STRICTLY EQUIVALENT A POSTERIORI ESTIMATORS FOR QUASI-OPTIMAL NONCONFORMING METHODS: ABSTRACT THEORY AND SECOND-ORDER PROBLEMS

CHRISTIAN KREUZER, MATTHIAS ROTT, ANDREAS VEESER, AND PIETRO ZANOTTI

Abstract. We devise guidelines for deriving a posteriori error estimators for quasi-optimal nonconforming finite element methods and apply them in the case of symmetric elliptic problems of second order. The resulting estimators are defined for all source terms that are admissible to the underlying weak formulations. More importantly, they are equivalent to the error in a strict sense. In particular, their data oscillation part is bounded by the error and, furthermore, can be designed to be bounded by classical data oscillations. The estimators are computable, except for the data oscillation part. Since even the computation of some bound of the oscillation part is not possible in general, we advocate handling it on a case-by-case basis. We illustrate the practical use of two estimators obtained for the Crouzeix-Raviart method applied to the Poisson problem with a source term that is not a function and whose singular part with respect to the Lebesgue measure is not aligned with the mesh.

1. Introduction

Nonconforming finite element methods are a well-established technique for the approximate solution of partial differential equations (PDEs). Classical nonconforming elements, like the ones of Morley [Mor68] and Crouzeix-Raviart [CR73], were originally proposed as valuable alternatives to conforming ones for the biharmonic and Stokes equations. Later, other nonconforming techniques, such as discontinuous Galerkin and C^0 interior penalty methods, have proved to be competitive for a wide range of problems.

Nonconforming methods weaken the coupling between elements so that at least some discrete functions are not admissible to the weak formulation of the PDE of interest. On the one hand, this increases flexibility for approximation and accommodating desired structural properties, but on the other hand, it complicates the theoretical analysis and, without suitable measures, excludes data that are admissible in the weak PDE.

An a posteriori error analysis aims at devising a quantity, called estimator, that, ideally, is computable and equivalent to the error of the approximate solution. Such an estimator can be used to assess the quality of the approximate solution and, if it can be divided into local contributions, to guide adaptive mesh refinement. To outline the structure of most available results, consider a weak formulation of the form

$$u \in V \text{ such that } \forall v \in V \ a(u, v) = \langle f, v\rangle,$$

2010 *Mathematics Subject Classification.* 65N30, 65N15.

Key words and phrases. a posteriori analysis, nonconforming finite elements, error-dominated oscillation, quasi-optimality, discontinuous Galerkin.

where the bilinear form a is symmetric and V-coercive, and apply some nonconforming method based on a discrete space $V_h \not\subset V$. Then the error $\|u-u_h\|$ and the estimator est_h are related by the following equivalence, cf. [DDP95, KP03, Car05, Voh07, AR08, Bre15, CN22, CGN24]:

$$\|u-u_h\|^2 + \operatorname{osc}_h^2 \eqsim \operatorname{est}_h^2 = \operatorname{ncf}_h^2 + \eta_h^2 + \operatorname{osc}_h^2, \tag{1.1}$$

where ncf_h is an approximation of the distance of u_h to V, η_h is a PDE-specific part of the estimator depending on u_h, and osc_h is the so-called oscillation, which typically involves only data. Notice the presence of osc_h on both sides, which spoils the ideal equivalence of error and estimator. This flaw is due to the fact that in general the error cannot bound the oscillation osc_h on a given mesh, and this may even persist under adaptive refinement; cf. Remark 3.17 below. Often and related, η_h cannot be bounded by the error alone.

In contrast to (1.1), this article establishes the strict equivalence.

$$\|u-u_h\|^2 \eqsim \operatorname{est}_h^2 = \operatorname{ncf}_h^2 + \eta_h^2 + \operatorname{osc}_h^2, \tag{1.2}$$

where osc_h and often also η_h differ from their counterparts in (1.1) and are defined for all $f \in V'$. Moreover, the oscillation osc_h in (1.2) can be designed to be bounded by the one in (1.1). Observe that now the error dominates both oscillation osc_h and η_h. We achieve this improvement by generalizing the new approach of [KV21], also used in the recent survey [BCNV24], to nonconforming methods. To shed some light on the derivation, we outline our approach.

We start by identifying a residual res_h and, similarly to existing approaches, obtain

$$\|u-u_h\|^2 = \|\operatorname{res}_h\|^2_{(V+V_h)'} = \inf_{v \in V} \|u_h - v\|^2 + \|\operatorname{res}_h^{\mathsf{C}}\|^2_{V'}$$

with the conforming part $\operatorname{res}_h^{\mathsf{C}} := \operatorname{res}_{h\,|V}$ of the residual. Note that the infinite dimension of V obstructs the computation of both terms on the right-hand side. For the first term, averaging operators allow one to derive indicators ncf_h that are strictly equivalent, computable, and can easily be divided into local contributions.

The second term with the conforming residual $\operatorname{res}_h^{\mathsf{C}}$ is more delicate in view of the dual norm and the fact that the data, say only the source f, are typically taken from an infinite-dimensional space. Considering just for a moment a simplifying special case, we address these two issues by the following two steps:

$$\|\operatorname{res}_h^{\mathsf{C}}\|^2_{V'} \eqsim \sum_{z \in \mathcal{V}} \|\operatorname{res}_h^{\mathsf{C}}\|^2_{V_z'} \eqsim \sum_{z \in \mathcal{V}} \|\mathcal{P}_z \operatorname{res}_h^{\mathsf{C}}\|^2_{V_z'} + \sum_{z \in \mathcal{V}} \|f - \mathcal{P}_z f\|^2_{V_z'}, \tag{1.3}$$

where $\mathcal{V}$ are the vertices of the underlying mesh, V_z are local counterparts of V, and $\mathcal{P}_z$ are local projections onto finite-dimensional counterparts D_z of V_z'.

The general version of the first step in (1.3) is given in Lemma 3.7. It hinges on (near) orthogonality properties of the conforming residual $\operatorname{res}_h^{\mathsf{C}}$, which originate in the nonconforming discrete test functions. If the nonconforming method is quasi-optimal in the sense of

$$\|u_h - u\| \leqslant C_{\mathsf{qo}} \inf_{v_h \in V_h} \|u - v_h\|, \tag{1.4}$$

they are verified in Lemma 3.6, else they have to be measured by an additional indicator in the spirit of Remark 4.13.

The second step in (1.3) further decomposes the residual, adapting the ideas in [KV21] to new features that arise from nonconformity. Thanks to $\dim D_z < \infty$, classical techniques allow the derivation of strictly equivalent and computable indicators

η_h for the sum with the approximate residuals $\mathcal{P}_z \operatorname{res}_h^{\mathsf{C}}$, while the computability issue due to infinite-dimensional data is isolated in the second sum, which coincides with osc_h^2 in (1.2). The latter decomposition is also useful for adaptivity; see Remark 3.19 below. It is also the crucial step for the difference between the two equivalences (1.1) and (1.2). For the latter, the projections $\mathcal{P}_z$ have to be stable in a V'-like manner, while classical techniques (implicitly) use projections that are stable only in proper subspaces; see Remark 3.17.

We advocate to first apply the outlined approach and then address the computability of the oscillation osc_h on a case-by-case basis, exploiting all the available information on the structure and regularity of the given data. We postpone a more detailed discussion to the Remarks 3.15, 3.16, and 3.18 and only mention here that an a posteriori analysis based upon a 'more computable' oscillation may lead to indicators η_h that cannot be bounded by the error alone, as in the case of several classical techniques.

To illustrate the outlined approach, we apply it to various quasi-optimal nonconforming methods for second-order problems, covering arbitrary approximation orders. For further applications to fourth-order problems, which underline the generality and suitability of the approach, see [KRVZ]. The approach also allows for estimator simplifications thanks to nonconformity as in [DDP95, DDPV96]. We restrict ourselves to equations with data-free symmetric principal terms in order to avoid the technical complications which are addressed in [BCNV24, Section 4] in the case of conforming methods. Furthermore, for the sake of simplicity, we do not cover estimators based on flux equilibration; the related changes are, however, illustrated for conforming methods in [KV21, Section 4.4] (lowest-order case) and in [BCNV24, Section 4.9] (higher-order cases).

Organization. In Section 2, we briefly recall the framework for quasi-optimal nonconforming methods from [VZ18a]. Section 3 provides the framework for the a posteriori analysis of this class of methods, based on the principles in [KV21]. Section 4 illustrates the application of the abstract results to lowest- and higher-order quasi-optimal methods for the Poisson problem. Finally, we illustrate the practical use of two derived estimators in Section 5.

Notation. Let X, Y be Hilbert spaces. We equip the dual space X' with the norm $\|\cdot\|_{X'} = \sup_{x\in X, \|x\|_X \leqslant 1}\langle \cdot, x\rangle$, where $\langle\cdot,\cdot\rangle$ is the dual pairing on $X' \times X$ and $\|\cdot\|_X$ is the norm on X. We denote the norm of a bounded linear operator $\mathcal{L} : X \to Y$ by $\|\mathcal{L}\|$. Moreover, we use standard notation for the Lebesgue and Sobolev spaces.

2. Quasi-optimal nonconforming methods

In this section, we summarize the framework for quasi-optimal nonconforming methods introduced in [VZ18a]. We report only the notions and results that are useful for our subsequent developments.

Let V be a Hilbert space with scalar product $a(\cdot,\cdot)$ and induced norm $\|\cdot\| := \|\cdot\|_V := \sqrt{a(\cdot,\cdot)}$. As in [VZ18a], we focus on PDEs whose weak formulation reads

$$\text{for } f \in V', \text{ find } u \in V \text{ such that } \forall v \in V \quad a(u,v) = \langle f, v\rangle. \tag{2.1}$$

Thanks to the Riesz representation theorem, the operator $A : V \to V'$ induced by a and defined as $Av := a(v,\cdot)$ is an isometry. Hence, for any source term $f \in V'$, there exists a unique solution $u = A^{-1}f$ of (2.1) with $\|u\| = \|f\|_{V'}$.

Let V_h be a finite-dimensional linear space, which is possibly *nonconforming*, i.e. not necessarily a subspace of V. We assume that a extends to a scalar product $\widetilde{a}$ on the space $\widetilde{V} := V + V_h$. The norm induced by $\widetilde{a}$ extends the one induced by a, therefore we still denote it by $\|\cdot\|$. Consequently, the operator $\widetilde{A} : \widetilde{V} \to \widetilde{V}'$ extending A via $\widetilde{A}\widetilde{v} = \widetilde{a}(\tilde{v}, \cdot)$ is an isometry as well. Writing $\Pi_V : \widetilde{V} \to V$ for the $\widetilde{a}$-orthogonal projection onto V, the *distance to conformity* of an element $\widetilde{v} \in \widetilde{V}$ is given by the distance to its projection

$$\|\widetilde{v} - \Pi_V \widetilde{v}\| = \inf_{v \in V} \|\widetilde{v} - v\|. \tag{2.2}$$

To discretize (2.1), we need a nondegenerate bilinear form $a_h : V_h \times V_h \to \mathbb{R}$ and a functional $f_h \in V_h'$, which approximate their counterparts a and f in a sense to be specified. According to [VZ18a, Theorem 4.7], a necessary condition for the quasi-optimality (1.4) is *full stability*, i.e. the mapping $V' \ni f \mapsto f_h \in V_h'$ induced by the discretization of the source term must be bounded. This, in turn, implies that we necessarily have $f_h = \mathcal{E}_h^* f$, where $\mathcal{E}_h^* : V' \to V_h'$ is the adjoint of some linear *smoothing* operator $\mathcal{E}_h : V_h \to V$. Therefore, we restrict our attention to nonconforming discretizations of (2.1) which read as follows:

$$\text{for } f \in V', \text{find } u_h \in V_h \text{ such that } \forall v_h \in V_h \quad a_h(u_h, v_h) = \langle f, \mathcal{E}_h v_h \rangle. \tag{2.3}$$

Since a_h is nondegenerate and $\dim V_h < \infty$, the operator $A_h : V_h \to V_h'$, defined by $A_h(v_h) := a_h(v_h, \cdot)$, is invertible. Then, for any source term $f \in V'$, there exists a unique solution of the discrete problem (2.3) with $u_h = A_h^{-1}\mathcal{E}_h^* f =: M_h f$. Henceforth, we assume that the evaluations of a_h and f in the discrete problem are computationally realizable, so that u_h is computable (in exact arithmetic). This entails that the evaluation of f for every test function in $\mathcal{E}_h(V_h)$ is known. We call the linear operator $M_h : V' \to V_h$ the *nonconforming method*, i.e. the mapping of the source term f to the approximation u_h of u, whereas $P_h : V \to V_h$ defined as

$$P_h := M_h A = A_h^{-1} \mathcal{E}_h^* A \tag{2.4}$$

is the associated approximation operator, i.e. the mapping of u to u_h.

The definitions (2.1) and (2.3) of the exact and approximate solution directly lead to the *generalized Galerkin orthogonality*

$$\forall v_h \in V_h \quad a_h(u_h, v_h) = a(u, \mathcal{E}_h v_h). \tag{2.5}$$

Note the presence of the smoothing operator $\mathcal{E}_h$ and the related possible difference between the test functions in the bilinear forms a and a_h.

In view of [VZ18a, Theorem 4.14], the compatibility condition

$$\forall u \in V \cap V_h,\ v_h \in V_h \quad a_h(u, v_h) = a(u, \mathcal{E}_h v_h), \tag{2.6}$$

called *full algebraic consistency*, is necessary and sufficient for the quasi-optimality of a fully stable nonconforming method. More precisely, the following result holds.

Proposition 2.1 (Quasi-optimality)**.** *For the fully stable method M_h, induced by the discretization* (2.3)*, the following statements are equivalent:*

(1) M_h is quasi-optimal, i.e. (1.4) *is verified for some constant $C_{\mathsf{qo}} \geqslant 1$.*
(2) M_h is fully consistent, i.e. (2.6) *is verified.*
(3) There is a constant $\delta_h \geqslant 0$ verifying

$$\forall w_h \in V_h \quad \sup_{v_h \in V_h} \frac{a_h(w_h, v_h) - a(\Pi_V w_h, \mathcal{E}_h v_h)}{\|A_h^* v_h\|_{V_h'}} \leqslant \delta_h \|w_h - \Pi_V w_h\|.$$

Moreover, if (1)-(3) *are verified, then the constant* $C_{\mathtt{qo}}$ *from* (1.4) *is such that*

$$\max\{C_{\mathtt{stab}}, \delta_h\} \leqslant C_{\mathtt{qo}} \leqslant \sqrt{C_{\mathtt{stab}}^2 + \delta_h^2}$$

with

$$C_{\mathtt{stab}} := \sup_{v_h \in V_h} \frac{\|\mathcal{E}_h v_h\|}{\|A_h^* v_h\|_{V_h'}}.$$

Proof. See [VZ18a, Theorem 4.14] and, for the gap between the barriers for $C_{\mathtt{qo}}$, [VZ18a, Remark 3.5]. □

We refer to the constants δ_h and $C_{\mathtt{stab}}$ in Proposition 2.1 as the *consistency measure* and the *stability constant* of the method M_h, respectively.

Remark 2.2 (Overconsistency). Quasi-optimal methods satisfying $\delta_h = 0$ are called *overconsistent* in [VZ19], hinting at the fact that this condition does not need to be satisfied for consistent quasi-optimality (or convergence for a suitable sequence of discrete spaces V_h). Such methods are closer to conforming methods in the following sense. First, their quasi-optimality constant equals the stability constant, $C_{\mathtt{qo}} = C_{\mathtt{stab}}$. Second, the error function is $\widetilde{a}$-orthogonal to $\mathcal{E}_h(V_h)$ because the generalized Galerkin orthogonality (2.5) can be rewritten as

$$\forall v_h \in V_h \quad \widetilde{a}(u - u_h, \mathcal{E}_h v_h) = 0. \tag{2.7}$$

3. Abstract a posteriori analysis

The aim of an a posteriori analysis is to devise a quantity, called *error estimator*, with the following properties: it is computable and equivalent to the typically unknown error and can be divided into local contributions, called *indicators*. Such an estimator allows one to assess the quality of the current approximate solution, and its indicators may be used as input for adaptive mesh refinement.

This section proposes a framework for the a posteriori analysis of quasi-optimal nonconforming methods. Conceptually speaking, it proceeds in three steps. The first step determines the *residual*, along with a convenient decomposition, and relates it to the error by means of a dual norm. This complements the previous frameworks used in [KP03, Voh07] and is an alternative to [Car05]. The second step provides a *localization*, i.e. it divides the global residual norm into local contributions. Finally, the third step addresses computability, generalizing the approach [KV21] to nonconforming methods.

3.1. **The residual of nonconforming approximations.**

We start by looking for a quantity that is equivalent to the error and given in terms of available information, viz. problem data f and approximate solution u_h in the setting of Section 2.

In the special case of a conforming Galerkin discretization, i.e. for $V_h \subseteq V$, $a_h = a_{|V_h \times V_h}$, and $\mathcal{E}_h = \mathrm{id}_V$ in (2.3), the residual $f - Au_h \in V'$ satisfies

$$\|u - u_h\| = \sup_{v \in V, \|v\| \leqslant 1} a(u - u_h, v) = \|f - Au_h\|_{V'} \tag{3.1}$$

thanks to (2.1) and the fact that A is a linear isometry.

In the general case, the possibility of $u_h \notin V$ and the identity

$$\|u - u_h\| = \sup_{\tilde{v} \in \widetilde{V}, \|\tilde{v}\| \leqslant 1} \widetilde{a}(u - u_h, \tilde{v}) \tag{3.2}$$

thus suggest to suitably extend the residual $f - Au_h \in V'$ in (3.1) to a functional in $\widetilde{V}'$ allowing for $u_h \notin V$. While $\widetilde{A}u_h \in \widetilde{V}'$ is such an extension for the term $Au_h \in V'$, the extension for $f \in V'$ is less obvious. Mimicking the role of (2.1) in (3.1) for the given nonconforming situation, we observe

$$\forall \tilde{v} \in \widetilde{V} \quad \widetilde{a}(u, \tilde{v}) = \widetilde{a}(u, \Pi_V \tilde{v}) = \langle f, \Pi_V \tilde{v} \rangle. \tag{3.3}$$

Note that these conditions are meaningful for $u \in \widetilde{V}$ and the original condition $u \in V$ is encoded in the new conditions

$$\forall v^\perp \in V^\perp \;\; \widetilde{a}(u, v^\perp) = 0$$

where $V^\perp$ denotes the $\widetilde{a}$-orthogonal complement of V in $\widetilde{V}$. These observations suggest to take $\langle f, \Pi_V \cdot \rangle \in \widetilde{V}'$ as extension for $f \in V'$ and motivate the following residual notions: given the solution u_h of the discrete problem (2.3), we define its residual as

$$\operatorname{res}_h := \operatorname{res}_h(f, u_h) := \langle f, \Pi_V \cdot \rangle - \widetilde{A}u_h \in \widetilde{V}', \tag{3.4}$$

which can be decomposed into a conforming and nonconforming component by

$$\operatorname{res}_h^{\mathtt{C}} := \operatorname{res}_h^{\mathtt{C}}(f, u_h) := \operatorname{res}_{h\,|V} = f - \widetilde{A}u_h \in V', \tag{3.5a}$$

$$\operatorname{res}_h^{\mathtt{NC}} := \operatorname{res}_h^{\mathtt{NC}}(u_h) := \operatorname{res}_{h\,|V^\perp} = -\widetilde{A}u_h \in \big(V^\perp\big)'. \tag{3.5b}$$

Thus, we have $\operatorname{res}_h = \langle \operatorname{res}_h^{\mathtt{C}}, \Pi_V \cdot \rangle + \langle \operatorname{res}_h^{\mathtt{NC}}, (\mathrm{id}_{\widetilde{V}} - \Pi_V)(\cdot) \rangle$ and, upon identifying a component with its respective term in this relationship, it can be viewed as an element of $\widetilde{V}'$.

The following lemma states a well-known orthogonal decomposition of the error, which is usually formulated not completely in terms of residuals.

Lemma 3.1 (Error, residual and their decompositions)**.** *For the respective solutions u and u_h of* (2.1) *and* (2.3)*, we have*

$$\|u - u_h\|^2 = \|\operatorname{res}_h\|_{\widetilde{V}'}^2 = \|\operatorname{res}_h^{\mathtt{C}}\|_{V'}^2 + \|\operatorname{res}_h^{\mathtt{NC}}\|_{(V^\perp)'}^2,$$

$$\|u - \Pi_V u_h\| = \|\operatorname{res}_h^{\mathtt{C}}\|_{V'}, \quad \textit{and} \quad \|u_h - \Pi_V u_h\| = \|\operatorname{res}_h^{\mathtt{NC}}\|_{(V^\perp)'}.$$

Proof. Let us write $e_h := u - u_h$ for the error function and observe $\Pi_V u = u$. In light of (3.2),

$$\|u - \Pi_V u_h\| = \sup_{v \in V, \|v\| \leqslant 1} \widetilde{a}(e_h, v), \quad \text{and} \quad \|u_h - \Pi_V u_h\| = \sup_{v^\perp \in V^\perp, \|v^\perp\| \leqslant 1} \widetilde{a}(e_h, v^\perp),$$

the identities for these errors follow from (3.3) and the fact that $\widetilde{A}$ is a linear isometry. The missing identity is then a consequence of the orthogonality

$$\widetilde{a}(u - \Pi_V u_h, u_h - \Pi_V u_h) = \widetilde{a}(\Pi_V e_h, e_h - \Pi_V e_h) = 0,$$

which expresses also the orthogonality of the residual components in $\widetilde{V}'$ endowed with the scalar product $\widetilde{a}(\widetilde{A}^{-1}\cdot, \widetilde{A}^{-1}\cdot)$. □

Remark 3.2 (Extended weak formulation). An equivalent formulation of the weak formulation (2.1) in the extended space $\widetilde{V}$ reads as follows: for $f \in V'$, find $u \in \widetilde{V}$ such that

$$\forall v \in V \quad a(u, v) = \langle f, v \rangle \qquad \text{and} \qquad \forall v^\perp \in V^\perp \quad \widetilde{a}(u, v^\perp) = 0.$$

This formulation is an alternative departure point for the definition of the residual res_h and its components $\operatorname{res}_h^{\mathtt{C}}$ and $\operatorname{res}_h^{\mathtt{NC}}$.

Remark 3.3 (Decomposition features). Orthogonal decompositions are 'constant-free' and therefore desirable in an a posteriori error analysis. The orthogonal decompositions in Lemma 3.1 are characterized by the property that one component essentially coincides with the conforming subcase. This prepares the ground for applying techniques that are available for conforming methods and will be exploited in the following steps concerning localization and computability.

We shall refer to the two components $\mathrm{res}_h^{\mathtt{C}}$ and $\mathrm{res}_h^{\mathtt{NC}}$ of the residual, respectively, as *conforming* and *nonconforming residual* for short and handle them differently.

3.2. **Indicators for the nonconforming residual.** We start with the nonconforming residual. According to Lemma 3.1, its norm $\|\mathrm{res}_h^{\mathtt{NC}}\|_{(V^\perp)'}$ equals the distance to conformity $\|u_h - \Pi_V u_h\|$ of the approximate solution u_h.

In applications, the abstract space V is a space of functions on some domain Ω and the extended bilinear form $\widetilde{a}$ is defined in terms of an integral over Ω. Thus, given a partition $\mathcal{T}$ of Ω, the extended energy norm can be written as

$$\|\cdot\|^2 = \sum_{T\in\mathcal{T}} |\cdot|_T^2, \tag{3.6}$$

where the local contributions $|\cdot|_T$ are, in general, only seminorms.

The distance to conformity $\|u_h - \Pi_V u_h\|$ can then easily be divided into the local quantities $|u_h - \Pi_V u_h|_T$, $T \in \mathcal{T}$. However, note that the presence of the projection Π_V and the infinite dimension of V typically entail that these quantities are not computable and depend on u_h in a global manner. We therefore seek an equivalent alternative. For this purpose, the next proposition individuates a simple algebraic condition, which involves the *conforming subspace* $V_h \cap V$.

Proposition 3.4 (Approximating the distance to conformity). *If*

(H1) $\quad \mathcal{A}_h : V_h \to V$ *is a linear operator with* $\forall w_h \in V \cap V_h \;\; \mathcal{A}_h w_h = w_h$,

then there exists a constant $C_{\mathtt{av}} > 0$ *such that*

$$\forall v_h \in V_h \quad C_{\mathtt{av}}\|v_h - \mathcal{A}_h v_h\| \leqslant \|v_h - \Pi_V v_h\| \leqslant \|v_h - \mathcal{A}_h v_h\|.$$

It is useful to note that the upper bound in Proposition 3.4 is constant-free. The dependencies of the constant $C_{\mathtt{av}}$ in the lower bound will be determined case-by-case in the applications.

Proposition 3.4 coincides with [CN22, Lemma 2.2]. Alternatively, it also follows from the proof of [VZ18a, Theorem 3.2], and this approach provides a formula for the constant $C_{\mathtt{av}}$.

Remark 3.5 (Computable realization of the distance to conformity). An operator $\mathcal{A}_h$ verifying (H1) provides a computable realization of the distance to conformity whenever it is itself computable and each $|\cdot|_T$ can be computationally evaluated on $V_h - \mathcal{A}_h(V_h)$. Note that even if $\mathcal{A}_h$ is a local operator, the indicators are still global in nature due to their dependence on u_h. In the applications of Section 4, $\mathcal{A}_h$ will be an *averaging operator*.

3.3. **Localization of the conforming residual.** Our next goal is to divide the dual norm $\|\mathrm{res}_h^{\mathtt{C}}\|_{V'} = \sup_{v\in V, \|v\|\leqslant 1}\langle \mathrm{res}_h^{\mathtt{C}}, v\rangle$ of the conforming residual into an equivalent 'sum' of local counterparts. This step is underlying most techniques. It is complicated by the fact that the functions in V and the norm $\|\cdot\|$ typically involve derivatives. More precisely, the partition of unity with characteristic functions underlying (3.6) cannot be applied directly. Instead, as for conforming methods, we use sufficiently smooth partitions of unity to decompose the test function v into local contributions. This decomposition increases the norm $\|\cdot\|$, which in turn has to be compensated with the help of some orthogonality property of the residual.

For nonconforming methods, it is convenient to formulate such orthogonality properties with the help of the smoothing operator $\mathcal{E}_h$. Indeed, the overconsistency of Remark 2.2 provides exact orthogonality in such a manner, while, in general, quasi-optimality turns out to be a useful ingredient.

Lemma 3.6 (Near orthogonality of conforming residual)**.** *If the discretization* (2.3) *induces a quasi-optimal method M_h, then we have, for all $v \in V$ and $K > 0$,*

$$\langle \mathrm{res}_h^{\mathtt{C}}, v\rangle \leqslant \inf_{v_h\in V_h, \|v_h\|\leqslant K\|v\|} \langle \mathrm{res}_h^{\mathtt{C}}, v - \mathcal{E}_h v_h\rangle + K\delta_h \|A_h\| \|u_h - \Pi_V u_h\| \|v\|. \tag{3.7}$$

Proof. Let $v_h \in V_h$ be such that $\|v_h\| \leqslant K\|v\|$. We add and subtract $\mathcal{E}_h v_h$ in the action of the conforming residual

$$\langle \mathrm{res}_h^{\mathtt{C}}, v\rangle = \langle \mathrm{res}_h^{\mathtt{C}}, v - \mathcal{E}_h v_h\rangle + \langle \mathrm{res}_h^{\mathtt{C}}, \mathcal{E}_h v_h\rangle.$$

For the second term on the right-hand side, we use the method (2.3), the $\widetilde{a}$-orthogonal projection Π_V, and the quasi-optimality of the method through Proposition 2.1(3), to derive

$$\langle \mathrm{res}_h^{\mathtt{C}}, \mathcal{E}_h v_h\rangle = a_h(u_h, v_h) - a(\Pi_V u_h, \mathcal{E}_h v_h) \leqslant \delta_h \|A_h^* v_h\|_{V_h'} \|u_h - \Pi_V u_h\|.$$

We thus conclude by recalling that the operator norm of the adjoint A_h^* equals the operator norm of A_h. □

Concerning the partition of unity, we assume the following properties:

(H2) There are a finite index set $\mathcal{Z}$, functions $(\Phi_z)_{z\in\mathcal{Z}}$, subspaces $(V_z)_{z\in\mathcal{Z}}$ of V, a bounded linear operator $\mathcal{I}_h : V \to V_h$ and constants $C_{\mathtt{loc}}, C_{\mathtt{col}} > 0$ such that
   (i) for $z \in \mathcal{Z}$ and $v \in V$, we have $\sum_{z\in\mathcal{Z}} v\Phi_z = v$ with $v\Phi_z \in V_z$,
   (ii) for $v_z \in V_z$, we have $\|\sum_{z\in\mathcal{Z}} v_z\|^2 \leqslant C_{\mathtt{col}}^2 \sum_{z\in\mathcal{Z}} \|v_z\|^2$, and
   (iii) for $v \in V$, we have $\sum_{z\in\mathcal{Z}} \|(v - \mathcal{E}_h\mathcal{I}_h v)\Phi_z\|^2 \leqslant C_{\mathtt{loc}}^2 \|v\|^2$.

The term 'localization' in the title of this subsection hints at the fact that in applications the functions Φ_z, $z \in \mathcal{Z}$, have compact supports. We therefore refer to the subspaces V_z, $z \in \mathcal{Z}$, as *local test spaces*. Properties (ii) and (iii) prescribe some sort of stability of the partition of unity, where the form of (iii) allows countering the effect of the cut-off.

Lemma 3.7 (Localizing the conforming residual norm)**.** *If the discretization* (2.3) *induces a quasi-optimal method M_h, then* (H2) *implies that*

$$\frac{1}{C_{\mathtt{col}}^2} \sum_{z\in\mathcal{Z}} \|\mathrm{res}_h^{\mathtt{C}}\|_{V_z'}^2 \leqslant \|\mathrm{res}_h^{\mathtt{C}}\|_{V'}^2 \leqslant 2C_{\mathtt{ort}}^2 \delta_h^2 \|u_h - \Pi_V u_h\|^2 + 2C_{\mathtt{loc}}^2 \sum_{z\in\mathcal{Z}} \|\mathrm{res}_h^{\mathtt{C}}\|_{V_z'}^2$$

holds for some constant $C_{\mathtt{ort}} \leqslant \|A_h\| \|\mathcal{I}_h\|$ arising through near orthogonality properties of the residual.

Recall that the distance to conformity $\|u_h - \Pi_V u_h\|$ is localized in Section 3.2. Furthermore, for overconsistent methods, we have $\delta_h = 0$ which means that the upper bound does not mix with the nonconforming residual; in this case, the factor 2 in the constant $2C_{\mathtt{loc}}^2$ of the upper bound can be removed.

Proof. We begin by verifying the lower bound. Given $z \in \mathcal{Z}$, let $v_z \in V_z$ be the Riesz representer of $\operatorname{res}^{\mathtt{C}}_{h\,|V_z}$ in V_z implying $\|v_z\| = \|\operatorname{res}^{\mathtt{C}}_h\|_{V_z'}$ and $\langle \operatorname{res}^{\mathtt{C}}_h, v_z\rangle = \|\operatorname{res}^{\mathtt{C}}_h\|^2_{V_z'}$. Defining $v := \sum_{z\in\mathcal{Z}} v_z \in V$, we infer

$$\sum_{z\in\mathcal{Z}} \|\operatorname{res}^{\mathtt{C}}_h\|^2_{V_z'} = \sum_{z\in\mathcal{Z}} \langle \operatorname{res}^{\mathtt{C}}_h, v_z\rangle = \langle \operatorname{res}^{\mathtt{C}}_h, v\rangle \leqslant \|\operatorname{res}^{\mathtt{C}}_h\|_{V'} \|v\|.$$

Thus, by invoking (ii) in (H2), we obtain the lower bound

$$\sum_{z\in\mathcal{Z}} \|\operatorname{res}^{\mathtt{C}}_h\|^2_{V_z'} \leqslant C^2_{\mathtt{col}} \|\operatorname{res}^{\mathtt{C}}_h\|^2_{V'}.$$

To prove the upper bound with Lemma 3.6, let $v \in V$ with $\|v\| \leqslant 1$ and set $v_h := \mathcal{I}_h v \in V_h$. The assumptions (i) and (iii) in (H2) provide

$$\langle \operatorname{res}^{\mathtt{C}}_h, v - \mathcal{E}_h v_h\rangle = \sum_{z\in\mathcal{Z}} \langle \operatorname{res}^{\mathtt{C}}_h, (v - \mathcal{E}_h \mathcal{I}_h v)\Phi_z\rangle \leqslant C_{\mathtt{loc}} \left(\sum_{z\in\mathcal{Z}} \|\operatorname{res}^{\mathtt{C}}_h\|^2_{V_z'} \right)^{\frac{1}{2}}.$$

We use this estimate in Lemma 3.6 with $K = \|\mathcal{I}_h\|$, then take the supremum over v and conclude the upper bound

$$\|\operatorname{res}^{\mathtt{C}}_h\|^2_{V'} \leqslant 2\delta_h^2 \|A_h\|^2 \|\mathcal{I}_h\|^2 \|u_h - \Pi_V u_h\|^2 + 2C^2_{\mathtt{loc}} \sum_{z\in\mathcal{Z}} \|\operatorname{res}^{\mathtt{C}}_h\|^2_{V_z'}. \qquad \square$$

3.4. **Computability and conforming residual norm.** Lemma 3.7 may suggest to employ the local quantities $\|\operatorname{res}^{\mathtt{C}}_h\|_{V_z'} = \|f - \widetilde{A}u_h\|_{V_z'}$, $z \in V_z$, as indicators. However, these quantities neither are computable nor can be bounded by a computable expression if we only know that f is taken from the infinite-dimensional space V'; cf. [KV21, Lemma 2]. We therefore adapt the approach in [KV21] to the nonconforming methods in Section 2, thus isolating this computability problem in data terms.

More precisely, for each $z \in \mathcal{Z}$, using a suitable *local projection* $\mathcal{P}_z$, we decompose the local conforming residual $(f - \widetilde{A}u_h)_{|V_z}$ into two parts with the following properties: The first part $(\mathcal{P}_z \operatorname{res}^{\mathtt{C}}_h)_{|V_z}$ is a local finite-dimensional approximation of the conforming residual, while the second part $(\operatorname{res}^{\mathtt{C}}_h - \mathcal{P}_z \operatorname{res}^{\mathtt{C}}_h)_{|V_z} = (f - \mathcal{P}_z f)_{|V_z}$ depends only on data and is of an oscillatory nature. Doing so, $\|\mathcal{P}_z \operatorname{res}^{\mathtt{C}}_h\|_{V_z'}$ is equivalent to a computable quantity and, therefore, the above computability problem gets isolated in the second part and may be handled on a case-by-case basis, depending on the knowledge of f. A further advantage of this decomposition is described in Remark 3.19 below.

To construct a local projection $\mathcal{P}_z$, we discretize the local test and dual spaces V_z and V_z' by respective local finite-dimensional spaces S_z and D_z. We refer to the elements of these discrete spaces as *simple (test) functions* and *simple functionals*, respectively. The simple test functions may be *locally nonconforming*, i.e. $S_z \nsubseteq V_z$ is allowed, but have to be *globally conforming*, i.e. $S_z \subseteq V$. Similarly to [DDP95, DDPV96], this additional freedom is exploited in Theorem 4.8 below to provide an estimator, which is simpler than the one in Theorem 4.4 based on locally conforming simple test functions. In view of possibly nonconforming simple test functions, we consider the local space $\widetilde{V}_z := V_z + S_z$ and require $D_z \subseteq \widetilde{V}_z'$. Thus, the simple functionals are locally nonconforming only if there are locally nonconforming simple test functions. The operator $\mathcal{P}_z$ will then project $\widetilde{V}_z'$ onto D_z orthogonal to S_z and we shall build on the following assumption:

(H3) There are subspaces $(D_z)_{z\in\mathcal{Z}}$, $(S_z)_{z\in\mathcal{Z}}$, and constants $C_{\overline{\mathtt{com}}}, C_{\underline{\mathtt{com}}} > 0$, such that, for all $z \in \mathcal{Z}$,

(i) $S_z \subseteq V$ and $D_z \subseteq \widetilde{V}_z'$, where $\widetilde{V}_z := V_z + S_z$,

(ii) $\widetilde{A}(V_h)_{|\widetilde{V}_z} \subseteq D_z$,

(iii) $\dim S_z \leqslant \dim D_z < \infty$,

(iv) for $\chi \in D_z$, we have $\|\chi\|_{V_z'} \leqslant C_{\overline{\mathtt{com}}}\|\chi\|_{S_z'}$,

(v) for $g \in V'$, we have $\sum_{z\in\mathcal{Z}} \|g\|^2_{S_z'} \leqslant C^2_{\underline{\mathtt{com}}} \sum_{z\in\mathcal{Z}} \|g\|^2_{V_z'}$.

If we consider the case of locally conforming simple test functions, i.e. $S_z \subseteq V_z$, the conditions (i) and (ii) are equivalent to

$$\widetilde{A}(V_h)_{|V_z} \subseteq D_z \subseteq V_z' \tag{3.8}$$

and (v) holds with $C_{\underline{\mathtt{com}}} = 1$ since $\|g\|^2_{S_z'} \leqslant \|g\|^2_{V_z'}$.

In any case, the simple functionals have to cover locally the part of the residual depending on the approximate solution. This is a crucial structural condition and the remaining conditions (iii)-(v) concern the relationship between simple functionals and simple test functions in order to ensure well-posedness and stability of the aforementioned operators $\mathcal{P}_z$, $z \in \mathcal{Z}$. Condition (iv) is equivalent to the inf-sup condition

$$\inf_{\chi\in D_z} \sup_{s\in S_z} \frac{\langle \chi, s\rangle}{\|\chi\|_{V_z'}\|s\|} \geqslant \frac{1}{C_{\overline{\mathtt{com}}}} > 0 \tag{3.9}$$

and implies $\dim D_z \leqslant \dim S_z$. Consequently, a necessary condition for (H3) is $\dim S_z = \dim D_z$. In Section 4 below, we exemplify how to choose simple functionals and local test functions in concrete settings.

The next result establishes the local projections, along with their stability properties and their interplay with the extended operator $\widetilde{A}$ and the discrete space V_h.

Proposition 3.8 (Local projections)**.** *Suppose assumption* (H3) *holds. Then the variational equations*

$$\forall s \in S_z \qquad \langle \mathcal{P}_z g, s\rangle = \langle g, s\rangle, \quad z \in \mathcal{Z}, \tag{3.10}$$

define linear projections $\mathcal{P}_z : \widetilde{V}_z' \to D_z$*, satisfying the invariances*

$$\forall z \in \mathcal{Z}, v_h \in V_h \qquad \big(\mathcal{P}_z\widetilde{A}v_h\big)_{|V_z} = \big(\widetilde{A}v_h\big)_{|V_z} \tag{3.11a}$$

and the collective stability bound

$$\forall g \in V' \qquad \sum_{z\in\mathcal{Z}} \|\mathcal{P}_z g\|^2_{V_z'} \leqslant C^2_{\overline{\mathtt{com}}} C^2_{\underline{\mathtt{com}}} \sum_{z\in\mathcal{Z}} \|g\|^2_{V_z'}. \tag{3.11b}$$

Hereafter, we also write simply $\mathcal{P}_z g$ *instead of* $\mathcal{P}_z(g_{|\widetilde{V}_z'})$ *for* $g \in V'$.

Proof. Fix any $z \in \mathcal{Z}$. Thanks to the rank-nullity theorem and condition (iii) in (H3), the well-posedness of the linear variational problem for $\mathcal{P}_z$ follows from its uniqueness. To show the latter, let $\chi \in D_z$ with $\langle \chi, s\rangle = 0$ for all $s \in S_z$, that is, $\chi_{|S_z} = 0$. Condition (iv) in (H3) then implies $\chi_{|V_z} = 0$. By its linearity, χ vanishes on the sum $V_z + S_z$, meaning $\chi = 0$ in light of (i) in (H3). Consequently, (3.10) has a unique solution in D_z and $\mathcal{P}_z$ is the linear projection from $\widetilde{V}_z'$ onto D_z with the orthogonality property

$$\forall s \in S_z \qquad \langle g - \mathcal{P}_z g, s\rangle = 0. \tag{3.12}$$

The invariance (3.11a) then follows from the observation that, for $v_h \in V_h$, the difference $\mathcal{P}_z \widetilde{A} v_h - (\widetilde{A} v_h)_{|\widetilde{V}_z}$ is in the image D_z of $\mathcal{P}_z$ thanks to (ii) in (H3).

Finally, given any $g \in V'$, the collective stability bound (3.11b) readily follows by combining (iv) and (v) in (H3) and the orthogonality (3.12) of $\mathcal{P}_z$:

$$\sum_{z \in \mathcal{Z}} \|\mathcal{P}_z g\|^2_{V'_z} \leqslant C^2_{\overline{\mathtt{com}}} \sum_{z \in \mathcal{Z}} \|\mathcal{P}_z g\|^2_{S'_z} = C^2_{\overline{\mathtt{com}}} \sum_{z \in \mathcal{Z}} \|g\|^2_{S'_z} \leqslant C^2_{\overline{\mathtt{com}}} C^2_{\underline{\mathtt{com}}} \sum_{z \in \mathcal{Z}} \|g\|^2_{V'_z}. \qquad \square$$

The use of local projections is motivated by difficulties with the computability of the local quantities $\|\operatorname{res}_h^{\mathtt{C}}\|_{V'_z}$. The next remark is therefore of primary interest.

Remark 3.9 (Computing the local projections). Let $z \in \mathcal{Z}$ and $g \in V'$. From an algebraic viewpoint, the projection $\mathcal{P}_z g$ of (3.10) can be computed by solving a linear system of equations. Moreover, we have for the S'_z-norm of $\mathcal{P}_z g$ that

$$\|\mathcal{P}_z g\|^2_{S'_z} = \mathbf{g}_z^T \mathbf{A}_z^{-1} \mathbf{g}_z,$$

where $\mathbf{g}_z$ is the column vector representing the action of g on some basis of S_z and $\mathbf{A}_z$ is the matrix representing the extended scalar product $\widetilde{a}$ on $S_z \times S_z$ with respect to the same basis.

The reader may skip the following two remarks, as they are not used in what follows. They, however, give instructive information on the approximation qualities of the local projections and how they are achieved.

Remark 3.10 (Near-best approximation of the local projections). The local projections provide the collective near-best approximation

$$\sum_{z \in \mathcal{Z}} \|g - \mathcal{P}_z g\|^2_{V'_z} \leqslant (2 + 2C^2_{\overline{\mathtt{com}}} C^2_{\underline{\mathtt{com}}}) \inf_{\chi \in \widetilde{A}(V_h)} \sum_{z \in \mathcal{Z}} \|g - \chi\|^2_{V'_z} \tag{3.13}$$

for all $g \in V'$. This follows from the local invariances (3.11a) for any fixed $\chi \in \widetilde{A}(V_h)$, the linearity of the local projections, and their collective stability (3.11b).

Remark 3.11 (Local projections as nonconforming approximations). Each projection $\mathcal{P}_z$ can be viewed as a *nonconforming Petrov-Galerkin approximation* of the identity $\mathrm{id}_{V'_z}$ on V'_z. Likewise, M_h in Section 2 is such an approximation of A^{-1}. This suggests a comparison of the approach based on (H3) and the one in [VZ18a] used in Section 2. To this end, we note that (ii) of (H3) is not of interest in this context and ignore that [VZ18a] does not cover the collective and product setting with different continuous trial and test spaces for the local projections.

A common feature of both approaches is that they invoke the same type of extended spaces. In fact, the model $\widetilde{V} = V + V_h$ is repeated in the context of (H3), on the test side, by $\widetilde{V}_z = V_z + S_z$, which we may also write as $\widetilde{V}_z = V_z \oplus S_z^{\mathrm{NC}}$, where S_z^{NC} is a complement of $S_z \cap V_z$ in S_z, i.e. we have $S_z = (S_z \cap V_z) \oplus S_z^{\mathrm{NC}}$. Similarly, for the more involved trial side, we can view $\widetilde{V}'_z$ as the direct sum of V'_z and the nonconforming part $(D_{z|S_z^{\mathrm{NC}}})$ of D_z. In fact, we have $\widetilde{V}'_z = (V_z \oplus S_z^{\mathrm{NC}})' \eqsim V'_z \oplus (S_z^{\mathrm{NC}})' \eqsim V'_z \oplus (D_{z|S_z^{\mathrm{NC}}})$ and $D_z \eqsim (D_{z|V_z}) \oplus (D_{z|S_z^{\mathrm{NC}}})$.

There is, however, an important and, at first glance, perhaps striking difference. The nonconforming discrete problems (3.10) do not invoke smoothers but exploit simple restriction. Recall that classical nonconforming methods using simple restriction cannot be quasi-optimal, i.e. cannot provide near-best approximation; cf. [VZ18a]. Nevertheless, the use of restriction in (3.10) does not obstruct the collective near-best approximation (3.13) in view of the fact that all simple test functions are globally conforming. In addition, restriction is preferable to smoothing in (3.10) as it simplifies the implementation of the local projections. The global conformity of all simple test functions also implies that each extended space $\widetilde{V}_z'$, like its counterpart V_z', can be viewed as a subspace of V', meaning that the simple functionals are globally conforming. Hence, the extension of the local projections to the extended local spaces is 'built-in', while the extension of the approximation operator P_h in (2.4) has to be constructed.

The next lemma applies the local projections in order to obtain the announced decomposition of the conforming residual. Therein, an indicator η_h is *computationally C-quantifiable* whenever it is equivalent to a computable quantity $\bar{\eta}_h$ up to a multiplicative constant $C > 0$, i.e., for all $f \in V'$, we have $C^{-1}\bar{\eta}_h \leqslant \eta_h \leqslant C\bar{\eta}_h$.

Lemma 3.12 (Conforming residual decomposition)**.** *Assumption* (H3) *ensures*

$$C_{\mathtt{dcmp}}^2 \left(\eta_h^2 + \operatorname{osc}_h^2\right) \leqslant \sum_{z \in \mathcal{Z}} \| \operatorname{res}_h^{\mathtt{C}} \|_{V_z'}^2 \leqslant 2 \left(\eta_h^2 + \operatorname{osc}_h^2\right)$$

with some $C_{\mathtt{dcmp}} \geqslant (1 + 2C_{\overline{\mathtt{com}}}^2 C_{\underline{\mathtt{com}}}^2)^{-1/2}$ *and the indicators*

$$\eta_h^2 := \sum_{z \in \mathcal{Z}} \|\mathcal{P}_z \operatorname{res}_h^{\mathtt{C}} \|_{V_z'}^2 \qquad \textit{and} \qquad \operatorname{osc}_h^2 := \sum_{z \in \mathcal{Z}} \|f - \mathcal{P}_z f\|_{V_z'}^2, \tag{3.14}$$

where the indicator η_h is computationally quantifiable up to $\max\{C_{\overline{\mathtt{com}}}, C_{\underline{\mathtt{com}}}^{-1}\}$ *whenever* $\operatorname{res}_h^{\mathtt{C}} = f - \widetilde{A}u_h$ *can be computationally evaluated for all simple test functions.*

Proof. The upper bound follows from the triangle inequality and the special case $(\mathcal{P}_z \widetilde{A}u_h)_{|V_z} = \widetilde{A}u_{h|V_z}$ of the invariance (3.11a):

$$\| \operatorname{res}_h^{\mathtt{C}} \|_{V_z'} \leqslant \|\mathcal{P}_z \operatorname{res}_h^{\mathtt{C}} \|_{V_z'} + \| \operatorname{res}_h^{\mathtt{C}} - \mathcal{P}_z \operatorname{res}_h^{\mathtt{C}} \|_{V_z'} = \|\mathcal{P}_z \operatorname{res}_h^{\mathtt{C}} \|_{V_z'} + \|f - \mathcal{P}_z f\|_{V_z'}. \tag{3.15}$$

The lower bound follows from using the last equality in the opposite direction and applying the collective stability (3.11b) of the local projections,

$$\sum_{z \in \mathcal{Z}} \|\mathcal{P}_z \operatorname{res}_h^{\mathtt{C}} \|_{V_z'}^2 + \|f - \mathcal{P}_z f\|_{V_z'}^2 \leqslant (2 + 3C_{\overline{\mathtt{com}}}^2 C_{\underline{\mathtt{com}}}^2) \sum_{z \in \mathcal{Z}} \| \operatorname{res}_h^{\mathtt{C}} \|_{V_z'}^2.$$

To show the quantifiability of η_h, observe that $\bar{\eta}_h^2 := \sum_{z \in \mathcal{Z}} \|\mathcal{P}_z \operatorname{res}_h^{\mathtt{C}} \|_{S_z'}^2$ is computable thanks to Remark 3.9 and the assumed knowledge on $\operatorname{res}_h^{\mathtt{C}}$. Furthermore, it is equivalent to η_h thanks to (iv) and (v) of (H3). □

It is important to note that, as a consequence of the structural condition (ii) in (H3), the oscillation osc_h is not just of an oscillatory nature, but a *data oscillation*, i.e., in the context of (2.1), it depends solely on the data f.

Remark 3.13 (Improving $C_{\mathtt{dcmp}}$). In the analysis [BCNV24, Section 4] for a conforming method, one has the counterpart of the improved inequality

$$C_{\mathtt{dcmp}} \geqslant \frac{1}{\sqrt{2} C_{\overline{\mathtt{com}}} C_{\underline{\mathtt{com}}}}$$

thanks to [Szy06, Lemma 2.1]. For the corresponding case when all simple test functions are locally conforming, i.e. $S_z \subseteq V_z$ for all $z \in \mathcal{Z}$, this improvement follows from applying the cited lemma on each projection $\mathcal{P}_z$. Similarly, it follows also if the local projections arise from a global one, viz. there is a projection on P_h on V' such that $P_{z|V_z} = P_{h|V_z}$ for all $z \in \mathcal{Z}$. In fact, we can then apply the cited lemma on P_h, endowing V' with the square root of $\sum_{z \in \mathcal{Z}} \| \cdot \|_{V_z'}^2$.

3.5. **Deriving a strictly equivalent error estimator.** We now combine the various ingredients and previous results to obtain the main result of our abstract a posteriori analysis: a guideline to derive an estimator bounding the error from above and below without any additional terms.

Theorem 3.14 (Abstract estimator)**.** *Let M_h be a quasi-optimal method, induced by the discretization* (2.3)*. Suppose that* (H1), (H2), *and* (H3) *hold and, given tuning constants $C_1, C_2 > 0$, define an abstract estimator by*

$$\begin{aligned}
\operatorname{est}_h^2 &:= (1 + C_1^2 \delta_h^2) \operatorname{ncf}_h^2 + C_2^2(\eta_h^2 + \operatorname{osc}_h^2), \quad \textit{where} \\
\operatorname{ncf}_h^2 &:= \|u_h - \mathcal{A}_h u_h\|^2, && \textit{(approximate distance to conformity)} \\
\eta_h^2 &:= \textstyle\sum_{z\in\mathcal{Z}} \|\mathcal{P}_z f - \widetilde{A} u_h\|_{V_z'}^2, && \textit{(approximate conforming residual)} \\
\operatorname{osc}_h^2 &:= \textstyle\sum_{z\in\mathcal{Z}} \|f - \mathcal{P}_z f\|_{V_z'}^2, && \textit{(data oscillation)}
\end{aligned}$$

and the projections $\mathcal{P}_z$ are defined by (3.10)*. Then this estimator quantifies the error by*

$$\underline{C} \operatorname{est}_h \leqslant \|u - u_h\| \leqslant \overline{C} \operatorname{est}_h$$

where the equivalence constants $\underline{C}$, $\overline{C}$ depend only on the norm $\|A_h\|$ of the discrete operator A_h, the constants arising from the assumptions (H1), (H2), *and* (H3), *and the tuning constants C_1, C_2.*

Proof. We first show the upper bound. Combining Lemmas 3.1, 3.7, 3.12, and Proposition 3.4 leads to

$$\begin{aligned}
\|u - u_h\|^2 = \|\operatorname{res}_h\|_{\tilde{V}'}^2 &= \|u_h - \Pi_V u_h\|^2 + \|\operatorname{res}_h^{\mathtt{C}}\|_{V'}^2 \\
&\leqslant (1 + 2C_{\mathtt{ort}}^2 \delta_h^2) \operatorname{ncf}_h^2 + 2C_{\mathtt{loc}}^2 \sum_{z\in\mathcal{Z}} \|\operatorname{res}_h^{\mathtt{C}}\|_{V_z'}^2 \\
&\leqslant (1 + 2C_{\mathtt{ort}}^2 \delta_h^2) \operatorname{ncf}_h^2 + 4C_{\mathtt{loc}}^2 (\eta_h^2 + \operatorname{osc}_h^2).
\end{aligned}$$

This implies the upper bound for some

$$\overline{C} \leqslant \max\left\{1, \frac{\sqrt{2}C_{\mathtt{ort}}}{C_1}, \frac{2C_{\mathtt{loc}}}{C_2}\right\}.$$

Regarding the lower bound, we apply the same auxiliary results to derive

$$\begin{aligned}
\operatorname{est}_h^2 &\leqslant (1 + C_1^2\delta_h^2) C_{\mathtt{av}}^{-2} \|u_h - \Pi_V u_h\|^2 + C_2^2 C_{\mathtt{dcmp}}^{-2} C_{\mathtt{col}}^2 \|\operatorname{res}_h^{\mathtt{C}}\|_{V'}^2 \\
&\leqslant \max\left\{(1 + C_1^2\delta_h^2) C_{\mathtt{av}}^{-2}, C_2^2 C_{\mathtt{dcmp}}^{-2} C_{\mathtt{col}}^2\right\} \|u - u_h\|^2,
\end{aligned}$$

whence the lower bound holds for some

$$\underline{C} \geqslant \max\left\{(1 + C_1^2\delta_h^2) C_{\mathtt{av}}^{-2}, C_2^2 C_{\mathtt{dcmp}}^{-2} C_{\mathtt{col}}^2\right\}^{-1}. \qquad \square$$

Several remarks about the properties of the estimator est_h in Theorem 3.14 are in order. We start by discussing the computability of est_h, compare with typical results from the literature, highlight its features for adaptivity, and conclude with a remark on how to choose simple functionals.

Remark 3.15 (On computability)*.* In Remark 3.5 and Lemma 3.12, we have already observed that, under mild assumptions, the indicators ncf_h and η_h are computable (in exact arithmetic), exploiting their finite-dimensional nature. The remaining oscillation indicator osc_h cannot be computationally bounded, not to mention computed, in general, viz. knowing only that the data f is from the infinite-dimensional space V'; cf. [KV21, Lemma 2 and Corollary 5]. We may neglect it (see Remark 3.16 for a discussion of this option) or resort to some approximation of osc_h, which is derived case-by-case, possibly exploiting additional information on the regularity and/or structure of f. The significance of Theorem 3.14 then depends on the quality of these approximations. In particular, to facilitate the approximation, it may be convenient to waive the lower bound with the data oscillation osc_h and replace it by a so-called *surrogate oscillation*, i.e. a quantity that provides an upper bound of osc_h and can be considered computable; cf. [CDN12, Section 7.1].

Remark 3.16 (Neglecting data oscillation?). Let us compare the decomposition of the conforming residual in Lemma 3.12 to the following formula in numerical integration:

$$\int_0^1 v = v(\tfrac{1}{2}) + \int_0^1 v(x) - v(\tfrac{1}{2})\,dx.$$

Similarly to Lemma 3.12 combined with Remark 3.9, but with an equality instead of an equivalence, this formula relates an exact non-computable value, a computable approximation and a representation of the error of the approximation. Since in numerical integration the error is often not considered on a computational level, one may be tempted to also neglect the oscillation osc_h. This approach amounts to use in computations simply the estimator

$$\widetilde{\operatorname{est}}_h^2 := (1 + C_1^2\delta_h^2)\operatorname{ncf}_h^2 + C_2^2\eta_h^2.$$

Note that this estimator depends on the data f only through the discrete solution u_h and the local projections $\mathcal{P}_z$, $z \in \mathcal{Z}$. Therefore, it cannot see components of f orthogonal to $\mathcal{E}_h(V_h) + \bigcup_{z\in\mathcal{Z}} S_z$, which may constitute the main part of the error; cf. [BCNV24, Section 5.3.1]. A reliable use of $\widetilde{\operatorname{est}}_h$ thus requires one to verify a smallness assumption of the type $\operatorname{osc}_h \ll \widetilde{\operatorname{est}}_h$. Apart from the fact that such a smallness condition would a priori restrict the admissible data, it is not clear to us that its verification is a simpler task than the alternatives outlined in Remark 3.15. Observe also that, in an adaptive context, the estimator $\widetilde{\operatorname{est}}_h$ does not provide information that is useful to adapt to the aforementioned orthogonal components of f and the smallness assumption will be needed for the initial mesh and all subsequent ones.

Remark 3.17 (Strict equivalence and error-dominated oscillation). The strict equivalence between the error and the estimator est_h and the related error-dominated oscillation osc_h are key features of Theorem 3.14. In fact, if we derive a posteriori bounds for (2.1) and (2.3) by means of classical techniques, the equivalence between error and estimator is spoiled.

To explain this defect in more detail, denote the classical estimator and oscillation by $\widehat{\operatorname{est}}_h$ and $\widehat{\operatorname{osc}}_h$, respectively. The classical oscillation requires $f \in L^2(\Omega) \subsetneq V'$ or other regularity assumptions beyond V'. The modified equivalence then reads

$$\|u - u_h\| \leqslant C_1\big(\widehat{\operatorname{est}}_h + \widehat{\operatorname{osc}}_h\big) \quad \text{and} \quad \widehat{\operatorname{est}}_h \leqslant C_1\big(\|u - u_h\| + \widehat{\operatorname{osc}}_h\big), \tag{3.16}$$

where, depending on the estimator type, $\widehat{\operatorname{osc}}_h$ may be not present in one of the bounds. Since the classical oscillation $\widehat{\operatorname{osc}}_h$ does not vanish whenever the error does, a bound of the type $\widehat{\operatorname{osc}}_h \leqslant C\|u - u_h\|$ does not hold. Furthermore, the fact that the classical oscillation $\widehat{\operatorname{osc}}_h$ is formally of higher order does not exclude that, under adaptive refinement, it converges slower than the error in certain cases; cf. [CDN12, Section 6.4] and [KV21, Section 3.8]. Consequently, a strict equivalence cannot be obtained from (3.16) and the defect may not mitigate under adaptive refinement. In contrast to this, Theorem 3.14 ensures $\operatorname{osc}_h \leqslant C\|u - u_h\|$ for some constant C.

The crucial step for this improvement is the decomposition of the conforming residual by means of the triangle inequality in (3.15). Therein, the decomposition is realized with the local projections that are collectively stable without changing the involved norm thanks to conditions (iv) and (v) of (H3). Classical techniques decompose, explicitly or implicitly, the conforming residual in the upper and/or lower bound in a less stable manner, passing to some norm that is strictly stronger than $\|\cdot\|_{V'}$. This difference in the stability of the decomposition of the conforming residual is the new twist. In view of the discussion in [BCNV24, Remark 4.19] and [KVZ], it is also a necessary ingredient for the strict equivalence.

Remark 3.18 (Classical oscillation as surrogate). A classical oscillation $\widehat{\operatorname{osc}}_h$ from Remark 3.17 may be used as a surrogate oscillation, as defined in Remark 3.15; see Theorems 4.4, 4.8, and 4.17 below. In this context, it is worth mentioning that the application of Theorem 3.14 is nevertheless advantageous. Indeed, Theorem 3.14 ensures the lower bound $\eta_h \leqslant C\|u - u_h\|$ with some constant C, while, for certain classical techniques, the corresponding lower bound gets spoiled, i.e. one has only $\widehat{\eta}_h \leqslant C(\|u - u_h\| + \widehat{\operatorname{osc}}_h)$ for the counterpart $\widehat{\eta}_h$ of η_h, with the disadvantages outlined in Remark 3.17; cf. [BCNV24, Remark 4.43].

Remark 3.19 (Conforming residual decomposition and adaptivity). We briefly discuss a further advantage of the decomposition of the conforming residual in an adaptive context. The quantity $\sum_{z\in\mathcal{Z}} \|\operatorname{res}_h^{\mathsf{C}}\|_{V_z'}^2 = \sum_{z\in\mathcal{Z}} \|f - \widetilde{A}u_h\|_{V_z'}^2$ is of infinite-dimensional and, although localized, of global nature. The infinite dimension arises from the data f, while the global dependence is due to the presence of the approximate solution u_h. These features lead to complications in adaptivity. Roughly speaking, infinite dimension in the absence of additional information means that a given pattern for local refinement cannot be expected to uniformly reduce the error, while global dependence complicates the behavior of indicators under local refinement. After the decomposition in Lemma 3.12, the indicator η_h is of finite-dimensional but global nature and the data oscillation osc_h of local but infinite-dimensional nature. In other words: the two difficulties get separated by this decomposition. This separation can be exploited in the design and analysis of adaptive algorithms; cf., e.g., [BCNV24, Section 5].

Remark 3.20 (Choosing the simple functionals). The crucial structural condition (ii) of (H3) prescribes a minimal size for the sets D_z, $z \in \mathcal{Z}$, of simple functionals. Furthermore, in view of the difficulties in computing osc_h, it appears convenient that its formal convergence order is larger than the one for the error and that it is smaller than available classical oscillations. It will then play, at least for sufficiently smooth sources f asymptotically, a less important role. Of course, the larger the sets D_z, $z \in \mathcal{Z}$, the greater the computational cost of the estimator. All this suggests to take the simple functionals not much larger than necessary for the aforementioned three conditions.

4. Estimators for the Poisson problem

In this section, we first exemplify the application of the guidelines developed in Section 3 in a simple case and then begin to illustrate their generality. For the first purpose, we consider the Poisson problem

$$-\Delta u = f \quad \text{in } \Omega \qquad \text{and} \qquad u = 0 \quad \text{on } \partial\Omega, \tag{4.1}$$

along with a first-order Crouzeix-Raviart method, which is quasi-optimal and overconsistent. Covering the use of locally conforming and locally nonconforming simple test functions, we derive and discuss two different strictly equivalent a posteriori error estimators, highlighting various aspects of the abstract framework.

To illustrate the generality of the guidelines, we then consider quasi-optimal discontinuous Galerkin methods of arbitrary but fixed order for the second-order problem (4.1). This will be complemented by dealing with nonconforming methods for a fourth-order problem [KRVZ].

4.1. **Domain, mesh, and polynomials.** Let Ω be a polyhedral Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \geqslant 2$. We denote by n the outward pointing unit normal vector field of $\partial\Omega$ and, correspondingly, $\partial_{\mathsf{n}} v := \nabla v \cdot \mathsf{n}$ stands for the normal derivative of a sufficiently smooth function v.

Let $\mathcal{T}$ be a simplicial face-to-face (conforming) mesh of the domain Ω. We denote, respectively, the set of its faces, of its interior faces, of its vertices, and of its interior vertices by $\mathcal{F}$, $\mathcal{F}^i$, $\mathcal{V}$, and $\mathcal{V}^i$. The meshsize function $h : \Omega \to [0, \infty)$ of $\mathcal{T}$ is given by

$$h_{|\mathrm{int}(T)} := h_T \quad \text{for } T \in \mathcal{T}, \qquad \text{and} \qquad h_{|F} := h_F, \quad \text{for } F \in \mathcal{F},$$

where h_T and h_F are the respective diameters of T and F. We extend the normal field n on the skeleton $\Sigma := \bigcup_{F \in \mathcal{F}} F$ by setting $\mathsf{n}_{|F} := n_F$ for $F \in \mathcal{F}^i$, where n_F is an arbitrary but fixed unit normal vector to the interior face F. The patches around a simplex $T \in \mathcal{T}$, a face $F \in \mathcal{F}$ and a vertex $z \in \mathcal{V}$ are the respective sets

$$\omega_T := \bigcup_{T' \in \mathcal{T}, T' \cap T \neq \varnothing} T', \qquad \omega_F := \bigcup_{T' \in \mathcal{T}, F \subseteq \partial T'} T', \quad \omega_z := \bigcup_{T' \in \mathcal{T}, z \in T'} T',$$
$$\text{and} \quad \omega_z^+ := \bigcup_{T' \in \mathcal{T}, T' \cap \omega_z \neq \varnothing} T'.$$

For convenience, we also introduce notation for the diameter, the induced mesh and the set of interior faces of the above patches around a vertex, namely

$$h_z := \mathrm{diam}(\omega_z), \qquad \mathcal{T}_z := \{T \in \mathcal{T} \mid z \in T\}, \qquad \mathcal{F}_z^i := \{F \in \mathcal{F}^i \mid z \in F\},$$
$$\mathcal{T}_z^+ := \{T \in \mathcal{T} \mid T \cap \omega_z \neq \varnothing\}, \qquad \mathcal{F}_z^{i+} := \{F \in \mathcal{F}^i \mid F \cap \omega_z \neq \varnothing\}.$$

For $\ell \in \mathbb{N}_0$, we denote by $\mathbb{P}_\ell(T)$ and $\mathbb{P}_\ell(F)$ the sets of polynomials with total degree not larger than ℓ on a simplex $T \in \mathcal{T}$ and a face $F \in \mathcal{F}$, respectively. For $\ell < 0$, we use the convention $\mathbb{P}_\ell(T) := \{0\} =: \mathbb{P}_\ell(F)$. Given $k \in \mathbb{N}_0$, we consider the following sets of piecewise polynomials on $\mathcal{T}$:

$$S_\ell^k := \{v \in H^k(\Omega) \mid \forall T \in \mathcal{T} \quad v_{|T} \in \mathbb{P}_\ell(T)\} \qquad \text{and} \qquad \mathring{S}_\ell^k := S_\ell^k \cap \mathring{H}^k(\Omega),$$

with the convention $H^0(\Omega) := L^2(\Omega) =: \mathring{H}^0(\Omega)$. We make use of the Courant basis functions $\{\Psi_z\}_{z \in \mathcal{V}}$ of the first-order space S_1^1, which are determined by the condition

$$\Psi_z(y) = \delta_{zy}, \quad y, z \in \mathcal{V}, \tag{4.2}$$

and of the bubble functions

$$\Psi_T := \prod_{z \in \mathcal{V}, z \in T} \Psi_z \in \mathring{S}_{d+1}^1 \qquad \text{and} \qquad \Psi_F := \prod_{z \in \mathcal{V}, z \in F} \Psi_z \in S_d^1 \tag{4.3}$$

associated with each element $T \in \mathcal{T}$ and face $F \in \mathcal{F}$.

For a piecewise smooth function $v : \Omega \to \mathbb{R}$, the jump $[\![v]\!] : \Sigma \to \mathbb{R}$ of v is defined on the skeleton Σ of $\mathcal{T}$ as

$$[\![v]\!]_{|F} := v_{|T_1} - v_{|T_2} \quad \text{for } F \in \mathcal{F}^i, \qquad \text{and} \qquad [\![v]\!]_{|F} := v_{|T} \quad \text{for } F \in \mathcal{F} \backslash \mathcal{F}^i,$$

where, for $F \in \mathcal{F}^i$, the simplices $T_1, T_2 \in \mathcal{T}$ are such that $T_1 \cap T_2 = F$ and n_F points from T_1 to T_2, and for $F \in \mathcal{F} \backslash \mathcal{F}^i$, $T \in \mathcal{T}$ is such that $T \supset F$. Similarly, the average $\{\!\{v\}\!\} : \Sigma \to \mathbb{R}$ of v is defined as

$$\{\!\{v\}\!\}_{|F} := \frac{v_{|T_1} + v_{|T_2}}{2} \quad \text{on } F \in \mathcal{F}^i, \qquad \text{and} \qquad \{\!\{v\}\!\}_{|F} := v_{|T} \quad \text{on } F \in \mathcal{F} \backslash \mathcal{F}^i.$$

Finally, we write $a \lesssim b$ if there exists a positive constant c such that $a \leqslant c\, b$ and $a \eqsim b$ when in addition $b \lesssim a$. The hidden constants possibly depend on the space dimension d, the polynomial degree ℓ and the shape constant

$$\gamma_{\mathcal{T}} := \max_{T \in \mathcal{T}} \frac{h_T}{\rho_T}, \tag{4.4}$$

of the mesh $\mathcal{T}$, where ρ_T is the diameter of the largest ball in T.

4.2. **Crouzeix-Raviart method as model example.** This section is devoted to the a posteriori analysis of the quasi-optimal lowest-order Crouzeix-Raviart method from [VZ19, section 3.2]. Exemplifying the guidelines in Section 3, two strictly equivalent error estimators of hierarchical type are derived in Theorems 4.4 and 4.8. The first estimator involves face-bubble functions, while the second one does not, in line with the results in [DDP95, DDPV96].

We start by recalling the aforementioned Crouzeix-Raviart method. The lowest-order finite element space of Crouzeix-Raviart [CR73] on $\mathcal{T}$ is defined as

$$CR_1 := \left\{ v \in S_1^0 \mid \forall F \in \mathcal{F} \ \int_F [\![v]\!] = 0 \right\}, \tag{4.5}$$

and is not contained in $\mathring{H}^1(\Omega)$, the trial and test space of the standard weak formulation of the Poisson problem (4.1). The corresponding discrete problem reads

$$\begin{aligned} &\text{for } f \in H^{-1}(\Omega), \text{ find } u_h \in CR_1 \text{ such that} \\ &\forall v_h \in CR_1 \quad \int_\Omega \nabla_h u_h \cdot \nabla_h v_h = \langle f, \mathcal{E}_{\mathsf{CR}} v_h \rangle. \end{aligned} \tag{4.6}$$

Hereafter, ∇_h denotes the broken gradient given by

$$(\nabla_h v)_{|T} := \nabla(v_{|T}) \quad \text{for } T \in \mathcal{T} \text{ and suitable } v.$$

The *smoothing operator* $\mathcal{E}_{\mathsf{CR}} : CR_1 \to \mathring{H}^1(\Omega)$ is defined in [VZ19, Proposition 3.3] and satisfies the following properties: for all $v_h \in CR_1$, we have

$$\forall F \in \mathcal{F} \qquad \int_F \mathcal{E}_{\mathsf{CR}} v_h = \int_F v_h, \tag{4.7a}$$

$$\forall T \in \mathcal{T} \qquad h_T^{-1} \|v_h - \mathcal{E}_{\mathsf{CR}} v_h\|_{L^2(T)} + \|\nabla \mathcal{E}_{\mathsf{CR}} v_h\|_{L^2(T)} \lesssim \|\nabla_h v_h\|_{L^2(\omega_T)}, \tag{4.7b}$$

which are related to consistency and stability of the method, respectively; see also below. Note that the bound (4.7b) entails that $\mathcal{E}_{\mathsf{CR}}$ is a local operator but may enlarge the support of its argument.

In order to apply Section 3, we first note that the Poisson problem (4.1) and the method (4.6) fit into the abstract framework of Section 2 with

$$
\begin{gathered}
\widetilde{V} = \mathring{H}^1(\Omega) + CR_1 \quad \text{with} \quad V = \mathring{H}^1(\Omega), \quad V_h = CR_1, \\
\widetilde{a}(v,w) = \int_\Omega \nabla_h v \cdot \nabla_h w, \quad \|v\| = \|\nabla_h v\|_{L^2(\Omega)}, \\
a_h = \widetilde{a}_{|CR_1 \times CR_1}, \quad \text{and} \quad \mathcal{E}_h = \mathcal{E}_{\mathsf{CR}}.
\end{gathered}
\tag{4.8}
$$

The conservation (4.7a) of the face means by the smoother $\mathcal{E}_{\mathsf{CR}}$ implies

$$
\forall v_h, w_h \in CR_1 \quad \int_\Omega \nabla_h w_h \cdot \nabla_h (v_h - \mathcal{E}_{\mathsf{CR}} v_h) = 0. \tag{4.9}
$$

Combining this with the fact that both, discrete and continuous bilinear forms, are restrictions of $\widetilde{a}$ implies that the consistency measure vanishes, i.e. $\delta_h = 0$. The method is thus overconsistent; cf. Remark 2.2. Hence, Proposition 2.1 and the smoother stability (4.7b) ensure that the method is quasi-optimal with the constant $C_{\mathsf{qo}} = \|\mathcal{E}_{\mathsf{CR}}\|$ depending on the shape constant $\gamma_{\mathcal{T}}$ and the dimension d.

Next, we verify the main assumptions of Section 3 for the setting (4.8). The derivations of the two estimators will differ only in the verification of (H3) for the construction of the local projections. Let us start with (H1) about the distance to conformity. According to Proposition 3.4, we need a linear operator that is invariant on the conforming subspace $\mathring{H}^1(\Omega) \cap CR_1 = \mathring{S}^1_1$. In line with Remark 3.5, a possible choice is the averaging operator $\mathcal{A}_{\mathsf{CR}} : CR_1 \to \mathring{S}^1_1$ defined by

$$
\mathcal{A}_{\mathsf{CR}} v_h(z) := \frac{1}{\#\mathcal{T}_z} \sum_{T \in \mathcal{T}_z} v_{h|T}(z), \qquad z \in \mathcal{V}^i,\ v_h \in CR_1, \tag{4.10}
$$

see [Bre96, CGN24].

Lemma 4.1 (Approximate $\|\nabla_h \cdot\|_{L^2(\Omega)}$-distance to $\mathring{H}^1$ by averaging or jumps)**.** *In the setting* (4.8)*, the operator* $\mathcal{A}_{\mathsf{CR}}$ *in* (4.10) *satisfies assumption* (H1) *and* $C_{\mathsf{av}} \gtrsim 1$. *In particular, for all* $v_h \in CR_1$*, we have that*

$$
\begin{aligned}
\|\nabla_h(v_h - \mathcal{A}_{\mathsf{CR}} v_h)\|_{L^2(\Omega)} \lesssim \int_\Sigma \frac{|\,[\![v_h]\!]\,|^2}{h} &\lesssim \inf_{v \in \mathring{H}^1(\Omega)} \|\nabla_h(v_h - v)\|_{L^2(\Omega)} \\
&\leqslant \|\nabla_h(v_h - \mathcal{A}_{\mathsf{CR}} v_h)\|_{L^2(\Omega)}.
\end{aligned}
$$

Proof. Clearly, $\mathcal{A}_{\mathsf{CR}}$ satisfies (H1) and so the third bound of the claimed equivalence holds. For its first bound, which is equivalent to $C_{\mathsf{av}} \gtrsim 1$, we have

$$
\|\nabla_h(v_h - \mathcal{A}_{\mathsf{CR}} v_h)\|^2_{L^2(\Omega)} \lesssim \int_\Sigma \frac{|\,[\![v_h]\!]\,|^2}{h};
$$

see, e.g., [KP03, Theorem 2.2]. Finally, the second bound follows by adapting ideas from [ABC03, Theorem 10]. □

We next establish assumption (H2). Recalling that the Courant basis functions (4.2) form a partition of unity in $H^1(\Omega)$, we take

$$
\mathcal{Z} = \mathcal{V}, \qquad \Phi_z = \Psi_z, \qquad V_z = \mathring{H}^1(\omega_z), \qquad \mathcal{I}_h = \mathcal{I}_{\mathsf{CR}} \tag{4.11}
$$

where $\mathcal{I}_{\mathsf{CR}} : \mathring{H}^1(\Omega) \to CR_1$ is the Crouzeix-Raviart interpolant; see, e.g., [Bre15, Section 2.1].

Lemma 4.2 (Partition of unity in H^1 for CR)**.** *In the setting* (4.8)*, the choices* (4.11) *satisfy assumption* (H2) *with constants* $C_{\mathtt{loc}}, C_{\mathtt{col}} \lesssim 1$.

Proof. All three properties of (H2) can be verified with standard arguments, see, e.g., [BCNV24]; we address property (iii) because its proof involves the smoother $\mathcal{E}_{\mathsf{CR}}$ as novelty. Let $v \in \mathring{H}^1(\Omega)$ and $z \in \mathcal{V}$. The scaling properties $\|\Psi_z\|_{L^\infty(\Omega)} = 1$ and $\|\nabla \Psi_z\|_{L^\infty(\Omega)} \eqsim h_z^{-1}$ of the Courant basis functions and a triangle inequality imply

$$\begin{aligned}\|\nabla\big((v - \mathcal{E}_{\mathsf{CR}}\mathcal{I}_{\mathsf{CR}}v)\Psi_z\big)\|_{L^2(\Omega)} \lesssim h_z^{-1}\big(&\|v - \mathcal{I}_{\mathsf{CR}}v\|_{L^2(\omega_z)} + \|\mathcal{I}_{\mathsf{CR}}v - \mathcal{E}_{\mathsf{CR}}\mathcal{I}_{\mathsf{CR}}v\|_{L^2(\omega_z)}\big) \\ &+ \|\nabla(v - \mathcal{E}_{\mathsf{CR}}\mathcal{I}_{\mathsf{CR}}v)\|_{L^2(\omega_z)}.\end{aligned}$$

Then, the stability properties of the smoother $\mathcal{E}_{\mathsf{CR}}$, see (4.7b), and of the interpolation $\mathcal{I}_{\mathsf{CR}}$, see [Bre15, eq. (2.14)], lead to

$$\sum_{z\in\mathcal{V}} \|\nabla((v - \mathcal{E}_{\mathsf{CR}}\mathcal{I}_{\mathsf{CR}}v)\Psi_z)\|^2_{L^2(\Omega)} \lesssim \sum_{z\in\mathcal{V}} \|\nabla v\|^2_{L^2(\omega_z^+)} \lesssim \|\nabla v\|^2_{L^2(\Omega)}.$$

Here the second inequality follows from the observation that the number of enlarged stars ω_z^+ containing a given simplex is bounded in terms of the shape constant $\gamma_{\mathcal{T}}$ and d. Hence $C_{\mathtt{loc}} \lesssim 1$. □

We finally turn to assumption (H3) enabling the construction of the local projections. In order to choose the simple functionals, we follow Remark 3.20. To this end, we observe that, for a discrete trial function $v_h \in CR_1$ and test function $v \in \mathring{H}^1(\Omega)$, element-wise integration by parts provides the representation

$$\int_\Omega \nabla_h v_h \cdot \nabla v = -\sum_{T\in\mathcal{T}} \int_T \Delta v_h v + \sum_{F\in\mathcal{F}^i} \int_F [\![\nabla v_h]\!] \cdot \mathsf{n} v \tag{4.12}$$

where we do not exploit $\Delta(v_{h|T}) = 0$ for $v_h \in CR_1$ and $T \in \mathcal{T}$, in order to hint at the higher-order case. Furthermore, we note that the classical oscillation in the case at hand arises from the approximation with functionals of the form

$$\mathring{H}^1(\Omega) \ni v \mapsto \int_\Omega \bar{f} v,$$

where $\bar{f} \in L^\infty(\Omega)$ is the piecewise function given by $\bar{f}_{|T} := |T|^{-1}\int_T f$ for suitable f. For local test functions $v \in \mathring{H}^1(\omega_z)$, $z \in \mathcal{V}$, the latter type of functionals and those of the second sum on the right-hand side of (4.12) are given by constants associated with the elements of the set $\mathcal{T}_z \cup \mathcal{F}_z^i$. We thus define

$$\begin{aligned}\widehat{D}_z = \Big\{\chi \in H^{-1}(\Omega) \mid \langle\chi, v\rangle = \sum_{T\in\mathcal{T}_z} \int_T r_T v + \sum_{F\in\mathcal{F}_z^i} \int_F r_F v \\ \text{with } r_T \in \mathbb{R},\, T \in \mathcal{T}_z, \text{ and } r_F \in \mathbb{R},\, F \in \mathcal{F}_z^i\Big\},\end{aligned} \tag{4.13}$$

where the degrees of freedom are decoupled in contrast to (4.12), and take

$$D_z = \widehat{D}_{z|\mathring{H}^1(\omega_z)} \tag{4.14a}$$

as simple functionals. The space $\widehat{D}_z$ is introduced to allow for a better comparison with the announced second approach to (H3). The fact that each member of the set $\mathcal{T}_z \cup \mathcal{F}_z^i$ corresponds to a degree of freedom in the given D_z suggests to establish an analogous correspondence for the simple test functions. A possible choice is

$$S_z = \operatorname{span}\{\Psi_T \mid T \in \mathcal{T}_z\} \oplus \operatorname{span}\{\Psi_F \mid F \in \mathcal{F}_z^i\}, \tag{4.14b}$$

where Ψ_T and Ψ_F are the bubble functions from (4.3). These choices for simple functionals and test functions are essentially also used in [KV21] for the conforming method based on the space $\mathring{S}_1^1$. The only difference is that here we pair simple functionals and test functions locally instead of globally.

Lemma 4.3 (Local projections for CR with face bubbles)**.** *In the settings* (4.8) *and* (4.11)*, the choices* (4.14) *verify assumption* (H3) *with* $C_{\overline{\mathtt{com}}} \lesssim 1$ *and* $C_{\underline{\mathtt{com}}} = 1$. *Therefore,* (3.10) *defines projections* $\mathcal{P}_z : H^{-1}(\omega_z) \to D_z \subset H^{-1}(\omega_z)$ *satisfying*

$$\forall g \in H^{-1}(\Omega) \qquad \sum_{z\in\mathcal{V}} \|\mathcal{P}_z g\|^2_{H^{-1}(\omega_z)} \lesssim \sum_{z\in\mathcal{V}} \|g\|^2_{H^{-1}(\omega_z)}.$$

Proof. Let $z \in \mathcal{V}$. As $S_z \subset \mathring{H}^1(\omega_z) = V_z$, the simple test functions are locally conforming. This readily yields condition (i) of (H3) with $\widetilde{V}_z = \mathring{H}^1(\omega_z)$, condition (v) with $C_{\underline{\mathtt{com}}} = 1$, and condition (ii) in light of (4.12). Condition (iii) holds because the functionals $\mathring{H}^1(\omega_z) \ni v \mapsto \int_K v$, $K \in \mathcal{T}_z \cup \mathcal{F}_z^i$, form a basis of D_z. Condition (iv) is proved in [KV21, Theorems 8-10] with $C_{\overline{\mathtt{com}}} \lesssim 1$; see also Lemma 4.7 below for a similar argument. Finally, Proposition 3.8 provides the existence of $\mathcal{P}_z$ and the claimed stability bound. □

The approach represented by Lemmas 4.1, 4.2 and 4.3 leads to the following strictly equivalent error estimator.

Theorem 4.4 (Estimator for CR)**.** *Let* $u \in \mathring{H}^1(\Omega)$ *be the weak solution of the Poisson problem* (4.1) *and* $u_h \in CR_1$ *its quasi-optimal Crouzeix-Raviart approximation from* (4.6)*. Given tuning constants* $C_1, C_2 > 0$*, define*

$$\begin{aligned}
\operatorname{est}^2_{\mathsf{CR}} &:= \operatorname{ncf}^2_{\mathsf{CR}} + C_1^2 \eta^2_{\mathsf{CR}} + C_2^2 \operatorname{osc}^2_{\mathsf{CR}}, \ \textit{where} \\
\operatorname{ncf}^2_{\mathsf{CR}} &:= \|\nabla_h (u_h - \mathcal{A}_{\mathsf{CR}} u_h)\|^2_{L^2(\Omega)}, \\
\eta^2_{\mathsf{CR}} &:= \sum_{T\in\mathcal{T}} \eta^2_{\mathsf{CR},T} \ \textit{with} \ \eta_{\mathsf{CR},T} := \max_{K\in\mathcal{K}_T} \frac{|\langle f, \Psi_K\rangle - \int_\Omega \nabla_h u_h \cdot \nabla \Psi_K|}{\|\nabla \Psi_K\|_{L^2(\Omega)}}, \\
\operatorname{osc}^2_{\mathsf{CR}} &:= \sum_{z\in\mathcal{V}} \|f - \mathcal{P}_z f\|^2_{H^{-1}(\omega_z)},
\end{aligned}$$

with the averaging operator $\mathcal{A}_{\mathsf{CR}}$ *of* (4.10)*,* $\mathcal{K}_T = \{T\} \cup \{F \in \mathcal{F}^i \mid F \subset T\}$*, the bubble functions* Ψ_T *and* Ψ_F *of* (4.3)*, and the local projections* $\mathcal{P}_z$ *from Lemma 4.3. This estimator quantifies the error by*

$$\underline{C} \operatorname{est}_{\mathsf{CR}} \leqslant \|\nabla_h (u - u_h)\|_{L^2(\Omega)} \leqslant \overline{C} \operatorname{est}_{\mathsf{CR}} \quad \textit{and} \quad \eta_{\mathsf{CR},T} \leqslant \|\nabla_h (u - u_h)\|_{L^2(\widetilde{\omega}_T)},$$

where $\widetilde{\omega}_T = \bigcup_{F\in\mathcal{F}^i, F\subset\partial T} \omega_F$ *and the equivalence constants* $\underline{C}$ *and* $\overline{C}$ *depend only on the dimension* d*, the shape constant* $\gamma_{\mathcal{T}}$*, and the tuning constants* C_1, C_2*.*

Furthermore, if $f \in L^2(\Omega)$*, the oscillation indicator is bounded in terms of the classical* L^2*-oscillation:*

$$\operatorname{osc}^2_{\mathsf{CR}} \lesssim \sum_{T\in\mathcal{T}} h_T^2 \inf_{c\in\mathbb{R}} \|f - c\|^2_{L^2(T)}.$$

Remark 4.5 (Estimator variants for CR). For the ease of implementation, the reduction of computational cost, or with the hope to improve the equivalence constants $\overline{C}, \underline{C}$, one may consider the following variants of the indicators in Theorem 4.4:

- In view of Lemma 4.1, we may replace $\operatorname{ncf}_{\mathsf{CR}}$ with properly scaled jumps and an additional tuning constant.

- A simplification of η_{CR} is derived in Theorem 4.8 below. [KV21, Section 4] discusses other alternatives for η_{CR} using so-called local problems (this type is also used in Theorem 4.17 below) and the standard residual technique. Moreover, in light of the equivalences $\|\nabla\psi_K\|_{L^2(\Omega)} \eqsim h_K^{d-2}$, one may replace $\eta_{\mathsf{CR},T}$ by
$$\max_{K\in\mathcal{K}_T} \frac{|\langle f, \Psi_K\rangle - \int_\Omega \nabla_h u_h \cdot \nabla\Psi_K|}{h_K^{d-2}},$$
which is particularly simple for $d = 2$. This simplified hierarchical indicator is closely related to Remark 3.9, which in turn connects to local problems on S_z. In fact, also taking into account $h_K \eqsim h_z$, we observe that the matrix $\mathbf{A}_z$ in Remark 3.9 is spectrally equivalent to the diagonal matrix $\mathbf{diag}(h_z^{d-2}, \dots, h_z^{d-2})$. We therefore have
$$\|\mathcal{P}_z \operatorname{res}_h^{\mathsf{C}}\|^2_{H^{-1}(\omega_z)} \eqsim \sum_{T\in\mathcal{T}_z} \frac{|\langle \operatorname{res}_h^{\mathsf{C}}, \Psi_T\rangle|^2}{h_T^{d-2}} + \sum_{F\in\mathcal{F}_z^i} \frac{|\langle \operatorname{res}_h^{\mathsf{C}}, \Psi_F\rangle|^2}{h_F^{d-2}}.$$
and see that the 'max' in the above hierarchical indicators is mainly motivated by the 'constant-free' lower bound in Theorem 4.4.
- For alternative oscillation indicators, see the discussion in Remarks 3.15, 3.16, and 3.18.

The hierarchical variant for η_{CR} in Theorem 4.4 is best suited for a comparison with Theorem 4.8.

Proof of Theorem 4.4. Thanks to Lemmas 4.1, 4.2, and 4.3, we can apply Theorem 3.14 with (4.8) and (4.11). The overconsistency of the Crouzeix-Raviart method (4.6), i.e. $\delta_h = 0$, simplifies the abstract estimator therein and the claimed equivalence follows with the help of

$$\sum_{z\in\mathcal{V}} \|\mathcal{P}_z \operatorname{res}_h^{\mathsf{C}}\|^2_{H^{-1}(\omega_z)} \eqsim \sum_{z\in\mathcal{V}} \|\operatorname{res}_h^{\mathsf{C}}\|^2_{S_z'} \eqsim \eta_{\mathsf{CR}}^2, \tag{4.15}$$

where the first equivalence is a consequence of properties (iv) and (v) of (H3) and the second one can be shown with the arguments in [KV21, §4.1].

To verify the claimed lower bound, let $T \in \mathcal{T}$ and $K \in \{T\} \cup \{F \in \mathcal{F}^i \mid F \subset \partial T\}$ be arbitrary. Thanks to the weak formulation of (4.1) and $\operatorname{supp}\Psi_K \subset \widetilde{\omega}_T$, we readily obtain the constant-free lower bound
$$\frac{|\langle f, \Psi_K\rangle - \int_\Omega \nabla_h u_h \cdot \nabla\Psi_K|}{\|\nabla\Psi_K\|_{L^2(\Omega)}} = \frac{|\int_\Omega \nabla_h(u - u_h)\cdot\nabla\Psi_K|}{\|\nabla\Psi_K\|_{L^2(\Omega)}} \leqslant \|\nabla_h(u-u_h)\|_{L^2(\widetilde{\omega}_T)}.$$

To show the bound for the oscillation, write $\bar{f}$ for the piecewise constant function with $\bar{f}_{|T} = |T|^{-1}\int_T f$ for all $T \in \mathcal{T}$. Since $\bar{f}_{|\omega_z} \in D_z$ for any vertex $z \in \mathcal{V}$, the properties of the local projections and the Poincaré inequality imply
$$\begin{aligned}\sum_{z\in\mathcal{V}} \|f - \mathcal{P}_z f\|^2_{H^{-1}(\omega_z)} &= \sum_{z\in\mathcal{V}} \|f - \bar{f} - \mathcal{P}_z(f-\bar{f})\|^2_{H^{-1}(\omega_z)} \lesssim \sum_{z\in\mathcal{V}} \|f-\bar{f}\|^2_{H^{-1}(\omega_z)}\\ &\lesssim \sum_{z\in\mathcal{V}} h_z^2 \|f-\bar{f}\|^2_{L^2(\omega_z)} \lesssim \sum_{T\in\mathcal{T}} h_T^2 \inf_{c\in\mathbb{R}} \|f-c\|^2_{L^2(T)}\end{aligned}$$
and the proof is finished. □

Remark 4.6 (Generalization to higher-order Crouzeix-Raviart elements). [VZ19, section 3.3] derives quasi-optimal methods based upon Crouzeix-Raviart spaces of any order and dimension. Using techniques similar to those in Section 4.3 for discontinuous Galerkin methods, Theorem 4.4 can be generalized to these methods.

We next derive an error estimator that does not need the evaluations $\langle \mathrm{res}_h^{\mathsf{C}}, \Psi_F\rangle$, $F \in \mathcal{F}^i$, of the conforming residual associated with the interior faces. This simplification is inspired by [DDP95, DDPV96] and, in the framework of Section 3, is mainly achieved by suitably replacing the face-bubble functions ψ_F, $F \in \mathcal{F}_z^i$, in the choice (4.14b) of the simple test functions. To identify their alternatives, denote by $(\Psi_F^{\mathsf{CR}})_{F\in\mathcal{F}^i} \subseteq CR_1$ the Crouzeix-Raviart basis functions, defined by the condition

$$\Psi_F^{\mathsf{CR}}(m_{F'}) = \delta_{FF'}, \quad \forall F' \in \mathcal{F}^i, \tag{4.16}$$

where $m_{F'}$ is the barycenter of F'. We observe from (4.6) that

$$\forall F \in \mathcal{F}^i \qquad \langle \mathrm{res}_h^{\mathsf{C}}, \mathcal{E}_{\mathsf{CR}}\Psi_F^{\mathsf{CR}}\rangle = \langle f, \mathcal{E}_{\mathsf{CR}}\Psi_F\rangle - \int_\Omega \nabla_h u_h \cdot \nabla_h \Psi_F^{\mathsf{CR}} = 0. \tag{4.17}$$

This orthogonality suggests replacing the face-bubble functions by the smoothed Crouzeix-Raviart basis functions, which have enlarged supports; cf. (4.7b). In order to maintain the boundedness in terms of classical oscillation, these enlarged supports need to be accompanied by additional element-bubble functions; see Remark 4.9 below. Thus, the new simple test functions and functionals are

$$S_z = \mathrm{span}\{\Psi_T \mid T \in \mathcal{T}_z^+\} \oplus \mathrm{span}\{\mathcal{E}_{\mathsf{CR}}\Psi_F^{\mathsf{CR}} \mid F \in \mathcal{F}_z^i\} \tag{4.18a}$$

$$\begin{aligned} D_z = \Big\{\chi \in \big(\mathring{H}^1(\omega_z) + S_z\big)' \mid \langle \chi, v\rangle = \sum_{T\in\mathcal{T}_z^+} \int_T r_T v + \sum_{F\in\mathcal{F}_z^i} \int_F r_F v \\ \text{with } r_T \in \mathbb{R},\ T \in \mathcal{T}_z^+, \text{ and } r_F \in \mathbb{R},\ F \in \mathcal{F}_z^i\Big\}. \end{aligned} \tag{4.18b}$$

The simple test functions (4.18a) are locally nonconforming, i.e. $S_z \not\subseteq \mathring{H}^1(\omega_z)$, and the simple functionals (4.18b) essentially differ from (4.14a) only by the additional characteristic functions χ_T, $T \in \mathcal{T}_z^+ \setminus \mathcal{T}_z$.

Lemma 4.7 (Local projections for CR with smoothed CR basis)**.** *In the settings* (4.8) *and* (4.11)*, the choices* (4.18) *satisfy assumption* (H3) *with* $C_{\overline{\mathtt{com}}} \lesssim 1$ *and* $C_{\underline{\mathtt{com}}} \lesssim 1$*. Hence,* (3.10) *defines projections* $\mathcal{P}_z : \widetilde{V}_z' \to D_z \subset \widetilde{V}_z'$ *with* $\widetilde{V}_z = \mathring{H}^1(\omega_z) + S_z$ *and we have*

$$\forall g \in H^{-1}(\Omega) \quad \sum_{z\in\mathcal{V}} \|\mathcal{P}_z g\|^2_{H^{-1}(\omega_z)} \lesssim \sum_{z\in\mathcal{V}} \|g\|^2_{H^{-1}(\omega_z)}.$$

Proof. Let $z \in \mathcal{V}$. Clearly, the choices (4.18) satisfy conditions (i) and (iii) of (H3). In order to verify (ii), let $v_h \in V_h$, $v \in \mathring{H}^1(\omega_z)$ and $s \in S_z$. Then the representation (4.12) and $\Delta(v_{h|T}) = 0$ in each $T \in \mathcal{T}$ yield

$$\int_\Omega \nabla_h v_h \cdot \nabla(v+s) = \sum_{F\in\mathcal{F}_z^i} \int_F [\![\nabla v_h]\!] \cdot \mathsf{n}(v+s) + \sum_{F'\in\mathcal{F}\setminus\mathcal{F}_z^i} \int_{F'} [\![\nabla v_h]\!] \cdot \mathsf{n}s$$

and we have to show that the second sum vanishes. Let $F' \in \mathcal{F}\setminus\mathcal{F}_z^i$ be any face appearing therein and observe, for any $T \in \mathcal{T}_z^+$ and any $F \in \mathcal{F}_z^i$, that

$$\Psi_{T|F'} = 0, \quad \int_{F'} \mathcal{E}_{\mathsf{CR}}\Psi_F^{\mathsf{CR}} = \int_{F'} \Psi_F^{\mathsf{CR}} = \Psi_F^{\mathsf{CR}}(m_{F'}) = 0 \tag{4.19}$$

thanks to the moment conservation (4.7a) and the definition of the Crouzeix-Raviart basis. Therefore, $\int_{F'} [\![\nabla v_h]\!] \cdot \mathsf{n}s = 0$ and condition (ii) is verified.

We next show condition (iv) of (H3). Given any $\chi \in D_z$ and $v \in \mathring{H}^1(\omega_z)$, we can write

$$\langle \chi, v \rangle = \sum_{T \in \mathcal{T}_z^+} \int_T r_T v + \sum_{F \in \mathcal{F}_z^i} \int_F r_F v$$

for some $r_T \in \mathbb{P}_0(T)$, $T \in \mathcal{T}_z^+$ and $r_F \in \mathbb{P}_0(F)$, $F \in \mathcal{F}_z^i$. Fixing a local test function v, we choose the simple test function

$$s = \sum_{T \in \mathcal{T}_z^+} s_T \Psi_T + \sum_{F \in \mathcal{F}_z^i} s_F \mathcal{E}_{\mathsf{CR}} \Psi_F^{\mathsf{CR}} \in S_z,$$

where $s_T \in \mathbb{P}_0(T)$, $T \in \mathcal{T}_z^+$, and $s_F \in \mathbb{P}_0(F)$, $F \in \mathcal{F}_z^i$, solve the problems

$$\begin{aligned} \forall p_F \in \mathbb{P}_0(F) \qquad & \int_F s_F p_F \Psi_F^{\mathsf{CR}} = \int_F v p_F, \\ \forall p_T \in \mathbb{P}_0(T) \qquad & \int_T s_T p_T \Psi_T = \int_T v p_T - \sum_{F \in \mathcal{F}_z^i} \int_T s_F p_T \mathcal{E}_{\mathsf{CR}} \Psi_F^{\mathsf{CR}}. \end{aligned}$$

These problems are uniquely solvable because Ψ_F^{CR} and Ψ_T are strictly positive in the interior of F and T, respectively. We then have $\langle \chi, v \rangle = \langle \chi, s \rangle$ and

$$\|\nabla s\|_{L^2(\omega_z^+)} \lesssim h_z^{-1} \sum_{T \in \mathcal{T}_z^+} \|s_T\|_{L^2(T)} + h_z^{-\frac{1}{2}} \sum_{F \in \mathcal{F}_z^i} \|s_F\|_{L^2(F)} \lesssim \|\nabla v\|_{L^2(\omega_z)},$$

where the first inequality follows from (4.7b) and the scaling properties of the basis functions of S_z and, for the second one, we use in addition norm equivalences with bubble functions [Ver13, Proposition 1.4] as well as trace and Poincaré inequalities. Hence, we arrive at

$$\frac{\langle \chi, v \rangle}{\|\nabla v\|_{L^2(\omega_z)}} \lesssim \frac{\langle \chi, s \rangle}{\|\nabla s\|_{L^2(\omega_z^+)}} \leqslant \|\chi\|_{S_z'}$$

and, taking the supremum over v, we obtain the desired condition (iv).

Finally, for verifying condition (v) in (H3), assume $g \in H^{-1}(\Omega)$. Since we have $S_z \subset \mathring{H}^1(\omega_z^+)$ and so $\|g\|_{S_z'} \leqslant \|g\|_{H^{-1}(\omega_z^+)}$ for all vertices $z \in \mathcal{V}$, the desired bound follows from

$$\sum_{z \in \mathcal{V}} \|g\|^2_{H^{-1}(\omega_z^+)} \lesssim \sum_{z \in \mathcal{V}} \|g\|^2_{H^{-1}(\omega_z)}. \tag{4.20}$$

Given any vertex $z \in \mathcal{V}$ and $v \in \mathring{H}^1(\omega_z^+)$, we can write $v = \sum_{y \in \mathcal{V} \cap \omega_z^+} v \Psi_y$ and have the stability bounds $\|\nabla(v\Psi_y)\|_{L^2(\omega_y)} \lesssim \|\nabla v\|_{L^2(\omega_z^+)}$ thanks to scaling properties of the Courant basis and the Poincaré inequality with zero boundary values on ω_z^+. Hence, similarly to the proof of Lemma 4.2, we derive

$$\begin{aligned} \frac{\langle g, v \rangle}{\|\nabla v\|_{L^2(\omega_z^+)}} &\lesssim \sum_{y \in \mathcal{V} \cap \omega_z^+} \frac{\langle g, v\Psi_y \rangle}{\|\nabla(v\Psi_y)\|_{L^2(\omega_y)}} \lesssim \sum_{y \in \mathcal{V} \cap \omega_z^+} \|g\|_{H^{-1}(\omega_y)} \\ &\lesssim \left(\sum_{y \in \mathcal{V} \cap \omega_z^+} \|g\|^2_{H^{-1}(\omega_y)} \right)^{1/2}, \end{aligned} \tag{4.21}$$

where in the last inequality we used that $\#(\mathcal{V} \cap \omega_z^+)$ is bounded in terms of the shape constant $\gamma_{\mathcal{T}}$ and d. We take the supremum over v, square and sum over all vertices z. This verifies (4.20) because the number of patches ω_z^+ containing a vertex y is again bounded in terms of the shape constant $\gamma_{\mathcal{T}}$ and d. $\square$

Theorem 4.8 (Simplified estimator for CR)**.** *Let $u \in \mathring{H}^1(\Omega)$ be the weak solution of the Poisson problem (4.1) and $u_h \in CR_1$ its quasi-optimal Crouzeix-Raviart approximation from (4.6). Given tuning constants $C_1, C_2 > 0$, define*

$$
\begin{aligned}
\operatorname{est}^2_{\widetilde{\mathsf{CR}}} &:= \operatorname{ncf}^2_{\widetilde{\mathsf{CR}}} + C_1^2 \eta^2_{\widetilde{\mathsf{CR}}} + C_2^2 \operatorname{osc}^2_{\widetilde{\mathsf{CR}}}, \quad \textit{where}\\
\operatorname{ncf}^2_{\widetilde{\mathsf{CR}}} &:= \|\nabla_h (u_h - \mathcal{A}_{\mathsf{CR}} u_h)\|^2_{L^2(\Omega)},\\
\eta^2_{\widetilde{\mathsf{CR}}} &:= \sum_{T \in \mathcal{T}} \eta^2_{\widetilde{\mathsf{CR}},T} \quad \textit{with } \eta_{\widetilde{\mathsf{CR}},T} := \frac{|\langle f, \Psi_T\rangle - \int_\Omega \nabla_h u_h \cdot \nabla \Psi_T|}{\|\nabla \Psi_T\|_{L^2(\Omega)}},\\
\operatorname{osc}^2_{\widetilde{\mathsf{CR}}} &:= \sum_{z \in \mathcal{V}} \|f - \mathcal{P}_z f\|^2_{H^{-1}(\omega_z)},
\end{aligned}
$$

with the averaging operator $\mathcal{A}_{\mathsf{CR}}$ of (4.10), the element-bubble functions Ψ_T of (4.3), and the local projections $\mathcal{P}_z$ from Lemma 4.7. This estimator quantifies the error by

$$
\underline{C} \operatorname{est}_{\widetilde{\mathsf{CR}}} \leqslant \|\nabla_h (u - u_h)\|_{L^2(\Omega)} \leqslant \overline{C} \operatorname{est}_{\widetilde{\mathsf{CR}}} \quad \textit{and} \quad \eta_{\widetilde{\mathsf{CR}},T} \leqslant \|\nabla_h (u - u_h)\|_{L^2(T)}
$$

and the equivalence constants $\underline{C}$ and $\overline{C}$ depend only on the dimension d, the shape constant $\gamma_{\mathcal{T}}$, and the tuning constants C_1, C_2.

Furthermore, if $f \in L^2(\Omega)$, the oscillation indicator is bounded in terms of the classical L^2-oscillation:

$$
\operatorname{osc}^2_{\widetilde{\mathsf{CR}}} \lesssim \sum_{T \in \mathcal{T}} h_T^2 \inf_{c \in \mathbb{R}} \|f - c\|^2_{L^2(T)}.
$$

Proof. The proof is very similar to the one of Theorem 4.4. Regarding the equivalence of the estimator, the main differences are that we invoke Lemma 4.7 instead of Lemma 4.3 and that the smoothed CR basis functions do not appear in the estimator due to their orthogonality (4.17) to the residual. Also the oscillation bound follows along the same lines, with the caveat that now the condition

$$
\bar{f}_{|\omega_z^+} \in D_z \tag{4.22}
$$

ensures the invariance $(\mathcal{P}_z \bar{f})_{|\mathring{H}^1(\omega_z)} = \bar{f}_{|\omega_z}$, where we identify $\bar{f}_{|\omega_z}$ with the functional $\mathring{H}^1(\omega_z) \ni v \mapsto \int_\Omega \bar{f} v$. $\square$

To discuss Theorem 4.8, two remarks address the design and generality of the approach behind it, and three remarks compare its estimator with alternatives.

Remark 4.9 (Nonconforming element-bubble functions). One may be tempted to build the second approach on the simpler pairs

$$
S_z = \operatorname{span}\{\Psi_T \mid T \in \mathcal{T}_z\} \oplus \operatorname{span}\{\mathcal{E}_{\mathsf{CR}} \Psi_F^{\mathsf{CR}} \mid F \in \mathcal{F}_z^i\}, \quad D_z = \widehat{D}_{z|(\mathring{H}^1(\omega_z) + S_z)}, \tag{4.23}
$$

which just replace the face-bubble functions in (4.14b) by the smoothed Crouzeix-Raviart basis functions, thus avoiding the nonconforming element-bubble functions in S_z along with their companions in D_z. In this case, the estimator equivalence in Theorem 4.8 still holds, but the oscillation bound is not valid in general. Indeed, assuming (4.23), testing (3.10) first with the element-bubble functions Ψ_T, $T \in \mathcal{T}_z$, and then with the smoothed Crouzeix-Raviart basis functions $\mathcal{E}_{\mathsf{CR}} \Psi_F^{\mathsf{CR}}$, $F \in \mathcal{F}_z$, reveals that $\mathcal{P}_z \bar{f} = \bar{f}$ holds only under additional assumptions on $\bar{f}_{|(\omega_z^+ \setminus \omega_z)}$ outside of the star ω_z. This drawback is overcome by including the nonconforming element-bubble functions Ψ_T, $T \in \mathcal{T}_z^+ \setminus \mathcal{T}_z$, and their companions χ_T, $T \in \mathcal{T}_z^+ \setminus \mathcal{T}_z$, ensuring (4.22), which in turn implies $(\mathcal{P}_z \bar{f})_{|\mathring{H}^1(\omega_z)} = \bar{f}_{|\omega_z}$.

On the other hand, if we set the oscillation bound in Theorem 4.8 aside, we can use the even simpler pair

$$\begin{aligned} S_z &= \operatorname{span}\{\mathcal{E}_{\mathsf{CR}} \Psi_F^{\mathsf{CR}} \mid F \in \mathcal{F}_z^i\}, \\ D_z &= \Big\{ \chi \in \big(\mathring{H}^1(\omega_z) + S_z\big)' \mid \langle \chi, v \rangle = \sum_{F \in \mathcal{F}_z^i} \int_F r_F v \text{ with } r_F \in \mathbb{R},\, F \in \mathcal{F}_z^i \Big\}. \end{aligned}$$

and obtain another strictly equivalent estimator, which consists only of an approximate distance to conformity and a data oscillation term and which is a lower bound of the estimators in [Bre15, Section 2.3] and [CGN24, Section 3.2.3], respectively.

Remark 4.10 (Simplification for some higher-order Crouzeix-Raviart methods). For $d = 2$ and odd polynomial degree ℓ, a similarly simplified estimator as in Theorem 4.8 can be derived for the corresponding Crouzeix-Raviart method in [VZ19, Section 3.3]. In this case, [SB06, Lemma 1] shows that point evaluations in the order ℓ Legendre nodes on the edges together with the order ℓ Lagrange points in the interior of a given element determine $\mathbb{P}_\ell$. The smoothed basis functions corresponding to the Legendre nodes on the edges are then orthogonal to the residual. Thus, the higher-order counterpart of (4.14) can be modified to a higher-order counterpart of (4.18) to obtain a generalization of Theorem 4.8.

Remark 4.11 (Comparison of strictly equivalent estimators with quasi-optimality). Let us compare the two estimators in Theorems 4.4 and 4.8. The differing indicators are related as follows:

$$\eta_{\mathsf{CR}} \geqslant \eta_{\widetilde{\mathsf{CR}}}, \quad \eta_{\mathsf{CR}} \not\lesssim \eta_{\widetilde{\mathsf{CR}}}, \quad \textit{and} \quad \operatorname{osc}_{\mathsf{CR}} \eqsim \operatorname{osc}_{\widetilde{\mathsf{CR}}}. \tag{4.24}$$

These properties are proved below.

Clearly, the implementation of the indicator $\eta_{\widetilde{\mathsf{CR}}}$ is easier than that of η_{CR}, while the one of $\operatorname{osc}_{\mathsf{CR}}$ is slightly easier than that of $\operatorname{osc}_{\widetilde{\mathsf{CR}}}$ because the latter also involves some smoothing; recall that smoothing itself has to be implemented in any case for assembling the discrete problem (2.3) and that oscillations may be replaced by surrogates. In terms of computational costs, the situation is quite similar in the sense that 'easier' is replaced with 'less costly'. Thus, we may see $\operatorname{est}_{\widetilde{\mathsf{CR}}}$ as a simplification of $\operatorname{est}_{\mathsf{CR}}$.

We finally inspect the (hidden) constants in the proof. Since the derivations of the two upper bounds differ only in the second equivalence of (4.15), the constants in front of the oscillations $\operatorname{osc}_{\mathsf{CR}}$ and $\operatorname{osc}_{\widetilde{\mathsf{CR}}}$ coincide, and we expect that those in front of η_{CR} and $\eta_{\widetilde{\mathsf{CR}}}$ are different. In light of the first inequality in (4.24), the latter should be larger. For the global lower bounds, the constants for $\eta_{\widetilde{\mathsf{CR}}}$ and $\operatorname{osc}_{\widetilde{\mathsf{CR}}}$ are expected to be larger in view of the increased overlap in the simple functions, as suggested by the proofs given for (v) in (H3).

Proof of (4.24). We use the superscript '$\sim$' to distinguish the spaces and operators related to the estimator $\operatorname{est}_{\widetilde{\mathsf{CR}}}$ from those related to $\operatorname{est}_{\mathsf{CR}}$.

$\boxed{1}$ The first claimed inequality readily follows by the respective definitions. Furthermore, the invariance and stability properties of $\mathcal{P}_z$ as well as $\widetilde{D}_{z|\mathring{H}^1(\omega_z)} = D_z$ imply

$$\|f - \mathcal{P}_z f\|_{H^{-1}(\omega_z)} = \|(\mathrm{id} - \mathcal{P}_z)(f - \widetilde{\mathcal{P}}_z f)\|_{H^{-1}(\omega_z)} \lesssim \|f - \widetilde{\mathcal{P}}_z f\|_{H^{-1}(\omega_z)}.$$

Hence, summing over $z \in \mathcal{V}$, gives the oscillation bound $\mathrm{osc}_{\mathsf{CR}} \lesssim \mathrm{osc}_{\widetilde{\mathsf{CR}}}$.

$\boxed{2}$ To prepare the proof of inequality $\mathrm{osc}_{\widetilde{\mathsf{CR}}} \lesssim \mathrm{osc}_{\mathsf{CR}}$, we introduce a further local projection acting over the extended star ω_z^+: for any vertex $z \in \mathcal{V}$, the pair

$$S_z^+ := \operatorname{span}\{\Psi_T \mid T \in \mathcal{T}_z^+\} \oplus \operatorname{span}\{\Psi_F \mid F \in \mathcal{F}_z^{i+}\},$$

$$D_z^+ := \Big\{\chi \in H^{-1}(\omega_z^+) \mid \langle \chi, v\rangle = \sum_{T\in\mathcal{T}_z^+} \int_T r_T v + \sum_{F\in\mathcal{F}_z^{i+}} \int_F r_F v \\ \text{with } r_T \in \mathbb{R},\ T \in \mathcal{T}_z^+, \text{ and } r_F \in \mathbb{R},\ F \in \mathcal{F}_z^{i+}\Big\},$$

induces through Proposition 3.8 a local projection $\mathcal{P}_z^+ : H^{-1}(\omega_z^+) \to D_z^+$. For a given vertex $y \in \mathcal{V} \cap \omega_z$, combining $D^+_{z\,|\mathring{H}^1(\omega_y)} = D_y$, $S_y \subset S_z^+$, and the uniqueness of (3.10) for (4.14) yield that, for any $g \in H^{-1}(\Omega)$,

$$(\mathcal{P}_z^+ g)_{|\mathring{H}^1(\omega_y)} = \mathcal{P}_y g. \tag{4.25}$$

Also, there are constants r_T, $T \in \mathcal{T}_z^+$, and r_F, $F \in \mathcal{F}_z^i$, such that, for all $v \in \mathring{H}^1(\omega_z)$ and $s \in \widetilde{S}_z$, we have

$$\langle \mathcal{P}_z^+ g, v + s\rangle = \sum_{T\in\mathcal{T}_z^+} \int_T r_T (v+s) + \sum_{F\in\mathcal{F}_z^i} \int_F r_F (v+s)$$

thanks to (4.19). In other words, $(\mathcal{P}_z^+ g)_{|(\mathring{H}^1(\omega_z)+\widetilde{S}_z)} \in \widetilde{D}_z$ and the invariance of $\widetilde{\mathcal{P}}_z$ ensure

$$\big(\mathcal{P}_z^+ g\big)_{|(\mathring{H}^1(\omega_z)+\widetilde{S}_z)} = \widetilde{\mathcal{P}}_z(\mathcal{P}_z^+ g). \tag{4.26}$$

Note that this property hinges on the companions of the nonconforming element-bubble functions in Remark 4.9.

$\boxed{3}$ We turn to the proper proof of $\mathrm{osc}_{\widetilde{\mathsf{CR}}} \lesssim \mathrm{osc}_{\mathsf{CR}}$. Employing (4.26) restricted to $\mathring{H}^1(\omega_z)$, the stability properties of $\widetilde{\mathcal{P}}_z$, $\widetilde{S}_z \subset \mathring{H}^1(\omega_z^+)$, (4.21) and (4.25), we obtain

$$\begin{aligned}
\|f - \widetilde{\mathcal{P}}_z f\|_{H^{-1}(\omega_z)} &= \|f - \mathcal{P}_z^+ f - \widetilde{\mathcal{P}}_z(f - \mathcal{P}_z^+ f)\|_{H^{-1}(\omega_z)} \\
&\lesssim \|f - \mathcal{P}_z^+ f\|_{H^{-1}(\omega_z^+)} \lesssim \left(\sum_{y\in\mathcal{V}\cap\omega_z^+} \|f - \mathcal{P}_y f\|^2_{H^{-1}(\omega_y)}\right)^{1/2},
\end{aligned}$$

and so squaring, and summing over all vertices, yields the desired bound.

$\boxed{4}$ Finally, we verify that η_{CR} cannot be bounded by $\eta_{\widetilde{\mathsf{CR}}}$. For $f \in H^{-1}(\Omega)$ and $T \in \mathcal{T}$, an integration by parts reveals

$$\langle f, \Psi_T\rangle - \int_\Omega \nabla_h u_h \cdot \nabla \Psi_T = \langle f, \Psi_T\rangle.$$

Moreover, for $F \in \mathcal{F}^i$, the orthogonality property of the interpolant $\mathcal{I}_{\mathsf{CR}}$ in (4.11), see [Bre15, Section 2.1], and (4.6) entail that

$$\langle f, \Psi_F\rangle - \int_\Omega \nabla_h u_h \cdot \nabla \Psi_F = \langle f, \Psi_F - \mathcal{E}_{\mathsf{CR}}\mathcal{I}_{\mathsf{CR}}\Psi_F\rangle.$$

The construction of $\mathcal{E}_{\mathsf{CR}}$ in [VZ19, Proposition 3.3] reveals that there is a face F such that $(\Psi_F - \mathcal{E}_{\mathsf{CR}}\mathcal{I}_{\mathsf{CR}}\Psi_F)_{|F}$ is not the zero function. Therefore, fixing that face and taking the source f so that $\langle f, v\rangle = \int_F (\Psi_F - \mathcal{E}_{\mathsf{CR}}\mathcal{I}_{\mathsf{CR}}\Psi_F)v$ for $v \in \mathring{H}^1(\Omega)$, we have $\eta_{\mathsf{CR}} \neq 0$, whereas $\eta_{\widetilde{\mathsf{CR}}} = 0$. $\square$

Remark 4.12 (Comparison with classical estimators). We compare the estimators in Theorems 4.4 and 4.8 with the classical ones in [DDPV96], for the classical Crouzeix-Raviart method

$$\text{find } u_h \in CR_1 \text{ such that } \forall v_h \in CR_1 \quad \int_\Omega \nabla_h u_h \cdot \nabla_h v_h = \langle f, v_h \rangle. \tag{4.27}$$

For $* \in \{\mathsf{CR}, \widetilde{\mathsf{CR}}\}$, the classical estimators are given by

$$\begin{aligned}
&\operatorname{est}_*^2 := \operatorname{ncf}_{\mathrm{cl},*}^2 + \eta_{\mathrm{cl},*}^2 + \operatorname{osc}_{\mathrm{cl},*}^2, \quad \text{where}\\
&\operatorname{ncf}_{\mathrm{cl},*}^2 := \int_\Sigma \frac{|\,[\![u_h]\!]\,|^2}{h}, \qquad \operatorname{osc}_{\mathrm{cl},*}^2 := \sum_{T\in\mathcal{T}} \inf_{c\in\mathbb{R}} \int_T h^2|f-c|^2\\
&\eta_{\mathrm{cl},*}^2 := \sum_{T\in\mathcal{T}} \eta_{\mathrm{cl},*}(T)^2 \quad \text{with}\\
&\eta_{\mathrm{cl},*}(T)^2 := \begin{cases} \int_T h^2|f|^2 + \sum_{F\subseteq\partial T}\int_F h|\,[\![\nabla u_h]\!]\cdot n|^2, & \text{if } * = \mathsf{CR},\\ \int_T h^2|f|^2, & \text{if } * = \widetilde{\mathsf{CR}}. \end{cases}
\end{aligned} \tag{4.28}$$

Note that both, the classical method (4.27) and the estimators (4.28), are not defined for all $f \in H^{-1}(\Omega)$. More crucially, there are no continuous extensions of their definitions to $H^{-1}(\Omega)$. As a consequence, the classical method (4.27) is not quasi-optimal and the estimators cannot be strictly equivalent; cf. [VZ18a, Remark 4.9] and [KVZ]. It is therefore no surprise that, if $f \in L^2(\Omega)$, we have

$$\operatorname{est}_* \lesssim \operatorname{est}_{\mathrm{cl},*} \quad * \in \{\mathsf{CR}, \widetilde{\mathsf{CR}}\}. \tag{4.29}$$

In fact, the respective inequalities for the oscillations indicators are already shown in Theorems 4.4 and 4.8, while those for the indicators of the approximate $\mathring{H}^1$-distance and the conforming residual follow from Lemma 4.1 and standard residual techniques. The fact that the dependence of $\operatorname{est}_{\mathrm{cl},*}$ on f does not continuously extend to $H^{-1}(\Omega)$ entails that there is a sequence $(f_k)_k \subset L^2(\Omega)$ such that the right-hand side of (4.29) tends to ∞, while the left-hand side remains bounded or even tends to 0.

Remark 4.13 (Quasi-optimality and strict estimator equivalence). The quasi-optimality of the nonconforming method is a useful but not necessary ingredient for the strict equivalence of error estimators. In fact, the usefulness is clear from Lemma 3.6 and, to show the non-necessity, we outline the derivation of a strictly equivalent error estimator for the classical Crouzeix-Raviart method (4.27). Since the quasi-optimality enters our approach only in Lemma 3.6, we can generalize the bounds therein by accepting the additional indicator

$$\|\operatorname{res}_h^{\mathsf{C}}\|_{\mathcal{E}_h(V_h)'} := \sup_{w_h\in\mathcal{E}_h(V_h), \|w_h\|\leqslant 1} \langle \operatorname{res}_h^{\mathsf{C}}, w_h\rangle.$$

This indicator can be computed by solving a discrete problem or quantified by using the correction of a solver with strict error reduction, or by a hierarchical approach as illustrated in [FV06]. Combining this idea, e.g., with Theorem 4.4 and using the notation therein shows that the estimator

$$\left(\operatorname{ncf}_{\mathsf{CR}}^2 + \eta_{\mathsf{CR}}^2 + \|\operatorname{res}_h^{\mathsf{C}}\|_{\mathcal{E}_{\mathsf{CR}}(CR_1)'}^2 + \operatorname{osc}_{\mathsf{CR}}^2\right)^{1/2}$$

is strictly equivalent to the error $\|\nabla_h(u-u_h)\|_{L^2(\Omega)}$ of the classical Crouzeix-Raviart method (4.27).

4.3. **Discontinuous Galerkin methods of fixed arbitrary order.** In this section we apply the guidelines of Section 3 to derive strictly equivalent error estimators for quasi-optimal discontinuous Galerkin methods of fixed arbitrary order. We shall restrict ourselves to the symmetric and nonsymmetric interior penalty methods in [VZ18b, Section 3.2] for the ease of presentation; see also Remark 4.19. To recall them, we use the notation of Section 4.1.

Given a polynomial degree $\ell \in \mathbb{N}$ and a stabilization parameter σ, we define a discrete bilinear form on the possibly discontinuous piecewise polynomials S^0_ℓ by

$$\begin{aligned} a_h(u_h, v_h) = &\int_\Omega \nabla_h u_h \cdot \nabla_h v_h + \int_\Sigma \frac{\sigma}{h} [\![u_h]\!] [\![v_h]\!] \\ &+ \int_\Sigma \theta [\![u_h]\!] \{\!\{\partial_n v_h\}\!\} - \{\!\{\partial_n u_h\}\!\} [\![v_h]\!] \end{aligned} \tag{4.30}$$

with $\theta = 1$ for the nonsymmetric case and $\theta = -1$ for the symmetric case. In order to ensure coercivity of the bilinear form, the former case requires $\sigma > 0$ and the latter $\sigma > \sigma^*$ for some sufficiently large σ^* depending on the dimension d, the shape constant $\gamma_\mathcal{T}$ and the polynomial degree ℓ; compare, e.g., with [VZ18b, Lemma 3.5 and (3.23)]. The discrete problem then reads

$$\text{for } f \in H^{-1}(\Omega), \text{ find } u_h \in S^0_\ell \text{ such that } \forall v_h \in S^0_\ell \quad a_h(u_h, v_h) = \langle f, \mathcal{E}_\ell v_h\rangle, \tag{4.31}$$

where the smoothing operator $\mathcal{E}_\ell : S^0_\ell \to \mathring{H}^1(\Omega)$ is defined in [VZ18b, Proposition 3.4] and satisfies the following *moment conditions* and *stability properties*: for all $v_h \in S^0_\ell$, $F \in \mathcal{F}^i$, $m_F \in \mathbb{P}_{\ell-1}(F)$, $T \in \mathcal{T}$, $m_T \in \mathbb{P}_{\ell-2}(T)$

$$\int_F \mathcal{E}_\ell(v_h) m_F = \int_F \{\!\{v_h\}\!\} m_F, \qquad \int_T \mathcal{E}_\ell(v_h) m_T = \int_T v_h m_T, \tag{4.32a}$$

$$\begin{aligned} h_T^{-2} \|v_h - \mathcal{E}_\ell v_h\|^2_{L^2(T)} &+ \|\nabla \mathcal{E}_\ell v_h\|^2_{L^2(T)} \\ &\lesssim \|\nabla_h v_h\|^2_{L^2(\omega_T)} + \sum_{F \in \mathcal{F}, F \cap T \neq \varnothing} \int_F \frac{1}{h} |[\![v_h]\!]|^2. \end{aligned} \tag{4.32b}$$

Together with the Poisson problem (4.1), method (4.31) fits into the framework of Section 2 by considering

$$\begin{gathered} \widetilde{V} = S^0_\ell + \mathring{H}^1(\Omega) \quad \text{with} \quad V = \mathring{H}^1(\Omega), \quad V_h = S^0_\ell, \\ \widetilde{a}(v, w) = \int_\Omega \nabla_h v \cdot \nabla_h w + \int_\Sigma \frac{\sigma}{h} [\![v]\!] [\![w]\!], \quad \|v\| = \|v\|_{\mathtt{dG}} := \widetilde{a}(v, v)^{1/2}, \\ a_h \text{ as in (4.30)}, \quad \text{and} \quad \mathcal{E}_h = \mathcal{E}_\ell. \end{gathered} \tag{4.33}$$

We note that these methods are quasi-optimal but not overconsistent, i.e. we have $\delta_h \lesssim 1$ in Proposition 2.1, where the hidden constants depend only on d, $\gamma_\mathcal{T}$, ℓ and σ; see [VZ18b, Theorems 3.6 and 3.7], which also shows that $\delta_h \searrow 0$ for $\sigma \nearrow \infty$.

To establish the main assumptions of Section 3, we follow the arguments of the first approach for the Crouzeix-Raviart method in Section 4.2, focusing on the changes. Regarding a suitable operator for (H1), recall that averaging the function values at the Lagrange points of order ℓ instead of the vertices of the mesh generalizes the operator of (4.10) to

$$\mathcal{A}_\ell : S^0_\ell \to \mathring{S}^1_\ell \qquad \text{such that} \qquad \forall w_h \in \mathring{S}^1_\ell = S^0_\ell \cap \mathring{H}^1(\Omega) \quad \mathcal{A}_\ell w_h = w_h; \tag{4.34}$$

compare, e.g., with [KP03, Section 2.1]. Using [VZ18b, (3.17)], we have the following result.

Lemma 4.14 (Approximate $\|\cdot\|_{\mathtt{dG}}$-distance to $\mathring{H}^1$ by averaging and jumps)**.** *In the setting* (4.33)*, the operator* $\mathcal{A}_\ell$ *of* (4.34) *satisfies assumption* (H1) *and we have* $C_{\mathtt{av}} \gtrsim \sqrt{\sigma/(1+\sigma)}$*. In particular, for all* $v_h \in S_\ell^0$*,*

$$\sqrt{\frac{\sigma}{1+\sigma}}\|v_h - \mathcal{A}_\ell v_h\|_{\mathtt{dG}} \lesssim \int_\Sigma \frac{\sigma}{h}|\,[\![v_h]\!]\,|^2 \leqslant \inf_{v\in\mathring{H}^1(\Omega)} \|v_h - v\|_{\mathtt{dG}} \leqslant \|v_h - \mathcal{A}_\ell v_h\|_{\mathtt{dG}}.$$

Next, we turn to assumption (H2). Motivated by the inclusion $CR_1 \subset S_\ell^0$, we choose

$$\mathcal{Z} = \mathcal{V}, \qquad \Phi_z = \Psi_z, \qquad V_z = \mathring{H}^1(\omega_z), \qquad \mathcal{I}_h = \mathcal{I}_{\mathsf{CR}}, \tag{4.35}$$

which coincides with (4.11) and, in particular, does not depend on the polynomial degree ℓ. We omit the proof of the next result as it is similar to the one of Lemma 4.2.

Lemma 4.15 (Partition of unity in H^1)**.** *In the setting* (4.33)*, the choices* (4.35) *satisfy assumption* (H2) *with constants* $C_{\mathtt{loc}}, C_{\mathtt{col}} \lesssim 1$*.*

To motivate the simple test functions and functionals for assumption (H3), we follow Remark 3.20 and observe that, for a discrete trial function $v_h \in S_\ell^0$ and a test function $v \in \mathring{H}^1(\Omega)$, we have

$$\widetilde{a}(v_h, v) = -\sum_{T\in\mathcal{T}} \int_T \Delta v_h \psi + \sum_{F\in\mathcal{F}^i} \int_F [\![\nabla v_h]\!]\cdot \mathsf{n} v \tag{4.36}$$

thanks to $[\![v]\!] = 0$ on all faces $F \in \mathcal{F}$. Thus, although the extended bilinear form differs from Section 4.2, the left-hand side has the same piecewise structure as the one in (4.12), only with polynomial densities of higher order. Given any vertex $z \in \mathcal{V}$, this suggests choosing the simple functionals

$$\begin{aligned} D_z = \Big\{\chi \in H^{-1}(\omega_z) \mid \chi(v) = \sum_{T\in\mathcal{T}_z} \int_T r_T v + \sum_{F\in\mathcal{F}_z^i} \int_F r_F v \\ \text{with } r_T \in \mathbb{P}_{\ell-1}(T),\, T \in \mathcal{T}_z, \text{ and } r_F \in \mathbb{P}_{\ell-1}(F), F \in \mathcal{F}_z^i\Big\}. \end{aligned} \tag{4.37a}$$

To define the accompanying simple functions, we need to extend the polynomials on faces into the volume. To this end, given a face $F = \operatorname{conv\,hull}\{z_0 \dots, z_{d-1}\} \in \mathcal{F}$ and its midpoint z_F, we recall that the extension operator $E_F : \mathbb{P}_{\ell-1}(F) \to H^1(\omega_F)$ given by

$$(E_F v)(x) := v\left(z_F + \sum_{i=1}^{d-1} \Psi_{z_i}(x)(z_i - z_F)\right), \quad x \in \omega_F,$$

satisfies the stability bound

$$\|E_F v\|_{L^2(\omega_F)} \lesssim h_F \|v\|_{L^2(F)};$$

see, e.g., [BCNV24, Lemma 4.20]. The simple functions are then

$$\begin{aligned} S_z &= \operatorname{span}\big\{p_T\Psi_T \mid p_T \in \mathbb{P}_{\ell-1}(T), T \in \mathcal{T}_z\big\} \\ &\quad \oplus \operatorname{span}\big\{(E_F p_F)\Psi_F \mid p_F \in \mathbb{P}_{\ell-1}(F), F \in \mathcal{F}_z^i\big\}. \end{aligned} \tag{4.37b}$$

These choices coincide with those for conforming methods in [KVZ].

Lemma 4.16 (Local projections for dG)**.** *In the settings* (4.33) *and* (4.35)*, the choices* (4.37) *satisfy assumption* (H3) *with* $C_{\overline{\mathtt{com}}} \lesssim 1$ *and* $C_{\underline{\mathtt{com}}} = 1$*. Hence,* (3.10) *defines projections* $\mathcal{P}_z : H^{-1}(\omega_z) \to D_z \subset H^{-1}(\omega_z)$ *with*

$$\forall g \in H^{-1}(\Omega) \qquad \sum_{z\in\mathcal{V}} \|\mathcal{P}_z g\|^2_{H^{-1}(\omega_z)} \lesssim \sum_{z\in\mathcal{V}} \|g\|^2_{H^{-1}(\omega_z)}.$$

Proof. Given any vertex $z \in \mathcal{V}$, we have $S_z \subset V_z$, viz. the simple functions are locally conforming, and therefore (i) of (H3) holds with $\widetilde{V}_z = \mathring{H}^1(\omega_z)$, (v) with $C_{\underline{\mathtt{com}}} = 1$, and (ii) in view of (4.36). Condition (iii) holds because D_z determines S_z. Condition (iv) with $C_{\overline{\mathtt{com}}} \lesssim 1$ is shown in [KVZ] by using an argument similar to the corresponding one in Lemma 4.7. Then Proposition 3.8 implies the existence of $\mathcal{P}_z$ and the collective stability bound. □

Using the above lemmas in the abstract a posteriori analysis of Section 3 leads to the following strictly equivalent estimator.

Theorem 4.17 (Estimator for dG)**.** *Let* $u \in \mathring{H}^1(\Omega)$ *be the weak solution of the Poisson problem* (4.1) *and* $u_h \in S^0_\ell$ *be its quasi-optimal discontinuous Galerkin approximation from* (4.31)*. Given tuning constants* $C_1, C_2, C_3 > 0$*, define*

$$\begin{aligned}
\operatorname{est}^2_\ell &:= C_1 \operatorname{ncf}^2_\ell + C_2 \eta^2_\ell + C_3 \operatorname{osc}^2_\ell, \quad \textit{where} \\
\operatorname{ncf}^2_\ell &:= \|u_h - \mathcal{A}_\ell u_h\|^2_{\mathsf{dG}}, \quad \operatorname{osc}^2_\ell := \sum_{z\in\mathcal{V}} \|f - \mathcal{P}_z f\|^2_{H^{-1}(\omega_z)}, \\
\eta^2_\ell &:= \sum_{z\in\mathcal{V}} \eta^2_{\ell,z} \quad \textit{with} \quad \eta_{\ell,z} := \sup_{s\in S_z} \frac{|\langle f, s\rangle - \int_\Omega \nabla_h u_h \cdot \nabla s|}{\|\nabla s\|_{L^2(\omega_z)}}
\end{aligned}$$

with the averaging operator $\mathcal{A}_\ell$ *of* (4.34)*, the local projections* $\mathcal{P}_z$ *from Lemma 4.16, the test simple functions* S_z *from* (4.37b)*, and* $\eta_{\ell,z}$ *can be computed as indicated in Remark 3.9. This estimator quantifies the error by*

$$\underline{C}\operatorname{est}_\ell \lesssim \|u_h - u\|_{\mathsf{dG}} \lesssim \underline{C}\operatorname{est}_\ell \qquad \eta_{\ell,z} \leqslant \|\nabla_h (u_h - u)\|_{L^2(\omega_z)},$$

where the equivalence constants $\underline{C}, \overline{C}$ *depend only on the dimension* d*, the shape constant* $\gamma_{\mathcal{T}}$*, the polynomial degree* ℓ*, the stabilization parameter* σ*, and the tuning constants* C_1, C_2 *and* C_3*.*

Furthermore, if $f \in L^2(\Omega)$*, the oscillation indicator is bounded in terms of the classical higher-order* L^2*-oscillation:*

$$\operatorname{osc}^2_\ell \lesssim \sum_{T\in\mathcal{T}} h^2_T \inf_{p\in\mathbb{P}_{\ell-1}} \|f - p\|^2_{L^2(T)}.$$

Remark 4.18 (Estimator variants for dG). The estimator in Theorem 4.17 amounts to an approach based on (discrete) local problems. Indeed, the solution of

$$\text{find } v_z \in S_z \text{ such that } \forall s \in S_z \int_{\omega_z} \nabla v_z \cdot \nabla s = \langle \operatorname{res}^{\mathsf{C}}_h, s\rangle$$

satisfies the identity $\eta_{\ell,z} = \|\nabla v_z\|_{L^2(\omega_z)}$. For alternative indicators, we refer to Remark 4.5 and [BCNV24, Section 4.9].

Proof of Theorem 4.17. Supposing the settings (4.33) and (4.35), Lemmas 4.14, 4.15, and 4.16 verify (H1), (H2) and (H3). Hence, we can apply Theorem 3.14 and the claimed equivalence follows from

$$\sum_{z\in\mathcal{V}} \|\mathcal{P}_z \operatorname{res}^{\mathsf{C}}_h\|^2_{H^{-1}(\omega_z)} \eqsim \sum_{z\in\mathcal{V}} \|\operatorname{res}^{\mathsf{C}}_h\|^2_{S'_z},$$

which is a consequence of (iv) and (v) in (H3). The lower bound is an immediate consequence of the definition of $\eta_{\ell,z}$ and $\langle \mathrm{res}_h^{\mathsf{C}}, s\rangle = \int_\Omega \nabla_h(u-u_h)\cdot\nabla s$ for all $s \in S_z$. □

Remark 4.19 (Assumptions for smoothing operator). Theorem 4.17 assumes that the smoothing operator in (4.31) verifies (4.32). We emphasize that this covers high-order smoothing operators, which are only invariant on $\mathring{S}^1_1$. Although such operators are not quasi-optimal in the sense of [CN22, Definition 2.1], they lead to quasi-optimal methods [VZ18b, Theorem 3.10]. In fact, the required stability property (4.32b) for quasi-optimality and Theorem 4.17 only needs a local invariance within $\mathring{S}^1_1$, which is a proper subspace of the full conforming subspace $\mathring{S}^1_\ell$ for $\ell \geqslant 2$.

5. A numerical example with rough source term

This section presents some numerical results obtained from the discretization of the Poisson problem (4.1), on an open polygon $\Omega \subseteq \mathbb{R}^2$, with Crouzeix-Raviart finite elements. Our implementation is realized with the library ALBERTA 3.1; see [HKK$^+$, SS05].

Recall that the Crouzeix-Raviart discretization (4.27) without smoothing is defined only for sufficiently smooth source terms like, e.g., $f \in L^2(\Omega)$. This assumption, in turn, implies that the solution of the Poisson problem is in $H^2(\Omega)$, provided that Ω is convex. Hence, adaptive mesh refinement can provide a higher error decay rate than uniform refinement only in presence of re-entrant corners, as for the so-called L-shaped or slit domains.

In contrast, the quasi-optimal discretization (4.6) is defined for general sources $f \in H^{-1}(\Omega)$, entailing that $H^2(\Omega)$-regularity of the solution cannot be expected in general. Therefore, adaptive mesh refinement can asymptotically outperform uniform refinement also for convex Ω. We propose one such example and test the ability of the estimators in Theorems 4.4 and 4.8 to quantify the error and to drive adaptive mesh refinement.

Let $\Omega = (0,1)^2$ be the unit square. For $\lambda \in (0,1)$, consider the manufactured weak solution of (4.1) defined as

$$u(x_1,x_2) := \begin{cases} x_1(\lambda - x_1)x_2(1-x_2), & \text{for } x_1 \in (0,\lambda),\\ (1-x_1)(x_1-\lambda)x_2(1-x_2), & \text{for } x_1 \in (\lambda,1). \end{cases} \tag{5.1a}$$

This function is in $\mathring{H}^1(\Omega)$ and the gradient is discontinuous along the 'critical' line $\Lambda = \{(x_1,x_2)\in\Omega \mid x_1 = \lambda\}$. Thus, the corresponding source $f = -\Delta u$ in (4.1) consists of a regular and a singular component and acts on $v \in \mathring{H}^1(\Omega)$ as follows

$$\langle f, v\rangle = \int_\Omega f_{\mathrm{reg}} v + \int_\Lambda w(x_2) v(\lambda, x_2)\mathrm{d}x_2, \tag{5.1b}$$

where $w(x_2) := x_2(1-x_2)$ and $f_{\mathrm{reg}} \in L^2(\Omega)$ is given by

$$f_{\mathrm{reg}}(x_1,x_2) = \begin{cases} 2x_1(\lambda - x_1) + 2x_2(1-x_2), & \text{for } x_1 \in (0,\lambda),\\ 2(1-x_1)(x_1-\lambda) + 2x_2(1-x_2), & \text{for } x_1 \in (\lambda,1). \end{cases} \tag{5.1c}$$

To obtain the initial mesh $\mathcal{T}_0$, we subdivide Ω into four triangles by drawing the two main diagonals, cf. Figure 5.2(A). Each successive mesh $\mathcal{T}_1, \mathcal{T}_2, \dots$ is obtained from the previous one by subdividing each marked triangle into four triangles with newest vertex bisection and performing completion to preserve mesh conformity.

We set $\lambda = \frac{2}{3}$ in (5.1), to make sure that the critical line Λ is not contained in the skeleton of any mesh $\mathcal{T}_k$, $k \geqslant 0$, cf. Remark 5.1 below. According to Remarks 3.15 and 3.18, we replace the data oscillation in the estimators of Theorems 4.4 and 4.8 by the surrogate

$$\text{sur}^2_{*,k} := \sum_{T \in \mathcal{T}_k} h_T^2 \inf_{c \in \mathbb{R}} \|f_{\text{reg}} - c\|^2_{L^2(T)} + \sum_{T \in \mathcal{T}_k, T \cap \Lambda \neq \varnothing} h_T^2 \left(\max_{T \cap \Lambda} w\right)^2 \tag{5.2}$$

where $* \in \{\mathsf{CR}, \widetilde{\mathsf{CR}}\}$. The second summand is inspired by [BCNV24, Lemma 7.19]. For both estimators, we make use of the tuning constants $C_1 = 1$ and $C_2 = 0.3$.

Mesh		Error		Theorem 4.4		Theorem 4.8	
k	$\#\mathcal{T}_k$	err_k	eoc_k	$\text{est}_{\mathsf{CR},k}$	$\text{eff}_{\mathsf{CR},k}$	$\text{est}_{\widetilde{\mathsf{CR}},k}$	$\text{eff}_{\widetilde{\mathsf{CR}},k}$
0	4	6.55e-02		1.37e-01	2.09	1.33e-01	2.03
1	16	5.91e-02	0.07	1.07e-01	1.81	1.05e-01	1.78
2	64	4.10e-02	0.26	7.33e-02	1.79	7.13e-02	1.74
3	256	2.71e-02	0.30	4.89e-02	1.80	4.77e-02	1.76
4	1024	1.86e-02	0.27	3.36e-02	1.81	3.30e-02	1.77
5	4096	1.30e-02	0.26	2.35e-02	1.81	2.31e-02	1.78
6	16384	9.11e-03	0.26	1.65e-02	1.82	1.62e-02	1.78
7	65536	6.42e-03	0.25	1.17e-02	1.82	1.15e-02	1.78
8	262144	4.53e-03	0.25	8.24e-03	1.82	8.09e-03	1.79
9	1048576	3.20e-03	0.25	5.82e-03	1.82	5.72e-03	1.79

TABLE 1. Errors, experimental orders of convergence, estimators and corresponding effectivity indices for uniform mesh refinement.

We first consider the uniform mesh refinement, i.e. we obtain $\mathcal{T}_{k+1}$ from $\mathcal{T}_k$, $k \geqslant 0$, by marking each triangle for refinement. Since $u \in H^{1+s}(T)$ for all $T \in \mathcal{T}_k$ only if $s < 0.5$, due to the choice of λ above, we expect that the error err_k on $\mathcal{T}_k$, measured in the broken H^1-norm from (4.8), decays to zero as $(\#\mathcal{T}_k)^{-0.25}$. We test our expectation by computing the experimental order of convergence

$$\text{eoc}_k := \log(\text{err}_k/\text{err}_{k-1})/\log(\#\mathcal{T}_{k-1}/\#\mathcal{T}_k), \quad k \geqslant 1.$$

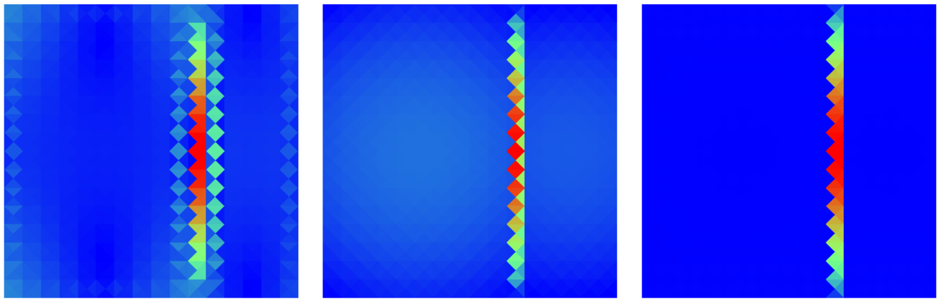
We compute also the effectivity index

$$\text{eff}_{*,k} := \text{est}_{*,k}/\text{err}_k, \quad k \geqslant 0$$

to assess the quality of the estimator $\text{est}_{*,k}$, with $* \in \{\mathsf{CR}, \widetilde{\mathsf{CR}}\}$, computed on $\mathcal{T}_k$. The data in Table 1 suggest that, for $k \to \infty$, we have $\text{eoc}_k \to 0.25$ as expected, the two estimators yield quite similar results and their effectivity indices are bounded between 1 and 2.

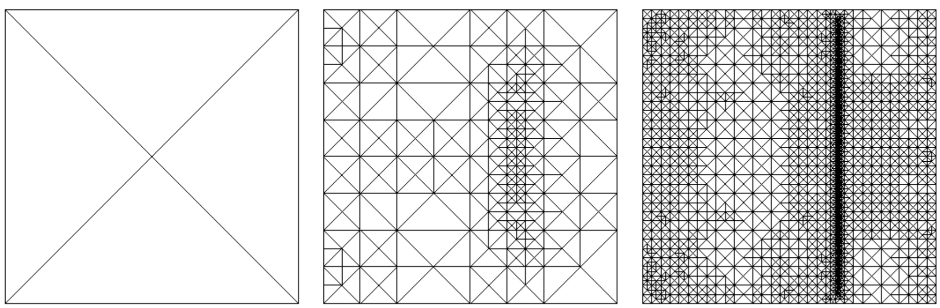
The decay rate $(\#\mathcal{T}_k)^{-0.25}$ of err_k above hinges on the discontinuity of ∇u along the critical line Λ. Therefore, the error is expected to concentrate along Λ. Remarkably, this property is captured by each indicator in the estimators, as illustrated in Figure 5.1 for $\text{est}_{\widetilde{\mathsf{CR}},k}$. In particular, all indicators decay to zero at the same rate and the surrogate oscillation is not of higher order compared to the error.

The above results suggest that adaptive mesh refinement could asymptotically outperform uniform refinement for this example. To check this, we restrict ourselves to the use of the simplified estimator $\text{est}_{\widetilde{\mathsf{CR}},k}$ as input for Dörfler marking with bulk



(A) $\mathrm{ncf}_{\widetilde{\mathsf{CR}},4}$ (B) $\eta_{\widetilde{\mathsf{CR}},4}$ (C) $\mathrm{sur}_{\widetilde{\mathsf{CR}},4}$

FIGURE 5.1. Distribution in space of the indicators in each component of the estimator $\mathrm{est}_{\widetilde{\mathsf{CR}},4}$, with oscillation replaced by surrogate, on the mesh $\mathcal{T}_4$ obtained by uniform refinement. Lower and higher values correspond to cold and warm colors, respectively.



(A) $\mathcal{T}_0$ (B) $\mathcal{T}_5$ (C) $\mathcal{T}_{10}$

FIGURE 5.2. Initial mesh $\mathcal{T}_0$ and subsequent meshes $\mathcal{T}_5$ and $\mathcal{T}_{10}$ generated by adaptive refinement.

parameter $\theta = 0.7$. As expected from Figure 5.1, such meshes are highly graded along the line Λ and relatively coarse elsewhere, see Figure 5.2. Moreover, owing to Figure 5.3, both the error and the estimator decay to zero as $(\#\mathcal{T}_k)^{-0.5}$, which is the best possible decay rate for a first-order method.

Remark 5.1 (Critical line). The position of the line Λ is crucial in this example. Indeed, if Λ was contained in the skeleton of, e.g., the initial mesh, then we would have $u \in H^2(T)$ for all $T \in \mathcal{T}_k$, $k \geqslant 0$, and we would observe the error decay rate $(\#\mathcal{T}_k)^{-0.5}$ also with uniform refinement. Moreover, we should modify the surrogate oscillation, because the simple functionals in (4.18b) would approximate the singular component of the source term on Λ more accurately than predicted by the last term in (5.2).

Funding. Christian Kreuzer acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 321270008. Pietro Zanotti was supported by the GNCS-INdAM project CUP E53C23001670001.

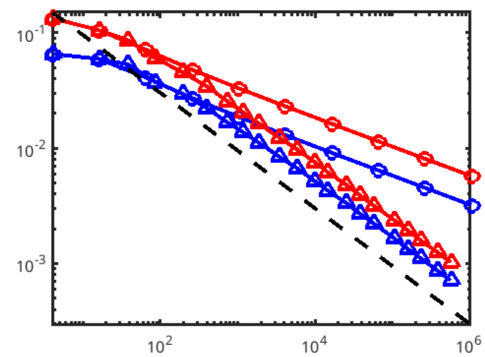


FIGURE 5.3. Error err_k (blue) and estimator $\mathrm{est}_{\widetilde{\mathsf{CR}},k}$ (red) for uniform (◯) and adaptive (△) mesh refinement versus $\#\mathcal{T}_k$. The dashed line indicates the decay rate $(\#\mathcal{T}_k)^{-0.5}$.

References

[ABC03] Y. Achdou, C. Bernardi, and F. Coquel. A priori and a posteriori analysis of finite volume discretizations of Darcy's equations. *Numer. Math.*, 96(1):17–42, 2003.

[AR08] M. Ainsworth and R. Rankin. Fully Computable Bounds for the Error in Nonconforming Finite Element Approximations of Arbitrary Order on Triangular Elements. *SIAM Journal on Numerical Analysis*, 46(6):3207–3232, January 2008. Publisher: Society for Industrial and Applied Mathematics.

[BCNV24] A. Bonito, C. Canuto, R. H. Nochetto, and A. Veeser. Adaptive finite element methods. *Acta Numer.*, 33:163–485, 2024.

[Bre96] S. Brenner. Two-level additive Schwarz preconditioners for nonconforming finite element methods. *Mathematics of Computation*, 65(215):897–921, 1996.

[Bre15] S. C. Brenner. Forty years of the Crouzeix-Raviart element. *Numer. Methods Partial Differential Equations*, 31(2):367–396, 2015.

[Car05] C. Carstensen. A unifying theory of a posteriori finite element error control. *Numer. Math.*, 100(4):617–637, 2005.

[CDN12] A. Cohen, R. DeVore, and R. H. Nochetto. Convergence rates of AFEM with H^{-1} data. *Found. Comput. Math.*, 12(5):671–718, 2012.

[CGN24] C. Carstensen, B. Gräßle, and N. Nataraj. Unifying a posteriori error analysis of five piecewise quadratic discretisations for the biharmonic equation. *J. Numer. Math.*, 32(1):77–109, 2024.

[CN22] C. Carstensen and N. Nataraj. Lowest-order equivalent nonstandard finite element methods for biharmonic plates. *ESAIM Math. Model. Numer. Anal.*, 56(1):41–78, 2022.

[CR73] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, 7:33–75, 1973.

[DDP95] E. Dari, R. Durán, and C. Padra. Error estimators for nonconforming finite element approximations of the Stokes problem. *Math. Comp.*, 64(211):1017–1033, 1995.

[DDPV96] E. Dari, R. Duran, C. Padra, and V. Vampa. A posteriori error estimators for nonconforming finite element methods. *RAIRO Modél. Math. Anal. Numér.*, 30(4):385–400, 1996.

[FV06] F. Fierro and A. Veeser. A posteriori error estimators, gradient recovery by averaging, and superconvergence. *Numerische Mathematik*, 103(2):267–298, April 2006.

[HKK$^+$] K.-J. Heine, D. Köster, O. Kriessl, A. Schmidt, and K. Siebert. Alberta - an adaptive hierachical finite element toolbox.

[KP03] 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(6):2374–2399, 2003.

[KRVZ] C. Kreuzer, M. Rott, A. Veeser, and P. Zanotti. Strictly equivalent a posteriori estimators for quasi-optimal nonconforming methods: fourth-order problems. In preparation.

[KV21] C. Kreuzer and A. Veeser. Oscillation in a posteriori error estimation. *Numer. Math.*, 148(1):43–78, 2021.

[KVZ] C. Kreuzer, A. Veeser, and P. Zanotti. Accurate error bounds for finite element methods. In preparation.

[Mor68] L. S. D. Morley. The triangular equilibrium element in the solution of plate bending problems. *The Aeronautical Quarterly*, 19:149–169, 1968.

[SB06] G. Stoyan and A. Baran. Crouzeix-Velte decompositions for higher-order finite elements. *Comput. Math. Appl.*, 51(6-7):967–986, 2006.

[SS05] A. Schmidt and K. G. Siebert. *Design of adaptive finite element software*, volume 42 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2005. The finite element toolbox ALBERTA, With 1 CD-ROM (Unix/Linux).

[Szy06] D. B. Szyld. The many proofs of an identity on the norm of oblique projections. *Numer. Algorithms*, 42(3-4):309–323, 2006.

[Ver13] R. Verfürth. *A posteriori error estimation techniques for finite element methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.

[Voh07] M. Vohralík. A posteriori error estimates for finite volume and mixed finite element discretizations of convection-diffusion-reaction equations. In *Paris-Sud Working Group on Modelling and Scientific Computing 2006–2007*, volume 18 of *ESAIM Proc.*, pages 57–69. EDP Sci., Les Ulis, 2007.

[VZ18a] A. Veeser and P. Zanotti. Quasi-optimal nonconforming methods for symmetric elliptic problems. I—Abstract theory. *SIAM J. Numer. Anal.*, 56(3):1621–1642, 2018.

[VZ18b] A. Veeser and P. Zanotti. Quasi-optimal nonconforming methods for symmetric elliptic problems. III—Discontinuous Galerkin and other interior penalty methods. *SIAM J. Numer. Anal.*, 56(5):2871–2894, 2018.

[VZ19] A. Veeser and P. Zanotti. Quasi-optimal nonconforming methods for symmetric elliptic problems. II—Overconsistency and classical nonconforming elements. *SIAM J. Numer. Anal.*, 57(1):266–292, 2019.

TU Dortmund, Fakultät für Mathematik, D-44221 Dortmund, Germany
Email address: christian.kreuzer@tu-dortmund.de

TU Dortmund, Fakultät für Mathematik, D-44221 Dortmund, Germany
Email address: matthias.rott@tu-dortmund.de

Università degli Studi di Milano, Dipartimento di Matematica, 20133 Milano, Italy
Email address: andreas.veeser@unimi.it

Università degli Studi di Milano, Dipartimento di Matematica, 20133 Milano, Italy
Email address: pietro.zanotti@unimi.it