© 2018 Society for Industrial and Applied Mathematics

QUASI-OPTIMAL NONCONFORMING METHODS FOR SYMMETRIC ELLIPTIC PROBLEMS. III—DISCONTINUOUS GALERKIN AND OTHER INTERIOR PENALTY METHODS[*]

ANDREAS VEESER[†] AND PIETRO ZANOTTI[†]

Abstract. We devise new variants of the following nonconforming finite element methods: discontinuous Galerkin methods of fixed arbitrary order for the Poisson problem, the Crouzeix–Raviart interior penalty method for linear elasticity, and the quadratic C^0 interior penalty method for the biharmonic problem. Each variant differs from the original method only in the discretization of the right-hand side. Before applying the load functional, a linear operator transforms nonconforming discrete test functions into conforming functions such that stability and consistency are improved. The new variants are thus quasi-optimal with respect to an extension of the energy norm. Furthermore, their quasi-optimality constants are uniformly bounded for shape regular meshes and tend to 1 as the penalty parameter increases.

Key words. quasi-optimality, discontinuous elements, Crouzeix–Raviart element, linear elasticity, C^0 interior penalty methods

AMS subject classifications. 65N30, 65N15, 65N12

DOI. 10.1137/17M1151675

1. Introduction. This article is the third in a series on quasi-optimal nonconforming methods for linear and symmetric elliptic problems. Here we apply the framework developed in the first part [23] to design and analyze quasi-optimal interior penalty methods. We illustrate our construction and main results in the case of approximating the Poisson problem with discontinuous linear elements via the symmetric interior penalty (SIP) method, which was first studied by Baker [5], Wheeler [26], and Arnold [1].

Let $u \in H_0^1(\Omega)$ be the weak solution of the Poisson problem

$$(1.1) \qquad -\Delta u = f \text{ in } \Omega, \qquad u = 0 \text{ on } \partial\Omega,$$

and let $\mathcal{M}$ be a simplicial, face-to-face mesh of the domain $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$. We write Σ for its skeleton and S_1^0 for the space of discontinuous $\mathcal{M}$-piecewise affine functions and use standard notation for piecewise gradients, jumps, averages, local meshsizes, etc. (cf. section 3.1 below). The SIP approximation $U \in S_1^0$ solves the discrete problem

$$(1.2) \qquad \forall \sigma \in S_1^0 \qquad b(U, \sigma) = \int_\Omega f\sigma,$$

where $f \in L^2(\Omega)$, the bilinear form $b := b_1 + b_2$ is given by

$$b_1(s, \sigma) := \int_\Omega \nabla_{\mathcal{M}} s \cdot \nabla_{\mathcal{M}} \sigma - \int_\Sigma \{\!\!\{\nabla s\}\!\!\} \cdot n \, [\![\sigma]\!] ,$$

$$b_2(s, \sigma) := \int_\Sigma \frac{\eta}{h} [\![s]\!] \, [\![\sigma]\!] - \int_\Sigma [\![s]\!] \{\!\!\{\nabla\sigma\}\!\!\} \cdot n,$$

[†]Dipartimento di Matematica, Università degli Studi di Milano, 20133 Milano, Italy (andreas.veeser@unimi.it, pietro.zanotti@unimi.it).

and the penalty parameter $\eta > 0$ is sufficiently large so that b is coercive. If we replace U by $u \in H_0^1(\Omega)$, we see that

$$(1.3) \quad u \in H^2(\Omega) \implies \forall \sigma \in S_1^0 \ b_1(u, \sigma) = \int_\Omega f\sigma, \quad \text{while} \quad \forall \sigma \in S_1^0 \ b_2(u, \sigma) = 0.$$

Hence, b_2 establishes symmetry and coercivity for shape regular meshes, without impairing the *consistency* provided by b_1. One therefore can derive the following abstract error bound (cf. Di Pietro and Ern [12, Theorem 4.17] and Gudi [14, section 3.2]):

$$(1.4) \qquad\qquad |u - U|_{1;\eta} \lesssim \inf_{s \in S_1^0} \left(|u - s|_{1;\eta}^2 + \mathrm{AG}(u - s)^2 \right)^{\frac{1}{2}},$$

where the norm

$$|v|_{1;\eta}^2 := \int_\Omega |\nabla_{\mathcal{M}} v|^2 + \int_\Sigma \frac{\eta}{h} |\, [\![v]\!] \,|^2, \quad v \in H_0^1(\Omega) + S_1^0,$$

extends the energy norm associated with (1.1) and is augmented with

$$\mathrm{AG}(v)^2 := \quad \int_\Sigma \frac{h}{\eta} |\{\!\{ \nabla v \}\!\}|^2 \quad \text{or} \quad \sum_{K \in \mathcal{M}} h_K^2 \inf_{c \in \mathbb{R}} \|\Delta v - c\|_{L^2(K)}^2$$

on the right-hand side. While (1.4) implies convergence of optimal order, the augmentation is an important difference to Céa's lemma. Indeed, since it is not meaningful for a generic solution in $H_0^1(\Omega)$, it cannot be bounded by the best error $\inf_{s \in S_1^0} |u - s|_{1;\eta}$ and, in addition, it restricts the applicability of (1.4). Notice that also the first part of (1.3) and the right-hand side of the discrete problem (1.2) require extra regularity of the solution and the load in (1.1), respectively. These observations suggest that the *stability* of the SIP method (1.2) is impaired. More precisely, if, e.g., the right-hand side cannot be boundedly extended to $H^{-1}(\Omega) = H_0^1(\Omega)'$, then $|U|_{1;\eta}$, or the error $|u - U|_{1;\eta}$, cannot be bounded in terms of $\|f\|_{H^{-1}(\Omega)}$. Since this "full stability" is necessary for removing the augmentation AG from (1.4), we thus expect that the SIP method (1.2) is not $|\cdot|_{1;\eta}$-*quasi-optimal* and so does not always fully exploit the approximation potential offered by its discrete space S_1^0. This suspicion is confirmed by Remark 4.9 in the first part [23] of this series.

In order to achieve quasi-optimality, we consider the following variant of the discrete problem (1.2): find $U_E \in S_1^0$ such that

$$(1.5) \qquad\qquad \forall \sigma \in S_1^0 \quad b(U_E, \sigma) = \langle f, E\sigma \rangle,$$

where the linear operator $E : S_1^0 \to H_0^1(\Omega)$ to be specified enables $f \in H^{-1}(\Omega)$. If we require that the means on internal faces are conserved as in Badia et al. [4, section 6],

$$(1.6) \qquad\qquad \forall \sigma \in S_1^0, F \in \mathcal{F}^i \quad \int_F E\sigma = \int_F \{\!\{ \sigma \}\!\},$$

then piecewise integrating by parts twice shows

$$\forall s, \sigma \in S_1^0 \quad b_1(s, \sigma) = \int_\Sigma [\![\nabla s]\!] \cdot n \,\{\!\{ \sigma \}\!\} = \int_\Omega \nabla_{\mathcal{M}} s \cdot \nabla(E\sigma).$$

The right-hand side of this identity provides an extension $\widetilde{b}_1$ of b_1 that does not involve

regularity beyond $H_0^1(\Omega)$, in contrast to the one tacitly employed in (1.3). Thus, we obtain the following regularity-free counterpart of (1.3):

$$\forall u \in H_0^1(\Omega),\ \sigma \in S_1^0 \qquad \widetilde{b}_1(u, \sigma) = \langle f, E\sigma \rangle, \quad \text{and} \quad b_2(u, \sigma) = 0.$$

In order to construct an "$H_0^1(\Omega)$-smoothing operator" that satisfies (1.6) and is computationally feasible, we extend a similar operator devised in the second part [24] of this series, ensuring that its operator norm $\|E\|_{\mathcal{L}(S_1^0, H_0^1(\Omega))}$ is bounded in terms of the shape coefficient $\gamma_{\mathcal{M}}$ of $\mathcal{M}$.

Exploiting the improved stability and consistency properties of (1.5), the abstract theory of [23] then yields

$$|u - U_E|_{1;\eta} \leq \left(1 + C\eta^{-1}\right)^{\frac{1}{2}} \inf_{s \in S_1^0} |u - s|_{1;\eta},$$

where C depends on d and $\gamma_{\mathcal{M}}$ and η is sufficiently large. Notably, as $\eta \to \infty$, the discontinuous space S_1^0 is replaced by the space S_1^1 of continuous piecewise affine functions and we end up exactly in Céa's lemma for the conforming Galerkin method with S_1^1.

It is worth comparing with the quasi-optimal Crouzeix–Raviart method for (1.1) of the second part [24] of this series. Thanks to the coupling between Crouzeix–Raviart elements, b_1 becomes symmetric and there is no need for b_2 and penalization. As a consequence, the ensuing quasi-optimality constant equals the operator norm with respect to the piecewise energy norm of the smoothing operator E.

The rest of this article is organized as follows. Section 2 provides a brief summary of the abstract results in [23] to be used here. In section 3, we introduce new variants of various interior penalty methods and prove their quasi-optimality. First, we design quasi-optimal DG methods of arbitrary fixed order for the Poisson problem, covering also the setting illustrated in this introduction. Second, we devise a quasi-optimal Crouzeix–Raviart interior penalty method for linear elasticity and establish a robust error bound for it in the nearly incompressible regime. Last, we conclude with a quasi-optimal variant of the quadratic C^0-interior penalty method for the biharmonic problem.

In these examples, we consider polyhedral domains with Lipschitz boundaries and homogeneous essential boundary conditions. An application of the presented approach to more general domains and boundary conditions is given in [25]. Numerical investigations will be presented elsewhere.

2. Stability and consistency for quasi-optimality. We briefly summarize the characterization of quasi-optimality in [23], adopting an approach to nonconforming consistency corresponding to the so-called second Strang lemma, which was introduced in [6].

A linear and symmetric elliptic problem can be written in the following abstract form: given $\ell \in V'$, find $u \in V$ such that

(2.1) $$\forall v \in V \quad a(u, v) = \langle \ell, v \rangle,$$

where V is an infinite-dimensional Hilbert space with scalar product $a(\cdot, \cdot)$, V' is its (topological) dual space, and $\langle \cdot, \cdot \rangle$ stands for the dual pairing of V and V'. We write $\|\cdot\| = \sqrt{a(\cdot, \cdot)}$ for the *energy norm*, which induces the *dual energy norm* $\|\ell\|_{V'} := \sup_{v \in V, \|v\|=1} \langle \ell, v \rangle$ on V'. Problem (2.1) is uniquely solvable and, introducing the Riesz isometry $A : V \to V'$, $v \mapsto a(v, \cdot)$, we have $u = A^{-1}\ell$ with $\|u\| = \|\ell\|_{V'}$.
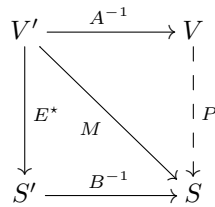
FIG. 1. *Diagram with operators A, B, E, nonconforming method M = (S, b, E), and induced approximation operator P.*

We shall design quasi-optimal methods $M : V' \to S$ with discrete problems of the following form: given $\ell \in V'$, find $M\ell \in S$ such that

$$(2.2) \qquad \forall \sigma \in S \quad b(M\ell, \sigma) = \langle \ell, E\sigma \rangle,$$

where S is a finite-dimensional linear space, $b : S \times S \to \mathbb{R}$ is a nondegenerate bilinear form, E is a linear operator from S to V, and $\langle \cdot, \cdot \rangle$ stands also for the pairing of S and S'. Although we do not require $S \subset V$, the operator E ensures that the method M is *entire*, i.e., defined for all $\ell \in V'$. In light of [23, Remark 2.4], this is a necessary condition for the kind of quasi-optimality we are interested in. We refer to E as a *smoothing operator* or *smoother*, because $S \not\subset V$ often arises for the lack of smoothness. Moreover, we identify the operator M with the triplet (S, b, E), ignoring some slight ambiguity; cf. [23, Remark 2.2].

The relationship between continuous and discrete problem is illustrated by the commutative diagram in Figure 1. This diagram introduces
- the adjoint $E^\star : V' \to S'$ given by $\langle E^\star \ell, \sigma \rangle = \langle \ell, E\sigma \rangle$ for $\ell \in V'$, $\sigma \in S$,
- the invertible map $B : S \to S'$, $s \mapsto b(s, \cdot)$,
- the approximation operator $P := MA$

and illustrates the representations

$$(2.3) \qquad M = B^{-1} E^\star \qquad \text{and} \qquad P = B^{-1} E^\star A.$$

The solution u of (2.1) is thus approximated by $M\ell$ with $\ell = Au$, that is, by Pu. To assess the quality of this approximation, we assume that a can be extended to a scalar product $\tilde{a}$ on the sum $\widetilde{V} := V + S$ and quantify the error with the *extended energy norm*

$$\| \cdot \| := \sqrt{\tilde{a}(\cdot, \cdot)} \quad \text{on } \widetilde{V},$$

using the same notation as for the original one. The best approximation error within S to u is then $\inf_{s \in S} \|u - s\|$ and attained by the $\tilde{a}$-orthogonal projection Π_S onto S. We say that the method M is *quasi-optimal* (for Problem (2.1) with respect to the extended energy norm) if there exists a constant $C \geq 1$ such that

$$(2.4) \qquad \forall u \in V \qquad \|u - Pu\| \leq C \inf_{s \in S} \|u - s\|.$$

The associated quasi-optimality constant C_{qopt} of M is then the smallest constant with this property. Notice that (2.4) involves all exact solutions of (2.1), not only certain smooth ones.

Remark 2.1 (equivalence to the best error). Notice that in (2.4) the same norm is used on both sides. Thus, $Pu \in S$ readily entails that the error $\|u - Pu\|$ is equivalent

to the best error $\inf_{s \in S} \|u - s\|$ for quasi-optimal methods. This property justifies the terminology *quasi-optimality* and distinguishes (2.4) from estimates like (1.4) in terms of other best errors.

THEOREM 2.2 (stability, consistency, and quasi-optimality). *Given a nonconforming method $M = (S, b, E)$ for* (2.1) *and an extended scalar product $\widetilde{a}$, introduce the bilinear form $d : V \times S \to \mathbb{R}$ by*

$$(2.5) \qquad\qquad d(v, \sigma) := b(\Pi_S v, \sigma) - a(v, E\sigma).$$

Then the following hold:

   (i) *M is bounded, or fully stable, with*

$$C_{\text{stab}} := \|M\|_{\mathcal{L}(V', S)} = \sup_{\sigma \in S} \frac{\|E\sigma\|}{\sup_{s \in S, \|s\|=1} b(s, \sigma)}.$$

  (ii) *M is quasi-optimal if and only if it is fully algebraically consistent in that*

$$\forall u \in S \cap V, \sigma \in S \quad 0 = d(u, \sigma) = b(u, \sigma) - a(u, E\sigma).$$

 (iii) *If M is quasi-optimal, then its quasi-optimality constant satisfies*

$$C_{\text{stab}} \le C_{\text{qopt}} = \sqrt{1 + \delta^2},$$

   where $\delta \in [0, \infty)$ is the consistency measure given by the smallest constant in

$$\forall v \in V, \sigma \in S \quad |d(v, \sigma)| \le \delta \sup_{\hat{s} \in S, \|\hat{s}\|=1} b(\hat{s}, \sigma) \inf_{s \in S} \|v - s\|.$$

Proof. Item (i) follows from [23, Theorem 4.7], while (ii) is a consequence of [23, Theorem 4.14] and (i). Finally, the first part of [23, Theorem 4.19] implies (iii).

Since the formula for C_{qopt} plays a key role in what follows, we provide a self-contained proof of it. Let $u \in V$. The Pythagoras theorem yields

$$(2.6) \qquad\qquad \|u - Pu\|^2 = \|u - \Pi_S u\|^2 + \|\Pi_S u - Pu\|^2.$$

To bound the second term, we observe that, owing to the nondegeneracy of b, $\|\sigma\|_b := \sup_{\hat{s} \in S, \|\hat{s}\|=1} b(\hat{s}, \sigma)$ defines norm on S with

$$(2.7) \qquad\qquad \forall s \in S \quad \|s\| = \sup_{\sigma \in S} \frac{b(s, \sigma)}{\|\sigma\|_b}.$$

Indeed, the definition of $\| \cdot \|_b$ readily implies "$\ge$" and the converse inequality follows from

$$\inf_{s \in S} \sup_{\sigma \in S} \frac{b(s, \sigma)}{\|s\|\|\sigma\|_b} = \inf_{\sigma \in S} \sup_{s \in S} \frac{b(s, \sigma)}{\|s\|\|\sigma\|_b} = \inf_{\sigma \in S} \frac{1}{\|\sigma\|_b} \sup_{s \in S} \frac{b(s, \sigma)}{\|s\|} = 1.$$

Employing (2.7) and the definitions of P and δ, we deduce

$$\|\Pi_S u - Pu\| = \sup_{\sigma \in S} \frac{d(u, \sigma)}{\|\sigma\|_b} \le \delta \|\Pi_S u - u\|.$$

Since u is arbitrary and the only inequality we have used is the one defining δ, this and (2.6) prove $C_{\text{qopt}} = \sqrt{1 + \delta^2}$. □

Some comments on Theorem 2.2 and the consequences of its proof are in order. The adverb "fully" in (i) means that M is defined on V'. (While, for instance, the method (1.2) is defined only on a strict subspace of $H^{-1}(\Omega)$.) The built-in *full stability* of the considered methods is a necessary condition for quasi-optimality. It has to be established by applying a smoothing operator E before evaluating the load functional. We refer to the constant C_{stab} as the stability constant of M.

The use of "fully" in (ii) indicates that no additional regularity on u is assumed. (In contrast, only sufficiently smooth solutions of (1.1) are considered in (1.3).) Notice that *full algebraic consistency* does not actually depend on the extension $\widetilde{a}$ of the scalar product a. In particular, it can be rephrased in the following manner: whenever an exact solution happens to be discrete, it has to be also the discrete solution. Natural candidates for full algebraic consistency are *nonconforming Galerkin methods* satisfying

$$(2.8) \qquad b_{|S_C \times S_C} = a_{|S_C \times S_C} \quad \text{and} \quad E_{|S_C} = \mathrm{Id}_{S_C},$$

where $S_C := S \cap V$ is the conforming subspace of S. Notice that this generalization of conforming Galerkin methods does not determine b and E if S is not a subspace of V. Furthermore, it may be weaker than full algebraic consistency, involving also nonconforming discrete test functions.

While full algebraic consistency involves only the conforming part S_C of the discrete space, the constant δ captures consistency properties of M for nonconforming directions in $S \setminus V$. We call M *(algebraically) overconsistent* whenever $d(\cdot, \cdot)$ vanishes, that is, whenever the discrete bilinear form b is $\widetilde{a}(\cdot, E\cdot)$. In this case, $C_{\mathrm{stab}} = C_{\mathrm{qopt}}$; see [24, Theorem 2.1]. The following simple consequence of the inf-sup theory implies that this appealing property requires a certain interplay of S, V, and $\widetilde{a}$; cf. also [24, Lemma 2.6], which strengthens the statement to a characterization.

LEMMA 2.3 (obstruction for nondegenerate $\widetilde{a}(\cdot, E\cdot)$). *Let S, V, and $\widetilde{a}$ be given as in Theorem 2.2 and assume that there exists some $\sigma \in S \setminus \{0\}$ such that $\widetilde{a}(\sigma, v) = 0$ for all $v \in V$. Then, for any smoother $E : S \to V$, the bilinear form $\widetilde{a}(\cdot, E\cdot)$ is degenerate.*

For a further discussion of the aforementioned notions and their role and properties, we refer to [23] and [24, section 2]. Here we continue by underlining that Theorem 2.2 was formulated with the following viewpoint: for quasi-optimality, the discrete bilinear form b has to be a perturbation of $\widetilde{a}(\cdot, E\cdot)$, which is fully algebraically consistent and affects the quasi-optimality via δ. This viewpoint will be our guiding principle for constructing quasi-optimal interior penalty methods. It is therefore of interest to bound C_{stab} and δ, connecting them to a well-known and important, but not yet mentioned, constant.

Remark 2.4 (stability, consistency, and inf-sup constants). Let $M = (S, b, E)$ be a nonconforming method. As S is finite-dimensional, the nondegeneracy of b entails that the inf-sup constant is positive:

$$\alpha := \inf_{\sigma \in S, \|\sigma\|=1} \sup_{s \in S, \|s\|=1} b(s, \sigma) > 0.$$

Then the definitions of C_{stab} and δ readily yield

$$(2.9) \qquad C_{\mathrm{stab}} \leq \frac{\|E\|_{\mathcal{L}(S,V)}}{\alpha} \quad \text{and} \quad \delta \leq \frac{\gamma}{\alpha},$$

where $\gamma \geq 0$ verifies $|d(s,\sigma)| \leq \gamma \inf_{s \in S} \|v-s\| \|\sigma\|$ for all $v \in V$ and $\sigma \in S$. Hence, up to the inverse of the inf-sup constant α, the constants C_{stab} and δ depend, respectively,

only on the smoothing operator E and the bilinear form $d(\cdot, \cdot)$. It is worth noting that these bounds may be pessimistic; see [24, Remark 2.5].

3. Applications to interior penalty methods. The goal of this section is to devise interior penalty methods that are based upon nonconforming finite elements and are quasi-optimal. In view of Theorem 2.2 and Lemma 2.3, we may achieve this by the following steps: given a continuous problem (2.1) and a nonconforming finite element space S,

- extend the scalar product a to the sum $V + S$,
- find a computationally feasible smoothing operator $E : S \to V$, possibly with $E_{|V \cap S} = \mathrm{Id}_{V \cap S}$,
- if necessary, use a bilinear form $d(\cdot, \cdot)$ with $d_{|(S \cap V) \times S} = 0$ to arrange that $b = \widetilde{a}(\cdot, E\cdot) + d(\cdot, \cdot)$ is nondegenerate and has other optional properties like symmetry, without losing control of δ in item (iii) of Theorem 2.2.

Notice that here the domain of the bilinear form $d(\cdot, \cdot)$ is $S \times S$, while in Theorem 2.2 it is $V \times S$. We use the same notation because both forms have a common, unique extension on $\widetilde{V} \times S$ thanks to $d_{|(S \cap V) \times S} = 0$.

Denoting by $\varphi_1, \ldots, \varphi_n$ the nodal basis of S, we consider a smoothing operator E to be *computationally feasible* if each $E\varphi_i$ is in some conforming finite element space and the number of elements in its support is bounded independently of n.

We shall carry out the aforementioned steps for three different settings, involving vector and fourth-order problems as well as various couplings between elements (completely discontinuous, Crouzeix–Raviart, continuous). In each case the nondegeneracy of b will be obtained by means of interior penalties.

3.1. Simplicial meshes and (broken) function spaces. We indicate Lebesgue and Sobolev spaces as usual (see, e.g., [9]) and adopt the following notations, mainly taken from [24].

Given $n \in \{0, \ldots, d\}$, an *n-simplex* $C \subseteq \mathbb{R}^d$ is the convex hull of $n + 1$ points $z_1, \ldots, z_{n+1} \in \mathbb{R}^d$ spanning an n-dimensional affine space. The uniquely determined points $z_1, \ldots, z_{n+1}$ are the vertices of C and form the set $\mathcal{L}_1(C)$. If $n \geq 1$, we let $\mathcal{F}_C$ denote the $(n-1)$-dimensional faces of C, which are the $(n-1)$-simplices arising by picking n distinct vertices from $\mathcal{L}_1(C)$. Given a vertex $z \in \mathcal{L}_1(C)$, its barycentric coordinate λ_z^C is the unique first-order polynomial on C such that $\lambda_z^C(y) = \delta_{zy}$ for all $y \in \mathcal{L}_1(C)$. Then $0 \leq \lambda_z^C \leq 1$ and $\sum_{z \in \mathcal{L}_1(C)} \lambda_z^C = 1$ in C and, if $\alpha = (\alpha_z)_{z \in \mathcal{L}_1(C)} \in \mathbb{N}_0^{n+1}$ is a multi-index,

$$(3.1) \qquad \int_C \prod_{z \in \mathcal{L}_1(C)} (\lambda_z^C)^{\alpha_z} = \frac{n! \alpha!}{(n + |\alpha|)!} |C|,$$

where $|C|$ stands also for the n-dimensional Hausdorff measure in $\mathbb{R}^d$. We write m_C for the barycenter of C, $h_C := \mathrm{diam}(C)$ for its diameter, ρ_C for the diameter of its largest inscribed n-dimensional ball, and γ_C for its shape coefficient $\gamma_C := h_C / \rho_C$.

Let $\mathcal{M}$ be a simplicial, face-to-face *mesh* of some open, bounded, connected, and polyhedral set $\Omega \subset \mathbb{R}^d$ with Lipschitz boundary $\partial\Omega$. More precisely, $\mathcal{M}$ is a finite collection of d-simplices in $\mathbb{R}^d$ such that $\overline{\Omega} = \bigcup_{K \in \mathcal{M}} K$ and the intersection of two arbitrary elements $K_1, K_2 \in \mathcal{M}$ is either empty or an n-simplex with $n \in \{0, \ldots, d\}$ and $\mathcal{L}_1(K_1 \cap K_2) = \mathcal{L}_1(K_1) \cap \mathcal{L}_1(K_2)$. We let $\mathcal{F} := \bigcup_{K \in \mathcal{M}} \mathcal{F}_K$ denote the $(d-1)$-dimensional faces of $\mathcal{M}$ and distinguish between boundary faces $\mathcal{F}^b := \{F \in \mathcal{F} \mid F \subseteq \partial\Omega\}$ and interior faces $\mathcal{F}^i := \mathcal{F} \setminus \mathcal{F}^b$. Moreover, let $\Sigma := \cup_{F \in \mathcal{F}} F$ be the skeleton of $\mathcal{M}$ and, fixing a unit normal n_F for each interior face $F \in \mathcal{F}^i$, extend the outer

normal n of $\partial\Omega$ to Σ by $n_{|F} = n_F$ for $F \in \mathcal{F}^i$. The ambiguity in the orientation of n_F is insignificant to our discussion. The meshsize h on Σ is given by $h_{|F} = h_F$ for all $F \in \mathcal{F}$ and the shape coefficient of $\mathcal{M}$ is

$$\gamma_{\mathcal{M}} := \max_{K \in \mathcal{M}} \gamma_K.$$

For $k \in \mathbb{N}$, the broken Sobolev space of order k is

$$H^k(\mathcal{M}) := \{v \in L^2(\Omega) \mid \forall K \in \mathcal{M} \ v_{|K} \in H^k(K)\}.$$

If $v \in H^k(\mathcal{M})$, we use the subscript $\mathcal{M}$ to indicate the piecewise variant of a differential operator. For instance, $\nabla_{\mathcal{M}} v$ is given by $(\nabla_{\mathcal{M}} v)_{|K} := \nabla(v_{|K})$ for all $K \in \mathcal{M}$. Jumps and averages are defined as follows. Given an interior face $F \in \mathcal{F}^i$, let $K_1, K_2 \in \mathcal{M}$ be the two elements such that $F = K_1 \cap K_2$ and the outer normal of K_1 coincides with $n_{|F}$ on F. Set

(3.2a) $$[\![v]\!] := v_{|K_1} - v_{|K_2}, \quad \{\!\{v\}\!\} := \frac{1}{2}\left(v_{|K_1} + v_{|K_2}\right) \quad \text{on } F.$$

The fact that the sign of $[\![v]\!]$ depends on the ordering of K_1 and K_2 will be insignificant to our discussion. It will be convenient to extend these definitions on $\partial\Omega$. Given $F \in \mathcal{F}^b$, let $K \in \mathcal{M}$ be the element such that $F = K \cap \partial\Omega$ and set

(3.2b) $$[\![v]\!] := \{\!\{v\}\!\} := v_{|K} \quad \text{on } F.$$

In this notation, piecewise integration by parts reads as follows: if $v, w \in H^1(\mathcal{M})$ and $j \in \{1, \ldots, d\}$, then

(3.3)
$$\int_{\Omega} (\partial_{j,\mathcal{M}} v)w - \int_{\Sigma \backslash \partial\Omega} [\![v]\!] \{\!\{w\}\!\} \, n \cdot e_j$$
$$= -\int_{\Omega} v(\partial_{j,\mathcal{M}} w) + \int_{\Sigma \backslash \partial\Omega} \{\!\{v\}\!\} [\![w]\!] \, n \cdot e_j + \int_{\partial\Omega} vw \, n \cdot e_j.$$

Notice that the surface integrals are independent of the orientation of n and that, e.g., the singular part of the distributional derivative $\partial_{j,\mathcal{M}} v$ is represented by means of the negative jumps $-[\![v]\!]_{|F}$, $F \in \mathcal{F}^i$.

Given $p \in \mathbb{N}_0$, we write $\mathbb{P}_p(C)$ for the linear space of *polynomials* on the n-simplex C with (total) degree $\leq p$. Consider $p \in \mathbb{N}$, excluding the trivial case $p = 0$. A polynomial in $\mathbb{P}_p(C)$ is determined by its point values at the Lagrange nodes $\mathcal{L}_p(C)$ of order p, which, for $p \geq 2$, are given by $\left\{x \in C \mid \forall z \in \mathcal{L}_1(C) \ \ p\lambda_z^C(x) \in \mathbb{N}_0\right\}$. We let $\Psi_{C,z}^p$, $z \in \mathcal{L}_p(C)$, denote the associated nodal basis in $\mathbb{P}_p(C)$ given by $\Psi_{C,z}^p(y) = \delta_{zy}$ for all $y, z \in \mathcal{L}_p(C)$. The Lagrange nodes are nested in that $\mathcal{L}_p(F) = \mathcal{L}_p(C) \cap F$ for any face $F \in \mathcal{F}_C$. Thus, the restriction $P_{|F}$ of $P \in \mathbb{P}_p(C)$ is determined by the "restriction" $\mathcal{L}_p(C) \cap F$ of the Lagrange nodes and we have $\Psi_{C,z|F}^p = \Psi_{F,z}^p$ for all $z \in \mathcal{L}_p(F)$.

Given $k, p \in \mathbb{N}_0$, the space of functions that are *piecewise polynomial* with degree $\leq p$ and are in $H_0^k(\Omega)$ (with the convention $H_0^0(\Omega) := L^2(\Omega)$ for $k = 0$) is

(3.4) $$S_p^k := \left\{s \in H_0^k(\Omega) \mid \forall K \in \mathcal{M} \ s_{|K} \in \mathbb{P}_p(K)\right\}.$$

The cases $p \in \mathbb{N}$ with $k \in \{0, 1\}$ are of particular interest.

Consider first S_p^0 with $p \in \mathbb{N}$ and extend each $\Psi_{K,z}^p$ outside of $K \in \mathcal{M}$ by 0. The functions $\{\Psi_{K,z}^p\}_{K \in \mathcal{M}, z \in \mathcal{L}_p(K)}$ form a basis of S_p^0 with $\Psi_{K,z|K'}^p(z') = \delta_{K,K'}\delta_{z,z'}$ for $K, K' \in \mathcal{M}$ and $z \in \mathcal{L}_p(K)$, $z' \in \mathcal{L}_p(K')$, which amounts to distinguishing Lagrange nodes from different elements.

The construction of a basis of S_p^1 is a little more involved. Here, identifying coinciding Lagrange nodes, we set $\mathcal{L}_p := \cup_{K \in \mathcal{M}}\mathcal{L}_p(K)$ as well as $\mathcal{L}_p^i := \mathcal{L}_p \setminus \partial\Omega$, and write Φ_z^p, $z \in \mathcal{L}_p$, for the function given piecewise by $\Phi_{z|K}^p := \Psi_{K,z}^p$ if $z \in K$ and $\Phi_{z|K}^p := 0$ otherwise. Then the nestedness of Lagrange nodes implies that $\{\Phi_z^p\}_{z \in \mathcal{L}_p^i}$ is a basis of S_p^1 satisfying $\Phi_z^p(y) = \delta_{zy}$ for all $y, z \in \mathcal{L}_p^i$. In connection with these basis functions, the following subdomains are useful. Let $\omega_z := \bigcup_{K' \ni z} K'$ be the star around $z \in \mathcal{L}_p$ and let $\omega_K := \bigcup_{K' \cap K \neq \emptyset} K'$ be the patch around $K \in \mathcal{M}$. Since $\partial\Omega$ is Lipschitz, stars are face-connected in the sense of [22]: given $z \in \mathcal{L}_p$ and any pair $K, K' \in \mathcal{M}$ with $z \in K \cap K'$, there exists a path $\{K_i\}_{i=1}^n \subset \mathcal{M}$ of elements containing z such that $K_1 = K$, $K_n = K'$, and each $K_i \cap K_{i+1} \in \mathcal{F}^i$.

If not specified differently, C_* stands for a function which is not necessarily the same at each occurrence and depends on a subset $*$ of $\{d, \gamma_{\mathcal{M}}, p\}$, increasing in $\gamma_{\mathcal{M}}$ and p if present. For instance, we have, for $K, K' \in \mathcal{M}$,

$$(3.5) \qquad K \cap K' \neq \emptyset \quad \Longrightarrow \quad |K| \leq C_{\gamma_{\mathcal{M}}} |K'| \text{ and } h_K \leq C_{\gamma_{\mathcal{M}}} \rho_{K'}$$

and, for $p \in \mathbb{N}$, $K \in \mathcal{M}$, and $z \in \mathcal{L}_p(K)$,

$$(3.6) \qquad c_{d,p}|K|^{\frac{1}{2}} h_K^{-1} \leq \|\nabla\Psi_{K,z}^p\|_{L^2(K)} \leq C_{d,p}|K|^{\frac{1}{2}} \rho_K^{-1}.$$

If there is no danger of confusion, $A \leq C_* B$ may be abbreviated as $A \lesssim B$.

3.2. Quasi-optimal DG methods for the Poisson problem. In this subsection we devise quasi-optimal DG methods for the Poisson problem, covering the results illustrated in the introduction (section 1).

Let Ω and $\mathcal{M}$ be as in section 3.1 and, with $\eta \geq 0$, define

$$(3.7) \qquad (v, w)_{1;\eta} := \int_\Omega \nabla_{\mathcal{M}} v \cdot \nabla_{\mathcal{M}} w + \sum_{F \in \mathcal{F}} \frac{\eta}{h_F} \int_F [\![v]\!] \, [\![w]\!], \quad |v|_{1;\eta} := (v, v)_{1;\eta}^{\frac{1}{2}}$$

on $H^1(\mathcal{M})$ and abbreviate $(\cdot, \cdot)_{1;0}$ to $(\cdot, \cdot)_1$. Recalling (3.4), we consider

$$(3.8) \qquad V = H_0^1(\Omega), \quad S = S_p^0 \text{ with } p \in \mathbb{N}, \quad \widetilde{a} = (\cdot, \cdot)_{1;\eta} \text{ on } \widetilde{V} = H_0^1(\Omega) + S_p^0.$$

Then $\widetilde{a}$ is a scalar product for $\eta > 0$ and the abstract problem (2.1) provides a weak formulation of (1.1). Our setting has two parameters: the polynomial degree p and the scaling factor η of the jumps. The latter will also be the penalty parameter and is essentially free to be specified by the user. In order to keep notation simple, we shall sometimes suppress the dependencies on p and η. The conforming part of S_p^0 is the strict subspace

$$(3.9) \qquad S_p^0 \cap H_0^1(\Omega) = S_p^1 = \{s \in S_p^0 \mid \forall F \in \mathcal{F} \ [\![s]\!] \equiv 0\}.$$

Moreover, we easily see that

$$(3.10) \qquad \emptyset \neq S_0^0 \subseteq S_p^0 \cap V^\perp,$$

which precludes overconsistency in light of Lemma 2.3.

To obtain hints for a suitable choice of the smoothing operator $E : S_p^0 \to H_0^1(\Omega)$, we invoke integration by parts and the piecewise structure of S_p^0. Let $s, \sigma \in S_p^0$ be arbitrary. On the one hand, piecewise integration by parts (3.3) yields

$$(s, E\sigma)_{1;\eta} = \sum_{K \in \mathcal{M}} \int_K (-\Delta s) E\sigma + \sum_{F \in \mathcal{F}^i} \int_F [\![\nabla s]\!] \cdot n E\sigma$$

due to $E\sigma \in H_0^1(\Omega)$. On the other hand, we want $\int_\Omega \nabla_\mathcal{M} s \cdot \nabla_\mathcal{M} \sigma = (s, \sigma)_1$ to appear in the discrete bilinear form. For this term, (3.2b) and (3.3) give

$$(s, \sigma)_1 = \sum_{K \in \mathcal{M}} \int_K (-\Delta s)\sigma + \sum_{F \in \mathcal{F}^i} \int_F [\![\nabla s]\!] \cdot n \{\!\{\sigma\}\!\} \; + \sum_{F \in \mathcal{F}} \int_F \{\!\{\nabla s\}\!\} \cdot n [\![\sigma]\!].$$

A comparison of these two identities suggests that the smoothing operator E should conserve certain moments on faces and elements and proves the following lemma. Such moment conservation was already used in Badia et al. [4, section 6] to design a DG method with a (partial) quasi-optimality result and in [24, sections 3.2 and 3.3] to construct overconsistent Crouzeix–Raviart-like methods of arbitrary fixed order.

LEMMA 3.1 (conservation of moments). *Let $p \in \mathbb{N}$ and, for notational conve-nience, set $\mathbb{P}_{-1}(K) = \emptyset$ for all $K \in \mathcal{M}$. If the linear operator $E : S_p^0 \to H_0^1(\Omega)$ satisfies*

$$(3.11) \qquad \int_F q(E\sigma) = \int_F q \{\!\{\sigma\}\!\} \quad and \quad \int_K r(E\sigma) = \int_K r\sigma$$

for all $F \in \mathcal{F}^i$, $q \in \mathbb{P}_{p-1}(F)$, $K \in \mathcal{M}$, $r \in \mathbb{P}_{p-2}(K)$, and $\sigma \in S_p^0$, then

$$(s, E\sigma)_{1;\eta} = \int_\Omega \nabla_\mathcal{M} s \cdot \nabla_\mathcal{M} \sigma - \sum_{F \in \mathcal{F}} \int_F \{\!\{\nabla s\}\!\} \cdot n [\![\sigma]\!]$$

for all $s, \sigma \in S_p^0$.

We adapt the construction of the smoothing operators in [24] to the given setting and begin with the so-called bubble smoother. It employs the following weighted L^2-projections associated to faces and elements. For every interior face $F \in \mathcal{F}^i$, let $Q_F : L^2(F) \to \mathbb{P}_{p-1}(F)$ be given by

$$(3.12) \qquad \forall q \in \mathbb{P}_{p-1}(F) \quad \int_F (Q_F v) q \, \Phi_F = \int_F vq,$$

where $\Phi_F := \prod_{z \in \mathcal{L}_1(F)} \Phi_z^1 \in S_d^1$ is the face bubble function supported in the two elements containing F. Moreover, for every mesh element $K \in \mathcal{M}$, set $Q_K = 0$ if $p = 1$, and otherwise let $Q_K : L^2(K) \to \mathbb{P}_{p-2}(K)$ be given by

$$(3.13) \qquad \forall r \in \mathbb{P}_{p-2}(K) \quad \int_K (Q_K v) r \, \Phi_K = \int_K vr,$$

where $\Phi_K := \prod_{z \in \mathcal{L}_1(K)} \Phi_z^1 \in S_{d+1}^1$ is the element bubble function with support K. For $v \in H^1(\mathcal{M})$, we then define the global bubble operators

$$B_{\mathcal{M},p} v := \sum_{K \in \mathcal{M}} (Q_K v) \Phi_K, \quad B_{\mathcal{F},p} v := \sum_{F \in \mathcal{F}^i} \sum_{z \in \mathcal{L}_{p-1}(F)} \left(Q_F \{\!\{v\}\!\} \right)(z) \Phi_z^{p-1} \Phi_F,$$

where, if $p = 1$, we set $\mathcal{L}_0(F) := \{m_F\}$ and $\Phi^0_{m_F} := 1$ with m_F being the barycenter of F. Notice that $B_{\mathcal{F},p}$ incorporates an extension by means of Lagrange basis functions in view of the partition of unity $\sum_{z \in \mathcal{L}_{p-1}(F)} \Phi^{p-1}_z = 1$ on F. The combination of these two operators provides the desired property and an extension of the operator with the same name in [24].

LEMMA 3.2 (bubble smoother). *For $p \in \mathbb{N}$, the linear operator $B_p : S^0_p \to H^1_0(\Omega)$ defined by*

$$B_p \sigma := B_{\mathcal{F},p}\sigma + B_{\mathcal{M},p}(\sigma - B_{\mathcal{F},p}\sigma)$$

satisfies (3.11) and, for any $K \in \mathcal{M}$, the local stability estimate

$$\|\nabla B_p \sigma\|_{L^2(K)} \leq \frac{C_{d,p}}{\rho_K} \left(\sup_{r \in \mathbb{P}_{p-2}(K)} \frac{\int_K \sigma r}{\|r\|_{L^2(K)}} + \sum_{F \in \mathcal{F}_K} \frac{|K|^{\frac{1}{2}}}{|F|^{\frac{1}{2}}} \sup_{q \in \mathbb{P}_{p-1}(F)} \frac{\int_F \{\!\!\{\sigma\}\!\!\} q}{\|q\|_{L^2(F)}} \right).$$

Proof. Proceed as in the proof of [24, Lemma 3.8]. □

The factor ρ_K^{-1} in the stability estimate in Lemma 3.2 suggests that B_p is not uniformly stable under refinement. The example in [24, Remark 3.5] confirms this also for the current setting. However, since the bound involves lower-order norms, we have the possibility to stabilize. This will be done with the help of the following variant $A_p : S^0_p \to S^1_p$ of nodal averaging. For every interior node $z \in \mathcal{L}^i_p$, fix some element $K_z \in \mathcal{M}$ containing z and set

(3.14) $$A_p \sigma := \sum_{z \in \mathcal{L}^i_p} \sigma_{|K_z}(z)\Phi^p_z, \qquad \sigma \in S^0_p.$$

Clearly, $A_p \sigma(z) = \sigma(z)$ whenever σ is continuous at $z \in \mathcal{L}^i_p$ and so A_p is a projection onto S^1_p. On the one hand, the operator A_q is a restriction of Scott–Zhang interpolation [20] defined for broken H^1-functions and, on the other hand, it is a simplified variant of nodal averaging in that it requires only one evaluation per degree of freedom. Nodal averaging has been used in various nonconforming contexts; see, e.g., Brenner [7], Karakashian and Pascal [16], and Oswald [18]. Here the following error bound for A_p will be instrumental.

LEMMA 3.3 (simplified nodal averaging and L^2-norms of jumps). *Let $p \in \mathbb{N}$, $\sigma \in S^0_p$ piecewise polynomial, $K \in \mathcal{M}$, and $z \in \mathcal{L}_p(K)$ be a Lagrange node. If $z \notin \partial K$, then $A_p\sigma(z) = \sigma_{|K}(z)$, else*

$$\left| \sigma_{|K}(z) - A_p\sigma(z) \right| \leq C_{d,p} \sum_{F \in \mathcal{F}: F \ni z} \frac{1}{|F|^{\frac{1}{2}}} \| [\![\sigma]\!] \|_{L^2(F)}.$$

Proof. The "then" part of the claim readily follows from the nonoverlapping of elements in $\mathcal{M}$. For the "else" part, we first recall that $[\![\sigma]\!]_{|F}$ denotes the jump across the face F and notice that its point values are well-defined. We thus can derive

$$\left| \sigma_{|K}(z) - A_p(z) \right| \leq \sum_{F \in \mathcal{F}: F \ni z} \left| [\![\sigma]\!]_{|F}(z) \right|$$

with the help of the face-connectedness of stars in $\mathcal{M}$; cf. [24, equation (3.6)]. Therefore, the inverse estimate $\| \cdot \|_{L^\infty(F)} \leq C_{d,p} |F|^{-\frac{1}{2}} \| \cdot \|_{L^2(F)}$ in $\mathbb{P}_p(F)$ finishes the proof. □

Stabilizing the bubble smoother B_p with simplified nodal averaging A_p, we obtain a smoothing operator with the desired properties.

PROPOSITION 3.4 (stable smoothing with moment conservation).   *The linear operator* $E_p : S_p^0 \to H_0^1(\Omega)$ *given by*

$$E_p\sigma := A_p\sigma + B_p(\sigma - A_p\sigma)$$

is invariant on S_p^1 *and satisfies* (3.11), *and for all* $\sigma \in S_p^0$,

$$\|\nabla_{\mathcal{M}}(\sigma - E_p\sigma)\|_{L^2(\Omega)} \leq C_{d,\gamma_{\mathcal{M}},p}\|h^{-\frac{1}{2}}[\![\sigma]\!]\|_{L^2(\Sigma)}.$$

Proof. We adapt the proof of [24, Propositions 3.3 and 3.9] to the current setting with jumps in the extended energy norm.

Clearly, the operator E_p is well-defined and maps into $H_0^1(\Omega)$. With A_p, also E_p is a projection onto S_p^1. We next show that E_p conserves the moments in (3.11). Given any $F \in \mathcal{F}^i$ and any $q \in \mathbb{P}_{p-1}(F)$, we arrange terms to exploit that B_p conserves moments and get

$$(3.15) \qquad \int_F (E_p\sigma)q = \int_F (B_p\sigma)q + \underbrace{\int_F (A_p\sigma - B_pA_p\sigma)q}_{=0} = \int_F \{\!\{\sigma\}\!\} q.$$

Arguing similarly, we obtain also that the element moments in (3.11) are conserved.

Finally, we turn to the claimed stability bound. Let $\sigma \in S_p^0$ and write

$$\|\nabla_{\mathcal{M}}(\sigma - E_p\sigma)\|_{L^2(\Omega)} \leq \|\nabla_{\mathcal{M}}(\sigma - A_p\sigma)\|_{L^2(\Omega)} + \|\nabla B_p(\sigma - A_p\sigma)\|_{L^2(\Omega)}.$$

In order to bound the right-hand side, we fix a mesh element $K \in \mathcal{M}$ and consider the first term. Employing $\Phi_{z|K}^p = \Psi_{K,z}^p$ and (3.6) and then Lemma 3.3, we obtain

$$\|\nabla(\sigma - A_p\sigma)\|_{L^2(K)} \leq \sum_{z \in \mathcal{L}_p(K)} |\sigma_{|K}(z) - A_p\sigma(z)| \, \|\nabla\Phi_z^p\|_{L^2(K)}$$

$$(3.16) \qquad \leq C_{d,\gamma_{\mathcal{M}},p} \sum_{z \in \mathcal{L}_p(K)} |\sigma_{|K}(z) - A_p\sigma(z)| \frac{|K|^{\frac{1}{2}}}{\rho_K}$$

$$\leq C_{d,\gamma_{\mathcal{M}},p} \sum_{z \in \mathcal{L}_p(K)} \sum_{F' \in \mathcal{F}, F' \ni z} \frac{|K|^{\frac{1}{2}}}{\rho_K |F'|^{\frac{1}{2}}} \|[\![\sigma]\!]\|_{L^2(F')}.$$

If $K' \in \mathcal{M}$ contains a face F' of the sum, then (3.5) implies

$$\frac{|K|^{\frac{1}{2}}}{\rho_K |F'|^{\frac{1}{2}}} \leq \frac{h_K}{\rho_K}\left(\frac{h_K^{d-2}}{\rho_{K'}^{d-1}}\right)^{\frac{1}{2}} \lesssim \rho_{K'}^{-\frac{1}{2}} \lesssim h_{F'}^{-\frac{1}{2}}.$$

Consequently, with the help of $\#\{K' \in \mathcal{M} \mid K' \subseteq \omega_K\} \leq C_{d,\gamma_{\mathcal{M}}}$, we arrive at

$$(3.17) \qquad \|\nabla(\sigma - A_p\sigma)\|_{L^2(K)} \lesssim \left(\sum_{F \in \mathcal{F}, F \cap K \neq \emptyset} h_F^{-1}\|[\![\sigma]\!]\|_{L^2(F)}^2\right)^{\frac{1}{2}}.$$

Next, consider the second term and observe that (3.1) gives

$$\sup_{r \in \mathbb{P}_{p-2}(K)} \frac{\int_K (\sigma - A_p\sigma)r}{\|r\|_{L^2(K)}} \leq C_{d,p} |K|^{\frac{1}{2}} \sum_{z \in \mathcal{L}_p(K)} |\sigma_{|K}(z) - A_p\sigma(z)|$$

and, for every $F \in \mathcal{F}_K$,

$$\sup_{q \in \mathbb{P}_{p-1}(F)} \frac{\int_F (\{\!\{\sigma\}\!\} - A_p \sigma) q}{\|q\|_{L^2(F)}} \le C_{d,p} |F|^{\frac{1}{2}} \sum_{K' \supset F} \sum_{z \in \mathcal{L}_p(F)} |\sigma_{|K'}(z) - A_p \sigma(z)| .$$

Inserting these two bounds in the stability estimate of Lemma 3.2, we find essentially the bound after the second inequality in (3.16) and so also

$$(3.18) \qquad \|\nabla B_p(\sigma - A_p \sigma)\|_{L^2(K)} \lesssim \left(\sum_{F \in \mathcal{F}, F \cap K \ne \emptyset} h_F^{-1} \| [\![\sigma]\!] \|_{L^2(F)}^2 \right)^{\frac{1}{2}} .$$

We arrive at the claimed inequality by summing (3.17) and (3.18) over all $K \in \mathcal{M}$, observing that the number of elements touching a given face is $\le C_{d,\gamma_\mathcal{M}}$. □

The smoothing operator E_p in Proposition 3.4 is computationally feasible. In fact, we have that

- it suffices to know the evaluations $\langle f, \Phi_z^p \rangle$ for $z \in \mathcal{L}_p^i$ as well as $\langle f, \Phi_z^{p-1} \Phi_F \rangle$ for $F \in \mathcal{F}^i$, $z \in \mathcal{L}_{p-1}(F)$, and $\langle f, \Phi_z^{p-2} \Phi_K \rangle$ for $K \in \mathcal{M}$, $z \in \mathcal{L}_{p-2}(K)$,
- the support of each $E_p \Psi_{K,z}^p$ is contained in ω_z,
- the operators Q_F and Q_K in (3.12) and (3.13) can be implemented via matrices associated with a reference element and, for $d = 2$, Q_F can be diagonalized by means of Legendre polynomials.

After having found a suitable smoothing operator, we now choose the bilinear form $d(\cdot, \cdot)$. Recall that, due to (3.10), the bilinear form $(\cdot, E_p \cdot)_{1;\eta}$ is degenerate and so $d(\cdot, \cdot)$ needs to be nontrivial. There are several choices; see, e.g., Arnold et al. [3]. Here we shall discuss the interplay between E_p and some of them.

A quasi-optimal NIP method. One possibility to achieve nondegeneracy is to employ the jump penalization in $(\cdot, \cdot)_{1;\eta}$. If, in addition, we neutralize the downgrading of coercivity due to $-\int_\Sigma \{\!\{\nabla s\}\!\} \cdot n \, [\![\sigma]\!]$ in $(\cdot, E_p \cdot)_{1;\eta}$, we end up with

$$(3.19) \quad b_{\mathrm{nip}} := (\cdot, E_p \cdot)_{1;\eta} + d_{\mathrm{nip}} \quad \text{with} \quad d_{\mathrm{nip}}(s, \sigma) := \int_\Sigma [\![s]\!] \{\!\{\nabla \sigma\}\!\} \cdot n + \frac{\eta}{h} [\![s]\!] \, [\![\sigma]\!]$$

and reestablish the bilinear form of the nonsymmetric interior penalty (NIP) method introduced in [19], according to Lemma 3.1 and Proposition 3.4.

LEMMA 3.5 (b_{nip} and extended energy norm). *For any penalty parameter $\eta > 0$, we have*

$$\forall s, \sigma \in S_p^0 \quad b_{\mathrm{nip}}(s, s) \ge |s|_{1;\eta}^2 \quad \text{and} \quad b_{\mathrm{nip}}(s, \sigma) \le \left(1 + \sqrt{\eta^{-1} \eta_*} \right) |s|_{1;\eta} \, |\sigma|_{1;\eta} ,$$

where $\eta_* > 0$ depends on d, p, and $\gamma_\mathcal{M}$.

Hence, if the penalty parameter η is not too small, we may consider $|\cdot|_{1;\eta}$ with the same η to be the discrete energy norm associated with b_{nip}. Remarkably, as $\eta \to \infty$, the coercivity and continuity constants tend to their respective counterparts of the limiting conforming Galerkin method in S_p^1.

Proof. The coercivity bound holds by construction. For the continuity bound, we observe that, if $F \in \mathcal{F}_K$ is a face of any $K \in \mathcal{M}$, we have the inverse estimate $\| \cdot \|_{L^2(F)} \le C_{d,\gamma_\mathcal{M},p} h_F^{-\frac{1}{2}} \| \cdot \|_{L^2(K)}$ in $\mathbb{P}_{p-1}(K)$ and set $\eta_* := (d+1) C_{d,\gamma_\mathcal{M},p}^2$. Then

$$(3.20) \qquad \| h^{\frac{1}{2}} \{\!\{\nabla \sigma\}\!\} \|_{L^2(\Sigma)}^2 \le \eta_* \| \nabla_\mathcal{M} \sigma \|_{L^2(\Omega)}^2$$

and the claimed continuity bound follows by standard steps. □

We thus arrive at $M_{\mathrm{nip}} = (S_p^0, b_{\mathrm{nip}}, E_p)$, a new variant of the NIP method of order p with the discrete problem

$$(3.21) \qquad U \in S_p^0 \quad \text{such that} \quad \forall \sigma \in S_p^0 \ b_{\mathrm{nip}}(U, \sigma) = \langle f, E_p\sigma \rangle.$$

Since $b_{\mathrm{nip}} = (\cdot, \cdot)_1$ on $S_p^1 \times S_p^1$ and $E = \mathrm{Id}$ on S_p^1, this is a nonconforming Galerkin method. In contrast to the original NIP method, it applies to any load $f \in H^{-1}(\Omega)$ and has the following property.

THEOREM 3.6 (quasi-optimality of M_{nip}).   *For any $\eta > 0$, the method M_{nip} is $|\cdot|_{1;\eta}$-quasi-optimal for the Poisson problem* (1.1) *with constant $\leq \sqrt{1 + C_{d,\gamma_{\mathcal{M}},p}\eta^{-1}}$.*

Proof. After using the combination of Lemma 3.1 and Proposition 3.4 in Theorem 2.2, it remains to bound the consistency measure δ from (iii). Let $v \in H_0^1(\Omega)$, $\sigma \in S_p^0$ and denote by $\Pi_{\eta,p}$ the $(\cdot, \cdot)_{1;\eta}$-orthogonal projection onto S_p^0. Then $[\![v]\!] = 0 = [\![E_p\sigma]\!]$ and the definition of $\Pi_{\eta,p}$ imply

$$b_{\mathrm{nip}}(\Pi_{\eta,p}v, \sigma) - (v, E_p\sigma)_1 = (\Pi_{\eta,p}v - v, E_p\sigma - \sigma)_1 + \int_\Sigma [\![\Pi_{\eta,p}v - v]\!] \{\!\{\nabla\sigma\}\!\} \cdot n,$$

whence Proposition 3.4 and (3.20) yield

$$(3.22) \qquad |b_{\mathrm{nip}}(\Pi_{\eta,p}v, \sigma) - (v, E_p\sigma)_1| \leq C_{d,\gamma_{\mathcal{M}},p}\eta^{-\frac{1}{2}} |\sigma|_{1;\eta} |\Pi_{\eta,p}v - v|_{1;\eta}.$$

We thus conclude $\delta \lesssim \eta^{-\frac{1}{2}}$ with the help of the coercivity bound in Lemma 3.5. $\square$

A quasi-optimal SIP method. The NIP bilinear form b_{nip} arises in particular by enforcing coercivity. As an alternative, one can achieve symmetry by changing the sign of the first term in d_{nip}. This leads to the SIP bilinear form b_{sip}; cf. (1.2). While b_{sip} verifies the same continuity bound as b_{nip}, the coercivity bound can be replaced as follows. Inequality (3.20) implies

$$\left| \int_\Sigma [\![s]\!] \{\!\{\nabla\sigma\}\!\} \cdot n \right| \leq \frac{1}{2}\sqrt{\eta_*\eta^{-1}} \left(\eta \|h^{-\frac{1}{2}} [\![s]\!] \|_{L^2(\Sigma)}^2 + \| \nabla_{\mathcal{M}} \sigma \|_{L^2(\Omega)}^2 \right),$$

from which we get

$$(3.23) \qquad \forall s \in S_p^0 \quad b_{\mathrm{sip}}(s, s) \geq \alpha(\eta_*\eta^{-1}) |s|_{1;\eta}^2 \quad \text{with} \quad \alpha(t) = 1 - \sqrt{t}.$$

Hence, if $\eta > \eta_*$, then the discrete problem

$$(3.24) \qquad U \in S_p^0 \quad \text{such that} \quad \forall \sigma \in S_p^0 \ b_{\mathrm{sip}}(U, \sigma) = \langle f, E_p\sigma \rangle$$

is well-posed and gives rise to a new variant of the SIP method, which is a nonconforming Galerkin method and denoted by M_{sip}. The following theorem covers the results illustrated in the introduction (section 1) and is proved as Theorem 3.6.

THEOREM 3.7 (quasi-optimality of M_{sip}).   *For any $\eta > \eta_*$, the method M_{sip} is $|\cdot|_{1,\eta}$-quasi-optimal for* (1.1) *with constant $\leq \sqrt{1 + C_{d,\gamma_{\mathcal{M}},p}(\alpha(\eta_*/\eta)^2\eta)^{-1}}$.*

For $\eta \to \infty$, we again end up in Céa's lemma for the limiting conforming Galerkin method in S_p^1.

High-order smoothing with first-order averaging. Assume that $p \geq 2$. The simplified averaging operator A_1 is defined also on S_p^0 and so we may consider

$$(3.25) \qquad \widetilde{E}_p \sigma := A_1 \sigma + B_p(\sigma - A_1 \sigma), \quad \sigma \in S_p^0,$$

which is cheaper to evaluate than E_p. In order to assess this idea, let us first check in which sense A_1 can provide stabilization. Since it is not a projection onto S_p^1 with $p \geq 2$, the conclusions in Lemma 3.3 have to be modified and, in particular, an error bound solely in jump terms is not possible.

LEMMA 3.8 (first-order averaging for higher-order piecewise polynomials). *Let $p \geq 2$, $K \in \mathcal{M}$, and $F \in \mathcal{F}_K$. For all $z \in \mathcal{L}_p(K) \cap F$ and all $\sigma \in S_p^0$, we have*

$$\left| \sigma_{|K}(z) - A_1 \sigma(z) \right| \leq C_{d,p} \left(\sum_{F' \cap F \neq \emptyset} \frac{1}{|F'|} \left| \int_{F'} [\![\sigma]\!] \right| + \sum_{K' \cap F \neq \emptyset} \frac{h_{K'}}{|K'|^{\frac{1}{2}}} \|\nabla \sigma\|_{L^2(K')} \right),$$

where F' and K' vary, respectively, in $\mathcal{F}$ and $\mathcal{M}$.

Proof. Given any $z \in \mathcal{L}_p(K) \cap F$, Lemma 3.1 in [24] ensures

$$(3.26) \qquad \left| \sigma_{|K}(z) - A_p \sigma(z) \right| \leq \sum_{F' \ni z} \frac{1}{|F'|} \left| \int_{F'} [\![\sigma]\!] \right| + C_{d,p} \sum_{K' \ni z} \frac{h_{K'}}{|K'|^{\frac{1}{2}}} \|\nabla \sigma\|_{L^2(K')}.$$

We distinguish two cases, depending on whether or not z is a vertex.

Case 1: $z \in \mathcal{L}_1(K)$. Then we have $A_1 \sigma(z) = A_p \sigma(z)$ and the claimed estimate follows from (3.26).

Case 2: $z \in \mathcal{L}_p(K) \setminus \mathcal{L}_1(K)$. Since $A_1 \sigma_{|F} \in \mathbb{P}_1(F)$ and $\sum_{y \in \mathcal{L}_1(F)} \lambda_y^F = 1$ on F, we may write

$$(3.27) \qquad |\sigma_{|K}(z) - A_1 \sigma(z)| \leq \sum_{y \in \mathcal{L}_1(F)} \left| \sigma_{|K}(z) - A_1 \sigma(y) \right| \lambda_y^F(z)$$

and, for any $y \in \mathcal{L}_1(F)$,

$$\left| \sigma_{|K}(z) - A_1 \sigma(y) \right| \leq \left| \sigma_{|K}(z) - \sigma_{|K}(y) \right| + \left| \sigma_{|K}(y) - A_1 \sigma(y) \right|.$$

As the second term of the right-hand side is already bounded in Case 1, it remains to bound the first term. Writing c for the mean value of σ in K, we deduce

$$\left| \sigma_{|K}(z) - \sigma_{|K}(y) \right| \leq \left| \sigma_{|K}(z) - c \right| + \left| \sigma_{|K}(y) - c \right|$$
$$\lesssim |F|^{-\frac{1}{2}} \|\sigma_{|K} - c\|_{L^2(F)} \lesssim h_K |K|^{-\frac{1}{2}} \|\nabla \sigma\|_{L^2(K)}$$

with the help of an inverse estimate in $\mathbb{P}_p(F)$ and [22, Lemma 3]. $\qquad \square$

Using this lemma in the proof of Proposition 3.4, we obtain the following properties of $\widetilde{E}_p$.

PROPOSITION 3.9 (moment conservation with first-order averaging). *The linear operator $\widetilde{E}_p$ from (3.25) is invariant on S_1^1 and satisfies (3.11), and for all $\sigma \in S_p^0$,*

$$\| \nabla_{\mathcal{M}} (\sigma - \widetilde{E}_p \sigma) \|_{L^2(\Omega)} \leq C_{d,\gamma_{\mathcal{M}},p} \left(\sum_{F \in \mathcal{F}} h_F^{-d} \left| \int_F [\![\sigma]\!] \right|^2 + \| \nabla_{\mathcal{M}} \sigma \|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}.$$

Combining the new smoothing operator $\widetilde{E}_p$ with one of the previous bilinear forms b_{var}, var $\in \{\text{nip}, \text{sip}\}$, leads to a nonconforming method $\widetilde{M}_{\text{var}}$ with discrete problem

$$(3.28) \qquad U \in S_p^0 \quad \text{such that} \quad \forall \sigma \in S_p^0 \ \ b_{\text{var}}(U, \sigma) = \langle f, \widetilde{E}_p \sigma \rangle,$$

which is well-posed for all $\eta > \eta_{\text{var}}$. Hereafter

$$\eta_{\text{var}} := \begin{cases} 0, & \text{if var} = \text{nip}, \\ \eta_*, & \text{if var} = \text{sip}, \end{cases} \quad \text{and} \quad \alpha_{\text{var}}(t) := \begin{cases} 1, & \text{if var} = \text{nip}, \\ 1 - \sqrt{t}, & \text{if var} = \text{sip}. \end{cases}$$

As $\widetilde{E}_p$ is only invariant on the strict subset S_1^1 of (3.9) for $p \geq 2$, the method $\widetilde{M}_{\text{var}}$ is *not* a nonconforming Galerkin method. Nevertheless, the following holds.

THEOREM 3.10 (quasi-optimality of $\widetilde{M}_{\text{var}}$).   *Let* var $\in \{\text{nip}, \text{sip}\}$. *If* $\eta > \eta_{\text{var}}$, *the method* $\widetilde{M}_{\text{var}}$ *is* $|\cdot|_{1;\eta}$*-quasi-optimal for the Poisson problem* (1.1) *with constant* $\leq C_{d, \gamma_{\mathcal{M}}, p} \sqrt{1 + (\alpha_{\text{var}}(\eta_*/\eta)^2 \eta)^{-1}}$.

Proof. Proceed as in the proof of Theorem 3.6 or as indicated for Theorem 3.7, replacing E_p by $\widetilde{E}_p$. The only difference is that, in the derivation of the counterpart of (3.22), we use

$$\sum_{F \in \mathcal{F}} h_F^{-d} \left| \int_F [\![\sigma]\!] \right|^2 \lesssim \sum_{F \in \mathcal{F}} h_F^{-1} \int_F |[\![\sigma]\!]|^2$$

and obtain only

$$|b_{\text{var}}(\Pi_{\eta, p} v, \sigma) - (v, \widetilde{E}_p \sigma)_1| \leq C_{d, \gamma_{\mathcal{M}}, p} \sqrt{1 + \left(\alpha_{\text{var}}(\eta_*/\eta)^2 \eta \right)^{-1}} \, |\sigma|_{1;\eta} \, |\Pi_{\eta, p} v - v|_{1;\eta}$$

because the stability bound in Proposition 3.9 involves gradient terms. □

3.3. A quasi-optimal and locking-free method for linear elasticity. The goal of this section is to conceive a quasi-optimal and locking-free method for linear elasticity.

Given $\Omega \subseteq \mathbb{R}^d$ as in section 3.1, we consider the displacement formulation of the linear elasticity problem with pure displacement boundary conditions: find $u \in H_0^1(\Omega)^d$ such that

$$(3.29) \qquad -\operatorname{div}\left(2\mu\, \varepsilon(u) + \lambda \operatorname{div}(u) \right) = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega.$$

Hereafter $\varepsilon(v) := (\nabla v + \nabla v^T)/2$ is the symmetric gradient and $\mu, \lambda > 0$ are the Lamé coefficients. We shall mostly suppress the dependencies on μ in the notation, while we trace the ones on λ.

Let $\mathcal{M}$ be a mesh of Ω as in section 3.1 and, for $\eta \geq 0$, define

$$(3.30) \qquad a_{\lambda;\eta}(v, w) := \int_\Omega 2\mu\, \varepsilon_{\mathcal{M}}(v) : \varepsilon_{\mathcal{M}}(w) + \lambda \operatorname{div}_{\mathcal{M}} v \operatorname{div}_{\mathcal{M}} w + \int_\Sigma \frac{\mu\eta}{h} [\![v]\!] \cdot [\![w]\!],$$

$$\|v\|_{\lambda;\eta} = a_{\lambda;\eta}(v, v)^{\frac{1}{2}}$$

for $v, w \in H^1(\mathcal{M})^d$ and abbreviate $a_{\lambda;0}$ to a_λ. The colon indicates the matrix scalar product $G : H = \sum_{j,\ell=1}^d G_{j\ell} H_{j\ell}$. We aim at applying Theorem 2.2 with the following setting:

$$(3.31) \qquad V = H_0^1(\Omega)^d, \quad S \subseteq (S_1^0)^d, \quad \widetilde{a} = a_{\lambda;\eta} \text{ on } \widetilde{V} = H_0^1(\Omega)^d + S,$$

where S will be specified below. Notice that $a_{\lambda;\eta}$ is then a scalar product and (2.1) provides a weak formulation of (3.29).

We readily deduce the following counterpart of Lemma 3.1.

LEMMA 3.11 (moment conservation). *If a linear operator $E : S \to H_0^1(\Omega)^d$ satisfies*

$$(3.32) \qquad \forall \sigma \in S, \ F \in \mathcal{F}^i \quad \int_F E\sigma = \int_F \{\!\!\{\sigma\}\!\!\},$$

then, for all $s, \sigma \in S$,

$$a_{\lambda;\eta}(s, E\sigma) = a_\lambda(s, \sigma) - \int_\Sigma \left(\{\!\!\{2\mu\,\varepsilon_{\mathcal{M}}(s) + \lambda\,\mathrm{div}_{\mathcal{M}}(s)I\}\!\!\} \right) n \cdot [\![\sigma]\!].$$

In section 3.2, the impact on coercivity or symmetry of the counterpart of the term $\int_\Sigma(\{\!\!\{2\mu\,\varepsilon_{\mathcal{M}}(s) + \lambda\,\mathrm{div}_{\mathcal{M}}(s)I\}\!\!\})n \cdot [\![\sigma]\!]$ was compensated only with the help of $d(\cdot,\cdot)$. Here we shall handle it also with the choice of the discrete space S. More precisely, if we choose the Crouzeix–Raviart space

$$(3.33) \qquad S = CR^d \quad \text{with} \quad CR := \{s \in S_1^0 \mid \forall F \in \mathcal{F} \int_F [\![s]\!] = 0\}$$

with homogeneous boundary conditions, then this term vanishes because, on each face $F \in \mathcal{F}$, the average $(\{\!\!\{2\mu\,\varepsilon_{\mathcal{M}}(s) + \lambda\,\mathrm{div}_{\mathcal{M}}(s)I\}\!\!\})n$ is a constant. Another advantage of setting $S = CR^d$ is that (3.32) becomes $\int_F E\sigma = \int_F \sigma$, $F \in \mathcal{F}^i$, since the integral on the faces of $\mathcal{M}$ is well-defined for functions in CR. We shall exploit this observation in Lemma 3.15 below. The conforming part of CR^d is

$$CR^d \cap H_0^1(\Omega)^d = (S_1^1)^d,$$

which is a strict subspace for $\#\mathcal{M} > 1$. Finally, Arnold [2] shows that, for certain choices of Ω and $\mathcal{M}$, there is a nonzero function

$$(3.34) \qquad s_0 \in CR^2 \setminus \{0\} \quad \text{with} \quad \varepsilon_{\mathcal{M}}(s_0) = 0 \text{ and } \mathrm{div}_{\mathcal{M}}\, s_0 = 0,$$

entailing, in contrast to the setting of [24, section 3.2] for the Poisson problem,

$$0 \neq s_0 \in CR^2 \cap (H_0^1(\Omega)^2)^\perp$$

so that overconsistency is in general ruled out by Lemma 2.3.

As (3.32) is the vector version of (3.11) for $p = 1$, we can take the computionally feasible smoothing operator E_1 from Proposition 3.4 componentwise. We denote this vector version again by E_1. Since $a_{\lambda;\eta}(\cdot, E_1\cdot)$ may be degenerate in view of (3.34), we take

$$(3.35) \quad b_{\mathrm{HL}} := a_{\lambda;\eta}(\cdot, E_1\cdot) + d_{\mathrm{HL}} \quad \text{with} \quad d_{\mathrm{HL}}(s, \sigma) = \int_\Sigma \frac{\mu\eta}{h}\,[\![s]\!] \cdot [\![\sigma]\!] \quad \text{with} \quad \eta > 0,$$

which is the discrete bilinear form in Hansbo and Larson [15, equation (26)]. We thus introduce a new penalized Crouzeix–Raviart method $M_{\mathrm{HL}} = (CR^d, b_{\mathrm{HL}}, E_1)$ given by the following discrete problem: find $U \in CR^d$ such that

$$(3.36) \quad \int_\Omega \left(2\mu\,\varepsilon_{\mathcal{M}}(U) : \varepsilon_{\mathcal{M}}(\sigma) + \lambda\,\mathrm{div}_{\mathcal{M}}\,U\,\mathrm{div}_{\mathcal{M}}\,\sigma\right) + \int_\Sigma \frac{\mu\eta}{h}\,[\![s]\!] \cdot [\![\sigma]\!] = \langle f, E_1\sigma\rangle$$

for all $\sigma \in CR^d$. The method M_{HL} is a nonconforming Galerkin method. The modification of the right-hand side with respect to [15] allows one to apply any $H^{-1}(\Omega)$-volume force with the following a priori error control.

THEOREM 3.12 (quasi-optimality of M_{HL}).   *The method M_{HL} is $\|\cdot\|_{\lambda;\eta}$-quasi-optimal for* (3.29) *with constant* $\leq \sqrt{1 + C_{d,\gamma_{\mathcal{M}}}(2\mu + \lambda)\eta^{-1}}$.

Proof. We first use Lemma 3.11 and Proposition 3.4 for $p = 1$ in Theorem 2.2. Then it remains to bound δ in item (iii) of Theorem 2.2. Let $v \in H_0^1(\Omega)^d$, $\sigma \in CR^d$, and denote by $\Pi_{\lambda;\eta}$ the $a_{\lambda;\eta}$-orthogonal projection onto CR^d. Lemma 3.11, the definition of CR, $[\![v]\!] = 0 = [\![E_1\sigma]\!]$, and the definition of $\Pi_{\lambda;\eta}$ imply

$$b_{\mathrm{HL}}(\Pi_{\lambda;\eta}v, \sigma) - a_\lambda(v, E_1\sigma) = a_\lambda(\Pi_{\lambda;\eta}v - v, E_1\sigma - \sigma)$$

and so Proposition 3.4 yields

$$|b_{\mathrm{HL}}(\Pi_{\lambda;\eta}v, \sigma) - a_\lambda(v, E_1\sigma)| \leq C_{d,\gamma_{\mathcal{M}}}\sqrt{2\mu + \lambda}\,\eta^{-1/2}\|\Pi_{\lambda;\eta}v - v\|_{\lambda;\eta}\|\sigma\|_{\lambda;\eta}.$$

Hence, we have $\delta \lesssim \sqrt{2\mu + \lambda}\,\eta^{-\frac{1}{2}}$ and the proof is finished.   □

The following remarks show that the upper bound of the quasi-optimality constant C_{qopt} in Theorem 3.12 captures the correct asymptotic behavior not only for the conforming limit $\eta \to \infty$.

Remark 3.13 (C_{qopt} as $\eta \to 0$). The degeneracy of the bilinear form $a_{\lambda;\eta}(\cdot, E_1\cdot)$ entails $C_{\mathrm{qopt}} \geq C_\lambda \eta^{-\frac{1}{2}}$. To see this, suppose that s_0 satisfies (3.34) and notice that identity (3.32) and [24, Lemma 3.2] guarantee that E_1 is injective. We then have that $\|E_1 s_0\|_{\lambda;\eta} = C_\lambda \neq 0$ and $\|s_0\|_{\lambda;\eta} = C\eta^{\frac{1}{2}}$. Hence, Theorem 2.2 yields $C_{\mathrm{qopt}} \geq C_{\mathrm{stab}} \geq C_\lambda \eta^{-\frac{1}{2}}$.

The following remark is closely connected with Linke [17, section 2] concerning incompressible flows.

Remark 3.14 (deterioration of C_{qopt} for nearly incompressible materials). The property

$$(3.37) \qquad E_1\big(\{s \in CR^d \mid \mathrm{div}_{\mathcal{M}}\, s = 0\}\big) \nsubseteq \{v \in H_0^1(\Omega)^d \mid \mathrm{div}\, v = 0\}$$

results in $C_{\mathrm{qopt}} \geq C_\eta \lambda^{\frac{1}{2}}$. Indeed, if $s \in CR^d$ satisfies $\mathrm{div}_{\mathcal{M}}\, s = 0$ and $\mathrm{div}(E_1 s) \neq 0$, we have $\|s\|_{\lambda;\eta} = C_\eta$ and $\|E_1 s\|_{\lambda;\eta} \approx C\lambda^{\frac{1}{2}}$ as $\lambda \to \infty$ and so Theorem 2.2 implies $C_{\mathrm{qopt}} \geq C_{\mathrm{stab}} \geq C_\eta \lambda^{\frac{1}{2}}$.

In order to verify (3.37), fix any face $F \in \mathcal{F}^i$ of a given mesh $\mathcal{M}$. Let Ψ_F be the associated basis function in CR with $\int_{F'} \Psi_F = \delta_{FF'}$ for all $F' \in \mathcal{F}$ and $\Psi_{F|K} = 0$ whenever $F \notin \mathcal{F}_K$. Then, appropriately picking the elements K_z in the definition (3.14) of A_1, we can arrange $A_1\Psi_F = 0$ and so $E_1\Psi_F = \beta\Phi_F$ with some $\beta > 0$ and Φ_F as in (3.12). Consider $\Psi_F t_F \in CR^d$, where t_F is a unit tangent vector of F. On the one hand, we have $\mathrm{div}_{\mathcal{M}}(t_F\Psi_F) = t_F \cdot \nabla_{\mathcal{M}}\Psi_F = 0$ and, on the other hand, $\mathrm{div}\, E_d(t_F\Psi_F) = \beta\,\mathrm{div}(t_F\Phi_F) = \beta t_F \cdot \nabla\Phi_F \neq 0$.

It is instructive to shed additional light on the performance of M_{HL} for nearly incompressible materials. First, recall that the space S_1^1 shows locking whenever $\{s \in S_1^1 \mid \mathrm{div}\, s = 0\}$ provides poor approximation; see [9, section 11.3]. Hence the choice $\eta \approx \lambda$ will also result in poor approximation for large λ. For fixed penalty parameter $\eta > 0$, the following lemma, which is also of interest by its own, will be

useful. It quantifies the difference between the original method $\hat{M}_{\mathrm{HL}}$ of Hansbo and Larson, and its new variant M_{HL}. Recall that, if $f \in L^2(\Omega)^d$, the discrete solution $\hat{U} \in CR^d$ of Hansbo and Larson is given by

$$(3.38) \qquad \forall \sigma \in CR^d \quad b_{\mathrm{HL}}(\hat{U}, \sigma) = \int_\Omega f \cdot \sigma.$$

LEMMA 3.15 (M_{HL} and $\hat{M}_{\mathrm{HL}}$). *Assume* $f \in L^2(\Omega)^d$ *and let* $U, \hat{U} \in CR^d$ *verify* (3.36) *and* (3.38), *respectively. Then*

$$\|U - \hat{U}\|_{\lambda;\eta} \le C_{d,\gamma_{\mathcal{M}}} \eta^{-\frac{1}{2}} \left(\sum_{K \in \mathcal{M}} h_K^2 \|f\|_{L^2(K)}^2 \right)^{\frac{1}{2}}.$$

Proof. The definitions of U and $\hat{U}$ immediately give

$$\|U - \hat{U}\|_{\lambda;\eta} = \sup_{\|\sigma\|_{\lambda;\eta}=1} \left| b_{\mathrm{HL}}(U - \hat{U}, \sigma) \right| = \sup_{\|\sigma\|_{\lambda;\eta}=1} \left| \int_\Omega f \cdot (E_1 \sigma - \sigma) \right|,$$

where σ varies in CR^d. For any element $K \in \mathcal{M}$, we have $\int_{\partial K} E_1 \sigma - \sigma = 0$ implying the Poincaré inequality $\|E_1 \sigma - \sigma\|_{L^2(K)} \lesssim h_K \|\nabla(E_1 \sigma - \sigma)\|_{L^2(K)}$. Therefore,

$$\left| \int_\Omega f \cdot (E_1 \sigma - \sigma) \right| \lesssim \left(\sum_{K \in \mathcal{M}} h_K^2 \|f\|_{L^2(K)}^2 \right)^{\frac{1}{2}} \|\nabla_{\mathcal{M}}(E_1 \sigma - \sigma)\|_{L^2(\Omega)}.$$

Hence, Proposition 3.4 and [24, Proposition 3.3] followed by Brenner [8, Theorem 3.1] finish the proof. □

We readily see from this proof that the asymptotic closeness of U and $\hat{U}$ could be increased by requiring that the smoothing operator also conserves element moments.

A consequence of Lemma 3.15 is the following equivalence concerning the asymptotic error bounds

$$\|u - U\|_{\lambda;\eta} \le Ch\|f\|_{L^2(\Omega)}, \quad \|u - \hat{U}\|_{\lambda;\eta} \le \hat{C}h\|f\|_{L^2(\Omega)}$$

with best constants C and $\hat{C}$ for all $h := \max_{K \in \mathcal{M}} h_K$ and $f \in L^2(\Omega)^d$:

$$(3.39) \qquad C \text{ is independent of } \lambda \iff \hat{C} \text{ is independent of } \lambda.$$

Therefore, the robustness result [15, Theorem 3.1], which ensures that $\hat{C}$ is independent of λ for polygons $\Omega \subseteq \mathbb{R}^2$, carries over to M_{HL}. In summary, for smooth volume forces, the method M_{HL} is locking-free. The nonrobustness of the quasi-optimality constant is thus due to rough volume forces, including forces for which the locking-free nonconforming methods in Falk [13], Brenner and Sung [10], and Hansbo and Larson [15] are not defined.

Let us conclude this section with a remark on the generalization to order $p \ge 2$, where CR is replaced by its higher-order counterpart CR_p from Stoyan and Baran [21]. This case is of different nature. In fact, the Korn inequalities of Brenner [8] ensure that $\|\cdot\|_{\lambda;\eta}$ is a norm on $H_0^1(\Omega)^d + CR_p^d$ even for $\eta = 0$. This allows one to construct overconsistent methods with the help of E_p from Proposition 3.4.

3.4. A quasi-optimal C^0 interior penalty method for the biharmonic problem. In this subsection, we introduce a new C^0 interior penalty method for the biharmonic problem with clamped boundary conditions,

$$(3.40) \qquad \Delta^2 u = f \text{ in } \Omega, \quad u = \partial_n u = 0 \text{ on } \partial\Omega,$$

and prove its quasi-optimality. We let Ω and $\mathcal{M}$ be as in section 3.1 with $d = 2$. Jumps and averages of vector- and matrix-valued maps are intended componentwise. Consequently, if $v \in H^2(\mathcal{M})$, then $[\![\nabla v]\!] \cdot n$ and $\{\!\{\nabla v\}\!\} \cdot n$ indicate, respectively, the jump and the average of the normal derivative of v on the skeleton Σ. We write also $[\![\partial^2 v/\partial n^2]\!]$ and $\{\!\{\partial^2 v/\partial n^2\}\!\}$ in place of $([\![D^2 v]\!]\, n)\cdot n$ and $(\{\!\{D^2 v\}\!\}\, n)\cdot n$, respectively. Given $\eta \geq 0$, set

$$(3.41) \qquad (v,w)_{2;\eta} := \int_\Omega \mathrm{D}^2_\mathcal{M} v : \mathrm{D}^2_\mathcal{M} w + \int_\Sigma \frac{\eta}{h} \left([\![\nabla v]\!] \cdot n\right)\left([\![\nabla w]\!] \cdot n\right),$$

$$|v|_{2;\eta} := (v,v)_{2;\eta}^{\frac{1}{2}}$$

for $v, w \in H^2(\mathcal{M})$ and abbreviate $(\cdot,\cdot)_{2;0}$ to $(\cdot,\cdot)_2$. Recalling (3.4), consider the following setting for Theorem 2.2:

$$(3.42) \qquad V = H_0^2(\Omega), \quad S = S_2^1, \quad \widetilde{a} = (\cdot,\cdot)_{2;\eta} \text{ on } \widetilde{V} = H_0^2(\Omega) + S_2^1,$$

where, for $\eta > 0$, the bilinear form $(\cdot,\cdot)_{2;\eta}$ is a scalar product on

$$(3.43) \qquad H_0^2(\Omega) + S_2^1 \subseteq \{v \in H^2(\mathcal{M}) \mid \forall F \in \mathcal{F} \quad [\![\nabla v]\!] \cdot t_F = 0\}$$

and t_F is a unit tangent vector of F. Then, the abstract problem (2.1) with (3.42) is a weak formulation of the biharmonic problem (3.40). The conforming part of S_2^1 is the strict subspace

$$(3.44) \qquad S_2^1 \cap H_0^2(\Omega) = \{s \in S_2^1 \mid [\![\nabla s]\!] \cdot n = 0\},$$

which may even be trivial; cf. [24, Remark 3.12]. Finally, we have

$$(3.45) \qquad \{0\} \neq S_1^1 \subseteq S_2^1 \cap H_0^2(\Omega)^\perp$$

and, therefore, Lemma 2.3 rules overconsistency out.

Let us turn to the choice of the smoothing operator. Interestingly, Brenner and Sung [11] propose a C^0 interior penalty method M_{BS} involving a smoothing operator based upon averaging. In contrast to similar methods, M_{BS} is well-defined for general loads $\ell \in H^{-2}(\Omega)$, fully stable according to Theorem 2.2(i), and, for any $\alpha > 0$ and all $\ell \in H^{-2+\alpha}(\Omega)$, its error in $|\cdot|_{2;\eta}$ with a suitable η decays at the optimal rate α. Nevertheless, M_{BS} is not guaranteed to be quasi-optimal with respect to $|\cdot|_{2;\eta}$, because it is not designed to be fully algebraically consistent.

To devise a method ensuring full algebraic consistency, we proceed as before and derive the following counterpart of Lemmas 3.1 and 3.11 with the help of integration by parts (3.3).

LEMMA 3.16 (moment conservation). *If the smoothing operator $E : S_2^1 \to H_0^2(\Omega)$ satisfies*

$$(3.46) \qquad \forall \sigma \in S_2^1, F \in \mathcal{F}^i \quad \int_F \nabla E\sigma = \int_F \{\!\{\nabla\sigma\}\!\},$$

then

$$\forall s, \sigma \in S_2^1 \quad (s, E\sigma)_{2;\eta} = \int_\Omega \mathrm{D}_\mathcal{M}^2 \, s : \mathrm{D}_\mathcal{M}^2 \, \sigma - \int_\Sigma \left\{\!\!\left\{ \frac{\partial^2 s}{\partial n^2} \right\}\!\!\right\} \llbracket \nabla \sigma \rrbracket \cdot n.$$

Thanks to $S_2^1 + H_0^2(\Omega) \subseteq C^0(\overline{\Omega})$ and the fundamental theorem of calculus, we may ensure the conservation (3.46) of the mean gradients on faces by

$$(3.47) \qquad \forall z \in \mathcal{L}_1^i \ E\sigma(z) = \sigma(z) \quad \text{and} \quad \forall F \in \mathcal{F}^i \ \int_F \nabla E\sigma \cdot n = \int_F \{\!\!\{ \nabla \sigma \}\!\!\} \cdot n.$$

The smoothing operator for Morley functions in [24] verifies these new requirements. We adapt its construction to the current setting, focusing on the modifications only. Let us begin with the (simplified) averaging operator mapping into the Hsieh–Clough–Tocher (HCT) space

$$HCT := \{s \in C^1(\overline{\Omega}) \mid \forall K \in \mathcal{M} \ s_{|K} \in C^1(K) \cap \mathbb{P}_3(\mathcal{M}_K), \ \ s = \partial_n s = 0 \text{ on } \partial\Omega\},$$

where $\mathcal{M}_K$ stands for the triangulation obtained by connecting each vertex of the triangle K with its barycenter m_K. For each vertex $z \in \mathcal{L}_1^i$ and edge $F \in \mathcal{F}^i$, we pick elements $K_z, K_F \in \mathcal{M}$ containing z or F, respectively, and define

$$(3.48) \quad A_{HCT}\sigma := \sum_{z \in \mathcal{L}_1^i} \left(\sigma(z)\Upsilon_z^0 + \sum_{j=1}^2 \partial_j\big(\sigma_{|K_z}\big)(z)\Upsilon_z^j \right) + \sum_{F \in \mathcal{F}^i} \frac{\partial\big(\sigma_{|K_F}\big)}{\partial n}(m_F)\Upsilon_F,$$

where Υ_z^j with $z \in \mathcal{L}_1^i$, $j \in \{0,1,2\}$ and Υ_F with $F \in \mathcal{F}^i$ form the nodal basis of HCT and m_F denotes the midpoint of F. Next, we introduce the bubble smoother. Given any interior edge $F \in \mathcal{F}^i$, let $K_1, K_2 \in \mathcal{M}$ be the two elements such that $F = K_1 \cap K_2$. Considering their barycentric coordinates $(\lambda_z^{K_i})_{z \in \mathcal{L}_1(K_i)}$, $i = 1, 2$, as first-order polynomials on $\mathbb{R}^2$, we set

$$\bar{\phi}_F := \frac{630}{|F|}\phi_F \quad \text{with} \quad \phi_F := \begin{cases} \prod\limits_{z \in \mathcal{L}_1(F)} \left(\lambda_z^{K_1}\lambda_z^{K_2}\right)^2 & \text{in } K_1 \cup K_2, \\ 0 & \text{in } \Omega \setminus (K_1 \cup K_2). \end{cases}$$

For all $F' \in \mathcal{F}$, we have $\int_{F'} \bar{\phi}_F = \delta_{FF'}$ according to (3.1) and $(\nabla\bar{\phi}_F)_{|F'} = 0$ if $F' \neq F$. We define also $\zeta_F(x) := (x - m_F) \cdot n_F$ for $x \in \mathbb{R}^2$, which satisfies $\zeta_F = 0$ on F and $\nabla\zeta_F = n_F$. Then $\bar{\Phi}_{n_F} := \zeta_F\bar{\phi}_F$ is in $H_0^2(\Omega)$ and we have $\int_{F'} \nabla\bar{\Phi}_{n_F} \cdot n_{F'} = \int_{F'} \bar{\phi}_F \nabla\zeta_F \cdot n_{F'} = \delta_{F,F'}$ for all $F' \in \mathcal{F}^i$. Hence,

$$B_{\partial_n}\sigma := \sum_{F \in \mathcal{F}^i} \left(\int_F \{\!\!\{ \nabla \sigma \}\!\!\} \cdot n \right) \bar{\Phi}_{n_F}$$

maps $S_2^1 + HCT$ into $H_0^2(\Omega)$, verifying $B_{\partial_n}\sigma(z) = 0$ for all $z \in \mathcal{L}_1^i$ and the second part of (3.47). The combination of bubble smoother and averaging thus yields the desired moment conservation in a stable manner.

PROPOSITION 3.17 (stable smoothing with moment conservation). *The linear operator $E_{\mathrm{C0}} : S_2^1 \to H_0^2(\Omega)$ given by*

$$E_{\mathrm{C0}}\sigma := A_{HCT}\sigma + B_{\partial_n}(\sigma - A_{HCT}\sigma)$$

is invariant on $S_2^1 \cap H_0^2(\Omega)$ and verifies (3.46), and for all $\sigma \in S_2^1$,

$$\| \mathrm{D}_{\mathcal{M}}^2(\sigma - E_{\mathrm{C0}}\sigma)\|_{L^2(\Omega)} \leq C_{\gamma_{\mathcal{M}}} \left(\sum_{F \in \mathcal{F}} h_F^{-1} \| [\![\nabla\sigma]\!] \cdot n\|_{L^2(F)}^2 \right)^{\frac{1}{2}}.$$

Proof. We proceed as in [24, Proposition 3.17] with the following difference. For the HCT averaging, we use the following bound: if $K \in \mathcal{M}$ and $z \in \mathcal{L}_2(K)$, then

$$\left|\nabla(\sigma_{|K})(z) - \nabla A_{HCT}\sigma(z)\right| \leq C \sum_{F \in \mathcal{F}: F \ni z} h_F^{-\frac{1}{2}} \| [\![\nabla\sigma]\!] \cdot n\|_{L^2(F)},$$

which follows along the lines of the proof of Lemma 3.3 and from (3.43). □

It remains to choose the bilinear form $d(\cdot, \cdot)$. In view of (3.45), we need to establish nondegeneracy, for example, in the vein of the extended energy norm $|\cdot|_{2;\eta}$. Requiring also symmetry for the resulting discrete bilinear form then leads to

$$d_{\mathrm{BS}}(s, \sigma) = \int_\Sigma [\![\nabla s]\!] \cdot n \left(-\left\{\!\!\left\{\frac{\partial^2 \sigma}{\partial n^2}\right\}\!\!\right\} + \frac{\eta}{h} [\![\nabla\sigma]\!] \cdot n \right)$$

and the discrete bilinear form of Brenner and Sung [11]:

$$(3.49) \qquad b_{\mathrm{BS}}(s, \sigma) = (s, \sigma)_{2;\eta} - \int_\Sigma \left(\left\{\!\!\left\{\frac{\partial^2 s}{\partial n^2}\right\}\!\!\right\} [\![\nabla\sigma]\!] \cdot n + [\![\nabla s]\!] \cdot n \left\{\!\!\left\{\frac{\partial^2 \sigma}{\partial n^2}\right\}\!\!\right\} \right).$$

Similarly to the SIP bilinear form, there is $\eta_* > 0$ depending on $\gamma_{\mathcal{M}}$ such that

$$(3.50) \qquad \|h^{-\frac{1}{2}} \left\{\!\!\left\{\partial^2\sigma/\partial^2 n\right\}\!\!\right\} \|_{L^2(\Sigma)} \leq \eta_* \| \mathrm{D}_{\mathcal{M}}^2 \sigma\|_{L^2(\Omega)}$$

and therefore b_{BS} is $|\cdot|_{2,\eta}$-coercive with constant $\alpha(\eta_*/\eta)$ whenever $\eta > \eta_*$; cf. (3.23) and [11, Lemma 7]. Under this assumption, the discrete problem

$$(3.51) \qquad U \in S_2^1 \text{ such that } \forall \sigma \in S_2^1 \ \ b_{\mathrm{BS}}(U, \sigma) = \langle f, E_{\mathrm{C0}}\sigma\rangle$$

is well-posed and introduces a new C^0 interior penalty method M_{C0} for the biharmonic problem (3.40). Inspecting b_{BS}, E_{C0} and recalling Proposition 3.17, we see that $M_{\mathrm{C0}} = (S_2^1, b_{\mathrm{BS}}, E_{\mathrm{C0}})$ is a nonconforming Galerkin method with a computationally feasible smoothing operator. It differs from the original method of Brenner and Sung [11] in the choice of the smoother and the following property.

THEOREM 3.18 (quasi-optimality of M_{C0}). *For any penalty parameter $\eta > \eta_*$, the method M_{C0} is $|\cdot|_{2;\eta}$-quasi-optimal for the biharmonic problem (3.40) with constant $\leq \sqrt{1 + C_{\gamma_{\mathcal{M}}}(\alpha(\eta_*/\eta)^2\eta)^{-1}}$.*

Proof. Assume $\eta > \eta_*$. Hence b_{BS} is coercive and Theorem 2.2 applies. After making use of Lemma 3.16, Proposition 3.17, and (3.44), it remains to bound δ in (iii) of Theorem 2.2. To this end, we let Π_η denote the $(\cdot, \cdot)_{2;\eta}$-orthogonal projection onto S_2^1 and derive, for all $v \in H_0^2(\Omega)$ and $\sigma \in S_2^1$,

$$b_{\mathrm{BS}}(\Pi_\eta v, \sigma) - (v, E_{\mathrm{C0}}\sigma)_2 = (\Pi_\eta v - v, E_{\mathrm{C0}}\sigma - \sigma)_2 - \int_\Sigma [\![\nabla(\Pi_\eta v - v)]\!] \cdot n \left\{\!\!\left\{\frac{\partial^2 \sigma}{\partial n^2}\right\}\!\!\right\}$$

with the help of $[\![\nabla E_{\mathrm{C0}}\sigma]\!] = 0 = [\![\nabla v]\!]$, Lemma 3.16, and the definition of Π_η. Con-

sequently, Proposition 3.17 and (3.50) yield

$$|b_{\mathrm{BS}}(\Pi_\eta v, \sigma) - (v, E_{\mathrm{C0}}\sigma)_2| \le C_{\gamma_{\mathcal{M}}} \eta^{-\frac{1}{2}} |\Pi_\eta v - v|_{2;\eta} |\sigma|_{2;\eta}.$$

The coercivity of b_{BS} thus implies $\delta^2 \lesssim \left(\alpha(\eta_*/\eta)^2\eta\right)^{-1}$ and the proof is finished.   □

The presented approach may be extended to design quasi-optimal methods of order $p \ge 3$. Perhaps the simplest manner is to keep the HCT averaging A_{HCT} and to construct a higher-order version of the bubble smoother similar to B_p in section 3.2. This will not result in a nonconforming Galerkin method, but achieves quasi-optimality.

Acknowledgment. The authors thank the referees for suggesting improvements of the presentation.

REFERENCES

[1] D. N. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.

[2] D. N. Arnold, *On nonconforming linear-constant elements for some variants of the Stokes equations*, Istit. Lombardo Accad. Sci. Lett. Rend. A, 127 (1993), pp. 83–93.

[3] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2001/02), pp. 1749–1779.

[4] S. Badia, R. Codina, T. Gudi, and J. Guzmán, *Error analysis of discontinuous Galerkin methods for the Stokes problem under minimal regularity*, IMA J. Numer. Anal., 34 (2014), pp. 800–819.

[5] G. A. Baker, *Finite element methods for elliptic equations using nonconforming elements*, Math. Comp., 31 (1977), pp. 45–59.

[6] A. Berger, R. Scott, and G. Strang, *Approximate boundary conditions in the finite element method*, in Symposia Mathematica, Vol. X, Academic Press, London, 1972, pp. 295–313.

[7] S. C. Brenner, *Two-level additive Schwarz preconditioners for nonconforming finite element methods*, Math. Comp., 65 (1996), pp. 897–921.

[8] S. C. Brenner, *Korn's inequalities for piecewise H^1 vector fields*, Math. Comp., 73 (2004), pp. 1067–1087.

[9] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts Appl. Math. 15, Springer, New York, 2008.

[10] S. C. Brenner and L.-Y. Sung, *Linear finite element methods for planar linear elasticity*, Math. Comp., 59 (1992), pp. 321–338.

[11] S. C. Brenner and L.-Y. Sung, *C^0 interior penalty methods for fourth order elliptic boundary value problems on polygonal domains*, J. Sci. Comput., 22/23 (2005), pp. 83–118.

[12] D. A. Di Pietro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, Math. Appl. (Berlin) 69, Springer, New York, 2012.

[13] R. S. Falk, *Nonconforming finite element methods for the equations of linear elasticity*, Math. Comp., 57 (1991), pp. 529–550.

[14] T. Gudi, *A new error analysis for discontinuous finite element methods for linear elliptic problems*, Math. Comp., 79 (2010), pp. 2169–2189.

[15] P. Hansbo and M. G. Larson, *Discontinuous Galerkin and the Crouzeix–Raviart element: Application to elasticity*, ESAIM Math. Model. Numer. Anal., 37 (2003), pp. 63–72.

[16] O. A. Karakashian and F. Pascal, *A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 2374–2399.

[17] A. Linke, *On the role of the Helmholtz decomposition in mixed methods for incompressible flows and a new variational crime*, Comput. Methods Appl. Mech. Engrg., 268 (2014), pp. 782–800.

[18] P. Oswald, *On a BPX-preconditioner for P1 elements*, Computing, 51 (1993), pp. 125–133.

[19] B. Rivière, M. F. Wheeler, and V. Girault, *A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems*, SIAM J. Numer. Anal., 39 (2001), pp. 902–931.

[20] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[21] G. Stoyan and Á. Baran, *Crouzeix–Velte decompositions for higher-order finite elements*, Comput. Math. Appl., 51 (2006), pp. 967–986.

[22] A. Veeser, *Approximating gradients with continuous piecewise polynomial functions*, Found. Comput. Math., 16 (2016), pp. 723–750.

[23] A. Veeser and P. Zanotti, *Quasi-optimal nonconforming methods for symmetric elliptic problems. I—Abstract theory*, SIAM J. Numer. Anal., 56 (2018), pp. 1621–1642.

[24] A. Veeser and P. Zanotti, *Quasi-Optimal Nonconforming Methods for Symmetric Elliptic Problems. II—Overconsistency and Classical Nonconforming Elements*, arXiv:1710.03447 [math.NA], 2017.

[25] A. Veeser and P. Zanotti, *Quasi-optimal nonconforming methods for second-order problems on domains with non-Lipschitz boundary*, in Proceedings of Numerical Mathematics and Advanced Applications - ENUMATH 2017, to appear.

[26] M. F. Wheeler, *An elliptic collocation-finite element method with interior penalties*, SIAM J. Numer. Anal., 15 (1978), pp. 152–161.