximation properties of $I_h^k$, (5.13) and (3.10c),

$$\begin{aligned}h^{-2\alpha}\|\delta^{1/2+\alpha}(\lambda-\lambda_h)\|_0^2&\leqslant Ch^{4-2\alpha}\|\delta^{1/2+\alpha}\nabla^2\lambda\|_0^2+h^{-1}\|w_h-\Pi_h^* w\|_0^2\\&\leqslant Ch^{4-2\alpha}\|\delta^{-1/2+\alpha}\nabla^2 w\|_0^2+h^{-1}\|w_h-\Pi_h^* w\|_0^2\\&\leqslant Ch^{4-2\alpha}h^{-1+2\alpha}\|\nabla^2 w\|_0^2+h^{-1}\|w_h-\Pi_h^* w\|_0^2.\end{aligned}$$

Recalling (3.5) gives $\|\nabla^2 w\|_0\leqslant C\|\operatorname{div}(\chi(\sigma-\sigma_h))\|_0$. Then (3.4), (5.9) yield the desired estimate. □

REMARK 5.5 We note that in Lemma 5.4, the log factor appears only for $\mathrm{RT}_h^0$ elements, whereas, in Lemma 5.2, the log factor arises for both $\mathrm{RT}_h^0$ and $\mathrm{BDM}_h^1$ elements.

### 5.4 *Proof of the main result, Theorem* 2.3

Theorem 2.3 is obtained by combining Lemmas 2.2, 5.1, 5.2 and 5.4. Since the case of higher-order elements is proved very similarly to the case of lowest-order elements, we restrict the presentation to that case and leave the other one to the reader.

Inserting the result of Lemma 5.4 and the bound (5.4) into Lemma 5.1, we find

$$\|\sqrt{\chi}(\sigma-\sigma_h)\|_0^4\leqslant Ch^{2k+3}|\log h|^2\|\sigma\|^2_{B^{k+3/2}_{2,1}}(\|\chi(\sigma-\sigma_h)\|_0^2+h^2\|\operatorname{div}(\chi(\sigma-\sigma_h))\|_0^2).$$

Since $\operatorname{div}(\chi(\sigma-\sigma_h))=\nabla\chi\cdot(\sigma-\sigma_h)+\chi\operatorname{div}(\sigma-\sigma_h)$, we get, in view of (3.2), (2.8) and (2.10),

$$\|\operatorname{div}(\chi(\sigma-\sigma_h))\|_0^2\leqslant C\left(\frac{1}{h^2}\|\sqrt{\chi}(\sigma-\sigma_h)\|_0^2+\|\chi(\operatorname{div}\sigma-\Pi_h^*\operatorname{div}\sigma)\|_0^2\right)$$
$$\leqslant C\left(\frac{1}{h^2}\|\sqrt{\chi}(\sigma-\sigma_h)\|_0^2+h^{2k+1}\|\sigma\|^2_{B^{k+3/2}_{2,1}}\right).$$

Finally this bound results in

$$\|\sqrt{\chi}(\sigma - \sigma_h)\|_0^4 \leqslant Ch^{2k+3}|\log h|^2 \|\sigma\|^2_{B^{k+3/2}_{2,1}}(\|\sqrt{\chi}(\sigma - \sigma_h)\|_0^2 + h^{2k+3}\|\sigma\|^2_{B^{k+3/2}_{2,1}})$$

and thus $\|(\sigma - \sigma_h)\|_{0;S_h} \leqslant \|\sqrt{\chi}(\sigma - \sigma_h)\|_0 \leqslant Ch^{k+3/2}|\log h|\|\sigma\|_{B^{k+3/2}_{2,1}}$. Now, our main result, the *a priori* bound (2.13), follows from Lemma 2.2.

## 6. Numerical results

In this section, we present two examples to confirm the theoretical convergence rates for the Laplace operator and one example with an application to the Stokes–Darcy coupling. In all three examples, we consider problem settings with a given solution on $\Omega \subset \mathbb{R}^2$. In addition to the finite elements on triangles, which we introduced in Section 2.1, we also consider finite elements on quadrilaterals such as $\mathrm{RT}_h^{[k]}$, $\mathrm{BDM}_h^{[k]}$ and define Brezzi–Douglas–Fortin–Marini $(\mathrm{BDFM})_h^{[k]}$, $k \in \mathbb{N}_0$ (Fig. 1). Following the notation of Brezzi & Fortin (1991, Chapter V.2), the superscript $[\cdot]$ indicates the association to quadrilateral elements. Note that, unlike the triangular case where we have that $\mathrm{RT}_h^k \subset \mathrm{BDM}_h^{k+1} \subset \mathrm{RT}_h^{k+1}$, this relationship does not hold any more for the quadrilateral case, since $\mathrm{div}(\mathrm{RT}_h^{[k]}) = Q_k \nsubseteq P_k = \mathrm{div}(\mathrm{BDM}_h^{[k+1]})$. Figure 1 illustrates the elements employed and shows the number of degrees of freedom associated with the edges and elements.

### 6.1 *Two-dimensional model problems*

We set $\Omega = (0, 1) \times (0, 1)$. The exact solution $u$ is prescribed in terms of polar coordinates $(r, \phi)$ as

$$u = r^\alpha \sin \phi;$$

here, the origin of the coordinate system is taken to be the point $(0.75, 0.5)$ and the angular variable $\phi$ measures the angle from the line $y = 0.5$. The interface $\Gamma$ is placed at $y = 0.5$ and is resolved by the mesh. The initial mesh, however, is such that $(0.75, 0.5)$ is not a mesh point.

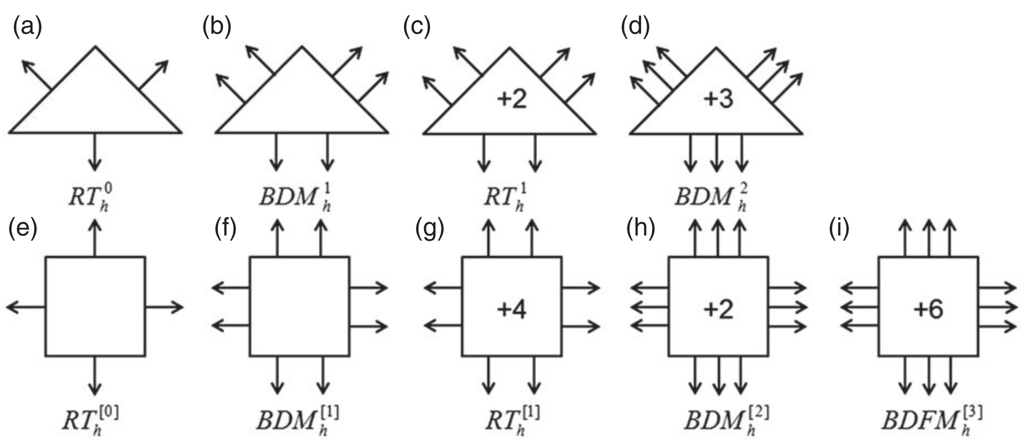


FIG. 1. Degrees of freedom of RT and BDM elements on triangles and quadrilaterals.

TABLE 1 *$L^2$ error in the flux for $\alpha = 1.5$*

| Level | $RT_h^0$ | Rate | $BDM_h^1$ | Rate | $RT_h^1$ | Rate | $BDM_h^2$ | Rate |
|---|---|---|---|---|---|---|---|---|
| 1 | 9.42e−2 | — | 4.84e−2 | — | 2.20e−2 | — | 1.75e−2 | — |
| 2 | 5.03e−2 | 0.904 | 2.48e−2 | 0.964 | 1.07e−2 | 1.040 | 8.76e−3 | 1.000 |
| 3 | 2.70e−2 | 0.896 | 1.24e−2 | 1.003 | 5.36e−3 | 1.003 | 4.37e−3 | 1.002 |
| 4 | 1.44e−2 | 0.905 | 6.19e−3 | 1.001 | 2.68e−3 | 1.001 | 2.19e−3 | 1.001 |
| 5 | 7.65e−3 | 0.916 | 3.09e−3 | 1.000 | 1.34e−3 | 1.001 | 1.09e−3 | 1.001 |
| 6 | 4.03e−3 | 0.925 | 1.55e−3 | 1.000 | 6.64e−4 | 1.000 | 5.45e−4 | 1.002 |

TABLE 2 *$L^2$ error in the flux for $\alpha = 1.75$*

| Level | $RT_h^0$ | Rate | $BDM_h^1$ | Rate | $RT_h^1$ | Rate | $BDM_h^2$ | Rate |
|---|---|---|---|---|---|---|---|---|
| 1 | 8.44e−2 | — | 3.36e−2 | — | 1.12e−2 | — | 8.35e−3 | — |
| 2 | 4.23e−2 | 0.996 | 1.44e−2 | 1.225 | 4.40e−3 | 1.349 | 3.51e−3 | 1.252 |
| 3 | 2.14e−2 | 0.985 | 5.92e−3 | 1.280 | 1.84e−3 | 1.257 | 1.48e−3 | 1.248 |
| 4 | 1.08e−2 | 0.987 | 2.46e−3 | 1.265 | 7.72e−4 | 1.252 | 6.21e−4 | 1.250 |
| 5 | 5.43e−3 | 0.990 | 1.03e−3 | 1.256 | 3.24e−4 | 1.252 | 2.61e−4 | 1.249 |
| 6 | 2.73e−3 | 0.993 | 4.33e−4 | 1.252 | 1.36e−4 | 1.257 | 1.10e−4 | 1.250 |

We focus on the choices $\alpha = 1.5$ and $\alpha = 1.75$. We note that, for $\frac{3}{2} < \alpha < 2$, the solution is in $B^{3/2}_{2,1}(\Omega)$ but not in $W^{2,\infty}(\Omega)$, and that full convergence rates cannot be expected owing to limited regularity. The numerical solution $(\sigma_h, u_h) \in V_h^k \times M_h^k$ is uniquely characterized by (2.7). We note that inhomogeneous Dirichlet boundary conditions $g_D$ enter as $\int_{\partial\Omega} g_D \tilde{\sigma}\, \mathbf{n}\, ds$ on the right-hand side of (2.1a). Here, we set $g_D$ equal to the trace of the given exact solution. The accuracy of the numerical solution for different choices of the pairings $(V_h^k, M_h^k)$ and on different types of meshes is compared.

6.1.1 *Simplicial mesh.* In Tables 1 and 2, the $L^2$ errors of the normal fluxes across the interface are presented for the limiting case $\alpha = \frac{3}{2}$ and the case $\alpha = 1.75$ for various levels of uniform refinement. On unstructured simplicial meshes we use $RT_h^0$, $RT_h^1$, $BDM_h^1$ and $BDM_h^2$ elements. As expected from the theory, the asymptotic convergence rates are determined by the low regularity of the problem. We point out that the singularity is not placed at a vertex of the initial mesh.

In Figs 2 and 3, the normal flux of the numerical solutions is plotted against that of the exact solution $u$ for both choices of $\alpha$. We point out that, only for $RT_h^0$, the flux is approximated by a piecewise constant, whereas for the cases $RT_h^1$ and $BDM_h^1$ the flux is approximated by linears and for $BDM_h^2$ by quadratics.

6.1.2 *Quadrilateral mesh.* Tables 3 and 4 show the same type of results but for mixed finite elements on a quadrilateral mesh, in particular for $RT_h^{[0]}$, $RT_h^{[1]}$, $BDM_h^{[1]}$, $BDM_h^{[2]}$ and $BDFM_h^{[3]}$ elements. In Table 3, the case $\alpha = 1.5$ is displayed. Here, we cannot expect a higher convergence rate than 1 independently of the choice of the finite element order. We recall that $BDM_h^{[1]}$ and $RT_h^{[1]}$ have the same degrees of freedom per edge but $RT_h^{[1]}$ has additional interior degrees of freedom, and thus has a significantly

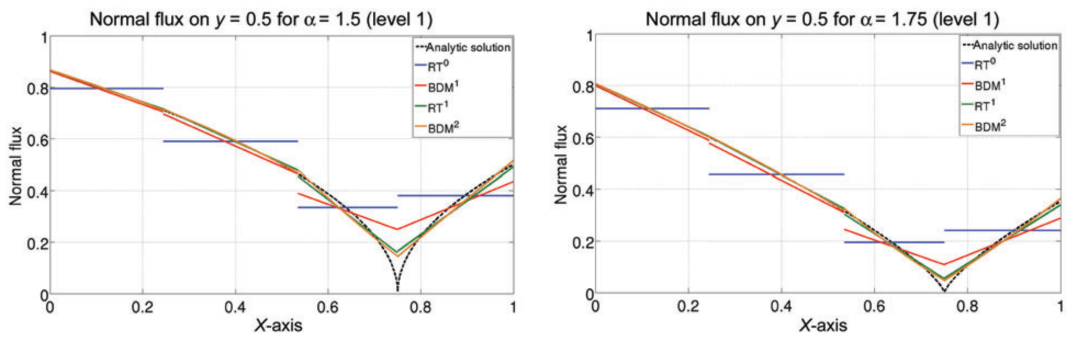


FIG. 2. Numerical approximation of the normal flux across the interface $\Gamma$ for $\alpha = 1.5$ (left) and $\alpha = 1.75$ (right) on refinement level 1 (simplicial mesh).

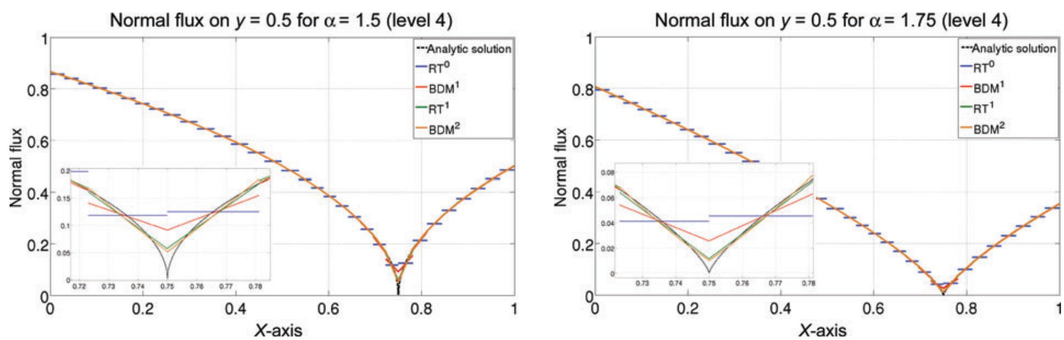


FIG. 3. Numerical approximation of the normal flux across the interface $\Gamma$ for $\alpha = 1.5$ (left) and $\alpha = 1.75$ (right) on refinement level 4 (simplicial mesh).

TABLE 3 *$L^2$ error in the flux for $\alpha = 1.5$, quadrilateral mesh*

| Level | $\mathrm{RT}_h^{[0]}$ | Rate | $\mathrm{BDM}_h^{[1]}$ | Rate | $\mathrm{RT}_h^{[1]}$ | Rate | $\mathrm{BDM}_h^{[2]}$ | Rate | $\mathrm{BDFM}_h^{[3]}$ | Rate |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1.18e−1 | — | 8.98e−2 | — | 8.41e−2 | — | 6.59e−2 | — | 5.28e−2 | — |
| 1 | 9.28e−2 | 0.348 | 5.97e−2 | 0.589 | 1.81e−2 | 2.218 | 3.40e−2 | 0.955 | 1.88e−2 | 1.486 |
| 2 | 5.14e−2 | 0.852 | 3.07e−2 | 0.956 | 8.73e−3 | 1.051 | 1.65e−2 | 1.038 | 9.45e−3 | 0.996 |
| 3 | 2.77e−2 | 0.890 | 1.54e−2 | 0.998 | 4.37e−3 | 0.999 | 8.28e−3 | 1.000 | 4.72e−3 | 1.000 |
| 4 | 1.48e−2 | 0.902 | 7.68e−3 | 1.004 | 2.18e−3 | 1.000 | 4.14e−3 | 1.000 | 2.36e−3 | 1.000 |
| 5 | 7.88e−3 | 0.913 | 3.84e−3 | 1.001 | 1.09e−3 | 1.000 | 2.07e−3 | 1.000 | 1.18e−3 | 1.000 |
| 6 | 4.16e−3 | 0.923 | 1.92e−3 | 1.000 | 5.46e−4 | 1.000 | 1.03e−3 | 1.000 | 5.91e−4 | 1.000 |

smaller error on all levels. Although $\mathrm{BDM}_h^{[2]}$ and $\mathrm{BDFM}_h^{[3]}$ have more degrees of freedom per edge than $\mathrm{RT}_h^{[1]}$, the quantitative errors are sensitive to the number of degrees of freedom per element.

Compared with $\alpha = 1.5$, the solution for $\alpha = 1.75$ is more regular, and thus we only expect for the lowest-order $\mathrm{RT}_h^{[0]}/\mathrm{RT}_h^0$ discretization an asymptotic rate of 1. In all other cases, we observe asymptotically a rate of $\sim 1.25$. There is no qualitative difference between an unstructured simplicial mesh and a regular quadrilateral mesh.

TABLE 4 $L^2$ *error in the flux for* $\alpha = 1.75$, *quadrilateral mesh*

| Level | $RT_h^{[0]}$ | Rate | $BDM_h^{[1]}$ | Rate | $RT_h^{[1]}$ | Rate | $BDM_h^{[2]}$ | Rate | $BDFM_h^{[3]}$ | Rate |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1.21e−1 | — | 8.36e−2 | — | 6.82e−2 | — | 4.79e−2 | — | 3.55e−2 | — |
| 1 | 8.33e−2 | 0.540 | 4.41e−2 | 0.923 | 7.81e−3 | 3.125 | 2.07e−2 | 1.212 | 1.05e−2 | 1.758 |
| 2 | 4.29e−2 | 0.959 | 1.89e−2 | 1.222 | 2.92e−3 | 1.418 | 8.35e−3 | 1.307 | 4.44e−3 | 1.240 |
| 3 | 2.16e−2 | 0.986 | 7.83e−3 | 1.272 | 1.23e−3 | 1.247 | 3.52e−3 | 1.246 | 1.87e−3 | 1.250 |
| 4 | 1.09e−2 | 0.987 | 3.25e−3 | 1.267 | 5.21e−4 | 1.243 | 1.48e−3 | 1.249 | 7.85e−4 | 1.250 |
| 5 | 5.49e−3 | 0.990 | 1.36e−3 | 1.257 | 2.19e−4 | 1.249 | 6.23e−4 | 1.250 | 3.30e−4 | 1.250 |
| 6 | 2.76e−3 | 0.993 | 5.71e−4 | 1.253 | 9.26e−5 | 1.250 | 2.62e−4 | 1.250 | 1.39e−4 | 1.250 |

TABLE 5 $L^2$ *error in the flux, quadrilateral mesh*

| Level | $RT_h^{[0]}$ | Rate | $BDM_h^{[1]}$ | Rate | $RT_h^{[1]}$ | Rate | $BDM_h^{[2]}$ | Rate | $BDFM_h^{[3]}$ | Rate |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.69e−0 | — | 1.55e−0 | — | 1.10e−0 | — | 2.64e−0 | — | 1.49e−0 | — |
| 2 | 8.13e−1 | 1.053 | 8.45e−1 | 0.871 | 4.15e−1 | 1.408 | 7.49e−1 | 1.819 | 3.32e−1 | 2.167 |
| 3 | 4.08e−1 | 0.996 | 2.11e−1 | 2.004 | 4.83e−2 | 3.103 | 1.41e−1 | 2.414 | 2.43e−2 | 3.774 |
| 4 | 1.81e−1 | 1.168 | 6.51e−2 | 1.695 | 8.55e−3 | 2.499 | 1.86e−2 | 2.915 | 1.24e−3 | 4.291 |
| 5 | 8.51e−2 | 1.091 | 1.75e−2 | 1.890 | 2.08e−3 | 2.042 | 2.33e−3 | 3.001 | 7.11e−5 | 4.123 |
| 6 | 4.18e−2 | 1.027 | 4.53e−3 | 1.953 | 5.19e−4 | 2.001 | 2.91e−4 | 3.001 | 5.78e−6 | 3.622 |
| 7 | 2.08e−2 | 1.012 | 1.15e−3 | 1.979 | 5.19e−4 | 2.001 | 3.63e−5 | 3.000 | 6.16e−7 | 3.230 |

TABLE 6 $L^2$ *error in the flux, simplicial mesh*

| Level | $RT_h^0$ | Rate | $BDM_h^1$ | Rate | $RT_h^1$ | Rate | $BDM_h^2$ | Rate |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.83e−0 | — | 1.67e−0 | — | 1.57e−0 | — | 1.90e−0 | — |
| 2 | 9.26e−1 | 0.987 | 8.35e−1 | 1.004 | 4.10e−1 | 1.934 | 2.83e−1 | 2.749 |
| 3 | 4.19e−1 | 1.145 | 2.26e−1 | 1.886 | 8.16e−2 | 2.329 | 5.32e−2 | 2.410 |
| 4 | 1.86e−1 | 1.173 | 5.45e−2 | 2.051 | 1.41e−2 | 2.534 | 7.42e−3 | 2.842 |
| 5 | 8.62e−2 | 1.108 | 1.33e−2 | 2.039 | 2.67e−3 | 2.398 | 9.54e−4 | 2.960 |
| 6 | 4.20e−2 | 1.038 | 3.33e−3 | 1.995 | 5.79e−4 | 2.206 | 1.20e−4 | 2.990 |
| 7 | 2.08e−2 | 1.010 | 8.38e−4 | 1.989 | 1.36e−4 | 2.089 | 1.50e−5 | 2.997 |

6.1.3 *Higher-order convergence.* In the second example, we consider the piecewise smooth solution

$$u = \begin{cases} e^{-x}\sin(2\pi y)^2 - (y-\frac{1}{2})[xy(x-\frac{1}{2})^2 + e^y\cos(8\pi xy)(x+\frac{1}{2})^2] & \text{if } y < \frac{1}{2}, \\ e^{-x}(y-\frac{1}{2})^2 4\pi^2 - (y-\frac{1}{2})[xy(x-\frac{1}{2})^2 + e^y\cos(8\pi xy)(x+\frac{1}{2})^2] & \text{if } y \geqslant \frac{1}{2} \end{cases}$$

on $\Omega = (0,1)\times(0,1)$. The interface $\Gamma$ is placed at $y = 0.5$ and resolved by the mesh. Since the exact solution is sufficiently smooth, we expect full convergence rates. Tables 5 and 6 show the numerical results for a uniform quadrilateral and a simplicial mesh, respectively. The observed convergence rates confirm our theoretical result. Note that the absolute errors for $BDM_h^{[2]}$ are larger than for $RT^{[1]}$ on the first four refinement levels although it has an additional degree of freedom per edge.

### 6.2 *Stokes–Darcy coupling*

In this subsection, we consider a more general problem setting that is not covered by our theoretical results. The coupling of the Stokes problem with the Laplace equation plays an important role in many applications. Of special interest are porous media applications where the Darcy velocity can be used to describe a single-phase single-component transport. On the pore scale, the pore structure is resolved and the Navier–Stokes equations model the flow in the free-flow region and within the pores. On the 'representative elementary volume' scale, however, the mathematical model can be considerably simplified by applying the potential theory resulting in Darcy's law in the porous media. Two-domain models exploit this observation and use suitable transfer conditions at the interface to couple the simple Darcy model for porous media with, for example, the simplified Stokes equation in the free-flow domain.

6.2.1 *Model equations and coupling conditions.* The unknown quantities in the Darcy domain $\Omega_D$ are the pressure $p_D$ and the Darcy velocity $\mathbf{u}_D$, whereas in the Stokes domain $\Omega_S$, the velocity $\mathbf{u}_S$ and the pressure $p_S$ are sought. In $\Omega_D$, we consider the first-order velocity system

$$\mathbf{u}_D = -\mathbf{K}\nabla p_D, \quad \operatorname{div}\mathbf{u}_D = f_D,$$

where $f_D$ stands for a source term and $\mathbf{K}$ is a positive-definite tensor that characterizes the intrinsic permeability of the porous medium. In $\Omega_S$, we consider the second-order velocity system

$$-\operatorname{div}(2\mu\mathbf{D}(\mathbf{u}_S)) + \nabla p_S = -\mathbf{f}_S, \quad \operatorname{div}\mathbf{u}_S = \mathbf{0},$$

where $\mathbf{f}_S$ stands for a source term, $\mathbf{D}(\mathbf{u}_S) := (\nabla\mathbf{u}_S + \nabla\mathbf{u}_S^T)/2$ denotes the deformation tensor and $\mu$ is the fluid viscosity. To close the system, we have to specify boundary and coupling conditions. In theory, the coupling conditions can be derived by applying volume-averaging techniques as described by Gray *et al.* (1993) and Whitaker (1999). In practice, however, simplified coupling conditions are often used. Here, we apply in a tangential direction the Beavers–Joseph velocity-jump condition due to Beavers & Joseph (1967) in combination with the modification by Saffman (1971) (see also Jäger & Mikelic, 2000, 2009). This condition can be written as

$$\mathbf{u}_S \cdot \boldsymbol{\tau} - \frac{\sqrt{k}}{\gamma} 2\mathbf{n} \cdot \mathbf{D}(\mathbf{u}_S) \cdot \boldsymbol{\tau} = 0. \tag{6.1}$$

The unit normal vector $\mathbf{n}$ points from $\Omega_S$ to $\Omega_D$, and $\boldsymbol{\tau}$ stands for the tangential vector on the interface. For a parameter $\mu > 0$ the value $k := \boldsymbol{\tau} \cdot \mu\mathbf{K} \cdot \boldsymbol{\tau}$ describes the dynamic viscosity. The parameter $\gamma > 0$ is a dimensionless constant that has to be determined experimentally. In the normal direction, the continuity of normal forces and mass conservation across the interface is assumed:

$$\mathbf{u}_S \cdot \mathbf{n} = \mathbf{u}_D \cdot \mathbf{n}, \tag{6.2}$$

$$p_S - 2\mu\mathbf{n} \cdot \mathbf{D}(\mathbf{u}_S) \cdot \mathbf{n} = p_D; \tag{6.3}$$

see also Layton *et al.* (2002).

6.2.2 *Numerical results.* For the numerical discretization, we follow a hybrid discontinuous Galerkin approach based on mixed finite elements of possibly different orders in both subdomains given by

TABLE 7 *$L^2$-error and convergence rate of velocities on the interface $\Gamma$ for the lowest-order $RT_h^{[0]}/BDM_h^{[1]}$ coupling*

| Level | $\mathbf{u}_S \cdot \mathbf{n}$ error | Rate | $\mathbf{u}_S \cdot \boldsymbol{\tau}$ error | Rate | $\mathbf{u}_D \cdot \mathbf{n}$ error | Rate |
|---|---|---|---|---|---|---|
| 1 | 1.48e−02 | — | 3.06e−02 | — | 2.77e−01 | — |
| 2 | 1.69e−03 | 3.125 | 3.48e−03 | 3.137 | 1.39e−01 | 1.000 |
| 3 | 2.68e−04 | 2.662 | 7.20e−04 | 2.273 | 6.94e−02 | 1.000 |
| 4 | 5.58e−05 | 2.262 | 1.63e−04 | 2.142 | 3.47e−02 | 1.000 |
| 5 | 1.33e−05 | 2.068 | 3.96e−05 | 2.045 | 1.73e−02 | 1.000 |
| 6 | 3.29e−06 | 2.016 | 9.81e−06 | 2.012 | 8.67e−03 | 1.000 |
| 7 | 8.20e−07 | 2.004 | 2.45e−06 | 2.002 | 4.33e−03 | 1.000 |

TABLE 8 *$L^2$-error and convergence rate of velocities on the interface $\Gamma$ for the $RT_h^{[1]}/BDM_h^{[2]}$ coupling*

| Level | $\mathbf{u}_S \cdot \mathbf{n}$ error | Rate | $\mathbf{u}_S \cdot \boldsymbol{\tau}$ error | Rate | $\mathbf{u}_D \cdot \mathbf{n}$ error | Rate |
|---|---|---|---|---|---|---|
| 1 | 1.55e−02 | — | 2.29e−02 | — | 1.54e−02 | — |
| 2 | 1.68e−03 | 3.205 | 1.81e−03 | 3.656 | 1.71e−03 | 3.169 |
| 3 | 2.06e−04 | 3.028 | 1.96e−04 | 3.211 | 2.24e−04 | 2.930 |
| 4 | 2.53e−05 | 3.020 | 2.19e−05 | 3.162 | 3.38e−05 | 2.727 |
| 5 | 3.13e−06 | 3.014 | 2.54e−06 | 3.108 | 6.42e−06 | 2.396 |
| 6 | 3.89e−07 | 3.009 | 3.04e−07 | 3.064 | 1.45e−06 | 2.143 |
| 7 | 4.85e−08 | 3.005 | 3.70e−08 | 3.036 | 3.54e−07 | 2.040 |

Girault & Rivière (2009) and Kanschat & Rivière (2010). Here, we consider the coupled Stokes–Darcy system on the unit square $\Omega := (0,1) \times (0,1)$ which is subdivided into two subdomains,

$$\Omega_S := (0,1) \times (0.5,1), \quad \Omega_D := (0,1) \times (0,0.5),$$

with the exact solution

$$\mathbf{u}_S = (\omega e^{\omega y} \sin(\omega x) + 2x^2(y-\xi) - 2G^3,\ -\omega e^{\omega y}\cos(\omega x) - 2x(y-\xi)^2)^T,$$
$$\mathbf{u}_D = (\omega^2 e^{\omega/2} \sin(\omega x) y - G^2(y+0.5)^2,\ -\omega e^{\omega/2}\cos(\omega x) - 2x(y+0.5)G^2)^T,$$
$$p_S = \omega e^{\omega/2}\cos(\omega x)(K^{-1}y - 2\omega\mu) + Gx(G - 8\mu),$$
$$p_D = K^{-1}(\omega e^{\omega/2}\cos(\omega x) y + x(y+0.5)^2 G^2),$$

where $\mu := 1.0$ , $K := 1.0$ , $\gamma := 2.0$ , $G := \sqrt{K\mu}/\gamma$, $\omega := 1/2G$ and $\xi := 0.5 - G$. The solution is chosen to fulfil the coupling conditions (6.1–6.3) on the interface $\Gamma = \bar{\Omega}_S \cap \bar{\Omega}_D$, and we assume that $\mathbf{K} = \mathrm{diag}(K)$. Note that the velocity has a continuous normal component on $\Gamma$ but is discontinuous in the tangential direction. The domain is discretized by a sequence of uniformly refined quadrilateral meshes, where the numerical solution in $\Omega_S$ is computed by the symmetric interior penalty Galerkin method using $\mathrm{BDM}_h^{[k+1]}$ elements, and in $\Omega_D$ we apply a mixed FEM using $\mathrm{RT}_h^{[k]}$ elements. We mention that the choice of the pairing is motivated by the idea of having the same order of convergence for the two subdomains. In Tables 7 and 8, the $L^2$ errors of the velocities on $\Gamma$ are listed for all refinement

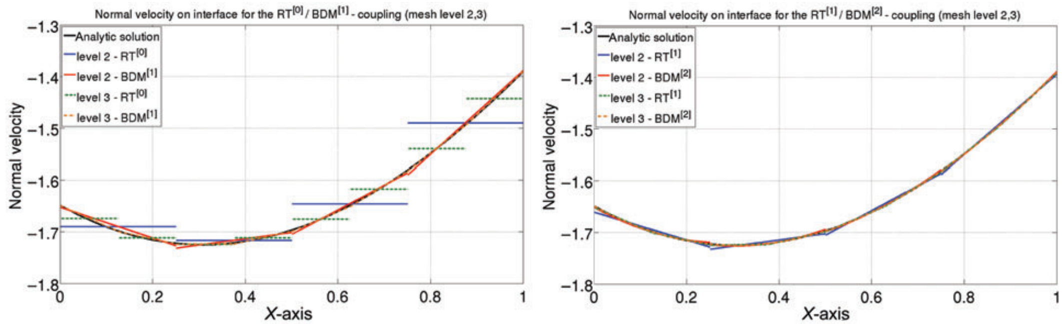


FIG. 4. Numerical approximation of the normal velocity across the interface $\Gamma$ for the $\mathrm{RT}_h^{[0]}/\mathrm{BDM}_h^{[1]}$ coupling (left) and the $\mathrm{RT}_h^{[1]}/\mathrm{BDM}_h^{[2]}$ coupling (right).

TABLE 9 *$L^2(\Omega_\mathrm{D})$ and $H^1(\Omega_\mathrm{S})$ errors for the $\mathrm{RT}_h^{[k]}/\mathrm{BDM}_h^{[k+1]}$ coupling, $k=0$ and $k=1$*

| | $k=0$ | | | | $k=1$ | | | |
|---|---|---|---|---|---|---|---|---|
| Level | $\mathbf{u}_\mathrm{S}$ error in $H^1$ | Rate | $\mathbf{u}_\mathrm{D}$ error in $L^2$ | Rate | $\mathbf{u}_\mathrm{S}$ error in $H^1$ | Rate | $\mathbf{u}_\mathrm{D}$ error in $L^2$ | Rate |
| 1 | 1.05e−01 | — | 2.08e−01 | — | 1.06e−01 | — | 1.40e−02 | — |
| 2 | 2.57e−02 | 2.031 | 1.05e−01 | 0.991 | 2.36e−02 | 2.169 | 3.32e−03 | 2.077 |
| 3 | 7.68e−03 | 1.744 | 5.23e−02 | 0.998 | 5.74e−03 | 2.038 | 8.27e−04 | 2.007 |
| 4 | 2.80e−03 | 1.454 | 2.62e−02 | 0.999 | 1.42e−03 | 2.012 | 2.07e−04 | 2.001 |
| 5 | 1.23e−03 | 1.185 | 1.31e−02 | 1.000 | 3.55e−04 | 2.003 | 5.16e−05 | 2.000 |
| 6 | 5.92e−04 | 1.057 | 6.55e−03 | 1.000 | 8.87e−05 | 2.001 | 1.29e−05 | 2.000 |
| 7 | 2.93e−04 | 1.016 | 3.27e−03 | 1.000 | 2.22e−05 | 2.000 | 3.23e−06 | 2.000 |

levels for $k=0$, 1. In Fig. 4, the normal velocities of the numerical solutions are plotted against the exact solution on two consecutive grid levels. Table 9 shows the $H^1(\Omega_\mathrm{S})$ and the $L^2(\Omega_\mathrm{D})$ errors for the velocities since these norms are natural, given that the Stokes system is (essentially) a second-order equation, whereas the Darcy equation is a first-order system. The results show full convergence rates for both choices of $k$.

**Funding**

The research was carried out while J.M.M. and B.W. attended the 'special semester' on 'Multiscale Simulation & Analysis in Energy and the Environment' hosted by the Johann Radon Institute for Computational and Applied Mathematics (RICAM) in Linz, 3 October 2011–15 December 2011. The hospitality and financial support of the RICAM is very gratefully acknowledged.

## REFERENCES

ARNOLD, D. & BREZZI, F. (1985) Mixed and nonconforming finite element methods: implementation, post-processing and error estimates. *RAIRO, Modél. Math. Anal. Numér.*, **19**, 7–32.

BEAVERS, G. & JOSEPH, D. (1967) Boundary conditions at a naturally permeable wall. *J. Fluid Mech.*, **30**, 197–207.

BREZZI, F. (1989) A survey of mixed finite element methods. *Finite Elements. Theory and Application, Proceeding of the ICASE Workshop*. Hampton, VA 1986. New York: Springer, pp. 34–49.

BREZZI, F. (2003) Stability of saddle-points in finite dimensions. *Frontiers in Numerical Analysis. 10th LMS–EPSRC Numerical Analysis Summer School*, Durham, UK, July 7–19, 2002. Berlin: Springer, pp. 17–61.

BREZZI, F., DOUGLAS, J. & MARINI, L. (1985) Two families of mixed finite elements for second order elliptic problems. *Numer. Math.*, **47**, 217–235.

BREZZI, F., DOUGLAS, J. & MARINI, L. (1986) Recent results on mixed finite element methods for second order elliptic problems. *Vistas in Applied Mathematics. Numerical Analysis, Atmospheric Sciences, Immunology*, pp. 25–43. dedicated G. I. Marchuk occasion of his 60th birthday.

BREZZI, F. & FORTIN, M. (1991) *Mixed and Hybrid Finite Element Methods*. New York: Springer.

GASTALDI, L. & NOCHETTO, R. (1989) Sharp maximum norm error estimates for general mixed finite element approximations to second order elliptic equations. *RAIRO Modél. Math. Anal. Numér.*, **23**, 103–128.

GIRAULT, V. & RIVIÈRE, B. (2009) DG approximation of coupled Navier–Stokes and Darcy equations by Beaver–Joseph–Saffman interface condition. *J. Numer. Anal.*, **47**, 2052–2089.

GRAY, W., LEIJNSE, A., KOLAR, R. & BLAIN, C. (1993) *Mathematical Tools for Changing Spatial Scales in the Analysis of Physical Systems*. Boca Raton, FL: CRC Press.

JÄGER, W. & MIKELIC, A. (2000) On the interface boundary condition of Beavers, Joseph and Saffman. *SIAM J. Appl. Math.*, **60**, 1111–1127.

JÄGER, W. & MIKELIC, A. (2009) Modeling effective interface laws for transport phenomena between an unconfined fluid and a porous medium using homogenization. *Transp. Porous Media*, **78**, 489–508.

KANSCHAT, G. & RIVIÈRE, B. (2010) A strongly conservative finite element method for the coupling of Stokes and Darcy flow. *J. Comput. Phys.*, **229**, 5933–5943.

LAYTON, W., SCHIEWECK, F. & YOTOV, I. (2002) Coupling fluid flow with porous media flow. *SIAM J. Numer. Anal.*, **40**, 2195–2218.

LI, J., MELENK, J., WOHLMUTH, B. & ZOU, J. (2010) Optimal a priori estimates for higher order finite elements for elliptic interface problems. *Appl. Numer. Math.*, **60**, 19–37.

MELENK, J. & WOHLMUTH, B. (2012) Quasi-optimal approximation of surface based Lagrange multipliers in finite element methods. *Siam J. Numer. Anal.*, **50**, 2064–2087.

NÉDÉLEC, J. (1980) Mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, **35**, 315–341.

NICOLAIDES, R. (1982) Existence, uniqueness and approximation for generalized saddle point problems. *SIAM J. Numer. Anal.*, **19**, 349–357.

RAVIART, P. & THOMAS, J. (1977) A mixed finite element method for 2nd order elliptic problems. *Mathematical Aspects of Finite Element Methods, Proceeding of the Conference Held in Rome 1975*. Lecture Notes in Mathematics, vol. 606. Berlin: Springer, pp. 292–315.

SAFFMAN, R. (1971) On the boundary condition at the surface of a porous medium. *Stud. Appl. Math.*, **50**, 93–101.

TARTAR, L. (2007) *An Introduction to Sobolev Spaces and Interpolation Spaces*. Lecture Notes of the Unione Matematica Italiana, vol. 3. Berlin: Springer, pp. xxvi+218.

TRIEBEL, H. (1995) *Interpolation Theory, Function Spaces, Differential Operators*, 2nd edn. Heidelberg: Johann Ambrosius Barth, p. 532.

WAHLBIN, L. (1995) *Superconvergence in Galerkin Finite Element Methods*. Lecture Notes in Mathematics, vol. 1605. Berlin: Springer.

WANG, J. (1989) Asymptotic expansions and $L^\infty$-error estimates for mixed finite element methods for second order elliptic problems. *Numer. Math.*, **55**, 401–430.

WHITAKER, S. (1999) *The Method of Volume Averaging*, vol. 240. Dordrecht: Kluwer Academic Publishers.

WRIGGERS, P. & CARSTENSEN, C. (eds) (2009) *Mixed Finite Element Technologies*. CISM Courses and Lectures, vol. 509. Wien: Springer.

**Appendix A. Regularity**

In Lemma 3.1, we showed assertion (iv) under the assumption of the convexity of $\Omega$, i.e., under the assumption of full regularity. This assumption can be weakened: it suffices that $\Omega$ admits a shift theorem by more than $1/2$ as we now show.

For simplicity of exposition, we formulate this shift theorem as an assumption but point out that, for example, for $d=2$ it is valid for polygonal Lipschitz domains $\Omega$.

ASSUMPTION A1 There exists $s_0 > 1/2$ such that $\tilde{T}^D$ is a bounded linear operator $\tilde{T}^D : H^{-1+s_0}(\Omega) \to H^{1+s_0}(\Omega)$. Here, $\tilde{T}^D : H^{-1}(\Omega) \to H_0^1(\Omega)$ is the solution operator for the Poisson problem: given $g \in H^{-1}(\Omega)$, find $y \in H_0^1(\Omega)$ such that

$$-\Delta y = g \quad \text{in } \Omega, \quad y = 0 \quad \text{on } \partial\Omega.$$

We now show an analogue of Melenk & Wohlmuth (2012, Lemma 5.2).

LEMMA A2 Assume the validity of the shift theorem of Assumption A1. Consider the following variational problem: given $g \in (L^2(\Omega))^d$, find $y \in H_0^1(\Omega)$ such that

$$(\nabla y, \nabla z)_{0,\Omega} = (g, \nabla z)_{0,\Omega} \quad \forall z \in H_0^1(\Omega).$$

Then, the solution operator $T^D : (L^2(\Omega))^d \to H_0^1(\Omega)$, given by $g \mapsto y$, extends to a bounded linear map

$$T^D : ((B^{1/2}_{2,1}(\Omega))^d)' \to B^{1/2}_{2,\infty}(\Omega).$$

*Proof. Step* 1. Our starting point is a very weak formulation. Fix $\varepsilon > 0$ such that $1/2 + \varepsilon < s_0$. We introduce the bilinear form $B$ on $H^{1/2-\varepsilon}(\Omega) \times (H^{3/2+\varepsilon}(\Omega) \cap H_0^1(\Omega))$ by

$$B(y, v) := \langle y, -\Delta v \rangle_{1/2-\varepsilon,-1/2+\varepsilon}.$$

A few comments are in order: first, $\langle \cdot, \cdot \rangle_{1/2-\varepsilon,-1/2+\varepsilon}$ stands for the duality pairing between $H^{1/2-\varepsilon}(\Omega)$ and $H^{-1/2+\varepsilon}(\Omega)$. We point out that the assumption $\varepsilon > 0$ implies that $H^{1/2-\varepsilon}(\Omega) = H_0^{1/2-\varepsilon}(\Omega)$ so that the duality pairing is indeed well defined. Second, from the mapping properties of $-\Delta$ (taken in the distributional sense), $-\Delta : H^2(\Omega) \to L^2(\Omega)$ and $-\Delta : H^1(\Omega) \to H^{-1}(\Omega)$, we obtain by interpolation that $-\Delta : H^{3/2+\varepsilon}(\Omega) \to H^{-1/2+\varepsilon}(\Omega)$ so that $B$ is indeed well defined.

We claim that $B$ satisfies an inf–sup condition. To that end, let $u' \in C_0^\infty(\Omega)$ be arbitrary. By our assumptions on the mapping properties of $\tilde{T}^D$ stated in Assumption A1, there exists $v \in H^{3/2+\varepsilon}(\Omega) \cap H_0^1(\Omega)$ such that

$$(\nabla v, \nabla z)_{0,\Omega} = (u', z)_{0,\Omega} = \langle u', z \rangle_{-1/2+\varepsilon,1/2-\varepsilon} \quad \forall z \in C_0^\infty(\Omega),$$

together with the bound $\|v\|_{3/2+\varepsilon;\Omega} \leqslant C\|u'\|_{-1/2+\varepsilon,\Omega}$. By the definition of the distributional Laplacian, we obtain

$$B(z, v) = \langle u', z \rangle_{-1/2+\varepsilon,1/2-\varepsilon} \quad \forall z \in C_0^\infty(\Omega).$$

Taking the supremum over $z \in C_0^\infty(\Omega)$, recalling the density of $C_0^\infty(\Omega)$ in $H^{1/2-\varepsilon}(\Omega) = H_0^{1/2-\varepsilon}(\Omega)$ and using the bound $\|v\|_{3/2+\varepsilon;\Omega} \leqslant C\|u'\|_{-1/2+\varepsilon;\Omega}$, we obtain

$$\sup_{z \in C_0^\infty(\Omega)} \frac{B(z, v)}{\|z\|_{1/2-\varepsilon;\Omega}\|v\|_{3/2+\varepsilon;\Omega}} = \sup_{z \in C_0^\infty(\Omega)} \frac{\langle u', z \rangle_{-1/2+\varepsilon,1/2-\varepsilon}}{\|z\|_{1/2-\varepsilon;\Omega}\|v\|_{3/2+\varepsilon;\Omega}} = \frac{\|u'\|_{-1/2+\varepsilon;\Omega}}{\|v\|_{3/2+\varepsilon;\Omega}} \geqslant C > 0.$$

Furthermore, the bilinear form $B$ satisfies the 'sup–sup' condition so that the bilinear form $B$ induces an isomorphism between $(H^{3/2+\varepsilon}(\Omega)\cap H^1_0(\Omega))'$ and $H^{1/2-\varepsilon}(\Omega)$.

*Step* 2. Consider the following problem: given $g\in((H^{1/2+\varepsilon}(\Omega))^d)'$, find $y\in H^{1/2-\varepsilon}(\Omega)$ such that

$$B(y,z)=\langle g,\nabla z\rangle_{(H^{1/2+\varepsilon})'\times H^{1/2+\varepsilon}} \qquad \forall\, z\in H^{3/2+\varepsilon}(\Omega)\cap H^1_0(\Omega). \tag{A.1}$$

By the first step, the solution operator

$$T^D_{vw}:((H^{1/2+\varepsilon}(\Omega))^d)'\to H^{1/2-\varepsilon}(\Omega)$$

given by $g\mapsto u$ is well defined and a bounded linear operator. We next claim that $T^D_{vw}$ also has the mapping property

$$(L^2(\Omega))^d\to H^1_0(\Omega).$$

In fact, we will show the stronger statement

$$T^D_{vw}=T^D \quad \text{on } (L^2(\Omega))^d.$$

To see this, let $g\in(L^2(\Omega))^d$. In order to see $y:=T^D_{vw}g\in H^1(\Omega)$, let $\varphi\in(C^\infty_0(\Omega))^d$ and define $z_\varphi:=\tilde{T}^D(\nabla\cdot\varphi)\in H^{3/2+\varepsilon}(\Omega)\cap H^1_0(\Omega)$. We note the classical estimate $\|z_\varphi\|_{1,\Omega}\leqslant C\|\varphi\|_{0;\Omega}$. We also observe that we have pointwise $-\Delta z_\varphi=\nabla\cdot\varphi$. Hence, an integration by parts together with the definition of the weak gradient $\nabla y$ yields

$$\langle\nabla y,\varphi\rangle\overset{\text{def}}{=}-(y,\nabla\cdot\varphi)_{0,\Omega}=(y,\Delta z_\varphi)_{0,\Omega}=-B(y,z_\varphi)\overset{(\text{A.1})}{=}\langle g,\nabla z_\varphi\rangle_{(H^{1/2+\varepsilon})'\times H^{1/2+\varepsilon}}=(g,\nabla z_\varphi)_{0,\Omega}.$$

For the right-hand side, we have $|(g,\nabla z_\varphi)_{0,\Omega}|\leqslant\|g\|_{0;\Omega}\|\nabla z_\varphi\|_{0;\Omega}\leqslant\|g\|_{0;\Omega}\|\varphi\|_{0;\Omega}$. Hence, $\nabla y\in L^2(\Omega)$.

As a next step towards showing that $y=T^D_{wv}g=T^Dg=:\tilde{y}$, we show that $y-\tilde{y}$ is harmonic. To that end, let $\varphi\in C^\infty_0(\Omega)$. Then

$$\langle-\Delta(y-\tilde{y}),\varphi\rangle\overset{\text{def}}{=}(y-\tilde{y},-\Delta\varphi)_{0,\Omega}=B(y,\varphi)-(\nabla\tilde{y},\nabla\varphi)_{0,\Omega}=(g,\nabla\varphi)_{0,\Omega}-(g,\nabla\varphi)_{0,\Omega}=0.$$

Next, we show $y\in H^1_0(\Omega)$. In order to establish this, we note that, since $-\Delta(y-\tilde{y})=0$, the co-normal derivative $\partial_n(y-\tilde{y})$ is a well-defined element of $H^{-1/2}(\partial\Omega)$. For all $\varphi\in C^\infty_0(\bar{\Omega})$ we have

$$\begin{aligned}0&=\langle\partial_n(y-\tilde{y}),\varphi\rangle=(\nabla(y-\tilde{y}),\nabla\varphi)_{0,\Omega}=(y,\partial_n\varphi)_{0,\partial\Omega}-(y,\Delta\varphi)_{0,\Omega}-(\nabla\tilde{y},\nabla\varphi)_{0,\Omega}\\&=(y,\partial_n\varphi)_{0,\partial\Omega}+B(y,\varphi)-(g,\nabla\varphi)_{0,\Omega}=(y,\partial_n\varphi)_{0,\partial\Omega}.\end{aligned}$$

By varying $\varphi\in C^\infty_0(\Omega)$, we conclude that $y=0$ on $\partial\Omega$.

We have thus shown that the very weak solution $y=T^D_{vw}g\in H^1_0(\Omega)$ if $g\in L^2(\Omega)$. An integration by parts then shows that $y$ solves the weak formulation, and the uniqueness of the weak solution thus provides $y=\tilde{y}$. This shows that $T^D_{vw}$ is the unique extension of $T^D$ to $((H^{+1/2+\varepsilon}(\Omega))^d)'$.

*Step* 3. The above steps have shown that $T^D_{vw}$ has the following mapping properties:

$$T^D_{vw}:((H^{1/2+\varepsilon}(\Omega))^d)'\to H^{1/2-\varepsilon}(\Omega), \qquad T^D_{vw}:(L^2(\Omega))^d\to H^1_0(\Omega).$$

By a standard interpolation argument, $T^D_{vw}$ is a bounded linear operator,

$$T_{vw}^D : ((L^2(\Omega))^d, ((H^{1/2+\varepsilon}(\Omega))^d)')_{\theta,\infty} \to (H_0^1(\Omega), H^{1/2-\varepsilon}(\Omega))_{\theta,\infty} \subset (H^1(\Omega), H^{1/2-\varepsilon}(\Omega))_{\theta,\infty}$$

for every $\theta \in (0,1)$. Select $\theta \in (0,1)$ such that $1/2 = \theta(1/2+\varepsilon)$. Then $(H^1, H^{1/2-\varepsilon})_{\theta,\infty} = B_{2,\infty}^{1/2}(\Omega)$. Furthermore, by Triebel (1995, Theorem 1.11.2) or Tartar (2007, Lemma 41.3),

$$((B_{2,1}^{1/2}(\Omega))^d)' = (((L^2(\Omega))^d, (H^{1/2+\varepsilon}(\Omega))^d)_{\theta,1})' = ((L^2(\Omega))^d, ((H^{1/2+\varepsilon}(\Omega))^d)')_{\theta,\infty}.$$

We conclude that $T_{vw}^D$ is a bounded linear operator from $((B_{2,1}^{1/2}(\Omega))^d)'$ to $B_{2,\infty}^{1/2}(\Omega)$. As we have already ascertained that $T_{vw}^D = T^D$ on $(L^2(\Omega))^d$, the proof is complete. □

LEMMA A3 Let $\Omega$ satisfy Assumption A1. Then the operator $T_w^M$ of Lemma 3.1 is a bounded linear operator,

$$T_w^M : ((B_{2,1}^{1/2}(\Omega))^d)' \to B_{2,\infty}^{1/2}(\Omega).$$

*Proof.* The proof follows by observing that, on $(L^2(\Omega))^d$, the operator $T_w^M$ coincides with the weak solution operator $T^D$ of Lemma A2, which has the stated mapping property. □