

No. 597

March 2019

**OSCILLATION IN A POSTERIORI
ERROR ESTIMATION**

C. Kreuzer, A. Veeger

ISSN: 2190-1767

OSCILLATION IN A POSTERIORI ERROR ESTIMATION

CHRISTIAN KREUZER AND ANDREAS VEESER

ABSTRACT. In a posteriori error analysis, the relationship between error and estimator is usually spoiled by so-called oscillation terms, which cannot be bounded by the error. In order to remedy, we devise a new approach where the oscillation has the following two properties. First, it is dominated by the error, irrespective of mesh fineness and the regularity of data and the exact solution. Second, it captures in terms of data the part of the residual that, in general, cannot be quantified with finite information. The new twist in our approach is a locally stable projection onto discretized residuals.

1. INTRODUCTION

Finite element methods are a successful and well-established technique for the solution of partial differential equations. A key tool for the quality assessment of a given finite element approximation and the application of adaptive techniques are so-called *a posteriori error estimators*. These are functionals that are computable in terms of data and the finite element approximation and aim at quantifying the approximation error. For all known estimators, their actual relationship to the error is spoiled by *oscillation*, i.e., by some additive terms measuring distances between non-discrete and discrete data. Remarkably, oscillation may be even greater than the error. This flaw directly interferes with the quality assessment and, on top of that, it weakens results on adaptive methods and complicates their proofs.

In this article we introduce a new approach to a posteriori error estimation, where oscillation is *error-dominated*, i.e. it is bounded by the error of the finite element approximation, up to a multiplicative constant depending on the shape-regularity of the underlying mesh.

We illustrate this new approach in the simplest case, where the weak solution $u \in H_0^1(\Omega)$ of the Dirichlet-Poisson problem

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

is approximated by the Galerkin approximation U that is continuous and piecewise affine over some simplicial mesh \mathcal{M} of Ω . It is instructive to start by recalling the a posteriori error bounds in terms of the standard residual estimator

$$(1.2) \quad \mathcal{E}_R(U, f, \mathcal{M}) := \left(\sum_{K \in \mathcal{M}} h_K \|J(U)\|_{L^2(\partial K)}^2 + h_K^2 \|f\|_{L^2(K)}^2 \right)^{1/2};$$

see, e.g., Ainsworth and Oden [2] or Verfürth [25]. If $f \in L^2(\Omega)$, then the energy norm error $\|u - U\|_{H_0^1(\Omega)}$ and the estimator are almost equivalent. More precisely, we have

$$(1.3) \quad \|u - U\|_{H_0^1(\Omega)} \lesssim \mathcal{E}_R(U, f, \mathcal{M}), \quad \mathcal{E}_R(U, f, \mathcal{M}) \lesssim \|u - U\|_{H_0^1(\Omega)} + \text{osc}_0(f, \mathcal{M}),$$

Date: March 12, 2019.

Key words and phrases. Galerkin finite element methods, a posteriori error bounds, computability, error-dominated oscillation.

where the interfering oscillation is given by

$$(1.4) \quad \text{osc}_0(f, \mathcal{M})^2 := \sum_{K \in \mathcal{M}} h_K^2 \|f - P_{0, \mathcal{M}} f\|_{L^2(K)}^2 \quad \text{with} \quad P_{0, \mathcal{M}} f|_K := \frac{1}{|K|} \int_K f.$$

Let us discuss the relationship of this classical L^2 -oscillation and the energy norm error; for the proofs of the nontrivial statements, see §3.8. Customarily, oscillation is associated with higher order. This idea is supported by the following observation: if f is actually in $H^1(\Omega)$, then $\text{osc}_0(f, \mathcal{M}) = O(h_{\mathcal{M}}^2)$ as $h_{\mathcal{M}} := \max_{K \in \mathcal{M}} h_K \searrow 0$.

On any *fixed* mesh however, the oscillation $\text{osc}_0(f, \mathcal{M})$ may be arbitrarily greater than the energy norm error $\|u - U\|_{H_0^1(\Omega)}$. This is a consequence of the fact that the L^2 -norm is strictly stronger than the H^{-1} -norm. The use of the L^2 -norm in (1.4) can be traced back to its use in the element residual $h_K \|f\|_{L^2(K)}$ in (1.2) and so it can be motivated by the request for the computability of the estimator. In fact, in contrast to an element residual based upon some local H^{-1} -norm of f , this form reduces to the (approximate) computation of an integral.

One may think that the use of the L^2 -norm is the only reason for the possible relative largeness of oscillations like $\text{osc}_0(f, \mathcal{M})$. Yet, Cohen, DeVore and Nochetto present in [11] a striking example which entails that even the H^{-1} -oscillation

$$(1.5) \quad \min_{g \in \mathbb{P}_0(\mathcal{M})} \|f - g\|_{H^{-1}(\Omega)}^2 \quad \text{with} \quad \mathbb{P}_0(\mathcal{M}) := \{g \in L^\infty(\Omega) \mid \forall K \in \mathcal{M} \ g|_K \text{ is constant}\}$$

from Braess [7] and Stevenson [22] may converge slower than the error; see Lemma 21 below. Notice that this contradicts the aforementioned idea that $\text{osc}_0(f, \mathcal{M})$ is always of higher order and, moreover, in view of $\text{osc}_0(f, \mathcal{M}) \lesssim \mathcal{E}_R(U, f, \mathcal{M})$, it entails that also the estimator $\mathcal{E}_R(U, f, \mathcal{M})$ may decrease slightly slower than the error.

The key tool to overcome the shortcomings of the above oscillations is a new projection operator $\mathcal{P}_{\mathcal{M}}$ enjoying the following properties; see §§3.3-3.5:

- $\mathcal{P}_{\mathcal{M}} f$ is discrete for any functional $f \in H^{-1}(\Omega)$. In comparison to $P_{0, \mathcal{M}}$, the image of $\mathcal{P}_{\mathcal{M}}$ is enriched by the span of the face-supported Dirac distributions and so contains true functionals.
- $\mathcal{P}_{\mathcal{M}} f$ is computable in a local manner. Here computable means that it can be determined from the information available in the linear systems for finite element approximations.
- The local dual norms of the new oscillation $f - \mathcal{P}_{\mathcal{M}} f$ are dominated by corresponding local errors. This property hinges on the face-supported Dirac distributions and on local H^{-1} -stability of $\mathcal{P}_{\mathcal{M}} f$.
- In contrast to the local dual norms of the residual $f + \Delta U$, the local dual norms of the discretized residual $\mathcal{P}_{\mathcal{M}} f + \Delta U$ can be estimated from below and above in a computable manner.

Thanks to these properties, we derive in §§3.6-3.7 abstract a posteriori bounds such that the oscillation is bounded by the error. In §4 we provide several realizations leading to hierarchical estimators and estimators based on local problems or based on equilibrated fluxes. Furthermore, in §4.2 we show that an extension of the standard residual estimator (1.2) onto the image of $\mathcal{P}_{\mathcal{M}}$ satisfies

$$(1.6) \quad \|u - U\|_{H_0^1(\Omega)}^2 \approx \mathcal{E}_R(U, \mathcal{P}_{\mathcal{M}} f, \mathcal{M})^2 + \sum_{z \in \mathcal{V}} \|f - \mathcal{P}_{\mathcal{M}} f\|_{H^{-1}(\omega_z)}^2,$$

where \mathcal{V} stands for the set of vertices of \mathcal{M} and ω_z is the star around the vertex z . A comparison with (1.3) immediately yields:

- Both $\mathcal{E}_R(U, f, \mathcal{M})$ and the right-hand side of (1.6) bound the energy norm error in terms of U , f , and \mathcal{M} . However, while the latter one is free of overestimation, the first one may overestimate, even asymptotically.

- Since $\mathcal{P}_{\mathcal{M}}f$ is discrete and computable in the aforementioned sense, we have that $\mathcal{E}_{\mathbb{R}}(U, \mathcal{P}_{\mathcal{M}}f, \mathcal{M})$ is also computable, while $\mathcal{E}_{\mathbb{R}}(U, f, \mathcal{M})$ is not.
- Equivalence (1.6) thus splits the estimation of the error in two parts, reflecting the spirit of Verfürth [25, Remark 1.8] and Ainsworth [1, Section 3.1]: One part is computable and related to the underlying differential operator. The other one depends solely on data; its computation, or rather estimation, hinges on a priori knowledge.

2. MODEL PROBLEM AND DISCRETIZATION

In order to exemplify our new approach to a posteriori error estimation, we consider the homogeneous Dirichlet problem for Poisson's equation and the energy norm error of the associated linear finite element solution. The purpose of this section is to recall the relevant properties of this boundary value problem and discretization.

We shall use the following notation associated with a (Lebesgue) measurable set ω of \mathbb{R}^d , $d \in \mathbb{N}$. Given $m \in \mathbb{N}$, we let $L^2(\omega; \mathbb{R}^m)$ denote the Lebesgue space of square integrable functions over ω with values in \mathbb{R}^m . We write $\langle v, w \rangle_{\omega}$ and $\|\cdot\|_{\omega}^2$ for its scalar product and its induced norm. For $m = 1$, we abbreviate $L^2(\omega; \mathbb{R})$ to $L^2(\omega)$.

If $\omega \subset \mathbb{R}^d$ is non-empty and open, $H^1(\omega)$ stands for the Sobolev space of all functions in $L^2(\omega)$ whose distributional gradient is also in $L^2(\omega; \mathbb{R}^d)$. Moreover, we let $H_0^1(\omega)$ be the closure in $H^1(\omega)$ of all infinitely differentiable function with compact support in ω . If the boundary $\partial\omega$ of ω is sufficiently regular (e.g., Lipschitz), this are all functions in $H^1(\omega)$ with vanishing trace on the boundary $\partial\omega$. Thanks to Friedrichs' inequality, $H_0^1(\omega)$ is a Hilbert space with scalar product $\langle \nabla \cdot, \nabla \cdot \rangle_{\omega}$ and norm $\|\nabla \cdot\|_{\omega}$. As usual, $H^{-1}(\omega)$ indicates the dual space of $H_0^1(\omega)$, i.e. the space of linear and continuous functionals on $H_0^1(\omega)$. We identify $L^2(\omega)$ with its dual space and thus have

$$(2.1) \quad H_0^1(\omega) \subset L^2(\omega) \subset H^{-1}(\omega).$$

The norm of $H^{-1}(\omega)$ is given by

$$\|\ell\|_{H^{-1}(\omega)} := \sup_{w \in H_0^1(\omega)} \frac{\langle \ell, w \rangle_{\omega}}{\|\nabla w\|_{\omega}}, \quad \ell \in H^{-1}(\omega),$$

where the dual brackets $\langle \ell, w \rangle_{\omega} := \ell(w)$, $w \in H_0^1(\omega)$, extend-restrict the scalar product in $L^2(\omega)$. If $D \subset \mathbb{R}^d$ is a set such that $\overset{\circ}{D}$ is suitable for one of the preceding notations, we also use D instead of the more cumbersome $\overset{\circ}{D}$, e.g. we write also $H^1(D)$ instead of $H^1(\overset{\circ}{D})$.

Let Ω be an open, bounded and connected subset of \mathbb{R}^d whose closure can be subdivided into simplices. We shall omit Ω in the notation of dual pairings and norms. The weak formulation of (1.1) reads as follows:

$$(2.2) \quad \begin{aligned} &\text{Given } f \in H^{-1}(\Omega), \text{ find } u = u_f \in H_0^1(\Omega) \text{ such that} \\ &\forall v \in H_0^1(\Omega) \quad \langle \nabla u, \nabla v \rangle = \langle f, v \rangle. \end{aligned}$$

In other words: we are looking for the Riesz representation of f in $H_0^1(\Omega)$. Notice that the Riesz representation theorem establishes an isomorphism between the space $H_0^1(\Omega)$ of solutions and the space $H^{-1}(\Omega)$ of loads. In particular, a unique solution exists not only for $f \in L^2(\Omega)$ but for all $f \in H^{-1}(\Omega)$. This fact suggests that, at least conceptually, an approximation method for (2.2), along with its a posteriori analysis, should cover also loads in $H^{-1}(\Omega)$.

In order to approximate the solution of (2.2), we use a Galerkin approximation based upon finite elements. For the sake of simplicity, we restrict ourselves to simplicial meshes and lowest order.

Let \mathcal{M} be a simplicial, face-to-face (conforming) mesh of the domain Ω . Given an element $K \in \mathcal{M}$, we denote by $h_K := \text{diam } K := \sup_{x,y \in K} |x - y|$ its diameter and by $\rho_K := \sup\{\text{diam } B \mid B \text{ ball in } K\}$ the maximal diameter of inscribed balls. In what follows, ' \lesssim ' stands for ' $\leq C \cdot$ ', where the generic constant C may depend on d and the *shape coefficient*

$$\sigma(\mathcal{M}) := \max_{K \in \mathcal{M}} \sigma_K \quad \text{with} \quad \sigma_K := \frac{h_K}{\rho_K}.$$

In the case of both inequalities ' \lesssim ' and ' \gtrsim ', we shall use ' \simeq ' as shorthand.

An interelement face of \mathcal{M} is a simplex F with d vertices arising as the intersection $F = K_1 \cap K_2$ of two uniquely determined elements $K_1, K_2 \in \mathcal{M}$. Its associated patch is

$$(2.3) \quad \omega_F := K_1 \cup K_2.$$

We let $\mathcal{F} = \mathcal{F}(\mathcal{M})$ denote the set of all $(d-1)$ -dimensional interelement faces of \mathcal{M} . Given $F \in \mathcal{F}$ and $K \in \mathcal{M}$ with $F \subset K$, we write

$$(2.4) \quad h_{K;F} = \frac{d|K|}{|F|} \in [\rho_K, h_K]$$

for the height of K over F .

Furthermore, $\mathcal{V} = \mathcal{V}(\mathcal{M})$ stands for the set of all vertices of \mathcal{M} . To any vertex $z \in \mathcal{V}$, we associate the sets

$$\omega_z := \bigcup \{K \in \mathcal{M} : K \ni z\}, \quad \sigma_z := \bigcup \{F \in \mathcal{F} : F \ni z\},$$

for which we have

$$(2.5) \quad \#\{K \in \mathcal{M} \mid K \ni z\} \lesssim \#\{F \in \mathcal{F} \mid F \ni z\} \lesssim 1.$$

If $K \in \mathcal{M}$ with $K \subset \omega_z$ for some $z \in \mathcal{V}$, then the diameter h_z of ω_z verifies

$$(2.6) \quad h_K \leq h_z \lesssim h_K.$$

Moreover, if e is a direction, i.e. $e \in \mathbb{R}^d$ with $|e| = 1$, we write $h_{z;e}$ for the maximal length of a line segment in ω_z with direction e . Then

$$(2.7) \quad \tilde{\rho}_z := \inf_{|e|=1} h_{z;e}$$

verifies

$$(2.8) \quad \rho_K \leq \tilde{\rho}_z \lesssim \rho_K$$

whenever $K \in \mathcal{M}$ with $K \subset \omega_z$.

Let \mathbb{P}_k be the space of polynomials of degree at most $k \in \mathbb{N}$ over \mathbb{R}^d and let

$$\mathbb{P}_k(\mathcal{M}) := \{V \in L^\infty(\Omega) \mid V|_K \in \mathbb{P}_k(K) \text{ for all } K \in \mathcal{M}\}$$

be its piecewise counterpart over \mathcal{M} . The space of continuous, piecewise affine functions over \mathcal{M} is then

$$\mathbb{V}(\mathcal{M}) := \mathbb{P}_1(\mathcal{M}) \cap H^1(\Omega) = \mathbb{P}_1(\mathcal{M}) \cap C^0(\bar{\Omega}).$$

Its *nodal basis* $\{\phi_z\}_{z \in \mathcal{V}}$ is defined by

$$\phi_z \in \mathbb{V}(\mathcal{M}) \quad \text{such that} \quad \phi_z(y) := \delta_{zy} \text{ for all } z, y \in \mathcal{V}.$$

This basis provides the nodal value representation

$$V = \sum_{z \in \mathcal{V}} V(z) \phi_z$$

for any $V \in \mathbb{V}(\mathcal{M})$ and the partition of unity

$$(2.9) \quad \sum_{z \in \mathcal{V}} \phi_z = 1 \quad \text{in } \Omega,$$

where, for each vertex $z \in \mathcal{V}$, we have $\text{supp } \phi_z = \omega_z$, with skeleton σ_z . Finally, we recall that, for any element $K \in \mathcal{M}$ and any powers $\alpha_z \in \mathbb{N}_0$, $z \in \mathcal{V} \cap K$, we have

$$(2.10) \quad \int_K \prod_{z \in \mathcal{V} \cap K} \phi_z^{\alpha_z} = \frac{d! \prod_{z \in \mathcal{V} \cap K} \alpha_z!}{(\sum_{z \in \mathcal{V} \cap K} \alpha_z + d)!} |K|.$$

The finite element functions satisfying the boundary condition in (2.2) form the space

$$\mathbb{V}_0(\mathcal{M}) := \{V \in \mathbb{V}(\mathcal{M}) \mid V(z) = 0 \text{ for all } z \in \mathcal{V} \cap \partial\Omega\} = \mathbb{P}_1(\mathcal{M}) \cap H_0^1(\Omega).$$

The associated *Galerkin approximation* $U = U_{f;\mathcal{M}}$ is characterized by

$$(2.11) \quad U \in \mathbb{V}_0(\mathcal{M}) \quad \text{such that} \quad \forall V \in \mathbb{V}_0(\mathcal{M}) \quad \langle \nabla U, \nabla V \rangle = \langle f, V \rangle.$$

Notice that the right-hand side and so U are well-defined, also for $f \in H^{-1}(\Omega)$, thanks to the conformity of $\mathbb{V}_0(\mathcal{M})$. Céa's lemma states that the Galerkin approximation is the best approximation with respect to the energy norm error, i.e.,

$$(2.12) \quad \|\nabla u - \nabla U\|_\Omega \leq \|\nabla u - \nabla V\|_\Omega \quad \text{for all } V \in \mathbb{V}_0(\mathcal{M}).$$

In order to determine the Galerkin approximation U , one usually obtains its values at the interior vertices $\mathcal{V}_0 := \mathcal{V} \cap \Omega$ by solving the symmetric positive definite linear system

$$M\alpha = F,$$

where

$$(2.13) \quad \alpha = (U(z))_{z \in \mathcal{V}_0}, \quad M = (\langle \nabla \phi_z, \nabla \phi_y \rangle)_{y,z \in \mathcal{V}_0}, \quad F = (\langle f, \phi_y \rangle)_{y \in \mathcal{V}_0}.$$

We thus see that the Galerkin approximation U is computable whenever the load evaluations

$$(2.14) \quad \langle f, \phi_y \rangle, y \in \mathcal{V}_0, \text{ are known exactly.}$$

Strictly speaking, these evaluations are in general not computable. In fact, even if $f \in L^2(\Omega)$ is a function, the evaluation of $\langle f, \phi_y \rangle = \int_\Omega f \phi_y$ requires the computation of an integral, which in general can be done only approximately by means of numerical integration. Notwithstanding, error analyses of approximations like (2.11) have proved very useful for the theoretical understanding and underpinning of finite element methods and are therefore very common. Accordingly, we shall suppose that the evaluations (2.14) are known to us. In §3.6 below, we will discuss which kind of additional information is used in our a posteriori analysis.

3. A POSTERIORI ANALYSIS WITH ERROR-DOMINATED OSCILLATION

We present our new approach to a posteriori error analysis by deriving bounds for the energy norm error of the Galerkin approximation (2.11). The key feature of these bounds is that all involved terms are dominated by the error.

3.1. Residual norms. Given some load $f \in H^{-1}(\Omega)$ and a Galerkin approximation $U_{f;\mathcal{M}}$, we want to quantify the energy norm error $\|\nabla(u_f - U_{f;\mathcal{M}})\|$, where the exact solution u_f of (2.2) is typically unknown to us.

Our starting point is the so-called *residual* $\text{Res}(f; \mathcal{M}) \in H^{-1}(\Omega)$ given by

$$\langle \text{Res}(f; \mathcal{M}), v \rangle := \langle f, v \rangle - \langle \nabla U_{f;\mathcal{M}}, \nabla v \rangle \quad \text{for all } v \in H_0^1(\Omega).$$

It is defined in terms of data and the computable Galerkin approximation and vanishes if and only if the latter equals the exact solution. The following lemma shows that appropriately measuring the size of the residual relates to the error.

Lemma 1 (Error, residual and load). *We have*

$$\|\nabla(u_f - U_{f;\mathcal{M}})\| = \|\text{Res}(f; \mathcal{M})\|_{H^{-1}(\Omega)} \leq \|f\|_{H^{-1}(\Omega)}.$$

Proof. Thanks to the differential equation in (2.2), we have, for all $v \in H_0^1(\Omega)$,

$$(3.1) \quad \langle \text{Res}(f; \mathcal{M}), v \rangle = \langle \nabla(u_f - U_{f; \mathcal{M}}), \nabla v \rangle = \langle -\Delta(u_f - U_{f; \mathcal{M}}), v \rangle,$$

where $-\Delta$ indicates the distributional Laplacian. Consequently, the claimed equality follows from the fact that $-\Delta : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$ is an isometry (which follows from the Cauchy-Schwarz inequality in $L^2(\Omega)$ and from testing with $v = u_f - U_{f; \mathcal{M}}$). The claimed inequality follows by invoking also (2.12):

$$\|\nabla(u_f - U_{f; \mathcal{M}})\| \leq \|\nabla u_f\| = \|f\|_{H^{-1}(\Omega)}. \quad \square$$

Thus, we aim now at quantifying the dual norm $\|\text{Res}(f; \mathcal{M})\|_{H^{-1}(\Omega)}$. The following simple observation shows that this task requires much more information than computing the Galerkin approximation.

Lemma 2 (Bounding residual norms). *Without any a priori information on the load $f \in H^{-1}(\Omega)$, the residual norm $\|\text{Res}(f; \mathcal{M})\|_{H^{-1}(\Omega)}$ cannot be bounded in terms of a finite number of adaptive evaluations of the form: $\langle f, v \rangle$ with $v \in H_0^1(\Omega)$.*

Proof. Suppose that the claim is false. Then, for each $f \in H^{-1}(\Omega)$, there is a number $B(f) \geq \|\text{Res}(f; \mathcal{M})\|_{H^{-1}(\Omega)}$ which is given in terms of evaluations $\langle f, v_i \rangle$, $i = 1, \dots, n_f$, where the choice of v_i may depend deterministically on the previous evaluations $\langle f, v_1 \rangle, \dots, \langle f, v_{i-1} \rangle$. Fix some functional $0 \neq \ell \in H^{-1}(\Omega)$. Since $H_0^1(\Omega)$ is infinite-dimensional, we can choose a normalized $w \in H_0^1(\Omega)$ that is perpendicular to $\mathbb{V}_0(\mathcal{M})$ and all test functions v_i , $i = 1, \dots, n_\ell$ associated with ℓ . Set $\delta := 3B(\ell)(-\Delta)w$ and observe that $U_{\delta; \mathcal{M}} = 0$ and $\langle \delta, v_i \rangle = 0$ for all $i = 1, \dots, n_\ell$. Therefore $\langle \ell + \delta, v_i \rangle = \langle \ell, v_i \rangle$ and we obtain the contradiction

$$\begin{aligned} B(\ell) &= B(\ell + \delta) \geq \|\text{Res}(\ell + \delta; \mathcal{M})\|_{H^{-1}(\Omega)} = \|\delta + \text{Res}(\ell; \mathcal{M})\|_{H^{-1}(\Omega)} \\ &\geq \|\delta\|_{H^{-1}(\Omega)} - \|\text{Res}(\ell; \mathcal{M})\|_{H^{-1}(\Omega)} \geq 3B(\ell) - B(\ell) = 2B(\ell) > 0. \quad \square \end{aligned}$$

Remark 3 (Load evaluations vs exact integrals). A similar yet simpler argument shows that, without any a priori information on $f \in L^2(\Omega)$, also $\|f\|$ cannot be bounded in terms of adaptive evaluations $\int_\Omega f v$ with $v \in L^2(\Omega)$.

Before discussing in §3.3 repercussions of Lemma 2, it is useful to take into account a further requirement for a posteriori bounds.

3.2. Localized residual norm. Adaptive mesh refinement is an important application of a posteriori bounds. It is usually based upon the comparison of local quantities. Therefore, it is of interest to split a posteriori bounds, or the residual norm itself, into local contributions.

Such a localization appears implicitly, e.g., in the a posteriori error analysis of Babuška and Miller [3]. It is based upon the $W^{1,\infty}$ -partition of unity (2.9) and the orthogonality property:

$$\langle \text{Res}(f; \mathcal{M}), \phi_z \rangle = 0 \quad \text{for all } z \in \mathcal{V}_0 = \mathcal{V} \cap \Omega.$$

We thus introduce the subclass

$$\mathcal{R}_{\mathcal{M}} := \{\ell \in H^{-1}(\Omega) \mid \forall V \in \mathbb{V}_0(\mathcal{M}) \langle \ell, V \rangle = 0\}$$

of residuals associated with Galerkin approximations. Recall that $\text{supp } \phi_z = \omega_z$ and that $H^{-1}(\omega_z)$ is a shorthand for $H^{-1}(\dot{\omega}_z)$.

Lemma 4 (Localization). *Let $\ell \in H^{-1}(\Omega)$ be any functional.*

(i) *If $\ell \in \mathcal{R}_{\mathcal{M}}$, then*

$$\|\ell\|_{H^{-1}(\Omega)}^2 \lesssim \sum_{z \in \mathcal{V}} \|\ell\|_{H^{-1}(\omega_z)}^2,$$

where the hidden constant depends only on d and the shape coefficient $\sigma(\mathcal{M})$.

(ii) We have

$$\sum_{z \in \mathcal{V}} \|\ell\|_{H^{-1}(\omega_z)}^2 \leq (d+1) \|\ell\|_{H^{-1}(\Omega)}^2.$$

Proof. See also Cohen, DeVore, and Nochetto [11, §3.2 and §3.4], Ern and Guermond [12, Proposition 31.7] or Blechta, Málek, and Vohralík [5, Theorem 3.7]. For the sake of completeness, we provide details. In order to show (i), we fix an arbitrary $v \in H_0^1(\Omega)$. In view of the partition of unity (2.9) and $\ell \in \mathcal{R}_{\mathcal{M}}$, we can write

$$(3.2) \quad \langle \ell, v \rangle = \sum_{z \in \mathcal{V}} \langle \ell, v \phi_z \rangle = \sum_{z \in \mathcal{V}} \langle \ell, (v - c_z) \phi_z \rangle,$$

where the reals $c_z \in \mathbb{R}$ are given by

$$c_z := \frac{\int_{\omega_z} v \phi_z \, dx}{\int_{\omega_z} \phi_z \, dx} \text{ for } z \in \mathcal{V}_0, \quad \text{and } c_z = 0 \text{ for } z \in \mathcal{V} \setminus \mathcal{V}_0.$$

Thanks to $0 \leq \phi_z \leq 1$, the inverse estimate $\|\nabla \phi_z\|_{L^\infty(\omega_z)} \leq \max_{K \subset \omega_z} \rho_K^{-1} \lesssim h_z^{-1}$ and the Poincaré-Friedrichs inequality $\|v - c_z\|_{\omega_z} \lesssim h_z \|\nabla v\|_{\omega_z}$ (see, e.g., Nochetto and Veiser [20, Lemma 4]), we have, for any $z \in \mathcal{V}$,

$$(3.3) \quad \|\nabla((v - c_z) \phi_z)\|_{\omega_z} \leq \|\nabla v\|_{\omega_z} + \|v - c_z\|_{\omega_z} \|\nabla \phi_z\|_{L^\infty(\omega_z)} \leq C_{\sigma(\mathcal{M})} \|\nabla v\|_{\omega_z},$$

where the constant $C_{\sigma(\mathcal{M})}$ depends only on $\sigma(\mathcal{M})$. Thus, (3.2) leads to

$$|\langle \ell, v \rangle| \lesssim \sum_{z \in \mathcal{V}} \|\ell\|_{H^{-1}(\omega_z)} \|\nabla v\|_{\omega_z} \leq \sqrt{d+1} \left(\sum_{z \in \mathcal{V}} \|\ell\|_{H^{-1}(\omega_z)}^2 \right)^{1/2} \|\nabla v\|$$

and the proof of (i) is finished.

To prove (ii), we let $v_z \in H_0^1(\omega_z)$ with $\|\nabla v_z\|_{\omega_z} \leq 1$ for any node $z \in \mathcal{V}$ and set $v = \sum_{z \in \mathcal{V}} \langle \ell, v_z \rangle v_z \in H_0^1(\Omega)$. Then

$$\sum_{z \in \mathcal{V}} \langle \ell, v_z \rangle^2 = \langle \ell, v \rangle \leq \|\ell\|_{H^{-1}(\Omega)} \|\nabla v\|,$$

and, with the help of two Cauchy-Schwarz inequalities,

$$\begin{aligned} \|\nabla v\|^2 &= \sum_{K \in \mathcal{M}} \sum_{z, y \in \mathcal{V} \cap K} \langle \ell, v_z \rangle \langle \ell, v_y \rangle \int_K \nabla v_z \cdot \nabla v_y \\ &\leq \sum_{K \in \mathcal{M}} \sum_{z \in \mathcal{V} \cap K} (d+1) |\langle \ell, v_z \rangle|^2 \|\nabla v_z\|_K^2 = (d+1) \sum_{z \in \mathcal{V}} |\langle \ell, v_z \rangle|^2. \end{aligned}$$

Consequently, we conclude (ii) by taking the suprema over all v_z for all $z \in \mathcal{V}$. \square

Thus, in the context of adaptive mesh refinement, we are also interested in quantifying the single terms of the *localized residual norm*

$$(3.4) \quad \|\text{Res}(f; \mathcal{M})\|_{H^{-1}(\mathcal{M})}^2 := \sum_{z \in \mathcal{V}} \|\text{Res}(f; \mathcal{M})\|_{H^{-1}(\omega_z)}^2.$$

Of course, we face the same problem for the local residual norms as for the global one.

Corollary 5 (Bounding local residual norms). *Without any a priori information on $f \in H^{-1}(\Omega)$, each local residual norm $\|\text{Res}(f; \mathcal{M})\|_{H^{-1}(\omega_z)}$, $z \in \mathcal{V}$, cannot be bounded in terms of a finite number of adaptive evaluations of f .*

Proof. Replace the domain Ω by ω_z in the proof of Lemma 2 and extend functionals in $H^{-1}(\omega_z)$ by 0 on the orthogonal complement of $H_0^1(\omega_z)$ in $H_0^1(\Omega)$. \square

3.3. Towards error-dominated oscillation. In view of Lemma 2 and Corollary 5, a posteriori bounds for the residual norm or its localized variant require knowledge on the load f beyond a finite number of evaluations. The actual knowledge of f can be of different nature and, accordingly, may require different techniques. Here we want to address only aspects of a posteriori error estimation that are independent of the nature of this knowledge. Correspondingly, we split the residual into an *discretized residual* and *data approximation*:

$$(3.5) \quad \text{Res}(f; \mathcal{M}) = (\mathcal{P}_{\mathcal{M}}f + \Delta U_{f; \mathcal{M}}) + (f - \mathcal{P}_{\mathcal{M}}f)$$

where $\mathcal{P}_{\mathcal{M}}$ maps onto a subspace $\mathbb{D}(\mathcal{M})$ of $H^{-1}(\Omega)$ such that

- $\|\mathcal{P}_{\mathcal{M}}f + \Delta U_{f; \mathcal{M}}\|_{H^{-1}(\mathcal{M})}$ can be bounded with the help of a finite number of evaluations of f and
- the task of bounding $\|f - \mathcal{P}_{\mathcal{M}}f\|_{H^{-1}(\mathcal{M})}$ hinges only on knowledge of the load f ; this task may be viewed as a matter of approximation theory since, apart from the choice of the norm, it is independent of the boundary value problem (2.2).

Here we have used the localized dual norm $\|\cdot\|_{H^{-1}(\mathcal{M})}$ in order to allow for applications in mesh adaptivity. It is then desirable that both parts are dominated by the error, i. e., we have

$$(3.6a) \quad \|\mathcal{P}_{\mathcal{M}}f + \Delta U_{f; \mathcal{M}}\|_{H^{-1}(\mathcal{M})} \lesssim \|\nabla(u_f - U_{f; \mathcal{M}})\|,$$

$$(3.6b) \quad \|f - \mathcal{P}_{\mathcal{M}}f\|_{H^{-1}(\mathcal{M})} \lesssim \|\nabla(u_f - U_{f; \mathcal{M}})\|.$$

In view of Lemma 1 and Lemma 4, the two conditions are equivalent.

The construction of a suitable mapping $\mathcal{P}_{\mathcal{M}}$ is the new twist in our approach. In order to get first hints on this, let us test out several candidates with necessary conditions arising from (3.6b).

The proof of Corollary 5 suggests that the problem lies in the fact that f is taken from an infinite-dimensional space. The projection $\mathcal{P}_{0, \mathcal{M}}$ into discrete data from (1.4) is thus a candidate for $\mathcal{P}_{\mathcal{M}}$. This choice, however, does not verify (3.6). In fact, Lemma 1, Lemma 4 (ii), and (3.6b) imply the stability estimate

$$(3.7) \quad \|\mathcal{P}_{\mathcal{M}}f\|_{H^{-1}(\mathcal{M})} \lesssim \|f\|_{H^{-1}(\Omega)},$$

while $\mathcal{P}_{0, \mathcal{M}}f$ is not even defined for a general $f \in H^{-1}(\Omega)$ (and cannot be continuously extended; cf. Lemma 20).

This flaw is easily remedied. For any element $K \in \mathcal{M}$, we replace in (1.4) the characteristic function χ_K of K by the weighted mean

$$(3.8) \quad \psi_K := \frac{(2d+1)!}{d!|K|} \prod_{z \in \mathcal{V} \cap K} \phi_z \in H_0^1(K) \quad \text{with} \quad \int_K \psi_K = 1$$

thanks to (2.10) and consider

$$(3.9) \quad \tilde{\mathcal{P}}_{0, \mathcal{M}}f := \sum_{K \in \mathcal{M}} \langle f, \psi_K \rangle \chi_K.$$

Since $\psi_K \in H_0^1(K) \subset H_0^1(\Omega)$ is an admissible test function, the operator $\tilde{\mathcal{P}}_{0, \mathcal{M}}$ is defined for all functionals in $H^{-1}(\Omega)$ and satisfies the stability estimate (3.7); see Remark 11 below.

But still, the new operator $\tilde{\mathcal{P}}_{0, \mathcal{M}}$ does not verify (3.6). To see this, consider $f = -\Delta V$ with $V \in \mathbb{V}_0(\mathcal{M})$ arbitrary. We then have

$$u_f = U_{f; \mathcal{M}}$$

and therefore $\text{Res}(f; \mathcal{M}) = 0$ and property (3.6b) entails

$$(3.10) \quad \forall V \in \mathbb{V}_0(\mathcal{M}) \quad \mathcal{P}_{\mathcal{M}}(\Delta V) = \Delta V.$$

In addition, integration by parts yields that, for all $v \in H_0^1(\Omega)$,

$$(3.11) \quad \langle \Delta V, v \rangle = - \int_{\Omega} \nabla V \cdot \nabla v = \sum_{F \in \mathcal{F}} \int_F J(V) v \, ds,$$

where ds indicates the $(d-1)$ -dimensional Hausdorff measure in \mathbb{R}^d and $J(V)$ is the jump in the normal flux $\nabla V \cdot n$ across interelement sides. More precisely, if $F = K_1 \cap K_2$ is the intersection of the elements $K_1, K_2 \in \mathcal{M}$ with respective outer normals $\mathbf{n}_1, \mathbf{n}_2$, then $J(V)|_F := \nabla V|_{K_1} \cdot \mathbf{n}_1 + \nabla V|_{K_2} \cdot \mathbf{n}_2 \in \mathbb{R}$. If $V \neq 0$, then we have also $\Delta V \neq 0$, while (3.11) yields $\tilde{\mathcal{P}}_{0,\mathcal{M}}(\Delta V) = 0$, in contradiction with (3.10). Hence (3.6) does not hold for $\tilde{\mathcal{P}}_{0,\mathcal{M}}$.

The two conditions (3.7) and (3.10) are central to our goals. Although they can be checked without involving the Galerkin approximation (2.11), they are also sufficient for (3.6). Incidentally, they imply that $\mathcal{P}_{\mathcal{M}}$ has to be a near best ‘interpolation’ operator in light of the Lebesgue lemma.

The failure of (3.10) for $\tilde{\mathcal{P}}_{0,\mathcal{M}}$ is not related to the choice of the test functions ψ_K , $K \in \mathcal{M}$, but to its range. In fact, (3.11) and the fundamental lemma of calculus of variation show that $\Delta V \notin L^2(\Omega)$ whenever $V \neq 0$, while $\tilde{\mathcal{P}}_{0,\mathcal{M}}(\mathbb{V}_0(\mathcal{M})) \subset L^2(\Omega)$. In other words: to remedy, we have to change the range.

Finally, it is desirable that $\mathcal{P}_{\mathcal{M}}$ is a local operator for two reasons. First, this comes in useful when evaluating $\mathcal{P}_{\mathcal{M}}$. Second, since $-\Delta$ is a local operator, we have the following lower bound for the local error:

$$(3.12) \quad \|\text{Res}(f; \mathcal{M})\|_{H^{-1}(\omega_z)} \leq \|\nabla(u_f - U_{f;\mathcal{M}})\|_{\omega_z},$$

which follows from testing (3.1) with all v from $H_0^1(\omega_z)$. This bound can be exploited if we strengthen (3.6) to the local conditions

$$(3.13a) \quad \|\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)} \lesssim \|\text{Res}(f; \mathcal{M})\|_{H^{-1}(\omega_z)},$$

$$(3.13b) \quad \|f - \mathcal{P}_{\mathcal{M}}f\|_{H^{-1}(\omega_z)} \lesssim \|\text{Res}(f; \mathcal{M})\|_{H^{-1}(\omega_z)}$$

for all $z \in \mathcal{V}$. We shall therefore demand the stability (3.7) and invariance (3.10) in a suitable local manner.

In order to formulate local invariance, let us introduce the following notations associated with an open subset $\omega \subset \Omega$. If $\ell_1, \ell_2 \in H^{-1}(\Omega)$, we say $\ell_1 = \ell_2$ on ω whenever $\ell_1(v) = \ell_2(v)$ for all $v \in H_0^1(\omega)$. Moreover, we write $\ell_1 \in \underline{\mathbb{D}}(\mathcal{M})$ on ω when additionally ℓ_2 can be chosen such that $\ell_2 \in \underline{\mathbb{D}}(\mathcal{M})$. Notice that, thanks to the fundamental lemma of the calculus of variations, these notions reduce to the usual ones if $\ell \in L^2(\Omega)$, i.e. $\ell(v) = \int_{\Omega} g v$ for all $v \in H_0^1(\Omega)$.

Let us summarize our discussion by a list of desired properties for the operator $\mathcal{P}_{\mathcal{M}}$ and its range $\underline{\mathbb{D}}(\mathcal{M}) \subset H^{-1}(\Omega)$, which corresponds to the set of all possible discretized residuals. This list provides the guidelines for our approach and choices. Denoting by $\Delta(\mathbb{V}_0(\mathcal{M})) = \{\Delta V \mid V \in \mathbb{V}_0(\mathcal{M})\}$ the image of $\mathbb{V}_0(\mathcal{M})$ under the distributional Laplacian, we aim for the following properties:

$$(3.14a) \quad \Delta(\mathbb{V}_0(\mathcal{M})) \subset \underline{\mathbb{D}}(\mathcal{M}),$$

$$(3.14b) \quad \text{if } \ell \in \underline{\mathbb{D}}(\mathcal{M}) \text{ on } \hat{\omega}_z, \text{ then } \|\ell\|_{H^{-1}(\omega_z)} \text{ is quantifiable with a finite number of evaluations of } \ell,$$

$$(3.14c) \quad \mathcal{P}_{\mathcal{M}} \text{ is linear,}$$

$$(3.14d) \quad \mathcal{P}_{\mathcal{M}}f \text{ is locally computable in terms of a finite number of evaluations of } f,$$

$$(3.14e) \quad \text{if } \ell \in \underline{\mathbb{D}}(\mathcal{M}) \text{ on } \hat{\omega}_z, \text{ then } \mathcal{P}_{\mathcal{M}}\ell = \ell \text{ on } \hat{\omega}_z,$$

$$(3.14f) \quad \|\mathcal{P}_{\mathcal{M}}\ell\|_{H^{-1}(\omega_z)} \lesssim \|\ell\|_{H^{-1}(\omega_z)} \text{ for all } \ell \in H^{-1}(\omega_z).$$

Regarding the above discussion, we have that conditions (3.14f), (3.14e) and (3.14a) are equivalent to (3.13); cf. §3.7. Conditions (3.14d) and (3.14b) allow to quantify

the local dual norms of the approximate residual $\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}} \in \underline{\mathbb{D}}(\mathcal{M})$ in a computable manner; compare also with §3.6 below.

In the next three sections we construct two operators $\mathcal{P}_{\mathcal{M}}$ fulfilling (3.14).

3.4. Discretized residuals and a locally stable biorthogonal system. We present a possible choice of the set $\underline{\mathbb{D}}(\mathcal{M})$ of discretized residuals and introduce an associated biorthogonal system, which is instrumental in constructing a suitable operator $\mathcal{P}_{\mathcal{M}}$ with range $\underline{\mathbb{D}}(\mathcal{M})$.

We set

$$(3.15) \quad \underline{\mathbb{D}}(\mathcal{M}) := \{\ell \in H^{-1}(\Omega) \mid \langle \ell, v \rangle = \sum_{K \in \mathcal{M}} \int_K c_K v \, dx + \sum_{F \in \mathcal{F}} \int_F c_F v \, ds$$

$$\text{for all } v \in H_0^1(\Omega) \text{ with } c_K, c_F \in \mathbb{R} \text{ for } K \in \mathcal{M}, F \in \mathcal{F}\}.$$

Every functional $\ell \in \underline{\mathbb{D}}(\mathcal{M})$ is thus constant on each element and on each face. Obviously, condition (3.14a) is verified. More precisely, $\underline{\mathbb{D}}(\mathcal{M})$ is a strict superset of $\Delta(\mathbb{V}_0(\mathcal{M}))$, since in $\Delta(\mathbb{V}_0(\mathcal{M}))$ only certain linear combinations of the constants c_F , $F \in \mathcal{F}$ are allowed. The fact that these constants are independent in $\underline{\mathbb{D}}(\mathcal{M})$ facilitates the definition of $\mathcal{P}_{\mathcal{M}}$. Moreover, we have added the contributions given by the constants c_K , $K \in \mathcal{M}$, for comparability with the classical oscillations and a posteriori error estimators and because similar contributions will appear for higher order elements; cf. Kreuzer and Veeser [15]. In spite of these enlargements, we still have $\dim \underline{\mathbb{D}}(\mathcal{M}) < \infty$. Consequently, an argument as in the proof of Lemma 2, which hinges on infinite dimension, is ruled out.

Let us associate a biorthogonal system with $\underline{\mathbb{D}}(\mathcal{M})$. To this end, we introduce the surface Dirac distributions

$$(3.16a) \quad \chi_F : \begin{cases} H_0^1(\Omega) & \rightarrow \mathbb{R}, \\ v & \mapsto \int_F v \, ds, \end{cases} \quad F \in \mathcal{F},$$

and we identify the characteristic functions χ_K , $K \in \mathcal{M}$, with their associated distributions

$$(3.16b) \quad \chi_K : \begin{cases} H_0^1(\Omega) & \rightarrow \mathbb{R}, \\ v & \mapsto \int_K v \, dx, \end{cases} \quad K \in \mathcal{M}.$$

Notice that the definitions of χ_F and χ_K involve different measures for integration: the $(d-1)$ -dimensional Hausdorff measure for χ_F and the d -dimensional Lebesgue measure for χ_K . Correspondingly, each χ_K is absolutely continuous and each χ_F is singular with respect to the d -dimensional Lebesgue measure.

We collect all elements and interelement faces in the index set $\mathcal{I} = \mathcal{I}(\mathcal{M}) := \mathcal{M} \cup \mathcal{F}$ and derive in the next lemma the properties of the functionals χ_i , $i \in \mathcal{I}$, that are of interest to us.

Lemma 6 (Basis and scaling). *The functionals χ_i , $i \in \mathcal{I}$, are a basis of $\underline{\mathbb{D}}(\mathcal{M})$. For any element $K \in \mathcal{M}$ and any face $F \in \mathcal{F}$ containing a vertex $z \in \mathcal{V}$, we have*

$$\|\chi_K\|_{H^{-1}(\omega_z)} \leq |K|^{1/2} \tilde{\rho}_z \quad \text{and} \quad \|\chi_F\|_{H^{-1}(\omega_z)} \leq |F|^{1/2} \tilde{\rho}_z^{1/2}$$

with $\tilde{\rho}_z$ from (2.7).

Proof. We will use the Friedrichs inequality

$$(3.17) \quad \forall v \in H_0^1(\omega_z) \quad \|v\|_{\omega_z} \leq \tilde{\rho}_z \|\nabla v\|_{\omega_z}$$

and the following trace theorem: if $F \in \mathcal{F}$ with $F \ni z$ and \mathbf{n} denotes a normal of F , then

$$(3.18) \quad \forall w \in W_0^{1,1}(\omega_z) \quad \|w\|_{L^1(F)} \leq \frac{1}{2} \|\nabla w \cdot \mathbf{n}\|_{L^1(\omega_z)}.$$

Given $K \in \mathcal{M}$ with $K \ni z$ and any $v \in H_0^1(\omega_z)$, the Cauchy-Schwarz inequality and (3.17) yield

$$|\langle \chi_K, v \rangle| = \left| \int_K v \, dx \right| \leq |K|^{1/2} \|v\|_{\omega_z} \leq |K|^{1/2} \tilde{\rho}_z \|\nabla v\|_{\omega_z},$$

which verifies the first claimed inequality. To show the second one, fix $F \in \mathcal{F}$ with $F \ni z$ and let again $v \in H_0^1(\omega_z)$. Using (3.18) with $w = v^2$ and then again (3.17), we derive

$$|\langle \chi_F, v \rangle| = \left| \int_F v \, ds \right| \leq |F|^{1/2} \|v\|_F \leq |F|^{1/2} \|v\|_{\omega_z}^{1/2} \|\nabla v \cdot \mathbf{n}\|_{\omega_z}^{1/2} \leq |F|^{1/2} \tilde{\rho}_z^{1/2} \|\nabla v\|_{\omega_z}$$

and also the second claimed inequality is proved. \square

In order to complete the basis of Lemma 6 to a biorthogonal system, we use the following test functions: Given any element $K \in \mathcal{M}$, take

$$(3.19a) \quad \psi_K = \frac{(2d+1)!}{d!|K|} \prod_{z \in \mathcal{V} \cap K} \phi_z.$$

Given any interelement face $F \in \mathcal{F}$, let z_i , $i = 1, 2$, be the vertices in the patch ω_F , see (2.3), that are opposite to F and set

$$(3.19b) \quad \psi_F := \frac{(2d-1)!}{(d-1)!|F|} \left(\prod_{z \in \mathcal{V} \cap F} \phi_z \right) \left(1 - (2d+1) \sum_{i=1}^2 \phi_{z_i} \right).$$

Let us verify that the basis χ_i , $i \in \mathcal{I}$ and the test functions ψ_i , $i \in \mathcal{I}$, actually form a biorthogonal system with a crucial stability condition.

Lemma 7 (Locally stable biorthogonal system). *Together with the basis χ_i , $i \in \mathcal{I}$, the test functions ψ_i , $i \in \mathcal{I}$, form a locally stable biorthogonal system:*

(i) *We have*

$$\forall i, j \in \mathcal{I} \quad \langle \chi_i, \psi_j \rangle = \delta_{ij}.$$

(ii) *Let $\mathcal{I}_z := \{i \in \mathcal{I} \mid i \ni z\}$ denote the elements and faces containing a vertex $z \in \mathcal{V}$. Then*

$$\forall i \in \mathcal{I}_z \quad \|\chi_i\|_{H^{-1}(\omega_z)} \|\nabla \psi_i\|_{\omega_z} \leq C_\psi,$$

where the stability constant C_ψ only depends on d and the shape coefficient $\sigma(\mathcal{M})$.

Proof. To show (i), we consider the cases of elements $j \in \mathcal{M}$ and faces $j \in \mathcal{F}$ separately. First, let $K \in \mathcal{M}$ be an element. As already seen in (3.8), we have $\langle \chi_K, \psi_K \rangle = \int_K \psi_K = 1$. Moreover, since $\psi_K = 0$ in $\Omega \setminus \overset{\circ}{K}$, we infer $\langle \chi_{K'}, \psi_K \rangle = 0$ for any $K' \in \mathcal{M} \setminus \{K\}$ and $\langle \chi_F, \psi_K \rangle = 0$ for any $F \in \mathcal{F}$.

Second, fix a face $F \in \mathcal{F}$. Using (2.10), we obtain

$$\langle \chi_F, \psi_F \rangle = \frac{(2d-1)!}{(d-1)!|F|} \int_F \prod_{z \in \mathcal{V} \cap F} \phi_z \, ds = 1.$$

From $\psi_F = 0$ in $\Omega \setminus \overset{\circ}{\omega}_F$, where ω_F is the patch of the two elements containing the face F , we infer $\langle \chi_{F'}, \psi_F \rangle = 0$ for any $F' \in \mathcal{F} \setminus \{F\}$ and $\langle \chi_K, \psi_F \rangle = 0$ for any $K \in \mathcal{M}$ with $K \not\ni F$. Last, let $K \in \mathcal{M}$ such that $K \supset F$. Using again (2.10), we deduce

$$\langle \chi_K, \psi_F \rangle = \frac{(2d-1)!}{(d-1)!|F|} \left(\int_K \prod_{z \in \mathcal{V} \cap F} \phi_z \, dx - (2d+1) \int_K \prod_{z \in \mathcal{V} \cap K} \phi_z \, dx \right) = 0.$$

For (ii), we again treat elements and faces separately. Let $K \in \mathcal{M}$ be an element containing z . The well-known inverse estimate $\|\nabla \psi_K\|_K \leq C_d \rho_K^{-1} \|\psi_K\|_K$, $K \subset \omega_z$ and (2.10) imply

$$\|\nabla \psi_K\|_{\omega_z} = \|\nabla \psi_K\|_K \leq \frac{C_d}{|K|^{1/2} \rho_K}.$$

Combining this with the first inequality in Lemma 6 and (2.8), we obtain the claimed inequality for elements:

$$\|\chi_K\|_{H^{-1}(\omega_z)} \|\nabla \psi_K\|_{\omega_z} \leq C_d \frac{\tilde{\rho}_z}{\rho_K} \leq C_{d;\sigma(\mathcal{M})}.$$

Let $F \in \mathcal{F}$ be an interelement face containing z and write $F = K_1 \cap K_2$, where $K_1, K_2 \in \mathcal{M}$ are the two elements containing F . Proceeding as before, we deduce

$$(3.20) \quad \|\nabla \psi_F\|_{\omega_z}^2 = \sum_{n=1,2} \|\nabla \psi_F\|_{K_n}^2 \leq C_d^2 \sum_{n=1,2} \rho_{K_n}^{-2} \|\psi_F\|_{K_n}^2 \leq C_d^2 \sum_{n=1,2} \frac{|K_n|}{|F|^2 \rho_{K_n}^2}.$$

and

$$\|\chi_F\|_{H^{-1}(\omega_z)} \|\nabla \psi_F\|_{\omega_z} \leq C_d \left(\sum_{i=1,2} \frac{h_{K_n;F} \tilde{\rho}_z}{\rho_{K_n}^2} \right)^{1/2} \leq C_{d;\sigma(\mathcal{M})}. \quad \square$$

In what follows, we shall rely only on the properties of the test functions ψ_i , $i \in \mathcal{I}$ expressed in Lemma 7. In other words: what counts is not their special form, but the fact that they form a stable biorthogonal system with the basis χ_i , $i \in \mathcal{I}$, of $\mathbb{D}(\mathcal{M})$.

3.5. Construction and properties of $\mathcal{P}_{\mathcal{M}}$. We now propose a possible choice for the projection operator $\mathcal{P}_{\mathcal{M}}$ and verify the desired properties (3.14). Set

$$(3.21) \quad \mathcal{P}_{\mathcal{M}} \ell = \sum_{i \in \mathcal{I}} \langle \ell, \psi_i \rangle \chi_i,$$

where the functionals χ_i , $i \in \mathcal{I}$, are given by (3.16) and the test functions ψ_i , $i \in \mathcal{I}$, by (3.19). Clearly, $\mathcal{P}_{\mathcal{M}}$ is linear and $\mathcal{P}_{\mathcal{M}} f$ is locally computable in terms of a finite number of evaluations of f , i. e., we have (3.14c) and (3.14d).

The biorthogonality of these functionals and test functions implies the following local counterparts of the algebraic condition (3.10).

Theorem 8 (Local invariance). *For any functional $\ell \in H^{-1}(\Omega)$, element $K \in \mathcal{M}$, and side $F \in \mathcal{F}$, the operator $\mathcal{P}_{\mathcal{M}}$ does not change the following discrete restrictions:*

- (i) *If $\ell \in \mathbb{D}(\mathcal{M})$ on \mathring{K} , then $\mathcal{P}_{\mathcal{M}} \ell = \ell$ on \mathring{K} .*
- (ii) *If $\ell \in \mathbb{D}(\mathcal{M})$ on $\mathring{\omega}_F$, then $\mathcal{P}_{\mathcal{M}} \ell = \ell$ on $\mathring{\omega}_F$.*

Proof. Let $\ell = c \chi_K$ on \mathring{K} with $c \in \mathbb{R}$. For any $i \in \mathcal{I}$, we have $\langle \ell, \psi_i \rangle = c \int_K \psi_i = c \delta_{K,i}$ by means of Lemma 7 (i). Consequently, $\mathcal{P}_{\mathcal{M}} \ell = c \chi_K$ on \mathring{K} , which proves (i).

To show (ii), let $K_1, K_2 \in \mathcal{M}$ be the two elements containing F and let $\ell = c \chi_F + \sum_{i=1,2} c_i \chi_{K_i}$ on $\mathring{\omega}_F$ with $c, c_1, c_2 \in \mathbb{R}$. Using again Lemma 7 (i), we observe

$$\langle \ell, \psi_F \rangle = c \langle \chi_F, \psi_F \rangle + \sum_{i=1,2} c_i \langle \chi_{K_i}, \psi_F \rangle = c \quad \text{and} \quad \langle \ell, \psi_{K_i} \rangle = c_i \quad \text{for } i = 1, 2$$

and $\langle \ell, \psi_i \rangle = 0$ for all $i \in \mathcal{I} \setminus \{F, K_1, K_2\}$. Consequently,

$$\mathcal{P}_{\mathcal{M}} \ell = c \chi_F + \sum_{i=1,2} c_i \chi_{K_i} = \ell \quad \text{on } \mathring{\omega}_F$$

and also (ii) is verified. \square

Theorem 8 implies in particular (3.14e). Moreover, it has the following global consequences.

Corollary 9 (Global invariance). *The operator $\mathcal{P}_{\mathcal{M}}$ is a linear projection onto the discretized residuals $\mathbb{D}(\mathcal{M})$ from (3.15). In particular, we have*

$$\mathcal{P}_{\mathcal{M}}(\Delta V) = \Delta V \quad \text{and} \quad \mathcal{P}_{\mathcal{M}}(f) = f.$$

for any $V \in \mathbb{V}_0(\mathcal{M})$ and any \mathcal{M} -piecewise constant function $f \in \mathbb{P}_0(\mathcal{M})$.

Next, we verify the local stability (3.14f) of $\mathcal{P}_{\mathcal{M}}$. As a side product, we also obtain the local stability of the operator $\tilde{\mathcal{P}}_{0,\mathcal{M}}$, which was left open in §3.3.

Theorem 10 (Local stability). *The linear projection $\mathcal{P}_{\mathcal{M}}$ is locally H^{-1} -stable: for any functional $\ell \in H^{-1}(\Omega)$ and any vertex $z \in \mathcal{V}$, we have*

$$\|\mathcal{P}_{\mathcal{M}}\ell\|_{H^{-1}(\omega_z)} \lesssim \|\ell\|_{H^{-1}(\omega_z)},$$

where the hidden constant depends only on d and $\sigma(\mathcal{M})$.

Proof. Given $v \in H_0^1(\omega_z)$, we derive

$$\begin{aligned} |\langle \mathcal{P}_{\mathcal{M}}\ell, v \rangle| &\leq \sum_{i \in \mathcal{I}_z} |\langle \ell, \psi_i \rangle \langle \chi_i, v \rangle| \leq \sum_{i \in \mathcal{I}_z} \|\ell\|_{H^{-1}(\omega_z)} \|\nabla \psi_i\|_{\omega_z} \|\chi_i\|_{H^{-1}(\omega_z)} \|\nabla v\|_{\omega_z} \\ &\lesssim \|\ell\|_{H^{-1}(\omega_z)} \|\nabla v\|_{\omega_z}, \end{aligned}$$

where we used Lemma 7 (ii) and $\#\mathcal{I}_z \lesssim 1$. \square

Remark 11 (Stability of $\tilde{\mathcal{P}}_{0,\mathcal{M}}$). The argument in the proof of Theorem 10 also shows that $\tilde{\mathcal{P}}_{0,\mathcal{M}}$ is locally H^{-1} -stable. In fact, one simply replaces $\mathcal{P}_{\mathcal{M}}$ by $\tilde{\mathcal{P}}_{0,\mathcal{M}}$ and the index set \mathcal{I}_z by $\mathcal{I}_z \cap \mathcal{M}$.

Let us conclude this section with the following further remarks on the linear projection $\mathcal{P}_{\mathcal{M}}$.

Remark 12 (Orthogonality). For any $\ell \in H^{-1}(\Omega)$, the functional $\ell - \mathcal{P}_{\mathcal{M}}\ell$ is orthogonal to $\text{span}\{\psi_i \mid i \in \mathcal{I}\}$. This is an immediate consequence of Lemma 7 (i).

Remark 13 (Adjoint of $\mathcal{P}_{\mathcal{M}}$). Formally, the adjoint of $\mathcal{P}_{\mathcal{M}}$ is given by

$$\mathcal{P}_{\mathcal{M}}^*v = \sum_{i \in \mathcal{I}} \langle \chi_i, v \rangle \psi_i, \quad v \in H_0^1(\Omega).$$

Here Lemma 7 (i) implies

$$(3.22) \quad \int_K \mathcal{P}_{\mathcal{M}}^*v = \int_K v \quad \text{and} \quad \int_F \mathcal{P}_{\mathcal{M}}^*v = \int_F v$$

for all elements $K \in \mathcal{M}$ and interelement faces $F \in \mathcal{F}$. The operator $\mathcal{P}_{\mathcal{M}}^*$ and these conditions, which characterize it, were used in Veiser [23] to derive an a posteriori error upper bound in terms of a hierarchical estimator. That argument, as well as Morin, Nochetto, and Siebert [18, Theorem 3.6] and Verfürth [24, (3.14)], is closely related to Theorem 15 below.

3.6. Required a priori information, an alternative to $\mathcal{P}_{\mathcal{M}}$, and quantification of the discretized residual. The purpose of this section is twofold. First, we illustrate which type of a priori information on f in (2.2) is needed to carry out our approach, presenting also a possible alternative to $\mathcal{P}_{\mathcal{M}}$. Second, we show that a stable biorthogonal system is not only useful to construct $\mathcal{P}_{\mathcal{M}}$, but also to quantify the local dual norms of discretized residuals.

Clearly, the operator $\mathcal{P}_{\mathcal{M}}$ of §3.5 can be applied to the right-hand side f of (2.2) whenever

$$(3.23) \quad \langle f, \psi_i \rangle, \quad i \in \mathcal{I}, \quad \text{are known exactly.}$$

In order to ensure a meaningful discretized residual, this information goes beyond (2.14), the information necessary for the Galerkin approximation (2.11) on the mesh \mathcal{M} ; it is available, e.g., when one is able to compute the counterpart of (2.11) of order $d+1$ over \mathcal{M} .

There are other possibilities to obtain a meaningful discretized residual. The following one fits particularly well to (2.14) in the context of mesh adaptivity. Suppose that we are given an initial mesh and a refinement procedure such that the set \mathbb{M} of all refined meshes form a shape-regular family. Furthermore, suppose that, for

any mesh $\mathcal{M} \in \mathbb{M}$, there is a refinement $\widetilde{\mathcal{M}} \in \mathbb{M}$ with vertices $\mathcal{V}(\widetilde{\mathcal{M}})$ that satisfies the following properties:

$$(3.24a) \quad \forall \widetilde{K} \in \widetilde{\mathcal{M}} \exists K \in \mathcal{M} \quad \text{with} \quad \widetilde{K} \subset K \text{ and } h_K \lesssim h_{\widetilde{K}},$$

$$(3.24b) \quad \forall i \in \mathcal{I}(\mathcal{M}) \exists \widetilde{z} \in \mathcal{V}(\widetilde{\mathcal{M}}) \quad \text{such that} \quad \widetilde{z} \text{ is interior to } i.$$

Let us now fix a mesh $\mathcal{M} \in \mathbb{M}$ and a refinement $\widetilde{\mathcal{M}} \in \mathbb{M}$ satisfying (3.24). For any $i \in \mathcal{I}(\mathcal{M})$, using (3.24b), we fix a vertex $\widetilde{z} \in \mathcal{V}(\widetilde{\mathcal{M}})$ interior to i and denote by $\widetilde{\phi}_z$ its associated hat function in $\mathbb{V}(\widetilde{\mathcal{M}})$. We then obtain counterparts $\widetilde{\psi}_i$, $i \in \mathcal{I}$, of the test functions ψ_i , $i \in \mathcal{I}$, by using these hat functions with a suitable scaling in place of the element and faces bubble functions in (3.19) such that the following lemma holds. We skip the technical details, referring to Morin, Nocketto and Siebert [17] and Veeseer [23].

Lemma 14 (Another locally stable biorthogonal system). *Together with the basis χ_i , $i \in \mathcal{I}$, the test functions $\widetilde{\psi}_i$, $i \in \mathcal{I}$, form a locally stable biorthogonal system:*

(i) *We have*

$$\forall i, j \in \mathcal{I} \quad \langle \chi_i, \widetilde{\psi}_j \rangle = \delta_{ij}.$$

(ii) *Let $\mathcal{I}_z = \{i \in \mathcal{I} \mid i \ni z\}$ denote the elements and faces containing a vertex $z \in \mathcal{V}$. Then*

$$\forall i \in \mathcal{I}_z \quad \|\chi_i\|_{H^{-1}(\omega_z)} \|\nabla \widetilde{\psi}_i\|_{\omega_z} \leq C_{\widetilde{\psi}},$$

where the stability constant $C_{\widetilde{\psi}}$ only depends on d and the shape coefficient $\sigma(\mathcal{M})$.

Thus, the operator

$$(3.25) \quad \widetilde{\mathcal{P}}_{\mathcal{M}} \ell := \sum_{i \in \mathcal{I}} \langle \ell, \widetilde{\psi}_i \rangle \chi_i$$

defines an alternative to $\mathcal{P}_{\mathcal{M}}$ and the properties (3.14) without (3.14b) can be established as for $\mathcal{P}_{\mathcal{M}}$. The operator $\widetilde{\mathcal{P}}_{\mathcal{M}}$ can be evaluated on any mesh $\mathcal{M} \in \mathbb{M}$ whenever

$$(3.26) \quad \forall \widetilde{\mathcal{M}} \in \mathbb{M} \forall z \in \mathcal{V}_0(\widetilde{\mathcal{M}}) \quad \langle f, \widetilde{\phi}_z \rangle \text{ are known exactly,}$$

where $\{\widetilde{\phi}_z\}_{z \in \mathcal{V}_0(\widetilde{\mathcal{M}})}$ denotes the nodal basis of $\mathbb{V}_0(\widetilde{\mathcal{M}})$. This is exactly (2.14) for all meshes in \mathbb{M} . Consequently, it is also needed to ensure that an adaptive algorithm with the above refinement procedure can always compute the Galerkin approximation (2.11).

Let us now turn to the quantification of the discretized residual and verify (3.14b), considering a general locally stable biorthogonal system.

Theorem 15 (Quantifying local dual norms). *Let ψ_i , $i \in \mathcal{I}$, be the test functions from Lemma 7 or Lemma 14. If $\ell \in \mathbb{D}(\mathcal{M})$ on a star ω_z , then the corresponding local dual norm can be quantified by a finite number of evaluations:*

$$\frac{1}{d+1} \sum_{i \in \mathcal{I}_z} \left| \left\langle \ell, \frac{\psi_i}{\|\nabla \psi_i\|} \right\rangle \right|^2 \leq \|\ell\|_{H^{-1}(\omega_z)}^2 \lesssim \sum_{i \in \mathcal{I}_z} \left| \left\langle \ell, \frac{\psi_i}{\|\nabla \psi_i\|} \right\rangle \right|^2$$

where the hidden constant depends on d , $\sigma(\mathcal{M})$, and C_{ψ} .

Proof. Let us first prove the lower bound, which holds for any arbitrary functional $\ell \in H^{-1}(\Omega)$. In fact, the definition of the dual norm readily yields

$$(3.27) \quad \left| \left\langle \ell, \frac{\psi_i}{\|\nabla \psi_i\|} \right\rangle \right| \leq \|\ell\|_{H^{-1}(\text{supp } \psi_i)}$$

for any $i \in \mathcal{I}_z$. Notice that the essential supremum of $x \mapsto \#\{i \in \mathcal{I}_z \mid \text{supp } \psi_i \ni x\}$ is bounded by $d + 1$. Arguing as in the proof of Lemma 4 (ii), we therefore obtain

$$(3.28) \quad \sum_{i \in \mathcal{I}_z} \|\ell\|_{H^{-1}(\text{supp } \psi_i)}^2 \leq (d + 1) \|\ell\|_{H^{-1}(\omega_z)}^2$$

and the proof of the lower bound is finished.

To show the upper bound, we (need to) assume that $\ell \in \mathbb{D}(\mathcal{M})$ on ω_z . Given $v \in H_0^1(\omega_z)$, we can then write

$$\langle \ell, v \rangle = \sum_{i \in \mathcal{I}_z} c_i \langle \chi_i, v \rangle \quad \text{with } c_i \in \mathbb{R}.$$

In light of the biorthogonality, we have $c_i = \langle \ell, \psi_i \rangle$. Using also the local stability of the biorthogonal system, we infer

$$\begin{aligned} |\langle \ell, v \rangle| &\leq \sum_{i \in \mathcal{I}_z} |\langle \ell, \psi_i \rangle \langle \chi_i, v \rangle| \\ &\leq \sum_{i \in \mathcal{I}_z} \|\nabla \psi_i\|_{\omega_z} \|\chi_i\|_{H^{-1}(\omega_z)} \left| \left\langle \ell, \frac{\psi_i}{\|\nabla \psi_i\|} \right\rangle \right| \|\nabla v\|_{\omega_z} \\ &\leq C_\psi \left(\sum_{i \in \mathcal{I}_z} \left| \left\langle \ell, \frac{\psi_i}{\|\nabla \psi_i\|} \right\rangle \right| \right) \|\nabla v\|_{\omega_z}. \end{aligned}$$

Since the solid angle of every simplex containing z is bounded away from 0 in terms of d and the shape coefficient $\sigma(\mathcal{M})$, we have $\#\mathcal{I}_z \leq C_{\sigma(\mathcal{M})}$. Consequently, the Cauchy-Schwarz inequality on the sum implies the desired upper bound. \square

Theorem 15 implies the missing (3.14b) for both operators $\mathcal{P}_\mathcal{M}$ and $\tilde{\mathcal{P}}_\mathcal{M}$ and, in accordance with §3.3, we have splittings of the local residual norms with the desired properties. Notice that, in view of the discussion of this section and Corollary 5, bounding the terms

$$\|\mathcal{P}_\mathcal{M}f - f\|_{H^{-1}(\omega_z)} \quad \text{or} \quad \|\tilde{\mathcal{P}}_\mathcal{M}f - f\|_{H^{-1}(\omega_z)}$$

cannot be done in general with a finite number of evaluations of the load f . Notably, these terms involve only the load and the discretized residuals

$$\|\mathcal{P}_\mathcal{M}f + \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)} \quad \text{or} \quad \|\tilde{\mathcal{P}}_\mathcal{M}f + \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)}$$

can be quantified with finite information, which, in light of Remark 3, is less than the information required for evaluating local L^2 -norms of the load f .

3.7. A posteriori error bounds. We now summarize our preceding results by deriving a posteriori error bounds. The resulting bounds are defined for any load $f \in H^{-1}(\Omega)$ and the oscillation is dominated by the error.

The following statements remain correct if $\mathcal{P}_\mathcal{M}$ is replaced by $\tilde{\mathcal{P}}_\mathcal{M}$ from (3.25).

Theorem 16 (Abstract upper bound). *For any functional $f \in H^{-1}(\Omega)$ and any conforming mesh \mathcal{M} , we have*

$$\|\nabla(u_f - U_{f;\mathcal{M}})\|^2 \lesssim \sum_{z \in \mathcal{V}} \|\mathcal{P}_\mathcal{M}f + \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)}^2 + \|\mathcal{P}_\mathcal{M}f - f\|_{H^{-1}(\omega_z)}^2.$$

Each local dual norm $\|\mathcal{P}_\mathcal{M}f - \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)}$ of the discretized residual can be quantified with a finite number of evaluations of f , while the quantification of the local dual norms $\|\mathcal{P}_\mathcal{M}f - f\|_{H^{-1}(\omega_z)}$ of the oscillation requires additional a priori information on f .

Proof. Lemma 1, Lemma 4 and a triangle inequality imply the claimed bound. Recalling that

$$\mathcal{P}_\mathcal{M}f + \Delta U_{f;\mathcal{M}} \in \mathbb{D}(\mathcal{M}),$$

Lemma 15 and Corollary 5 ensure the statements about the quantification of the two parts of the bound. \square

In contrast to previous results available in literature, the complete upper bound in Theorem 16 is also a lower bound, even locally.

Theorem 17 (Abstract local lower bounds). *For any functional $f \in H^{-1}(\Omega)$ and any conforming mesh \mathcal{M} , the discretized residual and the oscillation are locally dominated by the error: for every vertex $z \in \mathcal{V}$, we have*

$$\|\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)} \lesssim \|\nabla(u_f - U_{f;\mathcal{M}})\|_{\omega_z}$$

and

$$\|\mathcal{P}_{\mathcal{M}}f - f\|_{H^{-1}(\omega_z)} \lesssim \|\nabla(u_f - U_{f;\mathcal{M}})\|_{\omega_z}.$$

Proof. In light of (3.12), the first claimed inequality follows from the triangle inequality and the second one. The latter is a consequence of Theorems 8 and 10 and (3.12):

$$\begin{aligned} \|\mathcal{P}_{\mathcal{M}}f - f\|_{H^{-1}(\omega_z)} &\leq \|\mathcal{P}_{\mathcal{M}}(f + \Delta U_{f;\mathcal{M}})\|_{H^{-1}(\omega_z)} + \|f - \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)} \\ &\lesssim \|f - \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)} \lesssim \|\nabla(u_f - U_{f;\mathcal{M}})\|_{\omega_z}. \end{aligned} \quad \square$$

Squaring and summing, we readily get the global lower bounds.

Corollary 18 (Abstract global lower bounds). *For any functional $f \in H^{-1}(\Omega)$ and any conforming mesh \mathcal{M} , the discretized residual and the oscillation are globally dominated by the error in that*

$$\sum_{z \in \mathcal{V}} \|\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)}^2 \lesssim \|\nabla(u_f - U_{f;\mathcal{M}})\|^2$$

and

$$\sum_{z \in \mathcal{V}} \|\mathcal{P}_{\mathcal{M}}f - f\|_{H^{-1}(\omega_z)}^2 \lesssim \|\nabla(u_f - U_{f;\mathcal{M}})\|^2.$$

To summarize: if we are able to quantify the oscillation terms $\|\mathcal{P}_{\mathcal{M}}f - f\|_{H^{-1}(\omega_z)}$, $z \in \mathcal{V}$, then the right-hand side in Theorem 16 is a truly equivalent a posteriori error estimator.

Remark 19 (Surrogate oscillation). The quantification of the local dual norms $\|\mathcal{P}_{\mathcal{M}}f - f\|_{H^{-1}(\omega_z)}$, $z \in \mathcal{V}$, of the oscillation appears to be a difficult matter. In [11, Section 7], Cohen, DeVore, and Nochetto consider similar terms for special f and resort to surrogates that can be approximated with the help of numerical integration. Those surrogates hinge on additional regularity of f , which entails the risk of overestimation; cf. Lemma 20 below.

3.8. Classical versus error-dominated oscillation. In this section we compare the error-dominated oscillation $(\sum_{z \in \mathcal{V}} \|\mathcal{P}_{\mathcal{M}}f - f\|_{H^{-1}(\omega_z)}^2)^{1/2}$ with the classical L^2 - and H^{-1} -oscillation,

$$\text{osc}_0(f, \mathcal{M}) \quad \text{and} \quad \min_{g \in \mathbb{P}_0(\mathcal{M})} \|f - g\|_{H^{-1}(\Omega)},$$

from (1.4) and (1.5) in the introduction. Doing so, we verify statements of the introduction and substantiate the advantages of the stability and invariance properties of the operator $\mathcal{P}_{\mathcal{M}}$.

Let us first show that the error-dominated oscillation is always smaller, up to a multiplicative constant, than both classical oscillations. To this end, let $f \in H^{-1}(\Omega)$ and let $g \in \mathbb{P}_0(\mathcal{M})$ be an arbitrary piecewise constant approximation over \mathcal{M} . The local invariance and stability properties of $\mathcal{P}_{\mathcal{M}}$ in Theorems 8 and 10 imply that, for all $z \in \mathcal{V}$,

$$(3.29) \quad \begin{aligned} \|f - \mathcal{P}_{\mathcal{M}}f\|_{H^{-1}(\omega_z)} &\leq \|f - g\|_{H^{-1}(\omega_z)} + \|\mathcal{P}_{\mathcal{M}}(g - f)\|_{H^{-1}(\omega_z)} \\ &\lesssim \|f - g\|_{H^{-1}(\omega_z)}. \end{aligned}$$

Combining this with Lemma 4 (ii) and minimizing over g , we obtain the bound in terms of the classical H^{-1} -oscillation:

$$(3.30a) \quad \sum_{z \in \mathcal{V}} \|f - \mathcal{P}_{\mathcal{M}} f\|_{H^{-1}(\omega_z)}^2 \lesssim \min_{g \in \mathbb{P}_0(\mathcal{M})} \|f - g\|_{H^{-1}(\Omega)}^2.$$

To show the other bound, suppose $f \in L^2(\Omega)$. Making use of the orthogonality of $P_{0,\mathcal{M}}$ and Poincaré inequalities in the elements of ω_z , we deduce

$$\|f - P_{0,\mathcal{M}} f\|_{H^{-1}(\omega_z)}^2 \lesssim \sum_{K \subset \omega_z} h_K^2 \|f - P_{0,\mathcal{M}} f\|_K^2,$$

which together with (3.29) gives the bound in terms of the L^2 -oscillation:

$$(3.30b) \quad \sum_{z \in \mathcal{V}} \|f - \mathcal{P}_{\mathcal{M}} f\|_{H^{-1}(\omega_z)}^2 \lesssim \sum_{K \in \mathcal{M}} h_K^2 \|f - P_{0,\mathcal{M}} f\|_K^2 = \text{osc}_0(f, \mathcal{M})^2.$$

The converse bounds of (3.30) do not hold. For the classical L^2 -oscillation, this applies even on a fixed mesh and is in particular due to stability issues. The following lemma provides an illustration, relating directly to the error instead of the error-dominated oscillation.

Lemma 20 (Overestimation of classical L^2 -oscillation). *For any conforming mesh \mathcal{M} , there exists a sequence $(f_k)_k \subset L^2(\Omega)$ such that*

$$\frac{\text{osc}_0(f_k, \mathcal{M})}{\|\nabla(u_{f_k} - U_{f_k;\mathcal{M}})\|} \rightarrow \infty \quad \text{as } k \rightarrow \infty.$$

Proof. Choose $f \in H^{-1}(\Omega) \setminus L^2(\Omega)$. Since $L^2(\Omega)$ is dense in $H^{-1}(\Omega)$, there exists a sequence $(f_k)_k \subset L^2(\Omega)$ such that $f_k \rightarrow f$ in $H^{-1}(\Omega)$. On the one hand, the energy norm errors $\|\nabla(u_{f_k} - U_{f_k;\mathcal{M}})\|$ are uniformly bounded with respect to k . On the other hand, in view of $\lim_{k \rightarrow \infty} \|f_k\|_{L^2(\Omega)} = \infty$, the oscillation $\text{osc}_0(f_k, \mathcal{M})$ becomes arbitrarily large for $k \rightarrow \infty$. \square

In the case of the classical H^{-1} -oscillation, (3.30a) cannot be inverted because of invariance issues. Let us illustrate this again by the relationship to the Galerkin error. Consider

$$(3.31) \quad f = -\Delta V \text{ for some } V \in \mathbb{V}_0(\mathcal{M}^\dagger) \setminus \{0\},$$

where \mathcal{M}^\dagger is some conforming simplicial mesh of Ω . For any conforming refinement \mathcal{M} of \mathcal{M}^\dagger , we then have $u_f = V = U_{f;\mathcal{M}}$ and $f \notin \mathbb{P}_0(\mathcal{M})$. Hence

$$\|\nabla(u_f - U_{f;\mathcal{M}})\| = 0 < \min_{g \in \mathbb{P}_0(\mathcal{M})} \|f - g\|_{H^{-1}(\Omega)},$$

where the classical H^{-1} -oscillation can be made arbitrarily large for a given \mathcal{M} but decreases to 0 under suitable refinement. One could argue that the (neighborhoods of the) loads (3.31) are very special, in particular because the optimal convergence rate of (3.31) is formally ∞ . Here is another example based upon Cohen, DeVore, and Nochetto [11, Section 6.4], where the optimal nonlinear convergence rate for the error is finite and often encountered in practice.

Lemma 21 (Another overestimation of classical H^{-1} -oscillation). *Let $\Omega = (0, 1)^2$. There is a functional $f \in H^{-1}(\Omega)$ and a sequence $(L_n)_n$ with $\log n \gtrsim L_n \rightarrow \infty$ as $n \rightarrow \infty$ such that*

$$(3.32a) \quad \min_{\#\mathcal{M} \leq n} \|\nabla(u_f - U_{f;\mathcal{M}})\| \lesssim n^{-1/2},$$

and

$$(3.32b) \quad \min_{\#\mathcal{M} \leq n} \min_{g \in \mathbb{P}_0(\mathcal{M})} \left(\sum_{z \in \mathcal{V}(\mathcal{M})} \|f - g\|_{H^{-1}(\omega_z)}^2 \right)^{1/2} \geq L_n n^{-1/2},$$

where \mathcal{M} varies in all meshes created by recursive or iterative newest vertex bisection of some conforming initial mesh \mathcal{M}_0 of Ω .

Proof. In [11, Section 6.4] Cohen, DeVore and Nocketto construct some function $u_f \in H_0^1(\Omega)$ and a sequence L_n as claimed for which (3.32a) and

$$(3.33) \quad \min_{\#\mathcal{M} \leq n} \left(\sum_{z \in \mathcal{V}(\mathcal{M})} \|f\|_{H^{-1}(\omega_z)}^2 \right)^{1/2} \geq L_n n^{-1/2}$$

hold. It thus remains to establish (3.32b). To this end, we fix temporarily an arbitrary vertex $z \in \mathcal{V}$ of a conforming mesh \mathcal{M} and let $g \in \mathbb{P}_0(\mathcal{M})$. The inverse triangle and (3.12) yield

$$\begin{aligned} \|f - g\|_{H^{-1}(\omega_z)} &\geq \|\Delta U_{f;\mathcal{M}} + g\|_{H^{-1}(\omega_z)} - \|f + \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)} \\ &\geq \|\Delta U_{f;\mathcal{M}} + g\|_{H^{-1}(\omega_z)} - \|\nabla(u_f - U_{f;\mathcal{M}})\|_{\omega_z}. \end{aligned}$$

By Lemma 7, we have, for all $K \in \mathcal{M}$,

$$\langle \Delta U_{f;\mathcal{M}}, \psi_K \rangle = \sum_{F \in \mathcal{F}} J(U_{f;\mathcal{M}})|_F \int_F \chi_F \psi_K \, ds = 0$$

and, for all $F \in \mathcal{F}$ and $K_1, K_2 \in \mathcal{M}$ with $K_1 \cap K_2 = F$,

$$\begin{aligned} \langle \Delta U_{f;\mathcal{M}} + g, \psi_F \rangle &= \int_F J(U_{f;\mathcal{M}}) \psi_F \, ds + \sum_{i=1,2} g|_{K_i} \int_{K_i} \chi_{K_i} \psi_F \, dx \\ &= \langle \Delta U_{f;\mathcal{M}}, \psi_F \rangle. \end{aligned}$$

Theorem 15 therefore implies

$$\begin{aligned} \|\Delta U_{f;\mathcal{M}} + g\|_{H^{-1}(\omega_z)} &\gtrsim \sum_{i \in \mathcal{I}_z \cap \mathcal{F}} \left| \left\langle \Delta U_{f;\mathcal{M}} + g, \frac{\psi_i}{\|\nabla \psi_i\|} \right\rangle \right| \\ &= \sum_{i \in \mathcal{I}_z \cap \mathcal{F}} \left| \left\langle \Delta U_{f;\mathcal{M}}, \frac{\psi_i}{\|\nabla \psi_i\|} \right\rangle \right| \\ &= \sum_{i \in \mathcal{I}_z} \left| \left\langle \Delta U_{f;\mathcal{M}}, \frac{\psi_i}{\|\nabla \psi_i\|} \right\rangle \right| \gtrsim \|\Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)}. \end{aligned}$$

Exploiting also Lemma 4, we arrive at

$$\begin{aligned} &\left(\sum_{z \in \mathcal{V}} \|\Delta U_{f;\mathcal{M}} + g\|_{H^{-1}(\omega_z)}^2 \right)^{1/2} \\ &\quad \gtrsim \left(\sum_{z \in \mathcal{V}} \|\Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)}^2 \right)^{1/2} \\ &\quad \gtrsim \left(\sum_{z \in \mathcal{V}} \|f\|_{H^{-1}(\omega_z)}^2 \right)^{1/2} - \left(\sum_{z \in \mathcal{V}} \|f + \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)}^2 \right)^{1/2} \\ &\quad \geq \left(\sum_{z \in \mathcal{V}} \|f\|_{H^{-1}(\omega_z)}^2 \right)^{1/2} - C \|\nabla(u_f - U_{f;\mathcal{M}})\|. \end{aligned}$$

Consequently, (3.32a) and (3.33) lead to

$$\min_{\#\mathcal{M} \leq n} \min_{g \in \mathbb{P}_0(\mathcal{M})} \left(\sum_{z \in \mathcal{V}} \|f - g\|_{H^{-1}(\omega_z)}^2 \right)^{1/2} \geq (L_n - C) n^{-1/2},$$

which, upon redefining $(L_n)_n$, implies (3.32b) and the proof is finished. \square

Remark 22 (Overestimation of H^{-1} -variant of standard residual estimator). As pointed out by Cohen, DeVore, and Nochetto [11], the example of Lemma 21 entails that the right-hand side of

$$\|\nabla(u_f - U_{f;\mathcal{M}})\|^2 \lesssim \sum_{z \in \mathcal{V}(\mathcal{M})} \|\Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)}^2 + \|f\|_{H^{-1}(\omega_z)}^2,$$

a variant of the standard residual estimator defined for all loads $f \in H^{-1}(\Omega)$, is overestimating. In §4.2 below, we propose through our new approach another variant that is free of overestimation.

4. REALIZATIONS WITH CLASSICAL TECHNIQUES

The a posteriori error bounds in §3.7 are abstract in that they are given in terms of the local dual norms $\|\cdot\|_{H^{-1}(\omega_z)}$, $z \in \mathcal{V}$, of the discretized residual and the oscillation. For the norms $\|\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)}$, $z \in \mathcal{V}$, of the discretized residual, we required a quantification in terms of finite information on the load and provided a possible realization in Theorem 15. In this section we discuss a selection of alternative realizations. All realizations are motivated by classical approaches to a posteriori analysis and cover two explicit and two implicit techniques. It is worth making the following observations:

- Hierarchical estimators and estimators based upon local problems implicitly introduce a splitting of the residual like the one proposed in §3.3.
- The overestimation of the standard residual estimator in Remark 22 can be cured with the help of the splitting of the residual in §3.3.
- Employing different local dual norms, the approach of §3 can be extended to estimators based on flux equilibration.
- Each realization quantifies a local dual norm of the discretized residual by a computable, equivalent norm. Both equivalence and computability hinge on the finite-dimensional nature of the discretized residual.

4.1. An hierarchical estimator. Hierarchical estimators investigate the residual on an extension of the given finite element space. While higher order extensions were used originally, Bornemann, Erdmann, and Kornhuber show in [6] that an extension containing the functions

$$(4.1) \quad \lambda_K := \prod_{z \in \mathcal{V} \cap K} \phi_z, \quad K \in \mathcal{M}, \quad \text{and} \quad \lambda_F := \prod_{z \in \mathcal{V} \cap F} \phi_z, \quad F \in \mathcal{F},$$

already ensures reliability for piecewise constant loads $f \in \mathbb{P}_0(\mathcal{M})$. The indicators of a corresponding, ‘minimal’ hierarchical estimator are given by

$$\mathcal{E}_H(f, \mathcal{M}, i) := \left\langle \text{Res}(f; \mathcal{M}), \frac{\lambda_i}{\|\nabla \lambda_i\|} \right\rangle, \quad i \in \mathcal{I} = \mathcal{M} \cup \mathcal{F},$$

and computable in terms of $U_{f;\mathcal{M}}$ and the evaluations $\langle f, \lambda_i \rangle$, $i \in \mathcal{I}$. This definition implies the constant-free local lower bounds

$$\mathcal{E}_H(f, \mathcal{M}, i) \leq \|\text{Res}(f; \mathcal{M})\|_{H^{-1}(\text{supp } \lambda_i)}$$

and therefore, cf. (3.28), we have that, for every $z \in \mathcal{V}$ and $I_z = \{i \in \mathcal{I} \mid i \ni z\}$,

$$(4.2) \quad \left(\sum_{i \in I_z} \mathcal{E}_H(f, \mathcal{M}, i)^2 \right)^{1/2} \leq \sqrt{d+1} \|\text{Res}(f, \mathcal{M})\|_{H^{-1}(\omega_z)},$$

which is a local counterpart of the global lower bound in Veeger [23, Lemma 3.3].

This estimator is very closely related to the discretized residuals of §3.4 and Theorem 15. Indeed, if $K \in \mathcal{M}$ and $F \in \mathcal{F}$, $K_1, K_2 \in \mathcal{M}$ such that $F = K_1 \cap K_2$,

we have

$$(4.3) \quad \psi_K = \frac{(2d+1)!}{d!|K|} \lambda_K \quad \text{and} \quad \psi_F = \frac{(2d-1)!}{(d-1)!|F|} \left(\lambda_F - (2d+1) \sum_{i=1}^2 \lambda_{K_i} \right).$$

in view of (3.19). Hence $\text{span}\{\psi_i \mid i \in \mathcal{I}\} = \text{span}\{\lambda_i \mid i \in \mathcal{I}\}$ and Remark 12 yields $\langle f, \lambda_i \rangle = \langle \mathcal{P}_{\mathcal{M}} f, \lambda_i \rangle$, $i \in \mathcal{I}$, and the indicators may be viewed also as evaluations of the discretized residual: for $i \in \mathcal{I}$,

$$\mathcal{E}_{\text{H}}(f, \mathcal{M}, i) = \left\langle \mathcal{P}_{\mathcal{M}} f + \Delta U_{f, \mathcal{M}}, \frac{\lambda_i}{\|\nabla \lambda_i\|} \right\rangle.$$

As a consequence, we also have the following counterpart of (4.2):

$$(4.4) \quad \left(\sum_{i \in \mathcal{I}_z} \mathcal{E}_{\text{H}}(f, \mathcal{M}, i)^2 \right)^{1/2} \leq \sqrt{d+1} \|\mathcal{P}_{\mathcal{M}} f + \Delta U_{f, \mathcal{M}}\|_{H^{-1}(\omega_z)}.$$

In order to prove the converse bound, we may proceed with the help of $\mathcal{P}_{\mathcal{M}}^*$ as in [23]. However, having Theorem 15 at our disposal, it is simpler to exploit (4.3). We immediately see

$$(4.5a) \quad \mathcal{E}_{\text{H}}(f, \mathcal{M}, K) = \left\langle \mathcal{P}_{\mathcal{M}} f + \Delta U_{f, \mathcal{M}}, \frac{\psi_K}{\|\nabla \psi_K\|} \right\rangle.$$

Moreover, given $F \in \mathcal{F}$, $K_1, K_2 \in \mathcal{M}$ with $F = K_1 \cap K_2$, we deduce

$$C_d |F|^{-1} \leq \max_F \psi_F \leq h_F \max_{K_1} |\nabla \psi_F| \lesssim h_F |K|^{-1/2} \|\nabla \psi_F\|_{K_1}$$

with $h_F := \text{diam } F$ and, for $i \in \{F, K_1, K_2\}$

$$\|\nabla \lambda_i\|_{\omega_F} \leq C_d \max_{i=1,2} \rho_K^{-1} |\omega_F|^{1/2}.$$

We therefore obtain $\|\nabla \psi_F\|^{-1} \|\nabla \lambda_i\| \lesssim |F|$ and

$$(4.5b) \quad \left\langle \mathcal{P}_{\mathcal{M}} f + \Delta U_{f, \mathcal{M}}, \frac{\psi_F}{\|\nabla \psi_F\|} \right\rangle \lesssim \sum_{i \in \{F, K_1, K_2\}} \mathcal{E}_{\text{H}}(f, \mathcal{M}, i)$$

Summing up, the hierarchical estimator quantifies the local discretized residual,

$$\sum_{i \in \mathcal{I}_z} \mathcal{E}_{\text{H}}(f, \mathcal{M}, i)^2 \simeq \|\mathcal{P}_{\mathcal{M}} f + \Delta U_{f, \mathcal{M}}\|_{H^{-1}(\omega_z)}^2, \quad z \in \mathcal{V},$$

and we have the following a posteriori bounds.

Theorem 23 (Hierarchical estimator with error-dominated oscillation). *For any functional $f \in H^{-1}(\Omega)$ and any conforming mesh \mathcal{M} , we have the global equivalence*

$$\|\nabla(u_f - U_{f; \mathcal{M}})\|^2 \approx \sum_{i \in \mathcal{I}} \mathcal{E}_{\text{H}}(f, \mathcal{M}, i)^2 + \sum_{z \in \mathcal{V}} \|\mathcal{P}_{\mathcal{M}} f - f\|_{H^{-1}(\omega_z)}^2,$$

as well as the following local lower bounds: for every $z \in \mathcal{V}$,

$$\sum_{i \in \mathcal{I}_z} \mathcal{E}_{\text{H}}(f, \mathcal{M}, i)^2 \leq (d+1) \|\nabla(u_f - U_{f; \mathcal{M}})\|_{\omega_z}^2,$$

$$\sum_{i \in \mathcal{I}_z} \|\mathcal{P}_{\mathcal{M}} f - f\|_{H^{-1}(\omega_z)}^2 \lesssim \|\nabla(u_f - U_{f; \mathcal{M}})\|_{\omega_z}^2.$$

The hidden constants depend only on d and $\sigma(\mathcal{M})$.

Proof. Combine Theorem 16, Theorem 17, Corollary 18, (3.12), (4.2), and (4.5). \square

4.2. An improved standard residual estimator. The standard residual estimator applies suitably scaled norms to the jump and element residual; see, e.g., Verfürth [25, Section 1.4]. In the case of the discretized residual

$$\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}} = \sum_{F \in \mathcal{F}} (\langle f, \psi_F \rangle + J(U_{f;\mathcal{M}})|_F) \chi_F + \sum_{K \in \mathcal{M}} \langle f, \psi_K \rangle \chi_K,$$

this leads to the following indicators:

$$\begin{aligned} \mathcal{E}_{\mathbb{R}}(U_{f;\mathcal{M}}, \mathcal{P}_{\mathcal{M}}f, F) &:= h_F^{1/2} \|\langle f, \psi_F \rangle + J(U_{f;\mathcal{M}})|_F\|_F & F \in \mathcal{F}, \\ \mathcal{E}_{\mathbb{R}}(U_{f;\mathcal{M}}, \mathcal{P}_{\mathcal{M}}f, K) &:= h_K \|\langle f, \psi_K \rangle\|_K, & K \in \mathcal{M}, \end{aligned}$$

where h_F and h_K denote, respectively, the diameters of F and K and computability is given in terms of $U_{f;\mathcal{M}}$ and (3.23).

These indicators actually quantify the discretized residual and in a way that is very tight to Theorem 15: for any interelement face $F \in \mathcal{F}$,

$$(4.6a) \quad \mathcal{E}_{\mathbb{R}}(U_{f;\mathcal{M}}, \mathcal{P}_{\mathcal{M}}f, F) \approx \left\langle \mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}, \frac{\psi_F}{\|\nabla \psi_F\|} \right\rangle$$

and, for any element $K \in \mathcal{M}$,

$$(4.6b) \quad \mathcal{E}_{\mathbb{R}}(U_{f;\mathcal{M}}, \mathcal{P}_{\mathcal{M}}f, K) \approx \left\langle \mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}, \frac{\psi_K}{\|\nabla \psi_K\|} \right\rangle,$$

where the hidden constants depend only on d and $\sigma(\mathcal{M})$. To see (4.6a), let $F \in \mathcal{F}$ be any interelement face. Lemma 7 (i), the trace inequality (3.18) for $w = \psi_F^2$ and the Friedrichs inequality (3.17) for $v = \psi_F$, both with ω_F in place of ω_z , give

$$\begin{aligned} \left\langle \mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}, \frac{\psi_F}{\|\nabla \psi_F\|} \right\rangle &= \left\langle (\langle f, \psi_F \rangle + J(U_{f;\mathcal{M}})|_F) \chi_F, \frac{\psi_F}{\|\nabla \psi_F\|} \right\rangle \\ &\leq \|\langle f, \psi_F \rangle + J(U_{f;\mathcal{M}})|_F\|_F \frac{\|\psi_F\|_F}{\|\nabla \psi_F\|} \leq h_F^{1/2} \|\langle f, \psi_F \rangle + J(U_{f;\mathcal{M}})|_F\|_F, \end{aligned}$$

while (3.20) yields $\|\nabla \psi_F\|_{\Omega} \lesssim (h_F |F|)^{-1/2}$ and so

$$\begin{aligned} \left\langle \mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}, \frac{\psi_F}{\|\nabla \psi_F\|} \right\rangle &= \frac{\|\langle f, \psi_F \rangle + J(U_{f;\mathcal{M}})|_F\|_F}{|F|^{1/2} \|\nabla \psi_F\|} \\ &\gtrsim h_F^{1/2} \|\langle f, \psi_F \rangle + J(U_{f;\mathcal{M}})|_F\|_F. \end{aligned}$$

Similarly, we obtain (4.6b).

Inserting the combination of Theorem 15 and (4.6) in the abstract a posteriori analysis of §3.7, we obtain the following result.

Theorem 24 (Standard residual estimator with error-dominated oscillation). *For any functional $f \in H^{-1}(\Omega)$ and any conforming mesh \mathcal{M} , we have the global equivalence*

$$\|\nabla(u_f - U_{f;\mathcal{M}})\|^2 \approx \sum_{i \in \mathcal{I}} \mathcal{E}_{\mathbb{R}}(U_{f;\mathcal{M}}, \mathcal{P}_{\mathcal{M}}f, i)^2 + \sum_{z \in \mathcal{V}} \|\mathcal{P}_{\mathcal{M}}f - f\|_{H^{-1}(\omega_z)}^2,$$

as well as the following local lower bounds: for $z \in \mathcal{V}$,

$$\sum_{i \in \mathcal{I}_z} \mathcal{E}_{\mathbb{R}}(U_{f;\mathcal{M}}, \mathcal{P}_{\mathcal{M}}f, i)^2 + \|\mathcal{P}_{\mathcal{M}}f - f\|_{H^{-1}(\omega_z)}^2 \lesssim \|\nabla(u_f - U_{f;\mathcal{M}})\|_{\omega_z}^2.$$

The hidden constants depend only on d and $\sigma(\mathcal{M})$.

Theorem 24 relies on key features of the approach in §3, which the following remark elaborates on.

Remark 25 (Classical vs new standard residual estimator). In contrast to the classical standard residual estimator (1.2) and its H^{-1} -variant in Remark 22, the variant of Theorem 24 is completely equivalent to the error. The reason for this improvement lies in a suitable correction of the original jump residual. To elucidate this, remember that both the classical standard residual estimator and its H^{-1} -variant in Remark 22 do not discretize the residual and therefore compare them to

$$\sum_{F \in \mathcal{F}} h_F^{1/2} \|J(U_{f;\mathcal{M}}) + \langle f, \psi_F \rangle\|_F^2 + \sum_{z \in \mathcal{V}} \left\| f - \sum_{F \in \mathcal{F}} \langle f, \psi_F \rangle \chi_F \right\|_{H^{-1}(\omega_z)}^2,$$

which also does not split off an infinite-dimensional part of the load f . The corrections $\langle f, \psi_F \rangle$, $F \in \mathcal{F}$, of the jump residual make sure that the new jump residual has the invariance properties necessary for avoiding overestimation, i. e., it vanishes whenever the exact solution happens to be discrete. Corrections with this property have been used previously. For example, Nochetto [19] considers the special case $f = f_1 + \operatorname{div} \mathbf{f}_2$, where f_1, \mathbf{f}_2 are suitable functions, and assigns $(\operatorname{div} \mathbf{f}_2)|_K$, $K \in \mathcal{M}$, to the element residual and the jumps in the normal trace of \mathbf{f}_2 across interelement sides correct the jump residual. Similarly, in standard residual estimators for the Stokes problem, pressure jumps correct the jump residual associated with the velocity. The novelty is that the corrections $\langle f, \psi_F \rangle$, $F \in \mathcal{F}$, are defined for an arbitrary $f \in H^{-1}(\Omega)$ and also locally H^{-1} -stable and so fulfill the second necessary condition to avoid local overestimation. Notably, the latter entails that, even if f is a smooth function, the jump residual will be corrected significantly in certain cases.

4.3. An estimator based on local problems. A local problem lifts the residual to a local extension of the given finite element space and so provides a local correction, the norm of which is used as an error indicator; cf. Babuška and Rheinboldt [4]. While computability requires finite-dimensional extensions, the higher cost with respect to the previous explicit estimators is tied up with the hope of improved accuracy.

The following instance from Verfürth [25, Section 1.7.1 and Remark 1.21] is vertex-based and uses the local extensions

$$\mathbb{U}_z := \operatorname{span}\{\lambda_i \mid i \in \mathcal{I}_z\} = \operatorname{span}\{\psi_i \mid i \in \mathcal{I}_z\}, \quad z \in \mathcal{V},$$

where the functions ψ_i and λ_i are defined, respectively, in (3.19) and (4.1). Given a vertex $z \in \mathcal{V}$, the indicator is then

$$\mathcal{E}_L(f, \mathcal{M}, z) := \|\nabla \nu_z\|,$$

where

$$\nu_z \in \mathbb{U}_z \quad \text{such that} \quad \forall \lambda \in \mathbb{U}_z \quad \int_{\Omega} \nabla \nu_z \cdot \nabla \lambda \, dx = \langle \operatorname{Res}(f; \mathcal{M}), \lambda \rangle.$$

Thus, ν_z is computable in terms of $U_{f;\mathcal{M}}$ and, e.g., (3.23). The indicator $\mathcal{E}_L(f, \mathcal{M}, z)$ may be viewed as an implicit counterpart of $(\sum_{i \in \mathcal{I}_z} \mathcal{E}_H(f, \mathcal{M}, i)^2)^{1/2}$ from §4.1. Taking $\lambda = \nu_z$, we immediately obtain the constant-free lower bound

$$(4.7) \quad \mathcal{E}_L(f, \mathcal{M}, z) \leq \|\operatorname{Res}(f; \mathcal{M})\|_{H^{-1}(\omega_z)},$$

which slightly improves upon (4.2).

Notice that, in light of Remark 12, the solution ν_z can be interpreted also as a lift of the discretized residual $\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}$. Consequently, the first inequality in

$$(4.8) \quad \mathcal{E}_L(f, \mathcal{M}, z) \leq \|\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}\|_{H^{-1}(\omega_z)} \lesssim \mathcal{E}_L(f, \mathcal{M}, z)$$

is correct. The second one follows from Remark 13 and Theorem 10 in the spirit of Morin, Nocketto and Siebert [18]. In fact, for $v \in H_0^1(\omega_z)$, we have

$$\begin{aligned} \langle \mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}, v \rangle &= \langle \text{Res}(f; \mathcal{M}), \mathcal{P}_{\mathcal{M}}^*v \rangle = \int_{\omega_z} \nabla \nu_z \cdot \nabla \mathcal{P}_{\mathcal{M}}v \, dx \\ &\leq \|\nabla \nu_z\| \|\nabla \mathcal{P}_{\mathcal{M}}^*v\|_{\omega_z} \lesssim \mathcal{E}_L(f, \mathcal{M}, z) \|\nabla v\|_{\omega_z}. \end{aligned}$$

Theorem 26 (Estimator based on local problems with error-dominated oscillation). *For any functional $f \in H^{-1}(\Omega)$ and any conforming mesh \mathcal{M} , we have the global equivalence*

$$\|\nabla(u_f - U_{f;\mathcal{M}})\|^2 \approx \sum_{z \in \mathcal{V}} \mathcal{E}_L(f, \mathcal{M}, z)^2 + \|\mathcal{P}_{\mathcal{M}}f - f\|_{H^{-1}(\omega_z)}^2,$$

as well as the following local lower bounds: for every $z \in \mathcal{V}$,

$$\mathcal{E}_L(f, \mathcal{M}, z) \leq \|\nabla(u_f - U_{f;\mathcal{M}})\|_{\omega_z} \quad \text{and} \quad \|\mathcal{P}_{\mathcal{M}}f - f\|_{H^{-1}(\omega_z)} \lesssim \|\nabla(u_f - U_{f;\mathcal{M}})\|_{\omega_z}.$$

The hidden constants depend only on d and $\sigma(\mathcal{M})$.

Proof. Combine Theorem 16, Theorem 17, Corollary 18, (3.12), (4.7) and (4.8). \square

4.4. An estimator based on flux equilibration. While indicators based on local problems provide constant-free local lower bounds, estimators based on flux equilibration aim for a constant-free, or at least explicit, global upper bound. This is achieved with the help of other, more sophisticated liftings within the framework of the fundamental theorem of Prager and Synge [21], which, for the homogeneous Dirichlet problem (1.1), can be formulated as follows: For any $v \in H_0^1(\Omega)$, we have

$$(4.9) \quad \|\nabla(v - u)\| = \min \{ \|\boldsymbol{\xi}\| \mid \boldsymbol{\xi} \in L^2(\Omega; \mathbb{R}^d) \text{ with } \text{div } \boldsymbol{\xi} = \Delta v + f \text{ in } H^{-1}(\Omega) \}.$$

Realizations of this idea in Ainsworth [1], Braess and Schöberl [9], Ern, Smears and Vohralik [13, 14], and Luce and Wohlmuth [16] make use of some classical oscillation. Its replacement by an error-dominated oscillation requires some adjustment to the approach of §3.

The upper bound in the localization of Lemma 4 involves a non-explicit multiplicative constant. In order to improve on this, we replace the local spaces $H_0^1(\omega_z)$, $z \in \mathcal{V}$, with

$$H_z := \begin{cases} \{v \in H^1(\omega_z) \mid \int_{\omega_z} v = 0\}, & \text{if } z \in \mathcal{V}_0 = \mathcal{V} \cap \Omega, \\ \{v \in H^1(\omega_z) \mid v|_{\partial\omega_z \cap \partial\Omega} = 0\}, & \text{if } z \in \mathcal{V} \setminus \mathcal{V}_0, \end{cases}$$

equip them with the norm $\|\nabla \cdot\|_{\omega_z}$, and denote the respective dual spaces by H_z^* .

Lemma 27 (Alternative localization). *Let $\ell \in H^{-1}(\Omega)$ be any functional.*

(i) *If $\ell \in \mathcal{R}_{\mathcal{M}}$, then*

$$\|\ell\|_{H^{-1}(\Omega)}^2 \leq (d+1) \sum_{z \in \mathcal{V}} \|\phi_z \ell\|_{H_z^*}^2.$$

(ii) *We have*

$$\sum_{z \in \mathcal{V}} \|\phi_z \ell\|_{H_z^*}^2 \lesssim \|\ell\|_{H^{-1}(\Omega)}^2,$$

where the hidden constant depends only on d and the shape coefficient $\sigma(\mathcal{M})$.

Proof. The proof is essentially a regrouping of the arguments proving Lemma 4, where (3.3) slips into the proof of (ii); cf. Canuto et al. [10, Proposition 3.1]. \square

Splitting the residual up in discretized residual and oscillation, we then obtain the following abstract error bounds; we do not state the global lower bound as it is immediate consequence of the local one.

Lemma 28 (Alternative abstract error bounds). *For any functional $f \in H^{-1}(\Omega)$ and any conforming mesh \mathcal{M} , we have the global upper bound*

$$\|\nabla(u_f - U_{f;\mathcal{M}})\|^2 \leq \sum_{z \in \mathcal{V}} \left(\|\phi_z(\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}})\|_{H_z^*} + \|\phi_z(\mathcal{P}_{\mathcal{M}}f - f)\|_{H_z^*} \right)^2,$$

as well as the following local lower bounds: for every vertex $z \in \mathcal{V}$,

$$\|\phi_z(\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}})\|_{H_z^*} + \|\phi_z(\mathcal{P}_{\mathcal{M}}f - f)\|_{H_z^*} \lesssim \|\nabla(u_f - U_{f;\mathcal{M}})\|_{\omega_z}$$

The hidden constants depend only on d and $\sigma(\mathcal{M})$.

Proof. The global upper bound follows from Lemma 27 (ii) and the triangle inequality. To prove the local lower bounds, we recall Theorem 17 and take $\ell = \mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}$ and $\ell = \mathcal{P}_{\mathcal{M}}f - f$ in

$$(4.10) \quad \langle \phi_z \ell, v_z \rangle = \langle \ell, \phi_z v_z \rangle \leq \|\ell\|_{H^{-1}(\omega_z)} \|\nabla(v_z \phi_z)\|_{\omega_z} \lesssim \|\ell\|_{H^{-1}(\omega_z)} \|\nabla v_z\|_{\omega_z},$$

which exploits (3.3) for $v_z \in H_z$ and $z \in \mathcal{V}$. \square

In order to quantify the local discretized residual, we construct local equilibrated fluxes following the ideas of Braess, Pillwein, and Schöberl [8] and Ern, Smears, and Vohralík [13]. To this end, fix any vertex $z \in \mathcal{V}$ and define the operator $\pi_z : \{\phi_z v \mid v \in H^1(\omega_z)\} \rightarrow H_z^*$ by

$$(4.11) \quad \pi_z(\phi_z \ell) := \begin{cases} \phi_z \ell - \frac{\langle \phi_z \ell, 1 \rangle}{|\omega_z|} & \text{if } z \in \mathcal{V}_0, \\ \phi_z \ell & \text{if } z \in \mathcal{V} \setminus \mathcal{V}_0. \end{cases}$$

We emphasize that $\pi_z(\phi_z(\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}}))$ can be computed in terms of $U_{f;\mathcal{M}}$ and (3.23). Thanks to the definition of the spaces H_z , $z \in \mathcal{V}$, and the general form of the theorem of Prager and Synge (see, e.g., Verfürth [25, Proposition 1.40]), we have

$$(4.12) \quad \|\phi_z(\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}})\|_{H_z^*} = \|\pi_z \phi_z(\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}})\|_{H_z^*} = \min_{\xi \in \mathbb{W}_z} \|\xi\|_{\omega_z}$$

with the affine space

$$\begin{aligned} \mathbb{W}_z &:= \{ \xi \in L^2(\omega_z; \mathbb{R}^d) \mid \operatorname{div} \xi = \pi_z(\phi_z(\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}})) \in H_z^* \\ &\quad \text{and } \xi \cdot \mathbf{n} = 0 \text{ on } \partial\omega_z \text{ if } z \in \mathcal{V}_0 \\ &\quad \text{and } \xi \cdot \mathbf{n} = 0 \text{ on } \partial\omega_z \setminus \partial\Omega \text{ if } z \in \mathcal{V} \setminus \mathcal{V}_0 \}, \end{aligned}$$

and the equalities in the definition of \mathbb{W}_z have to be understood in the sense of distributions; the space \mathbb{W}_z is not empty since $\langle \pi_z(\phi_z(\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}})), 1 \rangle = 0$ for every $z \in \mathcal{V}_0$.

In order to introduce a discrete counterpart of \mathbb{W}_z in (4.12), we employ the Raviart-Thomas-Nédélec spaces

$$\operatorname{RTN}(K) := \{ \Xi : K \rightarrow \mathbb{R}^d \mid \Xi(x) = \mathbf{a} + bx \text{ for some } \mathbf{a} \in \mathbb{P}_1^d, b \in \mathbb{P}_1 \}, \quad K \in \mathcal{M},$$

and define

$$\begin{aligned} \mathbb{W}_z(\mathcal{M}) &:= \{ \Xi \in L^2(\omega_z) \mid \Xi|_K \in \operatorname{RTN}(K) \text{ for all } K \in \mathcal{M} \text{ with } K \subset \omega_z \\ &\quad \text{and } \operatorname{div} \Xi = \pi_z(\phi_z(\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}})) \in H_z^* \\ &\quad \text{and } \Xi \cdot \mathbf{n} = 0 \text{ on } \partial\omega_z \text{ if } z \in \mathcal{V}_0 \\ &\quad \text{and } \Xi \cdot \mathbf{n} = 0 \text{ on } \partial\omega_z \setminus \partial\Omega \text{ if } z \in \mathcal{V} \setminus \mathcal{V}_0 \}, \end{aligned}$$

which satisfies

$$(4.13) \quad \min_{\Xi \in \mathbb{W}_z(\mathcal{M})} \|\Xi\|_{\omega_z} \lesssim \min_{\xi \in \mathbb{W}_z} \|\xi\|_{\omega_z} \leq \min_{\Xi \in \mathbb{W}_z(\mathcal{M})} \|\Xi\|_{\omega_z}$$

and the hidden constant depends only on d and $\sigma(\mathcal{M})$. Indeed, the right inequality is obvious because of $\mathbb{W}_z(\mathcal{M}) \subset \mathbb{W}_z$. The left inequality can be proved by an

explicit construction; see, e.g., [8, 13]. For the ease of presentation, however, we shall assume

$$\mathfrak{E}_z := \arg \min_{\mathfrak{E} \in \mathbb{W}_z(\mathcal{M})} \|\mathfrak{E}\|_{\omega_z}$$

and note

$$\|\mathfrak{E}_z\|_{\omega_z} \lesssim \|\phi_z(\mathcal{P}_{\mathcal{M}}f + \Delta U_{f;\mathcal{M}})\|_{H_z^*} \leq \|\mathfrak{E}_z\|_{\omega_z}.$$

in view of (4.12) and (4.13). Inserting this in the abstract bounds of Lemma 28, we readily obtain the following a posteriori bounds; as before, we suppress the global lower bound.

Theorem 29 (Equilibrated flux estimator with error-dominated oscillation). *For any functional $f \in H^{-1}(\Omega)$ and any conforming mesh \mathcal{M} , we have the global upper bound*

$$\|\nabla(u_f - U_{f;\mathcal{M}})\|^2 \leq (d+1) \sum_{z \in \mathcal{V}} (\|\mathfrak{E}_z\|_{\omega_z} + \|\phi_z(\mathcal{P}_{\mathcal{M}}f - f)\|_{H_z^*})^2$$

as well as the following local lower bounds: for every vertex $z \in \mathcal{V}$,

$$\|\mathfrak{E}_z\|_{\omega_z}^2 + \|\phi_z(\mathcal{P}_{\mathcal{M}}f - f)\|_{H_z^*}^2 \lesssim \|\nabla(u_f - U_{f;\mathcal{M}})\|_{\omega_z}^2.$$

The hidden constant depends only on d and $\sigma(\mathcal{M})$.

In contrast to the cited previous bounds, the upper bound in Theorem 29 contains the multiplicative constant $d+1$. This constant arises from the localization in Lemma 27. As an alternative to this localization, one may use the constant-free upper bound in the following remark and split the estimator part $\|\mathfrak{E}\|$ therein into local L^2 -contributions.

Remark 30 (Alternative upper bound). Observing that

$$\sum_{z \in \mathcal{V}} \operatorname{div} \mathfrak{E}_z = f + \Delta U_{f;\mathcal{M}} + \sum_{z \in \mathcal{V}} \pi_z(\phi_z(\mathcal{P}_{\mathcal{M}}f - f)),$$

we set $\mathfrak{E} := \sum_{z \in \mathcal{V}} \mathfrak{E}_z$ and apply the theorem of Prager and Synge (4.9) globally and Lemma 27 to obtain

$$\|\nabla(u_f - U_{f;\mathcal{M}})\| \leq \|\mathfrak{E}\| + \sqrt{d+1} \left(\sum_{z \in \mathcal{V}} \|\phi_z(\mathcal{P}_{\mathcal{M}}f - f)\|_{H_z^*}^2 \right)^{1/2}.$$

Acknowledgment. The authors would like to thank Rüdiger Verfürth for the proposal to define a precursor of the operator $\mathcal{P}_{\mathcal{M}}$ in terms of a biorthogonal system.

REFERENCES

- [1] M. AINSWORTH, *A framework for obtaining guaranteed error bounds for finite element approximations*, Journal of Computational and Applied Mathematics, 234 (2010), pp. 2618–2632.
- [2] M. AINSWORTH AND J. T. ODEN, *A Posteriori Error Estimation in Finite Element Analysis*, Pure and Applied Mathematics, Wiley-Interscience, New York, 2000.
- [3] I. BABUŠKA AND A. MILLER, *A feedback finite element method with a posteriori error estimation: I. The finite element method and some basic properties of the a posteriori error estimator*, Comput. Methods Appl. Mech. Engrg., 61 (1987), pp. 1–40.
- [4] I. BABUŠKA AND W. RHEINBOLDT, *Error estimates for adaptive finite element computations*, SIAM J. Numer. Anal., 15 (1978), pp. 736–754.
- [5] J. BLECHTA, J. MÁLEK, AND M. VOHRALÍK, *Localization of the $W^{-1,q}$ -norm for local a posteriori efficiency*. IMA J. Numer. Anal., to appear.
- [6] F. A. BORNEMANN, B. ERDMANN, AND R. KORNUBER, *A posteriori error estimates for elliptic problems in two and three space dimensions*, SIAM J. Numer. Anal., 33 (1996), pp. 1188–1204.
- [7] D. BRAESS, *Finite elements*, Cambridge University Press, Cambridge, third ed., 2007. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker.
- [8] D. BRAESS, V. PILLWEIN, AND J. SCHÖBERL, *Equilibrated residual error estimates are p-robust*, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 1189–1197.

- [9] D. BRAESS AND J. SCHÖBERL, *Equilibrated residual error estimator for edge elements*, Math. Comp., 77 (2008), pp. 651–672.
- [10] C. CANUTO, R. H. NOCHETTO, R. STEVENSON, AND M. VERANI, *On p -robust saturation for hp-afem*, Computers & Mathematics with Applications, 73 (2017), pp. 2004–2022.
- [11] A. COHEN, R. DEVORE, AND R. H. NOCHETTO, *Convergence rates of AFEM with H^{-1} data*, Found. Comput. Math., 12 (2012), pp. 671–718.
- [12] A. ERN AND J.-L. GUERMOND, *Finite elements: II. Applications to PDEs*. in preparation.
- [13] A. ERN, I. SMEARS, AND M. VOHRALÍK, *Discrete p -robust $H(\text{div})$ -liftings and a posteriori estimates for elliptic problems with H^{-1} source terms*, Calcolo, (2017), pp. 1–17.
- [14] A. ERN AND M. VOHRALÍK, *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.
- [15] C. KREUZER AND A. VEESER, *A posteriori analysis with error-dominated oscillation for higher order methods*. in preparation.
- [16] 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.
- [17] P. MORIN, R. H. NOCHETTO, AND K. G. SIEBERT, *Convergence of adaptive finite element methods*, SIAM Review, 44 (2002), pp. 631–658.
- [18] ———, *Local problems on stars: A posteriori error estimators, convergence, and performance*, Math. Comp., 72 (2003), pp. 1067–1097.
- [19] R. H. NOCHETTO, *Pointwise a posteriori error estimates for elliptic problems on highly graded meshes*, Math. Comp., 64 (1995), pp. 1–22.
- [20] R. H. NOCHETTO AND A. VEESER, *Primer of adaptive finite element methods*, in Multiscale and adaptivity: modeling, numerics and applications, vol. 2040 of Lecture Notes in Math., Springer, Heidelberg, 2012, pp. 125–225.
- [21] W. PRAGER AND J. L. SYNGE, *Approximations in elasticity based on the concept of function space*, Quart. Appl. Math., 5 (1947), pp. 241–269.
- [22] R. STEVENSON, *Optimality of a standard adaptive finite element method*, Found. Comput. Math., 7 (2007), pp. 245–269.
- [23] A. VEESER, *Convergent adaptive finite elements for the nonlinear Laplacian*, Numer. Math., 92 (2002), pp. 743–770.
- [24] R. VERFÜRTH, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Adv. Numer. Math., John Wiley, Chichester, UK, 1996.
- [25] R. VERFÜRTH, *A posteriori error estimation techniques for finite element methods*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2013.

CHRISTIAN KREUZER, FAKULTÄT FÜR MATHEMATIK, TECHNISCHE UNIVERSITÄT DORTMUND, VOGELPOTHSWEG 87, D-44227 DORTMUND, GERMANY

URL: <http://www.mathematik.tu-dortmund.de/lxx/cms/de/mitarbeiter/ckreuzer.html>

Email address: christian.kreuzer@tu-dortmund.de

ANDREAS VEESER, DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DEGLI STUDI DI MILANO, VIA C. SALDINI 50, I-20133 MILANO, ITALY

URL: <http://mat.unimi.it/users/veeser/>

Email address: andreas.veeser@unimi.it