Sharp algebraic and total a posteriori error bounds for h and p finite elements via a multilevel approach

Jan Papež, Ulrich Rüde, Martin Vohralík, Barbara Wohlmuth

Sharp algebraic and total a posteriori error bounds for h and p finite elements via a multilevel approach*

Jan Papež† Ulrich Rüde‡ Martin Vohralík§ Barbara Wohlmuth¶

February 7, 2018

Abstract

We derive guaranteed, fully computable, constant-free, and sharp upper and lower a posteriori estimates on the algebraic, total, and discretization errors of finite element approximations of the Poisson equation obtained by an arbitrary iterative solver. The estimators are computed locally over patches of mesh elements around vertices and are based on suitable liftings of the total and algebraic residuals. The key ingredient is the decomposition of the algebraic error over a hierarchy of meshes, with a global solve on the coarsest mesh. Distinguishing the algebraic and discretization error components allows us to formulate safe stopping criteria ensuring that the algebraic error does not dominate the total error. We also prove equivalence of our total estimate with the total error, up to a generic polynomial-degree-independent constant. Numerical experiments illustrate sharp control of all error components and accurate prediction of their spatial distribution in several test problems. These include smooth and singular solutions, higher-order conforming finite elements, and different multigrid methods as well as the preconditioned conjugate gradient method as the iterative solver.

Key words: finite element method, algebraic error, a posteriori error estimate, p-robustness, stopping criterion, hierarchical splitting

1 Introduction

Numerical approximation of partial differential equations often gives rise to large sparse systems of linear algebraic equations. Their efficient solution is often achieved by iterative solvers. Many are based on a hierarchy of nested meshes such as, e.g., multilevel and multigrid methods. Asymptotically optimal complexity has been originally shown on uniformly refined meshes, see, e.g., the monographs [15, 29] and the references therein. For convergence results on graded meshes, we refer to more recent contributions [54, 31, 55, 19]. Multigrid solvers are nowadays especially attractive in the context of massively parallel large scale simulations on modern architectures, see [26, 56, 27, 37] and the references therein. In many cases, however, other techniques like preconditioned Krylov solvers remain a vital and widely used alternative.

Two intriguing questions are what is the *error* in a *given step* of the iterative solver and how to realize *adaptive stopping criteria* which balance algebraic and discretization errors. Pioneering answers were proposed by Brandt [14], Bank and Sherman [6], Bai and Brandt [5], Bank and Smith [7], Rüde [46, 47, 48], or Oswald [38], or more recently by, e.g., Janssen and Kanschat [32]. Rigorous upper a posteriori error estimates of the total error valid at each step of the geometric multigrid were designed in the work of Becker

*This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 647134 GATIPOR) and by the German Research Foundation (DFG) through WO 671/11-1 and WO 671/15-1 (within the Priority Programme SPP 1748, "Reliable Simulation Techniques in Solid Mechanics. Development of Non-standard Discretisation Methods, Mechanical and Mathematical Analysis"). It was also supported by the ERC-CZ project LL1202 financed by the MŠMT of the Czech Republic.

†Inria Paris, 2 rue Simone Iff, 75589 Paris, France and Sorbonne Université, Université Paris-Diderot SPC, CNRS, Laboratoire Jacques-Louis Lions, équipe ALPINES. The work of this author was done during his doctoral studies at Faculty of Mathematics and Physics, Charles University, and at Institute of Computer Science, the Czech Academy of Sciences. (jan@papez.org).

‡Informatik 10, System Simulation, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany (ulrich.ruede@fau.de).

§Inria Paris, 2 rue Simone Iff, 75589 Paris, France & Université Paris-Est, CERMICS (ENPC), 77455 Marne-la-Vallée, France (martin.vohralik@inria.fr).

¶Fakultät für Mathematik, Lehrstuhl für Numerische Mathematik, Boltzmannstrasse 3, 85748 Garching bei München, Germany (wohlmuth@ma.tum.de).

et al. [9]. The estimates are then decomposed into parts that are identified with the algebraic and discretization error components, though not necessarily their upper or lower bounds. Consequently, adaptive stopping criteria are proposed and numerical experiments show that numerous iterations may be saved with respect to classical stopping criteria. These estimates involve some generic constants that are not at disposal in practical computations. Later, the development of a posteriori error estimates including algebraic error continued, e.g., in Arioli *et al.* [3, 2], Meidner *et al.* [35], Jiránek *et al.* [33], Ern and Vohralík [23], or Papež *et al.* [41], see also the references therein. Typically, some limitations still appear: the estimates in [33, 23, 41] apply to an arbitrary iterative solver, but additional iteration steps are required; the estimates in [3, 35, 2] provide a total bound which is not fully computable and/or the results only apply to one specific solver.

The present contribution develops a unified approach which seems rather flexible. Our main results can be summarized as follows:

1. We provide *constant-free* fully computable *upper* and *lower* bounds for all the *algebraic*, *discretization*, and *total* errors, independent of the choice of the iterative solver and including higher-order (conforming) finite elements. We also prove that our total upper bound is *efficient*, i.e., equivalent to the total error up to a generic constant, independent of the approximation polynomial degree.

2. The evaluation of the estimates is based on a *global coarse solve* and independent *local solves* associated with vertex patches, based on a hierarchy of nested meshes.

3. *Safe stopping criteria* for iterative solvers are proposed. They ensure that the algebraic system will neither be over-solved, nor under-solved. *Mass balance* is simultaneously guaranteed on each iteration by a post-processing of the current solution.

Property 1 states that all total, algebraic, and discretization errors are fully certified[1]. We derive the upper bounds on the algebraic and total errors in the spirit of Prager and Synge [44], relying on the a posteriori methodology of *equilibrated flux* reconstructions that can be seen as $\mathbf{H}(\mathrm{div},\Omega)$-liftings of the algebraic and total residuals. These have been proposed and studied for finite elements with an exact algebraic solve in Destuynder and Métivet [20], Luce and Wohlmuth [34], Braess and Schöberl [12], Nicaise *et al.* [36], Ainsworth [1], Vohralík [53], and Ern and Vohralík [23, 24, 25], see also the references therein. Relying on the recent results of Braess *et al.* [11] and Ern and Vohralík [25], *p-robust* efficiency is shown which means that the derived total upper bound will be of the same quality for any order finite elements. Our lower bounds on the algebraic and total errors are then obtained via $H^1_0(\Omega)$-liftings of the algebraic and total residuals.

Property 2 improves the approaches of [33, Sec. 7.2] and [23, 41], where an a priori unknown number of additional steps of the iterative algebraic solver needs to be performed so as to evaluate the estimates. In contrast, the present approach is typically of linear cost in terms of the number of mesh vertices. More precisely, a sweep through all the vertices on all levels is necessary, where small *mixed finite element problems* on patches of elements are solved to obtain the $\mathbf{H}(\mathrm{div},\Omega)$-*liftings* and small *conforming finite element problems* on patches of elements are solved to obtain the $H^1_0(\Omega)$-*liftings*. In terms of computational efficiency, this may be acceptable; in particular, when used in an integrated strategy where solver and estimator go hand in hand, it may be performed simply as part of a multigrid solver. Implementation-wise, the matrices corresponding to the above-described local problems can be assembled and inverted only once, prior to the start of the iterative solver. In each iteration, the estimators are then evaluated merely by simple matrix-vector multiplications. Actually, any solution of local patch problems (local matrices assembly and inversion) can be avoided by an explicit variant that we discuss in details in Section 8.3 below.

Finally, two salient consequences of Property 3 are as follows: it typically brings an important economy in terms of the number of algebraic solver iterations and, second, a very desirable feature in many applications such as porous media flows gets satisfied.

Our hierarchical decomposition of the algebraic error shares some ideas with those of stable splittings [7, 39, 38, 47, 48, 28] for multigrid and multilevel methods. The key differences of our approach, though, seem to be: a) our estimates give guaranteed and computable bounds; b) no generic multiplicative constant appears; c) no regularity beyond $H^1(\Omega)$ of the weak solution is requested; d) the hierarchy of meshes needs to be neither structured nor uniformly refined; e) any polynomial degree $p \geq 1$ is allowed; f) no inverse inequality is employed in our theory; h) the derived estimates hold true for any iterative solver. In contrast to these splitting techniques, however, our approach is typically more expensive.

The rest of this contribution is organized as follows: In Sec. 2, we introduce the model problem, set up the notation, and define discrete spaces. Sec. 3 introduces the concept of algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-liftings used to estimate the algebraic error from above, as well as of algebraic residual $H^1_0(\Omega)$-liftings used

[1] Supposing that computer evaluation of the data oscillation term for total and discretization errors is possible.

to estimate the algebraic error from below. Next, total residual broken $\mathbf{H}(\mathrm{div},\Omega)$- and total residual $H^1_0(\Omega)$-liftings are introduced in Sec. 4. The main results of our paper are summarized in Sec. 5, where we include short (instructive) proofs. Sec. 6 consequently brings details of the construction of the algebraic liftings, and Sec. 7 describes the total liftings. We discuss in Sec. 8 efficient implementation of the local linear systems and possible reductions of the induced cost. Numerical performance of the estimates is next examined in Sec. 9 for V- and full geometric multigrid cycles, as well as for the preconditioned conjugate gradient method. A concluding discussion is presented in Sec. 10, pointing to some current developments concerning implementation and relationships of the present methodology to estimates based on stable splittings. Finally, in Appendix A we demonstrate the importance of the hierarchical multilevel construction of the $H^1_0(\Omega)$-lifting of the algebraic residual by its comparison with a simpler finest-level-only procedure, and in Appendix B, we give the proof of the local and global polynomial-degree-robust efficiency of the total error bound.

2 Model problem and setting

This section introduces the considered model Poisson problem, its finite element discretization, the solution of the arising system of linear equations by an arbitrary algebraic solver, the notion of the algebraic and total residuals, and finally defines discrete spaces used in the construction of our estimates.

2.1 Model problem

Let $\Omega \subset \mathbb{R}^d$, $1 \le d \le 3$, be an open bounded polytope with a Lipschitz-continuous boundary and let the source term $f \in L^2(\Omega)$. We look for a function $u : \Omega \to \mathbb{R}$ such that

$$-\Delta u = f \qquad \text{in } \Omega, \tag{2.1a}$$
$$u = 0 \qquad \text{on } \partial\Omega. \tag{2.1b}$$

In weak form, problem (2.1) reads: find $u \in H^1_0(\Omega)$ such that

$$(\nabla u, \nabla v) = (f, v) \qquad \forall v \in H^1_0(\Omega), \tag{2.2}$$

with $(\cdot,\cdot)$ standing for the $L^2(\Omega)$ or $[L^2(\Omega)]^d$ scalar product. There exists one and only one solution to (2.2) by the Riesz representation theorem.

Let $\mathcal{T}_h$ be a simplicial mesh of Ω, matching in the sense that for two distinct elements K of $\mathcal{T}_h$, their intersection is either an empty set or a common vertex, edge, or face. Associated with $\mathcal{T}_h$, let there be a hierarchy of meshes $\{\mathcal{T}_j\}_{0\le j\le J}$. These are again matching simplicial partitions of the domain Ω, *nested* in the sense that $\mathcal{T}_j$ is a refinement of $\mathcal{T}_{j-1}$, $1 \le j \le J$, and satisfying $\mathcal{T}_h = \mathcal{T}_J$. The set of vertices of $\mathcal{T}_j$ is denoted by $\mathcal{V}_j$, and it is decomposed into interior vertices $\mathcal{V}_j^{\mathrm{int}}$ and boundary vertices $\mathcal{V}_j^{\mathrm{ext}}$. By $\psi_j^{\mathbf{a}}$ we denote the standard hat function associated with the vertex $\mathbf{a} \in \mathcal{V}_j$, $0 \le j \le J$, i.e., the function that is piecewise affine with respect to the j-th level mesh $\mathcal{T}_j$, taking the value 1 at the vertex $\mathbf{a}$ and zero at all other j-th level vertices of $\mathcal{V}_j$. The support of $\psi_j^{\mathbf{a}}$ is denoted by $\omega_j^{\mathbf{a}}$ and it corresponds to the patch of elements of $\mathcal{T}_j$ which share the vertex $\mathbf{a} \in \mathcal{V}_j$. We identify $\omega_h^{\mathbf{a}} := \omega_J^{\mathbf{a}}$ and $\mathcal{V}_h := \mathcal{V}_J$, $\mathcal{V}_h^{\mathrm{int}} := \mathcal{V}_J^{\mathrm{int}}$, $\mathcal{V}_h^{\mathrm{ext}} := \mathcal{V}_J^{\mathrm{ext}}$. The pth-order conforming finite element space associated with the mesh $\mathcal{T}_j$ is denoted by V_j^p and defined as

$$V_j^p := \mathbb{P}_p(\mathcal{T}_j) \cap H^1_0(\Omega), \quad \mathbb{P}_p(\mathcal{T}_j) := \{v_h \in L^2(\Omega), v_h|_K \in \mathbb{P}_p(K) \quad \forall K \in \mathcal{T}_j\}, \quad 1 \le p,\ 0 \le j \le J.$$

In the following, p will be fixed, and we set $V_h := V_J^p$ and $N_h := \dim V_h$. This choice is adopted for the simplicity of presentation; taking into account a variable polynomial degree in the developments below is possible along the lines of [22] and will be addressed elsewhere.

The Galerkin finite element approximation of (2.2) consists in finding $u_h^{\mathrm{ex}} \in V_h$ such that

$$(\nabla u_h^{\mathrm{ex}}, \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_h. \tag{2.3}$$

Let ψ_h^l, $1 \le l \le N_h$, form a basis of V_h. Then problem (2.3) is equivalent to solving a system of linear algebraic equations with a symmetric positive definite matrix: find $\mathrm{U}_h^{\mathrm{ex}} \in \mathbb{R}^{N_h}$ such that

$$\mathbb{A}_h \mathrm{U}_h^{\mathrm{ex}} = \mathrm{F}_h, \tag{2.4}$$

where $(\mathbb{A}_h)_{lm} := (\nabla\psi_h^m, \nabla\psi_h^l)$, $(\mathrm{F}_h)_l := (f, \psi_h^l)$, and $u_h^{\mathrm{ex}} = \sum_{m=1}^{N_h} (\mathrm{U}_h^{\mathrm{ex}})_m \, \psi_h^m$. We note that $\mathrm{U}_h^{\mathrm{ex}}$ depends on the choice of the basis of V_h while $u_h^{\mathrm{ex}} \in V_h$ does not.

2.2 Algebraic and total residuals

Let $\mathrm{U}_h \in \mathbb{R}^{N_h}$ be an *arbitrary approximation* to the exact solution $\mathrm{U}_h^{\mathrm{ex}}$ of (2.4), corresponding to

$$u_h = \sum_{m=1}^{N_h} (\mathrm{U}_h)_m \, \psi_h^m \in V_h. \tag{2.5}$$

The *algebraic residual vector* is then

$$\mathrm{R}_h := \mathrm{F}_h - \mathbb{A}_h \mathrm{U}_h. \tag{2.6}$$

Following Papež *et al.* [41], we associate with R_h a *discontinuous elementwise polynomial* r_h of degree p, vanishing on the boundary of Ω, i.e., $r_h \in \mathbb{P}_p(\mathcal{T}_h)$, $r_h|_{\partial\Omega} = 0$. Denote by N_h^l the number of elements forming the support of the basis function ψ_h^l, $1 \le l \le N_h$. In two space dimensions and for affine elements with $p = 1$, one typically has $N_h^l = 6$, depending on the structure of the mesh and of the applied refinement rules. Then, for each fixed element $K \in \mathcal{T}_h$, we define $r_h|_K \in \mathbb{P}_p(K)$ by

$$(r_h, \psi_h^l)_K = \frac{(\mathrm{R}_h)_l}{N_h^l}, \quad r_h|_{\partial K \cap \partial\Omega} = 0, \tag{2.7}$$

for all basis functions ψ_h^l of the space V_h non-vanishing on K. Such r_h satisfies obviously $(\mathrm{R}_h)_l = (r_h, \psi_h^l)$, $1 \le l \le N_h$, and the algebraic relation (2.6) yields

$$(r_h, v_h) = (f, v_h) - (\nabla u_h, \nabla v_h) \qquad \forall v_h \in V_h. \tag{2.8}$$

We point out that although r_h is uniquely defined by (2.7), it is not the unique element in $\mathbb{P}_p(\mathcal{T}_h)$ which satisfies (2.8). It is also possible to define $r_h \in V_h$ directly by (2.8), as in, e.g., [9]. However, this requires the solution of a globally coupled mass system, see also [41, Sec. 5.1]. Thus we prefer to work with (2.7).

The *total residual* $\mathcal{R}_h$ for $u_h \in V_h$ is a functional on $H_0^1(\Omega)$. More precisely, $\mathcal{R}_h \in H^{-1}(\Omega)$ is given by

$$\langle \mathcal{R}_h, v \rangle := (f, v) - (\nabla u_h, \nabla v) \qquad \forall v \in H_0^1(\Omega). \tag{2.9}$$

Note that in contrast to the algebraic residual R_h (respectively its representer r_h) that are fully at our disposal, we can only access the total residual $\mathcal{R}_h$ via its action on test functions $v \in H_0^1(\Omega)$.

2.3 Finite element spaces for $\mathbf{H}(\mathrm{div}, \Omega)$- and $H_0^1(\Omega)$-liftings

For our estimates, two types of liftings of the algebraic and total residuals will be needed. First, we construct $\mathbf{H}(\mathrm{div}, \Omega)$-conforming vector-valued liftings in Raviart–Thomas (–Nédélec in three space dimensions)(RTN) mixed finite element spaces $\mathbf{V}_j^q \times Q_j^q \subset \mathbf{H}(\mathrm{div}, \Omega) \times L^2(\Omega)$, $q \ge 0$. These contain vector-valued piecewise polynomials with continuous normal trace and are given by $\mathbf{V}_j^q := \{\mathbf{v}_j \in \mathbf{H}(\mathrm{div}, \Omega); \mathbf{v}_j|_K \in [\mathbb{P}_q(K)]^d + \mathbb{P}_q(K)\mathbf{x} \quad \forall K \in \mathcal{T}_j\}$ and $Q_j^q := \mathbb{P}_q(\mathcal{T}_j)$, see, e.g., Brezzi and Fortin [17]. We denote $\mathbf{V}_h := \mathbf{V}_J^q$ and $Q_h := Q_J^q$ for a fixed polynomial degree q that we choose as either $q = p$ or $q = p + 1$, where, recall, p is the polynomial degree of the finite element approximation space $V_h = V_J^p$. Second, $H_0^1(\Omega)$-conforming scalar-valued liftings of the algebraic and total residuals will be performed in subspaces of the finite element spaces V_j^p. Henceforth, for a given space X, we use $X(\omega_j^{\mathbf{a}})$ to denote its restriction to the subdomain $\omega_j^{\mathbf{a}} \subset \Omega$.

3 Hierarchical liftings of the algebraic residual

Locally defined discrete $\mathbf{H}(\mathrm{div}, \Omega)$- and $H_0^1(\Omega)$-conforming *liftings* of the *algebraic residual* vector R_h of (2.6) play a crucial role in the definition of our error estimators. We now describe their concept. We start from the coarsest level *Riesz representer* $\rho_{0,\mathrm{alg}} \in V_0^p$ of the algebraic error given by

$$(\nabla \rho_{0,\mathrm{alg}}, \nabla v_0) = (r_h, v_0) \qquad \forall v_0 \in V_0^p, \tag{3.1}$$

familiar from the coarsest-grid residual solve in multigrid methods.[2] We start by a motivation.

[2] From the theoretical viewpoint, it would be sufficient to consider piecewise affine functions from V_0^1 in (3.1) in place of V_0^p.

3.1 Motivation

Consider a hierarchical decomposition of the algebraic error

$$\tilde{\rho}_{h,\mathrm{alg}} := \sum_{j=0}^{J} \tilde{\rho}_{j,\mathrm{alg}} \ \in V_h, \tag{3.2a}$$

$$\tilde{\rho}_{j,\mathrm{alg}} \in V_j^p, \qquad (\nabla\tilde{\rho}_{j,\mathrm{alg}}, \nabla v_j) = (r_h, v_j) - \sum_{i=0}^{j-1} (\nabla\tilde{\rho}_{i,\mathrm{alg}}, \nabla v_j) \qquad \forall v_j \in V_j^p, \quad \forall j = 0 \dots J. \tag{3.2b}$$

Here $\tilde{\rho}_{0,\mathrm{alg}} = \rho_{0,\mathrm{alg}}$ and $\tilde{\rho}_{j,\mathrm{alg}}$, $j \geq 1$, are corrections on increasing mesh levels recursively built from the coarsest-level. From definition (3.2b) with $j = J$ and from (3.2a), it is obvious that

$$(\nabla\tilde{\rho}_{h,\mathrm{alg}}, \nabla v_h) = (r_h, v_h) \qquad \forall v_h \in V_h,$$

i.e., $\tilde{\rho}_{h,\mathrm{alg}}$ given by (3.2a) is the Riesz representer of the algebraic residual r_h on the fine finite element space V_h. It is thus immediate from the characterizations (2.8) of r_h and (2.3) of u_h^{ex} that $\tilde{\rho}_{h,\mathrm{alg}} = u_h^{\mathrm{ex}} - u_h$, i.e., $\tilde{\rho}_{h,\mathrm{alg}}$ represents the exact algebraic error. Moreover, we find by induction, using the definition of $\tilde{\rho}_{i,\mathrm{alg}}$ and taking in (3.2b) $v_j = \tilde{\rho}_{i,\mathrm{alg}}$, $i < j$, the orthogonality relation

$$(\nabla\tilde{\rho}_{j,\mathrm{alg}}, \nabla\tilde{\rho}_{i,\mathrm{alg}}) = 0 \qquad 0 \leq i, j \leq J,\ j \neq i. \tag{3.3}$$

Consequently, the algebraic error admits the following hierarchical decomposition

$$\|\nabla(u_h^{\mathrm{ex}} - u_h)\|^2 = \|\nabla\tilde{\rho}_{h,\mathrm{alg}}\|^2 = \sum_{j=0}^{J} \|\nabla\tilde{\rho}_{j,\mathrm{alg}}\|^2. \tag{3.4}$$

3.2 $\mathbf{H}(\mathrm{div},\Omega)$- and $H_0^1(\Omega)$-liftings of the algebraic residual

Our $\mathbf{H}(\mathrm{div},\Omega)$- and $H_0^1(\Omega)$-liftings $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ and $\rho_{h,\mathrm{alg}}$, respectively, are approximate versions of (3.2b) via localization over patches of elements. When the approximation is done by mixed finite elements, we obtain an algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-conforming lifting $\boldsymbol{\sigma}_{h,\mathrm{alg}} \in \mathbf{V}_h \subset \mathbf{H}(\mathrm{div},\Omega)$ such that $\nabla\cdot\boldsymbol{\sigma}_{h,\mathrm{alg}} = r_h$. This will lead to an upper algebraic error bound $\|\nabla(u_h^{\mathrm{ex}} - u_h)\| \leq \|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|$. When conforming finite elements are employed, an algebraic residual $H_0^1(\Omega)$-conforming lifting $\rho_{h,\mathrm{alg}} \in V_h \subset H_0^1(\Omega)$, approximating $\tilde{\rho}_{h,\mathrm{alg}}$, will be obtained, yielding a lower bound. The construction of $\rho_{h,\mathrm{alg}}$ can in particular be regarded as applying the restrictive additive Schwarz method for the problems (3.2b); see Sec. 6 for the discussion and the references.

Let a mesh level $1 \leq j \leq J$ be fixed, together with a vertex $\mathbf{a} \in \mathcal{V}_{j-1}$, i.e., a vertex from the next coarser mesh. The key ingredient for our algebraic residual liftings are local spaces associated with the mesh of level j but defined on the subdomain $\omega_{j-1}^{\mathbf{a}}$ of all simplices of level $j-1$ sharing the vertex $\mathbf{a}$. $\mathbf{H}(\mathrm{div},\Omega)$-conforming liftings are in particular be constructed in

$$\begin{aligned} &\mathbf{V}_{j,j-1}^{\mathbf{a}} := \{\mathbf{v}_j \in \mathbf{V}_j^q(\omega_{j-1}^{\mathbf{a}});\ \mathbf{v}_j\cdot\mathbf{n}_{\omega_{j-1}^{\mathbf{a}}} = 0 \text{ on } \partial\omega_{j-1}^{\mathbf{a}}\},\\ &Q_{j,j-1}^{\mathbf{a}} := \{q_j \in Q_j^q(\omega_{j-1}^{\mathbf{a}});\ (q_j, 1)_{\omega_{j-1}^{\mathbf{a}}} = 0\}, \end{aligned} \qquad \mathbf{a} \in \mathcal{V}_{j-1}^{\mathrm{int}}, \tag{3.5a}$$

$$\begin{aligned} &\mathbf{V}_{j,j-1}^{\mathbf{a}} := \{\mathbf{v}_j \in \mathbf{V}_j^q(\omega_{j-1}^{\mathbf{a}});\ \mathbf{v}_j\cdot\mathbf{n}_{\omega_{j-1}^{\mathbf{a}}} = 0 \text{ on } \partial\omega_{j-1}^{\mathbf{a}} \setminus \partial\Omega\},\\ &Q_{j,j-1}^{\mathbf{a}} := Q_j^q(\omega_{j-1}^{\mathbf{a}}), \end{aligned} \qquad \mathbf{a} \in \mathcal{V}_{j-1}^{\mathrm{ext}}, \tag{3.5b}$$

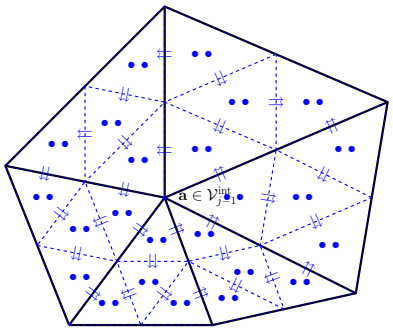
where $\mathbf{n}_{\omega_{j-1}^{\mathbf{a}}}$ stands for the outward unit normal of the domain $\omega_{j-1}^{\mathbf{a}}$. The degrees of freedom of the spaces $\mathbf{V}_{j,j-1}^{\mathbf{a}}$ for an interior vertex with polynomial degrees $p = q = 1$ are illustrated in Fig. 1, left. Note that all functions in spaces $\mathbf{V}_{j,j-1}^{\mathbf{a}}$ and $Q_{j,j-1}^{\mathbf{a}}$ extended by zero outside of the patch domain $\omega_{j-1}^{\mathbf{a}}$ belong to the spaces $\mathbf{V}_j^q$ and Q_j^q defined over the entire computational domain Ω, respectively. By Π_j we denote the $L^2(\Omega)$-orthogonal projection onto Q_j^q, we set $\Pi_h := \Pi_J$, and let $\Pi_{Q_{j,j-1}^{\mathbf{a}}}$ stand for the $L^2(\omega_{j-1}^{\mathbf{a}})$-orthogonal projection onto $Q_{j,j-1}^{\mathbf{a}}$.

The $H_0^1(\Omega)$-conforming liftings are in their turn constructed in the spaces

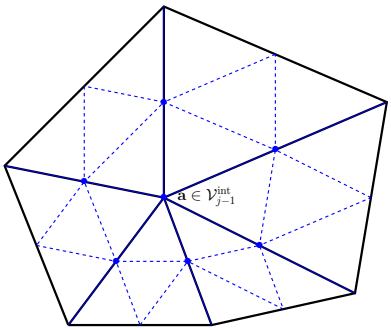
$$V_{j,j-1}^{\mathbf{a}} := V_j^p(\omega_{j-1}^{\mathbf{a}}) \cap H_0^1(\omega_{j-1}^{\mathbf{a}}) = \{v_h \in H_0^1(\omega_{j-1}^{\mathbf{a}}), v_h|_K \in \mathbb{P}_p(K) \quad \forall K \in \mathcal{T}_j \subset \omega_{j-1}^{\mathbf{a}}\}. \tag{3.6}$$

Again, these are level j spaces defined over the coarser patches $\omega_{j-1}^{\mathbf{a}}$, and its elements extended by zero belong to V_j^p. An illustration of the setting for an interior vertex and $p = 1$ is given in Fig. 1, right.

We present the liftings in Concepts 3.1 and 3.2 which clearly reflect the hierarchical structure and the local patch-wise construction on all but the coarsest level. Details are postponed to Sec. 6.



coarse patch $\omega_{j-1}^{\mathbf{a}}$ for $\mathbf{a} \in \mathcal{V}_{j-1}^{\text{int}}$ (full line)
fine mesh $\mathcal{T}_j$ of $\omega_{j-1}^{\mathbf{a}}$ (dashed line)
degrees of freedom of $\mathbf{V}_{j,j-1}^{\mathbf{a}}$ (arrows and bullets)



coarse patch $\omega_{j-1}^{\mathbf{a}}$ for $\mathbf{a} \in \mathcal{V}_{j-1}^{\text{int}}$ (full line)
fine mesh $\mathcal{T}_j$ of $\omega_{j-1}^{\mathbf{a}}$ (dashed line)
degrees of freedom of $V_{j,j-1}^{\mathbf{a}}$ (bullets)

Figure 1: Degrees of freedom of the local Raviart–Thomas spaces $\mathbf{V}_{j,j-1}^{\mathbf{a}}$ (left) and of the local Lagrange spaces $V_{j,j-1}^{\mathbf{a}}$ (right) for algebraic error liftings, $p = q = 1$, interior vertex $\mathbf{a} \in \mathcal{V}_{j-1}^{\text{int}}$

Basic property:

$$\boldsymbol{\sigma}_{h,\text{alg}} \in \mathbf{V}_h \subset \mathbf{H}(\text{div}, \Omega), \tag{3.7a}$$
$$\nabla{\cdot}\boldsymbol{\sigma}_{h,\text{alg}} = r_h. \tag{3.7b}$$

Construction (Sec. 6.1, Definition 6.3):

1. Solve the coarse problem (3.1).
2. Run through mesh levels $j = 1 \dots J$:
   (a) Run through all coarser vertices $\mathbf{a} \in \mathcal{V}_{j-1}$.
   (b) Assemble the RTN spaces $\mathbf{V}_{j,j-1}^{\mathbf{a}}$ and $Q_{j,j-1}^{\mathbf{a}}$ from (3.5), see Fig. 1, left.
   (c) Solve the local Neumann mixed finite element problems (6.6) to obtain $\boldsymbol{\sigma}_{j,\text{alg}}^{\mathbf{a}} \in \mathbf{V}_{j,j-1}^{\mathbf{a}}$.
3. Combine

$$\boldsymbol{\sigma}_{h,\text{alg}} := \sum_{j=1}^{J} \sum_{\mathbf{a} \in \mathcal{V}_{j-1}} \boldsymbol{\sigma}_{j,\text{alg}}^{\mathbf{a}}.$$

Simplifications:

- Explicit construction avoiding local patchwise problems (6.6): Sec. 8.3
- Deriving a bound avoiding physical construction of $\boldsymbol{\sigma}_{h,\text{alg}}$: possible.

Concept 3.1: Algebraic residual $\mathbf{H}(\text{div}, \Omega)$-lifting

Basic property:

$$\rho_{h,\text{alg}} \in V_h \subset H_0^1(\Omega). \tag{3.8}$$

Construction (Sec. 6.2, Definition 6.5):

1. Solve the coarse problem (3.1).
2. Run through mesh levels $j = 1 \dots J$:
   (a) Run through all coarser vertices $\mathbf{a} \in \mathcal{V}_{j-1}$.
   (b) Assemble the Lagrange spaces $V_{j,j-1}^{\mathbf{a}}$ from (3.6), see Fig. 1, right.
   (c) Solve the local Dirichlet conforming finite element problems (6.8) to obtain $\rho_{j,\text{alg}}^{\mathbf{a}} \in V_{j,j-1}^{\mathbf{a}}$.
3. Combine

$$\rho_{h,\text{alg}} := \rho_{0,\text{alg}} + \sum_{j=1}^{J} \sum_{\mathbf{a} \in \mathcal{V}_{j-1}} w_{\psi_{j-1}^{\mathbf{a}}}(\rho_{j,\text{alg}}^{\mathbf{a}}),$$

where $w_{\psi_{j-1}^{\mathbf{a}}}(\cdot)$ is a weighting by the partition of unity such that $w_{\psi_{j-1}^{\mathbf{a}}}(\rho_{j,\text{alg}}^{\mathbf{a}}) \in V_{j,j-1}^{\mathbf{a}} \subset V_j^p$.

Simplifications:

- Explicit construction avoiding local patchwise problems (6.8): Sec. 8.3
- Deriving a bound avoiding physical construction of $\rho_{h,\text{alg}}$: possible.

Concept 3.2: Algebraic residual $H_0^1(\Omega)$-lifting

4 $\mathbf{H}(\text{div}, \Omega)$- and $H_0^1(\Omega)$-liftings of the total residual

We now introduce the concepts of the total residual broken $\mathbf{H}(\text{div}, \Omega)$-lifting $\boldsymbol{\sigma}_{h,\text{dis}}$, and of the total residual $H_0^1(\Omega)$-lifting $\rho_{h,\text{tot}}$. Details are again postponed, to Sec. 7. In contrast to (3.5)–(3.6), our total residual liftings will only be defined on spaces associated with the finest mesh $\mathcal{T}_h$ and finest mesh patches $\omega_h^{\mathbf{a}}$, $\mathbf{a} \in \mathcal{V}_h$. We will in particular employ Raviart–Thomas(–Nédélec) $\mathbf{H}(\text{div}, \Omega)$-conforming vector-valued spaces with homogeneous Neumann boundary condition given by

$$\begin{aligned} &\mathbf{V}_h^{\mathbf{a}} := \{\mathbf{v}_h \in \mathbf{V}_h(\omega_h^{\mathbf{a}});\ \mathbf{v}_h{\cdot}\mathbf{n}_{\omega_h^{\mathbf{a}}} = 0 \text{ on } \partial\omega_h^{\mathbf{a}}\}, \\ &Q_h^{\mathbf{a}} := \{q_h \in Q_h(\omega_h^{\mathbf{a}});\ (q_h, 1)_{\omega_h^{\mathbf{a}}} = 0\}, \end{aligned} \qquad \mathbf{a} \in \mathcal{V}_h^{\text{int}}, \tag{4.1a}$$

$$\begin{aligned} &\mathbf{V}_h^{\mathbf{a}} := \{\mathbf{v}_h \in \mathbf{V}_h(\omega_h^{\mathbf{a}});\ \mathbf{v}_h{\cdot}\mathbf{n}_{\omega_h^{\mathbf{a}}} = 0 \text{ on } \partial\omega_h^{\mathbf{a}} \setminus \partial\Omega\}, \\ &Q_h^{\mathbf{a}} := Q_h(\omega_h^{\mathbf{a}}), \end{aligned} \qquad \mathbf{a} \in \mathcal{V}_h^{\text{ext}}, \tag{4.1b}$$

see Fig. 2, left, for $p = q = 1$. The corresponding $H^1(\omega_h^{\mathbf{a}})$-conforming scalar-valued spaces are given by

$$V_h^{\mathbf{a}} := \{v_h \in H^1(\omega_h^{\mathbf{a}}),\ v_h|_K \in \mathbb{P}_p(K) \quad \forall K \in \mathcal{T}_h \subset \omega_h^{\mathbf{a}},\ (v_h, 1)_{\omega_h^{\mathbf{a}}} = 0\} \qquad \mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}, \tag{4.2a}$$

$$V_h^{\mathbf{a}} := \{v_h \in H^1(\omega_h^{\mathbf{a}}),\ v_h|_K \in \mathbb{P}_p(K) \quad \forall K \in \mathcal{T}_h \subset \omega_h^{\mathbf{a}},\ v_h = 0 \text{ on } \partial\omega_h^{\mathbf{a}} \cap \partial\Omega\} \qquad \mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}, \tag{4.2b}$$

see Fig. 2, right, for $p = q = 1$.

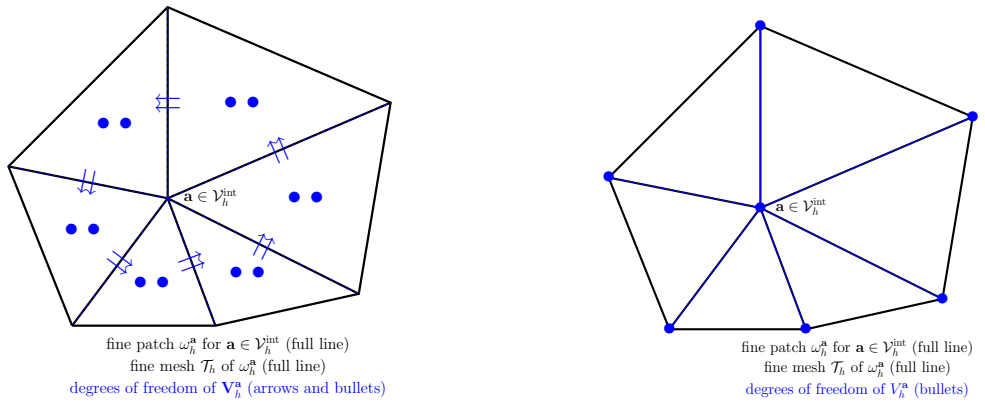


Figure 2: Degrees of freedom of the local Raviart–Thomas spaces $\mathbf{V}_h^{\mathbf{a}}$ (left) and of the local Lagrange spaces $V_h^{\mathbf{a}}$ (right) for total error liftings, $p = q = 1$, interior vertex $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$

We present the liftings of the total residual in Concepts 4.1 and 4.2. The total lifting $\boldsymbol{\sigma}_{h,\mathrm{tot}}$ actually lies in the broken $\mathbf{H}(\mathrm{div}, \Omega)$ space, or more precisely in its broken Raviart–Thomas(–Nédélec) finite-dimensional subspace. Indeed, the normal component $\boldsymbol{\sigma}_{h,\mathrm{tot}}$ is by (4.3b) below not continuous, which is linked to the fact that by a simple integration by parts, we find from (2.9)

$$\langle \mathcal{R}_h, v\rangle := (f, v) - (\nabla u_h, \nabla v) = \sum_{K\in\mathcal{T}_h} (f + \Delta u_h, v)_K - \sum_{e\in\mathcal{E}_h^{\mathrm{int}}} (\llbracket \nabla u_h \rrbracket \cdot \mathbf{n}_e, v)_e \qquad \forall v \in H_0^1(\Omega).$$

Here $\mathcal{E}_h^{\mathrm{int}}$ denotes the inner vertices (for $d = 1$), edges ($d = 2$) or faces ($d = 3$).

Basic property:

$$\boldsymbol{\sigma}_{h,\mathrm{tot}}|_K \in \mathbf{V}_h(K) \quad \forall K \in \mathcal{T}_h, \tag{4.3a}$$

$$\llbracket \boldsymbol{\sigma}_{h,\mathrm{tot}} \rrbracket \cdot \mathbf{n}_e = \llbracket \nabla u_h \rrbracket \cdot \mathbf{n}_e \quad \forall e \in \mathcal{E}_h^{\mathrm{int}}, \tag{4.3b}$$

$$(\nabla \cdot \boldsymbol{\sigma}_{h,\mathrm{tot}})|_K = (\Pi_h f + \Delta u_h)|_K \ \forall K \in \mathcal{T}_h. \tag{4.3c}$$

Construction (Sec. 7.1, Definition 7.1):

1. Run through all finest vertices $\mathbf{a} \in \mathcal{V}_h$.
2. Assemble the RTN spaces $\mathbf{V}_h^{\mathbf{a}}$ and $Q_h^{\mathbf{a}}$ from (4.1), see Fig. 2, left.
3. Solve the local Neumann mixed finite element problems (7.1) to obtain $\boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}} \in \mathbf{V}_h^{\mathbf{a}}$.
4. Combine $\boldsymbol{\sigma}_{h,\mathrm{dis}} := \sum_{\mathbf{a}\in\mathcal{V}_h} \boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}} \in \mathbf{V}_h$, yielding $\nabla\cdot\boldsymbol{\sigma}_{h,\mathrm{dis}} = \Pi_h f - r_h$.
5. Obtain $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ following Concept 3.1.
6. Set

$$\boldsymbol{\sigma}_{h,\mathrm{tot}} := \nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{alg}} + \boldsymbol{\sigma}_{h,\mathrm{dis}}. \tag{4.4}$$

Simplifications:

- Explicit construction avoiding local patchwise problems (7.1): Sec. 8.3.
- Deriving a bound avoiding physical construction of $\boldsymbol{\sigma}_{h,\mathrm{dis}}$: possible.

Concept 4.1: Total residual broken $\mathbf{H}(\mathrm{div}, \Omega)$-lifting

Basic property:

$$\rho_{h,\mathrm{tot}} \in V_J^{p+1} \subset H_0^1(\Omega). \tag{4.5}$$

Construction (Sec. 7.2, Definition 7.3):

1. Run through all finest vertices $\mathbf{a} \in \mathcal{V}_h$.
2. Assemble the Lagrange spaces $V_h^{\mathbf{a}}$ from (4.2), see Fig. 2, right.
3. Solve the local Neumann conforming finite element problems (7.2) to obtain $\rho_{h,\mathrm{tot}}^{\mathbf{a}} \in V_h^{\mathbf{a}}$.
4. Combine

$$\rho_{h,\mathrm{tot}} := \sum_{\mathbf{a}\in\mathcal{V}_h} \psi_h^{\mathbf{a}} \rho_{h,\mathrm{tot}}^{\mathbf{a}}.$$

Simplifications:

- Explicit construction avoiding local patchwise problems (7.2): Sec. 8.3.
- Deriving a bound avoiding physical construction of $\rho_{h,\mathrm{tot}}$: possible.

Concept 4.2: Total residual $H_0^1(\Omega)$-lifting

5 Main theoretical results

This section summarizes our main theoretical results. Guaranteed upper and lower bounds are presented in Sec. 5.1. We subsequently use these bounds to design mathematically sound stopping criteria in Sec. 5.2. Under these criteria, efficiency, i.e., optimal properties of our upper estimates on the total error, are stated in Sec. 5.3.

All the presented results are valid for *any* $u_h \in V_h$, in particular on *any step* of *any algebraic iterative solver* applied to the linear system (2.4). We recall that $u \in H^1_0(\Omega)$ is the weak solution given by (2.2), $u_h^{\mathrm{ex}} \in V_h$ its (unavailable) finite element approximation of order p given by (2.3), and $r_h \in \mathbb{P}_p(\mathcal{T}_h)$ is the corresponding algebraic residual representer of (2.8). Our first observation is that our $\mathbf{H}(\mathrm{div},\Omega)$-liftings automatically guarantee mass balance at any time.

Lemma 5.1 (Local mass balance for an arbitrary $u_h \in V_h$)**.** *For $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ and $\boldsymbol{\sigma}_{h,\mathrm{dis}}$ of Concepts 3.1 and 4.1, there holds*

$$\boldsymbol{\sigma}_{h,\mathrm{alg}} + \boldsymbol{\sigma}_{h,\mathrm{dis}} \in \mathbf{V}_h \subset \mathbf{H}(\mathrm{div},\Omega), \tag{5.1a}$$

$$\nabla\cdot(\boldsymbol{\sigma}_{h,\mathrm{alg}} + \boldsymbol{\sigma}_{h,\mathrm{dis}}) = \Pi_h f. \tag{5.1b}$$

The lemma follows from Lemmas 6.4 and 7.2 below.

5.1 Guaranteed upper and lower error bounds

Relying on the (broken) $\mathbf{H}(\mathrm{div},\Omega)$-*liftings* of Concepts 3.1 and 4.1, with either $q = p$ or $q = p+1$, *upper bounds* on the algebraic and total errors are obtained:

Theorem 5.2 (Guaranteed upper bounds on algebraic and total errors)**.** *There holds*

$$\|\nabla(u_h^{\mathrm{ex}} - u_h)\| \le \|\boldsymbol{\sigma}_{h,\mathrm{alg}}\| =: \eta_{\mathrm{alg}} \qquad \textit{(algebraic error upper bound)}, \tag{5.2a}$$

$$\|\nabla(u - u_h)\| \le \|\boldsymbol{\sigma}_{h,\mathrm{tot}}\| + \underbrace{\left(\sum_{K\in\mathcal{T}_h} \frac{h_K^2}{\pi^2}\|f - \Pi_h f\|_K^2\right)^{\frac12}}_{\eta_{\mathrm{osc}}}$$

$$\le \underbrace{\|\nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}\|}_{\widetilde{\eta}_{\mathrm{dis}}} + \eta_{\mathrm{alg}} + \eta_{\mathrm{osc}} =: \eta \qquad \textit{(total error upper bound)}. \tag{5.2b}$$

Proof. Using that $(u_h^{\mathrm{ex}} - u_h) \in V_h$,

$$\|\nabla(u_h^{\mathrm{ex}} - u_h)\| = \sup_{v_h\in V_h,\, \|\nabla v_h\|=1} (\nabla(u_h^{\mathrm{ex}} - u_h), \nabla v_h). \tag{5.3}$$

Fix $v_h \in V_h$ with $\|\nabla v_h\| = 1$. The exact finite element solution definition (2.3), the residual equation (2.8), and the fact that $\nabla\cdot\boldsymbol{\sigma}_{h,\mathrm{alg}} = r_h$ by Lemma 6.4 below give

$$(\nabla(u_h^{\mathrm{ex}} - u_h), \nabla v_h) = (f, v_h) - (\nabla u_h, \nabla v_h) = (r_h, v_h) = (\nabla\cdot\boldsymbol{\sigma}_{h,\mathrm{alg}}, v_h) = -(\boldsymbol{\sigma}_{h,\mathrm{alg}}, \nabla v_h),$$

where we have also employed the Green theorem as $\boldsymbol{\sigma}_{h,\mathrm{alg}} \in \mathbf{H}(\mathrm{div},\Omega)$ and $v_h \in H^1_0(\Omega)$. Thus the Cauchy–Schwarz inequality proves (5.2a).

Similarly, as $(u - u_h) \in H^1_0(\Omega)$, there holds $\|\nabla(u - u_h)\| = \sup_{v\in H^1_0(\Omega),\, \|\nabla v\|=1}(\nabla(u - u_h), \nabla v)$. Fix $v \in H^1_0(\Omega)$ with $\|\nabla v\| = 1$. Employing (2.2), adding and subtracting $(\boldsymbol{\sigma}_{h,\mathrm{alg}} + \boldsymbol{\sigma}_{h,\mathrm{dis}}, \nabla v)$, and using the Green theorem, we get

$$(\nabla(u - u_h), \nabla v) = (f, v) - (\nabla u_h, \nabla v) = (f - \nabla\cdot(\boldsymbol{\sigma}_{h,\mathrm{alg}} + \boldsymbol{\sigma}_{h,\mathrm{dis}}), v) - (\boldsymbol{\sigma}_{h,\mathrm{alg}} + \boldsymbol{\sigma}_{h,\mathrm{dis}} + \nabla u_h, \nabla v). \tag{5.4}$$

Recall the Poincaré inequality

$$\|v - \Pi_{\mathbb{P}_0(K)} v\|_K \le \frac{h_K}{\pi}\|\nabla v\|_K \qquad \forall K \in \mathcal{T}_h,$$

see [43, 8]. Employing also Lemma 5.1 and the Cauchy–Schwarz inequality, the first term of right-hand side of (5.4) can be estimated via

$$(f - \nabla\cdot(\boldsymbol{\sigma}_{h,\mathrm{alg}} + \boldsymbol{\sigma}_{h,\mathrm{dis}}), v) = (f - \Pi_h f, v) = \sum_{K\in\mathcal{T}_h} (f - \Pi_h f, v - \Pi_{Q_h^0} v)_K \le \eta_{\mathrm{osc}}.$$

Relation (4.4) and the Cauchy–Schwarz and triangle inequalities for the second term then prove (5.2b). □

Relying on $H^1_0(\Omega)$-*liftings* of Concepts 3.2 and 4.2, we obtain *lower bounds* on the algebraic and total errors:

Theorem 5.3 (Guaranteed lower bounds on algebraic and total errors)**.** *There holds*

$$\|\nabla(u_h^{\mathrm{ex}} - u_h)\| \geq \frac{(r_h, \rho_{h,\mathrm{alg}})}{\|\nabla \rho_{h,\mathrm{alg}}\|} =: \underline{\eta}_{\mathrm{alg}} \qquad \textit{(algebraic error lower bound)}, \tag{5.5a}$$

$$\|\nabla(u - u_h)\| \geq \frac{\sum_{\mathbf{a}\in\mathcal{V}_h} \|\nabla \rho^{\mathbf{a}}_{h,\mathrm{tot}}\|^2_{\omega_h^{\mathbf{a}}}}{\|\nabla \rho_{h,\mathrm{tot}}\|} =: \underline{\eta} \qquad \textit{(total error lower bound)}. \tag{5.5b}$$

Proof. To prove the lower bound (5.5a) on the algebraic error, we use the norm characterization (5.3), the exact finite element solution definition (2.3), the fact that $\rho_{h,\mathrm{alg}} \in V_h$, and the algebraic residual representer r_h characterization (2.8) to see that

$$\|\nabla(u_h^{\mathrm{ex}} - u_h)\| \geq \frac{(f, \rho_{h,\mathrm{alg}}) - (\nabla u_h, \nabla \rho_{h,\mathrm{alg}})}{\|\nabla \rho_{h,\mathrm{alg}}\|} = \frac{(r_h, \rho_{h,\mathrm{alg}})}{\|\nabla \rho_{h,\mathrm{alg}}\|}.$$

The lower bound (5.5b) follows as in [41, Theorem 2]. □

Remark 5.4 (Lower bound on the total error)**.** *As shown in [41, Remark 2], there holds*

$$\underline{\eta} \geq \frac{\left(\sum_{\mathbf{a}\in\mathcal{V}_h} \|\nabla \rho^{\mathbf{a}}_{h,\mathrm{tot}}\|^2_{\omega_h^{\mathbf{a}}}\right)^{\frac{1}{2}}}{C_{\mathrm{cont,PF}}\sqrt{d+1}},$$

where $C_{\mathrm{cont,PF}}$ is the constant from the inequality (5.11) below. This lower bound on the total error has an ℓ^2 structure with respect to the patches around mesh vertices (can be written as a square root of a sum of squares) but is less precise than $\underline{\eta}$.

We will finally estimate the discretization error $\|\nabla(u - u_h^{\mathrm{ex}})\|$, proceeding as in [41, Sec. 6]. We point out that neither u nor u_h^{ex} is known, so that this is the hardest estimate. The Galerkin orthogonality of u_h^{ex}, though, guarantees

$$\|\nabla(u - u_h)\|^2 = \|\nabla(u - u_h^{\mathrm{ex}})\|^2 + \|\nabla(u_h^{\mathrm{ex}} - u_h)\|^2,$$

so that we can employ the bounds of Theorems 5.2 and 5.3 to get:

Corollary 5.5 (Guaranteed bounds on the discretization error)**.** *There holds*

$$\|\nabla(u - u_h^{\mathrm{ex}})\| \leq \sqrt{\eta^2 - \underline{\eta}^2_{\mathrm{alg}}} =: \eta_{\mathrm{dis}} \qquad \textit{(discretization error upper bound)}. \tag{5.6}$$

If $\underline{\eta} \geq \eta_{\mathrm{alg}}$, there also holds

$$\|\nabla(u - u_h^{\mathrm{ex}})\| \geq \sqrt{\underline{\eta}^2 - \eta^2_{\mathrm{alg}}} =: \underline{\eta}_{\mathrm{dis}} \qquad \textit{(discretization error lower bound)}. \tag{5.7}$$

5.2 Safe global and local stopping criteria

We propose here adaptive stopping criteria, balancing (locally) the algebraic and discretization error components. Apart from the estimators η_{dis} and $\underline{\eta}_{\mathrm{dis}}$ of Corollary 5.5, all the global estimators η_* can be expressed as the sum of their elementwise contributions $\eta_{*,K}$, i.e., $\eta_* = \left\{\sum_{K\in\mathcal{T}_h} (\eta_{*,K})^2\right\}^{1/2}$, the notation that we employ henceforth. We also denote $\eta_{\mathrm{alg},\omega_h^{\mathbf{a}}} := \|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|_{\omega_h^{\mathbf{a}}}$, $\mathbf{a} \in \mathcal{V}_h$.

The second estimate in (5.2b) distinguishes the *algebraic* and *discretization* error *indicators*. It is thus straightforward, see, e.g., [35, 10, 33, 2, 23, 41] and the references therein, to propose the following stopping criteria. Let $\gamma > 0$ and $\gamma_K > 0$, $\forall K \in \mathcal{T}_h$, be fixed parameters, typically of order 0.1:

$$\eta_{\mathrm{alg}} \leq \gamma(\widetilde{\eta}_{\mathrm{dis}} + \eta_{\mathrm{osc}}) \qquad \text{(global stopping criterion)}, \tag{5.8a}$$

$$\eta_{\mathrm{alg},\omega_h^{\mathbf{a}}} \leq \min_{K\subset\omega_h^{\mathbf{a}}} \gamma_K(\widetilde{\eta}_{\mathrm{dis},K} + \eta_{\mathrm{osc},K}), \quad \forall \mathbf{a} \in \mathcal{V}_h \qquad \text{(local stopping criterion)}. \tag{5.8b}$$

The above criteria only rely on the ingredients of Theorem 5.2, where, in particular, $\widetilde{\eta}_{\mathrm{dis}}$ is not proven to provide a lower bound on the discretization error. Thus criterion (5.8a) does not guarantee that the algebraic error is by the factor of γ smaller than the discretization one, i.e., it is not guaranteed that

$$\|\nabla(u_h^{\mathrm{ex}} - u_h)\| \leq \gamma \|\nabla(u - u_h^{\mathrm{ex}})\|.$$

In order to ensure this desirable balance and to *avoid a possible too early stopping* of the iterative solver (see the discussion in [41, Sec. 6.3]), we thus also propose alternative, safe stopping criteria:

$$\eta_{\mathrm{alg}} \leq \gamma\, \underline{\eta}_{\mathrm{dis}} \qquad \text{(global safe stopping criterion);} \tag{5.9a}$$

$$\eta_{\mathrm{alg},\omega_h^{\mathbf{a}}} \leq \min_{K\in\omega_h^{\mathbf{a}}} \frac{\gamma_K}{(1+\gamma_K^2)^{\frac{1}{2}}} \frac{\|\nabla \rho_{h,\mathrm{tot}}^{\mathbf{a}}\|_{\omega_h^{\mathbf{a}}}}{C_{\mathrm{cont,PF}}}, \quad \forall \mathbf{a}\in\mathcal{V}_h \qquad \text{(local safe stopping criterion).} \tag{5.9b}$$

Criteria (5.8a) and (5.9a) are global criteria and consequently local effects might be missed. The *local criteria* (5.8b) and (5.9b) ask that patch by patch around the nodes of the finest mesh $\mathcal{T}_h$, the local algebraic estimators are small enough. They in particular allow to prove the *local efficiency* result of Theorem 5.6 below.

5.3 Efficiency and robustness of the upper total estimates

We state here the efficiency of the upper bound η on the total error of Theorem 5.2 under both criteria (5.8) and (5.9), proceeding as in [23, 24, 41]. The proof is technical and is postponed to Appendix B below.

Let

$$H^1_*(\omega_h^{\mathbf{a}}) := \{v\in H^1(\omega_h^{\mathbf{a}});\, (v,1)_{\omega_h^{\mathbf{a}}} = 0\}, \qquad \mathbf{a}\in\mathcal{V}_h^{\mathrm{int}}, \tag{5.10a}$$

$$H^1_*(\omega_h^{\mathbf{a}}) := \{v\in H^1(\omega_h^{\mathbf{a}});\, v = 0 \text{ on } \partial\omega_h^{\mathbf{a}}\cap\partial\Omega\}, \quad \mathbf{a}\in\mathcal{V}_h^{\mathrm{ext}}. \tag{5.10b}$$

Standard scaling arguments and the Poincaré–Friedrichs inequality yield

$$\|\nabla(\psi_h^{\mathbf{a}} v)\|_{\omega_h^{\mathbf{a}}} \leq C_{\mathrm{cont,PF}} \|\nabla v\|_{\omega_h^{\mathbf{a}}} \qquad \forall v\in H^1_*(\omega_h^{\mathbf{a}}), \tag{5.11}$$

where the constant $C_{\mathrm{cont,PF}}$ only depends on the shape regularity of the mesh $\mathcal{T}_h$, see, e.g., [18, Theorem 3.1], [11, Sec. 3], or [24, proof of Lemma 3.12]. A crucial result that we rely on is the stability of the mixed finite element problem (7.1) in the sense that

$$\|\psi_h^{\mathbf{a}}\nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}\|_{\omega_h^{\mathbf{a}}} \leq C_{\mathrm{st}} \sup_{v\in H^1_*(\omega_h^{\mathbf{a}});\, \|\nabla v\|_{\omega_h^{\mathbf{a}}}=1} \big\{(f-r_h, \psi_h^{\mathbf{a}} v)_{\omega_h^{\mathbf{a}}} - (\nabla u_h, \nabla(\psi_h^{\mathbf{a}} v))_{\omega_h^{\mathbf{a}}}\big\}, \tag{5.12}$$

for $(f-r_h)\psi_h^{\mathbf{a}} \in Q_h$. This result has been proven, in two space dimensions, in Braess *et al.* [11, Theorem 7], and extension to three space dimensions is given in [25, Corollaries 3.3 and 3.6]. Importantly, the constant $C_{\mathrm{st}} > 0$ only depends on the shape regularity of the mesh $\mathcal{T}_h$; a computable upper bound on C_{st} is given in [24, Corollary 3.16 and Lemma 3.23].

Our first result establishes a local converse inequality to (5.2b), stating that our estimators on the total error are elementwise below the total error when the local stopping criteria (5.8b) or (5.9b) are satisfied:

Theorem 5.6 (Local efficiency)**.** *Let $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ and $\boldsymbol{\sigma}_{h,\mathrm{dis}}$ of Concepts 3.1 and 4.1 be of polynomial degree $q = p+1$. Let either the local stopping criterion (5.8b) be satisfied with parameters γ_K such that*

$$\gamma_K \leq \frac{1}{2(d+1)C_{\mathrm{st}}C_{\mathrm{cont,PF}}} \qquad \forall K\in\mathcal{T}_h, \tag{5.13}$$

or let (5.9b) hold, without any requirement on γ_K. Let finally f be a piecewise polynomial, $f\in\mathbb{P}_p(\mathcal{T}_h)$. Then

$$\widetilde{\eta}_{\mathrm{dis},K} + \eta_{\mathrm{alg},K} \leq 2(1+\gamma_K)C_{\mathrm{st}}C_{\mathrm{cont,PF}} \sum_{\mathbf{a}\in\mathcal{V}_h,\, \mathbf{a}\subset\partial K} \|\nabla(u-u_h)\|_{\omega_h^{\mathbf{a}}} \qquad \forall K\in\mathcal{T}_h. \tag{5.14}$$

The global lower counterpart to (5.2b), assessing the maximal theoretical overestimation of the total error $\|\nabla(u-u_h)\|$ by the total estimator η of Theorem 5.2 under the global stopping criteria (5.8a) or (5.9a), is:

Theorem 5.7 (Global efficiency)**.** *Let $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ and $\boldsymbol{\sigma}_{h,\mathrm{dis}}$ of Concepts 3.1 and 4.1 be of polynomial degree $q = p+1$. Let either the global stopping criterion (5.8a) be satisfied with parameter γ such that*

$$\gamma \leq \frac{1}{2(d+1)C_{\mathrm{st}}C_{\mathrm{cont,PF}}}, \tag{5.15}$$

or let (5.9a) hold, without any requirement on γ. Let finally f be a piecewise polynomial, $f\in\mathbb{P}_p(\mathcal{T}_h)$. Then

$$\widetilde{\eta}_{\mathrm{dis}} + \eta_{\mathrm{alg}} \leq 2(1+\gamma)(d+1)C_{\mathrm{st}}C_{\mathrm{cont,PF}}\|\nabla(u-u_h)\|.$$

Theorems 5.6 and 5.7 only concern the upper bound on the total error.

Remark 5.8 (Efficiency and robustness of the upper algebraic error estimator and quality of the lower bounds)**.** *We remark that we do not presently have a proof that* $\eta_{\mathrm{alg}} \leq C\|\nabla(u_h^{\mathrm{ex}} - u_h)\|$ *also holds, in addition to* $\|\nabla(u_h^{\mathrm{ex}} - u_h)\| \leq \eta_{\mathrm{alg}}$ *of (5.2a), which would be the efficiency of our upper estimate on the algebraic error. We also do not dispose of any efficiency results of the lower bounds* $\underline{\eta}$ *and* $\underline{\eta}_{\mathrm{alg}}$*. Numerical experiments in Sec. 9, though, show a very satisfactory behavior.*

6 Construction of the $\mathbf{H}(\mathrm{div},\Omega)$- and $H_0^1(\Omega)$-liftings of the algebraic residual

This section provides all details for the construction of the algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-lifting $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ of Concept 3.1, and of the algebraic residual $H_0^1(\Omega)$-lifting $\rho_{h,\mathrm{alg}}$ of Concept 3.2.

6.1 Algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-lifting

We first present the construction of the *algebraic residual lifting* $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ of Concept 3.1. These are obtained by dual mixed finite element approximation of problem inspired by (3.2b). It proceeds on patches around vertices from $\mathcal{V}_{j-1}$, with finer meshes induced by $\mathcal{T}_j$, $1 \leq j \leq J$, and with homogeneous Neumann boundary conditions. We first detail the two-level setting, and then generalize the construction to an arbitrary number of mesh levels.

6.1.1 Two-level algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-lifting

Let $J = 1$, with coarse mesh $\mathcal{T}_0$ and fine mesh $\mathcal{T}_h = \mathcal{T}_1$. Our algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-lifting $\boldsymbol{\sigma}_{h,\mathrm{alg}} \in \mathbf{V}_h$ is based on the mixed finite element problems on each *coarsest patch* $\omega_0^{\mathbf{a}}$, $\mathbf{a} \in \mathcal{V}_0$. The spaces $\mathbf{V}_{1,0}^{\mathbf{a}} \times Q_{1,0}^{\mathbf{a}}$ associated with fine mesh $\mathcal{T}_1$ are used, see (3.5) and Fig. 1, left, for an illustration. In particular, following (3.5), homogeneous Neumann boundary conditions are prescribed on $\partial\omega_0^{\mathbf{a}}$ for interior vertices $\mathbf{a} \in \mathcal{V}_0^{\mathrm{int}}$ and on $\partial\omega_0^{\mathbf{a}} \setminus \partial\Omega$ for boundary vertices $\mathbf{a} \in \mathcal{V}_0^{\mathrm{ext}}$; on $\partial\omega_0^{\mathbf{a}} \cap \partial\Omega$ a homogeneous Dirichlet boundary condition is imposed for $\mathbf{a} \in \mathcal{V}_0^{\mathrm{ext}}$. The actual construction reads:

Definition 6.1 (Two-level algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-lifting; $J = 1$)**.** *Define*

$$\boldsymbol{\sigma}_{h,\mathrm{alg}} := \sum_{\mathbf{a}\in\mathcal{V}_0} \boldsymbol{\sigma}_{1,\mathrm{alg}}^{\mathbf{a}}, \tag{6.1}$$

where the vertex contributions are defined as the solution of: find $(\boldsymbol{\sigma}_{1,\mathrm{alg}}^{\mathbf{a}}, \gamma_1^{\mathbf{a}}) \in \mathbf{V}_{1,0}^{\mathbf{a}} \times Q_{1,0}^{\mathbf{a}}$ *such that*

$$(\boldsymbol{\sigma}_{1,\mathrm{alg}}^{\mathbf{a}}, \mathbf{v}_1)_{\omega_0^{\mathbf{a}}} - (\gamma_1^{\mathbf{a}}, \nabla\cdot\mathbf{v}_1)_{\omega_0^{\mathbf{a}}} = 0 \qquad \forall \mathbf{v}_1 \in \mathbf{V}_{1,0}^{\mathbf{a}}, \tag{6.2a}$$

$$(\nabla\cdot\boldsymbol{\sigma}_{1,\mathrm{alg}}^{\mathbf{a}}, q_1)_{\omega_0^{\mathbf{a}}} = (r_h\psi_0^{\mathbf{a}} - \nabla\rho_{0,\mathrm{alg}}\cdot\nabla\psi_0^{\mathbf{a}}, q_1)_{\omega_0^{\mathbf{a}}} \qquad \forall q_1 \in Q_{1,0}^{\mathbf{a}}. \tag{6.2b}$$

Problems (6.2) are well-posed, see, e.g., Brezzi and Fortin [17] or [23, 24] and the references therein. In particular, the Neumann compatibility condition

$$(r_h\psi_0^{\mathbf{a}} - \nabla\rho_{0,\mathrm{alg}} \cdot \nabla\psi_0^{\mathbf{a}}, 1)_{\omega_0^{\mathbf{a}}} = 0 \tag{6.3}$$

for all coarsest interior vertices $\mathbf{a} \in \mathcal{V}_0^{\mathrm{int}}$ directly follows from (3.1). Note also that (6.2) are the Euler–Lagrange conditions for the constrained minimization problem

$$\boldsymbol{\sigma}_{1,\mathrm{alg}}^{\mathbf{a}} := \arg \min_{\mathbf{v}_1\in\mathbf{V}_{1,0}^{\mathbf{a}},\, \nabla\cdot\mathbf{v}_1 = \Pi_{Q_{1,0}^{\mathbf{a}}}(r_h\psi_0^{\mathbf{a}} - \nabla\rho_{0,\mathrm{alg}}\cdot\nabla\psi_0^{\mathbf{a}})} \|\mathbf{v}_1\|_{\omega_0^{\mathbf{a}}}. \tag{6.4}$$

It is actually the $L^2(\Omega)$ norm of the algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-lifting $\|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|$ that bounds the algebraic error, see (5.2a). Thus, our construction finds the best local contributions $\boldsymbol{\sigma}_{1,\mathrm{alg}}^{\mathbf{a}}$ which lead to the satisfaction of the following crucial divergence constraint:

Lemma 6.2 (Properties of $\boldsymbol{\sigma}_{h,\mathrm{alg}}$; $J = 1$)**.** *The algebraic residual lifting of Definition 6.1 satisfies* $\boldsymbol{\sigma}_{h,\mathrm{alg}} \in \mathbf{V}_h$ *and*

$$\nabla\cdot\boldsymbol{\sigma}_{h,\mathrm{alg}} = r_h.$$

Proof. All contributions $\boldsymbol{\sigma}^{\mathbf{a}}_{1,\mathrm{alg}}$ extended by zero outside of the patches $\omega^{\mathbf{a}}_0$ belong to $\mathbf{V}_h$, so that $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ does so by (6.1) as well. We next use the local divergence constraints (6.2b) in combination with the Neumann compatibility condition (6.3) which gives $\nabla{\cdot}\boldsymbol{\sigma}^{\mathbf{a}}_{1,\mathrm{alg}} = \Pi_h(r_h\psi^{\mathbf{a}}_0 - \nabla\rho_{0,\mathrm{alg}}{\cdot}\nabla\psi^{\mathbf{a}}_0)$. Consequently, the fact that $\psi^{\mathbf{a}}_0$, $\mathbf{a}\in\mathcal{V}_0$, form a partition of unity and the definition (6.1) lead to

$$\begin{aligned}\nabla{\cdot}\boldsymbol{\sigma}_{h,\mathrm{alg}} &= \sum_{\mathbf{a}\in\mathcal{V}_0}\nabla{\cdot}\boldsymbol{\sigma}^{\mathbf{a}}_{1,\mathrm{alg}} = \sum_{\mathbf{a}\in\mathcal{V}_0}\Pi_h(r_h\psi^{\mathbf{a}}_0 - \nabla\rho_{0,\mathrm{alg}}{\cdot}\nabla\psi^{\mathbf{a}}_0)\\ &= \Pi_h\left(\sum_{\mathbf{a}\in\mathcal{V}_0}(r_h\psi^{\mathbf{a}}_0 - \nabla\rho_{0,\mathrm{alg}}{\cdot}\nabla\psi^{\mathbf{a}}_0)\right) = \Pi_h r_h = r_h,\end{aligned}$$

which finishes the proof. □

6.1.2 Multilevel hierarchical algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-lifting

We now extend Definition 6.1 to an arbitrary number of mesh levels. To do so, we exploit the *hierarchical* mesh structure and use a telescopic sum argument:

Definition 6.3 (Algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-lifting)**.** *Define*

$$\boldsymbol{\sigma}_{h,\mathrm{alg}} := \sum_{j=1}^{J}\sum_{\mathbf{a}\in\mathcal{V}_{j-1}}\boldsymbol{\sigma}^{\mathbf{a}}_{j,\mathrm{alg}}, \tag{6.5}$$

where the vertex contributions $\boldsymbol{\sigma}^{\mathbf{a}}_{j,\mathrm{alg}}$ on level j, $j=1,\dots,J$, are defined as the solution of: find $(\boldsymbol{\sigma}^{\mathbf{a}}_{j,\mathrm{alg}},\gamma^{\mathbf{a}}_j)\in\mathbf{V}^{\mathbf{a}}_{j,j-1}\times Q^{\mathbf{a}}_{j,j-1}$ such that

$$(\boldsymbol{\sigma}^{\mathbf{a}}_{j,\mathrm{alg}},\mathbf{v}_j)_{\omega^{\mathbf{a}}_{j-1}} - (\gamma^{\mathbf{a}}_j,\nabla{\cdot}\mathbf{v}_j)_{\omega^{\mathbf{a}}_{j-1}} = 0 \qquad \forall\mathbf{v}_j\in\mathbf{V}^{\mathbf{a}}_{j,j-1}, \tag{6.6a}$$

$$(\nabla{\cdot}\boldsymbol{\sigma}^{\mathbf{a}}_{j,\mathrm{alg}},q_j)_{\omega^{\mathbf{a}}_{j-1}} = ((\mathrm{Id}-\Pi_{j-1})(r_h\psi^{\mathbf{a}}_{j-1} - \nabla\rho_{0,\mathrm{alg}}{\cdot}\nabla\psi^{\mathbf{a}}_{j-1}),q_j)_{\omega^{\mathbf{a}}_{j-1}} \qquad \forall q_j\in Q^{\mathbf{a}}_{j,j-1}; \tag{6.6b}$$

here, with an abuse of notation, we redefine $\Pi_0 := 0$.

Problems (6.6) are well-posed for all levels $1\le j\le J$ and all vertices $\mathbf{a}\in\mathcal{V}_{j-1}$. In particular, for $j=1$, they coincide with those of Definition 6.1. The Neumann compatibility conditions for $\mathbf{a}\in\mathcal{V}^{\mathrm{int}}_{j-1}$, $j>1$, follow from $((\mathrm{Id}-\Pi_{j-1})(r_h\psi^{\mathbf{a}}_{j-1} - \nabla\rho_{0,\mathrm{alg}}{\cdot}\nabla\psi^{\mathbf{a}}_{j-1}),1)_{\omega^{\mathbf{a}}_{j-1}} = 0$. In generalization of Lemma 6.2, we now have:

Lemma 6.4 (Properties of $\boldsymbol{\sigma}_{h,\mathrm{alg}}$)**.** *The algebraic residual lifting of Definition 6.3 satisfies $\boldsymbol{\sigma}_{h,\mathrm{alg}}\in\mathbf{V}_h$ and*

$$\nabla{\cdot}\boldsymbol{\sigma}_{h,\mathrm{alg}} = r_h.$$

Proof. The fact that $\boldsymbol{\sigma}_{h,\mathrm{alg}}\in\mathbf{V}_h\subset\mathbf{H}(\mathrm{div},\Omega)$ is immediate as all the summands $\boldsymbol{\sigma}^{\mathbf{a}}_{j,\mathrm{alg}}$ in (6.5) (when extended by zero outside of $\omega^{\mathbf{a}}_{j-1}$) belong to $\mathbf{V}_h\subset\mathbf{H}(\mathrm{div},\Omega)$. As for the divergence, using (6.5), the fact that the hat functions $\psi^{\mathbf{a}}_{j-1}$ for vertices $\mathbf{a}\in\mathcal{V}_{j-1}$ on level $j-1$ form a partition of unity on the entire domain Ω, the fact that $\Pi_j(\mathrm{Id}-\Pi_{j-1}) = \Pi_j - \Pi_{j-1}$, the local divergence constraint (6.6b), and since $\Pi_0 = 0$, we obtain

$$\begin{aligned}\nabla{\cdot}\boldsymbol{\sigma}_{h,\mathrm{alg}} &= \sum_{j=1}^{J}\sum_{\mathbf{a}\in\mathcal{V}_{j-1}}\nabla{\cdot}\boldsymbol{\sigma}^{\mathbf{a}}_{j,\mathrm{alg}} = \sum_{j=1}^{J}\sum_{\mathbf{a}\in\mathcal{V}_{j-1}}(\Pi_j-\Pi_{j-1})(r_h\psi^{\mathbf{a}}_{j-1} - \nabla\rho_{0,\mathrm{alg}}{\cdot}\nabla\psi^{\mathbf{a}}_{j-1})\\ &= \sum_{j=1}^{J}(\Pi_j-\Pi_{j-1})r_h = \Pi_h r_h - \Pi_0 r_h = r_h.\end{aligned}$$

□

6.2 Algebraic residual $H^1_0(\Omega)$-lifting

We now pass to the *algebraic residual lifting* $\rho_{h,\mathrm{alg}}$ of Concept 3.2. The construction below is motivated by problems (3.2b) and uses local primal conforming finite element approximations as a counterpart to the dual mixed problems (6.6). It proceeds on patches around vertices from $\mathcal{V}_{j-1}$, with finer meshes induced by $\mathcal{T}_j$,

$1 \le j \le J$, and with homogeneous Dirichlet boundary conditions imposed. It can be equivalently regarded as setting $\rho_{h,\mathrm{alg}} := u_h^{\mathrm{new}} - u_h$, where u_h^{new} has been computed from u_h as an improvement by one step of a (multigrid) solver. Recall that $(r_h, v_j)_{\omega_{j-1}^{\mathbf{a}}} = (f, v_j)_{\omega_{j-1}^{\mathbf{a}}} - (\nabla u_h, \nabla v_j)_{\omega_{j-1}^{\mathbf{a}}}$ by (2.8); using spaces (3.6), see Fig. 1, right, we define:

Definition 6.5 (Algebraic residual $H_0^1(\Omega)$-lifting)**.** *Define*

$$\rho_{h,\mathrm{alg}} := \rho_{0,\mathrm{alg}} + \sum_{j=1}^{J} \rho_{j,\mathrm{alg}}, \tag{6.7}$$

where $\rho_{0,\mathrm{alg}}$ solves (3.1) and $\rho_{j,\mathrm{alg}} := \sum_{\mathbf{a}\in\mathcal{V}_{j-1}} w_{\psi_{j-1}^{\mathbf{a}}}\big(\rho_{j,\mathrm{alg}}^{\mathbf{a}}\big) \in V_j^p$, $1 \le j \le J$. Here the vertex contributions $\rho_{j,\mathrm{alg}}^{\mathbf{a}}$ on patches $\omega_{j-1}^{\mathbf{a}}$ are defined as the solution of: find $\rho_{j,\mathrm{alg}}^{\mathbf{a}} \in V_{j,j-1}^{\mathbf{a}}$ such that

$$(\nabla\rho_{j,\mathrm{alg}}^{\mathbf{a}}, \nabla v_j)_{\omega_{j-1}^{\mathbf{a}}} = (r_h, v_j)_{\omega_{j-1}^{\mathbf{a}}} - \sum_{i=0}^{j-1} (\nabla\rho_{i,\mathrm{alg}}, \nabla v_j)_{\omega_{j-1}^{\mathbf{a}}} \qquad \forall v_j \in V_{j,j-1}^{\mathbf{a}}, \tag{6.8}$$

and $w_{\psi_{j-1}^{\mathbf{a}}}(\cdot)$ stands for weighting a function $v_h \in V_{j,j-1}^{\mathbf{a}}$ by the hat function $\psi_{j-1}^{\mathbf{a}}$ as follows

$$w_{\psi_{j-1}^{\mathbf{a}}}(v_h) \in V_{j,j-1}^{\mathbf{a}}, \quad w_{\psi_{j-1}^{\mathbf{a}}}(v_h)(\mathbf{x}) = \psi_{j-1}^{\mathbf{a}}(\mathbf{x}) \cdot (v_h)(\mathbf{x}), \quad \mathbf{x}\ldots \textit{ Lagrangian points}^{3} \textit{ of } K \in \mathcal{T}_j, K \subset \omega_{j-1}^{\mathbf{a}}.$$

We note that the above construction of $\rho_{j,\mathrm{alg}}$ can be regarded as applying the restricted additive Schwarz (RAS) method (see, e.g., [21, Chapter 1]) to problems (3.2b) with the domain Ω decomposed into (overlapping) patches $\omega_{j-1}^{\mathbf{a}}$. Indeed, $w_{\psi_{j-1}^{\mathbf{a}}}\big(\rho_{j,\mathrm{alg}}^{\mathbf{a}}\big)$ corresponds to multiplication of the Lagrangian degrees of freedom by a discrete partition of unity induced by the values of the hat functions $\psi_{j-1}^{\mathbf{a}}$. The whole hierarchical construction of $\rho_{h,\mathrm{alg}}$ can then be interpreted as one step of multigrid $V(0,1)$-cycle with RAS smoothing on each level. Other options like $\rho_{j,\mathrm{alg}} := \sum_{\mathbf{a}\in\mathcal{V}_{j-1}} \rho_{j,\mathrm{alg}}^{\mathbf{a}} \in V_j^p$ (additive Schwarz) performed in our numerical experiments significantly worse.

7 Construction of the $\mathbf{H}(\mathrm{div},\Omega)$- and $H_0^1(\Omega)$-liftings of the total residual

We give here details on the discrete broken $\mathbf{H}(\mathrm{div},\Omega)$-lifting $\boldsymbol{\sigma}_{h,\mathrm{tot}}$ of Concept 4.1 and the discrete $H_0^1(\Omega)$-lifting $\rho_{h,\mathrm{tot}}$ of Concept 4.2. Both stem from the total residual $\mathcal{R}_h$ given by (2.9) and serve to estimate the total error. In contrast to Sec. 6, the constructions here are localized over patches of mesh elements on the finest mesh only.

7.1 Total residual broken $\mathbf{H}(\mathrm{div},\Omega)$-lifting

We first define a *discretization flux reconstruction* $\boldsymbol{\sigma}_{h,\mathrm{dis}} \in \mathbf{V}_h$. We proceed as in Sec. 6.1 via solution of local mixed finite element problems, but this time only on the *finest patches* $\omega_h^{\mathbf{a}}$ around the *finest mesh vertices* $\mathbf{a} \in \mathcal{V}_h$, similarly to [23, Definition 6.9] and [41, Sec. 4.4]:

Definition 7.1 (Discretization flux reconstruction)**.** *Define*

$$\boldsymbol{\sigma}_{h,\mathrm{dis}} := \sum_{\mathbf{a}\in\mathcal{V}_h} \boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}},$$

where the vertex contributions $\boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}$ are defined as the solution of: find $(\boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}, \gamma_h^{\mathbf{a}}) \in \mathbf{V}_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}$ such that

$$(\boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}, \mathbf{v}_h)_{\omega_h^{\mathbf{a}}} - (\gamma_h^{\mathbf{a}}, \nabla\cdot\mathbf{v}_h)_{\omega_h^{\mathbf{a}}} = -(\psi_h^{\mathbf{a}}\nabla u_h, \mathbf{v}_h)_{\omega_h^{\mathbf{a}}} \qquad \forall \mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}}, \tag{7.1a}$$

$$(\nabla\cdot\boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}, q_h)_{\omega_h^{\mathbf{a}}} = (f\psi_h^{\mathbf{a}} - \nabla u_h\cdot\nabla\psi_h^{\mathbf{a}} - r_h\psi_h^{\mathbf{a}}, q_h)_{\omega_h^{\mathbf{a}}} \qquad \forall q_h \in Q_h^{\mathbf{a}}. \tag{7.1b}$$

The Neumann compatibility condition for (7.1) amounts to

$$(f - r_h, \psi_h^{\mathbf{a}})_{\omega_h^{\mathbf{a}}} - (\nabla u_h, \nabla\psi_h^{\mathbf{a}})_{\omega_h^{\mathbf{a}}} = 0 \qquad \forall \mathbf{a} \in \mathcal{V}_h^{\mathrm{int}},$$

and is a direct consequence of (2.8). Similarly as in Lemmas 6.2 and 6.4, we can characterize the divergence of $\boldsymbol{\sigma}_{h,\mathrm{dis}}$ (cf. [23, Lemma 6.12] and [41, Sec. 4.4]):

Lemma 7.2 (Properties of $\boldsymbol{\sigma}_{h,\mathrm{dis}}$)**.** *The discretization flux reconstruction of Definition 7.1 satisfies $\boldsymbol{\sigma}_{h,\mathrm{dis}} \in \mathbf{V}_h$ and*

$$\nabla\cdot\boldsymbol{\sigma}_{h,\mathrm{dis}} = \Pi_h f - r_h.$$

[3] Or any other points in each triangle/tetrahedron that uniquely determine a function of $V_{j,j-1}^{\mathbf{a}}$, cf. [16, Sec. 3.2].

7.2 Total residual $H^1_0(\Omega)$-lifting

For bounding the total error from below, we follow some basic ideas in Bramble [13], Babuška and Strouboulis [4, Sec. 5.1], Repin [45, Sec. 4.1.1], Ern and Vohralík [24, Sec. 3.3]. Recall the definition of the spaces $H^1_*(\omega_h^{\mathbf{a}})$ from (5.10). Similarly to [41], the construction is:

Definition 7.3 (Total residual $H^1_0(\Omega)$-lifting)**.** *Define*

$$\rho_{h,\mathrm{tot}} := \sum_{\mathbf{a}\in\mathcal{V}_h} \psi_h^{\mathbf{a}} \rho_{h,\mathrm{tot}}^{\mathbf{a}},$$

where the vertex contributions $\rho_{h,\mathrm{tot}}^{\mathbf{a}} \in V_h^{\mathbf{a}}$ *solve*

$$(\nabla\rho_{h,\mathrm{tot}}^{\mathbf{a}}, \nabla v_h)_{\omega_h^{\mathbf{a}}} = (f, \psi_h^{\mathbf{a}} v_h)_{\omega_h^{\mathbf{a}}} - (\nabla u_h, \nabla(\psi_h^{\mathbf{a}} v_h))_{\omega_h^{\mathbf{a}}} \qquad \forall v_h \in V_h^{\mathbf{a}}\,. \tag{7.2}$$

The local problems (7.2) are associated with homogeneous Neumann data on $\partial\omega_h^{\mathbf{a}}$ for interior vertices $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$ and with homogeneous Dirichlet–Neumann boundary data for boundary vertices $\mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}$. In both cases, they are well-posed due to the definition (5.10) of the space $H^1_*(\omega_h^{\mathbf{a}})$, where the zero mean value condition is imposed. Even though $\rho_{h,\mathrm{tot}}^{\mathbf{a}}$ are in general nonzero on the patch boundary $\omega_h^{\mathbf{a}}$, the usage of the hat functions $\psi_h^{\mathbf{a}}$ in the definition of $\rho_{h,\mathrm{tot}}$ results in an $H^1_0(\Omega)$-conforming lifting of the total residual.

8 Implementation, evaluation cost, and equivalent explicit version without patchwise local problems

In this section, we comment on the cost of evaluating the error estimators and discuss how it can be reduced in comparison to a straightforward implementation of the presented definitions. We also show how to modify the constructions of Secs. 6–7 to avoid any need to solve patchwise linear systems, while still maintaining all the results of Secs. 5.1–5.3, in particular the guaranteed constant-free bounds, efficiency, and robustness. We restrict the presentation on how to evaluate the algebraic error upper bound $\|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|$ in Theorem 5.2; a similar approach, however, applies to $\|\boldsymbol{\sigma}_{h,\mathrm{tot}}\|$ or $\|\nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}\|$ of Theorem 5.2 and, in the $H^1_0(\Omega)$-setting, to $\underline{\eta}_{\mathrm{alg}}$ and $\underline{\eta}$ of Theorem 5.3 as well.

8.1 The coarsest level problem

In order to obtain our algebraic error estimators, we first need to find the solution $\rho_{0,\mathrm{alg}}$ of (3.1). The cost of the coarse grid solve, with the number of unknowns given by the dimension of the coarsest-level space V_0^p, can be considered as negligible, but obtaining the actual piecewise polynomial $\rho_{h,\mathrm{alg}}$ as a function on the finest mesh $\mathcal{T}_h$ brings the algorithmic price of problem (3.1) to the order of one geometric multigrid cycle without any smoothing. One important exception exists here: whenever the considered iterative solver itself is a geometric multigrid without any post-smoothing, any iterate u_h satisfies

$$(f, v_0) - (\nabla u_h, \nabla v_0) = (r_h, v_0) = 0 \qquad \forall v_0 \in V_0^p, \tag{8.1}$$

so that $\rho_{0,\mathrm{alg}} = 0$ and the coarse level solve (3.1) is not necessary at all. An example is described in more details below in Sec. 9.

8.2 Efficient implementation of the proposed estimators

Let $\boldsymbol{\psi}_h^l$, $1 \le l \le \dim(\mathbf{V}_h)$, denote basis functions of the Raviart–Thomas(–Nédélec) space $\mathbf{V}_h = \mathbf{V}_J^q$. Then developing $\boldsymbol{\sigma}_{h,\mathrm{alg}} = \sum_{l=1}^{\dim(\mathbf{V}_h)} (\mathrm{S}_h)_l \boldsymbol{\psi}_h^l$, the algebraic error estimator $\eta_{\mathrm{alg}} = \|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|$ from Theorem 5.2 takes the form

$$\|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|^2 = \mathrm{S}_h^t \mathbb{M}_h \mathrm{S}_h, \tag{8.2}$$

where the global mixed finite element matrix $\mathbb{M}_h$ is given by $(\mathbb{M}_h)_{lm} := (\boldsymbol{\psi}_h^m, \boldsymbol{\psi}_h^l)$. For the coefficients S_h, we see from (6.5) that

$$\mathrm{S}_h = \sum_{j=1}^{J} \sum_{\mathbf{a}\in\mathcal{V}_{j-1}} \mathbb{P}_j^J \mathrm{S}_j^{\mathbf{a}} \tag{8.3}$$

with an appropriate prolongation matrix $\mathbb{P}_j^J$. These two operations are of optimal complexity, though the passages over different mesh levels via the prolongation matrix $\mathbb{P}_j^J$ again bring the algorithmic price to the

order of one geometric multigrid cycle without any smoothing. The main question is the cost of obtaining the local coefficients $\mathrm{S}_j^{\mathbf{a}}$. Here, the key advantage of the problems (6.6) which define $\mathrm{S}_j^{\mathbf{a}}$ is that they are local to the patch $\omega_{j-1}^{\mathbf{a}}$ of all elements of $\mathcal{T}_{j-1}$ sharing the vertex $\mathbf{a} \in \mathcal{V}_{j-1}$, with a given *local matrix*, and independent one from each other, enabling a *perfectly scalable parallelization* of the local problems. Unfortunately, in a direct implementation of (6.6), presented in more details in [41, Appendix B], the size of the local matrices is not very small. To give a concrete example, consider the lowest order finite element case $p = 1$ in two space dimensions, with flux equilibration employing the polynomial degree $q = 1$, in the two-level setting $J = 1$. Then, for an interior vertex $\mathbf{a} \in \mathcal{V}_0$ connected to 6 other ones in the situation of Fig. 1 and with the spaces $\mathbf{V}_{1,0}^{\mathbf{a}}$ and $Q_{1,0}^{\mathbf{a}}$ of (3.5), the local matrix stemming from (6.6) has 179 degrees of freedom, see, e.g., [17, Sec. III.3.1]: 2 per interior face flux, 2 per element for the flux interior moments, and 3 per element for the Lagrange multiplier space $Q_{1,0}^{\mathbf{a}}$, all this minus one for the mean value constraint. Thus a straightforward solution of (6.6) via the LU-factorization will cost approximately $\frac{2}{3} \times 179^3 \approx 3\ 824\ 000$ FLOPS per vertex $\mathbf{a} \in \mathcal{V}_0$. This roughly corresponds to 956 000 FLOPS per one unknown with respect to the piecewise affine finite element discretization on the mesh $\mathcal{T}_1$. Though the operations for solving dense systems are generally cheaper on many modern computer systems, these numbers are much higher than the typical cost of 30 FLOPS per unknown to compute a second order approximation to the Laplace equation with a full multigrid algorithm [50]. It is therefore essential that more efficient implementations are considered. In particular, the above cost can be reduced radically by noticing that:

- Using an equivalent implementation of (6.6) via *hybridization* (static condensation) reduces the size of the local matrices (see, e.g., [17, Sec. V.2]). In particular, in the above example, hybridization brings the local matrix size from 179 down to 84, i.e., the solution cost from cca 956 000 to cca 99 000 FLOPS per unknown (there are two degrees of freedom per face here).
- The local matrices can be assembled and factorized only *once* in a preparatory step before the iterative solver applied to (2.4) is started. In each iteration, the estimators are then evaluated by forward-backward substitution, i.e., the biggest part of the local systems price is only to be paid once and not on each iteration.
- On a *structured mesh*, the local matrices associated with (6.6), as well as their equivalent hybridized formulations, are the *same for all interior vertices*. Also, on a typical locally refined unstructured triangular mesh created by newest-vertex bisection from an initial structured mesh, only a *handful of different patch geometries* and associated matrices exists, so that only a few local matrices need to be assembled and factorized.

The cost of obtaining $\boldsymbol{\sigma}_{h,\mathrm{dis}}$ is much smaller since:

- Only the finest mesh $\mathcal{T}_h = \mathcal{T}_J$ and its vertices from the set $\mathcal{V}_h = \mathcal{V}_J$ are employed, there is no run through all the mesh levels.
- The sizes of the local matrices are much smaller (42 and 24 instead of the above 179 and 84, respectively), as the spaces $\mathbf{V}_h^{\mathbf{a}}$ and $Q_h^{\mathbf{a}}$ of (4.1) associated with local problems (7.1) are defined on the patches of finest mesh elements only (there is no dashed submesh in the setting of Fig. 2 in comparison with Fig. 1).

Remark, however, that estimating the total error in Theorem 5.2 requires both $\boldsymbol{\sigma}_{h,\mathrm{dis}}$ and $\boldsymbol{\sigma}_{h,\mathrm{alg}}$.

8.3 Equivalent explicit version without local patchwise problems

We finally discuss a procedure to construct local $\mathbf{H}(\mathrm{div}, \Omega)$-residual liftings replacing $\boldsymbol{\sigma}_{j,\mathrm{alg}}^{\mathbf{a}}$ given by (6.6) without *any local linear solve on the patch of elements* (the same applies to $\boldsymbol{\sigma}_{h,\mathrm{dis}}$). This construction still provides guaranteed upper bounds on the total and algebraic errors as in Theorem 5.2, as well as efficiency and polynomial robustness as in Theorems 5.6 and 5.7.

The procedure, following [11, Theorem 7] and described in detail in [25, Sec. 6] for the construction of an approximation of $\boldsymbol{\sigma}_{h,\mathrm{dis}}$ and evaluating the discretization error in absence of the algebraic one, replaces the minimizer corresponding to finding the solution of (7.1) by a simplification obtained by a run through the patch of the elements sharing the given vertex. For each element, one is then left to determine the interior moments and the face moments of not previously visited interior faces. For $p = 1$ with flux equilibration using $q = 1$ in two space dimensions, this typically leaves, for each element in the patch, a linear system with 4 unknowns, whose solution can be found explicitly, or as in Sec. 8.2 by the LU-factorization with approximate cost of $\frac{2}{3} \times 4^3 + 2 \times 4^2 \approx 85$ FLOPS. The sweeps for different patches are independent,

p (N_h)	MG iter	algebraic error	eff. index UB	eff. index LB	total error	eff. index UB	eff. index LB	discr. error	eff. index UB	eff. index LB
1 (3.5×10^4)	1	1.2	1.11	1.03^{-1}	1.3	1.46	1.67^{-1}	2.4×10^{-1}	5.84	—
	2	8.0×10^{-2}	1.13	1.04^{-1}	2.5×10^{-1}	1.35	1.05^{-1}		1.39	1.08^{-1}
	3	5.1×10^{-3}	1.15	1.06^{-1}	2.4×10^{-1}	1.06	1.03^{-1}		1.06	1.03^{-1}
2 (1.4×10^5)	1	1.2	1.10	1.00^{-1}	1.2	1.48	1.46^{-1}	2.9×10^{-3}	4.50×10^2	—
	2	1.1×10^{-1}	1.18	1.00^{-1}	1.1×10^{-1}	1.77	1.07^{-1}		5.29×10^1	—
	3	2.8×10^{-3}	1.18	1.01^{-1}	4.1×10^{-3}	1.66	1.19^{-1}		2.10	4.32^{-1}
	4	9.6×10^{-5}	1.20	1.02^{-1}	2.9×10^{-3}	1.05	1.03^{-1}		1.05	1.03^{-1}
3 (3.2×10^5)	1	5.9×10^{-1}	1.09	1.00^{-1}	5.9×10^{-1}	1.33	1.69^{-1}	2.2×10^{-5}	2.36×10^4	—
	3	2.2×10^{-3}	1.19	1.00^{-1}	2.2×10^{-3}	1.75	1.07^{-1}		1.42×10^2	—
	5	1.0×10^{-5}	1.19	1.01^{-1}	2.4×10^{-5}	1.44	1.37^{-1}		1.52	1.70^{-1}
	6	1.4×10^{-6}	1.18	1.01^{-1}	2.2×10^{-5}	1.08	1.12^{-1}		1.08	1.12^{-1}
4 (5.6×10^5)	1	4.5×10^{-1}	1.08	1.00^{-1}	4.5×10^{-1}	1.39	1.48^{-1}	1.5×10^{-7}	2.90×10^6	—
	4	1.2×10^{-4}	1.13	1.00^{-1}	1.2×10^{-4}	1.58	1.10^{-1}		9.63×10^2	—
	7	6.4×10^{-7}	1.11	1.01^{-1}	6.5×10^{-7}	1.53	1.14^{-1}		5.20	—
	10	8.3×10^{-9}	1.12	1.01^{-1}	1.5×10^{-7}	1.06	1.16^{-1}		1.06	1.16^{-1}

Table 1: Sinus problem, multigrid V-cycles: effectivity of the upper (UB) and lower (LB) error bounds

p (N_h)	algebraic error	eff. index UB	eff. index LB	total error	eff. index UB	eff. index LB	discretization error	eff. index UB	eff. index LB
1 (3.5×10^4)	8.7×10^{-4}	1.03	1.01^{-1}	2.4×10^{-1}	1.04	1.03^{-1}	2.4×10^{-1}	1.04	1.03^{-1}
2 (1.4×10^5)	1.9×10^{-5}	1.10	1.01^{-1}	2.9×10^{-3}	1.02	1.02^{-1}	2.9×10^{-3}	1.02	1.02^{-1}
3 (3.2×10^5)	2.6×10^{-6}	1.08	1.00^{-1}	2.2×10^{-5}	1.12	1.13^{-1}	2.2×10^{-5}	1.12	1.14^{-1}
4 (5.6×10^5)	5.1×10^{-8}	1.03	1.00^{-1}	1.6×10^{-7}	1.29	1.38^{-1}	1.5×10^{-7}	1.32	1.47^{-1}

Table 2: Sinus problem, one full multigrid cycle: effectivity of the upper (UB) and lower (LB) error bounds

and therefore the procedure is completely parallelizable. In the lowest-order case $q = 0$, similar explicit procedures were in the past proposed, e.g., in [20, 34, 33, 53, 30]; in particular, following [33, Sec. 7.3], the present construction can be extended to the patches of the form of Fig. 1 for the construction of an approximation of $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ which avoids the solution of the local systems (6.6). A numerical comparison of [53, Sec. 7] in particular witnesses an almost indistinguishable accuracy of the bounds constructed without any local patch solve, with a drastically reduced cost.

One can similarly approximate explicitly the linear systems (6.8) and (7.2), with in particular Theorem 5.3 still holding true. Indeed, no specific property of the type (5.12) is in the present analysis needed for the $H^1_0(\Omega)$-liftings $\rho_{h,\mathrm{alg}}$ and $\rho_{h,\mathrm{tot}}$.

Finally, for the price of introduction of (known) generic constants, yet much cheaper variant of the derived estimates is possible, which avoids any construction of the various liftings. This variant ultimately takes the form of the well-known stable-splittings estimates and will be treated in a forthcoming contribution.

9 Numerical results

We illustrate in this section the performance of our a posteriori error estimates and of the corresponding adaptive stopping criteria. We consider the model problem (2.1) with three different choices of the domain $\Omega \in \mathbb{R}^2$ and of the exact solution u:

$$\begin{aligned}
&\Omega := (-1,1)^2, && u(x,y) := \sin(2\pi x)\sin(2\pi y) && \text{sinus},\\
&\Omega := (0,1)^2, && u(x,y) := x(x-1)y(y-1)\exp\big(-100((x-\tfrac{1}{2})^2-(y-\tfrac{117}{1000})^2)\big) && \text{peak},\\
&\Omega := (-1,1)^2 \setminus [0,1]\times[-1,0], && u(r,\theta) := r^{2/3}\sin(2\theta/3) && \text{L-shape}.
\end{aligned}$$

In the last case, we impose an inhomogeneous Dirichlet boundary condition corresponding to the prescribed exact solution. The additional boundary approximation error is neglected since it is of higher order with respect to the $H^{1/2}(\partial\Omega)$-norm compared to the $H^1(\Omega)$-seminorm on the domain.

We consider the finite element method (2.3) with the polynomial degrees $p = 1, \dots, 4$. For each test problem, we start from an initial Delaunay triangulation of the domain Ω and consider four uniform refinement

p (N_h)	PCG iter	algebraic error	eff. index UB	eff. index LB	total error	eff. index UB	eff. index LB	discretization error	eff. index UB	eff. index LB
1 (3.5×10^4)	3	6.4×10^{-2}	1.01	1.00^{-1}	2.5×10^{-1}	1.26	1.06^{-1}	2.4×10^{-1}	1.28	1.07^{-1}
	6	1.3×10^{-2}	1.01	1.00^{-1}	2.4×10^{-1}	1.09	1.03^{-1}		1.09	1.03^{-1}
2 (1.4×10^5)	5	5.2×10^{-2}	1.01	1.00^{-1}	5.2×10^{-2}	1.09	8.71^{-1}	2.9×10^{-3}	7.88	—
	10	5.6×10^{-3}	1.00	1.00^{-1}	6.3×10^{-3}	1.36	2.31^{-1}		2.23	—
	15	1.2×10^{-4}	1.01	1.00^{-1}	2.9×10^{-3}	1.05	1.02^{-1}		1.05	1.02^{-1}
3 (3.2×10^5)	11	4.2×10^{-2}	1.00	1.00^{-1}	4.2×10^{-2}	1.05	11.2^{-1}	2.2×10^{-5}	6.11×10^2	—
	22	6.3×10^{-3}	1.00	1.00^{-1}	6.3×10^{-3}	1.05	9.80^{-1}		9.46×10^1	—
	33	6.7×10^{-5}	1.01	1.00^{-1}	7.0×10^{-5}	1.29	3.85^{-1}		2.76	—
	44	8.0×10^{-7}	1.00	1.00^{-1}	2.2×10^{-5}	1.04	1.10^{-1}		1.04	1.10^{-1}
4 (5.6×10^5)	15	2.6×10^{-2}	1.00	1.00^{-1}	2.6×10^{-2}	1.05	11.5^{-1}	1.5×10^{-7}	5.52×10^4	—
	30	7.0×10^{-4}	1.01	1.00^{-1}	7.0×10^{-4}	1.08	7.99^{-1}		1.88×10^3	—
	45	9.1×10^{-7}	1.00	1.00^{-1}	9.3×10^{-7}	1.16	6.98^{-1}		3.80	—
	60	2.5×10^{-9}	1.01	1.00^{-1}	1.5×10^{-7}	1.02	1.15^{-1}		1.02	1.15^{-1}

Table 3: Sinus problem, PCG iterations: effectivity of the upper (UB) and lower (LB) error bounds

p (N_h)	MG iter	algebraic error	eff. index UB	eff. index LB	total error	eff. index UB	eff. index LB	discretization error	eff. index UB	eff. index LB
1 (9.3×10^3)	1	6.1×10^{-3}	1.13	1.02^{-1}	6.9×10^{-3}	1.61	1.21^{-1}	3.3×10^{-3}	2.84	—
	2	1.9×10^{-4}	1.13	1.03^{-1}	3.3×10^{-3}	1.10	1.03^{-1}		1.10	1.03^{-1}
2 (3.8×10^4)	1	7.5×10^{-3}	1.13	1.00^{-1}	7.5×10^{-3}	1.61	1.23^{-1}	1.1×10^{-4}	8.53×10^1	—
	2	4.5×10^{-4}	1.17	1.01^{-1}	4.6×10^{-4}	1.76	1.06^{-1}		6.13	—
	3	8.1×10^{-6}	1.17	1.01^{-1}	1.1×10^{-4}	1.10	1.03^{-1}		1.10	1.03^{-1}
3 (8.5×10^4)	1	4.9×10^{-3}	1.10	1.00^{-1}	4.9×10^{-3}	1.40	1.44^{-1}	2.9×10^{-6}	1.68×10^3	—
	3	1.3×10^{-5}	1.18	1.00^{-1}	1.3×10^{-5}	1.75	1.07^{-1}		6.66	—
	5	7.8×10^{-9}	1.17	1.00^{-1}	2.9×10^{-6}	1.01	1.11^{-1}		1.01	1.11^{-1}
4 (1.5×10^5)	1	4.4×10^{-3}	1.09	1.00^{-1}	4.4×10^{-3}	1.44	1.37^{-1}	6.3×10^{-8}	7.28×10^4	—
	3	1.8×10^{-5}	1.15	1.00^{-1}	1.8×10^{-5}	1.67	1.06^{-1}		3.72×10^2	—
	5	2.4×10^{-8}	1.11	1.00^{-1}	6.8×10^{-8}	1.34	1.35^{-1}		1.38	1.49^{-1}
	6	1.1×10^{-9}	1.11	1.00^{-1}	6.3×10^{-8}	1.02	1.15^{-1}		1.02	1.15^{-1}

Table 4: Peak problem, multigrid V-cycles: effectivity of the upper (UB) and lower (LB) error bounds

steps, so that $J = 4$. In each step, each triangle is decomposed into four congruent subtriangles. The arising algebraic systems (2.4) are solved iteratively by three different solvers: 1) a geometric multigrid method (MG) with V(5,0)-cycles, i.e. employing 5 Gauss–Seidel pre-smoothing iterations and no post-smoothing; 2) a full multigrid (FMG) method using a single V(3,3)-cycle on each level, i.e., with 3 Gauss–Seidel pre- and post-smoothing iterations; 3) a preconditioned conjugate gradient method (PCG) with an incomplete Cholesky preconditioner with the relative drop-off tolerance 10^{-4} (see, e.g., [49, Sec. 10.4]). The multigrid transfer operators are the canonical ones, exploiting the nestedness of the finite element spaces associated with the different mesh levels.

For this work, we use the standard Lagrange bases and mesh-based nestedness only, we do not employ the p-multigrid techniques of, e.g., Pasquetti and Rapetti [42] or Sundar *et al.* [51], or hierarchical bases, see, e.g., Vassilevski [52] and the references therein. Although the choice of the basis and of the MG type may significantly influence the convergence rates or the cost per iteration, the presented a posteriori bounds are independent of it. Our starting iterate is the zero vector, and we use the stopping criterion (5.8a) with the parameter $\gamma = 0.1$, except for the FMG method with V(3,3)-cycle, for which one single iteration is performed. For the multigrid iteration using V(5,0)-cycles, all our iterates satisfy (up to possible round-off errors in the direct solver on the coarsest level) the coarsest-level orthogonality property (8.1) so that the coarsest level solve (3.1) is not required. This is not the case for the FMG method and the PCG solver. Here, after each iteration, we solve the coarsest-level defect problem (3.1) to construct the Riesz representer $\rho_{0,\mathrm{alg}} \in V_0^p$. The algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-lifting $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ and the discretization flux reconstruction $\boldsymbol{\sigma}_{h,\mathrm{dis}}$ are constructed by Definitions 6.3 and 7.1, respectively, with $q = p$, and the algebraic and total residual liftings $\rho_{h,\mathrm{alg}}$ and $\rho_{h,\mathrm{tot}}$ are obtained by Definitions 6.5 and 7.3, respectively.

Tables 1–8 show the values of the total, algebraic, and discretization errors and the effectivity indices of

p (N_h)	algebraic error	eff. index UB	eff. index LB	total error	eff. index UB	eff. index LB	discretization error	eff. index UB	eff. index LB
1 (9.3×10^3)	1.8×10^{-5}	1.02	1.00^{-1}	3.3×10^{-3}	1.04	1.03^{-1}	3.3×10^{-3}	1.04	1.03^{-1}
2 (3.8×10^4)	1.9×10^{-7}	1.07	1.00^{-1}	1.1×10^{-4}	1.01	1.01^{-1}	1.1×10^{-4}	1.01	1.01^{-1}
3 (8.5×10^4)	2.2×10^{-7}	1.08	1.00^{-1}	2.9×10^{-6}	1.08	1.12^{-1}	2.9×10^{-6}	1.08	1.13^{-1}
4 (1.5×10^5)	9.1×10^{-9}	1.05	1.00^{-1}	6.4×10^{-8}	1.14	1.25^{-1}	6.3×10^{-8}	1.15	1.26^{-1}

Table 5: Peak problem, one full multigrid cycle: effectivity of the upper (UB) and lower (LB) error bounds

p (N_h)	PCG iter	algebraic error	eff. index UB	eff. index LB	total error	eff. index UB	eff. index LB	discretization error	eff. index UB	eff. index LB
1 (9.3×10^3)	2	1.0×10^{-3}	1.01	1.00^{-1}	3.5×10^{-3}	1.29	1.09^{-1}	3.3×10^{-3}	1.32	1.10^{-1}
	4	9.1×10^{-5}	1.02	1.00^{-1}	3.3×10^{-3}	1.07	1.03^{-1}		1.07	1.03^{-1}
2 (3.8×10^4)	4	6.1×10^{-4}	1.01	1.00^{-1}	6.2×10^{-4}	1.20	4.57^{-1}	1.1×10^{-4}	3.81	—
	8	3.2×10^{-6}	1.01	1.00^{-1}	1.1×10^{-4}	1.04	1.01^{-1}		1.04	1.01^{-1}
3 (8.5×10^4)	7	1.1×10^{-3}	1.00	1.00^{-1}	1.1×10^{-3}	1.04	14.0^{-1}	2.9×10^{-6}	1.00×10^2	—
	14	2.2×10^{-5}	1.02	1.00^{-1}	2.2×10^{-5}	1.22	3.61^{-1}		5.41	—
	21	4.8×10^{-8}	1.01	1.00^{-1}	2.9×10^{-6}	1.02	1.11^{-1}		1.02	1.11^{-1}
4 (1.5×10^5)	7	1.1×10^{-3}	1.00	1.00^{-1}	1.1×10^{-3}	1.06	10.5^{-1}	6.3×10^{-8}	5.82×10^3	—
	14	4.9×10^{-5}	1.01	1.00^{-1}	4.9×10^{-5}	1.11	5.31^{-1}		3.80×10^2	—
	21	2.0×10^{-7}	1.01	1.00^{-1}	2.1×10^{-7}	1.28	3.71^{-1}		2.81	—
	28	1.9×10^{-10}	1.01	1.00^{-1}	6.3×10^{-8}	1.01	1.15^{-1}		1.01	1.15^{-1}

Table 6: Peak problem, PCG iterations: effectivity of the upper (UB) and lower (LB) error bounds

the upper and lower bounds of the estimators η, η_{alg}, η_{dis}, $\underline{\eta}$, $\underline{\eta}_{\text{alg}}$, and $\underline{\eta}_{\text{dis}}$ of Theorems 5.2 and 5.3 and of Corollary 5.5, at specified iterations of the multigrid and conjugate gradient solvers. The effectivity indices for the different terms are computed by

$$\text{effectivity index} := \frac{\text{bound}}{\text{error}}.$$

As for estimating the discretization error, the lower bound $\underline{\eta}_{\text{dis}}$ is not computable in the iterations where the algebraic error is the dominating component of the total one; here the condition $\eta_{\text{alg}} \leq \underline{\eta}$ of Corollary 5.5 is not satisfied. Actually, while the algebraic error is the dominating factor, the upper bound η_{dis} may not provide a relevant information about the discretization error. Deriving a bound on the discretization error that provides an accurate estimate also for approximations with dominating algebraic error is, to the best of our knowledge, an unresolved issue. However, when the stopping criterion (5.8a) is satisfied, all our upper and lower bounds on all the algebraic, total, and discretization errors are very precise, with effectivity indices always below 1.7. Note finally that in the full multigrid case, the algebraic error is typically reduced to the discretization error within one step. In addition to the presented results, we also tested the accuracy of the approximate discretization error estimator $\widetilde{\eta}_{\text{dis}}$, see (5.2b). In our experiments it actually bounds the discretization error from above and is always smaller than the guaranteed discretization error upper bound η_{dis} given in Corollary 5.5.

The upper bounds (5.2b) and (5.2a) allow estimating the local distribution of the algebraic and total errors. In the numerical experiments, we observe that $\eta_K := \widetilde{\eta}_{\text{dis},K} + \eta_{\text{alg},K} + \eta_{\text{osc},K}$ and $\eta_{\text{alg},K}$ show an excellent agreement for the local error distribution of the algebraic and total errors, respectively, see Figures 3–8. Note that the local distribution of the algebraic and discretization errors can be very different (cf. [40]), see in particular Figures 7–8, where the discretization error is concentrated in the re-entrant corner, whereas the algebraic error is increased in other parts of the domain Ω. Especially in such situations, the local stopping criteria (5.8b) and (5.9b) may be very much relevant not to let the total error be locally dominated by the algebraic one, even if the globally measured algebraic error is small, see the discussion in [40, 41].

10 Conclusion

The proposed a posteriori error estimators enable to monitor the algebraic and total errors in each iteration of an arbitrary algebraic solver, and consequently also the discretization one. The construction of the

p (N_h)	MG iter	algebraic error	eff. index UB	eff. index LB	total error	eff. index UB	eff. index LB	discretization error	eff. index UB	eff. index LB
1 (2.5×10^4)	1	1.4	1.14	1.03^{-1}	1.4	1.60	1.27^{-1}	2.2×10^{-2}	8.31×10^{1}	—
	2	6.7×10^{-2}	1.14	1.04^{-1}	7.0×10^{-2}	1.61	1.29^{-1}		4.22	—
	3	4.3×10^{-3}	1.16	1.07^{-1}	2.3×10^{-2}	1.37	1.16^{-1}		1.38	1.17^{-1}
	4	4.1×10^{-4}	1.17	1.09^{-1}	2.2×10^{-2}	1.22	1.13^{-1}		1.22	1.13^{-1}
2 (1.0×10^5)	1	2.6	1.19	1.01^{-1}	2.6	1.78	1.08^{-1}	8.9×10^{-3}	4.31×10^{2}	—
	2	8.9×10^{-2}	1.19	1.01^{-1}	8.9×10^{-2}	1.79	1.05^{-1}		1.49×10^{1}	—
	3	2.2×10^{-3}	1.18	1.01^{-1}	9.2×10^{-3}	1.55	1.42^{-1}		1.58	1.50^{-1}
	4	8.6×10^{-5}	1.19	1.02^{-1}	8.9×10^{-3}	1.32	1.29^{-1}		1.32	1.29^{-1}
3 (2.3×10^5)	1	2.4	1.19	1.00^{-1}	2.4	1.72	1.11^{-1}	5.3×10^{-3}	6.29×10^{2}	—
	2	1.1×10^{-1}	1.20	1.00^{-1}	1.1×10^{-1}	1.76	1.07^{-1}		2.92×10^{1}	—
	3	3.6×10^{-3}	1.18	1.00^{-1}	6.4×10^{-3}	1.89	1.47^{-1}		2.19	6.42^{-1}
	4	1.8×10^{-4}	1.17	1.01^{-1}	5.3×10^{-3}	1.48	1.42^{-1}		1.48	1.42^{-1}
4 (4.0×10^5)	1	2.6	1.18	1.00^{-1}	2.6	1.68	1.09^{-1}	3.8×10^{-3}	9.42×10^{2}	—
	2	1.3×10^{-1}	1.18	1.00^{-1}	1.3×10^{-1}	1.71	1.07^{-1}		4.93×10^{1}	—
	3	6.0×10^{-3}	1.16	1.00^{-1}	7.1×10^{-3}	1.87	1.20^{-1}		3.13	—
	4	3.5×10^{-4}	1.13	1.00^{-1}	3.8×10^{-3}	1.57	1.66^{-1}		1.57	1.67^{-1}

Table 7: L-shape problem, multigrid V-cycles: effectivity of the upper (UB) and lower (LB) error bounds

p (N_h)	algebraic error	eff. index UB	eff. index LB	total error	eff. index UB	eff. index LB	discretization error	eff. index UB	eff. index LB
1 (2.5×10^4)	4.4×10^{-4}	1.11	1.01^{-1}	2.2×10^{-2}	1.22	1.13^{-1}	2.2×10^{-2}	1.22	1.13^{-1}
2 (1.0×10^5)	8.0×10^{-5}	1.12	1.01^{-1}	8.9×10^{-3}	1.32	1.28^{-1}	8.9×10^{-3}	1.32	1.28^{-1}
3 (2.3×10^5)	5.5×10^{-5}	1.09	1.00^{-1}	5.3×10^{-3}	1.45	1.42^{-1}	5.3×10^{-3}	1.45	1.42^{-1}
4 (4.0×10^5)	7.2×10^{-5}	1.08	1.00^{-1}	3.8×10^{-3}	1.49	1.62^{-1}	3.8×10^{-3}	1.49	1.62^{-1}

Table 8: L-shape problem, one full multigrid cycle: effectivity of the upper (UB) and lower (LB) error bounds

algebraic bounds relies on computing the Riesz projection of the algebraic residual onto the coarsest level and exploits a hierarchy of nested meshes. In the special case of a geometric multigrid solver and in absence of post-smoothing, the Riesz projection is actually not necessary. Both algebraic and total error estimators are based on solving small independent local problems on patches of elements, which can naturally be done in parallel. In our test cases, we observed excellent effectivity indices for all the tested finite element polynomial degrees $p = 1, \dots, 4$, V-cycle multigrid, full multigrid, and preconditioned conjugate gradients solvers, and all algebraic, total, and discretization errors.

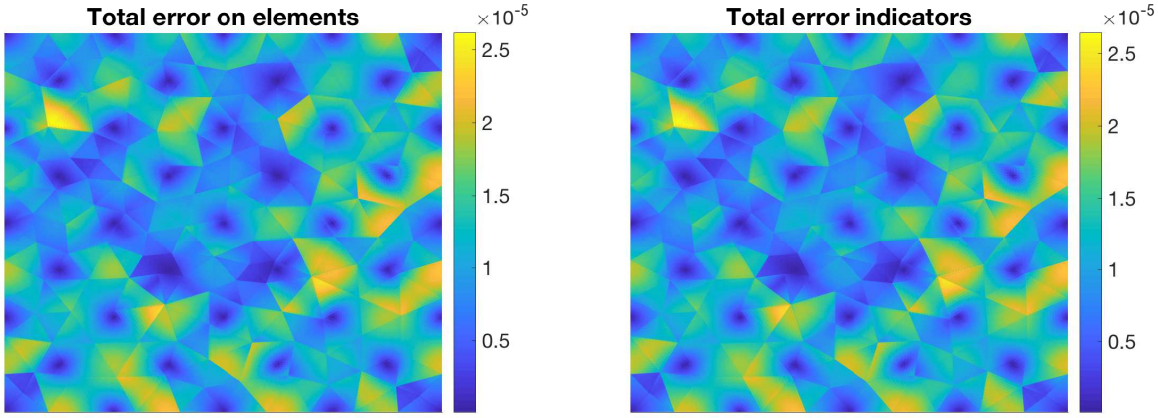


Figure 3: Sinus problem, $p = 2$: elementwise distribution of the total energy error $\|\nabla(u - u_h)\|_K$ (left) and the of local error indicators η_K (right) after one full multigrid cycle

p (N_h)	PCG iter	algebraic error	eff. index UB	eff. index LB	total error	eff. index UB	eff. index LB	discretization error	eff. index UB	eff. index LB
1 (2.5×10^4)	4	8.9×10^{-2}	1.02	1.00^{-1}	9.1×10^{-2}	1.26	4.33^{-1}	2.2×10^{-2}	3.35	—
	8	3.8×10^{-4}	1.01	1.00^{-1}	2.2×10^{-2}	1.22	1.12^{-1}		1.22	1.12^{-1}
2 (1.0×10^5)	4	6.2×10^{-1}	1.01	1.00^{-1}	6.2×10^{-1}	1.07	9.06^{-1}	8.9×10^{-3}	2.61×10^1	—
	8	6.0×10^{-3}	1.01	1.00^{-1}	1.1×10^{-2}	1.65	1.58^{-1}		1.88	2.86^{-1}
	12	1.9×10^{-4}	1.01	1.00^{-1}	8.9×10^{-3}	1.33	1.28^{-1}		1.33	1.28^{-1}
3 (2.3×10^5)	7	1.0	1.00	1.00^{-1}	1.0	1.05	10.0^{-1}	5.3×10^{-3}	6.29×10^1	—
	14	3.1×10^{-2}	1.01	1.00^{-1}	3.1×10^{-2}	1.24	6.25^{-1}		4.48	—
	21	1.7×10^{-3}	1.00	1.00^{-1}	5.6×10^{-3}	1.68	1.48^{-1}		1.74	1.59^{-1}
	28	9.6×10^{-5}	1.00	1.00^{-1}	5.3×10^{-3}	1.46	1.41^{-1}		1.46	1.41^{-1}
4 (4.0×10^5)	7	1.2	1.01	1.00^{-1}	1.2	1.08	7.56^{-1}	3.8×10^{-3}	1.30×10^2	—
	14	5.0×10^{-2}	1.01	1.00^{-1}	5.1×10^{-2}	1.14	6.77^{-1}		7.34	—
	21	3.4×10^{-3}	1.00	1.00^{-1}	5.0×10^{-3}	1.77	2.23^{-1}		2.19	—
	28	1.8×10^{-4}	1.01	1.00^{-1}	3.8×10^{-3}	1.52	1.60^{-1}		1.52	1.60^{-1}

Table 9: L-shape problem, PCG iterations: effectivity of the upper (UB) and lower (LB) error bounds

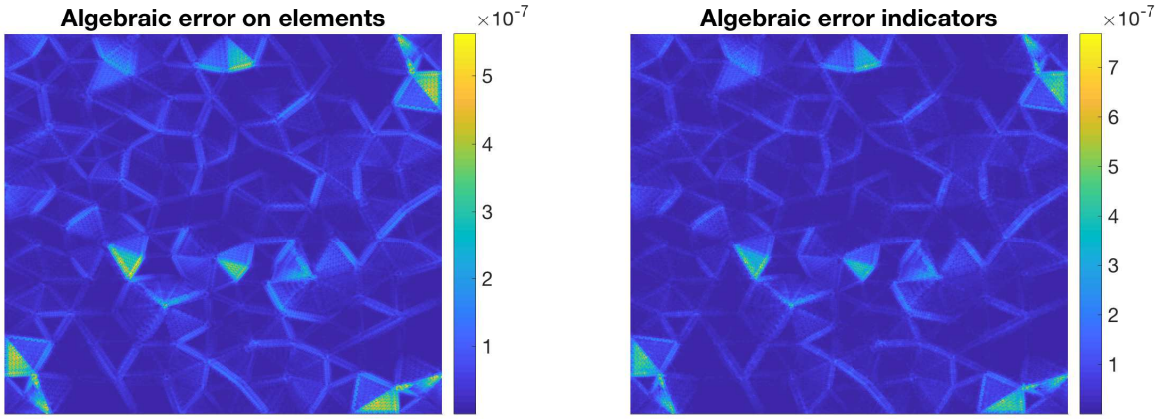


Figure 4: Sinus problem, $p = 2$: elementwise distribution of the algebraic energy error $\|\nabla(u_h^{\mathrm{ex}} - u_h)\|_K$ (left) and of the local error indicators $\eta_{\mathrm{alg},K}$ (right) after one full multigrid cycle

Controlling all error components at any point opens several new perspectives. First, the present unknown-constant-free approach can significantly reduce the overall cost of iterative solutions of large algebraic systems. Even though a non-negligible effort is spent on the evaluation of the estimators, no (unknown) "safety multiplicative factor" is necessary to ensure reaching the given precision and typically many additional iteration steps and mesh refinements can be saved. Second, as the estimates come with a *construction independent* of both the *numerical scheme* and the *algebraic solver*, they open the way for rigorous "fault control". Note that it is immediate to check, for the constructed fields $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ and $\boldsymbol{\sigma}_{h,\mathrm{dis}}$, whether the crucial property (5.1b) is satisfied exactly/compute the misfit in it. Then even rounding errors and numerical stability issues connected with the computation of the approximate solution u_h given by (2.5) can be addressed. Indeed, computing of the estimators of Theorems 5.2 and 5.3 only involves elementwise evaluations of norms of polynomials, whose exactness or confidence interval inclusion can be ensured. In a more distant future, the developed methodology prones for development of integrated solvers with all adaptive algebraic operations like multigrid restriction, prolongation, cycling, and relaxation; with justified stopping criteria; and coupling with adaptive *hp* refinement on the discretization side.

The derivation of the present estimates allows for several simplifications. This typically introduces various (known) multiplicative constants and decreases the precision of the estimates (sometimes even the guaranteed character is lost) but brings radical cuts in the cost of their evaluation. In particular, it is possible to obtain a fully algebraic version of the estimates on the algebraic error. Moreover, such estimates are in the form analogous to the estimators based on the stable splittings, which opens door to study their mutual relationships. These developments are, however, out of scope of the present manuscript and will be described in detail in a forthcoming contribution.

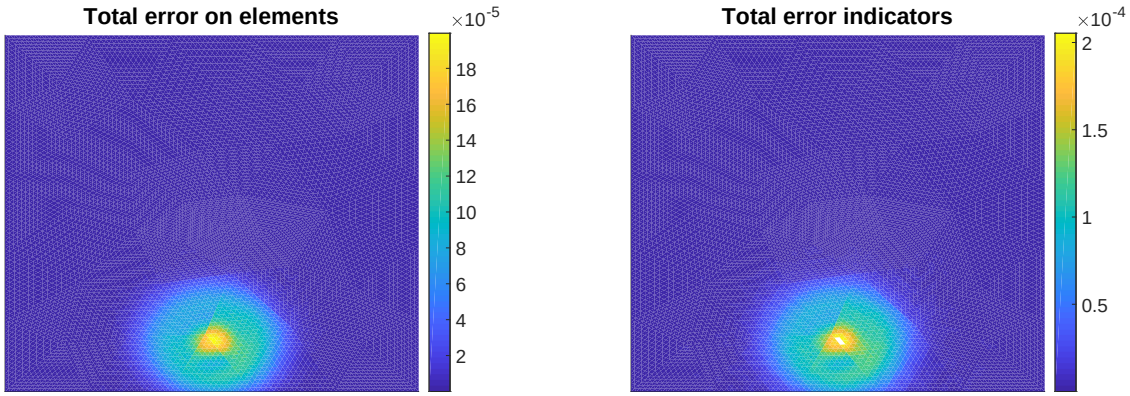


Figure 5: Peak problem, $p = 1$: elementwise distribution of the total energy error $\|\nabla(u - u_h)\|_K$ (left) and of the local error indicators η_K (right) after 2 iterations of the multigrid V-cycle

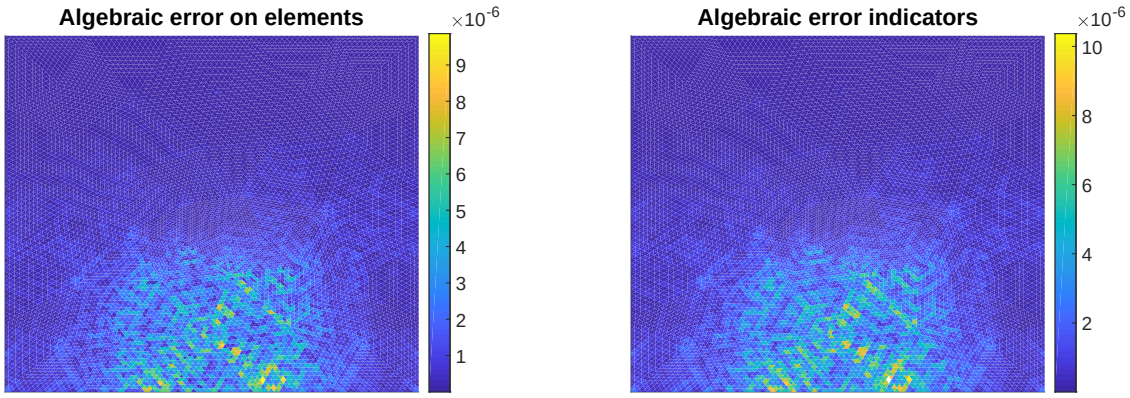


Figure 6: Peak problem, $p = 1$: elementwise distribution of the algebraic energy error $\|\nabla(u_h^{\text{ex}} - u_h)\|_K$ (left) and of the local error indicators $\eta_{\text{alg},K}$ (right) after 2 iterations of the multigrid V-cycle

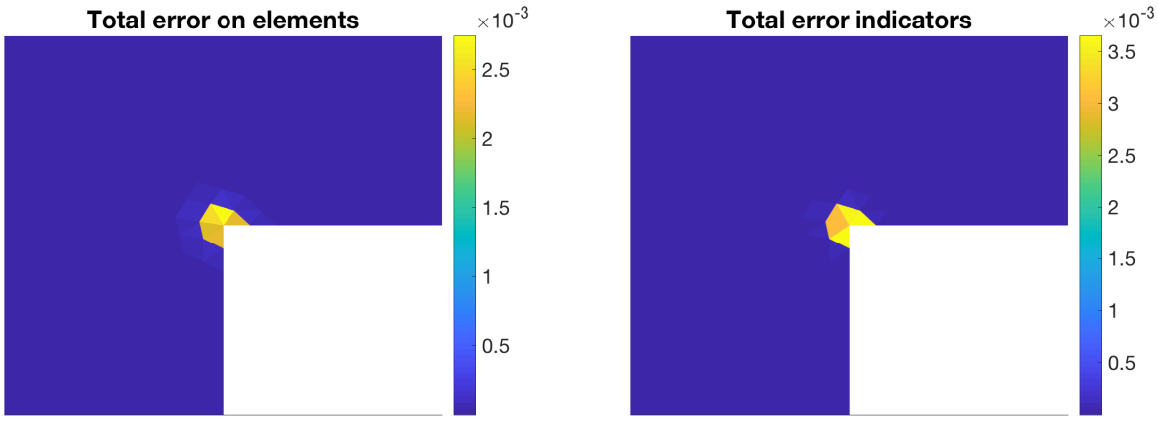


Figure 7: L-shape problem, $p = 3$: elementwise distribution of the total energy error $\|\nabla(u - u_h)\|_K$ (left) and of the local error indicators η_K (right) after 28 PCG iterations. We plot in both figures the part $[-0.1, 0.1] \times [-0.1, 0.1]$ of the discretization domain Ω

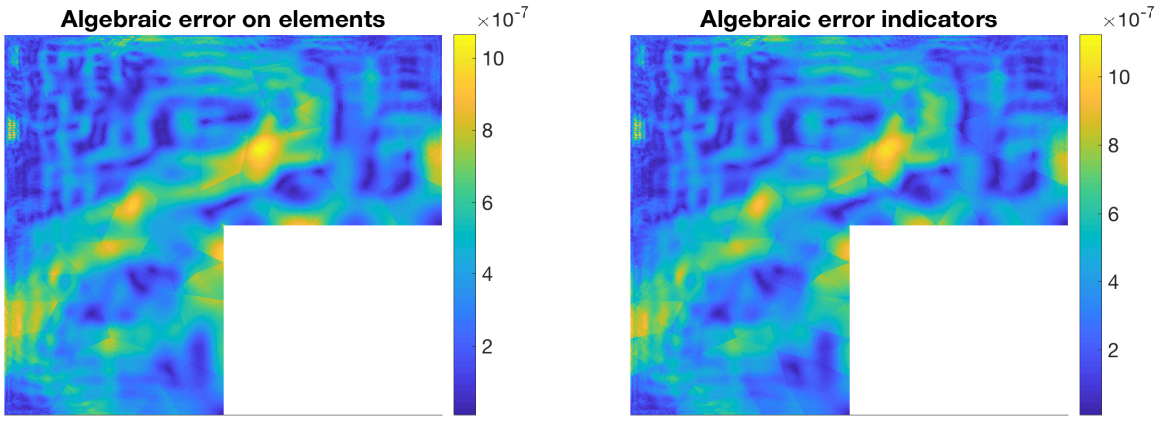


Figure 8: L-shape problem, $p = 3$: elementwise distribution of the algebraic energy error $\|\nabla(u_h^{\mathrm{ex}} - u_h)\|_K$ (left) and of the local error indicators $\eta_{\mathrm{alg},K}$ (right) after 28 PCG iterations

A Comparison of the hierarchical algebraic residual $H_0^1(\Omega)$-lifting with a simpler finest-level lifting

In this appendix, we compare the algebraic residual lifting $\rho_{h,\mathrm{alg}}$ of Definition 6.5 with a simpler construction in the spirit of [4, Sec. 5.1], [45, Sec. 4.1.1], [24, Sec. 3.3], and most precisely of [41, Sec. 5.4]. Variant of Definition 6.5 not using mesh hierarchy but only relying on the two finest meshes $\mathcal{T}_h = \mathcal{T}_J$ and $\mathcal{T}_{J-1}$ reads:

Definition A.1 (Finest-level-only algebraic residual lifting)**.** *Define the algebraic residual lifting*

$$\widehat{\rho}_{h,\mathrm{alg}} := \sum_{\mathbf{a}\in\mathcal{V}_{J-1}} w_{\psi_{J-1}^{\mathbf{a}}}(\widehat{\rho}_{h,\mathrm{alg}}^{\mathbf{a}}) \in V_h \subset H_0^1(\Omega),$$

where the vertex contributions $\widehat{\rho}_{h,\mathrm{alg}}^{\mathbf{a}} \in V_h(\omega_{J-1}^{\mathbf{a}}) \cap H_0^1(\omega_{J-1}^{\mathbf{a}})$ *solve*

$$(\nabla\widehat{\rho}_{h,\mathrm{alg}}^{\mathbf{a}}, \nabla v_h)_{\omega_{J-1}^{\mathbf{a}}} = (f, v_h)_{\omega_{J-1}^{\mathbf{a}}} - (\nabla u_h, \nabla v_h)_{\omega_{J-1}^{\mathbf{a}}} = (r_h, v_h)_{\omega_{J-1}^{\mathbf{a}}} \qquad \forall v_h \in V_h(\omega_{J-1}^{\mathbf{a}}) \cap H_0^1(\omega_{J-1}^{\mathbf{a}}) \quad \text{(A.1)}$$

and the weighting $w_{\psi_{J-1}^{\mathbf{a}}}(\cdot)$ *by the hat functions* $\psi_{J-1}^{\mathbf{a}}$ *is as in Definition 6.5.*

We now construct an example where the lower bound on the algebraic error obtained with $\widehat{\rho}_{h,\mathrm{alg}}$ of Definition A.1 employed in Theorem 5.3 in place of $\rho_{h,\mathrm{alg}}$, i.e., $(r_h, \widehat{\rho}_{h,\mathrm{alg}})/\|\nabla\widehat{\rho}_{h,\mathrm{alg}}\|$, significantly deteriorates with increasing condition number of the stiffness matrix $\mathbb{A}_h$ (either by increasing number of mesh elements, or by increasing polynomial degree), while using the Gauss–Seidel solver. Such a non-robust behavior is not observed for the lower bound $\underline{\eta}_{\mathrm{alg}}$ obtained with the hierarchically constructed lifting $\rho_{h,\mathrm{alg}}$ of Definition 6.5, put forward in this paper. It is, though, interesting to note that both lower bounds with $\widehat{\rho}_{h,\mathrm{alg}}$ and $\rho_{h,\mathrm{alg}}$ behave similarly in the numerical experiments in Sec. 9 for the multigrid solvers (not for the PCG solver).

The precise setting we employ is as follows: we consider the L-shape test problem of Sec. 9, on the initial mesh with element size $h \approx 0.026$, and on the mesh with element size $h/4$ obtained by two uniform refinements. The initial approximation U_h is chosen such that the initial algebraic error $\mathrm{U}_h^{\mathrm{ex}} - \mathrm{U}_h$ is a vector with random elements uniformly distributed in the interval $[-0.5, 0.5] \cdot \|\mathrm{U}_h^{\mathrm{ex}}\|$. Then we use the Gauss–Seidel solver that, after fast reduction of oscillating components of the error in the initial iterations, leads to a very slow decrease of both the energy norm of the algebraic error $\|\mathrm{U}_h^{\mathrm{ex}} - \mathrm{U}_h\|_{\mathbb{A}_h} = \|\nabla(u_h^{\mathrm{ex}} - u_h)\|$ and the Euclidean norm of the algebraic error $\|\mathrm{U}_h^{\mathrm{ex}} - \mathrm{U}_h\|$, as well as to a large ratio between these norms and the Euclidean norm of the algebraic residual $\|\mathrm{R}_h\|$. The algebraic error is getting dominated by smooth components and the Euclidean norm of the algebraic residual may not be descriptive. This situation is depicted in Fig. 9. Fig. 10 then gives the underestimating factor

$$\frac{\|\nabla(u_h^{\mathrm{ex}} - u_h)\|}{\text{lower bound}} \geq 1$$

for the algebraic error lower bound $\underline{\eta}_{\mathrm{alg}}$ given by the hierarchical multilevel lifting $\rho_{h,\mathrm{alg}}$ of Definition 6.5 (with $J = 4$) in comparison to the lower bound corresponding to $\widehat{\rho}_{h,\mathrm{alg}}$ of Definition A.1.

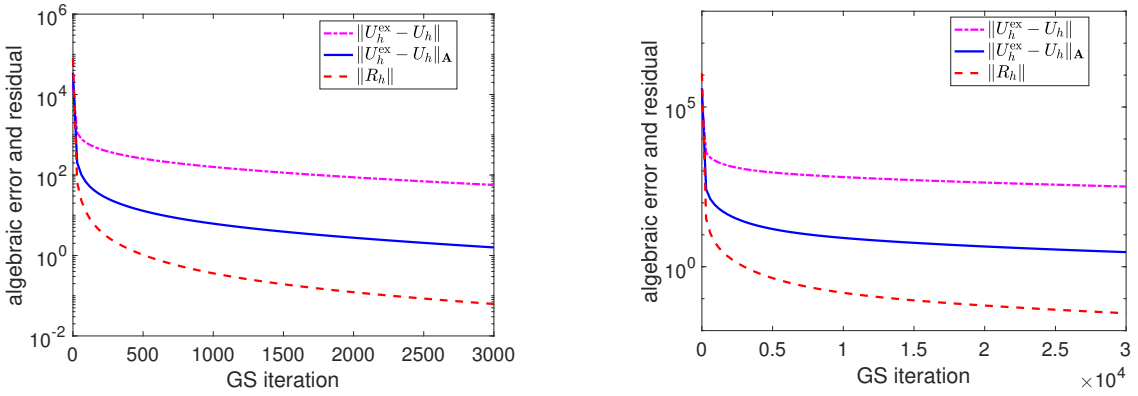


Figure 9: L-shape problem with random initial error and Gauss–Seidel iterative solver: the Euclidean norm $\|\mathrm{U}_h^{\mathrm{ex}} - \mathrm{U}_h\|$ of the algebraic error, the energy norm $\|\mathrm{U}_h^{\mathrm{ex}} - \mathrm{U}_h\|_{\mathbb{A}}$ of the algebraic error, and the Euclidean norm $\|\mathrm{R}_h\|$ of the algebraic residual with the polynomial degrees $p = 1$ (left) and $p = 3$ (right)

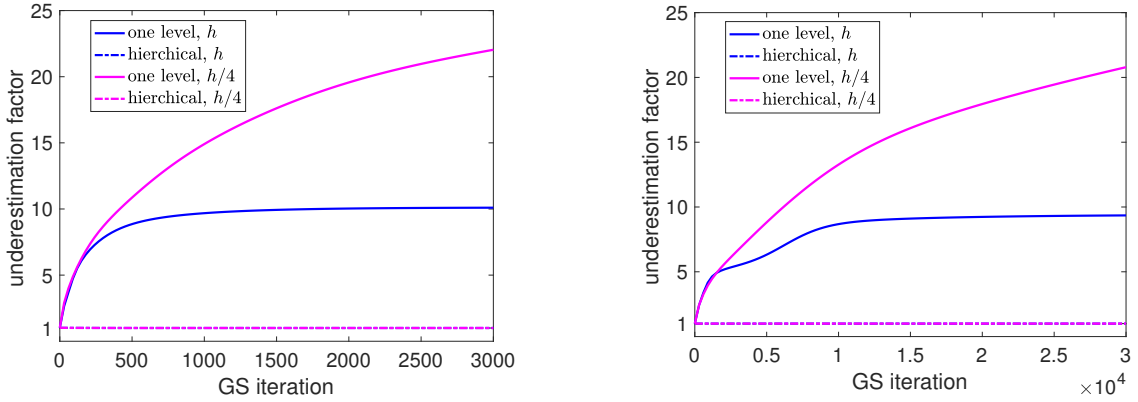


Figure 10: L-shape problem with random initial error and Gauss–Seidel iterative solver: the underestimation of the lower bound given by the hierarchical multilevel construction of Definition 6.5 and the one-level construction of Definition A.1 of the algebraic residual lifting ($\rho_{h,\mathrm{alg}}$, $\widehat{\rho}_{h,\mathrm{alg}}$ respectively) with varying meshsize and with the polynomial degrees $p = 1$ (left) and $p = 3$ (right)

The chosen numerical experiment indicates from Fig. 10 a significant difference between the two approaches, where only the hierarchical multilevel construction of Definition 6.5 is robust with respect to the number of mesh elements and the polynomial degree, yielding underestimation factors very close to the optimal value of one. Clearly, regions of influence are better covered in Definition 6.5 and the algebraic residual is more precisely represented by the lifting in this case. Let us just note that the effectivity index (overestimation factor) of the upper bound η_{alg} of Theorem 5.2 on the algebraic error (which is also constructed hierarchically) stays in this experiment below 1.1.

Remark A.2 (Multigrid solvers and reduction of the algebraic residual on all levels)**.** *It is one of the advantages of a multigrid solver that by the traversal through all meshes in the hierarchy, the residuals on all levels are reduced. Thus, heuristically, when a multigrid solver finishes, the residual is approximately of the same size on all levels, so that evaluating it on the finest level only is enough to provide a good estimate of the energy norm of the algebraic error. This may in particular explain the above-mentioned very good behavior of $\widehat{\rho}_{h,\mathrm{alg}}$ of Definition A.1 for a multigrid solver, in contrast to the Gauss–Seidel solver of Fig. 10. This is also the key for some theoretical considerations where estimates on the energy norm of the algebraic error independent of the number of mesh levels can be shown in an "ideal setting". In practice, though, all conditions may not be met, and consequently a mild dependency on the number of mesh levels may appear. In case of the Gauss–Seidel method tested in this appendix, the residuals are only reduced on the finest level but they stay large on the coarser levels, which implies that the overall algebraic energy error remains large, and also that $\widehat{\rho}_{h,\mathrm{alg}}$ behaves badly whereas $\rho_{h,\mathrm{alg}}$ does not. The results of numerical experiments allow to commend the estimators based on $\boldsymbol{\sigma}_{h,\mathrm{alg}}$ and $\rho_{h,\mathrm{alg}}$ developed here, as we observe a behavior independent of the number of elements and levels, as well as of polynomial degrees, and this not only for multigrid, but also for the PCG and Gauss–Seidel solvers.*

B Proofs of local and global efficiency

We collect in this second appendix the proofs of the results presented in Sec. 5.3 and a closely related final comment.

Proof of Theorem 5.6. Let $K \in \mathcal{T}_h$ be fixed. Definition $\widetilde{\eta}_{\mathrm{dis},K} = \|\nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}\|_K$, Definition 7.1 of $\boldsymbol{\sigma}_{h,\mathrm{dis}}$, the partition of unity on the simplex K by the hat functions $\psi_h^{\mathbf{a}}$ associated with its vertices, and the triangle inequality yield

$$\widetilde{\eta}_{\mathrm{dis},K} = \|\nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}\|_K = \Big\| \sum_{\mathbf{a}\in\mathcal{V}_h,\, \mathbf{a}\subset\partial K} \big(\psi_h^{\mathbf{a}} \nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}\big) \Big\|_K \leq \sum_{\mathbf{a}\in\mathcal{V}_h,\, \mathbf{a}\subset\partial K} \|\psi_h^{\mathbf{a}} \nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}\|_{\omega_h^{\mathbf{a}}}. \tag{B.1}$$

Now, for any vertex $\mathbf{a}$ of the simplex K, using the assumptions on the polynomial degree q and the datum f and crucially the stability bound (5.12),

$$\|\psi_h^{\mathbf{a}} \nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}\|_{\omega_h^{\mathbf{a}}} \leq C_{\mathrm{st}} \sup_{v\in H^1_*(\omega_h^{\mathbf{a}});\, \|\nabla v\|_{\omega_h^{\mathbf{a}}}=1} \big((\nabla(u-u_h), \nabla(\psi_h^{\mathbf{a}} v))_{\omega_h^{\mathbf{a}}} + (\boldsymbol{\sigma}_{h,\mathrm{alg}}, \nabla(\psi_h^{\mathbf{a}} v))_{\omega_h^{\mathbf{a}}} \big).$$

Here we have also employed the weak formulation (2.2), the divergence property of Lemma 6.4, and the Green theorem, since $\boldsymbol{\sigma}_{h,\mathrm{alg}} \in \mathbf{H}(\mathrm{div}, \omega_h^{\mathbf{a}})$ and $(\psi_h^{\mathbf{a}} v) \in H^1_0(\omega_h^{\mathbf{a}})$. Thus, the Cauchy–Schwarz and triangle inequalities together with (5.11) give

$$\|\psi_h^{\mathbf{a}} \nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}\|_{\omega_h^{\mathbf{a}}} \leq C_{\mathrm{st}} C_{\mathrm{cont,PF}} \big(\|\nabla(u-u_h)\|_{\omega_h^{\mathbf{a}}} + \|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|_{\omega_h^{\mathbf{a}}} \big). \tag{B.2}$$

By assumption, we have $\eta_{\mathrm{osc}} = 0$, and thus the local stopping criterion (5.8b) implies

$$\widetilde{\eta}_{\mathrm{dis},K} + \eta_{\mathrm{alg},K} \leq (1+\gamma_K)\widetilde{\eta}_{\mathrm{dis},K} = (1+\gamma_K)\|\nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}\|_K$$

and

$$\sum_{\mathbf{a}\in\mathcal{V}_h,\, \mathbf{a}\subset\partial K} \|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|_{\omega_h^{\mathbf{a}}} = \sum_{\mathbf{a}\in\mathcal{V}_h,\, \mathbf{a}\subset\partial K} \eta_{\mathrm{alg},\omega_h^{\mathbf{a}}} \leq (d+1)\gamma_K \widetilde{\eta}_{\mathrm{dis},K},$$

since any simplex has $(d+1)$ vertices. Using (B.1), (B.2), and the above inequalities we infer,

$$(1+\gamma_K)\widetilde{\eta}_{\mathrm{dis},K} \leq (1+\gamma_K) C_{\mathrm{st}} C_{\mathrm{cont,PF}} \left(\sum_{\mathbf{a}\in\mathcal{V}_h,\, \mathbf{a}\subset\partial K} \|\nabla(u-u_h)\|_{\omega_h^{\mathbf{a}}} + (d+1)\gamma_K \widetilde{\eta}_{\mathrm{dis},K} \right).$$

Consequently, the requirement on γ_K (5.13) yields

$$(1+\gamma_K)\widetilde{\eta}_{\mathrm{dis},K} \leq 2(1+\gamma_K) C_{\mathrm{st}} C_{\mathrm{cont,PF}} \sum_{\mathbf{a}\in\mathcal{V}_h,\, \mathbf{a}\subset\partial K} \|\nabla(u-u_h)\|_{\omega_h^{\mathbf{a}}},$$

and we conclude (5.14).

Under the criterion (5.9b), employing the bound $\|\nabla \rho_{h,\mathrm{tot}}^{\mathbf{a}}\|_{\omega_h^{\mathbf{a}}} \leq C_{\mathrm{cont,PF}} \|\nabla(u-u_h)\|_{\omega_h^{\mathbf{a}}}$ which immediately follows from Definition 7.3 as in [41, proof of Theorem 8],

$$\|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|_{\omega_h^{\mathbf{a}}} \leq \frac{\gamma_K}{(1+\gamma_K^2)^{\frac{1}{2}}} \|\nabla(u-u_h)\|_{\omega_h^{\mathbf{a}}} \leq \gamma_K \|\nabla(u-u_h)\|_{\omega_h^{\mathbf{a}}}.$$

Thus (B.2) gives

$$\|\psi_h^{\mathbf{a}} \nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}\|_{\omega_h^{\mathbf{a}}} \leq (1+\gamma_K) C_{\mathrm{st}} C_{\mathrm{cont,PF}} \|\nabla(u-u_h)\|_{\omega_h^{\mathbf{a}}}$$

and the claim follows from (B.1), $\|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|_K \leq \|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|_{\omega_h^{\mathbf{a}}}$, and the fact that $C_{\mathrm{st}} \geq 1$, $C_{\mathrm{cont,PF}} \geq 1$. □

Proof of Theorem 5.7. The proof follows that of Theorem 5.6. The overlap in the construction of $\boldsymbol{\sigma}_{h,\mathrm{dis}}$ by $\boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}$ leads to

$$\|\nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}\|^2 \leq (d+1) \sum_{\mathbf{a}\in\mathcal{V}_h} \|\psi_h^{\mathbf{a}} \nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}^{\mathbf{a}}\|_{\omega_h^{\mathbf{a}}}^2,$$

whence (B.2) results in

$$\|\nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}\|^2 \leq (d+1) C_{\mathrm{st}}^2 C_{\mathrm{cont,PF}}^2 \sum_{\mathbf{a}\in\mathcal{V}_h} \big(\|\nabla(u-u_h)\|_{\omega_h^{\mathbf{a}}} + \|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|_{\omega_h^{\mathbf{a}}} \big)^2,$$

so that

$$\|\nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}\| \leq (d+1) C_{\mathrm{st}} C_{\mathrm{cont,PF}} (\|\nabla(u-u_h)\| + \|\boldsymbol{\sigma}_{h,\mathrm{alg}}\|). \tag{B.3}$$

Using the global stopping criterion (5.8a),

$$\widetilde{\eta}_{\mathrm{dis}} + \eta_{\mathrm{alg}} \leq (1+\gamma)\widetilde{\eta}_{\mathrm{dis}} \leq (1+\gamma)(d+1) C_{\mathrm{st}} C_{\mathrm{cont,PF}} (\|\nabla(u-u_h)\| + \gamma \widetilde{\eta}_{\mathrm{dis}}),$$

and finally (5.15) leads to

$$(1+\gamma)\widetilde{\eta}_{\mathrm{dis}} \leq 2(1+\gamma)(d+1) C_{\mathrm{st}} C_{\mathrm{cont,PF}} \|\nabla(u-u_h)\|$$

and to the assertion.

It is immediate, see the discussion in [41, Sec. 6.3], that the safe criterion (5.9a) is equivalent to $\eta_{\mathrm{alg}} \leq \gamma/(1+\gamma^2)^{\frac{1}{2}} \underline{\eta}$ with $\underline{\eta}$ the guaranteed total error lower bound of Theorem 5.3. Thus, using (5.5b),

$$\eta_{\mathrm{alg}} = \|\boldsymbol{\sigma}_{h,\mathrm{alg}}\| \leq \frac{\gamma}{(1+\gamma^2)^{\frac{1}{2}}} \|\nabla(u-u_h)\| \leq \gamma \|\nabla(u-u_h)\|,$$

and, in combination with (B.3),

$$\widetilde{\eta}_{\mathrm{dis}} = \|\nabla u_h + \boldsymbol{\sigma}_{h,\mathrm{dis}}\| \leq (1+\gamma)(d+1) C_{\mathrm{st}} C_{\mathrm{cont,PF}} \|\nabla(u-u_h)\|,$$

so that the assertion follows. □

Remark B.1 ($\mathbf{H}(\mathrm{div},\Omega)$-liftings with minimal algebraic error resolution)**.** *In [23, Definition 6.9] a different definition for $\boldsymbol{\sigma}^{\mathbf{a}}_{h,\mathrm{dis}}$ is used, where $r_h \psi^{\mathbf{a}}_h$ in (7.1b) is replaced by the constant $(r_h \psi^{\mathbf{a}}_h, 1)_{\omega^{\mathbf{a}}_h}/|\omega^{\mathbf{a}}_h|$ and similarly in (6.2b) and (6.6b). Though still $\nabla\cdot(\boldsymbol{\sigma}_{h,\mathrm{alg}} + \boldsymbol{\sigma}_{h,\mathrm{dis}}) = \Pi_h f$ as in Lemma 5.1, so that the total upper bound (5.2b) of Theorem 5.2 stays true, $\nabla\cdot\boldsymbol{\sigma}_{h,\mathrm{alg}} \neq r_h$, as the algebraic residual representer r_h is only seen via a combination of its averages. The local efficiency of Theorem 5.6 of the upper bound on the total error is actually improved, in the sense that there is no need for the assumption (5.13) and the choice $q = p$ is sufficient, and similarly in Theorem 5.7. Unfortunately, there is no algebraic error control in the sense of (5.2a) of Theorem 5.2, and, moreover, the local problems corresponding to coarser patches $\omega^{\mathbf{a}}_{J-1}$, $\mathbf{a} \in \mathcal{V}_{J-1}$, also appear in the discretization flux reconstruction and not only in the algebraic residual $\mathbf{H}(\mathrm{div},\Omega)$-lifting.*

References

[1] M. Ainsworth, *A framework for obtaining guaranteed error bounds for finite element approximations*, J. Comput. Appl. Math., 234 (2010), pp. 2618–2632.

[2] M. Arioli, E. H. Georgoulis, and D. Loghin, *Stopping criteria for adaptive finite element solvers*, SIAM J. Sci. Comput., 35 (2013), pp. A1537–A1559.

[3] M. Arioli, D. Loghin, and A. J. Wathen, *Stopping criteria for iterations in finite element methods*, Numer. Math., 99 (2005), pp. 381–410.

[4] I. Babuška and T. Strouboulis, *The finite element method and its reliability*, Numerical Mathematics and Scientific Computation, The Clarendon Press Oxford University Press, New York, 2001.

[5] D. Bai and A. Brandt, *Local mesh refinement multilevel techniques*, SIAM J. Sci. Statist. Comput., 8 (1987), pp. 109–134.

[6] R. E. Bank and A. H. Sherman, *An adaptive, multilevel method for elliptic boundary value problems*, Computing, 26 (1981), pp. 91–105.

[7] R. E. Bank and R. K. Smith, *A posteriori error estimates based on hierarchical bases*, SIAM J. Numer. Anal., 30 (1993), pp. 921–935.

[8] M. Bebendorf, *A note on the Poincaré inequality for convex domains*, Z. Anal. Anwendungen, 22 (2003), pp. 751–756.

[9] R. Becker, C. Johnson, and R. Rannacher, *Adaptive error control for multigrid finite element methods*, Computing, 55 (1995), pp. 271–288.

[10] R. Becker and S. Mao, *Convergence and quasi-optimal complexity of a simple adaptive finite element method*, M2AN Math. Model. Numer. Anal., 43 (2009), pp. 1203–1219.

[11] D. Braess, V. Pillwein, and J. Schöberl, *Equilibrated residual error estimates are p-robust*, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 1189–1197.

[12] D. Braess and J. Schöberl, *Equilibrated residual error estimator for edge elements*, Math. Comp., 77 (2008), pp. 651–672.

[13] J. H. Bramble, *Multigrid methods*, vol. 294 of Pitman Research Notes in Mathematics Series, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1993.

[14] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390.

[15] ——, *Guide to multigrid development*, in Multigrid methods (Cologne, 1981), vol. 960 of Lecture Notes in Math., Springer, Berlin-New York, 1982, pp. 220–312.

[16] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, vol. 15 of Texts in Applied Mathematics, Springer, New York, third ed., 2008.

[17] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, vol. 15 of Springer Series in Computational Mathematics, Springer-Verlag, New York, 1991.

[18] C. Carstensen and S. A. Funken, *Fully reliable localized error control in the FEM*, SIAM J. Sci. Comput., 21 (1999/00), pp. 1465–1484.

[19] L. Chen, R. H. Nochetto, and J. Xu, *Optimal multilevel methods for graded bisection grids*, Numer. Math., 120 (2012), pp. 1–34.

[20] P. Destuynder and B. Métivet, *Explicit error bounds in a conforming finite element method*, Math. Comp., 68 (1999), pp. 1379–1396.

[21] V. Dolean, P. Jolivet, and F. Nataf, *An introduction to domain decomposition methods*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2015. Algorithms, theory, and parallel implementation.

[22] V. Dolejší, A. Ern, and M. Vohralík, *hp-adaptation driven by polynomial-degree-robust a posteriori error estimates for elliptic problems*, SIAM J. Sci. Comput., 38 (2016), pp. A3220–A3246.

[23] A. Ern and M. Vohralík, *Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs*, SIAM J. Sci. Comput., 35 (2013), pp. A1761–A1791.

[24] ——, *Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations*, SIAM J. Numer. Anal., 53 (2015), pp. 1058–1081.

[25] ——, *Stable broken H^1 and $\boldsymbol{H}(\mathrm{div})$ polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions.* HAL Preprint 01422204, submitted for publication, 2016.

[26] R. D. Falgout and J. E. Jones, *Multigrid on massively parallel architectures*, in Multigrid methods, VI (Gent, 1999), vol. 14 of Lect. Notes Comput. Sci. Eng., Springer, Berlin, 2000, pp. 101–107.

[27] B. Gmeiner, U. Rüde, H. Stengel, C. Waluga, and B. Wohlmuth, *Performance and scalability of hierarchical hybrid multigrid solvers for Stokes systems*, SIAM J. Sci. Comput., 37 (2015), pp. C143–C168.

[28] M. Griebel and P. Oswald, *On the abstract theory of additive and multiplicative Schwarz algorithms*, Numer. Math., 70 (1995), pp. 163–180.

[29] W. Hackbusch, *Multigrid methods and applications*, vol. 4 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1985.

[30] A. Hannukainen, R. Stenberg, and M. Vohralík, *A unified framework for a posteriori error estimation for the Stokes problem*, Numer. Math., 122 (2012), pp. 725–769.

[31] R. Hiptmair, H. Wu, and W. Zheng, *Uniform convergence of adaptive multigrid methods for elliptic problems and Maxwell's equations*, Numer. Math. Theory Methods Appl., 5 (2012), pp. 297–332.

[32] B. Janssen and G. Kanschat, *Adaptive multilevel methods with local smoothing for H^1- and H^{curl}-conforming high order finite element methods*, SIAM J. Sci. Comput., 33 (2011), pp. 2095–2114.

[33] P. Jiránek, Z. Strakoš, and M. Vohralík, *A posteriori error estimates including algebraic error and stopping criteria for iterative solvers*, SIAM J. Sci. Comput., 32 (2010), pp. 1567–1590.

[34] R. Luce and B. I. Wohlmuth, *A local a posteriori error estimator based on equilibrated fluxes*, SIAM J. Numer. Anal., 42 (2004), pp. 1394–1414.

[35] D. Meidner, R. Rannacher, and J. Vihharev, *Goal-oriented error control of the iterative solution of finite element equations*, J. Numer. Math., 17 (2009), pp. 143–172.

[36] S. Nicaise, K. Witowski, and B. I. Wohlmuth, *An a posteriori error estimator for the Lamé equation based on equilibrated fluxes*, IMA J. Numer. Anal., 28 (2008), pp. 331–353.

[37] Y. Notay and A. Napov, *A massively parallel solver for discrete Poisson-like problems*, J. Comput. Phys., 281 (2015), pp. 237–250.

[38] P. Oswald, *Multilevel finite element approximation*, Teubner Skripten zur Numerik. [Teubner Scripts on Numerical Mathematics], B. G. Teubner, Stuttgart, 1994. Theory and applications.

[39] P. Oswald, *Stable subspace splittings for Sobolev spaces and domain decomposition algorithms*, in Domain decomposition methods in scientific and engineering computing (University Park, PA, 1993), vol. 180 of Contemp. Math., Amer. Math. Soc., Providence, RI, 1994, pp. 87–98.

[40] J. Papež, J. Liesen, and Z. Strakoš, *Distribution of the discretization and algebraic error in numerical solution of partial differential equations*, Linear Algebra Appl., 449 (2014), pp. 89–114.

[41] J. Papež, Z. Strakoš, and M. Vohralík, *Estimating and localizing the algebraic and total numerical errors using flux reconstructions*, Numer. Math., (2017). DOI 10.1007/s00211-017-0915-5.

[42] R. Pasquetti and F. Rapetti, *Spectral element methods on simplicial meshes*, in Spectral and high order methods for partial differential equations—ICOSAHOM 2012, vol. 95 of Lect. Notes Comput. Sci. Eng., Springer, Cham, 2014, pp. 37–55.

[43] L. E. Payne and H. F. Weinberger, *An optimal Poincaré inequality for convex domains*, Arch. Rational Mech. Anal., 5 (1960), pp. 286–292.

[44] W. Prager and J. L. Synge, *Approximations in elasticity based on the concept of function space*, Quart. Appl. Math., 5 (1947), pp. 241–269.

[45] S. Repin, *A posteriori estimates for partial differential equations*, vol. 4 of Radon Series on Computational and Applied Mathematics, Walter de Gruyter GmbH & Co. KG, Berlin, 2008.

[46] U. Rüde, *Fully adaptive multigrid methods*, SIAM J. Numer. Anal., 30 (1993), pp. 230–248.

[47] ——, *Mathematical and computational techniques for multilevel adaptive methods*, vol. 13 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1993.

[48] ——, *Error estimates based on stable splittings*, in Domain decomposition methods in scientific and engineering computing (University Park, PA, 1993), vol. 180 of Contemp. Math., Amer. Math. Soc., Providence, RI, 1994, pp. 111–118.

[49] Y. Saad, *Iterative methods for sparse linear systems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, second ed., 2003.

[50] K. Stüben and U. Trottenberg, *Multigrid methods: Fundamental algorithms, model problem analysis and applications*, in Multigrid methods, Springer, 1982, pp. 1–176.

[51] H. Sundar, G. Stadler, and G. Biros, *Comparison of multigrid algorithms for high-order continuous finite element discretizations*, Numer. Linear Algebra Appl., 22 (2015), pp. 664–680.

[52] P. S. Vassilevski, *Multilevel block factorization preconditioners*, Springer, New York, 2008. Matrix-based analysis and algorithms for solving finite element equations.

[53] M. Vohralík, *Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients*, J. Sci. Comput., 46 (2011), pp. 397–438.

[54] H. Wu and Z. Chen, *Uniform convergence of multigrid V-cycle on adaptively refined finite element meshes for second order elliptic problems*, Sci. China Ser. A, 49 (2006), pp. 1405–1429.

[55] J. Xu, L. Chen, and R. H. Nochetto, *Optimal multilevel methods for H(grad), H(curl), and H(div) systems on graded and unstructured grids*, in Multiscale, nonlinear and adaptive approximation, Springer, Berlin, 2009, pp. 599–659.

[56] U. M. Yang, *Parallel algebraic multigrid methods—high performance preconditioners*, in Numerical solution of partial differential equations on parallel computers, vol. 51 of Lect. Notes Comput. Sci. Eng., Springer, Berlin, 2006, pp. 209–236.