

Supercloseness of Primal-Dual Galerkin Approximations for Second Order Elliptic Problems

Bernardo Cockburn¹ · Manuel A. Sánchez^{1,2} ·
Chunguang Xiong³

Received: 13 March 2017 / Revised: 24 July 2017 / Accepted: 7 August 2017 /
Published online: 18 August 2017
© Springer Science+Business Media, LLC 2017

Abstract We show that two widely used Galerkin formulations for second-order elliptic problems provide approximations which are actually superclose, that is, their difference converges faster than the corresponding errors. In the framework of linear elasticity, the two formulations correspond to using either the stiffness tensor or its inverse the compliance tensor. We find sufficient conditions, for a wide class of methods (including mixed and discontinuous Galerkin methods), which guarantee a supercloseness result. For example, for the HDG_k method using polynomial approximations of degree $k > 0$, we find that the difference of approximate fluxes superconverges with order $k + 2$ and that the difference of the scalar approximations superconverges with order $k + 3$. We provide numerical results verifying our theoretical results.

Keywords Hybridizable discontinuous Galerkin · Superclose · Mixed methods · Tensor coefficient

Mathematics Subject Classification 65F10 · 65N30

✉ Manuel A. Sánchez
sanchez@umn.edu

Bernardo Cockburn
cockburn@math.umn.edu

Chunguang Xiong
xiongchg@bit.edu.cn

¹ School of Mathematics, 127 Vincent Hall, University of Minnesota, Minneapolis, MN 55455, USA

² Institute for Mathematical and Computational Engineering, School of Engineering and Faculty of Mathematics, Pontificia Universidad Católica de Chile, Santiago, Chile

³ Department of Mathematics, Beijing Institute of Technology, Beijing 100081, China

1 Introduction

We show that approximate solutions given by two Galerkin formulations for the model second-order elliptic problem

$$-\nabla \cdot (a \nabla u) = f \quad \text{in } \Omega, \tag{1a}$$

$$u = u_D \quad \text{on } \partial\Omega_D, \tag{1b}$$

$$-a \nabla u \cdot \mathbf{n} = \mathbf{q}_N \quad \text{on } \partial\Omega_N, \tag{1c}$$

where $a = a(x)$ is a bounded, symmetric and uniformly positive-definite $d \times d$ matrix-valued function in Ω , with inverse $c(x)$, $f \in L^2(\Omega)$, $u_D \in H^{1/2}(\partial\Omega_D)$ and $\mathbf{q}_N \in H^{1/2}(\partial\Omega_N)$, can be *superclose*, that is, the difference converges *faster* than the corresponding errors.

These Galerkin formulations differ *only* in the use of the tensor a or its inverse c in their formulations. Indeed, the formulations are based on the following equivalent rewritings of our model problem:

$$\left\{ \begin{array}{l} \mathbf{g} + \nabla u = \mathbf{0} \quad \text{in } \Omega, \\ c\mathbf{q} - \mathbf{g} = \mathbf{0} \quad \text{in } \Omega, \\ \nabla \cdot \mathbf{q} = f \quad \text{in } \Omega, \\ u = u_D \quad \text{on } \partial\Omega_D, \\ \mathbf{q} \cdot \mathbf{n} = \mathbf{q}_N \quad \text{on } \partial\Omega_N, \end{array} \right. \tag{A_1}$$

$$\left\{ \begin{array}{l} \mathbf{g} + \nabla u = \mathbf{0} \quad \text{in } \Omega, \\ \mathbf{q} - a\mathbf{g} = \mathbf{0} \quad \text{in } \Omega, \\ \nabla \cdot \mathbf{q} = f \quad \text{in } \Omega, \\ u = u_D \quad \text{on } \partial\Omega_D, \\ \mathbf{q} \cdot \mathbf{n} = \mathbf{q}_N \quad \text{on } \partial\Omega_N. \end{array} \right. \tag{A_2}$$

Thus, they differ only in the way the second equation is discretized. We observe that, when the tensors a and c are constant on the mesh, there is no difference between the corresponding Galerkin approximations. In the general case, we find sufficient conditions on the definition of the Galerkin methods which guarantee that their approximations are superclose.

The first formulation seems to be natural when mixed-like methods are defined. In contrast, in most cases, the tensor a is the natural data of the model which might be difficult or computationally expensive to actually invert. This is not only true for the simple model problem we are considering here but for more involved elliptic problems like the equations of linear elasticity where a corresponds to the *stiffness* tensor and c to the so-called *compliance* tensor. For the second-order elliptic problem under consideration, Galerkin formulations using the first set of equations have been used to define mixed methods [4] and have also been used for the original introduction of the hybridizable discontinuous Galerkin (HDG) methods [11]; see also the HDG methods for linear elasticity in [15]. On the other hand, there are many methods whose formulation is based on the second set of equations. For instance, in [1] the authors develop a so-called *expandedmixed* finite element method and give a finite difference interpretation. The HDG methods based on this set of equations are fully discussed in [7], where it is noted that they come directly from the the HDG methods for linear elasticity proposed in [21]. We also highlight that, as pointed out above, in the nonlinear case or when the tensor a can not be easily inverted, the second formulation becomes more relevant, see for example the HDG methods for the p -Laplacian equations [14] and the HDG methods for the equations of nonlinear elasticity [20]. Let us end by pointing out that the so-called Hybrid High-Order method, for diffusion [17, 18] and for elasticity, [16], uses a primal formulation which does use the tensor a . And yet, it can be re-interpreted as an HDG method using the first formulation, see [8]. Roughly speaking, this happens because the space of fluxes depends on the tensor a in a suitable manner.

In this paper, we want to address the question of how different are the numerical approximations based on the forms (A₁) and (A₂). To do that, we first prove estimates for the

difference of the approximations to the negative gradient variable \mathbf{g} and the flux variable \mathbf{q} by a classical energy argument. Then we prove an estimate for the difference of the approximations to the scalar variable u by using a standard duality argument and by using a *non-standard* approximation result. In particular, for the HDG method using polynomial approximations of degree $k > 0$ for all the variables, we obtain that the difference of the approximations of the negative gradient and the flux variable converge with an order of $k + 2$, and the scalar variable with an order of $k + 3$, where k is the polynomial degree associated to the local finite element space. This is, in general, 1 and 2 degrees extra than the convergence of each numerical approximation. A practical consequence of these results is that using one or the other formulation is essentially immaterial.

The remainder of the paper is structured as follows. In Sect. 2, we introduce, the general properties satisfied by the finite element approximations based on the Eqs. (A₁) and (A₂). We then state and discuss our supercloseness result, Theorem 1; we prove it in Sect. 3. In Sect. 4, we present numerical experiments validating our theoretical findings. We end in Sect. 5 with some concluding remarks.

2 The Finite Element Approximations and Their Supercloseness Properties

In this section, we state and discuss our main results. We begin by introducing the Galerkin methods we consider. We then state our main results, that is, the supercloseness properties between their approximations.

2.1 The Numerical Schemes

Notation In order to define the discrete primal-dual formulations we first introduce some notation. Let $\mathcal{T}_h = \{K\}$ be a conforming partition of $\bar{\Omega}$ into elements K , and let $\mathcal{F}_h = \{F \in \partial K : K \in \mathcal{T}_h\}$ be the set of all the faces ($d = 3$) or edges ($d = 2$) F of the partition. We assume that \mathcal{T}_h satisfies standard finite element assumptions, see [5] and [6]. The numerical methods that we will introduce next seek for a finite element approximation to the vector fields \mathbf{g} and \mathbf{q} , and the scalar field u . These numerical approximations will be defined on the following discontinuous piecewise polynomial spaces:

$$\begin{aligned} \mathbf{V}_h &= \{ \mathbf{v} \in [L^2(\Omega)]^d : \mathbf{v}|_K \in \mathbf{V}(K), \forall K \in \mathcal{T}_h \}, \\ W_h &= \{ w \in L^2(\Omega) : w|_K \in W(K), \forall K \in \mathcal{T}_h \}, \end{aligned}$$

where $\mathbf{V}(K)$ and $W(K)$ are local spaces, each one contained in a polynomial space.

In addition, we are using the following standard notation:

$$\begin{aligned} (\boldsymbol{\sigma}, \mathbf{v})_{\mathcal{T}_h} &= \sum_{K \in \mathcal{T}_h} \int_K \boldsymbol{\sigma}(x) \cdot \mathbf{v}(x) dx, \quad (\zeta, w)_{\mathcal{T}_h} = \sum_{K \in \mathcal{T}_h} \int_K \zeta(x) w(x) dx, \\ \langle \zeta, w \rangle_{\partial \mathcal{T}_h} &= \sum_{K \in \mathcal{T}_h} \int_{\partial K} \zeta(s) w(s) ds. \end{aligned}$$

The Formulations Next, we define the different formulations satisfied by the finite element approximations we are considering. Note that although the methods are assumed to *satisfy* the equations of each formulation, they are not necessarily *defined* by them.

The first formulation is based on the Eq. (A₁): The approximation $(\mathbf{g}_h^1, \mathbf{q}_h^1, u_h^1) \in \mathbf{V}_h \times \mathbf{V}_h \times W_h$ satisfies

$$(\mathbf{g}_h^1, \mathbf{v})_{\mathcal{T}_h} - (u_h^1, \nabla \cdot \mathbf{v})_{\mathcal{T}_h} + \langle \widehat{u}_h^1, \mathbf{v} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} = 0, \quad \forall \mathbf{v} \in \mathbf{V}_h, \tag{2a}$$

$$(c\mathbf{q}_h^1 - \mathbf{g}_h^1, \mathbf{v})_{\mathcal{T}_h} = 0, \quad \forall \mathbf{v} \in \mathbf{V}_h, \tag{2b}$$

$$-(\mathbf{q}_h^1, \nabla w)_{\mathcal{T}_h} + \langle \widehat{\mathbf{q}}_h^1 \cdot \mathbf{n}, w \rangle_{\partial \mathcal{T}_h} = (f, w)_{\mathcal{T}_h}, \quad \forall w \in W_h. \tag{2c}$$

The second formulation is based on the equations (A₂): The approximation $(\mathbf{g}_h^2, \mathbf{q}_h^2, u_h^2) \in \mathbf{V}_h \times \mathbf{V}_h \times W_h$ satisfies

$$(\mathbf{g}_h^2, \mathbf{v})_{\mathcal{T}_h} - (u_h^2, \nabla \cdot \mathbf{v})_{\mathcal{T}_h} + \langle \widehat{u}_h^2, \mathbf{v} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} = 0, \quad \forall \mathbf{v} \in \mathbf{V}_h, \tag{3a}$$

$$(\mathbf{q}_h^2 - a\mathbf{g}_h^2, \mathbf{v})_{\mathcal{T}_h} = 0, \quad \forall \mathbf{v} \in \mathbf{V}_h, \tag{3b}$$

$$-(\mathbf{q}_h^2, \nabla w)_{\mathcal{T}_h} + \langle \widehat{\mathbf{q}}_h^2 \cdot \mathbf{n}, w \rangle_{\partial \mathcal{T}_h} = (f, w)_{\mathcal{T}_h}, \quad \forall w \in W_h. \tag{3c}$$

Note that to complete the definition of the numerical methods, additional information about the local spaces and the *numerical traces* \widehat{u}_h^i and $\widehat{\mathbf{q}}_h^i \cdot \mathbf{n}$ is required, for $i = 1, 2$. To obtain our results, only very general conditions need to be imposed which we gather in the assumptions we display next. We first introduce an auxiliary dual problem and its Galerkin approximation by form (A₁) needed for the estimates by the scalar approximations.

The Auxiliary Dual Problem To prove the estimates for the scalar variable of Theorem 1, we use a duality argument. So, we need to introduce the following dual problem

$$c\boldsymbol{\psi} + \nabla\varphi = \mathbf{0} \quad \text{in } \Omega, \tag{4a}$$

$$\nabla \cdot \boldsymbol{\psi} = \theta \quad \text{in } \Omega, \tag{4b}$$

$$\varphi = 0 \quad \text{on } \partial\Omega_D, \tag{4c}$$

$$\boldsymbol{\psi} \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega_N. \tag{4d}$$

and the approximation $(\boldsymbol{\psi}_h, \varphi_h) \in \mathbf{V}_h \times W_h$ of (4) satisfying Eq. (2), that is,

$$(c\boldsymbol{\psi}_h, \mathbf{v})_{\mathcal{T}_h} - (\varphi_h, \nabla \cdot \mathbf{v})_{\mathcal{T}_h} + \langle \widehat{\varphi}_h, \mathbf{v} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} = 0 \quad \forall \mathbf{v} \in \mathbf{V}_h, \tag{5a}$$

$$-(\boldsymbol{\psi}_h, \nabla w)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{\psi}}_h \cdot \mathbf{n}, w \rangle_{\partial \mathcal{T}_h} = (\theta, w)_{\mathcal{T}_h} \quad \forall w \in W_h. \tag{5b}$$

Assumption We make the following assumptions on

(A) the local space $\mathbf{V}(K)$, $K \in \mathcal{T}_h$:

(i) $\bar{a}|_K(\mathbf{g}_h^1 - \mathbf{g}_h^2)|_K \in \mathbf{V}(K)$, where $\bar{a}|_K$ is the average of tensor a on K .

(ii) $\bar{c}|_K(\mathbf{q}_h^1 - \mathbf{q}_h^2)|_K \in \mathbf{V}(K)$, where $\bar{c}|_K$ is the average of tensor c on K .

(iii) $\bar{a}|_K \mathbf{v}|_K, \bar{c}|_K \mathbf{v}|_K \in \mathbf{V}(K)$ for all $\mathbf{v} \in \mathbf{V}(K)$.

(B) the numerical traces

(i) Single-valuedness: \widehat{u}_h and $\widehat{\mathbf{q}}_h \cdot \mathbf{n}$ are single valued on \mathcal{F}_h .

(ii) Non-negativity: $\langle \widehat{u}_h - u_h, (\mathbf{q}_h - \widehat{\mathbf{q}}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \geq 0$.

(iii) Cancellation: $\langle u_h - \widehat{u}_h, (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} + \langle \varphi_h - \widehat{\varphi}_h, (\mathbf{q}_h - \widehat{\mathbf{q}}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} = 0$.

(C) the approximation properties the flux:

(i) $\mathbf{V}(K) \supset [\mathcal{P}^0(K)]^d$,

Table 1 Local spaces of some mixed and DG methods

Method	$V(K)$	$W(K)$	$M(F)$
RT_k <i>simplexes</i>	$[\mathcal{P}^k(K)]^d \oplus x\mathcal{P}^k(K)$	$\mathcal{P}^k(K)$	$\mathcal{P}^k(F)$
RT_k <i>squares</i>	$\mathcal{P}^{k+1,k}(K) \times \mathcal{P}^{k,k+1}(K)$	$\mathcal{P}^{k,k}(K)$	$\mathcal{P}^k(F)$
HDG_k <i>polyhedra</i>	$[\mathcal{P}^k(K)]^d$	$\mathcal{P}^k(K)$	$\mathcal{P}^k(F)$
DG_k <i>polyhedra</i>	$[\mathcal{P}^k(K)]^d$	$\mathcal{P}^k(K)$	–
BDM_k <i>simplexes</i>	$[\mathcal{P}^k(K)]^d$	$\mathcal{P}^{k-1}(K)$	$\mathcal{P}^k(F)$
BDM_k <i>squares</i>	$[\mathcal{P}^k(K)]^2 \oplus \text{span}\{\nabla^\perp x y^{k+1}, \nabla^\perp x^{k+1} y\}$	$\mathcal{P}^{k-1}(K)$	$\mathcal{P}^k(F)$

(ii) $\|q - q_h^1\|_{L^2(\mathcal{T}_h)} \leq C_a h^\alpha \|u\|_{H^2(\mathcal{T}_h)}$ for some $\alpha \in (0, 1]$.

(D) the elliptic regularity of the dual problem:

$$\|\varphi\|_{H^2(\Omega)} + \|\psi\|_{H^1(\Omega;c)} \leq C \|\theta\|_{L^2(\Omega)},$$

where the norm $\|\cdot\|_{H^1(\Omega;c)}$ is the H^1 -norm weighted with $c^{1/2}$. See Sect. 2.2. This inequality is satisfied, for example, if the domain Ω is convex and either $\partial\Omega_D$ or $\partial\Omega_N$ vanishes.

Example The main examples of methods satisfying the above weak formulations are the hybridized version of the Raviart–Thomas (RT) [19] and Brezzi–Douglas–Marini (BDM) [3] mixed methods, the Discontinuous Galerkin (DG) methods and the so-called hybridizable Discontinuous Galerkin (HDG) [9] methods. In Table 1, we display the choices of the local spaces $V(K)$, $W(K)$. For the hybridized version of the mixed methods and for the HDG methods, \widehat{u}_h is an additional unknown. This is why we also describe the space $M(F)$ to which the restriction of \widehat{u}_h to the face F belongs. Next, we briefly discuss the satisfaction of the assumptions (A), (B) and (C) by these methods.

2.1.1 The Local Vector Spaces: Assumption (A)

For the DG and HDG_k methods, we see that assumption (Aiii), and hence assumptions (Ai) and (Aii), are satisfied. Assumption (Aiii) is satisfied by the BDM_k method for simplexes, but not for squares (or cubes). The RT_k method does not satisfy condition (Aiii) neither; it does not satisfy condition (Aii) though. See Table 2.

2.1.2 The Numerical Traces: Assumption (B)

For the DG methods, the numerical traces \widehat{u}_h and \widehat{q}_h are explicitly defined in terms of the original unknowns of the problem, u_h and q_h . To describe them, let us recall the standard DG notation for the averages and jumps on the interior faces

$$\begin{aligned} \{\{u\}\} &:= \frac{1}{2}(u^+ + u^-), & \{\{q\}\} &:= \frac{1}{2}(q^+ + q^-), \\ \llbracket u \rrbracket &:= u^+ n^+ + u^- n^-, & \llbracket q \rrbracket &:= q^+ \cdot n^+ + q^- \cdot n^-. \end{aligned}$$

The numerical traces are then defined on each $F \in \mathcal{F}_h \setminus \partial\Omega$ by:

$$\widehat{\mathbf{q}}_h^i := \{\{\mathbf{q}_h^i\}\} + C_{11} \llbracket u_h^i \rrbracket - C_{12} \llbracket \mathbf{q}_h^i \rrbracket, \quad \widehat{u}_h^i := \{\{u_h^i\}\} + C_{12} \cdot \llbracket u_h^i \rrbracket + C_{22} \llbracket \mathbf{q}_h^i \rrbracket,$$

for $i = 1, 2$, where the auxiliary parameters C_{11} , C_{12} and C_{22} might depend on x . On the boundary faces we imposed the boundary conditions of the problems by

$$\widehat{\mathbf{q}}_h^i \cdot \mathbf{n} := \begin{cases} \mathbf{q}_h^i \cdot \mathbf{n} + C_{11}(u_h^i - u_D) & \text{on } \partial\Omega_D, \\ \mathbf{q}_N & \text{on } \partial\Omega_N, \end{cases}$$

and

$$\widehat{u}_h^i := \begin{cases} u_D & \text{on } \partial\Omega_D, \\ u_h^i + C_{22}(\mathbf{q}_h^i \cdot \mathbf{n} - \mathbf{q}_N) & \text{on } \partial\Omega_N, \end{cases}$$

for $i = 1, 2$. In this way, the DG methods always satisfy assumptions (Bi) and (Biii). They satisfy (Bii) when C_{11} and C_{22} are nonnegative.

For the particular case

$$C_{11} = \tau^+ \tau^- / (\tau^{++} \tau^-), \quad C_{22} = 1 / (\tau^+ + \tau^-), \quad C_{12} = \frac{1}{2} \llbracket \tau \rrbracket / (\tau^+ + \tau^-),$$

we have

$$\begin{aligned} \widehat{u}_h &= \frac{\tau^+ u_h^+ + \tau^- u_h^-}{\tau^+ + \tau^-} + \frac{1}{\tau^+ + \tau^-} \llbracket \mathbf{q}_h \rrbracket, \\ \widehat{\mathbf{q}}_h &= \frac{\tau^- \mathbf{q}_h^+ + \tau^+ \mathbf{q}_h^-}{\tau^+ + \tau^-} + \frac{\tau^+ \tau^-}{\tau^{++} \tau^-} \llbracket u_h \rrbracket, \end{aligned}$$

on all interior faces, and it turns out that we can write

$$\widehat{\mathbf{q}}_h \cdot \mathbf{n}^\pm = \mathbf{q}_h^\pm \cdot \mathbf{n}^\pm + \tau^\pm (u_h^\pm - \widehat{u}_h).$$

We thus obtain an HDG method, see [7, 11]. For general HDG methods and the hybridized version of the mixed methods, the scalar numerical trace is a *single-valued*, new unknown on $\mathcal{F}_h \setminus \partial\Omega_D$; on $\widehat{u}_h = u_D$ on $\partial\Omega_D$ though. The numerical trace of the flux is defined as a linear combination of the other unknowns with a stabilization function τ ; for the mixed methods, $\tau = 0$. Specifically, we have

$$\widehat{\mathbf{q}}_h^i = \mathbf{q}_h^i + \tau(u_h^i - \widehat{u}_h^i) \mathbf{n} \quad \text{on } \mathcal{F}_h,$$

for $i = 1, 2$. To ensure the satisfaction of assumption (Bi), the methods impose the condition

$$\langle \widehat{\mathbf{q}}_h \cdot \mathbf{n}, m \rangle_{\partial\mathcal{T}_h} = \langle \mathbf{q}_N, m \rangle_{\partial\mathcal{T}_h}, \quad \forall m \in M_h. \tag{6}$$

where $M_h := \{m : m|_F \in M(F), \forall F \in \mathcal{F}_h\}$. Finally, when the stabilization function $\tau(\cdot)$ is just a multiplication operator, that is, $\tau(\mu) := \tau \cdot \mu$, the assumption (Biii) is satisfied and the assumption (Bii) are satisfied if $\tau \geq 0$. See Table 2.

2.1.3 The Approximation of the Flux: Assumption (C)

In Table 3, we display several cases in which assumptions (C) are satisfied.

We only consider the cases for which condition (Aii) is also satisfied.

Table 2 Assumptions (A) and (B)

Method <i>elements</i>	(Ai)	(Aii)	(Aiii)	(Bi)	(Bii)	(Biii)
RT_k <i>simplexes</i>	×	✓	×	✓	✓	✓
RT_k <i>squares,cubes</i>	×	×	×	✓	✓	✓
HDG_k <i>polyhedra</i>	✓	✓	✓	✓	✓ $\tau \geq 0$	✓
DG_k <i>polyhedra</i>	✓	✓	✓	✓	$C_{11}, C_{22} \geq 0$	✓
BDM_k <i>simplexes</i>	✓	✓	✓	✓	✓	✓
BDM_k <i>squares,cubes</i>	×	×	×	✓	✓	✓

A ✓ means that the assumption is satisfied by the numerical method. A × means that it is not

Table 3 Assumptions (C) for vector spaces satisfying (Aii)

Method <i>elements</i>	α	Condition	References
RT_k <i>simplexes</i>	1	$k \geq 0$	[4]
HDG_k <i>simplices</i>	1	$\tau _{\partial K} = 0$ on all faces of K but one	[10]
HDG_k <i>simplices</i>	1	$\tau^{\pm 1}$ of order one, $k \geq 0$	[12]
HDG_k <i>simplices</i>	1	$(h\tau)^{\pm 1}$ of order one, $k \geq 1$	[12]
HDG_k <i>polyhedra</i>	1	$(h\tau)^{\pm 1}$ of order one, $k \geq 1$	[5]
HDG_k <i>polyhedra</i>	$\frac{1}{2}$	τ is of order one	[5]
DG_k <i>simplices</i>	1	$C_{11}^{\pm 1}, C_{22}^{\pm 1}$ and C_{12} of order one	[13]
DG_k <i>polyhedra</i>	1	$(hC_{11})^{\pm 1}, (C_{22}/h)^{\pm 1}, C_{12}$ of order one, $k \geq 1$	[5]
DG_k <i>polyhedra</i>	$\frac{1}{2}$	$C_{11}^{\pm 1}, C_{22}^{\pm 1}, C_{12}$ of order one	[5]
BDM_k <i>simplexes</i>	1	$k \geq 1$	[4]

2.2 Supercloseness of the Approximations

To state our results, we use the following notation.

We define the Sobolev space $X(\mathcal{T}_h) = \prod_{K \in \mathcal{T}_h} X(K)$, for any Sobolev space X , and its norm

$$\|\mu\|_{X(\mathcal{T}_h)}^2 := \sum_{K \in \mathcal{T}_h} \|\mu\|_{X(K)}^2, \quad \forall \mu \in X(\mathcal{T}_h).$$

Finally, for $\mathbf{v} \in [L^2(\mathcal{T}_h)]^d$ we define the norm weighted with a tensor \mathbf{c} by

$$\|\mathbf{v}\|_{L^2(\mathcal{T}_h; \mathbf{c})}^2 := (\mathbf{c}\mathbf{v}, \mathbf{v})_{\mathcal{T}_h}.$$

We let u be the solution of problem (1) and set $\mathbf{g} := -\nabla u$ and $\mathbf{q} := \mathbf{a}\mathbf{g}$. We also let $(\mathbf{g}_h^1, \mathbf{q}_h^1, u_h^1)$ and $(\mathbf{g}_h^2, \mathbf{q}_h^2, u_h^2)$ be numerical approximations satisfying (2) and (3), respec-

tively. for some estimates, we are going to use the following elliptic regularity inequality: Now we are ready to state the supercloseness properties of the approximations satisfying (2) and (3). Its proof is provided in the next section.

Theorem 1 *Suppose that the assumption (B) on the numerical traces hold. Then*

$$\begin{aligned} \|\mathbf{g}_h^1 - \mathbf{g}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{a})} &\leq \|c^{\frac{1}{2}}(a - \bar{a})c^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)} (\|\mathbf{q} - \mathbf{q}_h^1\|_{L^2(\mathcal{T}_h; \mathbf{c})} \\ &\quad + \|\mathbf{g} - \mathbf{g}_h^1\|_{L^2(\mathcal{T}_h; \mathbf{a})}) \quad \text{if (Ai) holds,} \\ \|\mathbf{q}_h^1 - \mathbf{q}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{c})} &\leq \|a^{\frac{1}{2}}(c - \bar{c})a^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)} (\|\mathbf{q} - \mathbf{q}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{c})} \\ &\quad + \|\mathbf{g} - \mathbf{g}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{a})}) \quad \text{if (Aii) holds,} \end{aligned}$$

Suppose now that assumptions (B) and (C) on the approximation of the flux of the first method hold. Then, if the elliptic regularity inequality of assumption (D) holds,

$$\|u_h^1 - u_h^2\|_{L^2(\mathcal{T}_h)} \leq C_1 h^\alpha (\|\mathbf{q} - \mathbf{q}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{c})} + \|\mathbf{g} - \mathbf{g}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{a})}).$$

Moreover, if assumption (Aiii) holds and if $[\mathcal{P}^1(K)]^d \subseteq \mathbf{V}(K)$,

$$\|u_h^1 - u_h^2\|_{L^2(\mathcal{T}_h)} \leq C_2 h^{1+\alpha} (\|\mathbf{q} - \mathbf{q}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{c})} + \|\mathbf{g} - \mathbf{g}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{a})}).$$

The constants C_1 and C_2 are independent of h and the solution. The constant C_1 depends on $\|c\|_{W^{1,\infty}(\mathcal{T}_h)}$ whereas the constant C_2 depends on $\|c\|_{W^{2,\infty}(\mathcal{T}_h)}$.

Theorem 2 *Assume that, for some positive constant κ , we have*

$$\max \{ \|c^{1/2}(a - \bar{a})c^{1/2}\|_{L^\infty(K)}, \|a^{1/2}(c - \bar{c})a^{1/2}\|_{L^\infty(K)} \} \leq \kappa < 1 \quad \forall K \in \mathcal{T}_h,$$

and set

$$\Upsilon_h := \min_{i=1,2} \{ \|\mathbf{q} - \mathbf{q}_h^i\|_{L^2(\mathcal{T}_h; \mathbf{c})} + \|\mathbf{g} - \mathbf{g}_h^i\|_{L^2(\mathcal{T}_h; \mathbf{a})} \}.$$

Then, the estimates of Theorem 1 become

$$\begin{aligned} \|\mathbf{g}_h^1 - \mathbf{g}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{a})} &\leq \|c^{\frac{1}{2}}(a - \bar{a})c^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)} \Upsilon_h \quad \text{if (B) and (Ai) hold,} \\ \|\mathbf{q}_h^1 - \mathbf{q}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{c})} &\leq \|a^{\frac{1}{2}}(c - \bar{c})a^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)} \Upsilon_h \quad \text{if (B) and (Aii) hold,} \\ \|u_h^1 - u_h^2\|_{L^2(\mathcal{T}_h)} &\leq C_1 h^\alpha \Upsilon_h \quad \text{if (B), (C) and (D) hold,} \\ \|u_h^1 - u_h^2\|_{L^2(\mathcal{T}_h)} &\leq C_2 h^{1+\alpha} \Upsilon_h \quad \text{if (B), (C), (D) hold and } [\mathcal{P}^1(K)]^d \subseteq \mathbf{V}(K). \end{aligned}$$

We summarize the application of this result to the numerical methods described in our examples in Table 4. There, we assume that the extra assumption of Theorem 2 holds.

A few remarks are in order. First, note that, when using simplexes, it is known that the first formulation of the HDG_k method with τ of order one, converges with order $k + 1$ in all the approximations; see [12]. Theorem 1 states the the second formulation and the first one are *superclose* in the sense that the order of convergence of the difference of their approximations converge with order $k + 2$, for the vector-valued approximations, and with order $k + 3$, for the scalar approximation when $c \in W^{2,\infty}(\mathcal{T}_h)$.

Note also that, *regardless* of the actual shape of the elements, when the values of the stabilization function $\tau^{\pm 1}$, respectively, $(h\tau)^{\pm 1}$, are of order one, respectively, the order of convergence of the difference of their approximations converge with an additional half an order, respectively, a full order, for the vector-valued approximations, and with three half

Table 4 Orders of convergence in h

Method <i>elements</i>	$\frac{1}{Y_h} \ e_g\ _{L^2(\mathcal{T}_h;a)}$	$\frac{1}{Y_h} \ e_q\ _{L^2(\mathcal{T}_h;c)}$	$\frac{1}{Y_h} \ e_u\ _{L^2(\mathcal{T}_h)}$	$\frac{1}{Y_h} \ e_u\ _{L^2(\mathcal{T}_h)}$
RT _k <i>simplexes</i>	0	1	1	1
HDG _k <i>polyhedra</i>	1	1	α	$1 + \alpha$
BDM _k <i>simplexes</i>	1	1	1	2
DG _k <i>polyhedra</i>	1	1	α	$1 + \alpha$

The value of α is provided in the previous table. Here $e_g := g_h^1 - g_h^2$, $e_q := q_h^1 - q_h^2$ and $e_u := u_h^1 - u_h^2$. For the first three columns, we assume that $c \in W^{1,\infty}(\mathcal{T}_h)$. For the last, that $c \in W^{2,\infty}(\mathcal{T}_h)$. For the last two columns, we assume that the elliptic regularity inequality holds

orders, respectively two full orders, for the scalar approximation whenever $c \in W^{2,\infty}(\mathcal{T}_h)$. This also applies to the DG_k methods.

Concerning the mixed methods, similar results are obtained for the BDM_k method. Interestingly enough, although the convergence of the method is of order k for the scalar approximations, their difference is of order $k + 3$. On the other hand, since the assumption (Ai) is *not* satisfied for the RT_k method, we see that the approximate fluxes, but not the approximate gradients, are superclose. Moreover, the scalar approximations are superclose with an extra power in h , but not two, like to the DG methods or three for the BDM_k method.

The numerical results presented in Sect. 4 confirm that all the results in the above table are sharp.

To end this section, we note that, besides the supercloseness result, Theorem 1 implies that the approximation properties of one scheme can be deduced for those of the other. In particular, if either approximate solution converges, then the other approximate solution converges too, and converges with the same rate.

3 Proofs

Here, we provide detailed proofs of our main result, Theorem 1.

3.1 The Error Equations

We begin by obtaining the error equations. Let (g_h^1, q_h^1, u_h^1) and (g_h^2, q_h^2, u_h^2) be functions satisfying (2) and (3), respectively. Then, if we set

$$e_g := g_h^1 - g_h^2, e_q := q_h^1 - q_h^2, e_u := u_h^1 - u_h^2, e_{\hat{q}} := \hat{q}_h^1 - \hat{q}_h^2, e_{\hat{u}} := \hat{u}_h^1 - \hat{u}_h^2,$$

we have, subtracting Eq. (3) from Eq. (2), that

$$(e_g, v)_{\mathcal{T}_h} - (e_u, \nabla \cdot v)_{\mathcal{T}_h} + \langle e_{\hat{u}}, v \cdot n \rangle_{\partial \mathcal{T}_h} = 0 \quad \forall v \in V_h, \tag{7a}$$

$$(c q_h^1 - g_h^1, v)_{\mathcal{T}_h} = (q_h^2 - a g_h^2, v)_{\mathcal{T}_h} = 0 \quad \forall v \in V_h, \tag{7b}$$

$$-(e_q, \nabla w)_{\mathcal{T}_h} + \langle e_{\hat{q}}, n \cdot w \rangle_{\partial \mathcal{T}_h} = 0 \quad \forall w \in W_h. \tag{7c}$$

3.2 Proof of Estimates of the Difference for the Vector Unknowns

Here, we prove the estimates for the difference of the approximations to the gradient and to the flux of Theorem 1. To do that, we proceed by using a variation on the classic energy argument.

Step 1: The energy argument Taking $v := e_q \in V_h$ in the first error Eq. (7a), $w := e_u \in W_h$ in the third error Eq. (7c), and adding the resulting equations, we get

$$(e_g, e_q)_{T_h} - (e_u, \nabla \cdot e_q)_{T_h} + \langle e_{\hat{u}}, e_q \cdot n \rangle_{\partial T_h} - (e_q, \nabla e_u)_{T_h} + \langle e_{\hat{q}} \cdot n, e_u \rangle_{\partial T_h} = 0.$$

After integrating by parts, and after adding and subtracting the term $\langle e_{\hat{u}}, e_{\hat{q}} \cdot n \rangle_{\partial T_h}$, we get

$$(e_g, e_q)_{T_h} + \langle e_{\hat{u}} - e_u, (e_q - e_{\hat{q}}) \cdot n \rangle_{\partial T_h} = 0.$$

where we have used the fact that $\langle e_{\hat{u}}, e_{\hat{q}} \cdot n \rangle_{\partial T_h} = 0$ since the numerical traces are single-valued by (Bi), $e_{\hat{u}} = 0$ on $\partial\Omega_D$ and $e_{\hat{q}} \cdot n = 0$ on $\partial\Omega_N$. Then, by the positivity property (Bii), we obtain

$$(e_g, e_q)_{T_h} \leq 0.$$

Step 2: The estimate of the difference in the gradient By the second of the equations in (7b),

$$(e_g, e_q)_{T_h} = (e_g, q_h^1 - q_h^2)_{T_h} = (e_g, q_h^1 - a g_h^2)_{T_h} = \|e_g\|_{L^2(T_h;a)}^2 + (q_h^1 - a g_h^1, e_g)_{T_h},$$

and, by the last inequality of Step 1,

$$\begin{aligned} \|e_g\|_{L^2(T_h;a)}^2 &\leq - (q_h^1 - a g_h^1, e_g)_{T_h} \\ &= - (c q_h^1 - g_h^1, a e_g)_{T_h} \\ &= - (c q_h^1 - g_h^1, (a - \bar{a}) e_g)_{T_h}, \end{aligned}$$

since $(c q_h^1 - g_h^1, \bar{a} e_g)_{T_h} = 0$. This holds because, by assumption (Ai), $\bar{a} e_g \in V_h$, and so we can use the first of the equations in (7b) with $v := \bar{a} e_g$. Thus,

$$\begin{aligned} \|e_g\|_{L^2(T_h;a)}^2 &\leq - (c q_h^1 - c q, (a - \bar{a}) e_g)_{T_h} - (g - g_h^1, (a - \bar{a}) e_g)_{T_h} \\ &\leq (\|q - q_h^1\|_{L^2(T_h;c)} + \|g - g_h^1\|_{L^2(T_h;a)}) \|c^{\frac{1}{2}}(a - \bar{a})c^{\frac{1}{2}}\|_{L^\infty(T_h)} \|e_g\|_{L^2(T_h;a)}, \end{aligned}$$

and we get our estimate

$$\|e_g\|_{L^2(T_h;a)} \leq (\|q - q_h^1\|_{L^2(T_h;c)} + \|g - g_h^1\|_{L^2(T_h;a)}) \|c^{\frac{1}{2}}(a - \bar{a})c^{\frac{1}{2}}\|_{L^\infty(T_h)}.$$

Step 3: The estimate of the difference in the flux By the first of the Eq. (7b) with $v := e_q$,

$$\begin{aligned} (e_g, e_q)_{T_h} &= (g_h^1 - g_h^2, e_q)_{T_h} = (c q_h^1 - g_h^2, e_q)_{T_h} \\ &= \|e_q\|_{L^2(T_h;c)}^2 + (c q_h^2 - g_h^2, e_q)_{T_h}, \end{aligned}$$

and by the last inequality of Step 1, we get

$$\begin{aligned} \|e_q\|_{L^2(T_h;c)}^2 &\leq - (c q_h^2 - g_h^2, e_q)_{T_h} \\ &= - (q_h^2 - a g_h^2, c e_q)_{T_h} \\ &= - (q_h^2 - a g_h^2, (c - \bar{c}) e_q)_{T_h}, \end{aligned}$$

since $(\mathbf{q}_h^2 - \mathbf{a}\mathbf{g}_h^2, \bar{\mathbf{c}}\mathbf{e}_q)_{\mathcal{T}_h} = 0$. Indeed, by assumption (Aii), $\bar{\mathbf{c}}\mathbf{e}_q \in \mathbf{V}_h$, and we can take $\mathbf{v} := \bar{\mathbf{c}}\mathbf{e}_q$ in the second of the equations in (7b). Thus,

$$\begin{aligned} \|\mathbf{e}_q\|_{L^2(\mathcal{T}_h; \mathbf{c})}^2 &\leq -(\mathbf{q}_h^2 - \mathbf{q}, (\mathbf{c} - \bar{\mathbf{c}})\mathbf{e}_q)_{\mathcal{T}_h} - (\mathbf{a}(\mathbf{g} - \mathbf{g}_h^2), (\mathbf{c} - \bar{\mathbf{c}})\mathbf{e}_q)_{\mathcal{T}_h} \\ &\leq (\|\mathbf{q} - \mathbf{q}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{c})} + \|\mathbf{g} - \mathbf{g}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{a})}) \|\mathbf{a}^{\frac{1}{2}}(\mathbf{c} - \bar{\mathbf{c}})\mathbf{a}^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)} \|\mathbf{e}_q\|_{L^2(\mathcal{T}_h; \mathbf{c})}, \end{aligned}$$

and we obtain our first estimate

$$\|\mathbf{e}_q\|_{L^2(\mathcal{T}_h; \mathbf{c})} \leq (\|\mathbf{q} - \mathbf{q}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{c})} + \|\mathbf{g} - \mathbf{g}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{a})}) \|\mathbf{a}^{\frac{1}{2}}(\mathbf{c} - \bar{\mathbf{c}})\mathbf{a}^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)}.$$

This completes the proof of the estimates of the difference in the vector unknowns.

3.3 Proof of the Estimates of the Difference for the Scalar Variable

Step 1: An identity for the difference Taking $w := e_u \in W_h$ in the second equation of the approximation to the dual solution, (5b), we get

$$\begin{aligned} (e_u, \theta)_{\mathcal{T}_h} &= -(\nabla e_u, \boldsymbol{\psi}_h)_{\mathcal{T}_h} + \langle \widehat{\boldsymbol{\psi}}_h \cdot \mathbf{n}, e_u \rangle_{\partial \mathcal{T}_h} \\ &= (e_u, \nabla \cdot \boldsymbol{\psi}_h)_{\mathcal{T}_h} + \langle (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n}, e_u \rangle_{\partial \mathcal{T}_h} \\ &= (\mathbf{e}_g, \boldsymbol{\psi}_h)_{\mathcal{T}_h} + \langle e_{\widehat{u}}, \boldsymbol{\psi}_h \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} + \langle (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n}, e_u \rangle_{\partial \mathcal{T}_h} \end{aligned}$$

by the error Eq. (7a) with $\mathbf{v} := \boldsymbol{\psi}_h \in \mathbf{V}_h$. Since $\langle \widehat{\boldsymbol{\psi}}_h \cdot \mathbf{n}, e_{\widehat{u}} \rangle_{\partial \mathcal{T}_h} = 0$ because of the single-valuedness of the numerical traces, assumption (Bi) and the fact that $e_{\widehat{u}} = 0$ on $\partial \Omega_D$ and $\widehat{\boldsymbol{\psi}}_h \cdot \mathbf{n} = 0$ on $\partial \Omega_N$, we obtain

$$\begin{aligned} (e_u, \theta)_{\mathcal{T}_h} &= (\mathbf{e}_g, \boldsymbol{\psi}_h)_{\mathcal{T}_h} + \langle e_u - e_{\widehat{u}}, (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \\ &= (\mathbf{c}\mathbf{q}_h^1 - \mathbf{g}_h^2, \boldsymbol{\psi}_h)_{\mathcal{T}_h} + \langle e_u - e_{\widehat{u}}, (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h}, \end{aligned}$$

by the first of the Eq. (7b) with $\mathbf{v} := \boldsymbol{\psi}_h \in \mathbf{V}_h$. Finally,

$$\begin{aligned} (e_u, \theta)_{\mathcal{T}_h} &= (\mathbf{c}\mathbf{e}_q, \boldsymbol{\psi}_h)_{\mathcal{T}_h} + (\mathbf{c}\mathbf{q}_h^2 - \mathbf{g}_h^2, \boldsymbol{\psi}_h)_{\mathcal{T}_h} + \langle e_u - e_{\widehat{u}}, (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \\ &= (\mathbf{c}\mathbf{q}_h^2 - \mathbf{g}_h^2, \boldsymbol{\psi}_h)_{\mathcal{T}_h}, \end{aligned}$$

because the term $\Theta_h := (\mathbf{c}\mathbf{e}_q, \boldsymbol{\psi}_h)_{\mathcal{T}_h} + \langle e_u - e_{\widehat{u}}, (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h}$ is equal to zero.

Let us prove this claim. We have

$$\begin{aligned} \Theta_h &= (\mathbf{e}_q, \mathbf{c}\boldsymbol{\psi}_h)_{\mathcal{T}_h} + \langle e_u - e_{\widehat{u}}, (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \\ &= (\nabla \cdot \mathbf{e}_q, \varphi_h)_{\mathcal{T}_h} - \langle \widehat{\varphi}_h, \mathbf{e}_q \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} + \langle e_u - e_{\widehat{u}}, (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \end{aligned}$$

by the Eq. (5a) with $\mathbf{v} := \mathbf{e}_q$. Integrating by parts, we get

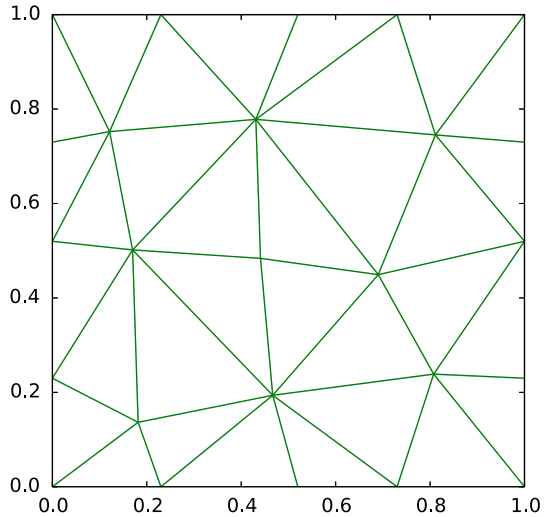
$$\begin{aligned} \Theta_h &= -(\mathbf{e}_q, \nabla \varphi_h)_{\mathcal{T}_h} + \langle \varphi_h - \widehat{\varphi}_h, \mathbf{e}_q \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} + \langle e_u - e_{\widehat{u}}, (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \\ &= -\langle \mathbf{e}\widehat{\mathbf{q}} \cdot \mathbf{n}, \varphi_h \rangle_{\partial \mathcal{T}_h} + \langle \varphi_h - \widehat{\varphi}_h, \mathbf{e}_q \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} + \langle e_u - e_{\widehat{u}}, (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h}, \end{aligned}$$

by Eq. (7c) with $w := \varphi_h$. Finally,

$$\begin{aligned} \Theta_h &= -\langle \mathbf{e}\widehat{\mathbf{q}} \cdot \mathbf{n}, \varphi_h - \widehat{\varphi}_h \rangle_{\partial \mathcal{T}_h} + \langle \varphi_h - \widehat{\varphi}_h, \mathbf{e}_q \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} + \langle e_u - e_{\widehat{u}}, (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \\ &= \langle (\mathbf{e}_q - \mathbf{e}\widehat{\mathbf{q}}) \cdot \mathbf{n}, \varphi_h - \widehat{\varphi}_h \rangle_{\partial \mathcal{T}_h} + \langle e_u - e_{\widehat{u}}, (\widehat{\boldsymbol{\psi}}_h - \boldsymbol{\psi}_h) \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} \\ &= 0, \end{aligned}$$

by the cancellation property of the traces, assumption (Biii).

Fig. 1 Initial unstructured triangulation



Step 2: The first estimate To obtain our first estimate, we first note that

$$\begin{aligned} (e_u, \theta)_{\mathcal{T}_h} &= (\mathbf{q}_h^2 - \mathbf{a}\mathbf{g}_h^2, \mathbf{c}\boldsymbol{\psi}_h)_{\mathcal{T}_h} \\ &= (\mathbf{q}_h^2 - \mathbf{a}\mathbf{g}_h^2, \mathbf{c}(\boldsymbol{\psi}_h - \boldsymbol{\psi}))_{\mathcal{T}_h} + (\mathbf{q}_h^2 - \mathbf{a}\mathbf{g}_h^2, \mathbf{c}\boldsymbol{\psi})_{\mathcal{T}_h} \\ &= (\mathbf{q}_h^2 - \mathbf{a}\mathbf{g}_h^2, \mathbf{c}(\boldsymbol{\psi}_h - \boldsymbol{\psi}))_{\mathcal{T}_h} + (\mathbf{q}_h^2 - \mathbf{a}\mathbf{g}_h^2, (I - P_{V_h})(\mathbf{c}\boldsymbol{\psi}))_{\mathcal{T}_h}, \end{aligned}$$

by the second of the equations in (7b) with $\mathbf{v} := P_{V_h}(\mathbf{c}\boldsymbol{\psi})$, where P_{V_h} is the L^2 -projection into V_h . Then we easily get that

$$\begin{aligned} (e_u, \theta)_{\mathcal{T}_h} &\leq \|\mathbf{q}_h^2 - \mathbf{a}\mathbf{g}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{c})} \|\boldsymbol{\psi}_h - \boldsymbol{\psi}\|_{L^2(\mathcal{T}_h; \mathbf{c})} + \|(I - P_{V_h})(\mathbf{c}\boldsymbol{\psi})\|_{L^2(\mathcal{T}_h; \mathbf{c})} \\ &\leq \|\mathbf{q}_h^2 - \mathbf{a}\mathbf{g}_h^2\|_{L^2(\mathcal{T}_h; \mathbf{c})} Ch^\alpha \|\boldsymbol{\psi}\|_{H^1(\mathcal{T}_h; \mathbf{c})}, \end{aligned}$$

by assumption (Cii) and the approximation properties of P_{V_h} in combination with assumption (Ci). The result now follows from the estimates of the errors in the vector unknowns and the elliptic regularity inequality.

Step 3: An auxiliary non-standard, approximation result The improved estimate of the error in the scalar unknown is more delicate to prove. To prove it, we are going to use the following simple but non-standard auxiliary result.

Lemma 3.1 Assume that $\mathcal{P}^1(K) \subseteq V(K)$. Then, the following estimate holds

$$\|(I - P_{V_h})((\mathbf{c} - \bar{\mathbf{c}})\boldsymbol{\psi})\|_{L^2(K; \mathbf{a})} \leq Ch^2 \|\mathbf{a}\|_{L^\infty(K)}^{1/2} \|\mathbf{c}\|_{W^{2,\infty}(K)} \|\boldsymbol{\psi}\|_{H^1(K)}.$$

Proof Since $\mathcal{P}^1(K) \subseteq V(K)$, we have that $(I - P_{V_h})\mathcal{P}^1(K) = \{\mathbf{0}\}$, and so

$$\|(I - P_{V_h})((\mathbf{c} - \bar{\mathbf{c}})\boldsymbol{\psi})\|_{L^2(K; \mathbf{a})} \leq \|\mathbf{a}\|_{L^\infty(K)}^{1/2} \inf_{\mathbf{v} \in \mathcal{P}^1(K)} \|(\mathbf{c} - \bar{\mathbf{c}})\boldsymbol{\psi} - \mathbf{v}\|_{L^2(K)}.$$

To estimate the right-hand side, we need a variation of Taylor’s expansion which would not use second-order derivatives of $\boldsymbol{\psi}$ (otherwise we would not be able to use the elliptic regularity inequality), but only those of \mathbf{c} . In the one-dimensional case, such variation is the

Table 5 Errors and estimated orders of convergence using RT_k on an unstructured triangulation

k	l	$\ u_h^1 - u_h^2\ _{L^2(\mathcal{T}_h)}$	e.o.c.	$\ q_h^1 - q_h^2\ _{L^2(\mathcal{T}_h;c)}$	e.o.c.	$\ g_h^1 - g_h^2\ _{L^2(\mathcal{T}_h;a)}$	e.o.c.
0	1	6.7e-02	–	3.3e-02	–	3.3e-01	–
	2	7.2e-01	-3.43	1.8e-02	0.90	7.9e-01	-1.25
	3	1.6e-01	2.21	4.5e-03	1.96	3.9e-01	1.00
	4	3.8e-02	2.04	9.9e-04	2.20	1.9e-01	1.06
	5	9.6e-03	1.98	2.4e-04	2.03	9.5e-02	0.99
1	1	1.8e-03	–	6.4e-03	–	6.4e-02	–
	2	6.0e-04	1.56	1.3e-03	2.35	2.6e-02	1.31
	3	8.1e-05	2.89	1.6e-04	2.94	6.7e-03	1.94
	4	9.4e-06	3.10	2.0e-05	3.04	1.6e-03	2.03
	5	1.2e-06	2.98	2.5e-06	2.99	4.1e-04	1.99
2	1	1.5e-03	–	2.0e-03	–	1.7e-02	–
	2	4.0e-03	-1.42	3.1e-04	2.72	7.6e-03	1.20
	3	2.3e-04	4.12	2.1e-05	3.88	1.0e-03	2.92
	4	1.5e-05	3.97	1.3e-06	3.96	1.3e-04	3.00
	5	9.3e-07	3.98	8.3e-08	4.00	1.6e-05	2.99
3	1	8.5e-05	–	3.0e-04	–	2.2e-03	–
	2	1.2e-05	2.86	1.8e-05	4.05	3.0e-04	2.87
	3	3.9e-07	4.92	6.0e-07	4.92	2.0e-05	3.92
	4	1.1e-08	5.15	1.8e-08	5.08	1.2e-06	4.02
	5	3.5e-10	4.98	5.6e-10	4.98	7.8e-08	3.99
4	1	3.5e-05	–	5.1e-05	–	5.6e-04	–
	2	2.1e-05	0.72	1.9e-06	4.75	3.8e-05	3.89
	3	3.2e-07	6.05	3.2e-08	5.88	1.2e-06	4.93
	4	5.1e-09	5.96	5.2e-10	5.95	4.0e-08	4.97

following identity:

$$f(s)g(s) = (f(0) + sf'(0))g(0) + \int_0^s [f(z)g'(z) + (s - z)(f''(z)g(z) + f'(z)g'(z))]dz.$$

Using this identity (with $f := (c - \bar{c})_{ij}$ and $g := \psi_j, i, j = 1, \dots, d$) and bounding each of the resulting terms using approximations in $\mathcal{P}^1(K)$, we easily get the following estimate for star-shaped, uniformly regular elements:

$$\inf_{v \in \mathcal{P}^1(K)} \|(c - \bar{c})\psi - v\|_{L^2(K)} \leq C h^2 \|c\|_{W^{2,\infty}} \|\psi\|_{H^1(K)}.$$

This completes the proof. □

Step 4: The improved estimate We are now ready to prove the last estimate. Since assumption (Aiii) holds, we have that $\bar{c}\psi_h \in V_h$, and we can take $v := \bar{c}\psi_h$ in the second of the

Table 6 Errors and estimated orders of convergence using BDM_k on an unstructured triangulation

k	l	$\ u_h^1 - u_h^2\ _{L^2(\mathcal{T}_h)}$	e.o.c.	$\ q_h^1 - q_h^2\ _{L^2(\mathcal{T}_h;c)}$	e.o.c.	$\ g_h^1 - g_h^2\ _{L^2(\mathcal{T}_h;a)}$	e.o.c.
1	1	3.7e-04	–	5.8e-03	–	7.4e-03	–
	2	2.3e-05	3.98	7.5e-04	2.95	8.8e-04	3.07
	3	1.5e-06	3.97	9.7e-05	2.95	1.1e-04	3.01
	4	9.5e-08	3.99	1.2e-05	2.98	1.4e-05	3.00
	5	5.9e-09	4.00	1.5e-06	2.99	1.7e-06	3.00
2	1	2.9e-05	–	1.1e-03	–	1.4e-03	–
	2	1.3e-06	4.49	1.1e-04	3.35	1.2e-04	3.57
	3	4.1e-08	4.98	7.1e-06	3.90	7.7e-06	3.93
	4	1.3e-09	5.00	4.5e-07	3.98	4.9e-07	3.98
	5	3.9e-11	5.00	2.8e-08	4.00	3.1e-08	3.99
3	1	8.8e-06	–	4.1e-04	–	5.2e-04	–
	2	1.3e-07	6.12	1.3e-05	5.01	1.5e-05	5.06
	3	1.9e-09	6.05	4.0e-07	4.99	4.8e-07	5.00
	4	2.9e-11	6.02	1.3e-08	5.00	1.5e-08	5.00
	5	4.5e-13	6.01	3.9e-10	5.00	4.7e-10	5.00
4	1	6.0e-07	–	3.4e-05	–	4.8e-05	–
	2	7.4e-09	6.32	9.4e-07	5.19	1.1e-06	5.40
	3	6.2e-11	6.90	1.6e-08	5.91	1.9e-08	5.92
	4	4.9e-13	6.98	2.5e-10	5.98	3.0e-10	5.98

Eq. (7b) to get that

$$\begin{aligned}
 (e_u, \theta)_{\mathcal{T}_h} &= (q_h^2 - a g_h^2, c \psi_h)_{\mathcal{T}_h} \\
 &= (q_h^2 - a g_h^2, (c - \bar{c}) \psi_h)_{\mathcal{T}_h} \\
 &= (q_h^2 - a g_h^2, (c - \bar{c}) \psi)_{\mathcal{T}_h} + (q_h^2 - a g_h^2, (c - \bar{c})(\psi_h - \psi))_{\mathcal{T}_h}, \\
 &= (q_h^2 - a g_h^2, (I - P_{V_h})((c - \bar{c}) \psi))_{\mathcal{T}_h} + (q_h^2 - a g_h^2, (c - \bar{c})(\psi_h - \psi))_{\mathcal{T}_h},
 \end{aligned}$$

by the second of the Eq. (7b) with $v := P_{V_h}((c - \bar{c}) \psi)$. Then,

$$\begin{aligned}
 (e_u, \theta)_{\mathcal{T}_h} &\leq \|q_h^2 - a g_h^2\|_{L^2(\mathcal{T}_h;c)} \|(I - P_{V_h})((c - \bar{c}) \psi)\|_{L^2(\mathcal{T}_h;a)} \\
 &\quad + \|q_h^2 - a g_h^2\|_{L^2(\mathcal{T}_h;c)} \|a^{\frac{1}{2}}(c - \bar{c})a^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)} \|\psi - \psi_h\|_{L^2(\mathcal{T}_h;c)} \\
 &\leq \|q_h^2 - a g_h^2\|_{L^2(\mathcal{T}_h;c)} Ch^2 \|a\|_{L^\infty(\mathcal{T}_h)}^{1/2} \|c\|_{W^{2,\infty}(\mathcal{T}_h)} \|\psi\|_{H^1(\mathcal{T}_h)} \\
 &\quad + \|q_h^2 - a g_h^2\|_{L^2(\mathcal{T}_h;c)} \|a^{\frac{1}{2}}(c - \bar{c})a^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)} Ch^\alpha \|\psi\|_{H^1(\mathcal{T}_h)},
 \end{aligned}$$

by Lemma 3.1 and by assumption (Cii). The improved estimate now follows by using the elliptic regularity inequality.

It remains to show the last statement of Theorem 1.

Table 7 Errors and estimated orders of convergence using DG_k , with $C_{11} = 1$, $C_{12} = [1, 1]^t$, $C_{22} = 1$ on an unstructured triangulation

k	l	$\ u_h^1 - u_h^2\ _{L^2(\mathcal{T}_h)}$	e.o.c.	$\ q_h^1 - q_h^2\ _{L^2(\mathcal{T}_h;c)}$	e.o.c.	$\ g_h^1 - g_h^2\ _{L^2(\mathcal{T}_h;a)}$	e.o.c.
0	1	1.9e-03	–	8.4e-03	–	3.6e-02	–
	2	6.3e-04	1.57	4.2e-03	1.02	1.1e-02	1.66
	3	5.0e-04	0.34	3.1e-03	0.41	7.9e-03	0.51
	4	2.0e-04	1.34	1.1e-03	1.54	2.6e-03	1.60
	5	5.8e-05	1.77	3.0e-04	1.84	7.1e-04	1.87
1	1	2.8e-03	–	3.4e-02	–	9.0e-02	–
	2	8.9e-04	1.64	1.1e-02	1.67	2.0e-02	2.15
	3	1.0e-04	3.14	2.1e-03	2.37	3.2e-03	2.68
	4	6.2e-06	4.02	2.8e-04	2.86	3.9e-04	3.02
	5	3.8e-07	4.01	3.7e-05	2.94	4.8e-05	3.02
2	1	4.5e-03	–	4.3e-02	–	7.7e-02	–
	2	2.7e-04	4.06	8.0e-03	2.44	9.6e-03	3.00
	3	8.6e-06	4.96	4.0e-04	4.33	5.3e-04	4.19
	4	3.0e-07	4.83	2.5e-05	3.98	3.5e-05	3.91
	5	9.8e-09	4.94	1.6e-06	4.00	2.2e-06	3.97
3	1	1.9e-03	–	4.4e-02	–	5.1e-02	–
	2	2.0e-05	6.56	4.5e-04	6.62	9.0e-04	5.84
	3	7.7e-07	4.67	5.2e-05	3.10	6.9e-05	3.69
	4	1.2e-08	5.99	1.7e-06	4.93	2.2e-06	4.99
	5	1.9e-10	6.00	5.4e-08	4.97	6.8e-08	5.01
4	1	4.3e-04	–	6.1e-03	–	1.2e-02	–
	2	8.7e-06	5.63	4.1e-04	3.88	5.3e-04	4.55
	3	5.8e-08	7.23	5.0e-06	6.38	6.5e-06	6.37
	4	4.8e-10	6.90	7.9e-08	5.99	1.0e-07	5.96

3.4 Proof of Theorem 2

Using the triangle inequality in the estimate of the error in the flux, we get

$$\begin{aligned} \|e_q\|_{L^2(\mathcal{T}_h;c)} &\leq (\|q - q_h^1\|_{L^2(\mathcal{T}_h;c)} + \|e_q\|_{L^2(\mathcal{T}_h;c)} \\ &\quad + \|g - g_h^1\|_{L^2(\mathcal{T}_h;a)} + \|e_g\|_{L^2(\mathcal{T}_h;a)}) \|a^{\frac{1}{2}}(c - \bar{c})a^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)}. \end{aligned}$$

But by **Step 2** of the estimates of the vector unknowns, we can obtain that

$$\|e_g\|_{L^2(\mathcal{T}_h;c)} \leq \|q_h^2 - a g_h^2\|_{L^2(\mathcal{T}_h;c)},$$

without having to use Assumption (Ai). As a consequence, we readily get that

$$\|e_q\|_{L^2(\mathcal{T}_h;c)} \leq C (\|q - q_h^1\|_{L^2(\mathcal{T}_h;c)} + \|g - g_h^1\|_{L^2(\mathcal{T}_h;a)}) \|a^{\frac{1}{2}}(c - \bar{c})a^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)},$$

where $C := 2/(1 - \|a^{\frac{1}{2}}(c - \bar{c})a^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)}) \leq 2/(1 - \kappa)$ by our assumption. We can then write that

Table 8 Errors and estimated orders of convergence using HDG_k, with stabilization parameter $\tau = 1$, using unstructured triangular meshes (top) and uniform Cartesian meshes (bottom)

k	l	$\ u_h^1 - u_h^2\ _{L^2(\mathcal{T}_h)}$	e.o.c.	$\ q_h^1 - q_h^2\ _{L^2(\mathcal{T}_h;c)}$	e.o.c.	$\ g_h^1 - g_h^2\ _{L^2(\mathcal{T}_h;a)}$	e.o.c.
<i>Triangles</i>							
0	1	2.3e-03	–	8.2e-03	–	2.7e-02	–
	2	5.1e-04	2.15	2.3e-03	1.82	6.2e-03	2.11
	3	1.3e-04	1.98	6.1e-04	1.93	1.6e-03	1.99
	4	3.3e-05	1.98	1.6e-04	1.97	3.9e-04	1.99
	5	8.2e-06	1.99	3.9e-05	1.99	9.9e-05	1.99
1	1	3.1e-04	–	4.9e-03	–	6.5e-03	–
	2	2.3e-05	3.79	7.3e-04	2.76	8.5e-04	2.94
	3	1.5e-06	3.92	9.6e-05	2.91	1.1e-04	2.98
	4	9.5e-08	3.98	1.2e-05	2.97	1.4e-05	3.00
	5	6.0e-09	3.99	1.5e-06	2.99	1.7e-06	3.00
2	1	3.0e-05	–	1.2e-03	–	1.5e-03	–
	2	1.4e-06	4.49	1.1e-04	3.43	1.2e-04	3.64
	3	4.3e-08	4.99	7.1e-06	3.91	7.7e-06	3.95
	4	1.3e-09	5.01	4.5e-07	3.98	4.9e-07	3.98
	5	4.1e-11	5.00	2.8e-08	4.00	3.1e-08	3.99
3	1	8.9e-06	–	4.0e-04	–	5.0e-04	–
	2	1.3e-07	6.09	1.3e-05	4.99	1.5e-05	5.04
	3	2.0e-09	6.03	4.0e-07	4.98	4.8e-07	4.99
	4	3.1e-11	6.02	1.3e-08	4.99	1.5e-08	4.99
	5	4.8e-13	6.01	3.9e-10	5.00	4.7e-10	5.00
4	1	6.2e-07	–	3.5e-05	–	5.0e-05	–
	2	7.8e-09	6.32	9.4e-07	5.23	1.1e-06	5.46
	3	6.5e-11	6.90	1.6e-08	5.91	1.9e-08	5.92
	4	5.2e-13	6.98	2.5e-10	5.98	3.0e-10	5.98
<i>Squares</i>							
0	1	1.5e-03	–	6.1e-03	–	1.5e-02	–
	2	1.1e-03	0.46	4.5e-03	0.43	1.9e-02	-0.29
	3	2.2e-04	2.36	7.2e-03	-0.67	4.0e-03	2.22
	4	3.2e-05	2.79	2.1e-03	1.78	7.6e-04	2.40
	5	4.6e-06	2.79	5.5e-04	1.93	1.4e-04	2.46
	6	6.6e-07	2.80	1.4e-04	1.97	2.4e-05	2.54
	7	9.2e-08	2.83	3.5e-05	1.98	3.9e-06	2.62
1	1	4.8e-03	–	3.2e-02	–	6.2e-02	–
	2	3.3e-04	3.87	3.0e-03	3.41	8.9e-03	2.81
	3	2.7e-05	3.61	6.7e-04	2.17	1.3e-03	2.81
	4	2.2e-06	3.63	1.4e-04	2.21	1.6e-04	2.97
	5	1.7e-07	3.71	2.2e-05	2.73	2.0e-05	3.02
	6	1.2e-08	3.84	2.9e-06	2.89	2.4e-06	3.02

Table 8 continued

k	l	$\ u_h^1 - u_h^2\ _{L^2(\mathcal{T}_h)}$	e.o.c.	$\ q_h^1 - q_h^2\ _{L^2(\mathcal{T}_h;c)}$	e.o.c.	$\ g_h^1 - g_h^2\ _{L^2(\mathcal{T}_h;a)}$	e.o.c.
	7	7.3e-10	4.00	3.8e-07	2.95	3.0e-07	3.01
2	1	1.3e-04	–	1.8e-03	–	3.5e-03	–
	2	9.0e-05	0.48	2.2e-03	-0.30	3.9e-03	-0.15
	3	2.4e-06	5.23	1.4e-04	3.97	2.9e-04	3.75
	4	8.5e-08	4.80	1.0e-05	3.79	1.8e-05	3.99
	5	2.7e-09	5.01	6.7e-07	3.90	1.1e-06	4.01
	6	8.0e-11	5.05	4.3e-08	3.96	7.1e-08	4.01
3	1	1.1e-03	–	3.3e-02	–	1.8e-02	–
	2	1.6e-05	6.08	5.3e-04	5.99	1.0e-03	4.13
	3	2.8e-07	5.84	1.8e-05	4.84	3.4e-05	4.92
	4	4.3e-09	6.05	5.8e-07	5.00	1.1e-06	4.97
	5	6.6e-11	6.01	1.8e-08	4.99	3.4e-08	4.99
4	1	2.4e-05	–	5.2e-04	–	1.0e-03	–
	2	2.3e-06	3.33	1.0e-04	2.32	1.8e-04	2.51
	3	2.3e-08	6.65	1.6e-06	6.00	3.1e-06	5.87
	4	1.9e-10	6.91	2.7e-08	5.92	4.9e-08	5.98
	5	1.5e-12	6.99	4.3e-10	5.96	7.6e-10	6.00

$$\|e_q\|_{L^2(\mathcal{T}_h;c)} \leq \frac{2}{1-\kappa} \|a^{\frac{1}{2}}(c - \bar{c})a^{\frac{1}{2}}\|_{L^\infty(\mathcal{T}_h)} \min_{i=1,2} \{\|q - q_h^i\|_{L^2(\mathcal{T}_h;c)} + \|g - g_h^i\|_{L^2(\mathcal{T}_h;a)}\}.$$

The other estimates follow in a similar manner.

This completes the Proof of Theorem 2.

4 Numerical Experiments

We present numerical experiments devised to corroborate our theoretical results on supercloseness. To do that, we take $\Omega := (0, 1)^2$, $\partial\Omega_D = \partial\Omega$, and set f and u_D such that the exact solution of our model problem with

$$a(x) = \begin{pmatrix} x_1 + x_2 + 1 & -x_2 \\ -x_2 & 2x_1 + x_2 + 1 \end{pmatrix},$$

is $u = \sin(2\pi x_1) \sin(2\pi x_2)$. We consider an initial unstructured triangulation (see Fig. 1) and we estimate the orders of convergence as we refine uniformly, indexing the meshes with a mesh parameter h_l , for $l = 1, 2, 3, 4, 5$. We do this piecewise polynomial approximations of degree $k = 12, 3, 4$; we also take $k = 0$ for the RT_k, DG_k and HDG_k methods).

The results for the RT_k, BDM_k, DG_k (with parameters $C_{11} = 1.0, C_{22} = 1.0$ and $C_{12} = [1.0, 1.0]^t$) and HDG_k method (with $\tau = 1.0$) are displayed on the Tables 5, 6, 7 and 8, respectively. In all cases, we observe the supercloseness orders predicted by Theorem 1 and displayed in Table 4. It is interesting to see that for the RT_k method, the difference of the approximations of the flux shows one order of convergence more than the for difference of the approximations of the gradient, which reflects the fact that for the RT_k method assumption (Bii) is satisfied but assumption (Bi) is not satisfied. This difference in the convergence

properties of the approximate fluxes and gradients does not appear in the BDM_k method as this method satisfies both assumptions.

For uniform meshes of squares of size h , we estimate the orders of convergence as we refine the mesh by taking $h = 2^{-l}$, for $l = 1, 2, 3, 4, 5$. The results for the HDG $_k$ method, with $\tau = 1.0$, are displayed in the bottom of Table 8. Again, we do observe the same orders of supercloseness as the ones predicted by Theorem 1 and displayed in Table 4.

5 Extensions and Concluding Remarks

We have proved the supercloseness property of two Galerkin formulations for second-order elliptic problems. Our analysis holds for a wide class of mixed finite element methods, as for instance the Raviart–Thomas [19] and Brezzi–Douglas–Marini [3] elements, discontinuous Galerkin methods, and hybridizable discontinuous Galerkin methods [11].

Although we have not treated the interior penalty (IP) method [2], it is easy to get even stronger results by using slight modifications to our approach. Indeed, even though our theory does not apply directly, since the definition of the numerical traces does not necessarily satisfy assumptions B (ii) and B (iii), it is not difficult to show, for the first two formulations of the IP method, that difference of the approximations converge with order $k + 3$, $k + 3$ and $k + 2$ for the scalar, the gradient and the flux approximations, respectively. Even more, it is a simple exercise to see that, if the tensor a is piecewise linear then the IP schemes give the same approximation for the scalar and gradient variables.

We believe that it is reasonable to expect that similar results hold for the corresponding formulations and numerical methods for linear elasticity.

Acknowledgements We wish to thank one of the referees for constructive criticism leading to a better presentation of the material, including the discussion about the IP methods which we originally did not consider.

References

1. Arbogast, T., Wheeler, M.F., Yotov, I.: Mixed finite elements for elliptic problems with tensor coefficients as cell-centered finite differences. *SIAM J. Numer. Anal.* **34**(2), 828–852 (1997)
2. Arnold, N.D.: An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.* **19**(4), 742–760 (1982)
3. Brezzi, F., Douglas Jr., J., Marini, L.D.: Two families of mixed finite elements for second order elliptic problems. *Numer. Math.* **47**(2), 217–235 (1985)
4. Brezzi, F., Fortin, M.: *Mixed and Hybrid Finite Element Methods*. Springer Series in Computational Mathematics, vol. 15. Springer, New York (1991)
5. Castillo, P., Cockburn, B., Perugia, I., Schötzau, D.: An a priori error analysis of the local discontinuous Galerkin method for elliptic problems. *SIAM J. Numer. Anal.* **38**(5), 1676–1706 (2000)
6. Ciarlet, P.G.: *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)] (2002)
7. Cockburn, B.: Static condensation, hybridization, and the devising of the HDG methods. In: Barrenechea, G.R., Brezzi, F., Cagniani, A., Georgoulis, E.H. (eds) *Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations*, volume 114 of *Lecture Notes of Computer Science Engineering*, pp. 129–177. Springer, Berlin, 2016. LMS Durham Symposia funded by the London Mathematical Society. Durham, 8–16 July (2014)
8. Cockburn, B., Di-Pietro, D.A., Ern, A.: Bridging the hybrid high-order and hybridizable discontinuous Galerkin methods. *ESAIM Math. Model. Numer. Anal.* **50**(3), 635–650 (2016)
9. Cockburn, B., Dong, B., Guzmán, J.: A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems. *Math. Comput.* **77**(264), 1887–1916 (2008)

10. Cockburn, B., Dong, B., Guzmán, J.: A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems. *Math. Comput.* **77**, 1887–1916 (2008)
11. Cockburn, B., Gopalakrishnan, J., Lazarov, R.: Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.* **47**(2), 1319–1365 (2009)
12. Cockburn, B., Gopalakrishnan, J., Sayas, F.-J.: A projection-based error analysis of HDG methods. *Math. Comput.* **79**, 1351–1367 (2010)
13. Cockburn, B., Guzmán, J., Wang, H.: Superconvergent discontinuous Galerkin methods for second-order elliptic problems. *Math. Comput.* **78**, 1–24 (2009)
14. Cockburn, B., Shen, J.: A hybridizable discontinuous Galerkin method for the p -Laplacian. *SIAM J. Sci. Comput.* **38**(1), A545–A566 (2016)
15. Cockburn, B., Shi, K.: Superconvergent HDG methods for linear elasticity with weakly symmetric stresses. *IMA J. Numer. Anal.* **33**(3), 747–770 (2013)
16. Di-Pietro, D.A., Ern, A.: A hybrid high-order locking-free method for linear elasticity on general meshes. *Comput. Methods Appl. Mech. Eng.* **283**, 1–21 (2015)
17. Di-Pietro, D.A., Ern, A.: Hybrid high-order methods for variable-diffusion problems on general meshes. *C. R. Acad. Sci Paris Ser. I* **353**, 31–34 (2015)
18. Di-Pietro, D.A., Ern, A., Lemaire, S.: An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators. *Comput. Methods Appl. Math.* **14**(4), 461–472 (2014)
19. Raviart, P.-A., Thomas, J.M.: A mixed finite element method for 2nd order elliptic problems. In: *Mathematical Aspects of Finite Element Methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975)*, Lecture Notes in Mathematics, Vol. 606, pp. 292–315. Springer, Berlin (1977)
20. Soon, S.-C.: *Hybridizable discontinuous Galerkin methods for solid mechanics*. Ph.D. thesis, University of Minnesota, Minneapolis (2008)
21. Soon, S.-C., Cockburn, B., Stolarski, H.K.: A hybridizable discontinuous Galerkin method for linear elasticity. *Int. J. Numer. Methods Eng.* **80**(8), 1058–1092 (2009)