A NOTE ON MULTILEVEL BASED ERROR ESTIMATION

HELMUT HARBRECHT AND REINHOLD SCHNEIDER

Abstract. By employing the infinite multilevel representation of the residual, we derive computable bounds to estimate the distance of finite element approximations to the solution of the Poisson equation. If the finite element approximation is a Galerkin solution, the derived error estimator coincides with the standard element and edge based estimator. If Galerkin orthogonality is not satisfied, then the discrete residual additionally appears in terms of the BPX preconditioner. As a by-product of the present analysis, conditions are derived such that the hierarchical error estimation is reliable and efficient.

1. Introduction

Residual based a posteriori error estimators are by now a well established tool for the efficient solution of boundary value problems [1, 2, 28]. They are used to automatically design optimized meshes in the sense of achieving a given target accuracy with a minimal number of degrees of freedom. The mesh generation is generally steered by trying to equidistribute the local error contributions of a residual error estimator. Convergence results, i.e., the error reduction by some convergence factor, are however rather new. In the pioneering work [16, 17], a strategy has been proposed which ensures convergence to any desired accuracy in a finite number of refinement steps. It has significantly been improved in [22] by the notion of data oscillation. Optimal computational complexity has firstly been provided by means of wavelet theory [10, 11, 12]. These results and techniques have been extended by many authors, see e.g. [6, 9, 25, 27, 29] and the references therein.

The basic problem in constructing a posteriori error estimators is that the residual has to be measured in the dual space $H^{-1}(\Omega)$. The standard trick to obtain upper estimates of the error is to employ Galerkin orthogonality by introducing a local (Clémont) interpolant. Duality together with an approximation property provides the desired estimate. Nevertheless, an exact solution of the corresponding discrete equations can be very costly, even though a linearly scaling iterative scheme like a multigrid method is used. Hence, to find ways of estimating the discretization error without having the discrete solution at hand is important, especially in situations where the accurate solution of the discrete problem is expensive, e.g. in case of eigenvalue and related problems.

If the discrete system is not solved exactly, Galerkin orthogonality does not hold and cannot be employed to derive a posteriori estimates. However, the $H^{-1}(\Omega)$-norm of the discrete residual, i.e., the residual with respect to the current Galerkin solution, is computable by a preconditioner which is spectrally equivalent to the

2000 *Mathematics Subject Classification.* 65N15, 65N30, 65N50.

Key words and phrases. A-posteriori error estimates, multilevel finite elements, violated Galerkin orthogonality, hierarchical error estimation.

underlying operator, for example the BPX scheme [7]. Often one thus combines the BPX preconditioner with the traditional error estimators to estimate the residual, see e.g. [24] and the references therein. The reliability and efficiency of this error estimator can be deduced with the help of the stability of the traditional error estimator, see [27, Proposition 4.6].

In this article, we will provide an alternative technique for proving that the combination of the traditional a posteriori error estimator with the BPX preconditioner yields a reliable and efficient error estimator. The proof is based on a hypothetical infinite and dense collection of nested finite element spaces. Then, according to e.g. [20, 21, 23], the infinite BPX scheme provides a frame in $H^{-1}(\Omega)$ which means that

$$\|r\|^2_{H^{-1}(\Omega)} \sim \sum_{j\in\mathbb{N}_0}\sum_{k\in\triangle_j} \langle r,\varphi_{j,k}\rangle^2_{L^2(\Omega)}. \tag{1.1}$$

Here, $\varphi_{j,k}$ is the nodal finite element basis function on level j, normalized in the $H^1(\Omega)$-norm, and the sum has to be taken over all levels. Moreover, the bilinear form $\langle\cdot,\cdot\rangle_{L^2(\Omega)}$ has to be understood as the continuous extension of the $L^2(\Omega)$-inner product to the duality product $H^{-1}(\Omega)\times H^1_0(\Omega)\to\mathbb{R}$.

The infinite sum on the right hand side of (1.1) could be truncated at a certain refinement level. In fact, the contribution of a sufficiently refined subsequent level provides a bound of the residual from above. It also provides a lower bound which is seen by using a technique as in [17, 18, 22]. It is some kind of stability estimate, which could be provided by local stability estimates using sufficiently many bubble functions. For technical reasons, to apply these arguments, we need two (three) refinement levels in our proofs for piecewise (bi-) linear finite elements on triangles (parallelograms).

It turns out, if we neglect data oscillations or if we have them under explicit control, the $H^{-1}(\Omega)$-norm of the residual is split in the discrete part, which is easily computed by means of the BPX preconditioner, and a continuous part, which yields the traditional upper bound. Hence, during an iterative procedure, the discrete residual can arbitrarily be balanced with the discretization error.

Furthermore, we can verify the hierarchical error estimation. If one incorporates the data oscillation in the right hand side of (1.1), it is reasonable that the $H^{-1}(\Omega)$-norm of the remaining residual can be recovered, up to a constant, by using a BPX scheme applied on an appropriately refined grid. In this article, we will show how this result can be proven rigorously.

To keep the arguments simple, we consider the Dirichlet problem for the Poisson equation on a two-dimensional polygonal domain, using linear or bilinear finite elements, as model problem for demonstrating our arguments. We, however, emphasize that the present concept is quite general and can also be used to derive error estimators in case of other (elliptic) partial differential operators.

We finally like to mention that we have not tried to optimize or sharpen the constants which appear in the estimates. Moreover, we do not touch the question of goal oriented error estimators, where in the framework of dual weighted residual method [4] Galerkin orthogonality plays a crucial role.

The present article is organized as follows. The problem under consideration is formulated in Section 2. Section 3 introduces briefly the multilevel frame construction. In Section 4 it is shown how to switch between adaptive finite element representations and frame representations. In Section 5 we prove the efficiency and

reliability of the a posteriori estimator without having Galerkin orthogonality. This result is used to derive the hierarchical estimator in Section 6.

Throughout the article, in order to avoid the repeated use of generic but unspecified constants, by $C \lesssim D$ we mean that C can be bounded by a multiple of D, independently of parameters which C and D may depend on. Obviously, $C \gtrsim D$ is defined as $D \lesssim C$, and $C \sim D$ as $C \lesssim D$ and $C \gtrsim D$.

2. Problem statement

Let $\Omega \subset \mathbb{R}^2$ denote a two-dimensional bounded polygonal domain with boundary $\partial\Omega$. In the present article, given $f \in L^2(\Omega)$, we intend to solve the Poisson equation

$$-\Delta u = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega \tag{2.1}$$

by adaptive finite element methods. To do so, for a given finite element solution u_h, we need an error estimator, which estimates the residual $r \in H^{-1}(\Omega)$, given by

$$\langle r, v\rangle_{L^2(\Omega)} = \int_\Omega f v \,\mathrm{d}\mathbf{x} - \int_\Omega \nabla u_h \nabla v \,\mathrm{d}\mathbf{x} \quad \text{for all } v \in H^1_0(\Omega),$$

in the $H^{-1}(\Omega)$-norm. In general, to derive such error estimators, Galerkin orthogonality is employed.

Consider a shape regular triangulation $\mathcal{K}_h$ of the domain Ω by triangular or parallelogrammic elements $\mathcal{K}_h$. The set of all interior edges is indicated by $\mathcal{E}_h$. The space of continuous piecewise (bi-) linear elements will be denoted by V_h. Defining for $u_h \in V_h$ the element residual

$$q := \Delta u_h + f \quad \text{for all elements } K \in \mathcal{K}_h$$

and the jumps at the edges

$$\delta := \left[\frac{\partial u_h}{\partial \mathbf{n}}\right] \quad \text{for all edges } E \in \mathcal{E}_h,$$

the standard error estimation in case of Galerkin orthogonality (cf. [1, 2, 28]) is

$$\|r\|_{H^{-1}(\Omega)} \sim \|u - u_h\|_{H^1(\Omega)} \lesssim \sum_{K\in\mathcal{K}_h} h_K^2 \|q\|^2_{L^2(K)} + \sum_{E\in\mathcal{E}_h} h_E \|\delta\|^2_{L^2(E)} \tag{2.2}$$

with h_K and h_E denoting the mesh sizes of the elements K and edges E, respectively.

To numerically compute the approximate solution u_h to (2.1), one employs a finite element basis Φ_Σ of the space V_h and obtains a linear system of equations

$$\mathbf{A}_\Sigma \mathbf{u}_\Sigma = \mathbf{f}_\Sigma. \tag{2.3}$$

If we solve this discrete system not exactly, Galerkin orthogonality cannot be employed if the discrete residual $\mathbf{r}_\Sigma := \mathbf{f}_\Sigma - \mathbf{A}_\Sigma \mathbf{u}_\Sigma$ is too large. We shall show in the sequel that (2.2) still holds if the discrete residual is incorporated in terms of the BPX preconditioner $\mathrm{BPX}(\mathbf{r}_\Sigma)$ (see (5.1) for a precise definition). By incorporating data oscillations in a proper way, see e.g. [22], this estimator is shown to be reliable and efficient.

Notice that for all $K \in \mathcal{K}_h$ the local lower estimate

$$h_K^2 \|q\|^2_{L^2(K)} + \sum_{E\in\partial K} h_E \|\delta\|^2_{L^2(E)} \lesssim \|u - u_h\|^2_{H^1(\omega_K)} + \sum_{\substack{K'\in\mathcal{K}_h:\ K = K' \\ \text{or } K \text{ and } K' \text{ have} \\ \text{a common edge}}} h_K^2 \|f - f_h\|^2_{L^2(K')}$$

does always hold, i.e., without having Galerkin orthogonality. Here, the domain ω_K consists of the union of the element K and all elements which share an interior edge with K. The function f_h denotes the projection into the space of piecewise constants with respect to the triangulation and thus the difference $f - f_h$ refers to the data oscillation. For the proof of this estimate, we refer to e.g. [8, 28].

3. Multilevel frames

Let $\Omega \subset \mathbb{R}^2$ be a polygonal and bounded domain with coarse triangulation or quadrangulation (by parallelograms) $\mathcal{T}_0 = \{\tau_{0,k}\}$. By dyadic refinement of each element on level $j-1$ into 4 elements on level j, we recursively obtain for any $j > 0$ the triangulation or quadrangulation $\mathcal{T}_j = \{\tau_{j,k}\}$. On the mesh $\mathcal{T}_j$, we define standard Lagrangian piecewise linear or bilinear continuous finite elements $\Phi_j = \{\varphi_{j,k} : k \in \triangle_j\}$. This yields a nested sequence of finite dimensional trial spaces

$$V_0 \subset V_1 \subset \cdots \subset V_j \subset \cdots \subset H_0^1(\Omega), \tag{3.1}$$

where

$$\begin{aligned} V_j &= \operatorname{span}\{\varphi_{j,k} : k \in \triangle_j\} \\ &= \{u \in C(\Omega) : u|_\tau \text{ is a (bi-) linear polynomial for all } \tau \in \mathcal{T}_j\} \end{aligned}$$

($\triangle_j$ denotes an appropriate index set) and $\dim V_j \sim 4^j$. For our approach it is convenient to normalize the Lagrangian finite elements with respect to the energy space, i.e.,

$$\|\varphi_{j,k}\|_{H^1(\Omega)} \sim 1. \tag{3.2}$$

In view of this normalization, the stability of the basis functions implies the norm equivalence

$$\sum_{k \in \triangle_j} \langle v_j, \varphi_{j,k} \rangle_{L^2(\Omega)}^2 \sim 2^{-2j} \|v_j\|_{L^2(\Omega)}^2 \quad \text{for all } v_j \in V_j, \tag{3.3}$$

which immediately yields

$$\sum_{k \in \triangle_j} \langle v, \varphi_{j,k} \rangle_{L^2(\Omega)}^2 \lesssim 2^{-2j} \|v\|_{L^2(\Omega)}^2 \quad \text{for all } v \in L^2(\Omega). \tag{3.4}$$

From now on, we will indicate the basis functions by the multi-index $\lambda = (j,k) \in \Lambda$, where $\Lambda = \{\lambda = (j,k) : k \in \triangle_j,\ j \in \mathbb{N}_0\}$ denotes the set of all such indices. The notion $|\lambda| := j$ indicates the scale.

In accordance with [21, 23], the collection of all basis functions $\Phi_\Lambda = \{\varphi_\lambda : \lambda \in \Lambda\}$ defines a frame for $H_0^1(\Omega)$, which means that

$$\sum_{\lambda \in \Lambda} \langle f, \varphi_\lambda \rangle_{L^2(\Omega)}^2 \sim \|f\|_{H^{-1}(\Omega)}^2 \quad \text{for all } f \in H^{-1}(\Omega). \tag{3.5}$$

Notice that this frame underlies the construction of the so-called BPX preconditioner, see e.g. [7, 15, 20].

A natural tree structure is imposed by the present multiscale hierarchy.

Definition 3.1 (Father-son relation)**.** *A basis function φ_λ is called a son of the basis function $\varphi_{\lambda'}$, if $|\lambda'| + 1 = |\lambda|$ and $\operatorname{supp}(\varphi_\lambda) \subset \operatorname{supp}(\varphi_{\lambda'})$. The basis function $\varphi_{\lambda'}$ is then called the* father *of φ_λ. We call a basis function φ_λ a predecessor of $\varphi_{\lambda'}$,*

if there exists a path in the directed father-son graph which connects both functions. Analogously the notion successor is defined.

Notice that this definition does not imply uniqueness of the father-son relation. There exist basis functions which own more than one father. Moreover, for sake of convenience, we will use the above father-son relation also in the context of the related index sets.

4. Adaptive representations

Of course, on a computer, we can use only finite representations of functions. Therefore, we shall consider functions $u_h = \sum_{\lambda\in\Sigma} u_\lambda \varphi_\lambda = \Phi_\Sigma \mathbf{u}_\Sigma$ with finite index sets $\Sigma \subset \Lambda$. We have to impose some restrictions in order to ensure that the set $\Phi_\Sigma = \{\varphi_\lambda : \lambda \in \Sigma\}$ generates a standard finite element space. To that end, we note that each basis function φ_λ is supported on a certain number of elements, i.e., it holds

$$\operatorname{supp}(\varphi_\lambda) = \bigcup_{k\in\mathcal{J}_\lambda} \overline{\tau}_{|\lambda|,k}$$

for a certain index set $\mathcal{J}_\lambda$. Hence, a finite collection Φ_Σ induces a mesh $\mathcal{K}_h$ on the domain Ω via the collection of elements since

$$\overline{\Omega} = \bigcup_{\lambda\in\Sigma} \bigcup_{k\in\mathcal{J}_\lambda} \overline{\tau}_{|\lambda|,k}.$$

Here and in what follows, $\overline{A}$ means the closure of the set A.

Definition 4.1 (Gradedness)**.** *The mesh $\mathcal{K}_h$ is called graded, if the levels of adjacent elements (that are elements with common edges or vertices) differ at most by one.*

Any non-graded mesh can be extended to a graded mesh by additional refinement steps, cf. Figure 1. Let us remark that the present notion of gradedness is somewhat more restrictive than the one-irregularity from [3] since there, in case of parallelograms, the elements which share a common vertex may differ by two levels.

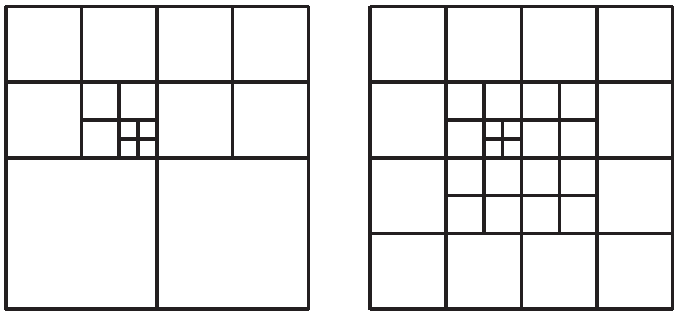

Figure 1. A nongraded mesh (left) and the corresponding graded mesh (right).

As one readily infers, by our definitions, a graded mesh induces always a finite element space V_h, and thus a basis Φ_Σ, which belongs to a mesh with at most one hanging node per edge. In Figure 2, we marked the degrees of freedom by "•" while

"$\circ$" indicates hanging nodes. Each degree of freedom is associated uniquely with a basis function φ_λ if we demand that the level $|\lambda|$ is the highest possible one relative to the given mesh. An illustration can be found in Figure 3. Notice that this choice induces that the value at the hanging node is given by linear interpolation from the two vertex values of the particular edge.

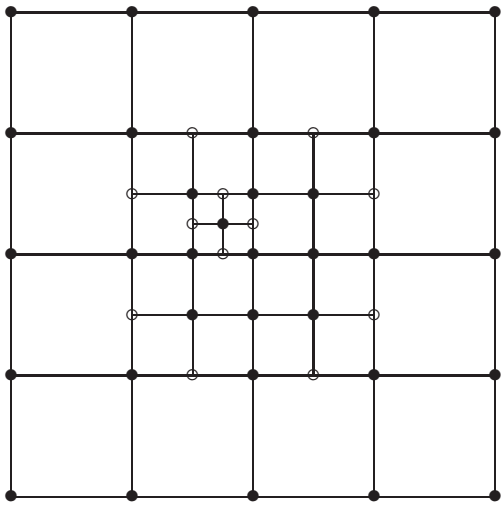

FIGURE 2. The degrees of freedom induced by a graded mesh.

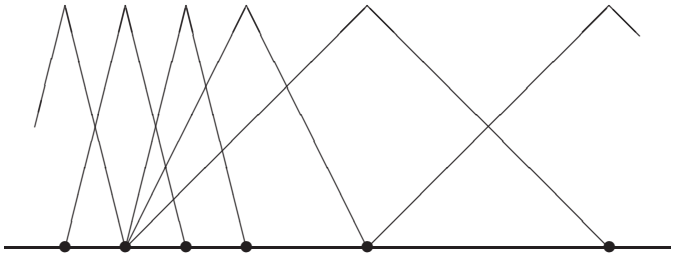

FIGURE 3. The finite element basis functions on a graded mesh.

5. Upper and lower error bounds

Consider a graded mesh and an associated finite element function $u_h \in V_h$ which has the basis representation $\sum_{\lambda\in\Sigma} u_\lambda\varphi_\lambda = \Phi_\Sigma \mathbf{u}_\Sigma$. Our first theorem states an upper bound of the residual $r = \Delta u_h + f$ for $f \in L^2(\Omega)$. Upper and lower bounds will include the discrete residual

$$\mathrm{BPX}(\mathbf{r}_\Sigma)^2 := \sum_{\substack{\lambda\in\{\Sigma \text{ and all its}\\ \text{predecessors}\}}} \langle r, \varphi_\lambda\rangle^2_{L^2(\Omega)}, \tag{5.1}$$

which can be computed by applying the standard BPX scheme [7] to the vector $\mathbf{r}_\Sigma$.

Theorem 5.1 (Upper bound)**.** *Let $r = q + \delta \in H^{-1}(\Omega)$ be split in the interior residual $q \in L^2(\Omega)$ and the jump terms at the edges $\delta \in H^{-1}(\Omega)$. Then, the residual is bounded by the sum of the discrete residual $\mathbf{r}_\Sigma$ plus the interior residual q and the jump δ across the edges:*

$$\|r\|^2_{H^{-1}(\Omega)} \lesssim \mathrm{BPX}(\mathbf{r}_\Sigma)^2 + \sum_{K\in\mathcal{K}_h} h_K^2 \|q\|^2_{L^2(K)} + \sum_{E\in\mathcal{E}_h} h_E \|\delta\|^2_{L^2(E)}.$$

Here, $\mathcal{K}_h$ denotes the set of elements and $\mathcal{E}_h$ the set of edges of the underlying mesh.[1] *Moreover, $h_K := \mathrm{diam}(K)$ for all $K \in \mathcal{K}_h$ and $h_E := \mathrm{diam}(E)$ for all $E \in \mathcal{E}_h$ are the related local mesh sizes.*

Proof. We need to estimate the residual $r = f + \Delta u_h$ in the space $H^{-1}(\Omega)$, which, on account of the frame property (3.5), is given by

$$\|r\|^2_{H^{-1}(\Omega)} \sim \sum_{\lambda\in\Lambda} \langle r, \varphi_\lambda\rangle^2_{L^2(\Omega)}.$$

We split the index set Λ into the set of *inactive* indices, that are all successors of Σ, and the set of *active* indices, which includes Σ and all its predecessors:

$$\begin{aligned}\|r\|^2_{H^{-1}(\Omega)} &\sim \underbrace{\sum_{\lambda \text{ is active}} \langle r, \varphi_\lambda\rangle^2_{L^2(\Omega)}}_{=\mathrm{BPX}(\mathbf{r}_\Sigma)^2} + \sum_{\lambda \text{ is inactive}} \langle r, \varphi_\lambda\rangle^2_{L^2(\Omega)} \\ &= \mathrm{BPX}(\mathbf{r}_\Sigma)^2 + \sum_{\lambda \text{ is inactive}} \left| \sum_{K\in\mathcal{K}_h} \langle q, \varphi_\lambda\rangle_{L^2(K)} + \sum_{E\in\mathcal{E}_h} \langle \delta, \varphi_\lambda\rangle_{L^2(E)} \right|^2.\end{aligned}$$

From $(a+b)^2 \le 2a^2 + 2b^2$ we conclude

$$\begin{aligned}\|r\|^2_{H^{-1}(\Omega)} \lesssim \mathrm{BPX}(\mathbf{r}_\Sigma)^2 &+ \sum_{\lambda \text{ is inactive}} \left| \sum_{K\in\mathcal{K}_h} \langle q, \varphi_\lambda\rangle_{L^2(K)} \right|^2 \\ &+ \sum_{\lambda \text{ is inactive}} \left| \sum_{E\in\mathcal{E}_h} \langle \delta, \varphi_\lambda\rangle_{L^2(E)} \right|^2.\end{aligned}$$

Since, due to the gradedness, each inactive basis function φ_λ is supported only on $\mathcal{O}(1)$ elements from $\mathcal{K}_h$, it follows that

$$\begin{aligned}\left| \sum_{K\in\mathcal{K}_h} \langle q, \varphi_\lambda\rangle_{L^2(K)} \right|^2 &\lesssim \sum_{K\in\mathcal{K}_h} \langle q, \varphi_\lambda\rangle^2_{L^2(K)}, \\ \left| \sum_{E\in\mathcal{E}_h} \langle \delta, \varphi_\lambda\rangle_{L^2(E)} \right|^2 &\lesssim \sum_{E\in\mathcal{E}_h} \langle \delta, \varphi_\lambda\rangle^2_{L^2(E)}.\end{aligned}$$

Thus, we arrive at

$$\begin{aligned}\|r\|^2_{H^{-1}(\Omega)} &\lesssim \mathrm{BPX}(\mathbf{r}_\Sigma)^2 + \sum_{K\in\mathcal{K}_h} \sum_{\lambda \text{ is inactive}} \langle q, \varphi_\lambda\rangle^2_{L^2(K)} + \sum_{E\in\mathcal{E}_h} \sum_{\lambda \text{ is inactive}} \langle \delta, \varphi_\lambda\rangle^2_{L^2(E)} \\ &=: \mathrm{BPX}(\mathbf{r}_\Sigma)^2 + \sum_{K\in\mathcal{K}_h} A_K + \sum_{E\in\mathcal{E}_h} B_E.\end{aligned}$$

[1] If an element $K \in \mathcal{K}_h$ has a hanging node, this part of the boundary consists of two edges.

Now, for fixed K and E, we estimate the terms A_K and B_E separately. To this end, we denote by $|K|$ and $|E|$ the associated refinement level of the elements and edges K and E, respectiveley. Then, by (3.4), we find

$$\begin{aligned} A_K &= \sum_{\lambda \text{ is inactive}} \langle q, \varphi_\lambda \rangle^2_{L^2(K)} \leq \sum_{j \geq |K|} \sum_{k \in \triangle_j} \langle q, \varphi_{j,k} \rangle^2_{L^2(K)} \lesssim \sum_{j \geq |K|} 2^{-2j} \|q\|^2_{L^2(K)} \\ &\sim 2^{-2|K|} \|q\|^2_{L^2(K)}. \end{aligned}$$

Similarly, by applying the argument (3.4) to the finite element spaces $V_j\big|_E$, we conclude

$$\begin{aligned} B_E &= \sum_{\lambda \text{ is inactive}} \langle \delta, \varphi_\lambda \rangle^2_{L^2(E)} \leq \sum_{j \geq |E|} \sum_{k \in \triangle_j} \langle \delta, \varphi_{j,k} \rangle^2_{L^2(E)} \lesssim \sum_{j \geq |E|} 2^{-j} \|\delta\|^2_{L^2(E)} \\ &\sim 2^{-|E|} \|\delta\|^2_{L^2(E)}. \end{aligned}$$

This finishes the proof due to $2^{-|K|} \sim h_K$ and $2^{-|E|} \sim h_E$. □

We emphasize that this error estimator is not computable since the norms $\|q\|_{L^2(K)}$ can generally not be computed exactly. To obtain computable bounds, we shall introduce the notion of L^2-data oscillation.

Definition 5.2 (Data oscillation)**.** *Let $V(K)$ denote the space of linear (in case of triangles) or bilinear (in case of parallelograms) polynomials on the element K. Then, for all $K \in \mathcal{K}_h$ we define*

$$\operatorname{osc}(f, K) := \|f - P_K f\|_{L^2(K)},$$

where $P_K : L^2(K) \to V(K)$ denotes the L^2-orthogonal projection onto the space $V(K)$.

Our definition of the data oscillation differs from the usual one where a projection onto piecewise constant functions is employed. But in order to ensure (5.3), we require a projection onto (possibly discontinuous) piecewise (bi-) linear functions. In case of Galerkin orthogonality, the discrete residual is not present and a projection onto piecewise constant functions will clearly be sufficient.

By using the data oscillations, we have the decomposition

$$\|q\|^2_{L^2(K)} = \|\Delta u_h + f\|^2_{L^2(K)} = \|\Delta u_h + P_K f\|^2_{L^2(K)} + \operatorname{osc}(f, K)^2. \tag{5.2}$$

Moreover, we shall introduce the projected residual

$$r_h := q_h + \delta, \quad \text{where } q_h = \sum_{K \in \mathcal{K}_h} \Delta u_h + P_K f.$$

In particular, there holds

$$\langle r, \varphi_\lambda \rangle_{L^2(\Omega)} = \langle r_h, \varphi_\lambda \rangle_{L^2(\Omega)} \quad \text{for all active } \lambda, \tag{5.3}$$

which means that r and r_h lead to the same discrete residual $\operatorname{BPX}(\mathbf{r}_\Sigma)$. Consequently, applying the identity (5.2) in each element $K \in \mathcal{K}_h$ yields immediately the following corollary.

Corollary 5.3. *It holds that*

$$\|r\|^2_{H^{-1}(\Omega)} \lesssim \mathrm{BPX}(\mathbf{r}_\Sigma)^2 + \sum_{K\in\mathcal{K}_h} h_K^2 \|q_h\|^2_{L^2(K)} + \sum_{E\in\mathcal{E}_h} h_E \|\delta\|^2_{L^2(E)} + \sum_{K\in\mathcal{K}_h} h_K^2 \operatorname{osc}(f,K)^2.$$

If the data oscillation is small, i.e., if the details of the right hand side f are resolved by the current finite element mesh, then one may neglect the last term in this bound. Then, the a-posteriori error estimator is computable by exploiting a quadrature formula which is exact for piecewise linear or bilinear functions, respectively.

To prove lower bounds, we need the following bubble property which is a well known technique (e.g. [18, 28]) in the derivation of error estimates.

Lemma 5.4 (Bubble property)**.** *Let $V(E)$ denote the space of linear polynomials on E. Then, for all edges E it holds that*

$$h_E\|\psi\|^2_{L^2(E)} \sim \sum_{\lambda\in\mathrm{bubble}(E)} \langle\psi,\varphi_\lambda\rangle^2_{L^2(E)} \quad \text{for all } \psi\in V(E) \tag{5.4}$$

with bubble(E) *denoting the indices of all functions on level $|E|+2$ which are completely supported inside the element, i.e.,*

$$\mathrm{bubble}(E) := \big\{|\lambda| = |E|+2 : (\operatorname{supp}\varphi_\lambda)\big|_E \subset \overline{E}\big\}.$$

The analogous result holds with respect to the elements K, i.e.,

$$h_K^2\|\psi\|^2_{L^2(K)} \sim \sum_{\lambda\in\mathrm{bubble}(K)} \langle\psi,\varphi_\lambda\rangle^2_{L^2(K)} \quad \text{for all } \psi\in V(K) \tag{5.5}$$

with

$$\mathrm{bubble}(K) := \{|\lambda| = |K|+2 : \operatorname{supp}\varphi_\lambda \subset \overline{K}\}.$$

Proof. The lower bounds in (5.4) and (5.5) are straightforward since the Riesz stability induces that

$$h_E\|\psi\|^2_{L^2(E)} \sim \sum_{|\lambda|=|E|+2} \langle\psi,\varphi_\lambda\rangle^2_{L^2(E)} \geq \sum_{\lambda\in\mathrm{bubble}(E)} \langle\psi,\varphi_\lambda\rangle^2_{L^2(E)}, \quad \psi\in V(E),$$

and likewise for elements K.

To show the upper bound of (5.4), we consider the edge $E := [0,4]$ and piecewise linear nodal ansatz functions $\{\varphi_0,\varphi_1,\ldots,\varphi_4\}$ with respect to the step width 1. Assume that there is a linear polynomial $p = \alpha+\beta(x-2)$ such that

$$\|p\|^2_{L^2([0,1])} = 1 \quad \text{and} \quad \sum_{i=1}^{3} \langle p,\varphi_i\rangle^2_{L^2(E)} = 0.$$

The latter condition leads to $\langle p,\varphi_i\rangle_{L^2(E)} = 0$, $i=1,2,3$, which, in view of

$$p = (\alpha-2\beta)\varphi_0 + (\alpha-\beta)\varphi_1 + \alpha\varphi_2 + (\alpha+\beta)\varphi_3 + (\alpha+2\beta)\varphi_4,$$

is equivalent to

$$\begin{bmatrix} 1/6 & 2/3 & 1/6 & & \\ & 1/6 & 2/3 & 1/6 & \\ & & 1/6 & 2/3 & 1/6 \end{bmatrix} \begin{bmatrix} \alpha-2\beta \\ \alpha-\beta \\ \alpha \\ \alpha+\beta \\ \alpha+2\beta \end{bmatrix} = \begin{bmatrix} \alpha-\beta \\ \alpha \\ \alpha+\beta \end{bmatrix} = \mathbf{0}.$$

However, this equation holds only if $\alpha = \beta = 0$, which is a contradiction to $\|p\|_{L^2([0,1])} = 1$. This proves that

$$\sum_{i=1}^{3} \langle p, \varphi_i \rangle^2_{L^2(E)} \geq \gamma \|p\|^2_{L^2(E)}$$

for some $\gamma > 0$, which is the desired upper bound in case of edges.

By using standard tensor product arguments, we obtain the assertion in case of parallelogrammic elements K. In case of triangular elements, a direct calculation yields the assertion likewise to above, since $V(K)$ is three dimensional while $\mathrm{bubble}(K)$ consists of three linearly independent ansatz functions. □

Similar results can be obtained without major difficulties for other kind of finite element basis functions as well. In this regard, one has only to consider a reference triangle and sufficiently many refinement steps. We would like to skip the details and confine ourselves to the present setting.

Theorem 5.5 (Lower bound)**.** *There holds*

$$\begin{aligned}\mathrm{BPX}(\mathbf{r}_\Sigma)^2 + \sum_{K\in\mathcal{K}_h} h_K^2 \|q_h\|^2_{L^2(K)} + \sum_{E\in\mathcal{E}_h} h_E \|\delta\|^2_{L^2(E)} \\ \lesssim \|r\|^2_{H^{-1}(\Omega)} + \sum_{K\in\mathcal{K}_h} h_K^2 \operatorname{osc}(f,K)^2.\end{aligned}$$

Proof. It holds due to the Lemma 5.4 that

$$\begin{aligned}\sum_{K\in\mathcal{K}_h} h_K^2 \|q_h\|^2_{L^2(K)} &\sim \sum_{K\in\mathcal{K}_h} \sum_{\lambda\in\mathrm{bubble}(K)} \langle q_h, \varphi_\lambda\rangle^2_{L^2(K)} \\ &= \sum_{K\in\mathcal{K}_h} \sum_{\lambda\in\mathrm{bubble}(K)} \Big(\langle q, \varphi_\lambda\rangle_{L^2(K)} - \langle q - q_h, \varphi_\lambda\rangle_{L^2(K)} \Big)^2 \\ &\lesssim \sum_{K\in\mathcal{K}_h} \sum_{\lambda\in\mathrm{bubble}(K)} \Big(\langle q, \varphi_\lambda\rangle^2_{L^2(K)} + \langle q - q_h, \varphi_\lambda\rangle^2_{L^2(K)} \Big).\end{aligned}$$

Herein, we can estimate

$$\langle q - q_h, \varphi_\lambda\rangle^2_{L^2(K)} \leq \|q - q_h\|^2_{L^2(K)} \|\varphi_\lambda\|^2_{L^2(K)} \lesssim h_K^2 \operatorname{osc}(f,K)^2$$

since $|\lambda| = |K| + 2$ for all $\lambda \in \mathrm{bubble}(K)$. Due to $\#\,\mathrm{bubble}(K) = \mathcal{O}(1)$ and

$$\langle q, \varphi_\lambda\rangle_{L^2(K)} = \langle r, \varphi_\lambda\rangle_{L^2(\Omega)} \quad \text{for all } \lambda \in \mathrm{bubble}(K),$$

we therefore conclude that

$$\begin{aligned}\sum_{K\in\mathcal{K}_h} h_K^2 \|q_h\|^2_{L^2(K)} &\lesssim \sum_{K\in\mathcal{K}_h} \sum_{\lambda\in\mathrm{bubble}(K)} \langle r, \varphi_\lambda\rangle^2_{L^2(\Omega)} + \sum_{K\in\mathcal{K}_h} h_K^2 \operatorname{osc}(f,K)^2 \\ &\leq \sum_{\lambda\in\Lambda} \langle r, \varphi_\lambda\rangle^2_{L^2(\Omega)} + \sum_{K\in\mathcal{K}_h} h_K^2 \operatorname{osc}(f,K)^2 \\ &\sim \|r\|^2_{H^{-1}(\Omega)} + \sum_{K\in\mathcal{K}_h} h_K^2 \operatorname{osc}(f,K)^2. \qquad (5.6)\end{aligned}$$

Since the jump terms at the edges are at most linear functions, the bubble property (5.4) applied to the spaces $V_j\big|_E$ yields

$$\sum_{E\in\mathcal{E}_h} h_E\|\delta\|^2_{L^2(E)} \sim \sum_{E\in\mathcal{E}_h}\sum_{\lambda\in\mathrm{bubble}(E)} \langle\delta,\varphi_\lambda\rangle^2_{L^2(E)}. \tag{5.7}$$

We shall now use that each edge E belongs to two neighbouring elements $K_1(E)$ and $K_2(E)$ whose levels differ at most by one from the edge level. Hence, for $\lambda\in\mathrm{bubble}(E)$, we arrive at the estimate

$$\begin{aligned}
\langle\delta,\varphi_\lambda\rangle^2_{L^2(E)} &= \langle\delta,\varphi_\lambda\rangle^2_{L^2(\Omega)}\\
&= \Big(\langle r,\varphi_\lambda\rangle_{L^2(\Omega)} - \langle q-q_h,\varphi_\lambda\rangle_{L^2(K_1(E)\cup K_2(E))} - \langle q_h,\varphi_\lambda\rangle_{L^2(K_1(E)\cup K_2(E))}\Big)^2\\
&\lesssim \langle r,\varphi_\lambda\rangle^2_{L^2(\Omega)} + \langle q-q_h,\varphi_\lambda\rangle^2_{L^2(K_1(E))} + \langle q-q_h,\varphi_\lambda\rangle^2_{L^2(K_2(E))}\\
&\qquad + \langle q_h,\varphi_\lambda\rangle^2_{L^2(K_1(E))} + \langle q_h,\varphi_\lambda\rangle^2_{L^2(K_2(E))}\\
&\lesssim \langle r,\varphi_\lambda\rangle^2_{L^2(\Omega)} + h^2_{K_1(E)}\|q-q_h\|^2_{L^2(K_1(E))} + h^2_{K_2(E)}\|q-q_h\|^2_{L^2(K_2(E))}\\
&\qquad + h^2_{K_1(E)}\|q_h\|^2_{L^2(K_1(E))} + h^2_{K_2(E)}\|q_h\|^2_{L^2(K_2(E))}.
\end{aligned}$$

Inserting this expression into (5.7) and replacing $\|q-q_h\|_{L^2(K)}$ by $\mathrm{osc}(f,K)$, we arrive at

$$\begin{aligned}
\sum_{E\in\mathcal{E}_h} h_E\|\delta\|^2_{L^2(E)} &\lesssim \sum_{E\in\mathcal{E}_h}\sum_{\lambda\in\mathrm{bubble}(E)} \Big(\langle r,\varphi_\lambda\rangle^2_{L^2(\Omega)} + h^2_{K_1(E)}\,\mathrm{osc}\big(f,K_1(E)\big)^2\\
&\qquad + h^2_{K_2(E)}\,\mathrm{osc}\big(f,K_2(E)\big)^2 + h^2_{K_1(E)}\|q_h\|^2_{L^2(K_1(E))} + h^2_{K_2(E)}\|q_h\|^2_{L^2(K_2(E))}\Big)\\
&\lesssim \sum_{\lambda\in\Lambda}\langle r,\varphi_\lambda\rangle^2_{L^2(\Omega)} + \sum_{K\in\mathcal{K}_h} h^2_K\|q_h\|^2_{L^2(K)} + \sum_{K\in\mathcal{K}_h} h^2_K\,\mathrm{osc}(f,K)^2.
\end{aligned}$$

By virtue of (3.5) and (5.6), we finally get

$$\sum_{E\in\mathcal{E}_h} h_E\|\delta\|^2_{L^2(E)} \lesssim \|r\|^2_{H^{-1}(\Omega)} + \sum_{K\in\mathcal{K}_h} h^2_K\,\mathrm{osc}(f,K)^2.$$

This finishes the proof in view of $\mathrm{BPX}(\mathbf{r}_\Sigma)\lesssim\|r\|_{H^{-1}(\Omega)}$. □

Hence, under the usual assumption that the data oscillation is resolved, we have proved a computable a-posteriori bound on the error for an arbitrarily given finite element function u_h which is efficient and reliable.

Remark 5.6. *The discrete BPX scheme for computing the term* $\mathrm{BPX}(\mathbf{r}_\Sigma)$ *can be replaced by any other method which provides the* $H^{-1}(\Omega)$*-norm of the discrete residual* $\mathbf{r}_\Sigma$*. Specifically, a multiplicative multigrid method can be used as well, see e.g.* [30] *and the references therein.*

6. Hierarchical error estimation

To derive a hierarchical error estimator we shall consider the union of the three subsequent generations Γ of Σ which stem from three uniform subdivisions of the actual mesh, hence $\Gamma\subset\Lambda$ and $\Gamma\cap\Sigma=\emptyset$. Then, despite of the presence of hanging nodes, there are at least five degrees of freedom per edge. Consequently, the set Γ contains for all $K\in\mathcal{K}_h$ and $E\in\mathcal{E}_h$ the bubble functions $\mathrm{bubble}(K)$ and $\mathrm{bubble}(E)$.

Theorem 6.1 (Estimator equivalence)**.** *It holds*

$$\sum_{K\in\mathcal{K}_h} h_K^2\|q_h\|_{L^2(K)}^2 + \sum_{E\in\mathcal{E}_h} h_E\|\delta\|_{L^2(E)}^2 \sim \sum_{\lambda\in\Gamma}\langle r_h,\varphi_\lambda\rangle_{L^2(\Omega)}^2. \tag{6.1}$$

Proof. We prove first that the right hand side is an upper bound of the left hand side.

By using the bubble property (Lemma 5.4) together with the identity

$$\langle q_h,\varphi_\lambda\rangle_{L^2(K)} = \langle r_h,\varphi_\lambda\rangle_{L^2(\Omega)} \quad \text{for all } \lambda\in\text{bubble}(K),$$

we arrive at

$$\sum_{K\in\mathcal{K}_h} h_K^2\|q_h\|_{L^2(K)}^2 \sim \sum_{K\in\mathcal{K}_h}\sum_{\lambda\in\text{bubble}(K)}\langle r_h,\varphi_\lambda\rangle_{L^2(\Omega)}^2 \lesssim \sum_{\lambda\in\Gamma}\langle r_h,\varphi_\lambda\rangle_{L^2(\Omega)}^2. \tag{6.2}$$

Next, mimicking the arguments of the proof of the corresponding bound in Theorem 5.5, we can estimate

$$\begin{aligned}
\sum_{E\in\mathcal{E}_h} h_E\|\delta\|_{L^2(E)}^2 &\sim \sum_{E\in\mathcal{E}_h}\sum_{\lambda\in\text{bubble}(E)}\langle\delta,\varphi_\lambda\rangle_{L^2(\Omega)}^2\\
&\lesssim \sum_{E\in\mathcal{E}_h}\sum_{\lambda\in\text{bubble}(E)}\Big(\langle r_h,\varphi_\lambda\rangle_{L^2(\Omega)} - \langle q_h,\varphi_\lambda\rangle_{L^2(K_1(E)\cup K_2(E))}\Big)^2\\
&\lesssim \sum_{E\in\mathcal{E}_h}\sum_{\lambda\in\text{bubble}(E)}\Big(\langle r_h,\varphi_\lambda\rangle_{L^2(\Omega)}^2 + \langle q_h,\varphi_\lambda\rangle_{L^2(K_1)}^2 + \langle q_h,\varphi_\lambda\rangle_{L^2(K_2)}^2\Big)\\
&\lesssim \sum_{\lambda\in\Gamma}\langle r_h,\varphi_\lambda\rangle_{L^2(\Omega)}^2 + \sum_{K\in\mathcal{K}_h} h_K^2\|q_h\|_{L^2(K)}^2.
\end{aligned}$$

By virtue of the above estimate (6.2), this gives again

$$\sum_{E\in\mathcal{E}_h} h_E\|\delta\|_{L^2(E)}^2 \lesssim \sum_{\lambda\in\Gamma}\langle r_h,\varphi_\lambda\rangle_{L^2(\Omega)}^2.$$

It remains to prove that the right hand side of (6.1) is also a lower bound of the left hand side. To this end, we estimate

$$\sum_{\lambda\in\Gamma}\langle r_h,\varphi_\lambda\rangle_{L^2(\Omega)}^2 \lesssim \sum_{\lambda\in\Gamma}\Big(\langle q_h,\varphi_\lambda\rangle_{L^2(\Omega)}^2 + \langle\delta,\varphi_\lambda\rangle_{L^2(\Omega)}^2\Big).$$

Like in the proof of Theorem 5.1, the gradedness implies

$$\begin{aligned}
\sum_{\lambda\in\Gamma}\langle r_h,\varphi_\lambda\rangle_{L^2(\Omega)}^2 &\lesssim \sum_{K\in\mathcal{K}_h}\sum_{\lambda\in\Gamma}\langle q_h,\varphi_\lambda\rangle_{L^2(K)}^2 + \sum_{E\in\mathcal{E}_h}\sum_{\lambda\in\Gamma}\langle\delta,\varphi_\lambda\rangle_{L^2(E)}^2\\
&\lesssim \sum_{K\in\mathcal{K}_h} h_K^2\|q_h\|_{L^2(K)}^2 + \sum_{E\in\mathcal{E}_h} h_E\|\delta\|_{L^2(E)}^2,
\end{aligned}$$

which completes the proof. □

Remark 6.2. *In case of triangles, the jump term δ is constant relative to the edges, whereas the interior residual q_h is an elementwise linear polynomial. Since only one refinement step is required to prove Lemma 5.4 in case of constant functions ψ on edges, in case of triangles two refinement steps would suffice to create all bubble functions that are necessary to catch the essential terms of the residual.*

Combining (6.1) with Corollary 5.3 and Theorem 5.5 implies the following statement for the computation of the complete residual.

Corollary 6.3 (Hierachical error estimator)**.** *Let Φ_Ξ be the finite element basis which stems from three dyadic subdivisions of the actual mesh, i.e., $\Xi \subset \Gamma$ is the finest mesh related to Γ. Let*

$$\mathbf{r}_\Xi := \Big[\langle r_h, \varphi_\lambda\rangle_{L^2(\Omega)}\Big]_{\lambda\in\Xi}$$

denote the discrete residual computed with respect to this finite element basis. Then, there holds

$$\begin{aligned}\mathrm{BPX}(\mathbf{r}_\Xi)^2 - \sum_{K\in\mathcal{K}_h} h_K^2\,\mathrm{osc}(f,K)^2 \\ \lesssim \|r\|^2_{H^{-1}(\Omega)} \lesssim \mathrm{BPX}(\mathbf{r}_\Xi)^2 + \sum_{K\in\mathcal{K}_h} h_K^2\,\mathrm{osc}(f,K)^2.\end{aligned}$$

Remark 6.4. *If we replace the right hand side f by its L^2-orthogonal projection f_h onto V_h, then the term with the data oscillation disappears while the finite element solution u_h does not change. It follows then simply*

$$\|r\|_{H^{-1}(\Omega)} \sim \mathrm{BPX}(\mathbf{r}_\Xi).$$

Due to the nonlocal nature of the BPX multilevel scheme, this error estimator is non-local, as it is the case for all estimators presented here if an algebraic residual is apparent. Therefore, we have to address how the present estimators could be used for an adaptive refinement scheme.

Of course, the refinement will be steered only by the error on the present finest grid Ξ. Perhaps there are several possibilities to perform this. For example, one may consider only the fine level portions, which are local, and use the usual marking strategy [17]. One may also iterate without ' mesh refinement as long as the algebraic error is dominating the estimated error.

Presently, we advice to a strategy which mimics the best N-term strategy of wavelet schemes, allowing also for a coarsening. Indeed, by going into frame coordinates one can directly apply optimal adaptive algorithms as in the multi-scale framework in terms of wavelet bases [10, 11, 19], see also [13, 14, 26] for a direct treatment in frame coordinates.

A simple strategy is given by switching between frame coordinates and finite element bases. One step of a possible realization might be formulated as follows:

(*i*) Refine the given mesh two times in case of triangles and three times in case of parallelograms, respectively, to get the finite element basis Ξ. For the tree

$$\Theta := \Xi \cup \{\text{all predecessors of } \Xi\}$$

compute the discrete residual in frame coordinates

$$\mathbf{r} := \big[\langle r, \varphi_\lambda\rangle_{L^2(\Omega)}\big]_{\lambda\in\Theta}.$$

(*ii*) For a given $0 < \vartheta < 1$ compute the smallest graded tree $\Sigma \subset \widehat{\Theta} \subset \Theta$ such that

$$\big\|\mathbf{r}|_{\widehat{\Theta}}\big\|_2 \geq \vartheta\|\mathbf{r}\|_2. \tag{6.3}$$

(*iii*) Refine the finite element mesh by taking into account the set of leaves of $\widehat{\Theta}$ and solve the associated Galerkin system with sufficient accuracy.

The tree computation in item (ii) can efficiently be performed by the tree algorithms proposed in [5]. The above algorithm will converge due to (6.3) and the fact that $\|\mathbf{r}\|_2 \sim \|r\|_{H^{-1}(\Omega)}$.

7. Conclusion

In this article, we provided a new proof for the standard error estimator if no Galerkin orthogonality is given. This proof is based on an infinite BPX scheme which is applied to the residual. The derived error estimator consists of the sum of the standard element and edge based estimator and H^{-1}-norm of the the discrete residual, given in terms of the BPX scheme. As a byproduct of the present analysis, hierarchical error estimation is shown for the first time to be reliable and efficient.

Acknowledgement. We acknowledge the support by the Sino-German Science Center (grant id 1228) on the occasion of the Chinese-German Workshop on Computational and Applied Mathematics in Augsburg 2015.

References

[1] M. Ainsworth and J.T. Oden. *A Posteriori Error Estimation in Finite Element Analysis.* Pure and Applied Mathematics. John Wiley & Sons, New York, 2000.
[2] I. Babuška and W.C. Rheinboldt. A-posteriori error estimates for the finite element method. *J. Numer. Meth. Engrg.* **12** (1978) 1597–1615.
[3] I. Babuška and A. Miller. A feedback finite element method with a-posteriori estimation, Part I: the finite element method and some basic properties of the a-posteriori error estimator. *Comput. Methods Appl. Mech. Engrg.* **61** (1987) 1–40.
[4] W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Differential Equations.* Lectures in Mathematics — ETH Zürich, Birkhäuser, Basel, 2003.
[5] P. Binev and R. DeVore. Fast computation in adaptive tree approximation. *Numer. Math.* **97** (2004) 193–217.
[6] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.* **97** (2004) 219–268.
[7] J. Bramble, J. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comput.* **55** (1990) 1–22.
[8] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics.* 3rd edition. Cambridge University Press, 2007.
[9] C. Carstensen. A unifying theory of a posteriori finite element error control. *Numer. Math.* **100** (2005) 617–637.
[10] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods for elliptic operator equations — Convergence rates. *Math. Comput.* **70** (2001) 27–75.
[11] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods II — Beyond the elliptic case. *Found. Comput. Math.* **2** (2002) 203–245.
[12] S. Dahlke, W. Dahmen, R. Hochmuth, and R. Schneider. Stable multiscale bases and local error estimation for elliptic problems. *Appl. Numer. Math.* **23** (1997), 21–47.
[13] S. Dahlke, M. Fornasier, and T. Raasch. Adaptive frame methods for elliptic operator equations *Adv. Comput. Math.* **27** (2007) 27–63.
[14] S. Dahlke, M. Fornasier, T. Raasch, R. Stevenson, and M. Werner. Adaptive frame methods for elliptic operator equations: The steepest descent approach. *IMA J. Numer. Math.* **27** (2007), 717–740.
[15] W. Dahmen. Wavelet and multiscale methods for operator equations. *Acta Numerica* **6** (1997) 55–228.
[16] W. Dörfler. A robust adaptive strategy for the nonlinear Poisson equation. *Computing* **55** (1995) 289–304.
[17] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.* **33** (1996) 1106–1124.
[18] W. Dörfler and V. Heuveline. Convergence of an adaptive *hp* finite element strategy in one space dimension. *Appl. Numer. Math.* **57** (2007) 1108–1124.

[19] T. Gantumur, H. Harbrecht, and R. Stevenson. An optimal adaptive wavelet method without coarsening of the iterands. *Math. Comput.* **76** (2007) 615–629.
[20] M. Griebel. Multilevel algorithms considered as iterative methods on semidefinite systems. *SIAM J. Sci. Comput.* **15** (1994), 547–565.
[21] H. Harbrecht, R. Schneider, and C. Schwab. Multilevel frames for sparse tensor product spaces. *Numer. Math.* **110** (2008) 199–220.
[22] P. Morin, R. Nochetto, and K. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.* **38** (2000) 466–488.
[23] P. Oswald. *Multilevel finite element approximation. Theory and applications.* Teubner Skripten zur Numerik. B.G. Teubner, Stuttgart, 1994.
[24] R. Rannacher, A. Westenberger, and W. Wollner. Adaptive finite element solution of eigenvalue problems: Balancing of discretization and iteration error. *J. Numer. Math.* **18** (2010) 303–327.
[25] R. Sacchi and A. Veeser. Locally efficient and reliable a posteriori error estimators for Dirichlet problems. *Math. Mod. Meth. Appl. Sci.* **16** (2006) 319–346.
[26] R. Stevenson. Adaptive solution of operator equations using wavelet frames. *SIAM J. Numer. Anal.* **41** (2003) 1074–1100.
[27] R. Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.* **7** (2007) 245–269.
[28] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinenment Techniques.* Wiley–Teubner, Chichester, 1996.
[29] R. Verfürth. *A Posteriori Error Estimation Techniques for Finite Element Methods.* Oxford University Press, Oxford, 2013.
[30] H. Wu and Z. Chen. Uniform convergence of multigrid V-cycle on adaptively refined finite element meshes for second order elliptic problems. *Sci. China Ser. A* **49** (2006) 1405–1429.

Helmut Harbrecht, Departement Mathematik und Informatik, Universität Basel, Spiegelgasse 1, 4051, Basel, Schweiz
E-mail address: helmut.harbrecht@unibas.ch

Reinhold Schneider, Institut für Mathematik, Technische Universität Berlin, Strasse des 17. Juni 136, 10623 Berlin, Deutschland
E-mail address: schneidr@math.tu-berlin.de