



HAL
open science

Quasi-optimal nonconforming approximation of elliptic PDES with contrasted coefficients and minimal regularity

Alexandre Ern, Jean-Luc Guermond

► **To cite this version:**

Alexandre Ern, Jean-Luc Guermond. Quasi-optimal nonconforming approximation of elliptic PDES with contrasted coefficients and minimal regularity. 2018. hal-01964299v1

HAL Id: hal-01964299

<https://hal.science/hal-01964299v1>

Preprint submitted on 21 Dec 2018 (v1), last revised 24 Nov 2021 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **QUASI-OPTIMAL NONCONFORMING APPROXIMATION OF**
2 **ELLIPTIC PDES WITH CONTRASTED COEFFICIENTS AND**
3 **MINIMAL REGULARITY***

4 ALEXANDRE ERN[†] AND JEAN-LUC GUERMOND[‡]

5 **Abstract.** In this paper we investigate the approximation of a diffusion model problem with
6 contrasted diffusivity and the error analysis of various nonconforming approximation methods. The
7 essential difficulty is that the Sobolev smoothness index of the exact solution may be just barely larger
8 than one. The lack of smoothness is handled by giving a weak meaning to the normal derivative of
9 the exact solution at the mesh faces. The error estimates are robust with respect to the diffusivity
10 contrast. We briefly show how the analysis can be extended to the Maxwell's equations.

11 **Key words.** Finite elements, Nonconforming methods, Error estimates, Minimal regularity,
12 Nitsche method, Boundary penalty, Elliptic equations, Maxwell's equations.

13 **AMS subject classifications.** 35J25, 65N15, 65N30

14 *This article is dedicated to the memory of Christine Bernardi.*

15 **1. Introduction.** The objective of the present paper is to revisit and unify
16 the error analysis of various nonconforming approximation techniques applied to a
17 diffusion model problem with contrasted diffusivity. We also briefly show how to
18 extend the analysis to Maxwell's equations.

19 **1.1. Content of the paper.** The nonconforming techniques we have in mind
20 are Crouzeix–Raviart finite elements [14], Nitsche's boundary penalty method [32],
21 the interior penalty discontinuous Galerkin (IPDG) method [2], and the hybrid high-
22 order (HHO) methods [16, 18] which are closely related to hybridizable discontinuous
23 Galerkin methods [13]. The main difficulty in the error analysis is that owing to
24 the contrast in the diffusivity, the Sobolev smoothness index of the exact solution
25 is barely larger than one. This makes the estimation of the consistency error in-
26 curred by nonconforming approximation techniques particularly challenging since the
27 normal derivative of the solution at the mesh faces is not integrable and it is thus
28 not straightforward to give a reasonable meaning to this quantity on each mesh face
29 independently.

30 The main goal of the present paper is to establish quasi-optimal error estimates by
31 using a mesh-dependent norm that remains bounded as long as the exact solution has
32 a Sobolev smoothness index strictly larger than one. By quasi-optimality, we mean
33 that the approximation error measured in the augmented norm is bounded, up to a
34 generic constant, by the best approximation error of the exact solution measured in
35 the same augmented norm by members of the discrete trial space. A key point in the
36 analysis is that the above generic constant is independent of the diffusivity contrast.
37 We emphasize that quasi-optimal error estimates are more informative than the more
38 traditional asymptotic error estimates, which bound the approximation error by terms

*This material is based upon work supported in part by the National Science Foundation grants DMS-1619892, DMS-1620058, by the Air Force Office of Scientific Research, USAF, under grant/contract number FA9550-18-1-0397, and by the Army Research Office under grant/contract number W911NF-15-1-0517.

[†]Department of Mathematics, Texas A&M University 3368 TAMU, College Station, TX 77843, USA.

[‡]Université Paris-Est, CERMICS (ENPC), 77455 Marne-la-Vallée cedex 2, France and INRIA Paris, 2 rue Simone Iff, 75589 Paris, France.

39 that optimally decay with the mesh size. Indeed, the former estimates cover the whole
 40 computational range whereas the latter estimates only cover the asymptotic range.
 41 One key novelty herein is the introduction of a weighted bilinear form that accounts
 42 for the default of consistency in all the cases (see (3.12)).

43 The paper is organized as follows. The model problem under consideration and
 44 the discrete setting are introduced in §2. The weighted bilinear form mentioned above
 45 which accounts for the consistency default at the mesh interfaces and boundary faces
 46 is defined in §3. The key results in this section are Lemma 3.3 and Lemma 3.5. We
 47 collect in §4 the error analyses of the approximation of the model problem with the
 48 Crouzeix–Raviart approximation, Nitsche’s boundary penalty method, the IPDG ap-
 49 proximation, and the HHO approximation. To avoid invoking Strang’s second Lemma,
 50 we introduce in §4.1 a linear form δ_h that measures consistency but does not need the
 51 exact solution to be inserted into the arguments of the discrete bilinear form at hand.
 52 The weighted bilinear form (3.12) turns out to an essential tool to deduce robust
 53 estimates of the norm of the consistency form δ_h for all the nonconforming methods
 54 considered. One originality of this paper is that all the error estimates provided in §4
 55 involve constants that are uniform with respect to the diffusivity contrast. Another
 56 salient feature is that the source term is assumed to be only in $L^q(D)$, where q is
 57 such that $L^q(D)$ is continuously embedded in $H^{-1}(D) := (H_0^1(D))'$; specifically, this
 58 means that $q > 2_* := \frac{2d}{2+d} \geq 1$ (here, $d \geq 2$ is the space dimension).

59 **1.2. Literature overview.** Let us put our work in perspective with the liter-
 60 ature. Perhaps a bit surprisingly, error estimates for nonconforming approximation
 61 methods are rarely presented in a quasi-optimal form in the literature. A key step to-
 62 ward achieving quasi-optimal error estimates has been achieved in Veerer and Zanotti
 63 [34, 35]. Therein, the approximation error and the best-approximation error are both
 64 measured using the energy norm and the source term is assumed to be just in the
 65 dual space $H^{-1}(D)$. However, at the time of this writing, this setting does not yet
 66 cover robust estimates w.r.t. the diffusivity contrast. In the present work, we proceed
 67 somewhat differently to obtain robust quasi-optimal error estimates. This is done at
 68 the following price: (i) We invoke augmented norms, which are, however, compatible
 69 with the elliptic regularity theory; (ii) We only consider source terms in the Lebesgue
 70 spaces $L^q(D)$ with $q > 2_* := \frac{2d}{2+d} \geq 1$; notice though that this regularity is weaker
 71 than assuming that source terms are in $L^2(D)$, as usually done in the literature.

72 The traditional approach to tackle the error analysis for nonconforming approxi-
 73 mation techniques are Strang’s lemmas. However, an important shortcoming of this
 74 approach whenever the Sobolev smoothness index of the exact solution is barely larger
 75 than one, is that it is not possible to insert the exact solution in the first argument of
 76 the discrete bilinear form. To do so, one needs to assume some additional regularity
 77 on the exact solution which often goes beyond the regularity provided by the prob-
 78 lem at hand. This approach has nevertheless been used by many authors to analyze
 79 discontinuous Galerkin (dG) methods (see, e.g., [15, 21] and the references therein).
 80 One way to overcome the limitations of Strang’s Second Lemma has been proposed
 81 by Gudi [29]. The key idea consists of introducing a mapping that transforms the
 82 discrete test functions into elements of the exact test space. An important property
 83 of this operator is that its kernel is composed of discrete (test) functions that are
 84 only needed to “stabilize” the discrete bilinear form, but do not contribute to the
 85 interpolation properties of the approximation setting. We refer to this mapping as
 86 trimming operator. The notion of trimming operator has been used in Li and Mao
 87 [31] to perform the analysis of the Crouzeix–Raviart approximation of the diffusion

88 problem and source term in $L^2(D)$ (see e.g., the definitions (5)–(7) and the identity
 89 (11) therein). The trimmed error estimate (which is sometimes referred to as “medius
 90 analysis” in the literature) has been applied in Gudi [29] to the IPDG approximation
 91 of the Laplace equation with a source term in $L^2(D)$ and to a fourth-order problem; it
 92 has been applied to the Stokes equations in Badia et al. [3] and to the linear elasticity
 93 equations in Carstensen and Schedensack [12]. One problem with methods using the
 94 trimming operator, though, is that they require constructing H^1 -conforming discrete
 95 quasi-approximation operators that do not account for the diffusivity contrast; this
 96 entails error estimates with constants that depend on the diffusivity contrast, i.e.,
 97 these error estimates are not robust.

98 It is shown in [25] in the case of Nitsche’s boundary penalty method that the
 99 dependency of the constants with respect to the diffusivity contrast can be eliminated
 100 by introducing an alternative technique based on mollification and an extension of
 101 the notion of the normal derivative. The objective of the present paper is to revisit
 102 and extend [25]. The analysis presented here is significantly simplified and modified
 103 to include the Crouzeix–Raviart approximation, the IPDG approximation, and the
 104 HHO approximation. One key novelty is the introduction of the weighted bilinear form
 105 (3.12) that accounts for the consistency default in all the cases. The present analysis
 106 hinges on two key ideas which are now part of the numerical analysis folklore. To
 107 the best of our knowledge, these ideas have been introduced/used in Lemma 4.7 in
 108 Amrouche et al. [1], Lemma 2.3 and Corollary 3.1 in Bernardi and Hecht [5] and
 109 Lemma 8.2 in Buffa and Perugia [9]. However, we believe that detailed proofs are
 110 seemingly missing in the literature, and another purpose of this paper is to fill this
 111 gap.

112 The first key idea is a face-to-cell lifting operator. Such an operator is mentioned
 113 in Lemma 4.7 in [1], and its construction is briefly discussed. The weights used in
 114 the norms therein, though, cannot give estimates that are uniform with respect to
 115 the mesh size. This operator is also mentioned in Lemma 2.3 in [5]. The authors
 116 claim that the face-to-cell operator has been constructed in Bernardi and Girault [4,
 117 Eq. (5.1)], which is unclear to us. A similar operator is invoked in Lemma 8.2 in
 118 [9]. The operator therein is constructed on the reference element \hat{K} and its stability
 119 properties are proved in the Sobolev scale $(H^s(\hat{K}))_{s \in (0,1)}$. The authors invoke also
 120 the Sobolev scale $(H^s(K))_{s \in (0,1)}$ for arbitrary cells K in a mesh \mathcal{T}_h belonging the
 121 shape-regular sequence $(\mathcal{T}_h)_{h \in \mathcal{H}}$. The norm equipping $H^s(K)$ is not explicitly defined
 122 therein, which leads to one statement that looks questionable (see e.g., Eq. (8.11)
 123 therein; a fix has been proposed in [8, Lem. A.3]). In particular, it is unclear how
 124 to keep track of constants that depend on K when one uses the real interpolation
 125 method to define $H^s(K)$. In order to clarify the status of this face-to-cell operator,
 126 which is essential for our analysis, and without claiming originality, we give (recall)
 127 all the details of its construction in the proof of Lemma 3.1. As in [1, Lem. 4.7], we
 128 use the Sobolev–Slobodeckij norm to equip the fractional-order Sobolev spaces; this
 129 allows us to track all the constants easily.

130 The second key idea introduced in the above papers is that of extending the notion
 131 of face integrals by using a duality argument together with the face-to-cell operator.
 132 The argument is deployed in Corollary 3.3 in [5], but the sketch of the proof has typos
 133 (e.g., an average has to be removed to make the inverse estimate in step (1) correct).
 134 This corollary is quoted and invoked in Cai et al. [11, Lem. 2.1]; it is the cornerstone
 135 of the argumentation therein. This argument is also deployed in Lemma 8.2 in [9].
 136 A similar argument is invoked in [1, Lem. 4.7] in a slightly different context. In all

137 the cases one must use a density argument to complete the proofs, but this argument
 138 is omitted and implicitly assumed to hold true in all the above references. We fill
 139 this gap in Lemma 3.3 and provide the full argumentation in the proof, including
 140 the passage to the limit by density. The proof invokes mollifiers that commute with
 141 differential operators and behave properly at the boundary of the domain; these tools
 142 have been recently revisited in [22] elaborating on seminal ideas from Schöberl [33].

143 **2. Preliminaries.** In this section we introduce the model problem and the dis-
 144 crete setting for the approximation.

145 **2.1. Model problem.** Let D be a Lipschitz domain in \mathbb{R}^d , which we assume for
 146 simplicity to be a polyhedron. We consider the following scalar model problem:

$$147 \quad (2.1) \quad -\nabla \cdot (\lambda \nabla u) = f \quad \text{in } D, \quad \gamma^{\mathfrak{g}}(u) = g \quad \text{on } \partial D,$$

148 where $\gamma^{\mathfrak{g}} : H^1(D) \rightarrow H^{\frac{1}{2}}(\partial D)$ is the usual trace map (the superscript \mathfrak{g} refers to
 149 the gradient), and $g \in H^{\frac{1}{2}}(\partial D)$ is the Dirichlet boundary data. The scalar-valued
 150 diffusion coefficient $\lambda \in L^\infty(D)$ is assumed to be uniformly bounded from below away
 151 from zero. For simplicity, we also assume that λ is piecewise constant in D , i.e., there
 152 is a partition of D into M disjoint Lipschitz polyhedra D_1, \dots, D_M s.t. $\lambda|_{D_i}$ is a
 153 positive real number for all $i \in \{1:M\}$.

154 It is standard in the literature to assume that $f \in L^2(D)$. We are going to relax
 155 this hypothesis in this paper by only assuming that $f \in L^q(D)$ with $q > \frac{2d}{2+d}$. Note
 156 that $q > 1$ since $d \geq 2$. Note also that $L^q(D) \hookrightarrow H^{-1}(D)$ since $H_0^1(D) \hookrightarrow H^{q'}(D)$
 157 with the convention that $\frac{1}{q} + \frac{1}{q'} = 1$. Since $\frac{2d}{2+d} < 2$, we are going to assume without
 158 loss of generality that $q \leq 2$.

159 In the case of the homogeneous Dirichlet condition ($g = 0$), the weak formulation
 160 of the model problem (2.1) is as follows:

$$161 \quad (2.2) \quad \begin{cases} \text{Find } u \in V := H_0^1(D) \text{ such that} \\ a(u, w) = \ell(w), \quad \forall w \in V, \end{cases}$$

162 with the bilinear and linear forms

$$163 \quad (2.3) \quad a(v, w) := \int_D \lambda \nabla v \cdot \nabla w \, dx, \quad \ell(w) := \int_D f w \, dx.$$

164 The bilinear form a is coercive in V owing to the Poincaré–Steklov inequality, and it
 165 is also bounded on $V \times V$ owing to the Cauchy–Schwarz inequality. The linear form ℓ
 166 is bounded on V since the Sobolev embedding theorem and Hölder’s inequality imply
 167 that $|\ell(w)| \leq \|f\|_{L^q(D)} \|w\|_{L^{q'}(D)} \leq c \|f\|_{L^q(D)} \|w\|_{H^1(D)}$. Note that $q \geq \frac{2d}{2+d}$ is the
 168 minimal integrability requirement on f for this boundedness property to hold true.
 169 The above coercivity and boundedness properties combined with the Lax–Milgram
 170 Lemma imply that (2.2) is well-posed. For the non-homogeneous Dirichlet boundary
 171 condition, one invokes the surjectivity of the trace map $\gamma^{\mathfrak{g}}$ to infer the existence of a
 172 lifting of g , say $u_g \in H^1(D)$ s.t. $\gamma^{\mathfrak{g}}(u_g) = g$, and one decomposes the exact solution
 173 as $u = u_g + u_0$ where $u_0 \in H_0^1(D)$ solves the weak problem (2.2) with $\ell(w)$ replaced
 174 by $\ell_g(w) = \ell(w) - a(u_g, w)$. The weak formulation thus modified is well-posed since
 175 ℓ_g is bounded on $H_0^1(D)$.

176 The notion of diffusive flux, which is defined as follows, will play an important
 177 role in the paper:

$$178 \quad (2.4) \quad \boldsymbol{\sigma}(v) := -\lambda \nabla v \in \mathbf{L}^2(D), \quad \forall v \in H^1(D).$$

179 We use boldface notation to denote vector-valued functions and vectors in \mathbb{R}^d .

180 LEMMA 2.1 (Exact solution). *Assume that there exist $r > 0$ and $q \in (\frac{2d}{2+d}, 2]$*
 181 *such that the exact solution u is in $H^{1+r}(D)$ and the source term f is in $L^q(D)$, then*

182 (2.5)
$$u \in V_s := \{v \in H_0^1(D) \mid \sigma(v) \in \mathbf{L}^p(D), \nabla \cdot \sigma(v) \in L^q(D)\},$$

183 *for some real number $p > 2$. \square*

184 *Proof.* The Sobolev embedding theorem implies that there is $p > 2$ s.t. $\mathbf{H}^r(D) \hookrightarrow$
 185 $\mathbf{L}^p(D)$. Indeed, if $2r < d$, we have $\mathbf{H}^r(D) \hookrightarrow \mathbf{L}^s(D)$ for all $s \in [2, \frac{2d}{d-2r}]$ and we can
 186 take $p = \frac{2d}{d-2r} > 2$, whereas if $2r \geq d$, we have $\mathbf{H}^r(D) \hookrightarrow \mathbf{H}^{\frac{d}{2}}(D) \hookrightarrow \mathbf{L}^s(D)$ for
 187 all $s \in [2, \infty)$, and we can take any $p > 2$. The above argument implies that $\nabla u \in$
 188 $\mathbf{L}^p(D)$, and since λ is piecewise constant and $\sigma(u) = -\lambda \nabla u$, we have $\sigma(u) \in \mathbf{L}^p(D)$.
 189 Moreover, since $\nabla \cdot \sigma(u) = f$ and $f \in L^q(D)$, we have $\nabla \cdot \sigma(u) \in L^q(D)$. \square

190 The regularity assumption $u \in H^{1+r}(D)$, $r > 0$, is reasonable owing to the elliptic
 191 regularity theory (see Theorem 3 in Jochmann [30], Lemma 3.2 in Bonito et al. [7] or
 192 Bernardi and Verfürth [6]). In general, one expects that $r \leq \frac{1}{2}$ whenever u is supported
 193 on at least two contiguous subdomains where λ takes different values; otherwise the
 194 normal derivative of u would be continuous across the interface separating the two
 195 subdomains in question, and owing to the discontinuity of λ , the normal component
 196 of the diffusive flux $\sigma(u)$ would be discontinuous across the interface, which would
 197 contradict the fact that $\sigma(u)$ has a weak divergence. It is however possible that $r > \frac{1}{2}$
 198 when the exact solution is supported on one subdomain only. If $r \geq 1$, we notice that
 199 one necessarily has $f \in L^2(D)$ (since $f|_{D_i} = \lambda|_{D_i}(\Delta u)_{D_i}$ for all $i \in \{1:M\}$), i.e., it is
 200 legitimate to assume that $q = 2$ if $r \geq 1$.

201 *Remark 2.2 (Extensions).* One could also consider lower-order terms in (2.1),
 202 e.g., $-\nabla \cdot (\lambda \nabla u) + \beta \cdot \nabla u + \mu u = f$ with $\beta \in \mathbf{W}^{1,\infty}(D)$ and $\mu \in L^\infty(D)$ s.t. $\mu - \frac{1}{2} \nabla \cdot \beta \geq 0$
 203 a.e. in D (for simplicity). The error analysis presented in this paper still applies pro-
 204 vided the lower-order terms are not too large, e.g., $\lambda \geq \max(h \|\beta\|_{\mathbf{L}^\infty(\Omega)}, h^2 \|\mu\|_{L^\infty(D)})$,
 205 where h denotes the mesh-size. Standard stabilization techniques have to be invoked
 206 if the lower-order terms are large when compared to the second-order diffusion op-
 207 erator. Furthermore, the error analysis can be extended to account for a piecewise
 208 constant tensor-valued diffusivity \mathfrak{d} ; then, the various constants in the error estimate
 209 depend on the square-root of the anisotropy ratios measuring the contrast between
 210 the largest and the smallest eigenvalue of \mathfrak{d} in each subdomain D_i . Finally, one can
 211 consider that the diffusion tensor \mathfrak{d} is piecewise smooth instead of being piecewise
 212 constant; a reasonable requirement is that $\mathfrak{d}|_{D_i}$ is Lipschitz for all $i \in \{1:M\}$. This
 213 last extension is, however, less straightforward because the discrete diffusive flux is
 214 no longer a piecewise polynomial function. \square

215 **2.2. Discrete setting.** We introduce in this section the discrete setting that we
 216 are going to use to approximate the solution to (2.2). Let \mathcal{T}_h be a mesh from a shape-
 217 regular sequence $(\mathcal{T}_h)_{h \in \mathcal{H}}$. Here, \mathcal{H} is a countable set with 0 as unique accumulation
 218 point. A generic mesh cell is denoted $K \in \mathcal{T}_h$ and is conventionally taken to be an
 219 open set. We also assume that \mathcal{T}_h covers each of the subdomains $\{D_i\}_{i \in \{1:M\}}$ exactly
 220 so that $\lambda_K := \lambda|_K$ is constant for all $K \in \mathcal{T}_h$. Let $(\widehat{K}, \widehat{P}, \widehat{\Sigma})$ be the reference finite
 221 element; we assume that $\mathbb{P}_{k,d} \subset \widehat{P} \subset W^{k+1,\infty}(\widehat{K})$ for some $k \geq 1$. Here, $\mathbb{P}_{k,d}$ is the
 222 (real) vector space composed of the d -variate polynomials of degree at most k . For all
 223 $K \in \mathcal{T}_h$, let $\mathbf{T}_K : \widehat{K} \rightarrow K$ be the geometric mapping and let $\psi_K^{\mathfrak{g}}(v) = v \circ \mathbf{T}_K$ be the
 224 pullback by the geometric mapping. We introduce the broken finite element space

225 (2.6)
$$P_h^{\mathfrak{b}}(\mathcal{T}_h) = \{v_h \in L^\infty(D) \mid v_h|_K \in P_K, \forall K \in \mathcal{T}_h\},$$

226 where $P_K := (\psi_K^g)^{-1}(\widehat{P}) \subset W^{k+1,\infty}(K)$. For any function $v_h \in P_k^b(\mathcal{T}_h)$, we define
 227 the broken diffusive flux $\boldsymbol{\sigma}(v_h) \in \mathbf{L}^2(D)$ by setting $\boldsymbol{\sigma}(v_h)|_K := -\lambda_K \nabla(v_h|_K)$ for all
 228 $K \in \mathcal{T}_h$. Upon introducing the notion of broken gradient $\nabla_h : W^{1,p}(\mathcal{T}_h) := \{v \in$
 229 $L^p(D) \mid \nabla(v|_K) \in L^p(K), \forall K \in \mathcal{T}_h\}$ by setting $(\nabla_h v)|_K := \nabla(v|_K)$ for all $K \in \mathcal{T}_h$
 230 and all $v \in W^{1,p}(\mathcal{T}_h)$, we have $\boldsymbol{\sigma}(v_h) = -\lambda \nabla_h v_h$.

231 For any cell $K \in \mathcal{T}_h$ we denote by \mathbf{n}_K the unit normal vector on ∂K pointing
 232 outward. We denote by \mathcal{F}_h° the collection of the mesh interfaces and \mathcal{F}_h^∂ the collection
 233 of the mesh faces at the boundary of D . We assume that \mathcal{T}_h is oriented in a generation-
 234 compatible way, and for each mesh face $F \in \mathcal{F}_h^\circ \cup \mathcal{F}_h^\partial$, we denote by \mathbf{n}_F the unit
 235 vector orienting F . For all $F \in \mathcal{F}_h^\circ$, we denote by $K_l, K_r \in \mathcal{T}_h$ the two cells s.t.
 236 $F = \partial K_l \cap \partial K_r$ and the unit normal vector \mathbf{n}_F orienting F points from K_l to K_r ,
 237 i.e., $\mathbf{n}_F = \mathbf{n}_{K_l} = -\mathbf{n}_{K_r}$. For all $F \in \mathcal{F}_h$, let \mathcal{T}_F be the collection of the one or two
 238 mesh cells sharing F . For all $K \in \mathcal{T}_h$, let \mathcal{F}_K be the collection of the faces of K and
 239 let $\epsilon_{K,F} = \mathbf{n}_F \cdot \mathbf{n}_K = \pm 1$. The jump across $F \in \mathcal{F}_h^\circ$ of any function $v \in W^{1,1}(\mathcal{T}_h)$ is
 240 defined by setting $\llbracket v \rrbracket_F(\mathbf{x}) = v|_{K_l}(\mathbf{x}) - v|_{K_r}(\mathbf{x})$ for a.e. $\mathbf{x} \in F$. If $F \in \mathcal{F}_h^\partial$, this jump is
 241 conventionally defined as the trace on F , i.e., $\llbracket v \rrbracket_F(\mathbf{x}) = v|_{K_l}(\mathbf{x})$ where $F = \partial K_l \cap \partial D$.
 242 We omit the subscript F in the jump whenever the context is unambiguous.

243 **3. The bilinear form n_\sharp .** In this section, we give a proper meaning to the
 244 normal trace of the diffusive flux of the solution to (2.2) over each mesh face. The
 245 material presented in §3.1 and §3.2 has been introduced in [25, §5.3] and is inspired
 246 from Amrouche et al. [1, Lem. 4.7], Bernardi and Hecht [5, Cor 3.3], and Buffa and
 247 Perugia [9, Lem. 8.2]; it is included here for the sake of completeness. The reader
 248 familiar with these techniques is invited to jump to §3.3 where the weighted bilinear
 249 form n_\sharp is introduced. This bilinear form is the main tool for the error analysis in §4.

250 **3.1. Face-to-cell lifting operator.** Let us first motivate our approach infor-
 251 mally. Let $K \in \mathcal{T}_h$ be a mesh cell, let \mathcal{F}_K be the collection of all the faces of K ,
 252 and let $F \in \mathcal{F}_K$ be a face of K . Let \mathbf{v} be a vector field defined on K . We are
 253 looking for (mild) regularity requirements on the field \mathbf{v} to give a meaning to the
 254 quantity $\int_F (\mathbf{v} \cdot \mathbf{n}_K) \phi \, ds$, where ϕ is a given smooth function on F (e.g., a poly-
 255 nomial function). It is well established that it is possible to give a weak meaning in
 256 $H^{-\frac{1}{2}}(\partial K)$ to the normal trace of \mathbf{v} on ∂K by means of an integration by parts for-
 257 mula if $\mathbf{v} \in \mathbf{H}(\text{div}; K) := \{\mathbf{v} \in \mathbf{L}^2(K) \mid \nabla \cdot \mathbf{v} \in L^2(K)\}$. In this situation, one can
 258 define the normal trace $\gamma_{\partial K}^d(\mathbf{v}) \in H^{-\frac{1}{2}}(\partial K)$ by setting

$$259 \quad (3.1) \quad \langle \gamma_{\partial K}^d(\mathbf{v}), \psi \rangle_{\partial K} := \int_K \left(\mathbf{v} \cdot \nabla w(\psi) + (\nabla \cdot \mathbf{v}) w(\psi) \right) dx,$$

260 for all $\psi \in H^{\frac{1}{2}}(\partial K)$, where $w(\psi) \in H^1(K)$ is a lifting of ψ , i.e., $\gamma_{\partial K}^g(w(\psi)) = \psi$, and
 261 $\gamma_{\partial K}^g : H^1(K) \rightarrow H^{\frac{1}{2}}(\partial K)$ is the trace map locally in K . Then, one has $\gamma_{\partial K}^d(\mathbf{v}) =$
 262 $\mathbf{v}|_{\partial K} \cdot \mathbf{n}_K$ whenever \mathbf{v} is smooth, e.g., if $\mathbf{v} \in \mathbf{H}(\text{div}; K) \cap \mathbf{C}^0(\overline{K})$. However, the above
 263 meaning is too weak for our purpose because we need to localize the action of the
 264 normal trace to functions ϕ only defined on a face F , i.e., ϕ may not be defined over
 265 the whole boundary ∂K . The key to achieve this is to extend ϕ by zero from F to
 266 ∂K . This obliges us to change the functional setting since the extended function is
 267 no longer in $H^{\frac{1}{2}}(\partial K)$. In what follows, we are going to use the fact that the zero-
 268 extension of a smooth function defined on a face F of ∂K is in $W^{1-\frac{1}{t},t}(\partial K)$ if $t < 2$,
 269 i.e., $t(1 - \frac{1}{t}) < 1$. Let us now present a rigorous construction.

270 Let p, q be two real numbers such that

$$271 \quad (3.2) \quad p > 2, \quad q > \frac{2d}{2+d}.$$

272 Notice that $q > 1$ since $d \geq 2$. Let $\tilde{p} \in (2, p]$ be such that $q \geq \frac{\tilde{p}d}{\tilde{p}+d}$; this is indeed
 273 possible since $p > 2$, $q > \frac{2d}{2+d}$, and the function $z \mapsto \frac{zd}{z+d}$ is increasing over \mathbb{R}_+ .
 274 Lemma 3.1 shows that there exists a bounded lifting operator

$$275 \quad (3.3) \quad L_F^K : W^{\frac{1}{\tilde{p}}, \tilde{p}'}(F) \longrightarrow W^{1, \tilde{p}'}(K),$$

276 with conjugate number \tilde{p}' s.t. $\frac{1}{\tilde{p}} + \frac{1}{\tilde{p}'} = 1$, so that for any $\phi \in W^{\frac{1}{\tilde{p}}, \tilde{p}'}(F)$, $L_F^K(\phi)$ is a
 277 lifting of the zero-extension of ϕ to ∂K , i.e.,

$$278 \quad (3.4) \quad \gamma_{\partial K}^g(L_F^K(\phi))|_{\partial K \setminus F} = 0, \quad \gamma_{\partial K}^g(L_F^K(\phi))|_F = \phi.$$

279 Notice that the domain of L_F^K is of the form $W^{1-\frac{1}{t}, t}(F)$ with $t = \tilde{p}' < 2$, which is
 280 consistent with the above observation regarding the zero-extension to ∂K of functions
 281 defined on F . We also observe that

$$282 \quad (3.5) \quad L_F^K(\phi) \in W^{1, p'}(K) \cap L^{q'}(K),$$

283 with conjugate numbers p', q' s.t. $\frac{1}{p} + \frac{1}{p'} = 1$, $\frac{1}{q} + \frac{1}{q'} = 1$. Indeed, $L_F^K(\phi) \in W^{1, p'}(K)$
 284 just follows from $p' \leq \tilde{p}'$ (i.e., $\tilde{p} \leq p$), whereas $L_F^K(\phi) \in L^{q'}(K)$ follows from
 285 $W^{1, \tilde{p}'}(K) \hookrightarrow L^{q'}(K)$ owing to the Sobolev Embedding Theorem (since $q' \leq \frac{\tilde{p}'d}{d-\tilde{p}'}$,
 286 as can be verified from $d \geq 2 > \tilde{p}'$ and $\frac{1}{\tilde{p}'} - \frac{1}{d} = 1 - (\frac{1}{\tilde{p}} + \frac{1}{d}) \leq 1 - \frac{1}{q} = \frac{1}{q'}$ because
 287 $q \geq \frac{\tilde{p}d}{\tilde{p}+d}$). We now state our main result on the lifting operator L_F^K .

288 LEMMA 3.1 (Face-to-cell lifting). *Let p and q satisfy (3.2). Let $\tilde{p} \in (2, p]$ be such
 289 that $q \geq \frac{\tilde{p}d}{\tilde{p}+d}$. Let $K \in \mathcal{T}_h$ be a mesh cell and let $F \in \mathcal{F}_K$ be a face of K . There
 290 exists a lifting operator $L_F^K : W^{\frac{1}{\tilde{p}}, \tilde{p}'}(F) \rightarrow W^{1, \tilde{p}'}(K)$ satisfying (3.4), and there exists
 291 c , uniform w.r.t. $h \in \mathcal{H}$, but depending on p and q , s.t. the following holds true:*

$$292 \quad (3.6) \quad h_K^{\frac{d}{p}} |L_F^K(\phi)|_{W^{1, p'}(K)} + h_K^{-1+\frac{d}{q}} \|L_F^K(\phi)\|_{L^{q'}(K)} \leq c h_K^{-\frac{1}{\tilde{p}}+\frac{d}{q}} \|\phi\|_{W^{\frac{1}{\tilde{p}}, \tilde{p}'}(F)},$$

293 for all $\phi \in W^{\frac{1}{\tilde{p}}, \tilde{p}'}(F)$ with the norm $\|\phi\|_{W^{\frac{1}{\tilde{p}}, \tilde{p}'}(F)} := \|\phi\|_{L^{\tilde{p}'}(F)} + h_F^{\frac{1}{\tilde{p}}} |\phi|_{W^{\frac{1}{\tilde{p}}, \tilde{p}'}(F)}$. \square

294 *Proof.* (1) The face-to-cell lifting operator L_F^K is constructed from a lifting op-
 295 erator $L_{\hat{F}}^{\hat{K}}$ on the reference cell. Let \hat{K} be the reference cell and let \hat{F} be one of
 296 its faces. Let us define the operator $L_{\hat{F}}^{\hat{K}} : W^{\frac{1}{\tilde{p}}, \tilde{p}'}(\hat{F}) \rightarrow W^{1, \tilde{p}'}(\hat{K})$. For any func-
 297 tion $\psi \in W^{\frac{1}{\tilde{p}}, \tilde{p}'}(\hat{F})$, let $\tilde{\psi}$ denote the zero-extension of ψ to $\partial \hat{K}$. Owing to Gris-
 298 vard [28, Thm. 1.4.2.4, Cor. 1.4.4.5], $\tilde{\psi}$ is in $W^{\frac{1}{\tilde{p}}, \tilde{p}'}(\partial \hat{K})$ since $\frac{\tilde{p}'}{\tilde{p}} = \frac{1}{\tilde{p}-1} < 1$ (i.e.,
 299 $\tilde{p} > 2$), and we have $\|\tilde{\psi}\|_{W^{\frac{1}{\tilde{p}}, \tilde{p}'}(\partial \hat{K})} \leq \hat{c}_1 \|\psi\|_{W^{\frac{1}{\tilde{p}}, \tilde{p}'}(\hat{F})}$ with the norm $\|\psi\|_{W^{\frac{1}{\tilde{p}}, \tilde{p}'}(\hat{F})} :=$
 300 $\|\psi\|_{L^{\tilde{p}'}(\hat{F})} + \ell_{\hat{K}}^{\frac{1}{\tilde{p}}} |\psi|_{W^{\frac{1}{\tilde{p}}, \tilde{p}'}(\hat{F})}$ where $\ell_{\hat{K}} = 1$ is a length scale associated with \hat{K} . Then we
 301 use the surjectivity of the trace map $\gamma_{\hat{K}}^g : W^{1, \tilde{p}'}(\hat{K}) \rightarrow W^{\frac{1}{\tilde{p}}, \tilde{p}'}(\partial \hat{K})$ (see Gagliardo [27,
 302 Thm. 1.I]) to define $L_{\hat{F}}^{\hat{K}}(\psi) \in W^{1, \tilde{p}'}(\hat{K})$ s.t. $\gamma_{\hat{K}}^g(L_{\hat{F}}^{\hat{K}}(\psi)) = \tilde{\psi}$ and $\|L_{\hat{F}}^{\hat{K}}(\psi)\|_{W^{1, \tilde{p}'}(\hat{K})} \leq$

303 $\widehat{c}_2 \|\widetilde{\psi}\|_{W^{\frac{1}{\widehat{p}}, \widehat{p}' }(\partial \widehat{K})}$, i.e., $\|L_{\widehat{F}}^{\widehat{K}}(\psi)\|_{W^{1, \widehat{p}' }(\widehat{K})} \leq \widehat{c} \|\psi\|_{W^{\frac{1}{\widehat{p}}, \widehat{p}' }(\widehat{F})}$, with $\widehat{c} = \widehat{c}_1 \widehat{c}_2$. By construc-
 304 tion, we have $\gamma_{\partial \widehat{K}}^g(L_{\widehat{F}}^{\widehat{K}}(\psi))|_{\widehat{F}} = \psi$ and $\gamma_{\partial \widehat{K}}^g(L_{\widehat{F}}^{\widehat{K}}(\psi))|_{\partial \widehat{K} \setminus \widehat{F}} = 0$.
 305 (2) We define the lifting operator $L_F^K : W^{\frac{1}{\widehat{p}}, \widehat{p}' }(\widehat{F}) \rightarrow W^{1, \widehat{p}' }(\widehat{K})$ by setting

$$306 \quad (3.7) \quad L_F^K(\phi)(\mathbf{x}) := L_{\widehat{F}}^{\widehat{K}}(\phi \circ \mathbf{T}_{K|\widehat{F}})(\mathbf{T}_K^{-1}(\mathbf{x})), \quad \forall \mathbf{x} \in K, \quad \forall \phi \in W^{\frac{1}{\widehat{p}}, \widehat{p}' }(\widehat{F}),$$

307 where $\mathbf{T}_K : \widehat{K} \rightarrow K$ is the geometric mapping and $\widehat{F} = \mathbf{T}_K^{-1}(F)$. By definition, if
 308 $\mathbf{x} \in F$, then $\widehat{\mathbf{x}} := \mathbf{T}_K^{-1}(\mathbf{x}) \in \widehat{F}$ and $\mathbf{T}_{K|\widehat{F}}(\widehat{\mathbf{x}}) = \mathbf{x}$, so that

$$309 \quad \gamma_{\partial K}^g(L_F^K(\phi))(\mathbf{x}) = \gamma_{\partial \widehat{K}}^g(L_{\widehat{F}}^{\widehat{K}}(\phi \circ \mathbf{T}_{K|\widehat{F}}))(\widehat{\mathbf{x}}) = \phi(\mathbf{T}_{K|\widehat{F}}(\widehat{\mathbf{x}})) = \phi(\mathbf{x}),$$

310 whereas if $\mathbf{x} \in \partial K \setminus F$, then $\widehat{\mathbf{x}} \in \partial \widehat{K} \setminus \widehat{F}$, so that $\gamma_{\partial \widehat{K}}^g(L_{\widehat{F}}^{\widehat{K}}(\phi \circ \mathbf{T}_{K|\widehat{F}}))(\widehat{\mathbf{x}}) = 0$. The
 311 above argument shows that (3.4) holds true.

312 (3) It remains to prove (3.6). Let us first bound $|L_F^K(\phi)|_{W^{1, p'}(K)}$. Notice that
 313 the definition of L_F^K is equivalent to $L_F^K(\phi) \circ \mathbf{T}_K(\widehat{\mathbf{x}}) := L_{\widehat{F}}^{\widehat{K}}(\phi \circ \mathbf{T}_{K|\widehat{F}})(\widehat{\mathbf{x}})$; that is,
 314 $\psi_K^g(L_F^K(\phi)) := L_{\widehat{F}}^{\widehat{K}}(\psi_F^g(\phi))$, where ψ_K^g is the pullback by \mathbf{T}_K , and ψ_F^g is the pullback
 315 by $\mathbf{T}_{K|\widehat{F}}$. Denoting by \mathbb{J}_K the Jacobian of the geometric mapping \mathbf{T}_K , we infer that

$$316 \quad |L_F^K(\phi)|_{W^{1, p'}(K)} \leq c \|\mathbb{J}_K^{-1}\|_{\ell^2} |\det(\mathbb{J}_K)|^{\frac{1}{p'}} |L_{\widehat{F}}^{\widehat{K}}(\psi_F^g(\phi))|_{W^{1, p'}(\widehat{K})}$$

$$317 \quad \leq c' \|\mathbb{J}_K^{-1}\|_{\ell^2} |\det(\mathbb{J}_K)|^{\frac{1}{p'}} |L_{\widehat{F}}^{\widehat{K}}(\psi_F^g(\phi))|_{W^{1, \widehat{p}' }(\widehat{K})}$$

$$318 \quad \leq c'' \|\mathbb{J}_K^{-1}\|_{\ell^2} |\det(\mathbb{J}_K)|^{\frac{1}{p'}} \|\psi_F^g(\phi)\|_{W^{\frac{1}{\widehat{p}}, \widehat{p}' }(\widehat{F})},$$

320 where the first inequality follows from the chain rule, the second is a consequence of
 321 $\widehat{p}' \geq p'$ (since $\widehat{p} \leq p$), and the third follows from the stability of the reference lifting
 322 operator $L_{\widehat{F}}^{\widehat{K}}$. Using now the chain rule and the shape-regularity of the mesh sequence,
 323 we infer that $\|\psi_F^g(\phi)\|_{W^{\frac{1}{\widehat{p}}, \widehat{p}' }(\widehat{F})} \leq c |\det(\mathbb{J}_F)|^{-\frac{1}{\widehat{p}'}} \|\phi\|_{W^{\frac{1}{\widehat{p}}, \widehat{p}' }(\widehat{F})}$, where \mathbb{J}_F is the Jacobian
 324 of the mapping $\mathbf{T}_{K|\widehat{F}} : \widehat{F} \rightarrow F$. Combining these bounds, we obtain

$$325 \quad |L_F^K(\phi)|_{W^{1, p'}(K)} \leq c \|\mathbb{J}_K^{-1}\|_{\ell^2} |\det(\mathbb{J}_K)|^{\frac{1}{p'}} |\det(\mathbb{J}_F)|^{-\frac{1}{\widehat{p}'}} \|\phi\|_{W^{\frac{1}{\widehat{p}}, \widehat{p}' }(\widehat{F})}$$

$$326 \quad \leq c' h_K^{-\frac{1}{\widehat{p}} + d(\frac{1}{\widehat{p}} - \frac{1}{p'})} \|\phi\|_{W^{\frac{1}{\widehat{p}}, \widehat{p}' }(\widehat{F})},$$

328 where the second bound follows from the shape-regularity of the mesh sequence.
 329 This proves the bound on $|L_F^K(\phi)|_{W^{1, p'}(K)}$ in (3.6). The proof of the bound on
 330 $\|L_F^K(\phi)\|_{L^{q'}(K)}$ uses similar arguments together with $W^{1, \widehat{p}' }(\widehat{K}) \hookrightarrow L^{q'}(\widehat{K})$ owing to
 331 the Sobolev Embedding Theorem and $q' \leq \frac{\widehat{p}' d}{d - \widehat{p}'}$ (as already shown above). \square

332 **3.2. Face localization of the normal diffusive flux.** Let $K \in \mathcal{T}_h$ be a mesh
 333 cell, $F \in \mathcal{F}_K$ be a face of K , and consider the following functional space:

$$334 \quad (3.8) \quad \mathbf{S}^d(K) := \{\boldsymbol{\tau} \in \mathbf{L}^p(K) \mid \nabla \cdot \boldsymbol{\tau} \in L^q(K)\},$$

335 equipped with the following dimensionally-consistent norm:

$$336 \quad (3.9) \quad \|\boldsymbol{\tau}\|_{\mathbf{S}^d(K)} := \|\boldsymbol{\tau}\|_{\mathbf{L}^p(K)} + h_K^{1 + d(\frac{1}{p} - \frac{1}{q})} \|\nabla \cdot \boldsymbol{\tau}\|_{L^q(K)}.$$

337 With the lifting operator L_F^K in hand, we now define the normal trace on the face F of
 338 K of any field $\boldsymbol{\tau} \in \mathbf{S}^d(K)$ to be the linear form in $(W^{\frac{1}{p}, \tilde{p}'}(F))'$ denoted by $(\boldsymbol{\tau} \cdot \mathbf{n}_K)|_F$
 339 and whose action on any function $\phi \in W^{\frac{1}{p}, \tilde{p}'}(F)$ is

$$340 \quad (3.10) \quad \langle (\boldsymbol{\tau} \cdot \mathbf{n}_K)|_F, \phi \rangle_F := \int_K \left(\boldsymbol{\tau} \cdot \nabla L_F^K(\phi) + (\nabla \cdot \boldsymbol{\tau}) L_F^K(\phi) \right) dx.$$

341 Here, $\langle \cdot, \cdot \rangle_F$ denotes the duality pairing between $(W^{\frac{1}{p}, \tilde{p}'}(F))'$ and $W^{\frac{1}{p}, \tilde{p}'}(F)$. Notice
 342 that the right-hand side of (3.10) is well-defined owing to Hölder's inequality and (3.6).
 343 Owing to (3.4), we readily verify that we have indeed defined an extension of the
 344 normal trace since we have $\langle (\boldsymbol{\tau} \cdot \mathbf{n}_K)|_F, \phi \rangle_F = \int_F (\boldsymbol{\tau} \cdot \mathbf{n}_K) \phi ds$ whenever the field $\boldsymbol{\tau}$ is
 345 smooth. Let us now derive an important bound on the linear form $(\boldsymbol{\tau} \cdot \mathbf{n}_K)|_F$ when
 346 acting on a function from the space P_F , which we define to be composed of the
 347 restrictions to F of the functions in P_K . Note that $P_F \subset W^{\frac{1}{p}, \tilde{p}'}(F)$.

348 **LEMMA 3.2** (Bound on normal component). *There exists a constant c , uniform*
 349 *w.r.t. $h \in \mathcal{H}$, but depending on p and q , s.t. the following holds true:*

$$350 \quad (3.11) \quad |\langle (\boldsymbol{\tau} \cdot \mathbf{n}_K)|_F, \phi_h \rangle_F| \leq c h_K^{d(\frac{1}{2} - \frac{1}{p})} \|\boldsymbol{\tau}\|_{\mathbf{S}^d(K)} h_F^{-\frac{1}{2}} \|\phi_h\|_{L^2(F)},$$

352 *for all $\boldsymbol{\tau} \in \mathbf{S}^d(K)$, all $\phi_h \in P_F$, all $K \in \mathcal{T}_h$, and all $F \in \mathcal{F}_K$. \square*

353 *Proof.* A direct consequence of (3.10), Hölder's inequality, and Lemma 3.1 is that

$$354 \quad |\langle (\boldsymbol{\tau} \cdot \mathbf{n}_K)|_F, \phi \rangle_F| \leq c h_K^{-\frac{1}{p} + d(\frac{1}{p} - \frac{1}{p})} \|\boldsymbol{\tau}\|_{\mathbf{S}^d(K)} \|\phi\|_{W^{\frac{1}{p}, \tilde{p}'}(F)},$$

355 for all $\phi \in W^{\frac{1}{p}, \tilde{p}'}(F)$. Recalling that $\|\phi\|_{W^{\frac{1}{p}, \tilde{p}'}(F)} = \|\phi\|_{L^{\tilde{p}'}(F)} + h_F^{\frac{1}{p}} |\phi|_{W^{\frac{1}{p}, \tilde{p}'}(F)}$, the
 356 shape-regularity of the mesh sequence implies that the following inverse inequality
 357 $\|\phi_h\|_{W^{\frac{1}{p}, \tilde{p}'}(F)} \leq c h_F^{(d-1)(\frac{1}{2} - \frac{1}{p})} \|\phi_h\|_{L^2(F)}$ holds true for all $\phi_h \in P_F$ (note that $\frac{1}{2} - \frac{1}{p} =$
 358 $\frac{1}{\tilde{p}'} - \frac{1}{2}$). The estimate (3.11) follows readily. \square

359 **3.3. Definition of n_{\sharp} and key identities.** Let us consider the functional space
 360 V_s defined in (2.5). For all $v \in V_s$, Lemma 2.1 shows that $\boldsymbol{\sigma}(v)|_K \in \mathbf{S}^d(K)$ for all
 361 $K \in \mathcal{T}_h$, and Lemma 3.2 implies that it is possible to give a meaning by duality
 362 to the normal component of $\boldsymbol{\sigma}(v)|_K$ on all the faces of K separately. Moreover,
 363 since we have set $\boldsymbol{\sigma}(v_h)|_K := -\lambda_K \nabla(v_h|_K)$ for all $v_h \in P_k^b(\mathcal{T}_h)$, and since we have
 364 $P_K \subset W^{k+1, \infty}(K)$ with $k \geq 1$, we infer that $\boldsymbol{\sigma}(v_h)|_K \in \mathbf{S}^d(K)$ as well. Thus,
 365 $\boldsymbol{\sigma}(v)|_K \in \mathbf{S}^d(K)$ for all $v \in (V_s + P_k^b(\mathcal{T}_h))$. Let us now introduce the bilinear form
 366 $n_{\sharp} : (V_s + P_k^b(\mathcal{T}_h)) \times P_k^b(\mathcal{T}_h) \rightarrow \mathbb{R}$ defined as follows:

$$367 \quad (3.12) \quad n_{\sharp}(v, w_h) := \sum_{F \in \mathcal{F}_h} \sum_{K \in \mathcal{T}_F} \epsilon_{K,F} \theta_{K,F} \langle (\boldsymbol{\sigma}(v)|_K \cdot \mathbf{n}_K)|_F, \llbracket w_h \rrbracket \rangle_F,$$

369 where the weights $\theta_{K,F}$ are still unspecified but are assumed to satisfy

$$370 \quad (3.13) \quad \theta_{K_l, F}, \theta_{K_r, F} \in [0, 1] \quad \text{and} \quad \theta_{K_l, F} + \theta_{K_r, F} = 1, \quad \forall F \in \mathcal{F}_h,$$

371 whereas for all $F \in \mathcal{F}_h^{\partial}$ with $F = \partial K_l \cap \partial D$, we set $\theta_{K_l, F} := 1$, $\theta_{K_r, F} := 0$. We
 372 will see in (3.19) below how these weights must depend on the diffusion coefficient to
 373 get a robust boundedness estimate on n_{\sharp} . The definition (3.12) is meaningful since

374 $\llbracket w_h \rrbracket_F \in P_F$ for all $w_h \in P_k^b(\mathcal{T}_h)$. The factor $\epsilon_{K,F}$ in (3.12) handles the relative
 375 orientation of \mathbf{n}_K and \mathbf{n}_F . For all $v \in W^{1,1}(\mathcal{T}_h)$, we define weighted averages as
 376 follows for a.e. $\mathbf{x} \in F \in \mathcal{F}_h^o$:

$$377 \quad (3.14a) \quad \{v\}_{F,\theta}(\mathbf{x}) := \theta_{K_l,F} v|_{K_l}(\mathbf{x}) + \theta_{K_r,F} v|_{K_r}(\mathbf{x}),$$

$$378 \quad (3.14b) \quad \{v\}_{F,\bar{\theta}}(\mathbf{x}) := \theta_{K_r,F} v|_{K_l}(\mathbf{x}) + \theta_{K_l,F} v|_{K_r}(\mathbf{x}).$$

380 Whenever $\theta_{K_l,F} = \theta_{K_r,F} = \frac{1}{2}$, these two definitions coincide with the usual arithmetic
 381 average. On boundary faces $F \in \mathcal{F}_h^\partial$, we have $\{v\}_{F,\theta}(\mathbf{x}) = v|_{K_l}(\mathbf{x})$, and $\{v\}_{F,\bar{\theta}}(\mathbf{x}) = 0$
 382 for a.e. $\mathbf{x} \in F$. We omit the subscript F whenever the context is unambiguous. The
 383 following identity will be useful:

$$384 \quad (3.15) \quad \llbracket vw \rrbracket = \{v\}_\theta \llbracket w \rrbracket + \llbracket v \rrbracket \{w\}_{\bar{\theta}}.$$

385 The following lemma is fundamental to understand the role that the bilinear form
 386 n_\sharp will play in the next section in the analysis of various nonconforming approximation
 387 methods.

388 **LEMMA 3.3** (Identities for n_\sharp). *The following holds true for any choice of weights*
 389 $\{\theta_{K,F}\}_{F \in \mathcal{F}_h, K \in \mathcal{T}_F}$ and for all $w_h \in P_k^b(\mathcal{T}_h)$, all $v_h \in P_k^b(\mathcal{T}_h)$, and all $v \in V_s$:

$$390 \quad (3.16a) \quad n_\sharp(v_h, w_h) = \sum_{F \in \mathcal{F}_h} \int_F \{\boldsymbol{\sigma}(v_h)\}_\theta \cdot \mathbf{n}_F \llbracket w_h \rrbracket \, ds,$$

$$391 \quad (3.16b) \quad n_\sharp(v, w_h) = \sum_{K \in \mathcal{T}_h} \int_K \left(\boldsymbol{\sigma}(v) \cdot \nabla w_h|_K + (\nabla \cdot \boldsymbol{\sigma}(v)) w_h|_K \right) dx. \quad \square$$

393 *Proof.* (1) Proof of (3.16a). Let $v_h, w_h \in P_k^b(\mathcal{T}_h)$. Since the restriction of $\boldsymbol{\sigma}(v_h)$
 394 to each mesh cell is smooth, and since the restriction of $L_F^K(\llbracket w_h \rrbracket)$ to ∂K is nonzero
 395 only on the face $F \in \mathcal{F}_K$ where it coincides with $\llbracket w_h \rrbracket$, we have

$$396 \quad \langle (\boldsymbol{\sigma}(v_h)|_K \cdot \mathbf{n}_K)|_F, \llbracket w_h \rrbracket \rangle_F = \int_K \left(\boldsymbol{\sigma}(v_h)|_K \cdot \nabla L_F^K(\llbracket w_h \rrbracket) + (\nabla \cdot \boldsymbol{\sigma}(v_h)|_K) L_F^K(\llbracket w_h \rrbracket) \right) dx$$

$$397 \quad = \int_{\partial K} \boldsymbol{\sigma}(v_h)|_K \cdot \mathbf{n}_K L_F^K(\llbracket w_h \rrbracket) \, ds = \int_F \boldsymbol{\sigma}(v_h)|_K \cdot \mathbf{n}_K \llbracket w_h \rrbracket \, ds,$$

399 where we used the divergence formula in K . Therefore, after using the definitions of
 400 $\epsilon_{K,F}$ and of $\theta_{K,F}$, we obtain

$$401 \quad n_\sharp(v_h, w_h) = \sum_{F \in \mathcal{F}_h} \sum_{K \in \mathcal{T}_F} \epsilon_{K,F} \theta_{K,F} \int_F \boldsymbol{\sigma}(v_h)|_K \cdot \mathbf{n}_K \llbracket w_h \rrbracket \, ds$$

$$402 \quad = \sum_{F \in \mathcal{F}_h} \int_F \{\boldsymbol{\sigma}(v_h)\}_\theta \cdot \mathbf{n}_F \llbracket w_h \rrbracket \, ds.$$

404 (2) Proof of (3.16b). Let $v \in V_s$ and $w_h \in P_k^b(\mathcal{T}_h)$. Let $\mathcal{K}_\delta^d : L^1(D) \rightarrow C^\infty(\bar{D})$ and
 405 $\mathcal{K}_\delta^b : L^1(D) \rightarrow C^\infty(\bar{D})$ be the mollification operators introduced in [22, §3.2]. These
 406 two operators satisfy the following key commuting property:

$$407 \quad (3.17) \quad \nabla \cdot (\mathcal{K}_\delta^d(\boldsymbol{\tau})) = \mathcal{K}_\delta^b(\nabla \cdot \boldsymbol{\tau}),$$

408 for all $\boldsymbol{\tau} \in L^1(D)$ s.t. $\nabla \cdot \boldsymbol{\tau} \in L^1(D)$. It is important to realize that this property can
 409 be applied to $\boldsymbol{\sigma}(v)$ for all $v \in V_s$ since $\nabla \cdot \boldsymbol{\sigma}(v) \in L^1(D)$ by definition of V_s . (Note

410 that this property cannot be applied to $\boldsymbol{\sigma}(v_h)$ with $v_h \in P_k^b(\mathcal{T}_h)$, since the normal
 411 component of $\boldsymbol{\sigma}(v_h)$ is in general discontinuous across the mesh interfaces, i.e., $\boldsymbol{\sigma}(v_h)$
 412 does not have a weak divergence.) Let us consider the mollified bilinear form

$$413 \quad n_{\sharp\delta}(v, w_h) := \sum_{F \in \mathcal{F}_h} \sum_{K \in \mathcal{T}_F} \epsilon_{K,F} \theta_{K,F} \langle (\mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)))|_K \cdot \mathbf{n}_K \rangle|_F, \llbracket w_h \rrbracket \rangle_F.$$

414 Owing to the commuting property (3.17), we infer that

$$415 \quad \langle (\mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)))|_K \cdot \mathbf{n}_K \rangle|_F, \llbracket w_h \rrbracket \rangle_F =$$

$$417 \quad \int_K \left(\mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \cdot L_F^K(\llbracket w_h \rrbracket) + \mathcal{K}_\delta^b(\nabla \cdot \boldsymbol{\sigma}(v)) L_F^K(\llbracket w_h \rrbracket) \right) dx.$$

419 Then Theorem 3.3 from [22] implies that

$$421 \quad \lim_{\delta \rightarrow 0} \int_K \left(\mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \cdot L_F^K(\llbracket w_h \rrbracket) + \mathcal{K}_\delta^b((\nabla \cdot \boldsymbol{\sigma}(v))) L_F^K(\llbracket w_h \rrbracket) \right) dx =$$

$$422 \quad \int_K \left(\boldsymbol{\sigma}(v) \cdot L_F^K(\llbracket w_h \rrbracket) + (\nabla \cdot \boldsymbol{\sigma}(v)) L_F^K(\llbracket w_h \rrbracket) \right) dx = \langle (\boldsymbol{\sigma}(v))|_K \cdot \mathbf{n}_K \rangle|_F, \llbracket w_h \rrbracket \rangle_F.$$

424 Summing over the mesh faces and the associated mesh cells, we infer that

$$425 \quad \lim_{\delta \rightarrow 0} n_{\sharp\delta}(v, w_h) = n_{\sharp}(v, w_h).$$

426 Moreover, since the mollified function $\mathcal{K}_\delta^d(\boldsymbol{\sigma}(v))$ is smooth, by repeating the calcula-
 427 tion done in Step (1), we also have

$$428 \quad n_{\sharp\delta}(v, w_h) = \sum_{F \in \mathcal{F}_h} \int_F \{ \mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \}_\theta \cdot \mathbf{n}_F \llbracket w_h \rrbracket ds.$$

430 Using the identity (3.15) with $\llbracket \mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \rrbracket \cdot \mathbf{n}_F = 0$ for all $F \in \mathcal{F}_h^\circ$, recalling that
 431 $\llbracket w_h \mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \rrbracket = w_h \mathcal{K}_\delta^d(\boldsymbol{\sigma}(v))|_F$ for all $F \in \mathcal{F}_h^\partial$, and using the divergence formula in
 432 K and the commuting property (3.17), we obtain

$$433 \quad n_{\sharp\delta}(v, w_h) = \sum_{F \in \mathcal{F}_h} \int_F \{ \mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \}_\theta \cdot \mathbf{n}_F \llbracket w_h \rrbracket ds + \sum_{F \in \mathcal{F}_h^\circ} \int_F \llbracket \mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \rrbracket \cdot \mathbf{n}_F \{ w_h \}_\theta ds$$

$$434 \quad = \sum_{F \in \mathcal{F}_h} \int_F \llbracket w_h \mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \rrbracket \cdot \mathbf{n}_F ds = \sum_{K \in \mathcal{T}_h} \int_{\partial K} \mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \cdot \mathbf{n}_K w_h|_K ds$$

$$435 \quad = \sum_{K \in \mathcal{T}_h} \int_K \left(\mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \cdot \nabla w_h|_K + \mathcal{K}_\delta^b(\nabla \cdot \boldsymbol{\sigma}(v)) w_h|_K \right) dx.$$

437 Invoking again Theorem 3.3 from [22] leads to the assertion since

$$438 \quad \lim_{\delta \rightarrow 0} n_{\sharp\delta}(v, w_h) = \sum_{K \in \mathcal{T}_h} \int_K \left(\boldsymbol{\sigma}(v) \cdot \nabla w_h|_K + (\nabla \cdot \boldsymbol{\sigma}(v)) w_h|_K \right) dx. \quad \square$$

439 *Remark 3.4* (Identity (3.16b)). The identity (3.16b) is the key tool to assert in
 440 a weak sense that $\boldsymbol{\sigma}(v) \cdot \mathbf{n}$ is continuous across the mesh interfaces without the need
 441 to assume that v is smooth, say $v \in H^{1+r}(D)$ with $r > \frac{1}{2}$. \square

442 We now establish an important boundedness estimate on the bilinear form n_\sharp .
 443 Since $\boldsymbol{\sigma}(v)|_K \in \mathbf{S}^d(K)$ for all $K \in \mathcal{T}_h$ and all $v \in V_s + P_k^b(\mathcal{T}_h)$, we can equip the space
 444 $V_s + P_k^b(\mathcal{T}_h)$ with the seminorm

$$445 \quad (3.18) \quad |v|_{n_\sharp}^2 := \sum_{K \in \mathcal{T}_h} \lambda_K^{-1} \left(h_K^{2d(\frac{1}{2} - \frac{1}{p})} \|\boldsymbol{\sigma}(v)|_K\|_{\mathbf{L}^p(K)}^2 + h_K^{2d(\frac{2+d}{2d} - \frac{1}{q})} \|\nabla \cdot \boldsymbol{\sigma}(v)|_K\|_{L^q(K)}^2 \right).$$

446 We notice that this seminorm is dimensionally-consistent with the classical energy-
 447 norm defined as $\sum_{K \in \mathcal{T}_h} \lambda_K \|\nabla v|_K\|_{\mathbf{L}^2(K)}^2$. Straightforward algebra shows that $|v|_{n_\sharp} \leq$
 448 $c \lambda_b^{-\frac{1}{2}} (\ell_D^{d(\frac{1}{2} - \frac{1}{p})} \|\boldsymbol{\sigma}(v)\|_{\mathbf{L}^p(D)} + \ell_D^{d(\frac{2+d}{2d} - \frac{1}{q})} \|\nabla \cdot \boldsymbol{\sigma}(v)\|_{L^q(D)})$, for all $v \in V_s$; here ℓ_D denotes
 449 a characteristic length of D . (Recall that $\|a\|_{\ell^s(\mathcal{I})} \leq \|a\|_{\ell^t(\mathcal{I})}$ for any finite sequence
 450 $(a_i)_{i \in \mathcal{I}}$ if $0 < t \leq s$, and we assumed that $q \leq 2$.)

451 In order to get robust error estimates with respect to λ , it is important to avoid
 452 any dependency on the ratio of the values taken by λ in two adjacent subdomains;
 453 otherwise, the error estimates become meaningless when the diffusion coefficient λ is
 454 highly contrasted. To avoid such dependencies, we introduce the following diffusion-
 455 dependent weights for all $F \in \mathcal{F}_h^\circ$, with $F = \partial K_l \cap \partial K_r$:

$$456 \quad (3.19) \quad \theta_{K_l, F} := \frac{\lambda_{K_r}}{\lambda_{K_l} + \lambda_{K_r}}, \quad \theta_{K_r, F} := \frac{\lambda_{K_l}}{\lambda_{K_l} + \lambda_{K_r}}.$$

457 We also define

$$458 \quad (3.20) \quad \lambda_F := \frac{2\lambda_{K_l}\lambda_{K_r}}{\lambda_{K_l} + \lambda_{K_r}} \text{ if } F \in \mathcal{F}_h^\circ \quad \text{and} \quad \lambda_F := \lambda_{K_l} \text{ if } F \in \mathcal{F}_h^\partial.$$

459 The two properties we are going to use are that $|\mathcal{T}_F| \lambda_K \theta_{K, F} = \lambda_F$, for all $K \in \mathcal{T}_F$,
 460 and $\lambda_F \leq \min_{K \in \mathcal{T}_F} \lambda_K$. (Here $|\mathcal{T}_F|$ denotes the cardinality of \mathcal{T}_F .)

461 LEMMA 3.5 (Boundedness of n_\sharp). *With the weights defined in (3.19) and λ_F*
 462 *defined in (3.20) for all $F \in \mathcal{F}_h$, there is c , uniform w.r.t. $h \in \mathcal{H}$ and λ , but depending*
 463 *on p and q , s.t. the following holds true for all $v \in V_s + P_k^b(\mathcal{T}_h)$ and all $w_h \in P_k^b(\mathcal{T}_h)$:*

$$464 \quad (3.21) \quad |n_\sharp(v, w_h)| \leq c |v|_{n_\sharp} \left(\sum_{F \in \mathcal{F}_h} \lambda_F h_F^{-1} \| [w_h] \|_{L^2(F)}^2 \right)^{\frac{1}{2}}. \quad \square$$

465 *Proof.* Let $v \in V_s + P_k^b(\mathcal{T}_h)$ and $w_h \in P_k^b(\mathcal{T}_h)$. Owing to the definition (3.12) of
 466 n_\sharp and the estimate (3.11) from Lemma 3.2, we infer that

$$467 \quad |n_\sharp(v, w_h)| \leq c \sum_{F \in \mathcal{F}_h} \sum_{K \in \mathcal{T}_F} \theta_{K, F} h_K^{d(\frac{1}{2} - \frac{1}{p})} \|\boldsymbol{\sigma}(v)|_K\|_{\mathbf{S}^d(K)} h_F^{-\frac{1}{2}} \| [w_h] \|_{L^2(F)} \\
 468 \quad \leq c \left(\sum_{F \in \mathcal{F}_h} \sum_{K \in \mathcal{T}_F} \lambda_K^{-\frac{1}{2}} h_K^{d(\frac{1}{2} - \frac{1}{p})} \|\boldsymbol{\sigma}(v)|_K\|_{\mathbf{L}^p(K)} |\mathcal{T}_F|^{-\frac{1}{2}} \lambda_F^{\frac{1}{2}} h_F^{-\frac{1}{2}} \| [w_h] \|_{L^2(F)} \right. \\
 469 \quad \left. + \sum_{F \in \mathcal{F}_h} \sum_{K \in \mathcal{T}_F} \lambda_K^{-\frac{1}{2}} h_K^{d(\frac{2+d}{2d} - \frac{1}{q})} \|\nabla \cdot \boldsymbol{\sigma}(v)|_K\|_{L^q(K)} |\mathcal{T}_F|^{-\frac{1}{2}} \lambda_F^{\frac{1}{2}} h_F^{-\frac{1}{2}} \| [w_h] \|_{L^2(F)} \right), \\
 470$$

471 where we used that $\theta_{K, F} \leq \theta_{K, F}^{\frac{1}{2}}$ (since $\theta_{K, F} \leq 1$), $|\mathcal{T}_F| \lambda_K \theta_{K, F} = \lambda_F$, the definition of
 472 $\|\cdot\|_{\mathbf{S}^d(K)}$, and $1 + d(\frac{1}{2} - \frac{1}{q}) = d(\frac{2+d}{2d} - \frac{1}{q})$. Owing to the Cauchy-Schwarz inequality, we
 473 infer that $\sum_{F \in \mathcal{F}_h} \sum_{K \in \mathcal{T}_F} a_K |\mathcal{T}_F|^{-\frac{1}{2}} b_F \leq (\sum_{K \in \mathcal{T}_h} |\mathcal{F}_K| a_K^2)^{\frac{1}{2}} (\sum_{F \in \mathcal{F}_h} b_F^2)^{\frac{1}{2}}$, for all real

474 numbers $\{a_K\}_{K \in \mathcal{T}_h}$, $\{b_F\}_{F \in \mathcal{F}_h}$, where we used $\sum_{F \in \mathcal{F}_h} \sum_{K \in \mathcal{T}_F} = \sum_{K \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_K}$
 475 for the term involving the a_K 's. Since $|\mathcal{F}_K|$ is uniformly bounded ($|\mathcal{F}_K| = d + 1$
 476 for simplicial meshes), applying this bound to the two terms composing the above
 477 estimate on $|n_{\sharp}(v, w_h)|$ leads to (3.21). \square

478 *Remark 3.6 (Literature).* Diffusion-dependent averages have been introduced in
 479 Dryja [19] for discontinuous Galerkin methods and have been analyzed, e.g., in Bur-
 480 man and Zunino [10], Dryja et al. [20], Di Pietro et al. [17], Ern et al. [26]. \square

481 **4. Applications.** The goal of this section is to perform a unified error analysis
 482 for the approximation of the model problem (2.1) with various nonconforming meth-
 483 ods: Crouzeix–Raviart finite elements, Nitsche’s boundary penalty, interior penalty
 484 discontinuous Galerkin, and hybrid high-order methods. We assume that the exact
 485 solution is in the functional space V_s defined in (2.5) with real numbers p, q satisfy-
 486 ing (3.2). Our unified analysis hinges on the dimensionally-consistent seminorm

$$487 \quad (4.1) \quad |v|_{\lambda, p, q}^2 := \|\lambda^{\frac{1}{2}} \nabla_h v\|_{\mathbf{L}^2(D)}^2 + |v|_{n_{\sharp}}^2, \quad \forall v \in V_s + P_k^b(\mathcal{T}_h),$$

488 with $|\cdot|_{n_{\sharp}}$ defined in (3.18). Since λ is piecewise constant, we have

$$489 \quad |v|_{\lambda, p, q}^2 := \sum_{K \in \mathcal{T}_h} \lambda_K \left(\|\nabla v|_K\|_{\mathbf{L}^2(K)}^2 + h_K^{2d(\frac{1}{2} - \frac{1}{p})} \|\nabla v|_K\|_{\mathbf{L}^p(K)}^2 \right. \\ 490 \quad (4.2) \quad \left. + h_K^{2d(\frac{d+2}{2d} - \frac{1}{q})} \|\Delta v|_K\|_{\mathbf{L}^q(K)}^2 \right).$$

492 Invoking inverse inequalities shows that there is c , uniform w.r.t. $h \in \mathcal{H}$, but depending
 493 on p and q , s.t.

$$494 \quad (4.3) \quad |v_h|_{\lambda, p, q} \leq c \|\lambda^{\frac{1}{2}} \nabla v_h\|_{\mathbf{L}^2(D)}, \quad \forall v_h \in P_k^b(\mathcal{T}_h).$$

495 **4.1. Abstract approximation result.** We start by recalling a general approx-
 496 imation result established in [25, Lem. 4.4]. Let V and W be two real Banach spaces.
 497 Let $a(\cdot, \cdot)$ be a bounded bilinear form on $V \times W$, and let $\ell(\cdot)$ be a bounded linear form
 498 on W , i.e., $\ell \in W'$. We consider the following abstract model problem:

$$499 \quad (4.4) \quad \begin{cases} \text{Find } u \in V \text{ such that} \\ a(u, w) = \ell(w), \quad \forall w \in W, \end{cases}$$

500 which we assume to be well-posed in the sense of Hadamard; that is to say, there is a
 501 unique solution and this solution depends continuously on the data.

502 We now formulate a discrete version of the problem (4.4) by using the Galerkin
 503 method. We replace the infinite-dimensional spaces V and W by finite-dimensional
 504 spaces V_h and W_h that are members of sequences of spaces $(V_h)_{h \in \mathcal{H}}$, $(W_h)_{h \in \mathcal{H}}$ en-
 505 dowed with some approximation properties as $h \rightarrow 0$. The norms in V_h and W_h are
 506 denoted by $\|\cdot\|_{V_h}$ and $\|\cdot\|_{W_h}$, respectively. The discrete version of (4.4) is formulated
 507 as follows:

$$508 \quad (4.5) \quad \begin{cases} \text{Find } u_h \in V_h \text{ such that} \\ a_h(u_h, w_h) = \ell_h(w_h), \quad \forall w_h \in W_h, \end{cases}$$

509 where $a_h(\cdot, \cdot)$ is a bounded bilinear form on $V_h \times W_h$ and $\ell_h(\cdot)$ is a bounded linear form
 510 on W_h ; note that $a_h(\cdot, \cdot)$ and $\ell_h(\cdot)$ possibly differ from $a(\cdot, \cdot)$ and $\ell(\cdot)$, respectively.
 511 We henceforth assume that $\dim(V_h) = \dim(W_h)$ and that

$$512 \quad (4.6) \quad \inf_{0 \neq v_h \in V_h} \sup_{0 \neq w_h \in W_h} \frac{|a_h(v_h, w_h)|}{\|v_h\|_{V_h} \|w_h\|_{W_h}} =: \alpha_h > 0, \quad \forall h > 0,$$

513 so that the discrete problem (4.5) is well-posed.

514 We formalize the fact that the error analysis requires the solution to (4.4) to be
515 slightly more regular than just being a member of V by introducing a functional space
516 V_s such that $u \in V_s \subsetneq V$. Our setting for the error analysis is therefore as follows:

$$517 \quad (4.7) \quad u \in V_s \subsetneq V, \quad u - u_h \in V_{\sharp} := V_s + V_h,$$

518 with the norm in V_{\sharp} denoted by $\|\cdot\|_{V_{\sharp}}$. Since V_h is finite-dimensional, we have

$$519 \quad (4.8) \quad c_{\sharp h} := \sup_{0 \neq v_h \in V_h} \frac{\|v_h\|_{V_{\sharp}}}{\|v_h\|_{V_h}} < \infty.$$

520 We now introduce the consistency error mapping $\delta_h : V_h \rightarrow W'_h := \mathcal{L}(W_h; \mathbb{R})$
521 defined for all $v_h \in V_h$ and all $w_h \in W_h$ by setting

$$522 \quad (4.9) \quad \langle \delta_h(v_h), w_h \rangle_{W'_h, W_h} := \ell_h(w_h) - a_h(v_h, w_h) = a_h(u_h - v_h, w_h).$$

523 We further assume that

$$524 \quad (4.10) \quad \omega_{\sharp h} := \sup_{u \in V_s} \sup_{v_h \in V_h \setminus \{u\}} \frac{\|\delta_h(v_h)\|_{W'_h}}{\|u - v_h\|_{V_{\sharp}}} < \infty.$$

525 *Example 4.1* (Conforming setting). Assume conformity, $a_h = a$, and $\ell_h = \ell$.
526 Take $V_s := V$, so that $V_{\sharp} = V$, and take $\|\cdot\|_{V_{\sharp}} := \|\cdot\|_V$. The consistency error (4.9) is
527 such that

$$528 \quad \langle \delta_h(v_h), w_h \rangle_{W'_h, W_h} = \ell(w_h) - a(v_h, w_h) = a(u - v_h, w_h),$$

529 where we used that $\ell(w_h) = a(u, w_h)$ (i.e., the Galerkin orthogonality property). Since
530 a is bounded on $V \times W$, (4.10) holds true with $\omega_{\sharp h} = \|a\|$; moreover, $c_{\sharp h} = 1$. \square

531 The main result we are going to invoke later is the following.

532 LEMMA 4.2 (Quasi-optimal error estimate). *If $u \in V_s$, then*

$$533 \quad (4.11) \quad \|u - u_h\|_{V_{\sharp}} \leq \left(1 + c_{\sharp h} \frac{\omega_{\sharp h}}{\alpha_h}\right) \inf_{v_h \in V_h} \|u - v_h\|_{V_{\sharp}}. \quad \square$$

534 *Proof.* The proof is classical; we sketch it for completeness. For all $v_h \in V_h$, we
535 have

$$536 \quad \|u_h - v_h\|_{V_{\sharp}} \leq c_{\sharp h} \|u_h - v_h\|_{V_h} \leq \frac{c_{\sharp h}}{\alpha_h} \sup_{0 \neq w_h \in W_h} \frac{|a_h(u_h - v_h, w_h)|}{\|w_h\|_{W_h}} \\ 537 \quad = \frac{c_{\sharp h}}{\alpha_h} \|\delta_h(v_h)\|_{W'_h} \leq \frac{c_{\sharp h} \omega_{\sharp h}}{\alpha_h} \|u - v_h\|_{V_{\sharp}}. \\ 538$$

539 We conclude by using the triangle inequality and taking the infimum over $v_h \in V_h$. \square

540 When the constants $c_{\sharp h}$ and $\omega_{\sharp h}$ can be bounded from above uniformly w.r.t.
541 $h \in \mathcal{H}$, we denote by c_{\sharp} and ω_{\sharp} any constant such that $c_{\sharp} \geq \sup_{h \in \mathcal{H}} c_{\sharp h}$ and $\omega_{\sharp} \geq$
542 $\sup_{h \in \mathcal{H}} \omega_{\sharp h}$.

543 **4.2. Crouzeix–Raviart approximation.** We consider in this section the ap-
 544 proximation of the model problem (2.2) with a homogeneous Dirichlet condition (for
 545 simplicity) using the Crouzeix–Raviart finite element space

$$546 \quad (4.12) \quad P_{1,0}^{\text{CR}}(\mathcal{T}_h) := \{v_h \in P_1^{\text{b}}(\mathcal{T}_h) \mid \int_F \llbracket v_h \rrbracket_F \, ds = 0, \forall F \in \mathcal{F}_h\}.$$

547 The discrete problem (4.5) is formulated with $V_h := P_{1,0}^{\text{CR}}(\mathcal{T}_h)$ and the following forms:

$$548 \quad (4.13) \quad a_h(v_h, w_h) := \int_D \lambda \nabla_h v_h \cdot \nabla_h w_h \, dx, \quad \ell_h(w_h) = \int_D f w_h \, dx.$$

549 We equip V_h with the norm $\|v_h\|_{V_h} := \|\lambda^{\frac{1}{2}} \nabla_h v_h\|_{\mathbf{L}^2(D)}$. The following result is stan-
 550 dard.

551 **LEMMA 4.3** (Coercivity, well-posedness). *The bilinear form a_h is coercive on V_h*
 552 *with coercivity constant $\alpha = 1$, and the discrete problem (4.5) is well-posed.* \square

553 Let $V_{\sharp} := V_s + V_h$ be equipped with the norm $\|v\|_{V_{\sharp}} := |v|_{\lambda,p,q}$ with $|v|_{\lambda,p,q}$ defined
 554 in (4.2) (this is indeed a norm on V_{\sharp} since $|v|_{\lambda,p,q} = 0$ implies that v is piecewise
 555 constant and hence vanishes identically owing to the definition of V_h). Owing to (4.3),
 556 there is c_{\sharp} , uniform w.r.t. $h \in \mathcal{H}$, but depending on p and q , s.t. $\|v_h\|_{V_{\sharp}} \leq c_{\sharp} \|v_h\|_{V_h}$,
 557 for all $v_h \in V_h$.

558 **LEMMA 4.4** (Consistency/boundedness). *There is ω_{\sharp} , uniform w.r.t. $h \in \mathcal{H}$, λ ,*
 559 *and $u \in V_s$, but depending on p and q , s.t. $\|\delta_h(v_h)\|_{V_h'} \leq \omega_{\sharp} \|u - v_h\|_{V_{\sharp}}$, for all*
 560 *$v_h \in V_h$.* \square

561 *Proof.* Let $v_h, w_h \in V_h$. Since $V_h \subset P_k^{\text{b}}(\mathcal{T}_h)$, the identity (3.16a) implies that

$$562 \quad n_{\sharp}(v_h, w_h) = \sum_{F \in \mathcal{F}_h} \int_F \{\boldsymbol{\sigma}(v_h)\}_{\theta} \cdot \mathbf{n}_F \llbracket w_h \rrbracket \, ds = 0,$$

563 because $\{\boldsymbol{\sigma}(v_h)\}_{\theta} \cdot \mathbf{n}_F$ is constant over F . Moreover, invoking the identity (3.16b) with
 564 $v = u$ and since $f = \nabla \cdot \boldsymbol{\sigma}(u)$, we have

$$565 \quad \ell_h(w_h) = n_{\sharp}(u, w_h) - \int_D \boldsymbol{\sigma}(u) \cdot \nabla_h w_h \, dx.$$

566 Combining the two above identities and letting $\eta := u - v_h$, we obtain

$$567 \quad \langle \delta_h(v_h), w_h \rangle_{V_h', V_h} = n_{\sharp}(u, w_h) + \int_D \lambda \nabla_h \eta \cdot \nabla_h w_h \, dx = n_{\sharp}(\eta, w_h) + \int_D \lambda \nabla_h \eta \cdot \nabla_h w_h \, dx.$$

569 The first term on the right-hand side is estimated by invoking the boundedness of
 570 n_{\sharp} (Lemma 3.5), the inequality $\lambda_F \leq \min_{K \in \mathcal{T}_F} \lambda_K$ (see (3.20)), and the bound
 571 $\sum_{F \in \mathcal{F}_h} \lambda_F h_F^{-1} \|\llbracket w_h \rrbracket\|_{L^2(F)}^2 \leq c \|w_h\|_{V_h}^2$, which is standard for Crouzeix–Raviart el-
 572 ements. The second term is estimated by using the Cauchy–Schwarz inequality. \square

573 **THEOREM 4.5** (Error estimate). *Let u solve (2.2) and u_h solve (4.5) with a_h*
 574 *and ℓ_h defined in (4.13). Assume that there is $r > 0$ s.t. $u \in H^{1+r}(D)$. There is*
 575 *c , uniform w.r.t. $h \in \mathcal{H}$, λ , and $u \in H^{1+r}(D)$, but depending on r , s.t. the following*
 576 *quasi-optimal error estimate holds true:*

$$577 \quad (4.14) \quad \|u - u_h\|_{V_{\sharp}} \leq c \inf_{v_h \in V_h} \|u - v_h\|_{V_{\sharp}}.$$

578 Moreover, letting $t := \min(1, r)$, where $1 = k$ is the degree of the Crouzeix–Raviart
579 finite element, we have

$$580 \quad (4.15) \quad \|u - u_h\|_{V_h} \leq c \left(\sum_{K \in \mathcal{T}_h} \lambda_K h_K^{2t} |u|_{H^{1+t}(K)}^2 + \lambda_K^{-1} h_K^{2d(\frac{d+2}{2d} - \frac{1}{q})} \|f\|_{L^q(K)}^2 \right)^{\frac{1}{2}}. \quad \square$$

581 *Proof.* The error estimate (4.14) follows from Lemma 4.2 combined with stability
582 (Lemma 4.3) and consistency/boundedness (Lemma 4.4). We now bound the infimum
583 in (4.14) by considering $\eta := u - \mathcal{I}_h^{\text{CR}}(u)$, where $\mathcal{I}_h^{\text{CR}}$ is the Crouzeix–Raviart interpo-
584 lation operator using averages over the faces as degrees of freedom. It is a standard
585 approximation result that there is c , uniform w.r.t. $u \in H^{1+t}(K)$, $t \geq 0$, and $h \in \mathcal{H}$,
586 s.t. $\|\nabla \eta|_K\|_{L^2(K)} \leq c h_K^t |u|_{H^{1+t}(K)}$ for all $K \in \mathcal{T}_h$. Moreover, invoking the embedding
587 $\mathbf{H}^t(\widehat{K}) \hookrightarrow \mathbf{L}^p(\widehat{K})$ and classical results on the transformation of Sobolev norms by the
588 geometric mapping, we obtain the bound

$$589 \quad (4.16) \quad h_K^{d(\frac{1}{2} - \frac{1}{p})} \|\nabla \eta|_K\|_{\mathbf{L}^p(K)} \leq c (\|\nabla \eta|_K\|_{L^2(K)} + h_K^t |\nabla \eta|_K|_{\mathbf{H}^t(K)}).$$

590 Observing that $|\nabla \eta|_K|_{\mathbf{H}^t(K)} = |u|_{H^{1+t}(K)}$ since $\mathcal{I}_h^{\text{CR}}(u)$ is affine on K and using
591 again the approximation properties of $\mathcal{I}_h^{\text{CR}}$, we infer that $h_K^{d(\frac{1}{2} - \frac{1}{p})} \|\nabla \eta|_K\|_{\mathbf{L}^p(K)} \leq$
592 $c h_K^t |u|_{H^{1+t}(K)}$. Finally, we have $\Delta \eta|_K = \lambda_K^{-1} f$ in K . \square

593 *Remark 4.6 (Convergence).* The rightmost term in (4.15) converges as $O(h)$
594 when $q = 2$. Moreover, convergence is lost when $q \leq \frac{2d}{d+2}$, which is somewhat natural
595 since in this case the linear form $w \mapsto \int_D f w \, dx$ is no longer bounded on $H^1(D)$. \square

596 *Remark 4.7 (Weights).* Although the weights introduced in (3.19) are not ex-
597 plicitly used in the Crouzeix–Raviart discretization, they play a role in the error
598 analysis. More precisely, we used the boundedness of the bilinear form n_{\sharp} together
599 with $\lambda_F \leq \min_{K \in \mathcal{T}_F} \lambda_K$ in the proof of Lemma 4.4. The present approach is some-
600 what more general than that in Li and Mao [31] since it delivers error estimates that
601 are robust with respect to the diffusivity contrast. The trimming operator invoked in
602 [31, Eq. (5)–(7)] cannot account for the diffusivity contrast. \square

603 **4.3. Nitsche’s boundary penalty method.** We consider in this section the
604 approximation of the model problem (2.1) by means of Nitsche’s boundary penalty
605 method. Now we set

$$606 \quad (4.17) \quad V_h := P_k^g(\mathcal{T}_h) := \{v_h \in P_k^b(\mathcal{T}_h) \mid \llbracket v_h \rrbracket_F = 0, \forall F \in \mathcal{F}_h^o\}, \quad k \geq 1,$$

607 i.e., V_h is H^1 -conforming. The discrete problem (4.5) is formulated with $V_h := P_k^g(\mathcal{T}_h)$
608 and the following forms:

$$609 \quad (4.18a) \quad a_h(v_h, w_h) := a(v_h, w_h) + \sum_{F \in \mathcal{F}_h^o} \int_F \left(\boldsymbol{\sigma}(v_h) \cdot \mathbf{n} + \varpi_0 \frac{\lambda_{K_l}}{h_F} v_h \right) w_h \, ds,$$

$$610 \quad (4.18b) \quad \ell_h(w_h) := \ell(w_h) + \sum_{F \in \mathcal{F}_h^o} \varpi_0 \frac{\lambda_{K_l}}{h_F} \int_F g w_h \, ds,$$

611
612 where the exact forms a and ℓ are defined in (2.3), K_l is the unique mesh cell s.t.
613 $F = \partial K_l \cap \partial D$, and the user-specified penalty parameter ϖ_0 is yet to be chosen large
614 enough. It is possible to add a symmetrizing term to the discrete bilinear form a_h .

615 We equip V_h with the norm $\|v_h\|_{V_h}^2 := \|\lambda^{\frac{1}{2}} \nabla v_h\|_{L^2(D)}^2 + |v_h|_{\partial}^2$ with $|v_h|_{\partial}^2 :=$
 616 $\sum_{F \in \mathcal{F}_h^\partial} \frac{\lambda_{K_l}}{h_F} \|v_h\|_{L^2(F)}^2$. Owing to the shape-regularity of the mesh sequence, there
 617 is c_I , uniform w.r.t. $h \in \mathcal{H}$ s.t.

$$618 \quad (4.19) \quad \|v_h\|_{L^2(F)} \leq c_I h_F^{-\frac{1}{2}} \|v_h\|_{L^2(K_l)},$$

619 for all $v_h \in V_h$ and all $F \in \mathcal{F}_h^\partial$. Let n_∂ denote the maximum number of boundary
 620 faces that a mesh cell can have ($n_\partial \leq d$ for simplicial meshes). The proof of the
 621 following result uses standard arguments.

622 LEMMA 4.8 (Coercivity, well-posedness). *Assume that the penalty parameter*
 623 *satisfies $\varpi_0 > \frac{1}{4} n_\partial c_I^2$. Then, a_h is coercive on V_h with constant $\alpha := \frac{\varpi_0 - \frac{1}{4} n_\partial c_I^2}{1 + \varpi_0} > 0$,*
 624 *and the discrete problem (4.5) is well-posed. \square*

625 Let $V_\sharp := V_S + V_h$. We equip the space V_\sharp with the norm $\|v\|_{V_\sharp}^2 := |v|_{\lambda, p, q}^2 + |v|_{\partial}^2$
 626 with

$$627 \quad |v|_{\lambda, p, q}^2 := \sum_{K \in \mathcal{T}_h} \lambda_K \|\nabla v|_K\|_{L^2(K)}^2$$

$$628 \quad (4.20) \quad + \sum_{K \in \overline{\mathcal{T}}_h^\partial} \lambda_K \left(h_K^{2d(\frac{1}{2} - \frac{1}{p})} \|\nabla v|_K\|_{L^p(K)}^2 + h_K^{2d(\frac{d+2}{2d} - \frac{1}{q})} \|\Delta v|_K\|_{L^q(K)}^2 \right),$$

629

630 where $\overline{\mathcal{T}}_h^\partial$ is the collection of the mesh cells having at least one boundary face, and
 631 $|v|_{\partial}^2 = \sum_{F \in \mathcal{F}_h^\partial} \frac{\lambda_{K_l}}{h_F} \|v\|_{L^2(F)}^2$. Owing to (4.3), there is c_\sharp , uniform w.r.t. $h \in \mathcal{H}$, but
 632 depending on p and q , s.t. $\|v_h\|_{V_\sharp} \leq c_\sharp \|v_h\|_{V_h}$, for all $v_h \in V_h$.

633 LEMMA 4.9 (Consistency/boundedness). *There is ω_\sharp , uniform w.r.t. $h \in \mathcal{H}$, λ ,*
 634 *and $u \in V_S$, but depending on p and q , s.t. $\|\delta_h(v_h)\|_{V'_h} \leq \omega_\sharp \|u - v_h\|_{V_\sharp}$, for all*
 635 *$v_h \in V_h$. \square*

636 *Proof.* Let $v_h, w_h \in V_h$. Using the identity (3.16a) for n_\sharp , $[[w_h]]_F = 0$ for all
 637 $F \in \mathcal{F}_h^\circ$ (since V_h is H^1 -conforming), and the definition of the weights at the bound-
 638 ary faces, we infer that $n_\sharp(v_h, w_h) = \sum_{F \in \mathcal{F}_h^\partial} \int_F \boldsymbol{\sigma}(v_h) \cdot \mathbf{n} w_h \, ds$. Hence, $a_h(v_h, w_h) =$
 639 $a(v_h, w_h) + n_\sharp(v_h, w_h) + \sum_{F \in \mathcal{F}_h^\partial} \varpi_0 \frac{\lambda_{K_l}}{h_F} \int_F v_h w_h \, ds$. Therefore, invoking the iden-
 640 tity (3.16b) for the exact solution u and observing that $f = \nabla \cdot \boldsymbol{\sigma}(u)$, we infer the
 641 important identity $\int_D f w_h \, dx = a(u, w_h) + n_\sharp(u, w_h)$. Then, recalling that $\gamma^g(u) = g$,
 642 and letting $\eta := u - v_h$, we obtain

$$643 \quad \langle \delta_h(v_h), w_h \rangle_{V'_h, V_h} = n_\sharp(\eta, w_h) + a(\eta, w_h) + \sum_{F \in \mathcal{F}_h^\partial} \varpi_0 \frac{\lambda_{K_l}}{h_F} \int_F \eta w_h \, ds.$$

644 We conclude by using the boundedness of n_\sharp from Lemma 3.5 and the Cauchy–Schwarz
 645 inequality. \square

646 THEOREM 4.10 (Error estimate). *Let u solve (2.1) and u_h solve (4.5) with a_h and*
 647 *ℓ_h defined in (4.18) and penalty parameter $\varpi_0 > \frac{1}{4} n_\partial c_I^2$. Assume that there is $r > 0$*
 648 *s.t. $u \in H^{1+r}(D)$. There is c , uniform with respect to $h \in \mathcal{H}$, λ , and $u \in H^{1+r}(D)$,*
 649 *but depending on r , s.t. the following quasi-optimal error estimate holds true:*

$$650 \quad (4.21) \quad \|u - u_h\|_{V_\sharp} \leq c \inf_{v_h \in V_h} \|u - v_h\|_{V_\sharp}.$$

651 Moreover, letting $t := \min(r, k)$, $\chi_t = 1$ if $t \leq 1$ and $\chi_t = 0$ if $t > 1$, we have

$$652 \quad (4.22) \quad \|u - u_h\|_{V_h} \leq c \left(\sum_{K \in \tilde{\mathcal{T}}_h} \lambda_K h_K^{2t} |u|_{H^{1+t}(\tilde{\mathcal{T}}_K)}^2 + \frac{\chi_t}{\lambda_K} h_K^{2d(\frac{d+2}{2d} - \frac{1}{q})} \|f\|_{L^q(K)}^2 \right)^{\frac{1}{2}},$$

653 where $\tilde{\mathcal{T}}_K$ is the collection of the mesh cells having at least a common vertex with K .
654 The broken Sobolev norm $|\cdot|_{H^{1+t}(\tilde{\mathcal{T}}_K)}$ can be replaced by $|\cdot|_{H^{1+t}(K)}$ if $1 + t > \frac{d}{2}$. \square

655 *Proof.* The error estimate (4.21) follows from Lemma 4.2 combined with stabil-
656 ity (Lemma 4.8) and consistency/boundedness (Lemma 4.9). We now bound the
657 infimum in (4.21) by using $\eta := u - \mathcal{I}_h^{\text{g,av}}(u)$, where $\mathcal{I}_h^{\text{g,av}}$ is the quasi-interpolation
658 operator introduced in [23, §5]. We take the polynomial degree of $\mathcal{I}_h^{\text{g,av}}$ to be $\ell := [t]$,
659 where $[t]$ denotes the smallest integer $n \in \mathbb{N}$ s.t. $n \geq t$. Notice that $\ell \geq 1$ be-
660 cause $r > 0$ and $k \geq 1$, and $\ell \leq k$ because $t \leq k$; hence, $\mathcal{I}_h^{\text{g,av}}(u) \in V_h$. We
661 need to bound all the terms composing the norm $\|\eta\|_{V_h}$. Owing to [23, Thm. 5.2]
662 with $m = 1$, we have $\|\nabla \eta\|_{L^2(K)} \leq c h_K^t |u|_{H^{1+t}(\tilde{\mathcal{T}}_K)}$ for all $K \in \mathcal{T}_h$. Moreover,
663 we have $h_F^{-\frac{1}{2}} \|\eta\|_{L^2(F)} \leq c h_{K_l}^t |u|_{H^{1+t}(\tilde{\mathcal{T}}_{K_l})}$ for all $F \in \mathcal{F}_h^\partial$. It remains to estimate
664 $h_K^{d(\frac{1}{2} - \frac{1}{p})} \|\nabla \eta|_K\|_{L^p(K)}$ and $h_K^{d(\frac{d+2}{2d} - \frac{1}{q})} \|\Delta \eta|_K\|_{L^q(K)}$ for all $K \in \overline{\mathcal{T}}_h^\partial$. Using (4.16), the
665 above bound on $\|\nabla \eta\|_{L^2(K)}$, and $|\nabla \eta|_{H^t(K)} = |\nabla u|_{H^t(K)} = |u|_{H^{1+t}(K)}$ since $\ell < 1 + t$,
666 we infer that $h_K^{d(\frac{1}{2} - \frac{1}{p})} \|\nabla \eta\|_{L^p(K)} \leq c h_K^t |u|_{H^{1+t}(\tilde{\mathcal{T}}_K)}$. Moreover, if $t \leq 1$, we have $\ell = 1$
667 so that $\|\Delta \eta|_K\|_{L^q(K)} = \|\Delta u\|_{L^q(K)} = \lambda_K^{-1} \|f\|_{L^q(K)}$. Instead, if $t > 1$, we infer that
668 $r > 1$ so that we can set $q = 2$ (recall that $f|_{D_i} = \lambda_{|D_i}(\Delta u)_{D_i}$ for all $i \in \{1:M\}$, and
669 $u \in H^2(D)$ if $r \geq 1$), and we estimate $\|\Delta \eta|_K\|_{L^2(K)}$ using [23, Thm. 5.2] with $m = 2$.
670 Finally, if $1 + t > \frac{d}{2}$, we can use the canonical Lagrange interpolation operator \mathcal{I}_h^{g}
671 instead of $\mathcal{I}_h^{\text{g,av}}$, and this allows us to replace $|\cdot|_{H^{1+t}(\tilde{\mathcal{T}}_K)}$ by $|\cdot|_{H^{1+t}(K)}$ in (4.22). \square

672 **4.4. Discontinuous Galerkin.** We consider in this section the approximation
673 of the model problem (2.1) by means of the symmetric interior penalty discontinuous
674 Galerkin method. The discrete problem (4.5) is formulated with $V_h := P_k^{\text{b}}(\mathcal{T}_h)$, $k \geq 1$,
675 the bilinear forms

$$676 \quad a_h(v_h, w_h) := \int_D \lambda \nabla_h v_h \cdot \nabla_h w_h \, dx + \sum_{F \in \mathcal{F}_h} \int_F \{\boldsymbol{\sigma}(v_h)\}_\theta \cdot \mathbf{n}_F \llbracket w_h \rrbracket \, ds$$

$$677 \quad (4.23a) \quad + \sum_{F \in \mathcal{F}_h} \int_F \llbracket v_h \rrbracket \{\boldsymbol{\sigma}(w_h)\}_\theta \cdot \mathbf{n}_F \, ds + \sum_{F \in \mathcal{F}_h} \varpi_0 \frac{\lambda_F}{h_F} \int_F \llbracket v_h \rrbracket \llbracket w_h \rrbracket \, ds,$$

$$678 \quad (4.23b) \quad \ell_h(w_h) := \ell(w_h) + \sum_{F \in \mathcal{F}_h^\partial} \varpi_0 \frac{\lambda_{K_l}}{h_F} \int_F g w_h \, ds,$$

680 where ℓ is defined in (2.3), λ_F in (3.20), and the user-specified penalty param-
681 eter ϖ_0 is yet to be chosen large enough. We equip V_h with the norm $\|v_h\|_{V_h}^2 :=$
682 $\|\lambda^{\frac{1}{2}} \nabla_h v_h\|_{L^2(D)}^2 + |v_h|_{\mathbb{J}}^2$ with $|v_h|_{\mathbb{J}}^2 := \sum_{F \in \mathcal{F}_h} \frac{\lambda_F}{h_F} \|\llbracket v_h \rrbracket\|_{L^2(F)}^2$. Recall the discrete trace
683 inequality (4.19) and recall that n_∂ denotes the maximum number of faces that a
684 mesh cell can have ($n_\partial \leq d + 1$ for simplicial meshes). The proof of the following
685 result uses standard arguments.

686 **LEMMA 4.11 (Coercivity, well-posedness).** *Assume that the penalty parameter*
687 *satisfies $\varpi_0 > n_\partial c_I^2$. Then, a_h is coercive on V_h with constant $\alpha := \frac{\varpi_0 - n_\partial c_I^2}{1 + \varpi_0} > 0$, and*
688 *the discrete problem (4.5) is well-posed. \square*

689 Let $V_{\sharp} := V_s + V_h$. We equip the space V_{\sharp} with the norm $\|v\|_{V_{\sharp}}^2 := |v|_{\lambda,p,q}^2 + |v|_J^2$
 690 with $|v|_{\lambda,p,q}$ defined in (4.2) and $|v|_J^2 := \sum_{F \in \mathcal{F}_h} \frac{\lambda_F}{h_F} \|\llbracket v \rrbracket\|_{L^2(F)}^2$. Owing to (4.3), there
 691 is c_{\sharp} , uniform w.r.t. $h \in \mathcal{H}$, but depending on p and q , s.t. $\|v_h\|_{V_{\sharp}} \leq c_{\sharp} \|v_h\|_{V_h}$, for all
 692 $v_h \in V_h$.

693 **LEMMA 4.12** (Consistency/boundedness). *There is ω_{\sharp} , uniform w.r.t. $h \in \mathcal{H}$,
 694 λ , and $u \in V_s$, but depending on p and q , s.t. $\|\delta_h(v_h)\|_{V_h'} \leq \omega_{\sharp} \|u - v_h\|_{V_{\sharp}}$, for all
 695 $v_h \in V_h$. \square*

696 *Proof.* Let $v_h, w_h \in V_h$. Owing to (3.16b) and since $f = \nabla \cdot \boldsymbol{\sigma}(u)$, we infer that
 697 $\int_D f w_h \, dx = \sum_{K \in \mathcal{T}_h} a_K(u, w_h) + n_{\sharp}(u, w_h)$ with $a_K(u, w_h) := -(\boldsymbol{\sigma}(u), \nabla_h w_h)_{L^2(K)}$.
 698 Using the identity (3.16a), we obtain

$$699 \quad \ell_h(w_h) = n_{\sharp}(u, w_h) - \int_D \boldsymbol{\sigma}(u) \cdot \nabla_h w_h \, dx + \sum_{F \in \mathcal{F}_h^{\partial}} \varpi_0 \frac{\lambda_F}{h_F} \int_F g w_h \, ds,$$

$$700 \quad a_h(v_h, w_h) = \int_D -\boldsymbol{\sigma}(v_h) \cdot \nabla_h w_h \, dx + n_{\sharp}(v_h, w_h)$$

$$701 \quad - \sum_{F \in \mathcal{F}_h} \int_F \llbracket v_h \rrbracket \{ \boldsymbol{\sigma}(w_h) \}_{\theta} \cdot \mathbf{n}_F \, ds + \sum_{F \in \mathcal{F}_h} \varpi_0 \frac{\lambda_F}{h_F} \int_F \llbracket v_h \rrbracket \llbracket w_h \rrbracket \, ds.$$

703 Then setting $\eta := u - v_h$ and using that $\llbracket u \rrbracket_F = 0$ for all $F \in \mathcal{F}_h^{\circ}$ and $\llbracket u \rrbracket_F = g$ for all
 704 $F \in \mathcal{F}_h^{\partial}$, we obtain the following representation of the consistency linear form $\delta_h(v_h)$:

$$705 \quad \langle \delta_h(v_h), w_h \rangle_{V_h', V_h} = n_{\sharp}(\eta, w_h) + \int_D \lambda \nabla \eta \cdot \nabla_h w_h \, dx$$

$$706 \quad - \sum_{F \in \mathcal{F}_h} \int_F \llbracket \eta \rrbracket \{ \boldsymbol{\sigma}(w_h) \}_{\theta} \cdot \mathbf{n}_F \, ds + \sum_{F \in \mathcal{F}_h} \varpi_0 \frac{\lambda_F}{h_F} \int_F \llbracket \eta \rrbracket \llbracket w_h \rrbracket \, ds.$$

708 Bounding the second, third and fourth terms uses standard arguments (see, e.g., [15]),
 709 whereas we invoke the boundedness estimate on n_{\sharp} from Lemma 3.5 for the first term. \square

710 **THEOREM 4.13** (Error estimate). *Let u solve (2.1) and u_h solve (4.5) with a_h
 711 and ℓ_h defined in (4.23) and penalty parameter $\varpi_0 > n_{\partial} c_1^2$. Assume that there is $r > 0$
 712 s.t. $u \in H^{1+r}(D)$. There is c , uniform with respect to $h \in \mathcal{H}$, λ , and $u \in H^{1+r}(D)$,
 713 but depending on r , s.t. the following quasi-optimal error estimate holds true:*

$$714 \quad (4.24) \quad \|u - u_h\|_{V_{\sharp}} \leq c \inf_{v_h \in V_h} \|u - v_h\|_{V_{\sharp}}.$$

715 *Moreover, letting $t := \min(r, k)$, $\chi_t = 1$ if $t \leq 1$ and $\chi_t = 0$ if $t > 1$, we have*

$$716 \quad (4.25) \quad \|u - u_h\|_{V_{\sharp}} \leq c \left(\sum_{K \in \mathcal{T}_h} \lambda_K h_K^{2t} |u|_{H^{1+t}(K)}^2 + \frac{\chi_t}{\lambda_K} h_K^{2d(\frac{d+2}{2d} - \frac{1}{q})} \|f\|_{L^q(K)}^2 \right)^{\frac{1}{2}}. \quad \square$$

717 *Proof.* We proceed as in the proof of Theorem 4.10, where we now use the L^1 -
 718 stable interpolation operator $\mathcal{I}_h^{\sharp} : L^1(D) \rightarrow P_k^b(\mathcal{T}_h)$ from [23, §3] to estimate the best
 719 approximation error. \square

720 **4.5. Hybrid high-order methods.** We consider in this section the approxi-
 721 mation of the model problem (2.1) with a homogeneous Dirichlet condition (for sim-
 722 plicity) by means of the hybrid high-order (HHO) method introduced in [16, 18]. We

762 The discrete problem is finally formulated as follows: Find $\hat{u}_h \in \hat{V}_{h,0}^k$ s.t.

$$763 \quad (4.34) \quad \hat{a}_h(\hat{u}_h, \hat{w}_h) = \hat{\ell}_h(\hat{w}_h), \quad \forall \hat{w}_h \in \hat{V}_{h,0}^k.$$

764 Notice that HHO methods are somewhat simpler than dG methods when it comes
765 to solving problems with contrasted coefficients. For HHO methods, one assembles
766 cellwise the local bilinear forms \hat{a}_K weighted by the local diffusion coefficient λ_K ,
767 whereas, for dG methods one has to invoke interface-based values of the diffusion
768 coefficient to construct the penalty term.

769 The following result is proved in [16, 18].

770 LEMMA 4.14 (Stability, boundedness, well-posedness). *There are $0 < \alpha \leq \omega$,*
771 *uniform w.r.t. $h \in \mathcal{H}$, such that*

$$772 \quad \alpha |\hat{v}_K|_{\hat{V}_K^k}^2 \leq \|\nabla \mathbf{R}_K^{k+1}(\hat{v}_K)\|_{L^2(K)}^2 + \|h_{\partial K}^{-\frac{1}{2}} \mathbf{S}_{\partial K}^k(\hat{v}_K)\|_{L^2(\partial K)}^2 = \hat{a}_K(\hat{v}_K, \hat{v}_K) \leq \omega |\hat{v}_K|_{\hat{V}_K^k}^2,$$

773 *for all $\hat{v}_K \in \hat{V}_K$ and all $K \in \mathcal{T}_h$, and the discrete problem (4.34) is well-posed.* \square

774 The two key tools in the error analysis of HHO methods are a local reduction
775 operator and the local elliptic projection. For all $K \in \mathcal{T}_h$, the local reduction operator
776 $\hat{\mathcal{I}}_K^k : H^1(K) \rightarrow \hat{V}_K^k$ is defined by $\hat{\mathcal{I}}_K^k(v) := (\Pi_K^k(v), \Pi_{\partial K}^k(\gamma_{\partial K}^g(v))) \in \hat{V}_K^k$, for all
777 $v \in H^1(K)$. The local elliptic projection $\mathcal{E}_K^{k+1} : H^1(K) \rightarrow \mathbb{P}_{k+1,d}$ is s.t. $(\nabla(\mathcal{E}_K^{k+1}(v) -$
778 $v), \nabla q)_{L^2(K)} = 0$, for all $q \in \mathbb{P}_{k+1,d}$, and $(\mathcal{E}_K^{k+1}(v) - v, 1)_{L^2(K)} = 0$. The following
779 result is established in [16, 18].

780 LEMMA 4.15 (Polynomial invariance). *The following holds true:*

$$781 \quad (4.35a) \quad \mathbf{R}_K^{k+1} \circ \hat{\mathcal{I}}_K^k = \mathcal{E}_K^{k+1},$$

$$782 \quad (4.35b) \quad \mathbf{S}_{\partial K}^k \circ \hat{\mathcal{I}}_K^k = (\gamma_{\partial K}^g \circ \Pi_K^k - \Pi_{\partial K}^k \circ \gamma_{\partial K}^g) \circ (I - \mathcal{E}_K^{k+1}).$$

784 *In particular, $\mathbf{R}_K^{k+1}(\hat{\mathcal{I}}_K^k(p)) = p$ and $\mathbf{S}_{\partial K}^k(\hat{\mathcal{I}}_K^k(p)) = 0$ for all $p \in \mathbb{P}_{k+1,d}$.* \square

785 Recalling the duality pairing $\langle \cdot, \cdot \rangle_F$ defined in (3.10), the generalization of the
786 bilinear form n_{\sharp} in the context of HHO methods is the bilinear form defined on
787 $(V_s + P_{k+1}^b(\mathcal{T}_h)) \times \hat{V}_{h,0}^k$ that acts as follows:

$$788 \quad (4.36) \quad n_{\sharp}(v, \hat{w}_h) := \sum_{K \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_K} \langle (\boldsymbol{\sigma}(v) \cdot \mathbf{n}_K)|_F, (w_K - w_{\partial K})|_F \rangle_F.$$

789 LEMMA 4.16 (Identities and boundedness for n_{\sharp}). *The following holds true for*
790 *all $\hat{w}_h \in \hat{V}_{h,0}^k$, all $v_h \in P_{k+1}^b(\mathcal{T}_h)$ and all $v \in V_s$:*

$$791 \quad (4.37a) \quad n_{\sharp}(v_h, \hat{w}_h) = \sum_{K \in \mathcal{T}_h} \int_K \lambda_K \nabla v_h|_K \cdot \nabla (\mathbf{R}_K^{k+1}(\hat{w}_K) - w_K) \, dx,$$

$$792 \quad (4.37b) \quad n_{\sharp}(v, \hat{w}_h) = \sum_{K \in \mathcal{T}_h} \int_K \left(\boldsymbol{\sigma}(v) \cdot \nabla w_K + (\nabla \cdot \boldsymbol{\sigma}(v)) w_K \right) \, dx.$$

794 *Moreover, there is c , uniform w.r.t. $h \in \mathcal{H}$ and λ , but depending on p and q , s.t. the*
795 *following holds true for all $v \in V_s + P_{k+1}^b(\mathcal{T}_h)$ and all $\hat{w}_h \in \hat{V}_{h,0}^k$:*

$$796 \quad (4.38) \quad |n_{\sharp}(v, \hat{w}_h)| \leq c |v|_{n_{\sharp}} \left(\sum_{K \in \mathcal{T}_h} \lambda_K h_K^{-1} \|w_K - w_{\partial K}\|_{L^2(\partial K)}^2 \right)^{\frac{1}{2}},$$

797 *with the $|\cdot|_{n_{\sharp}}$ -seminorm defined in (3.18).* \square

798 *Proof.* (i) We first prove (4.37a). Let $v_h \in P_{k+1}^b(\mathcal{T}_h)$ and $\hat{w}_h \in \hat{V}_{h,0}^k$. Since the
799 restriction of $\boldsymbol{\sigma}(v_h)$ to each mesh cell is smooth and since the trace on ∂K of the
800 face-to-cell lifting operator L_F^K is nonzero only on F , for all $F \in \mathcal{F}_K$, we have

$$\begin{aligned} 801 & \langle (\boldsymbol{\sigma}(v_h) \cdot \mathbf{n}_K)|_F, (w_K - w_{\partial K})|_F \rangle_F \\ 802 & = \int_K \boldsymbol{\sigma}(v_h)|_K \cdot \nabla L_F^K((w_K - w_{\partial K})|_F) + (\nabla \cdot \boldsymbol{\sigma}(v_h)|_K) L_F^K((w_K - w_{\partial K})|_F) \, dx \\ 803 & = \int_{\partial K} \boldsymbol{\sigma}(v_h)|_K \cdot \mathbf{n}_K L_F^K((w_K - w_{\partial K})|_F) \, ds = \int_F \boldsymbol{\sigma}(v_h)|_K \cdot \mathbf{n}_K (w_K - w_{\partial K}) \, ds, \end{aligned}$$

805 where we used the divergence formula in K . Therefore, we obtain

$$\begin{aligned} 806 & n_{\sharp}(v_h, \hat{w}_h) = \sum_{K \in \mathcal{T}_h} \int_{\partial K} \boldsymbol{\sigma}(v_h)|_K \cdot \mathbf{n}_K (w_K - w_{\partial K}) \, ds \\ 807 & = - \sum_{K \in \mathcal{T}_h} \lambda_K \int_{\partial K} \nabla v_h|_K \cdot \mathbf{n}_K (w_K - w_{\partial K}) \, ds \\ 808 & = \sum_{K \in \mathcal{T}_h} \lambda_K \int_K (\nabla v_h|_K \cdot \nabla (\mathbf{R}_K^{k+1}(\hat{w}_K) - w_K)) \, dx, \end{aligned}$$

810 where we used the definition (4.29) of the local reconstruction operator \mathbf{R}_K^{k+1} with the
811 test function $v_h|_K \in \mathbb{P}_{k,d} \subset \mathbb{P}_{k+1,d}$.

812 (ii) Let us now prove (4.37b). Let $v \in V_s$ and $\hat{w}_h \in \hat{V}_{h,0}^k$. We are going to proceed as in
813 the proof of (3.16b). We consider the mollification operators $\mathcal{K}_\delta^d : L^1(D) \rightarrow C^\infty(\bar{D})$
814 and $\mathcal{K}_\delta^b : L^1(D) \rightarrow C^\infty(\bar{D})$ introduced in [22, §3.2]. Let us consider the mollified
815 bilinear form

$$816 \quad n_{\sharp\delta}(v, \hat{w}_h) := \sum_{K \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_K} \langle (\mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \cdot \mathbf{n}_K)|_F, (w_K - w_{\partial K})|_F \rangle_F.$$

817 By using (3.10) and invoking the approximation properties of the mollification opera-
818 tors and the commuting property (3.17), we infer that $\lim_{\delta \rightarrow 0} n_{\sharp\delta}(v, \hat{w}_h) = n_{\sharp}(v, \hat{w}_h)$.
819 Since the restriction of $\mathcal{K}_\delta^d(\boldsymbol{\sigma}(v))$ to each mesh cell is smooth and since $\mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \in$
820 $C^0(\bar{D})$, we infer that

$$\begin{aligned} 821 & n_{\sharp\delta}(v, \hat{w}_h) = \sum_{K \in \mathcal{T}_h} \int_{\partial K} \mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \cdot \mathbf{n}_K (w_K - w_{\partial K}) \, ds = \sum_{K \in \mathcal{T}_h} \int_{\partial K} \mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \cdot \mathbf{n}_K w_K \, ds \\ 822 & = \sum_{K \in \mathcal{T}_h} \int_K (\mathcal{K}_\delta^d(\boldsymbol{\sigma}(v)) \cdot \nabla w_K + \mathcal{K}_\delta^b(\nabla \cdot \boldsymbol{\sigma}(v)) w_K) \, dx, \end{aligned}$$

824 where we used the divergence formula and the commuting property (3.17) in the last
825 line. Letting $\delta \rightarrow 0$, we conclude that $n_{\sharp\delta}(v, \hat{w}_h)$ also tends to the right-hand side
826 of (4.37b) as $\delta \rightarrow 0$. Hence, (4.37b) holds true.

827 (iii) The proof of (4.38) uses the same arguments as the proof of Lemma 3.5. \square

828 *Remark 4.17* ((4.37b)). The right-hand side of (4.37b) does not depend on the
829 face-based functions $w_{\partial K}$. This identity will replace the argument in [16, 18] invoking
830 the continuity of the normal component of $\boldsymbol{\sigma}(u)$ at the mesh interfaces, which makes
831 sense only when the exact solution is smooth enough, say $\boldsymbol{\sigma}(u) \in \mathbf{H}^r(D)$ with $r > \frac{1}{2}$. \square

832 Let $V_{\sharp} := V_s + P_{k+1}^b(\mathcal{T}_h)$ be equipped with the seminorm $\|v\|_{V_{\sharp}} := |v|_{\lambda,p,q}$ defined
 833 in (4.2). Notice that $\|v\|_{V_{\sharp}} = 0$ implies that $v = 0$ if v has zero mean-value in each
 834 mesh cell $K \in \mathcal{T}_h$; this is the case for instance if one takes $v = u - \mathcal{E}_h^{k+1}(u)$. We define
 835 the consistency error $\delta_h : \hat{V}_{h,0}^k \rightarrow (\hat{V}_{h,0}^k)'$ by setting, for all $\hat{w}_h \in V_{h,0}^k$,

$$836 \quad (4.39) \quad \langle \delta_h(\hat{v}_h), \hat{w}_h \rangle_{(\hat{V}_{h,0}^k)', \hat{V}_{h,0}^k} := \hat{\ell}_h(\hat{w}_h) - \hat{a}_h(\hat{v}_h, \hat{w}_h).$$

837 We define global counterparts of the local operators \mathbf{R}_K^{k+1} , $\hat{\mathcal{I}}_K^k$, and \mathcal{E}_K^{k+1} , namely
 838 $\mathbf{R}_h^{k+1} : \hat{V}_{h,0}^k \rightarrow P_{k+1}^b(\mathcal{T}_h)$, $\hat{\mathcal{I}}_h^k : H^1(D) \rightarrow \hat{V}_{h,0}^k$, and $\mathcal{E}_h^{k+1} : H^1(D) \rightarrow P_{k+1}^b(\mathcal{T}_h)$, by
 839 setting $\mathbf{R}_h^{k+1}(\hat{v}_h)|_K := \mathbf{R}_K^{k+1}(\hat{v}_K)$, $\hat{\mathcal{I}}_h^k(v)|_K := \hat{\mathcal{I}}_K^k(v|_K)$, and $\mathcal{E}_h^{k+1}(v)|_K := \mathcal{E}_K^{k+1}(v|_K)$,
 840 for all $\hat{v}_h \in \hat{V}_{h,0}^k$, all $v \in H^1(D)$, and all $K \in \mathcal{T}_h$.

841 LEMMA 4.18 (Consistency/boundedness). *There is ω_{\sharp} , uniform w.r.t. $h \in \mathcal{H}$, λ ,*
 842 *and $u \in V_s$, but depending on p and q , s.t.*

$$843 \quad (4.40) \quad \|\delta_h(\hat{\mathcal{I}}_h^k(u))\|_{(\hat{V}_{h,0}^k)'} \leq \omega_{\sharp} \|u - \mathcal{E}_h^{k+1}(u)\|_{V_{\sharp}}. \quad \square$$

844 *Proof.* Since $\sigma(u) = -\lambda \nabla u$, $\nabla \cdot \sigma(u) = f$, and $u \in V_s$, the identity (4.37b) yields
 845 $\hat{\ell}_h(\hat{w}_h) = \sum_{K \in \mathcal{T}_h} (f, w_K)_{L^2(K)} = \sum_{K \in \mathcal{T}_h} a_K(u, w_K) + n_{\sharp}(u, \hat{w}_h)$, where $a_K(u, w_K) :=$
 846 $\int_K -\sigma(u) \cdot \nabla w_K \, dx$. Using the definition of \hat{a}_h in (4.33), then the identity $\mathbf{R}_K^{k+1} \circ \hat{\mathcal{I}}_K^k =$
 847 \mathcal{E}_K^{k+1} (see (4.35a)), and finally (4.37a) with $v_h = \mathcal{E}_h^{k+1}(u)$, we obtain

$$848 \quad \hat{a}_h(\hat{\mathcal{I}}_h^k(u), \hat{w}_h) = \sum_{K \in \mathcal{T}_h} a_K(\mathcal{E}_K^{k+1}(u), w_K) + n_{\sharp}(\mathcal{E}_h^{k+1}(u), \hat{w}_h)$$

$$849 \quad + \sum_{K \in \mathcal{T}_h} \lambda_K (h_{\partial K}^{-1} \mathbf{S}_{\partial K}^k(\hat{\mathcal{I}}_K^k(u)), \mathbf{S}_{\partial K}^k(\hat{w}_K))_{L^2(\partial K)}.$$

851 Subtracting these two identities and using the definition of $\mathcal{E}_K^{k+1}(u)$, which implies that
 852 $a_K(u - \mathcal{E}_K^{k+1}(u), w_K) = 0$, for all $K \in \mathcal{T}_h$, leads to $\langle \delta_h(\hat{\mathcal{I}}_h^k(u)), \hat{w}_h \rangle_{(\hat{V}_{h,0}^k)', \hat{V}_{h,0}^k} = \mathfrak{T}_1 + \mathfrak{T}_2$
 853 with

$$854 \quad \mathfrak{T}_1 := n_{\sharp}(u - \mathcal{E}_h^{k+1}(u), \hat{w}_h), \quad \mathfrak{T}_2 := - \sum_{K \in \mathcal{T}_h} \lambda_K (h_{\partial K}^{-1} \mathbf{S}_{\partial K}^k(\hat{\mathcal{I}}_K^k(u)), \mathbf{S}_{\partial K}^k(\hat{w}_K))_{L^2(\partial K)}.$$

856 We invoke (4.38) to bound \mathfrak{T}_1 and observe that $\sum_{K \in \mathcal{T}_h} \lambda_K h_K^{-1} \|w_K - w_{\partial K}\|_{L^2(\partial K)}^2 \leq$
 857 $\|\hat{w}_h\|_{\hat{V}_{h,0}^k}^2$ owing to (4.28). For the bound on \mathfrak{T}_2 , we proceed as in [16, 18]. \square

858 THEOREM 4.19 (Error estimate). *Let u solve (2.1) and \hat{u}_h solve (4.34) with \hat{u}_h*
 859 *and $\hat{\ell}_h$ defined in (4.33). Assume that there is $r > 0$ s.t. $u \in H^{1+r}(D)$. There is*
 860 *c , uniform w.r.t. $h \in \mathcal{H}$, λ , and $u \in H^{1+r}(D)$, but depending on r , s.t. the following*
 861 *holds true:*

$$862 \quad (4.41) \quad \|\lambda^{\frac{1}{2}} \nabla_h(u - \mathbf{R}_h^{k+1}(\hat{u}_h))\|_{L^2(D)} \leq c \|u - \mathcal{E}_h^{k+1}(u)\|_{V_{\sharp}}.$$

863 Moreover, letting $t := \min(r, k+1)$, $\chi_t = 1$ if $t \leq 1$ and $\chi_t = 0$ if $t > 1$, we have

$$865 \quad (4.42) \quad \|\lambda^{\frac{1}{2}} \nabla_h(u - \mathbf{R}_h^{k+1}(\hat{u}_h))\|_{L^2(D)}$$

$$866 \quad \leq c \left(\sum_{K \in \mathcal{T}_h} \lambda_K h_K^{2t} |u|_{H^{1+t}(K)}^2 + \frac{\chi_t}{\lambda_K} h_K^{2d(\frac{d+2}{2d}-\frac{1}{q})} \|f\|_{L^q(K)}^2 \right)^{\frac{1}{2}}. \quad \square$$

867

868 *Proof.* (i) We adapt the proof of Lemma 4.2 to exploit the convergence order of the
 869 reconstruction operator. Let us set $\hat{\zeta}_h^k := \hat{\mathcal{I}}_K^k(u) - \hat{u}_h \in \hat{V}_{h,0}^k$ so that $\hat{\zeta}_K^k = \hat{\mathcal{I}}_K^k(u|_K) - \hat{u}_K$
 870 for all $K \in \mathcal{T}_h$. The coercivity property from Lemma 4.14 and the definition of the
 871 consistency error imply that

$$872 \quad \alpha \|\lambda^{\frac{1}{2}} \nabla_h \mathbf{R}_h^{k+1}(\hat{\zeta}_h^k)\|_{\mathbf{L}^2(D)}^2 \leq \frac{\hat{a}_h(\hat{\zeta}_h^k, \hat{\zeta}_h^k)}{\|\hat{\zeta}_h^k\|_{\hat{V}_{h,0}^k}^2} \|\lambda^{\frac{1}{2}} \nabla_h \mathbf{R}_h^{k+1}(\hat{\zeta}_h^k)\|_{\mathbf{L}^2(D)}^2$$

$$873 \quad \leq \frac{(\hat{a}_h(\hat{\zeta}_h^k, \hat{\zeta}_h^k))^2}{\|\hat{\zeta}_h^k\|_{\hat{V}_{h,0}^k}^2} = \frac{\langle \delta_h(\hat{\mathcal{I}}_h^k(u)), \hat{\zeta}_h^k \rangle_{(\hat{V}_{h,0}^k)', \hat{V}_{h,0}^k}}{\|\hat{\zeta}_h^k\|_{\hat{V}_{h,0}^k}^2} \leq \|\delta_h(\hat{\mathcal{I}}_h^k(u))\|_{(\hat{V}_{h,0}^k)'}^2.$$

875 Then, lemma 4.18 yields $\|\lambda^{\frac{1}{2}} \nabla \mathbf{R}_h^{k+1}(\hat{\zeta}_h^k)\|_{\mathbf{L}^2(D)} \leq c \|u - \mathcal{E}_h^{k+1}(u)\|_{V_{\sharp}}$. Moreover, since
 876 $\mathbf{R}_K^{k+1}(\hat{\mathcal{I}}_K^k(u)) = \mathcal{E}_K^{k+1}(u)$ for all $K \in \mathcal{T}_h$, see (4.35a), we have

$$877 \quad u - \mathbf{R}_h^{k+1}(\hat{u}_h) = u - \mathcal{E}_h^{k+1}(u) + \mathbf{R}_h^{k+1}(\hat{\zeta}_h^k).$$

878 The estimate (4.41) is now a consequence of the triangle inequality.

879 (ii) We now prove (4.42). Let us set $\eta^{k+1} := u - \mathcal{E}_h^{k+1}(u)$. We need to bound
 880 $\|\eta^{k+1}\|_{V_{\sharp}} = |\eta^{k+1}|_{\lambda,p,q}$, i.e., we must estimate $\|\nabla \eta^{k+1}\|_{\mathbf{L}^2(K)}$, $h_K^{d(\frac{1}{2}-\frac{1}{p})} \|\nabla \eta^{k+1}\|_{\mathbf{L}^p(K)}$,
 881 and $h_K^{d(\frac{d+2}{2d}-\frac{1}{q})} \|\Delta \eta^{k+1}\|_{L^q(K)}$ (see (4.2)). Owing to the optimality property of the
 882 elliptic projection and the approximation properties of Π_K^{k+1} , we have

$$883 \quad \|\nabla \eta^{k+1}\|_{\mathbf{L}^2(K)} \leq \|\nabla(u - \Pi_K^{k+1}(u))\|_{\mathbf{L}^2(K)} \leq c h_K^t |u|_{H^{1+t}(K)}.$$

884 for $t = \min(r, k+1)$. Let us now consider the other two terms. Let $\ell := \lceil t \rceil$, so that
 885 $t \leq \ell \leq 1+t$. Notice also that $\ell \leq k+1$, and $\ell \geq 1$ since we assumed that $r > 0$.
 886 Let us set $\eta^\ell := u - \mathcal{E}_h^\ell(u)$, then $\|\nabla \eta^\ell\|_{\mathbf{L}^2(K)} \leq c h_K^t |u|_{H^{1+t}(K)}$. Invoking the triangle
 887 inequality, an inverse inequality, and the triangle inequality again, we infer that

$$888 \quad h_K^{d(\frac{1}{2}-\frac{1}{p})} \|\nabla \eta^{k+1}\|_{\mathbf{L}^p(K)} \leq h_K^{d(\frac{1}{2}-\frac{1}{p})} \|\nabla \eta^\ell\|_{\mathbf{L}^p(K)} + c (\|\nabla \eta^{k+1}\|_{\mathbf{L}^2(K)} + \|\nabla \eta^\ell\|_{\mathbf{L}^2(K)}),$$

890 and the two terms between the parentheses are bounded by $c h_K^t |u|_{H^{1+t}(K)}$. Moreover,
 891 invoking (4.16), we obtain

$$892 \quad h_K^{d(\frac{1}{2}-\frac{1}{p})} \|\nabla \eta^\ell\|_{\mathbf{L}^p(K)} \leq c (\|\nabla \eta^\ell\|_{\mathbf{L}^2(K)} + h_K^t |\nabla \eta^\ell|_{\mathbf{H}^t(K)})$$

$$893 \quad = c (\|\nabla \eta^\ell\|_{\mathbf{L}^2(K)} + h_K^t |u|_{H^{1+t}(K)}) \leq c' h_K^t |u|_{H^{1+t}(K)},$$

895 since $t \leq \ell$. Similarly, we have

$$896 \quad h_K^{d(\frac{d+2}{2d}-\frac{1}{q})} \|\Delta \eta^{k+1}\|_{L^q(K)} \leq h_K^{d(\frac{d+2}{2d}-\frac{1}{q})} \|\Delta \eta^\ell\|_{L^q(K)} + c (\|\nabla \eta^{k+1}\|_{\mathbf{L}^2(K)} + \|\nabla \eta^\ell\|_{\mathbf{L}^2(K)}).$$

898 It remains to estimate $h_K^{d(\frac{d+2}{2d}-\frac{1}{q})} \|\Delta \eta^\ell\|_{L^q(K)}$. We proceed as in the end of the proof
 899 of Theorem 4.10. If $t \leq 1$ (so that $\chi_t = 1$), we have $\ell = 1$, and we infer that

$$900 \quad h_K^{d(\frac{d+2}{2d}-\frac{1}{q})} \|\Delta \eta^\ell\|_{L^q(K)} = \lambda_K^{-1} h_K^{d(\frac{d+2}{2d}-\frac{1}{q})} \|f\|_{L^q(K)}.$$

901 Otherwise, we have $t > 1$ (so that $\chi_t = 0$) and $\ell \geq 2$, and we take $q = 2$. Then,
 902 using the triangle inequality, an inverse inequality, and the triangle inequality again,

903 we obtain

$$904 \quad h_K \|\Delta \eta^\ell\|_{L^q(K)} \leq h_K \|\Delta(u - \Pi_K^\ell(u))\|_{L^q(K)} \\ 905 \quad + c(\|\nabla(u - \Pi_K^\ell(u))\|_{L^2(K)} + \|\nabla \eta^\ell\|_{L^2(K)}),$$

907 where Π_K^ℓ is the L^2 -orthogonal projection onto $\mathbb{P}_{\ell,d}$. We conclude by invoking the
908 approximation properties of Π_K^ℓ , recalling that $\|\nabla \eta^\ell\|_{L^2(K)} \leq ch_K^\ell |u|_{H^{1+\ell}(K)}$. \square

909 *Remark 4.20* (Supercloseness). Step (i) in the above proof actually shows that
910 $\|\hat{\zeta}_h^k\|_{\hat{V}_{h,0}^k} \leq c\|u - \mathcal{E}_h^{k+1}(u)\|_{V_2}$. Since $\zeta_K^k = \Pi_K^k(u) - u_K$ for all $K \in \mathcal{T}_h$, this implies the
911 supercloseness bound $(\sum_{K \in \mathcal{T}_h} \lambda_K \|\nabla(\Pi_K^k(u) - u_K)\|_{L^2(K)}^2)^{\frac{1}{2}} \leq c\|u - \mathcal{E}_h^{k+1}(u)\|_{V_2}$.

912 **5. Extensions to Maxwell's equations.** The various techniques presented in
913 this paper can be extended to the context of Maxwell's equations, since arguments
914 similar to those exposed in §3 can be deployed to define the tangential trace of vectors
915 fields on a face of K . Without going into the details, we show in this section how that
916 can be done.

917 **5.1. Lifting and tangential trace.** Let p, q be real numbers satisfying (3.2),
918 and let $\tilde{p} \in (2, p]$ be such that $q \geq \frac{\tilde{p}d}{\tilde{p}+d}$. Let K be a cell in \mathcal{T}_h , and let $F \in \mathcal{F}_K$ be a
919 face of K . Following [25], we introduce the space

$$920 \quad (5.1) \quad \mathbf{Y}^c(F) := \{\phi \in \mathbf{W}^{\frac{1}{\tilde{p}}, \tilde{p}'}(F) \mid \phi \cdot \mathbf{n}_F = 0\},$$

922 which we equip with the norm $\|\phi\|_{\mathbf{Y}^c(F)} := \|\phi\|_{L^{\tilde{p}'}(F)} + h_F^{\frac{1}{\tilde{p}}} |\phi|_{\mathbf{W}^{\frac{1}{\tilde{p}}, \tilde{p}'}(F)}$. Then the
923 following result can be established by proceeding as in the proof of Lemma 3.1.

924 **LEMMA 5.1** (Face-to-cell Lifting). *There exist a constant c , uniform w.r.t. h , but*
925 *depending on p and q , and a lifting operator $E_F^K : \mathbf{Y}^c(F) \rightarrow \mathbf{W}^{1, \tilde{p}'}(K)$ such that the*
926 *following holds true for any $\phi \in \mathbf{Y}^c(F)$: $E_F^K(\phi)|_{\partial K \setminus F} = \mathbf{0}$, $E_F^K(\phi)|_F = \phi$, and*

$$927 \quad (5.2) \quad |E_F^K(\phi)|_{\mathbf{W}^{1, \tilde{p}'}(K)} + h_K^{-1+d(\frac{1}{q}-\frac{1}{\tilde{p}})} \|E_F^K(\phi)\|_{L^{q'}(K)} \leq ch_K^{-\frac{1}{\tilde{p}}+d(\frac{1}{\tilde{p}}-\frac{1}{p})} \|\phi\|_{\mathbf{Y}^c(F)}. \quad \square$$

928 With this lifting operator in hand, we can define an extension to the notion of
929 the tangential trace on F of a vector field. To this end, we introduce the functional
930 space

$$931 \quad (5.3) \quad \mathbf{S}^c(K) := \{\tau \in L^p(K) \mid \nabla \times \tau \in L^q(K)\},$$

933 where the superscript c refers to the fact that the tangential trace is related to the
934 curl operator. We equip $\mathbf{S}^c(K)$ with the following dimensionally-consistent norm:

$$935 \quad (5.4) \quad \|\tau\|_{\mathbf{S}^c(K)} := \|\tau\|_{L^p(K)} + h_K^{1+d(\frac{1}{p}-\frac{1}{q})} \|\nabla \times \tau\|_{L^q(K)}.$$

936 We now define the tangential trace of any field τ in $\mathbf{S}^c(K)$ on the face F of K to be
937 the linear form $(\tau \times \mathbf{n}_K)|_F \in \mathbf{Y}^c(F)'$ such that

$$938 \quad (5.5) \quad \langle (\tau \times \mathbf{n}_K)|_F, \phi \rangle_F := \int_K \left(\tau \cdot \nabla \times E_F^K(\phi) - (\nabla \times \tau) \cdot E_F^K(\phi) \right) dx,$$

939 for all $\phi \in \mathbf{Y}^c(F)$, where $\langle \cdot, \cdot \rangle_F$ now denotes the duality pairing between $\mathbf{Y}^c(F)'$
940 and $\mathbf{Y}^c(F)$. Note that the right-hand side of (5.5) is well-defined owing to Hölder's
941 inequality and (5.2).

942 The discretization now involves the vector-valued broken finite element space

$$943 \quad (5.6) \quad \mathbf{P}_k^b(\mathcal{T}_h) = \{\mathbf{v}_h \in \mathbf{L}^\infty(D) \mid \mathbf{v}_h|_K \in \mathbf{P}_K, \forall K \in \mathcal{T}_h\},$$

944 where $\mathbf{P}_K := (\psi_K)^{-1}(\widehat{\mathbf{P}}) \subset \mathbf{W}^{k+1,\infty}(K)$, $(\widehat{K}, \widehat{\mathbf{P}}, \widehat{\Sigma})$ is the reference element, and ψ_K
 945 is an appropriate transformation. For instance, one can take $\psi_K(\mathbf{v}) = \psi_K^s(\mathbf{v}) := \mathbf{v} \circ \mathbf{T}_K$
 946 for continuous Lagrange elements or for dG approximation; one can also take $\psi_K(\mathbf{v}) =$
 947 $\psi_K^c(\mathbf{v}) := \mathbb{J}_K^T(\mathbf{v} \circ \mathbf{T}_K)$ for edge elements (ψ_K^c is covariant Piola transformation and
 948 \mathbb{J}_K the Jacobian of the geometric mapping). For any face $F \in \mathcal{F}_K$, we denote by \mathbf{P}_F
 949 the trace of \mathbf{P}_K on F . The following result is the counterpart of Lemma 3.2.

950 LEMMA 5.2 (Bound on tangential component). *There exists a constant c , uni-*
 951 *form w.r.t. h , but depending on p and q , so that the following estimate holds true for*
 952 *all $\mathbf{v} \in \mathbf{S}^c(K)$,*

$$953 \quad (5.7) \quad \|(\mathbf{v} \times \mathbf{n}_K)|_F\|_{\mathbf{Y}^c(F)'} \leq c h_K^{-\frac{1}{p} + d(\frac{1}{p} - \frac{1}{p})} \|\mathbf{v}\|_{\mathbf{S}^c(K)}.$$

954 Moreover, we have

$$955 \quad (5.8) \quad | \langle (\mathbf{v} \times \mathbf{n}_K)|_F, \boldsymbol{\phi}_h \rangle | \leq c h_K^{d(\frac{1}{2} - \frac{1}{p})} \|\mathbf{v}\|_{\mathbf{S}^c(K)} h_F^{-\frac{1}{2}} \|\boldsymbol{\phi}_h\|_{L^2(F)},$$

956 for all $\boldsymbol{\phi}_h \in \mathbf{P}_F$ s.t. $\boldsymbol{\phi}_h \cdot \mathbf{n}_F = 0$, all $K \in \mathcal{T}_h$, and all $F \in \mathcal{F}_K$. \square

957 Lemma 5.2 is essential for the error analysis of nonconforming approximation
 958 techniques of Maxwell's equations. It is a generalization of Bonito et al. [8, Lem. A3]
 959 and Buffa and Perugia [9, Lem. 8.2].

960 **5.2. Definition of n_{\sharp}^c and key identities.** The consistency analysis of Nitsche's
 961 boundary penalty method and of the dG approximation applied to Maxwell's equations
 962 can be done by introducing a bilinear form n_{\sharp} as in §3. We henceforth assume that
 963 the space dimension is either $d = 2$ or $d = 3$.

964 We define the notion of diffusive flux by introducing $\boldsymbol{\sigma} : \mathbf{H}(\text{curl}; D) \rightarrow \mathbf{L}^2(D)$
 965 such that $\boldsymbol{\sigma}(\mathbf{v}) := \lambda \nabla \times \mathbf{v}$, for any $\mathbf{v} \in \mathbf{H}(\text{curl}; D)$. Here, the diffusivity λ is either
 966 the reciprocal of the magnetic permeability or the reciprocal of electrical conductiv-
 967 ity, depending whether one works with the electric field or the magnetic field. The
 968 diffusivity is assumed to satisfy the hypotheses introduced in Section 2. We further
 969 define

$$970 \quad (5.9) \quad \mathbf{V}_s := \{\mathbf{v} \in \mathbf{H}(\text{curl}; D) \mid \boldsymbol{\sigma}(\mathbf{v}) \in \mathbf{L}^p(D), \nabla \times \boldsymbol{\sigma}(\mathbf{v}) \in \mathbf{L}^q(D)\},$$

971 and set $\mathbf{V}_{\sharp} := \mathbf{V}_s + \mathbf{P}_k^b(\mathcal{T}_h)$.

972 We adopt the same notation as in §3. Recall that for any $K \in \mathcal{T}_h$ and any $F \in \mathcal{F}_K$,
 973 we have defined $\epsilon_{K,F} = \mathbf{n}_F \cdot \mathbf{n}_K = \pm 1$. We consider arbitrary weights $\theta_{K,F}$ satisfying
 974 (3.13). We introduce the bilinear form $n_{\sharp}^c : (\mathbf{V}_s + \mathbf{P}_k^b(\mathcal{T}_h)) \times \mathbf{P}_k^b(\mathcal{T}_h) \rightarrow \mathbb{R}$ defined as
 975 follows:

$$976 \quad (5.10) \quad n_{\sharp}^c(\mathbf{v}, \mathbf{w}_h) := \sum_{F \in \mathcal{F}_h} \sum_{K \in \mathcal{T}_F} \epsilon_{K,F} \theta_{K,F} \langle (\boldsymbol{\sigma}(\mathbf{v})|_K \times \mathbf{n}_K)|_F, [\Pi_F(\mathbf{w}_h)] \rangle_F,$$

977
 978 where Π_F is the ℓ^2 -orthogonal projection onto the hyperplane tangent to F , i.e.,
 979 $\Pi_F(\mathbf{b}_h) := \mathbf{b}_h - (\mathbf{b}_h \cdot \mathbf{n}_K) \mathbf{n}_K = \mathbf{n}_K \times (\mathbf{b}_h \times \mathbf{n}_K)$. Notice that (5.10) is meaningful since
 980 $\Pi_F(\mathbf{b}_h)|_F$ is in $\mathbf{W}^{\frac{1}{p}, \vec{\nu}'}(F)$ and $\Pi_F(\mathbf{b}_h) \cdot \mathbf{n}_F = 0$, i.e., $\Pi_F(\mathbf{b}_h) \in \mathbf{Y}^c(F)$ for any $F \in \mathcal{F}_h$.
 981 The following result is the counterpart of Lemma 3.3.

982 LEMMA 5.3 (Identities for n_{\sharp}^c). *The following holds true for any choice of weights*
 983 $\{\theta_{K,F}\}_{F \in \mathcal{F}_h, K \in \mathcal{T}_F}$ and for all $\mathbf{w}_h \in \mathbf{P}_k^b(\mathcal{T}_h)$, all $\mathbf{v}_h \in \mathbf{P}_k^b(\mathcal{T}_h)$, and all $\mathbf{v} \in \mathbf{V}_s$:

984 (5.11a)
$$n_{\sharp}^c(\mathbf{v}_h, \mathbf{w}_h) = \sum_{F \in \mathcal{F}_h} \int_F (\{\boldsymbol{\sigma}(\mathbf{v}_h)\}_{\theta} \times \mathbf{n}_F) \cdot [\Pi_F(\mathbf{w}_h)] \, ds,$$

985 (5.11b)
$$n_{\sharp}^c(\mathbf{v}, \mathbf{w}_h) = \sum_{K \in \mathcal{T}_h} \int_K (\boldsymbol{\sigma}(\mathbf{v}) \cdot \nabla \times \mathbf{w}_{h|K} - (\nabla \times \boldsymbol{\sigma}(\mathbf{v})) \cdot \mathbf{w}_{h|K}) \, dx. \quad \square$$

 986

987 *Proof.* The proof is similar to that of Lemma 3.3. The proof of (5.11a) is quasi-
 988 identical to that of (3.16a). For the proof of (5.11b), one invokes the mollifying
 989 operators $\mathcal{K}_{\delta}^c : \mathbf{L}^1(D) \rightarrow \mathbf{C}^\infty(\overline{D})$ and $\mathcal{K}_{\delta}^d : \mathbf{L}^1(D) \rightarrow \mathbf{C}^\infty(\overline{D})$ introduced in [22, §3.2].
 990 These two operators satisfy the following key commuting property:

991 (5.12)
$$\nabla \times (\mathcal{K}_{\delta}^c(\boldsymbol{\tau})) = \mathcal{K}_{\delta}^d(\nabla \times \boldsymbol{\tau}),$$

992 for all $\boldsymbol{\tau} \in \mathbf{L}^1(D)$ s.t. $\nabla \times \boldsymbol{\tau} \in \mathbf{L}^1(D)$. Then one uses the identities $[\mathbf{v} \times \Pi_F(\mathbf{w})] =$
 993 $\{\mathbf{v}\}_{\theta} \times [\Pi_F(\mathbf{w})] + [\mathbf{v}] \times \{\Pi_F(\mathbf{w})\}_{\theta}$, $\mathbf{n}_K \times \Pi_F(\mathbf{w}_h) = \mathbf{n}_K \times \mathbf{w}_h$, and $\nabla \cdot (\mathbf{w}_h \times \boldsymbol{\sigma}(\mathbf{v})) =$
 994 $\boldsymbol{\sigma}(\mathbf{v}) \cdot (\nabla \times \mathbf{w}_h) - \mathbf{w}_h \cdot (\nabla \times \boldsymbol{\sigma}(\mathbf{v}))$. \square

995 We now establish the boundedness of the bilinear form n_{\sharp}^c . Since $\boldsymbol{\sigma}(\mathbf{v})|_K \in \mathcal{S}^c(K)$
 996 for all $K \in \mathcal{T}_h$ and all $\mathbf{v} \in \mathbf{V}_s + \mathbf{P}_k^b(\mathcal{T}_h)$, we equip the space $\mathbf{V}_s + \mathbf{P}_k^b(\mathcal{T}_h)$ with the
 997 seminorm
 998

999 (5.13)
$$|\mathbf{v}|_{n_{\sharp}^c}^2 := \sum_{K \in \mathcal{T}_h} \lambda_K^{-1} \left(h_K^{2d(\frac{1}{2} - \frac{1}{p})} \|\boldsymbol{\sigma}(\mathbf{v})|_K\|_{\mathbf{L}^p(K)}^2 \right. \\ \left. + h_K^{2d(\frac{2+d}{2d} - \frac{1}{q})} \|\nabla \times \boldsymbol{\sigma}(\mathbf{v})|_K\|_{\mathbf{L}^q(K)}^2 \right).$$

1000
1001

1002 LEMMA 5.4 (Boundedness of n_{\sharp}^c). *With the weights defined in (3.19) and λ_F de-*
 1003 *fined in (3.20) for all $F \in \mathcal{F}_h$, there is c , uniform w.r.t. $h \in \mathcal{H}$ and λ , but depending*
 1004 *on p and q , s.t. the following holds true for all $\mathbf{v} \in \mathbf{V}_s + \mathbf{P}_k^b(\mathcal{T}_h)$ and all $\mathbf{w}_h \in \mathbf{P}_k^b(\mathcal{T}_h)$:*

1005 (5.14)
$$|n_{\sharp}^c(\mathbf{v}, \mathbf{w}_h)| \leq c |\mathbf{v}|_{n_{\sharp}^c} \left(\sum_{F \in \mathcal{F}_h} \lambda_F h_F^{-1} \|\Pi_F(\mathbf{w}_h)\|_{\mathbf{L}^2(F)}^2 \right)^{\frac{1}{2}}. \quad \square$$

1006 With the above tools in hand, one can revisit Buffa and Perugia [9] and greatly
 1007 simplify the analysis of the dG approximation of Maxwell's equations. One can also
 1008 extend the work in [24] and analyze Nitsche's boundary penalty technique with edge
 1009 elements; one can also revisit Bonito et al. [7], where Nitsche's boundary penalty
 1010 technique has been used in conjunction with Lagrange elements. In all the cases one
 1011 then obtains error estimates that are robust with respect to the diffusivity contrast.

1012 **References.**

1013 [1] C. Amrouche, C. Bernardi, M. Dauge, and V. Girault. Vector potentials in
 1014 three-dimensional non-smooth domains. *Math. Methods Appl. Sci.*, 21(9):823–
 1015 864, 1998.
 1016 [2] D. N. Arnold. An interior penalty finite element method with discontinuous
 1017 elements. *SIAM J. Numer. Anal.*, 19:742–760, 1982.
 1018 [3] S. Badia, R. Codina, T. Gudi, and J. Guzmán. Error analysis of discontinuous
 1019 Galerkin methods for the Stokes problem under minimal regularity. *IMA J.*
 1020 *Numer. Anal.*, 34(2):800–819, 2014.

- 1021 [4] C. Bernardi and V. Girault. A local regularization operator for triangular and
1022 quadrilateral finite elements. *SIAM J. Numer. Anal.*, 35(5):1893–1916, 1998.
- 1023 [5] C. Bernardi and F. Hecht. Error indicators for the mortar finite element dis-
1024 cretization of the Laplace equation. *Math. Comp.*, 71(240):1371–1403, 2002.
- 1025 [6] C. Bernardi and R. Verfürth. Adaptive finite element methods for elliptic equa-
1026 tions with non-smooth coefficients. *Numer. Math.*, 85(4):579–608, 2000.
- 1027 [7] A. Bonito, J.-L. Guermond, and F. Luddens. Regularity of the maxwell equations
1028 in heterogeneous media and lipschitz domains. *J. Math. Anal. Appl.*, 408:498–
1029 512, 2013.
- 1030 [8] A. Bonito, J.-L. Guermond, and F. Luddens. An interior penalty method with
1031 C^0 finite elements for the approximation of the Maxwell equations in hetero-
1032 geneous media: convergence analysis with minimal regularity. *ESAIM Math.*
1033 *Model. Numer. Anal.*, 50(5):1457–1489, 2016.
- 1034 [9] A. Buffa and I. Perugia. Discontinuous Galerkin approximation of the Maxwell
1035 eigenproblem. *SIAM J. Numer. Anal.*, 44(5):2198–2226, 2006.
- 1036 [10] E. Burman and P. Zunino. A domain decomposition method for partial differ-
1037 ential equations with non-negative form based on interior penalties. *SIAM J.*
1038 *Numer. Anal.*, 44:1612–1638, 2006.
- 1039 [11] Z. Cai, X. Ye, and S. Zhang. Discontinuous Galerkin finite element methods for
1040 interface problems: a priori and a posteriori error estimations. *SIAM J. Numer.*
1041 *Anal.*, 49(5):1761–1787, 2011.
- 1042 [12] C. Carstensen and M. Schedensack. Medius analysis and comparison results for
1043 first-order finite element methods in linear elasticity. *IMA J. Numer. Anal.*, 35
1044 (4):1591–1621, 2015.
- 1045 [13] B. Cockburn, D. A. Di Pietro, and A. Ern. Bridging the Hybrid High-Order and
1046 Hybridizable Discontinuous Galerkin methods. *ESAIM: Math. Model Numer.*
1047 *Anal. (M2AN)*, 50(3):635–650, 2016.
- 1048 [14] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element
1049 methods for solving the stationary Stokes equations. I. *Rev. Française Automat.*
1050 *Informat. Recherche Opérationnelle Sér. Rouge*, 7(R-3):33–75, 1973.
- 1051 [15] D. A. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin*
1052 *Methods*, volume 69 of *Mathématiques & Applications*. Springer-Verlag, Berlin,
1053 2012.
- 1054 [16] D. A. Di Pietro and A. Ern. A Hybrid High-Order locking-free method for linear
1055 elasticity on general meshes. *Comput. Meth. Appl. Mech. Engrg.*, 283:1–21, 2015.
- 1056 [17] D. A. Di Pietro, A. Ern, and J.-L. Guermond. Discontinuous Galerkin methods
1057 for anisotropic semi-definite diffusion with advection. *SIAM J. Numer. Anal.*, 46
1058 (2):805–831, 2008.
- 1059 [18] D. A. Di Pietro, A. Ern, and S. Lemaire. An arbitrary-order and compact-
1060 stencil discretization of diffusion on general meshes based on local reconstruction
1061 operators. *Comput. Meth. Appl. Math.*, 14(4):461–472, 2014.
- 1062 [19] M. Dryja. On discontinuous Galerkin methods for elliptic problems with discon-
1063 tinuous coefficients. *Comput. Methods Appl. Math.*, 3(1):76–85, 2003.
- 1064 [20] M. Dryja, J. Galvis, and M. Sarkis. BDDC methods for discontinuous Galerkin
1065 discretization of elliptic problems. *J. Complexity*, 23(4-6):715–739, 2007.
- 1066 [21] A. Ern and J.-L. Guermond. Discontinuous Galerkin methods for Friedrichs’
1067 systems. I. General theory. *SIAM J. Numer. Anal.*, 44(2):753–778, 2006.
- 1068 [22] A. Ern and J.-L. Guermond. Mollification in strongly Lipschitz domains with
1069 application to continuous and discrete de Rham complexes. *Comput. Methods*
1070 *Appl. Math.*, 16(1):51–75, 2016.

- 1071 [23] A. Ern and J.-L. Guermond. Finite element quasi-interpolation and best approx-
1072 imation. *M2AN Math. Model. Numer. Anal.*, 51(4):1367–1385, 2017.
- 1073 [24] A. Ern and J.-L. Guermond. Analysis of the edge finite element approximation
1074 of the Maxwell equations with low regularity solutions. *Comput. Math. Appl.*, 75
1075 (3):918–932, 2018.
- 1076 [25] A. Ern and J.-L. Guermond. Abstract nonconforming error estimates and appli-
1077 cation to boundary penalty methods for diffusion equations and time-harmonic
1078 Maxwell’s equations. *Comput. Methods Appl. Math.*, 18(3):451–475, 2018.
- 1079 [26] A. Ern, A. F. Stephansen, and P. Zunino. A discontinuous Galerkin method
1080 with weighted averages for advection-diffusion equations with locally small and
1081 anisotropic diffusivity. *IMA J. Numer. Anal.*, 29(2):235–256, 2009.
- 1082 [27] E. Gagliardo. Caratterizzazioni delle tracce sulla frontiera relative ad alcune
1083 classi di funzioni in n variabili. *Rend. Sem. Mat. Univ. Padova*, 27:284–305,
1084 1957.
- 1085 [28] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs
1086 and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston,
1087 MA, 1985.
- 1088 [29] T. Gudi. A new error analysis for discontinuous finite element methods for linear
1089 elliptic problems. *Math. Comp.*, 79(272):2169–2189, 2010.
- 1090 [30] F. Jochmann. An H^s -regularity result for the gradient of solutions to elliptic
1091 equations with mixed boundary conditions. *J. Math. Anal. Appl.*, 238:429–450,
1092 1999.
- 1093 [31] M. Li and S. Mao. A new a priori error analysis of nonconforming and mixed
1094 finite element methods. *Appl. Math. Lett.*, 26(1):32–37, 2013.
- 1095 [32] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei
1096 Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind.
1097 *Abh. Math. Sem. Univ. Hamburg*, 36:9–15, 1971.
- 1098 [33] J. Schöberl. Commuting quasi-interpolation operators for mixed finite elements.
1099 Technical Report ISC-01-10-MATH, Texas A&M University, 2001. URL [www.
1100 isc.tamu.edu/publications-reports/tr/0110.pdf](http://www.isc.tamu.edu/publications-reports/tr/0110.pdf).
- 1101 [34] A. Veeseer and P. Zanotti. Quasi-optimal nonconforming methods for symmetric
1102 elliptic problems. I—Abstract theory. *SIAM J. Numer. Anal.*, 56(3):1621–1642,
1103 2018.
- 1104 [35] A. Veeseer and P. Zanotti. Quasi-optimal nonconforming methods for symmetric
1105 elliptic problems. III—Discontinuous Galerkin and other interior penalty meth-
1106 ods. *SIAM J. Numer. Anal.*, 56(5):2871–2894, 2018.