AN OPTIMAL ADAPTIVE MIXED FINITE ELEMENT METHOD

CARSTEN CARSTENSEN AND HELLA RABUS

Abstract. Various applications in fluid dynamics and computational continuum mechanics motivate the development of reliable and efficient adaptive algorithms for mixed finite element methods. In order to save degrees of freedom, not all but just a selection of finite element domains are refined. Hence the fundamental question of convergence as well as the question of optimality require new mathematical arguments. The presented adaptive algorithm for Raviart-Thomas mixed finite element methods solves the Poisson model problem, with optimal convergence rate.

1. Introduction

This paper suggests an optimal adaptive mixed finite element algorithm AMFEM for the Poisson model problem with unknown flux $p$ and primal variable $u$ with

$$ p + \nabla u = 0 \quad \text{and} \quad \text{div} \ p = f \quad \text{in} \ \Omega, \quad \text{while} \quad u = 0 \quad \text{on} \ \partial \Omega. $$

Given $f \in L^2(\Omega)$, the mixed variational formulation reads: Seek $(p, u) \in H(\text{div}, \Omega) \times L^2(\Omega)$ such that, for all $(q, v) \in H(\text{div}, \Omega) \times L^2(\Omega)$, it holds

$$ (p, q)_{L^2(\Omega)} = (\text{div} \ q, u)_{L^2(\Omega)} \quad \text{and} \quad (\text{div} \ p, v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)}. $$

Given finite-dimensional piecewise polynomial subspaces $RT_0(T_\ell) \subseteq H(\text{div}, \Omega)$ and $P_0(T_\ell) \subseteq L^2(\Omega)$, named after Raviart and Thomas and described in Section 2, the discrete problem reads: Seek $(p_\ell, u_\ell) \in RT_0(T_\ell) \times P_0(T_\ell)$ such that, for all $(q_\ell, v_\ell) \in RT_0(T_\ell) \times P_0(T_\ell)$,

$$ (p_\ell, q_\ell)_{L^2(\Omega)} - (\text{div} \ q_\ell, u_\ell)_{L^2(\Omega)} + (\text{div} \ p_\ell, v_\ell)_{L^2(\Omega)} = (f, v_\ell)_{L^2(\Omega)}. $$

The existence and uniqueness of a discrete solution follows from the inf-sup condition [12].

The a posteriori error control of mixed finite element methods dates back to the independent work of Alonso [2] and Carstensen [14]; the error reduction and convergence for adaptive mixed finite element methods was established by Carstensen and Hoppe [19, 18]. Chen, Holst, and Xu proved convergence and optimality [22] of...
adaptive mixed finite element methods following arguments of Rob Stevenson for the conforming finite element method. Their algorithm reduces oscillations separately, before approximating the solution by some adaptive algorithm in the spirit of W. Dörfler [23].

In an independent parallel work to this one here, Becker and Mao designed an alternative optimal algorithm [6], which switches between reduction of the edge-based error estimator and reduction of the oscillations via separate bulk chasing. Since this may lead to successive loops of oscillation reduction, a potentially more effective strategy is presented and analysed here. In the first case, when oscillations are small compared to the estimated error, edge-oriented Dörfler marking is applied, while in the second case we make use of an optimal algorithm, the Thresholding Second Algorithm (TSA) by Binev, Dahmen and DeVore [7] to reduce oscillations.

The main contribution of this paper is to show optimal convergence for a sequence of triangulations arising from two completely different strategies, i.e., Dörfler marking plus TSA. After each level of reducing oscillations, two triangulations refined independently from a common coarse triangulation have to be combined. This combination poses a new challenge for proving optimality as presented and beaten in the following sections.

The oscillations on level \( \ell \) are defined for a regular triangulation \( \mathcal{T}_\ell \) of \( \Omega \) into triangles by
\[
\text{osc}_\ell := \text{osc}(f, \mathcal{T}_\ell) := \left( \sum_{T \in \mathcal{T}_\ell} \text{osc}^2(f, T) \right)^{1/2}
\]
with local oscillations on the domain \( \omega \subseteq \Omega \) of area \( |\omega| \) defined as
\[
\text{osc}(f, \omega) := |\omega|^{1/2} \| f - f_\omega \|_{L^2(\omega)} \quad \text{with}
\]
\[
f_\omega := \int_\omega f \, dx := |\omega|^{-1} \int_\omega f(x) \, dx.
\]
For the set of all interior edges \( \mathcal{E}_\ell \) in the triangulation \( \mathcal{T}_\ell \), the edge-error estimator reads
\[
\eta_\ell := \eta_\ell(\mathcal{E}_\ell) \quad \text{with} \quad \eta_\ell^2(\mathcal{M}) := \sum_{E \in \mathcal{M}} \eta_\ell^2(E) \quad \text{for} \quad \mathcal{M} \subseteq \mathcal{E}_\ell
\]
and local contributions
\[
\eta_\ell(E) := |E|^{1/2} \| [p_\ell]_E \|_{L^2(E)} \quad \text{for all} \quad E \in \mathcal{E}_\ell.
\]
Here and in the sequel, \([q]_E := q|_{T_+} \! - \! q|_{T_-}\) denotes the jump of \(q\) across an edge \(E = T_+ \cap T_-\) shared by the two elements \(T_\pm \in \mathcal{T}_\ell\), and \(\nu_E = \nu_{T_+}\) is the unit normal vector exterior to \(T_+\) along \(E\). Note that the normal component \([p_\ell] \cdot \nu_E\) vanishes because of \(p_\ell \in H(\text{div}, \Omega)\), and so \(\eta_\ell(E)\) solely measures the jump parallel to \(E\).

The marking consists of the two alternatives \((A)\) and \((B)\) depending on the computable \(\text{osc}_\ell\) and \(\eta_\ell\) and some global parameter \(\kappa > 0\).

In Case \((A)\) \(\text{osc}^2_\ell \leq \kappa \eta^2_\ell\), compute \(\mathcal{M}_\ell \subseteq \mathcal{E}_\ell\) with minimal cardinality \(|\mathcal{M}_\ell|\) such that
\[
\theta_A \eta^2_\ell \leq \eta^2_\ell(\mathcal{M}_\ell)
\]
and compute a shape-regular triangulation $\mathcal{T}_{t+1}$, where the edges in $\mathcal{M}_t$ are bisected plus a minimal number of other edges in some closure algorithm to avoid hanging nodes.

In Case (B) $\text{osc}_t^2 > \kappa \eta^2$, run the Thresholding Second Algorithm from [7, 8] resulting in a regular triangulation $\mathcal{T}$ of nearly minimal cardinality $|\mathcal{T}|$ such that

$$\text{osc}(f, \mathcal{T}) \leq \text{ Tol}^2 := \rho_B \text{osc}_t^2,$$

and compute the overlay $\mathcal{T}_{t+1} := \mathcal{T}_t \oplus \mathcal{T}$.

The Thresholding Second Algorithm of Section 4.4 of [7] is one possible example; the point here is to enforce an oscillation reduction with optimal complexity independent of the refinements of Figure 2.1 to keep the number of levels small.

The algorithm is feasible in the sense that the decision in MARK is solely based on computed quantities and realises the simultaneous reduction of $\eta_t$ and $\text{osc}_t$. This avoids the computation of an initial triangulation $\mathcal{T}_0$, which approximates the data up to a given fixed tolerance as in [22].

The main theorem states optimal complexity (Theorem 5.8) for AMFEM as defined in detail in Section 2 for particular positive parameters $\alpha, \beta, \kappa$, and $0 < \theta_A, \rho_B < 1$. For $(p, f)$ in some approximation class $\mathcal{A}_s$ and the sequence of triangulations $(\mathcal{T}_t)_t$ from AMFEM with discrete fluxes $(p_t)_t$, Theorem 5.8 implies that

$$|\mathcal{T}| - |\mathcal{T}_0| \lesssim (\varepsilon_t^2 + \text{osc}_t^2)^{-1/(2s)} \approx \xi_t^{-1/s},$$

which is optimal with respect to $\mathcal{A}_s$ up to a multiplicative generic constant. The proof is based on overlay control (Theorem 3.3) and contraction (Lemma 5.2). In particular, for $\eta_t$, and $\text{osc}_t$ as defined above and the exact error $\varepsilon_t := \|p - p_t\|_{L^2(\Omega)}$ of the flux there exists $0 < \rho < 1$ such that contraction holds for the weighted term

$$\eta_{t+1}^2 + \alpha \varepsilon_{t+1}^2 + \beta \text{osc}_{t+1}^2 =: \xi_{t+1}^2 \leq \rho \xi_t^2.$$

Here and in the sequel, $|\mathcal{T}| := \text{card}(\mathcal{T}_t)$ denotes the number of elements in the finite set $\mathcal{T}_t$, and $A \leq B$ represents $A \leq CB$ for some mesh-independent, positive generic constant $C$, whereas $A \approx B$ represents $A \leq B \leq A$. Moreover, the standard notation of Lebesgue and Sobolev spaces is employed; e.g., the differential operators are defined for vector-valued functions $v(x) \in \mathbb{R}^2$ for all $x \in \Omega$ as

$$\text{div} v := \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2}, \quad \text{curl} v := \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2},$$

and for scalar-valued functions $v$ for all $x \in \Omega$

$$\text{Curl} v := \left( \frac{\partial v}{\partial x_2}, -\frac{\partial v}{\partial x_1} \right).$$

The remaining part of the paper is organised as follows. Section 2 introduces notation and describes further details of the proposed AMFEM. The focus is on the optimal oscillation reduction in MARK with the concept of the overlay $\mathcal{T}_t \oplus \mathcal{T}$, defined as the coarsest common regular refinement of both $\mathcal{T}_t$ and $\mathcal{T}$. Section 3 is based on the notion of trees and forests to represent refined meshes and overlays and to combine the control of the number of elements in both Cases (A) and (B). For the finite sequence $\mathcal{M}_t^{(0)}, \ldots, \mathcal{M}_t^{(K(t))}$ of sets of edges from Algorithm 3.2 with

$$\mathcal{T}_t \oplus \mathcal{T} = \text{Refine} \left( \mathcal{T}_t, (\mathcal{M}_t^{(k)})_{k=0,\ldots,K(t)} \right),$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
the key estimate in Theorem 3.3 guarantees that
\[ \sum_{k=0}^{K(\ell)} |M^{(k)}_\ell| \leq |T| - |T_0|. \]

Section 4 introduces the discrete stability and quasiorthogonality for the proof of contraction and optimality in Section 5. A numerical comparison of AMFEM with the adaptive algorithm in [6] concludes the paper.

2. Adaptive mixed finite element method algorithm (AMFEM)

This section is devoted to the design of an adaptive algorithm for the lowest-order mixed finite element method (MFEM) for solving the Poisson model problem (1.2).

2.1. Outline of the adaptive algorithm. Let \( T_0 \) be a regular, initial coarse triangulation of \( \Omega \) into closed triangles, where two distinct elements are either disjoint or share exactly one node or one common edge. Moreover, each element of \( T \in T_0 \) has at least one node in the interior of \( \Omega \). For any \( T \in T_0 \), one edge from the set of its interior edges \( E(T) \) is selected and called its reference edge \( E(T) \).

In successive loops of the basic steps Solve, Estimate, Mark and Refine, discrete solutions \( (p_\ell, u_\ell) \in RT_0(T_\ell) \times P_0(T_\ell) \) are computed on each level \( \ell \geq 0 \) based on the current shape-regular triangulation \( T_\ell \) of \( \Omega \) with the sets of its nodes \( N_\ell \), free nodes \( K_\ell := \Omega \cap N_\ell \) and interior edges \( E_\ell \). The adaptive algorithm is based on a combination of an edge-based error estimator and oscillation control in the step Mark described in the sequel.

2.2. Solve. The Poisson model problem (1.2) is solved on the current triangulation \( T_\ell \) with the space of Raviart-Thomas finite elements of lowest order for the triangulation \( T_\ell \) of level \( \ell \) [9, 12], namely
\[ RT_0(T_\ell) := \{ q \in \text{H(div, } \Omega) \mid \forall T \in T_\ell \exists a \in \mathbb{R}^2 \exists b \in \mathbb{R} \forall x \in T \ q(x) = a + bx \}. \]

Matlab implementations and documentations of Solve are provided in [5]. Notice that, in particular,
\[ \text{div } p_\ell + f_\ell = 0 \text{ a.e. in } \Omega \]
for the piecewise integral mean \( f_\ell \in P_0(T_\ell) \) defined by \( f_\ell|_T = f_T = \frac{1}{|T|} \int_T f(x) \, dx \) for all \( T \in T_\ell \).

2.3. Estimate. The error estimator \( \eta_\ell \) of (1.4) and the oscillations \( \text{osc}_\ell \) of (1.3) allow reliable and efficient error control on the given triangulation \( T_\ell \).

Theorem 2.1 ([2, 17, 14]). The error estimator \( \eta_\ell \) of (1.4) plus the oscillations \( \text{osc}_\ell \) of (1.3) are reliable and efficient in the sense that there exist positive constants \( C_{\text{eff}}, C_{\text{rel}} \), which depend on the shape but not on the size of the element domains, with
\[ C_{\text{eff}} \eta_\ell^2 \leq \| p - p_\ell \|^2_{L^2(\Omega)} \leq C_{\text{rel}}(\eta_\ell^2 + \text{osc}_\ell^2). \]
2.4. **Adaptive mesh-refinement:** Mark and Refine. Given parameters $0 < \theta_A, \rho_B < 1$, and positive $\kappa$, the algorithm distinguishes Cases (A) for $\text{osc}_\ell^2 \leq \kappa \eta_\ell^2$ and (B) for $\kappa \eta_\ell^2 < \text{osc}_\ell^2$.

The nonempty set $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell$ is specified by Mark and used in Refine to compute $\mathcal{T}_{\ell+1}$ by the Newest Vertex Bisection (NVB) and completion from [8, 7, 10, 32] with respect to $\mathcal{M}_\ell$.

2.5. **Case (A) for** $\text{osc}_\ell^2 \leq \kappa \eta_\ell^2$. Case (A) performs an error estimator reduction: Sort the set of all edges $\mathcal{E}_\ell$ in $(\mathcal{E}_1, \ldots, \mathcal{E}_N)$ with $\eta_{\mathcal{E}_1} \leq \cdots \leq \eta_{\mathcal{E}_N}$. Compute a set $\mathcal{M}_\ell := \{ E_{N-k}, \ldots, E_N \}$ of minimal cardinality $|\mathcal{M}_\ell| = k + 1$ with

$$\theta_A \eta_\ell^2 \leq \eta_{\mathcal{E}_1}^2 (\mathcal{M}_\ell).$$

A possible choice of refinement rules *green*, *blue*, and *bisec3* is depicted in Figure 2.1. The refined triangulation $\mathcal{T}_{\ell+1} := \text{REFINE}(\mathcal{T}_\ell, \mathcal{C}(\mathcal{M}_\ell))$ from $\mathcal{T}_\ell$ is uniquely defined in the way, such that exactly the edges in $\mathcal{C}(\mathcal{M}_\ell)$ are bisected. $\mathcal{C}(\mathcal{M}_\ell)$ is the minimal subset of $\mathcal{E}_\ell$ which includes $\mathcal{M}_\ell$ and is closed in the sense that

$$\{ E(T) \in \mathcal{E}_\ell \mid T \in \mathcal{T}_\ell \text{ and } \mathcal{E}(T) \cap \mathcal{C}(\mathcal{M}_\ell) \neq \emptyset \} \subseteq \mathcal{C}(\mathcal{M}_\ell).$$

2.6. **Case (B) for** $\text{osc}_\ell^2 > \kappa \eta_\ell^2$. Case (B) reduces the oscillations: Given $f, \mathcal{T}_0$, $0 < \rho_B < 1$, set $\text{Tol} := \rho_B^{1/2} \text{osc}_\ell$ and run Thresholding Second Algorithm plus completion [7] [8] to compute an optimal $\mathcal{T}$ with

$$\text{osc}^2(f, \mathcal{T}) \leq \text{Tol}^2 \text{ and } |\mathcal{T}| - |\mathcal{T}_0| \lesssim \text{Tol}^{-1/s}.$$ 

Hence, the regular overlay triangulation $\mathcal{T}_{\ell+1} := \mathcal{T} \oplus \mathcal{T}_\ell$, computed by means of the corresponding forests in Section 3 satisfies

$$\text{osc}_\ell^{\ell+1} \leq \text{Tol}^2.$$ 

By definition of Tol in each level of Case (B), an oscillation reduction with $0 < \rho_B < 1$ holds:

$$\text{osc}_\ell^{\ell+1} \leq \rho_B \text{osc}_\ell^2.$$ 

The refinement in (B) is not level-oriented in the sense that one element domain $K$ of $\mathcal{T}_\ell$ might contain a seemingly uncontrolled number of refined element domains.
in \( \{ T \in \mathcal{T}_{\ell+1} \mid T \subseteq K \} \subseteq \mathcal{T}_{\ell+1} \). The control requires the investigations of the subsequent section.

3. Combining the two cases in one optimal algorithm

This section is devoted to the overall control of the number of finite element domains treated in the two separate ways (A) and (B).

3.1. Forests representing triangulations. This subsection briefly recalls the concepts of trees and forests from [7] to clarify the notion of overlays and to embed oscillation reduction into the successive loops of the estimator reduction in both Cases (B) and (A).

A rooted tree is a graph, where one vertex is designated to be the root and any two vertices are connected by exactly one path. If two vertices are connected by an edge, the vertex closer to the root is called the parent, the other its child. A vertex with at least one child is called an interior vertex and otherwise a leaf. A pairwise disjoint set of trees is called a forest.

This paper focuses on regular triangulations generated from some coarse regular triangulation \( T_0 \) by NVB with refinements of Figure 2.1. Any possible triangulation \( T_\ell \) refined from \( T_0 \) is represented by one forest \( \mathcal{F}_\ell \), and the refinement of each triangle \( T \) of the initial triangulation \( T_0 \) corresponds to one tree with root \( T \) in \( \mathcal{F}_\ell \). The leaves of all trees of \( \mathcal{F}_\ell \) represent the elements of \( T_\ell \). Each vertex in \( \mathcal{F}_\ell \) has either two children or none and all but the root in \( T_0 \) have an ancestor.

In steps Mark and Refine of Case (A) in Amfem, a refined triangulation is computed by marking a set of edges, followed by Closure (cf. paragraph 3.4) and NVB. In Case (B), however, the current triangulation \( T_\ell \) is overlaid with a triangulation \( T \) represented by a forest \( \mathcal{F} \) with \( \text{osc}(f, T)^2 \leq \text{Tol}^2 \). The subsequent subsection explains the definition and the key estimate for the overlay of triangulations. Then, on each level \( \ell \), a triangulation \( T_\ell \) and its forest \( \mathcal{F}_\ell \) are available.

3.2. Overlay of two refinements of \( T_0 \). The coarsest common refinement \( T_\ell \oplus T \) of two regular triangulations \( T_\ell \) and \( T \) refined from \( T_0 \) by NVB, called overlay, is defined by the union of their forests \( \mathcal{F}_\ell \cup \mathcal{F} \). Its number of elements is bounded as stated in the following lemma.

**Lemma 3.1** (21). The overlay \( T_\ell \oplus T \) of two regular triangulations \( T_\ell \) and \( T \) refined from \( T_0 \) by NVB is regular and satisfies

\[
| T_\ell \oplus T | - | T_\ell | \leq | T | - | T_0 |.
\]

\( \square \)

The following algorithm is for theoretical purposes only and allows a common refinement control for both Cases (A) and (B).

**Algorithm 3.2** (Embed Oscillation Control). This algorithm provides a finite sequence of sets of successively marked reference edges \( (M^{(k)}_\ell)_{k=0,\ldots,K(\ell)} \) in order to embed Case (B) of Amfem in the standard level-oriented overall adaptive mesh-refinement. The output of the algorithm realises a finite number of successive refinements, written

\[
T_{\ell+1} := T_\ell \oplus T = \text{Refine}\left(T_\ell, (M^{(k)}_\ell)_{k=0,\ldots,K(\ell)}\right),
\]

where in each step \( k \) each triangulation is refined as shown in Figure 2.1 with respect to the set of marked edges \( M^{(k)}_\ell \) as follows.
Input: Given $\mathcal{T}_\ell$ and $\mathcal{T}_\ell \oplus \mathcal{T}$, set $T^{(0)}_\ell := \mathcal{T}_\ell$, $E^{(0)}_\ell := \mathcal{E}_\ell$, $k := 0$.

Loop: For $k = 0, 1, \ldots$ until termination for $k = K(\ell)$, set

$$\mathcal{M}^{(k)}_\ell := \{ E(T) \in \mathcal{E}^{(k)}_\ell \mid T \in T^{(k)}_\ell \setminus \mathcal{T}_{\ell+1} \}$$

and run NVB to refine $\mathcal{M}^{(k)}_\ell$ in $T^{(k)}_\ell$ with output

$$T^{(k+1)}_\ell := \text{Refine}(T^{(k)}_\ell, \mathcal{M}^{(k)}_\ell).$$

If $T^{(k+1)}_\ell \not\subseteq \mathcal{T}_\ell \oplus \mathcal{T}$, update $k$, else stop with $k = K(\ell)$.

Output: A sequence of reference edges $(\mathcal{M}^{(k)}_\ell)_{k=0, \ldots, K(\ell)}$.

The benefit of the artificial marked edges $\mathcal{M}^{(0)}_\ell, \ldots, \mathcal{M}^{(K(\ell))}_\ell$ in Case (B) is that the refinement (3.3) is level-oriented such that each triangle in $T^{(k)}_\ell$ is refined as shown in Figure 2.1 to obtain $T^{(k+1)}_\ell$.

3.3. Refinement control in Case (B).

Theorem 3.3. Given regular triangulations $\mathcal{T}_\ell$, $\mathcal{T}$ refined from $\mathcal{T}_0$ by NVB, Algorithm 3.2 stops after a finite number of $K(\ell) \geq 0$ steps with

$$T^{(K(\ell)+1)}_\ell = \mathcal{T}_\ell \oplus \mathcal{T}$$

and outputs a finite sequence of sets $(\mathcal{M}^{(k)}_\ell)_{k=0, \ldots, K(\ell)}$ with (3.2) and

$$\sum_{k=0}^{K(\ell)} |\mathcal{M}^{(k)}_\ell| \leq |\mathcal{T}| - |\mathcal{T}_0|.$$

Proof. Let $\mathcal{F}_\ell$, $\mathcal{F}_\ell^{(k)}$, and $\mathcal{F}_{\ell+1} := \mathcal{F}_\ell \cup \mathcal{F}$ denote the forests associated to the triangulations $\mathcal{T}$, $\mathcal{T}_\ell$, $T^{(k)}_\ell$, and $\mathcal{T}_{\ell+1} := \mathcal{T}_\ell \oplus \mathcal{T}$, respectively. By mathematical induction, one observes that NVB leads to a nested sequence

$$\mathcal{F}_\ell = \mathcal{F}_\ell^{(0)} \subseteq \mathcal{F}_\ell^{(1)} \subseteq \ldots \subseteq \mathcal{F}_\ell^{(K(\ell))} \subseteq \mathcal{F}_\ell^{(K(\ell)+1)} = \mathcal{F}_{\ell+1}.$$

In fact, $T^{(k)}_\ell \setminus \mathcal{T}_{\ell+1}$ denotes the triangles and $\mathcal{M}^{(k)}_\ell$ the marked edges to be refined in step $k$. Since $\mathcal{F}_{\ell+1} \setminus \mathcal{F}_\ell$ is finite, Algorithm 3.2 terminates after $K(\ell)$ steps with $\mathcal{F}_\ell^{(K(\ell)+1)} = \mathcal{F}_{\ell+1}$.

For each $E \in \mathcal{M}^{(k)}_\ell$ with $0 \leq k \leq K(\ell)$, at least one element in $T^{(k)}_\ell$ is refined into at least two new elements in $T^{(k+1)}_\ell$. Furthermore, if $E$ is an interior edge, at least two elements in $T^{(k)}_\ell$ are bisected to at least four new elements in $T^{(k+1)}_\ell$. Therefore it follows that

$$|\mathcal{M}^{(k)}_\ell| \leq |T^{(k+1)}_\ell| - |T^{(k)}_\ell|.$$

Recall that $T^{(0)}_\ell = \mathcal{T}_\ell$, $T^{(K(\ell)+1)}_\ell = \mathcal{T}_{\ell+1} = \mathcal{T}_\ell \oplus \mathcal{T}$ and apply (3.1), (3.3) to deduce

$$\sum_{k=0}^{K(\ell)} |\mathcal{M}^{(k)}_\ell| \leq |\mathcal{T}_\ell \oplus \mathcal{T}| - |\mathcal{T}_\ell| \leq |\mathcal{T}| - |\mathcal{T}_0|. \quad \square$$
3.4. Closure. There are several strategies to realise the implementation of refining a mesh by NVB with respect to a set of marked edges $\mathcal{M}_t$. One way is first to run some Closure algorithm to compute the smallest superset $C(\mathcal{M}_t)$ of $\mathcal{M}_t$ with \[2.3\]. Thereafter, second, refine each triangle according to Figure \[2.1\] and apply the indicated definition of reference edges.

The overheat of Closure is bounded in the following sense. A sequence of triangulations $(\mathcal{T}_t)_t$ and corresponding sets of marked edges $(\mathcal{M}_t)_t$ satisfy

$$|\mathcal{T}_t| - |\mathcal{T}_0| \leq C_0 \sum_{j=0}^{\ell-1} |\mathcal{M}_j|$$

for some $C_0 > 0$ depending solely on $\mathcal{T}_0$ \[7, 32\]. This estimate is usually employed in Case (A), but holds in Case (B) in the sense that

$$(3.5) \quad |\mathcal{T}_t| - |\mathcal{T}_0| \leq C_0 \sum_{j=0}^{\ell-1} \sum_{k=0}^{K(j)} |\mathcal{M}_j^{(k)}|,$$

where $\mathcal{M}_j^{(0)}, \ldots, \mathcal{M}_j^{K(j)}$ is the output of Algorithm \[3.2\] in Case (B), and where $\mathcal{M}_j^{(0)} := \mathcal{M}_j$ and $K(j) := 0$ in Case (A).

4. Further preliminaries

This section summarises some key arguments for the contraction property and optimal convergence. In contrast to \[22\], the proof of Lemma \[4.2\] presents a direct verification of the discrete stability based on the following nonstandard Poincaré inequality.

The subsequent analysis of MFEM employs the nonconforming finite element spaces

$$P_{1,0}^{NC}(\mathcal{T}_t) := \{ v \in P_1(\mathcal{T}_t) \mid \forall E \in \mathcal{E}_t, v \text{ is continuous at } \text{mid}(E) \},$$

$$P_{1,0}^{NC}(\mathcal{T}_t) := \{ v \in P_1^{NC}(\mathcal{T}_t) \mid \forall E \in \mathcal{E}_t^D, v(\text{mid}(E)) = 0 \}$$

for the midpoint mid($E$) of any edge $E$, and the set of edges $\mathcal{E}_t^D$ along the boundary $\partial\Omega$.

The following lemmas apply to triangulations $\mathcal{T}_{t+k}$ refined from $\mathcal{T}_t$ in possibly more than one level of refinements from Figure \[2.1\] and to their respective MFEM solutions $(p_t, u_k) \in RT_0(\mathcal{T}_t) \times P_0(\mathcal{T}_t)$, and $(p_{t+k}, u_{t+k}) \in RT_0(\mathcal{T}_{t+k}) \times P_0(\mathcal{T}_{t+k})$ of \[1.2\]. Besides, let $(p, u) \in H(\text{div}, \Omega) \times L^2(\Omega)$ denote the exact solution of \[1.1\].

Furthermore, let $\nabla_{t+k}$ denote the piecewise action of the gradient $\nabla$ on $\mathcal{T}_{t+k}$, $h_T = \text{diam}(T)$ and let $a_{t+k}$ denote the integral mean of some $a_{t+k} \in P_1^{NC}(\mathcal{T}_{t+k})$ on a coarser triangle $T \in \mathcal{T}_t$ used in the following lemmas.

**Lemma 4.1 (Poincaré inequality).** A nonstandard discrete Poincaré inequality for $a_{t+k} \in P_1^{NC}(\mathcal{T}_{t+k})$ and $T \in \mathcal{T}_t$ reads

$$\|a_{t+k} - a_T\|_{L^2(T)} \lesssim h_T \|\nabla_{t+k}a_{t+k}\|_{L^2(T)}.$$

**Proof.** The estimate \[4.1\] is a consequence of the work of Brenner. For a proof, one transforms $T \in \mathcal{T}_t$ and its refined mesh $\mathcal{T}_{t+k}|_T$ onto the reference triangle $T_{\text{ref}}$ with a refined mesh of right isosceles triangles. Furthermore, $a_{t+k}$ and $a_T$ are
transformed onto $T_{ref}$ to $\tilde{a}_{\ell+1}$ and $\tilde{a}_T$, respectively. Hence, the estimate from [11, Theorem 10.6.16] for arbitrary shape-regular meshes, cf. [11, p. 301f], simplifies to

$$
\|\tilde{a}_{\ell+1} - \tilde{a}_T\|_{L^2(T_{ref})} \lesssim \|\nabla_{\ell+1}\tilde{a}_{\ell+1}\|_{L^2(T_{ref})}.
$$

A careful transformation from $T_{ref}$ to $T$ yields the factor $h_T$ in [11].

Lemma 4.2 (Discrete Stability). Let $p_{\ell+1}$ and $\hat{p}_{\ell+1}$ be the flux-part of the discrete MFEM solutions of (1.2) in $RT_0(T_{\ell+1})$ with piecewise constant right-hand sides $f_{\ell+1} \in P_0(T_{\ell+1})$ and $f_{\ell} \in P_0(T_{\ell})$: e.g.,

$$
div p_{\ell+1} + f_{\ell+1} = 0 \quad \text{and} \quad div \hat{p}_{\ell+1} + f_{\ell} = 0.
$$

Then, there exists some constant $C_1 > 0$ (depending solely on the shape-regularity of $T_0$) such that

$$
C_1^{-1}\|p_{\ell+1} - \hat{p}_{\ell+1}\|_{L^2(\Omega)} \leq \left( \sum_{T \in T_{\ell}} h_T^2 \|f_{\ell+1} - f_{\ell}\|_{L^2(T)} \right)^{1/2} \leq \text{osc}_T.
$$

Proof. Let $x_{\ell+1} \in P_0(T_{\ell+1}; \mathbb{R}^2)$ denote the triangle midpoints, $x_{\ell+1}|_T := \text{mid}(T)$ for $T \in T_{\ell+1}$. It is well established [20] that

$$
p_{\ell+1}(x) = \nabla_{\ell+1}u_{\ell+1}^{NC} + (x - x_{\ell+1})f_{\ell+1}/2 \quad \text{for} \quad x \in T \in T_{\ell+1},
$$

$$
\hat{p}_{\ell+1}(x) = \nabla_{\ell+1}u_{\ell+1}^{NC} + (x - x_{\ell+1})f_{\ell}/2 \quad \text{for} \quad x \in T \in T_{\ell+1},
$$

while $u_{\ell+1}^{NC} \in P_{1,0}^{NC}(T_{\ell+1})$ solves

$$
\int_{\Omega} \nabla_{\ell+1}u_{\ell+1}^{NC} \cdot \nabla_{\ell+1}v_{\ell+1}^{NC} \, dx = \int_{\Omega} f_{\ell+1}v_{\ell+1}^{NC} \, dx \quad \text{for all} \quad v_{\ell+1}^{NC} \in P_{1,0}^{NC}(T_{\ell+1}).
$$

When, $(\bullet - x_{\ell+1})$ represents the factor $(x - x_{\ell+1})$ of $x \in \Omega$, the definition $a_{\ell+1} := u_{\ell+1}^{NC} - \tilde{a}_{\ell+1}$ yields

$$
\nabla_{\ell+1}a_{\ell+1} = \nabla_{\ell+1}u_{\ell+1}^{NC} - \nabla_{\ell+1}u_{\ell+1}^{NC} = p_{\ell+1} - \hat{p}_{\ell+1} - (\bullet - x_{\ell+1})(f_{\ell+1} - f_{\ell})/2.
$$

Since $(\bullet - x_{\ell+1}) \perp L^2(\Omega) P_0(T_{\ell})^2$ it follows that

$$
\|p_{\ell+1} - \hat{p}_{\ell+1}\|_{L^2(\Omega)} \leq \|(\bullet - x_{\ell+1})(f_{\ell+1} - f_{\ell})/2\|^2_{L^2(\Omega)} + \|\nabla_{\ell+1}a_{\ell+1}\|_{L^2(\Omega)}^2.
$$

(4.2)

Moreover, an elementwise integration by parts shows that

$$
\|\nabla_{\ell+1}a_{\ell+1}\|_{L^2(\Omega)}^2 = (p_{\ell+1} - \hat{p}_{\ell+1}, \nabla_{\ell+1}a_{\ell+1})_{L^2(\Omega)}
$$

$$
= - (div (p_{\ell+1} - \hat{p}_{\ell+1}), a_{\ell+1})_{L^2(\Omega)}
$$

$$
+ \sum_{E \in E_{\ell+1}} \int_E (p_{\ell+1} - \hat{p}_{\ell+1}) \cdot \nu_E |a_{\ell+1}|_E \, ds
$$

$$
= \sum_{T \in T_{\ell}} \int_T (f_{\ell+1} - f_{\ell})(a_{\ell+1} - a_T) \, dx.
$$

The last identity follows for $T \in T_{\ell}$ and $E \in E_{\ell+1}$ from

$$
\int_T (f_{\ell+1} - f_{\ell}) \, dx = 0 = \int_E |\tilde{a}_{\ell+1}|_E \, ds.
$$
The combination of (1.1) of Lemma 4.2 with the aforementioned arguments and orthogonality in (1.2) leads to

\[ \| \nabla \ell \kappa a \ell + k \|_{L^2(\Omega)}^2 \lesssim \left( \sum_{T \in T_0} h_T^2 \| f_{\ell+k} - f_\ell \|_{L^2(T)}^2 \right)^{1/2} \| \nabla \ell \kappa a \ell + k \|_{L^2(\Omega)}. \quad \square \]

Lemma 4.3 (Quasiorthogonality). Given \( C_1 > 0 \) from Lemma 4.2, quasiorthogonality holds in the sense of

\begin{align*}
(4.3) & \quad (p - p_{\ell+k}, p_{\ell+k} - p_\ell)_{L^2(\Omega)} \leq C_1 \varepsilon_{\ell+k} \operatorname{osc}_\ell; \\
(4.4) & \quad \| p_{\ell+k} - p_\ell \|_{L^2(\Omega)} \leq \varepsilon_\ell^2 - \varepsilon_{\ell+k}^2 + 2C_1 \varepsilon_{\ell+k} \operatorname{osc}_\ell; \\
(4.5) & \quad \varepsilon_\ell^2 - \varepsilon_{\ell+k}^2 \leq \| p_{\ell+k} - p_\ell \|_{L^2(\Omega)} + 2C_1 \varepsilon_{\ell+k} \operatorname{osc}_\ell.
\end{align*}

Proof. Let \( \hat{p}_{\ell+k} \in RT_0(T_{\ell+k}) \) be the intermediate solution of Lemma 4.2 with \( \text{div} \hat{p}_{\ell+k} + \ell = 0 \). An integration by parts and \( \text{div} (\hat{p}_{\ell+k} - p_\ell) = 0 \) show that \( (p - p_{\ell+k}, \hat{p}_{\ell+k} - p_\ell)_{L^2(\Omega)} = 0 \).

Hence an application of Lemma 4.2 yields

\[ \left| (p - p_{\ell+k}, p_{\ell+k} - p_\ell)_{L^2(\Omega)} \right| \leq \varepsilon_{\ell+k} \| p_{\ell+k} - \hat{p}_{\ell+k} \|_{L^2(\Omega)} \leq C_1 \varepsilon_{\ell+k} \operatorname{osc}_\ell. \]

This proves (4.3) and leads to (4.4) via

\[ \| p_{\ell+k} - p_\ell \|_{L^2(\Omega)}^2 = \varepsilon_\ell^2 - \varepsilon_{\ell+k}^2 - 2(p - p_{\ell+k}, p_{\ell+k} - p_\ell)_{L^2(\Omega)} \leq \varepsilon_\ell^2 - \varepsilon_{\ell+k}^2 + 2C_1 \varepsilon_{\ell+k} \operatorname{osc}_\ell. \]

The same arguments yield (4.5); namely,

\[ \varepsilon_\ell^2 - \varepsilon_{\ell+k}^2 = \| p_{\ell+k} - p_\ell \|_{L^2(\Omega)}^2 + 2(p - p_{\ell+k}, p_{\ell+k} - p_\ell)_{L^2(\Omega)} \leq \| p_{\ell+k} - p_\ell \|_{L^2(\Omega)}^2 + 2C_1 \varepsilon_{\ell+k} \operatorname{osc}_\ell. \quad \square \]

The following lemma is essentially contained in [22] and given here with a proof for convenient reading.

Lemma 4.4. Given \( C_1 > 0 \) from Lemma 4.2, there exists a constant \( C_2 > 0 \), which solely depends on \( T_0 \), such that

\begin{align*}
(4.6) & \quad |\mathcal{E}_\ell \setminus \mathcal{E}_{\ell+k}| \leq 3 \left( |T_{\ell+k}| - |T_\ell| \right), \\
(4.7) & \quad \| p_{\ell+k} - p_\ell \|_{L^2(\Omega)}^2 \leq C_2 \eta_\ell^2 (\mathcal{E}_\ell \setminus \mathcal{E}_{\ell+k}) + C_1^2 \varepsilon_\ell^2 \operatorname{osc}_\ell^2.
\end{align*}

Proof of (4.6) of Lemma 4.4. For each \( E \in \mathcal{E}_\ell \setminus \mathcal{E}_{\ell+k} \) there is a refinement of the neighbourhood \( \omega_E := T_+ \cup T_- \) of \( E \) and each neighbouring \( K \in \{ T_+, T_- \} \subset T_\ell \) with \( E \subseteq \partial K \) is, at least, bisected in the refinements from \( T_\ell \) to \( T_{\ell+k} \). Let

\[ \chi_j(K) := |\{ T \in T_j \mid T \subseteq \omega_E \}| \text{ for } K \in T_\ell \text{ and } j = \ell, \ell + 1, \ldots; \]

\[ m_E := \begin{cases} 1/2 & \text{if } E \text{ interior}, \ E \in \mathcal{E}_\ell \setminus \mathcal{E}_{\ell+k}, \\ 1 & \text{if } E \subseteq \partial \Omega, \ E \in \mathcal{E}_\ell \setminus \mathcal{E}_{\ell+k}, \\ 0 & \text{otherwise}. \end{cases} \]

Then, for all \( E \in \mathcal{E}_\ell \setminus \mathcal{E}_{\ell+k} \),

\[ 1 \leq m_E (\chi_{\ell+k}(\omega_E) - \chi_\ell(\omega_E)). \]
Proof of (4.7) of Lemma 4.4. Let \( \hat{p}_{\ell+k} \in RT_0(T_{\ell+k}) \) be the intermediate solution of Lemma 4.2 with \( \text{div} \, b_{\ell+k} + f_\ell = 0 \).

Then, \( \text{div} \, (\hat{p}_{\ell+k} - p_\ell) = 0 \) shows that \( (\hat{p}_{\ell+k} - p_{\ell+k}, \hat{p}_{\ell+k} - p_\ell)_{L^2(\Omega)} = 0 \). This and Lemma 4.2 yield

\[
\| p_{\ell+k} - p_\ell \|_{L^2(\Omega)} = \| \hat{p}_{\ell+k} - p_\ell \|_{L^2(\Omega)}^2 + \| p_{\ell+k} - \hat{p}_{\ell+k} \|_{L^2(\Omega)}^2 \\
\leq \| \hat{p}_{\ell+k} - p_\ell \|_{L^2(\Omega)}^2 + C_1 \text{osc}_f^2.
\]

It remains to bound \( \| \hat{p}_{\ell+k} - p_\ell \|_{L^2(\Omega)} \). The discrete orthogonal Helmholtz decomposition \( 3 \) yields \( b_{\ell+k} \in P_1(T_{\ell+k}) = \{ v \in P_1(T_{\ell+k}) \cap C(\Omega) \mid \int_\Omega v \, dx = 0 \} \) and \( a_{\ell+k} \in P_{NC}^{\text{osc}}(T_{\ell+k}) \) with

\[
\hat{p}_{\ell+k} - p_\ell = \nabla a_{\ell+k} + \text{Curl} \, b_{\ell+k}.
\]

Let \( b_\ell := \mathcal{I}_\ell b_{\ell+k} \) be some (e.g., Scott-Zhang) quasi-interpolation with

\[
\| b_{\ell+k} - \mathcal{I}_\ell b_{\ell+k} \|_{L^2(E)} \leq C b_{\ell+k}^{1/2} |b_{\ell+k}|_{H^1(\omega_E)}
\]

for \( E \in \mathcal{E}_\ell \) and its neighbourhood \( \omega_E \). Notice that

\[
\| b_{\ell+k} - \mathcal{I}_\ell b_{\ell+k} \|_{L^2(E)} = 0 \text{ if } E \in \mathcal{E}_{\ell+k} \cap \mathcal{E}_\ell.
\]

The \( L^2 \) orthogonality \( (\hat{p}_{\ell+k} - p_\ell) \perp_{L^2(\Omega)} \nabla a_{\ell+k} \) verifies

\[
\| \hat{p}_{\ell+k} - p_\ell \|_{L^2(\Omega)} = (\hat{p}_{\ell+k} - p_\ell, \text{Curl} \, b_{\ell+k})_{L^2(\Omega)},
\]

while the discrete conditions \( 1.2 \) with test function \( \text{Curl} \, b_\ell \in RT_0(T_\ell) \subseteq RT_0(T_{\ell+k}) \) lead to

\[
(\hat{p}_{\ell+k}, \text{Curl} \, b_{\ell+k})_{L^2(\Omega)} = 0 = (p_\ell, \text{Curl} \, b_\ell)_{L^2(\Omega)}.
\]

Hence,

\[
\| \hat{p}_{\ell+k} - p_\ell \|_{L^2(\Omega)} = (p_\ell, \text{Curl} (b_\ell - b_{\ell+k}))_{L^2(\Omega)}
\]

\[
= \sum_{E \in \mathcal{E}_\ell} \int_E |p_\ell|_E \cdot \tau_E (b_\ell - b_{\ell+k}) \, ds - \sum_{T \in \mathcal{T}_\ell} \int_T \text{curl}(p_\ell)(b_\ell - b_{\ell+k}) \, ds
\]

\[
= \sum_{E \in \mathcal{E}_{\ell+k} \setminus \mathcal{E}_\ell} \int_E |p_\ell|_E \cdot \tau_E (b_\ell - b_{\ell+k}) \, ds.
\]
In 2D it holds $|b_{t+k}|_{H^1(\Omega)} \leq \|\hat{p}_{t+k} - p_t\|_{L^2(\Omega)}$. Thus,
\[
\|\hat{p}_{t+k} - p_t\|_{L^2(\Omega)}^2 \leq \sum_{E \in \mathcal{E}_{t+k} \setminus \mathcal{E}_t} \|p_t\|_{E} L^2(E) \|b_t - b_{t+k}\|_{L^2(E)} \\
\leq C_{1/2}^3 \eta_t (\mathcal{E}_t \setminus \mathcal{E}_{t+k}) \|\hat{p}_{t+k} - p_t\|_{L^2(\Omega)}.
\]
A division by $\|\hat{p}_{t+k} - p_t\|_{L^2(\Omega)}$ plus $\mathbf{[18]}$ leads to the assertion $\mathbf{[17]}$. \hfill $\Box$

5. Convergence and Optimality of AMFEM

This section is devoted to the convergence analysis of the adaptive mixed finite element method (AMFEM). Again, let $T_{t+k}$ be a triangulation refined from $T_t$ in $k \geq 1$ levels of refinement and $p_{t+k}$, $p_t$ the flux part of their respective MFEM solutions.

5.1. Contraction property.

Lemma 5.1. There exists $C_3 > 0$, depending on $T_0$ only, such that for any $\delta > 0$, $C_\delta := C_3 (1 + 1/\delta)$ and on any level $\ell$ with Case (A) and $0 < \theta_A < 1$ or with Case (B) and $k \geq 1$, it follows that
\[
\eta_{t+1}^2 \leq (1 + \delta)(1 - \theta_A/2)\eta_t^2 + C_\delta \|p_{t+1} - p_t\|_{L^2(\Omega)}^2 \quad \text{in Case (A)},
\]
\[
\eta_{t+1}^2 \leq (1 + \delta)\eta_t^2 + C_\delta \|p_{t+1} - p_t\|_{L^2(\Omega)}^2 \quad \text{in Case (B)}.
\]
Proof. The estimates are proven by applying Young’s inequality and by exploiting $\mathbf{[16]}$. \hfill $\Box$

Lemma 5.2 (Contraction). Given $0 < \theta_A, \rho_B < 1$, there exist positive parameters $\alpha, \beta, \kappa_0$ and $0 < \rho < 1$, such that for all $0 < \kappa < \kappa_0$ and on any level with Case (A) or (B) of algorithm AMFEM the weighted term $\xi_t$ satisfies the contraction
\[
\xi_{t+1}^2 \leq \rho \xi_t^2, \quad \xi_t^2 := \eta_t^2 + \alpha \varepsilon_t^2 + \beta \text{osc}_t^2.
\]
Proof. Let $0 < \theta_A, \rho_B < 1$ and positive constants $C_1, C_\delta, C_{\text{rel}}$ be given from Lemmas $\mathbf{[4]}$ $\mathbf{[5]}$ and Theorem $\mathbf{[2]}$ respectively. Contraction is proven for the following choice of positive parameters $A, B, D, \alpha, \beta, \delta, \gamma, \kappa_0$, and $0 < \rho_A < 1$ and all $0 < \kappa < \kappa_0$:
\[
\delta := \theta_A/(4 - 2\theta_A), \quad \rho_A := (1 + \delta)(1 - \theta_A/2),
\]
\[
\gamma := \min \left\{ \frac{1}{2}, 1 - \frac{\rho_A}{C_\delta C_{\text{rel}}} \right\}, \quad B := \min \left\{ \frac{3}{4}, 1 - \frac{\rho_A}{C_\delta C_{\text{rel}}} \right\},
\]
\[
A := 2(C_1 C_\delta / \gamma + C_\delta C_{\text{rel}} B), \quad \alpha := C_\delta (1 - \gamma),
\]
\[
\kappa_0 := \frac{1 - \rho_A - C_\delta C_{\text{rel}} B}{A}, \quad \beta := 2 \max \left\{ \frac{C_1 C_\delta / \gamma + C_\delta C_{\text{rel}} B + D}{1 - \rho_B}, A - C_1 C_\delta / \gamma - C_\delta C_{\text{rel}} B \right\}.
\]
The application of Young’s inequality and reliability $\mathbf{[21]}$ to the quasiorthogonality $\mathbf{[43]}$ leads to
\[
\|p_{t+1} - p_t\|_{L^2(\Omega)}^2 \leq (1 - B)\varepsilon_t^2 - (1 - \gamma)\varepsilon_{t+1}^2 + (C_1/\gamma + C_{\text{rel}} B) \text{osc}_t^2 + C_{\text{rel}} B \eta_t^2.
\]
On each level with Case (A), substituting (5.4) in (5.1) reveals
\[
\eta^2_{\ell+1} \leq (\rho A + C_\delta C_{rel} B) \eta^2_{\ell} + C_\delta (1 - B) \varepsilon^2_{\ell} - C_\delta(1 - \gamma) \varepsilon^2_{\ell+k} + C_\delta (C_1/\gamma + C_{rel} B) \text{osc}^2_{\ell}.
\]
Together with reliability, \(\text{osc}^2_{\ell+1} \leq \text{osc}^2_{\ell}\), and \(\text{osc}^2_{\ell} \leq \kappa \eta^2_{\ell}\), this verifies
\[
\xi^2_{\ell+1} \leq (\rho A + A \kappa + C_\delta C_{rel} B) \eta^2_{\ell} + C_\delta (1 - B) \varepsilon^2_{\ell} + (\beta + C_1 C_\delta/\gamma - A + C_\delta C_{rel} B) \text{osc}^2_{\ell}.
\]
Thus, with the proposed choice of parameters, on any level with Case (A), contraction is realised, i.e., \(\xi^2_{\ell+1} \leq \rho_1 \xi^2_{\ell}\) holds with \(\rho_1\) defined by
\[
0 < \rho_1 := \max \{ (\rho A + A \kappa + C_\delta C_{rel} B), C_\delta (1 - B)/\alpha, (\beta + C_1 C_\delta/\gamma - A + C_\delta C_{rel} B)/\beta \} < 1.
\]
In Case (B), similarly to Case (A), substituting (5.4) in (5.2) and applying \(\text{osc}^2_{\ell+1} \leq \rho_B \text{osc}^2_{\ell}\), as well as \(\kappa \eta^2_{\ell} < \text{osc}^2_{\ell}\), proves
\[
\xi^2_{\ell+1} \leq (1 + \delta + C_\delta C_{rel} B - D \kappa) \eta^2_{\ell} + C_\delta (1 - B) \varepsilon^2_{\ell} + (C_1 C_\delta/\gamma + C_\delta C_{rel} B + \beta \rho_B + D) \text{osc}^2_{\ell}.
\]
Thus, for the special choice of parameters, contraction of \(\xi^2_{\ell}\) on any level with Case (B), i.e., \(\xi^2_{\ell+1} \leq \rho_2 \xi^2_{\ell}\), holds with \(\rho_2\) defined as
\[
0 < \rho_2 := \max \{ 1 + \delta + C_\delta C_{rel} B - D \kappa, C_\delta (1 - B)/\alpha, + (C_1 C_\delta/\gamma + C_\delta C_{rel} B + \beta \rho_B + D)/\beta \} < 1.
\]
Hence, there exist parameters \(\alpha, \beta, \kappa_0 > 0\) such that for all \(0 < \kappa < \kappa_0\) on any level, irrespective of the relation between \(\eta_{\ell}\) and \(\text{osc}_{\ell}\), the contraction (5.3) of the weighted term \(\xi_{\ell}\) with \(\rho := \max \{ \rho_1, \rho_2 \}\) is ensured. \(\square\)

Remark 5.3. Given \(C_{eff}, C_{rel}, C_1, \alpha, \beta, \kappa\) from Theorem 2.1 and Lemmas 4.2, 5.2 respectively, set
\[
C_A := 1 + \alpha C_{rel} + (\alpha C_{rel} + \beta) \kappa, \quad C_B := C_A/\kappa.
\]
Then, the weighted term \(\xi_{\ell}\) for triangulation \(T_{\ell}\) of the MFEM error \(\varepsilon_{\ell}\), the estimated error \(\eta_{\ell}\) and oscillations \(\text{osc}_{\ell}\) satisfies
\[
\xi^2_{\ell} \leq \begin{cases} C_A \eta^2_{\ell} & \text{if Case (A) applies,} \\ C_B \text{osc}^2_{\ell} & \text{if Case (B) applies,} \end{cases}
\]
(5.5)
(5.6)
\[
\xi^2_{\ell} \leq \varepsilon^2_{\ell} + \text{osc}^2_{\ell}.
\]
Proof. (5.5) is proven by applying reliability and the specific relation of the estimated error and oscillations in both Cases (A) and (B). (5.6) follows directly from the efficiency and reliability (2.1) of \(\eta_{\ell}\). \(\square\)

Lemma 5.4. Given \(C_{eff}\) from Theorem 2.1, \(C_1\) from Lemma 4.2 and \(\alpha, \beta\) as chosen in Lemma 5.2, let \(T\) be a triangulation refined from \(T_{\ell}\) by NVB and set
\[
C_C := \max \{ 4C_{eff}^{-1} + 2, 4C_1C_{eff}^{-1} + 2C_1 + 1 \}.
\]
The weighted terms $\xi_\ell$ and $\xi_T$ for $T$ of the MFEM error $\varepsilon_T$, the estimated error $\eta_T$, and oscillations $\text{osc}_T$ satisfy

\begin{align}
\varepsilon_T^2 &\leq 2\varepsilon_\ell^2 + 4C_1^2 \text{osc}_T^2, \\
\alpha \varepsilon_T^2 &\leq \xi_T^2 \leq C_C \xi_\ell^2. 
\end{align}

**Proof.** \[\text{(5.7)}\] follows from quasiorthogonality and Young’s inequality via

$$
\varepsilon_T^2 \leq \varepsilon_\ell^2 + 2C_1 \varepsilon_T \text{osc}_\ell - \|p_T - p_\ell\|^2_{L^2(\Omega)} 
\leq \varepsilon_\ell^2 + 2C_1^2 \text{osc}_\ell^2 + \varepsilon_T^2/2 - \|p_T - p_\ell\|^2_{L^2(\Omega)}.
$$

To prove \[\text{(5.8)}\], the estimates $1 \leq 2\alpha$ and $1 \leq \beta/C_1$ are verified by means of the definitions of $\alpha$ and $\beta$ in the proof of Lemma \[5.2\].

In addition to the previous estimates, efficiency is applied and yields

$$
\xi_T^2 \leq 2 \left( \frac{C_0}{\varepsilon_\ell^2} + \alpha \varepsilon_\ell^2 \right) + \left( \frac{C_0^2}{\varepsilon_\ell^2} (C_0^2 + \alpha) + \beta \right) \text{osc}_\ell^2 
\leq \left( \frac{2}{\alpha C_0^2 + 2} \right) \alpha \varepsilon_\ell^2 + \left( \frac{4C_0^2}{\beta C_0^2 + 2} + \frac{4C_0^2}{\beta} + 1 \right) \beta \text{osc}_\ell^2 
\leq C_C \xi_\ell^2, \quad \text{with} \quad C_C := \max \left\{ 4C_0^2 + 1, 4C_0 C_0^2 + 2C_1 + 1 \right\}.
$$

Note that $C_C > 0$ is independent of the special choice of parameters in the algorithm AMFEM, and in particular is independent of $\theta_A$. \hfill \Box

### 5.2. Optimal convergence.

This subsection is devoted to the optimal convergence rate \[7, 31\] of the adaptive algorithm AMFEM.

**Definition 5.5** (Approximation Class). Given an initial regular triangulation $T_0$ of $\Omega$ and $s > 0$, let

$$
A_s := \{(p, f) \in H(\text{div}, \Omega) \times L^2(\Omega) \mid \|(p, f)\|_{A_s} < \infty \} \quad \text{with}
$$

$$
\| (p, f) \|_{A_s} := \sup_{N \in \mathcal{P}} \left( \inf_{|T| \leq |T_0| \leq N} \left( \frac{\varepsilon(T) + \text{osc}(f, T)}{\varepsilon(T) + \text{osc}(f, T)} \right)^{1/2} \right),
$$

$$
\varepsilon(T) := \| p - p_T \|^2_{L^2(\Omega)}.
$$

The infimum is with respect to all regular and NVB-generated refinements $T$ of $T_0$, called admissible triangulations, with the number of element domains $|T| \leq N + |T_0|$ and with the exact error $\varepsilon(T)$ of the flux-part $p_T \in RT_0(T)$ of the MFEM solution.

**Remark 5.6.** The approximation class $A_s$ can be characterised \[31\] p. 255, Remark 5.1, p. 263, l. 17 by $(p, f) \in A_s$ if and only if for all $\varepsilon > 0$ there exists an admissible triangulation $T_\varepsilon$ such that the associated MFEM solution $(p_\varepsilon, u_\varepsilon)$ in $RT_0(T_\varepsilon) \times P_0(T_\varepsilon)$ satisfies

$$
\varepsilon^2(T_\varepsilon) + \text{osc}^2(f, T_\varepsilon) \leq \varepsilon^2 
\quad \text{and} \quad |T_\varepsilon| - |T_0| \lesssim \varepsilon^{-1/s} \|(p, f)\|_{A_s}^{1/s}.
$$

**Remark 5.7.** Our main result states optimal convergence of the AMFEM in the following sense: Given $(p, f) \in A_s$ and $\ell \in \mathcal{P}$, the algorithm generates a triangulation $T_\varepsilon$ with discrete solutions $(p_\varepsilon, u_\varepsilon)$ in $RT_0(T_\varepsilon) \times P_0(T_\varepsilon)$ such that

$$
|T_\varepsilon| - |T_0| \approx \xi_\ell^{-1/s} \approx \left( \varepsilon_\ell^2 + \text{osc}_\ell^2 \right)^{-1/(2s)} \approx \left( \varepsilon_\ell + \text{osc}_\ell \right)^{-1/s}.
$$
Proof. The key of the proof is to verify for any level \( \ell \) in either Case (A) or (B) that there exist \( K(\ell) \in \mathbb{N} \) sets of marked edges \( M_0^{(\ell)}, \ldots, M_{K(\ell)}^{(\ell)} \) such that

\[
T_{\ell+1} = \text{Refine} \left( T_\ell, (M_k^{(\ell)})_{0 \leq k \leq K(\ell)} \right),
\]

(5.9)

In fact, once (5.9) is verified, (3.5) with (5.9) and the contraction property (5.3) from Lemma 5.2 reveal optimal global convergence owing to the subsequent arguments

\[
|T_\ell| - |T_0| \leq \sum_{j=0}^{\ell-1} \sum_{k=0}^{\ell-1} |T_{j+1} - |T_j|| \leq C_0 \sum_{j=0}^{\ell-1} \sum_{k=0}^{\ell-1} |M_j(k)| \leq \sum_{j=0}^{\ell-1} \sum_{k=0}^{\ell-1} \beta j^{-1} = \xi_\ell^{-1/\alpha}. 
\]

This and the contraction property (5.3) from Lemma 5.2 reveal optimal global convergence owing to the subsequent arguments

\[
|T_\ell| - |T_0| \leq \xi_\ell^{-1/\alpha} \sum_{k=1}^{\ell} \beta^{-k/(2\alpha)} = \frac{1 - \beta^{-(\ell+1)/(2\alpha)}}{1 - \beta^{-1/(2\alpha)}} \xi_\ell^{-1/\alpha} \approx \xi_\ell^{-1/\alpha}. 
\]

Hence it remains to prove (5.9) in both Cases (A) and (B) for a given triangulation \( T_\ell \) and discrete solution \((p_\ell, u_\ell)\) of (1.2). Due to the choice of \( \kappa \), in Case (A), \( \tau > 0 \) can be chosen to satisfy

\[
0 < \tau^2 \leq \frac{\alpha(C_{\text{eff}} - 2C_1^2 \kappa - C_2 \theta_A)}{2C_AC_C},
\]

with positive constants \( C_A, C_C \) from Remark 5.3 and Lemma 5.4. Setting \( \epsilon := \tau \xi_\ell \), (5.0) and Remark 5.6 lead to the existence of some admissible triangulation \( T_\ell \) refined from \( T_0 \) such that

\[
\xi^2(T_\ell) := \eta^2(T_\ell) + \alpha \varepsilon^2(T_\ell) + \beta \text{osc}^2(f, T_\ell) \leq \epsilon^2, \quad |T_\ell| - |T_0| \leq \epsilon^{-1/\alpha}.
\]

with the exact flux-error \( \varepsilon(T_\ell) \), the estimated error \( \eta(T_\ell) \) and oscillations \( \text{osc}(f, T_\ell) \) on \( T_\ell \). For the overlay \( T := T_\ell \oplus T_\ell \) and its associated quantities \( \xi^2(T) = \eta^2(T) + \alpha \varepsilon^2(T) + \beta \text{osc}^2(f, T) \), with estimated error \( \eta(T) \), flux-part of the error \( \varepsilon(T) = \|p - p_T\|_{L^2(\Omega)} \), and oscillations \( \text{osc}(f, T) \), Lemma 5.1 verifies

\[
|T| - |T_\ell| = |T_\ell \oplus T_\ell| - |T_\ell| \leq |T_\ell| - |T_0| \leq \epsilon^{-1/\alpha} \approx \xi_\ell^{-1/\alpha}.
\]

In the next step we prove that \( \mathcal{E}_\ell \setminus \mathcal{E}_T \) fulfills the bulk criterion (2.2); namely, (5.10)

\[
\theta_A \eta^2_\ell \leq \eta^2_\ell (\mathcal{E}_\ell \setminus \mathcal{E}_T).
\]

Efficiency, \( \text{osc}^2_\ell \leq \kappa \eta^2_\ell \), and the following combination of (5.20)–(5.8),

\[
\alpha \varepsilon^2(T) \leq \xi^2(T) \leq C_C \xi^2(T_\ell) \leq C_C \tau^2 \xi^2_\ell \leq C_A C_C \tau^2 \eta^2_\ell,
\]
lead to
\[ C_2 \theta_A \eta^2 \leq \left( C_{\text{eff}} - 2 \kappa C_1^2 \right) \eta^2 - \left( 2 \tau^2 C_A \kappa \right) \eta^2 \leq \varepsilon^2 \leq 2 C^2 \varepsilon^2 - 2 \varepsilon^2(T), \]
while (4.5) and Lemma 4.3 show that
\[ \varepsilon^2 - 2 \varepsilon^2(T) \leq \| p_T - p_T \|^2_{L^2(\Omega)} + C^2 \varepsilon^2 \]
\[ \leq 2 C^2 \varepsilon^2 + C_2 \eta^2(\mathcal{E}_T \setminus \mathcal{E}_T). \]

The combination of the previous estimations results in (5.10). Since \( M_\ell \) was chosen with minimal cardinality and \( \theta_A \eta^2 \leq \eta^2(M_\ell) \), Lemma 4.4 yields
\[ |M_\ell| \leq |E_\ell \setminus E_T| \lesssim |T| - |T_\ell| \lesssim \xi^{-1/s}. \]

For each level \( \ell \) with Case (A), set \( K(\ell) := 0 \) and \( M^{(0)}_\ell := M_\ell \).

In Case (B), \( \kappa \eta^2 < \varepsilon^2 \), let \( T \) be some refinement of \( T_0 \) with \( \varepsilon^2(f, T) \leq \text{Tol}^2 \) and \( |T| - |T_0| \lesssim \text{Tol}^{-1/s}, \text{Tol}^2 := \rho_B \varepsilon^2 \).

Algorithm 3.2 computes a finite sequence \( M^{(0)}_\ell, \ldots, M^{K(\ell)}_\ell \) of marked reference edges such that
\[ T \oplus T_\ell = \text{Refine} \left( T_\ell, (M^{(k)}_\ell)_{k=0, \ldots, K(\ell)} \right). \]

Finally, Theorem 3.3 and 5.5 for Case (B) with \( \rho_B \varepsilon^2 = \text{Tol}^2 \) verify
\[ \sum_{k=0}^{K(\ell)} |M^{(k)}_\ell| \leq |T| - |T_0| \lesssim \text{Tol}^{-1/s} \lesssim \varepsilon^{-1/s} \lesssim \xi^{-1/s}, \]
which proves (5.10) in Case (B). This concludes the proof. \( \square \)

5.3. Numerical experiments. This subsection is devoted to a numerical comparison of the convergence behaviour of algorithm AMFEM of this paper and the Becker-Mao algorithm of [6] called BM-AMFEM in the example of [6]; namely,
\[ \Delta u + f = 0 \text{ in } \Omega \text{ and } u = u_D \text{ on } \partial \Omega \]
with data \( f \) and \( u_D \) and the exact solution
\[ u(x) := \exp(-100|x - (1/2, 1/2)|) \text{ for } x \in \Omega := (0, 1)^2. \]

The initial uniform mesh \( T_0 \) consists of 8 congruent, right isosceles triangles aligned to the principal diagonal. Table 5.1 displays the numerical outcome for the estimator \( \eta_\ell \), the oscillations \( \text{osc}_\ell \), and the flux part of the MFEM-error \( \varepsilon_\ell = \| p - p_T \|_{L^2(\Omega)} \) for several consecutive levels \( \ell \) with the number of degrees of freedom (ndof) of AMFEM and BM-AMFEM with \( \kappa = \theta_A = 0.1 \).

The reduction of large oscillations in the right-hand side \( f \) is optimal in AMFEM, and the steering parameter \( 0 < \rho_B < 1 \) is free.

The bulk-parameter \( 0 < \theta_B < 1 \) in Case (B) of BM-AMFEM influences the factor of reducing oscillations, i.e., \( \text{osc}_{\ell+1}^2 \leq (1 - \theta_B/2) \text{osc}_\ell^2 \), and leads to a reduction by \( (1 - \theta_B/2)^m \geq 2^{-m} \) in \( m \) steps. Unfortunately the value of \( \theta_B \) in the numerical example is not mentioned in [6]. Our choice \( \theta_B = 0.75 \) in BM-AMFEM seems to correspond to \( \rho_B = 1 - \theta_B/2 = 0.625 \) in AMFEM for an unbiased comparison displayed in the first two columns of Table 5.1.
Table 5.1. Numerical results for AMFEM and BM-Amfem with $\kappa = 0.1 = \theta_A$

| \(\ell\) | \(n_{\text{dof}}\) | \(\eta_{s}\) | \(\text{osc}_{\ell}\) | \(\varepsilon_{\ell}\) | Case | \(n_{\text{dof}}\) | \(\eta_{s}\) | \(\text{osc}_{\ell}\) | \(\varepsilon_{\ell}\) | Case | \(n_{\text{dof}}\) | \(\eta_{s}\) | \(\text{osc}_{\ell}\) | \(\varepsilon_{\ell}\) | Case |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 | 8.623e-8 | 1.308e+1 | 1.254 | B | 1 | 8 | 8.623e-8 | 1.308e+1 | 1.254 | B | 1 | 8 | 8.623e-8 | 1.308e+1 | 1.254 | B |
| 23 | 7.078e-6 | 1.097e+1 | 1.254 | B | 2 | 32 | 1.082e-5 | 7.444 | 1.253 | B | 2 | 826 | 8.810e-1 | 1.029e-1 | 2.643e-1 | A |
| 39 | 5.680e-4 | 6.704 | 1.253 | B | 3 | 70 | 1.324e-2 | 5.714 | 1.253 | B | 3 | 1567 | 4.935e-1 | 8.398e-2 | 1.588e-1 | A |
| 85 | 2.304e-2 | 2.837 | 1.254 | B | 4 | 101 | 1.814e-2 | 4.511 | 1.254 | B | 4 | 3622 | 2.983e-1 | 6.954e-2 | 1.047e-1 | A |
| 211 | 1.490 | 4.712 | 2.189 | B | 6 | 352 | 1.017 | 2.153 | 1.525 | B | 6 | 8986 | 1.896e-1 | 4.600e-2 | 6.864e-2 | A |
| 274 | 1.352 | 3.367 | 2.253 | B | 7 | 484 | 9.704e-1 | 1.193 | 9.171e-1 | B | 7 | 16189 | 1.404e-1 | 3.240e-2 | 5.018e-2 | A |
| 397 | 7.865e-1 | 1.376 | 1.059 | B | 9 | 655 | 9.126e-1 | 3.230e-1 | 3.618e-1 | B | 9 | 53332 | 7.735e-2 | 1.443e-2 | 2.626e-2 | A |
| 478 | 7.996e-1 | 7.190e-1 | 6.079e-1 | B | 10 | 745 | 8.787e-1 | 1.721e-1 | 2.863e-1 | A | 10 | 58301 | 7.882e-2 | 2.136e-2 | 3.697e-2 | A |
| 532 | 9.034e-1 | 3.726e-1 | 3.970e-1 | B | 11 | 1492 | 4.852e-1 | 1.598e-1 | 1.938e-1 | B | 11 | 1396 | 5.057e-1 | 2.720e-2 | 3.845e-2 | A |
| 586 | 9.713e-1 | 2.036e-1 | 3.209e-1 | A | 12 | 1540 | 5.057e-1 | 8.456e-2 | 1.619e-1 | A | 12 | 560 | 5.057e-1 | 2.720e-2 | 3.845e-2 | A |
| 892 | 6.470e-1 | 1.920e-1 | 2.455e-1 | A | 13 | 2902 | 3.363e-1 | 8.146e-2 | 1.185e-1 | A | 13 | 105 | 2.470e-2 | 2.495 | 1.254 | B |
| 1348 | 4.762e-1 | 1.854e-1 | 2.055e-1 | A | 14 | 7300 | 2.101e-1 | 4.635e-2 | 7.372e-2 | A | 14 | 346 | 5.464e-1 | 5.035 | 1.481 | B |
| 9727 | 1.786e-1 | 5.778e-2 | 7.351e-2 | B | 17 | 77161 | 6.454e-2 | 1.438e-2 | 2.320e-2 | A | 17 | 5332 | 5.514e-1 | 1.704 | 1.328 | B |
| 9769 | 1.824e-1 | 2.976e-2 | 5.956e-2 | B | 18 | | | | | | | | | | |
| 19702 | 1.283e-1 | 2.059e-2 | 4.185e-2 | A | 19 | | | | | | | | | | |
| 50230 | 8.024e-2 | 1.445e-2 | 2.692e-2 | A | 20 | | | | | | | | | | |

BM-Amfem with $\theta_B = 0.75$

AMfem with $\rho_B = 0.625$

AMfem with $\rho_B = 0.01$

BM-Amfem with $\theta_B = 1.0$
The parameter $0 < \theta_A \ll 1$ has to be sufficiently small and the value $\theta_A = 0.8$ in [6] is not at all supported by the analysis.

The two algorithms require about 10 levels to reduce the oscillations via Case (B). Our strategy aims to overcome this pre-asymptotic overhead by small values of $0 < \rho_B < 1$, and the last column in Table 5.1 displays the numerical outcome for $\rho_B = 0.01$ in AMFEM. The number of levels to overcome the pre-asymptotic regime in our proposed AMFEM is reduced dramatically to 2.

Recall that a larger $\theta_B < 1$ in BM-AMFEM may lead to fewer numbers of levels but more global refinements, so the optimal choice is less clear as well as the concept of optimality. For comparison, a choice near 1 leads to failure similar to the values displayed in the lower half of the last column in Table 5.1 for the extreme $\theta_B = 1.0$.

Each of the two adaptive algorithms is proven to be asymptotically optimal, and the numerical experiments clearly support this theoretical result (after a large or small pre-asymptotic regime).

REFERENCES


**HUMBOLDT-UNIVERSITÄT ZU BERLIN, UNTER DEN LINDEN 6, 10099 BERLIN, GERMANY AND DEPARTMENT OF COMPUTATIONAL SCIENCE AND ENGINEERING, YONSEI UNIVERSITY, 120-749 SEOUL, KOREA**

E-mail address: cc@mathematik.hu-berlin.de

Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

E-mail address: rabus@mathematik.hu-berlin.de