CONVERGENCE ANALYSIS OF ADAPTIVE FINITE ELEMENT METHODS

LONG CHEN

In this note, we introduce the convergence analysis of adaptive finite element methods (AFEMs) for the Poisson equation and refer to Nochetto, Siebert, and Veeser [10] for a comprehensive overview of the theory behind adaptive finite element methods.

1. Residual Type A Posteriori Error Estimate

For the sake of clarity, we consider the Poisson equation with homogeneous Dirichlet boundary conditions:

$$-\Delta u = f \quad \text{in} \quad \Omega, \qquad u = 0 \quad \text{on} \quad \partial\Omega. \tag{1}$$

Let $\mathcal{T}$ be a shape-regular triangulation of Ω, and $\mathcal{V}_{\mathcal{T}} \subset H_0^1(\Omega)$ be the linear finite element space based on $\mathcal{T}$. The linear finite element method for (1) is to find $u_{\mathcal{T}} \in \mathcal{V}_{\mathcal{T}}$ such that

$$(\nabla u_{\mathcal{T}}, \nabla v_{\mathcal{T}}) = (f, v_{\mathcal{T}}), \quad \text{for all } v_{\mathcal{T}} \in \mathcal{V}_{\mathcal{T}}. \tag{2}$$

Here, we assume $f \in L^2(\Omega)$, and $(\cdot,\cdot)$ denotes the L^2-inner product.

When the solution $u \in H^2(\Omega)$, we have the *a priori* error analysis:

$$|u - u_{\mathcal{T}}|_{1,\Omega} \lesssim h_{\mathcal{T}} \|u\|_{2,\Omega}.$$

However, such optimal convergence order may not hold when u is not in $H^2(\Omega)$. In this section, we derive a residual type *a posteriori* error estimate of the error $|u - u_{\mathcal{T}}|_{1,\Omega}$.

The H^1-norm of the error and the residual are connected through the differential operator and for the Poisson equation, which is the identity:

$$|u - u_{\mathcal{T}}|_1 = \| - \Delta(u - u_{\mathcal{T}})\|_{-1} = \sup_{w \in H_0^1(\Omega)} \frac{a(u - u_{\mathcal{T}}, w)}{|w|_1}, \tag{3}$$

where $a(u,v) = (\nabla u, \nabla v)$ is the bilinear form associated with the Poisson equation. The term $-\Delta u = f$ holds in L^2, but element-wise $\Delta_T u_{\mathcal{T}} = 0$ for a linear function. We need to understand it in the dual sense and provide a computable upper bound of the sup in (3).

1.1. A Local and Stable Quasi-Interpolation.

To define a function in the linear finite element space $\mathcal{V}_{\mathcal{T}}$, we only need to assign values at interior vertices. The nodal interpolation $u_I \in V_{\mathcal{T}}$ is defined as $u_I(x_i) = u(x_i)$ for $u \in C(\bar{\Omega})$, which is not well-defined for $u \in H^1(\Omega)$. For a vertex $x_i \in \mathcal{N}(\mathcal{T})$, recall that Ω_i consists of all simplices sharing this vertex, and for an element $\Omega_\tau = \cup_{x_i \in \tau} \Omega_i$. Instead of using nodal values of the function, we can use its average over Ω_i.

For an interior vertex x_i, we define $A_i u = |\Omega_i|^{-1} \int_{\Omega_i} u(x)\,\mathrm{d}x$. To incorporate the boundary condition, when $x_i \in \partial\Omega$, we define $A_i u = 0$. Define the averaged quasi-interpolation $\Pi_{\mathcal{T}} : L^1(\Omega) \mapsto \mathcal{V}_{\mathcal{T}}$ by

$$\Pi_{\mathcal{T}} u = \sum_{x_i \in \mathcal{N}(\mathcal{T})} A_i(u) \varphi_i,$$

where φ_i is the hat function (the basis of linear finite element space) at vertex x_i.

Lemma 1.1. *For $u \in H^1(\Omega_\tau)$, we have the error estimate*

$$\|u - \Pi_{\mathcal{T}} u\|_{0,\tau} \lesssim h_\tau |u|_{1,\Omega_\tau}.$$

Proof. For interior vertices, we use the average-type Poincaré inequality:

$$\|u - A_i u\|_{0,\Omega_i} \leq C h_\tau |u|_{1,\Omega_i}, \tag{4}$$

and for boundary vertices, we use Poincaré-Friedrichs since $u|_{\partial\Omega_i \cap \partial\Omega} = 0$ and the $\mathbb{R}^{d-1}$ Lebesgue measure of the set $\partial\Omega_i \cap \partial\Omega$ is non-zero. The constant C in the inequality (4) is independent of Ω_i since the mesh is shape regular. Then we use the partition of unity $\sum_{i=1}^{d+1} \varphi_i = 1$ restricted to one element τ to write

$$\begin{aligned}\int_\tau |u - \Pi_{\mathcal{T}} u|^2 &= \int_\tau \Big|\sum_{i=1}^{d+1} (u - A_i u)\varphi_i\Big|^2 \,\mathrm{d}x \\ &\lesssim \sum_{i=1}^{d+1} \int_{\Omega_i} |u - A_i u|^2 \,\mathrm{d}x \\ &\lesssim h_\tau \sum_{i=1}^{d+1} \int_{\Omega_i} |\nabla u|^2 \,\mathrm{d}x \lesssim h_\tau^2 \int_{\Omega_\tau} |\nabla u|^2 \,\mathrm{d}x.\end{aligned}$$

□

Exercise 1.2. *Prove that $\Pi_{\mathcal{T}}$ is stable in the L^2-norm:*

$$\|\Pi_{\mathcal{T}} u\|_{0,\tau} \lesssim \|u\|_{0,\Omega_\tau}.$$

Next, we prove that $\Pi_{\mathcal{T}}$ is stable in the H^1 norm. Let us introduce another average operator Q_τ: the L^2 projection to the piecewise constant function spaces on $\mathcal{T}$:

$$(Q_\tau u)|_\tau = |\tau|^{-1} \int_\tau u(x) \,\mathrm{d}x,$$

for each $\tau \in \mathcal{T}$.

Lemma 1.3. *For $u \in H^1(\Omega_\tau)$, we have stability*

$$|\Pi_{\mathcal{T}} u|_{1,\tau} \lesssim |u|_{1,\Omega_\tau}.$$

Proof. Using the Poincaré inequality, it is easy to see

$$\|u - Q_\tau u\|_{0,\tau} \lesssim h_\tau |u|_{1,\tau}.$$

We use the inverse inequality and the first-order approximation property of Q_τ and $\Pi_{\mathcal{T}}$ to obtain

$$\begin{aligned}|\Pi_{\mathcal{T}} u|_{1,\tau} &= |\Pi_{\mathcal{T}} u - A_\tau u|_{1,\tau} \\ &\leq h_\tau^{-1} \|\Pi_{\mathcal{T}} u - Q_\tau u\|_{0,\tau} \\ &\leq h_\tau^{-1} \left(\|u - \Pi_{\mathcal{T}} u\|_{0,\tau} + \|u - Q_\tau u\|_{0,\tau}\right) \\ &\lesssim |u|_{1,\Omega_\tau}.\end{aligned}$$

□

Summing over each element and using the finite overlapping property due to the shape regularity of the mesh, we obtain the stability and approximation property.

Lemma 1.4. *For $u \in H_0^1(\Omega)$, the quasi-interpolant $\Pi_{\mathcal{T}} u$ satisfies the following properties:*

(1) L^2 and H^1-stable:

$$\|\Pi_{\mathcal{T}} u\| \lesssim \|u\|, \quad |\Pi_{\mathcal{T}} u|_1 \lesssim |u|_1.$$

(2)

$$\sum_{\tau\in\mathcal{T}} \left(\|h^{-1}(u-\Pi_{\mathcal{T}} u)\|_{0,\tau}^2 + \|\nabla(u-\Pi_{\mathcal{T}} u)\|_{0,\tau}^2\right)^{1/2} \lesssim |u|_{1,\Omega}.$$

1.2. **Upper Bound.** The equidistribution principle suggests equidistributing the quantity $|u|_{2,1,\tau}$; see Introduction to Adaptive Finite Element Methods. However, it is not computable since u is unknown. One may want to approximate it by $|u_{\mathcal{T}}|_{2,1,\tau}$. For linear finite element function $u_{\mathcal{T}}$, we have $|u_{\mathcal{T}}|_{2,1,\tau} = 0$, thus providing no information about $|u|_{2,1,\tau}$.

The derivative of the piecewise constant vector function $\nabla u_{\mathcal{T}}$ will be delta distributions on edges with magnitudes equal to the jump of $\nabla u_{\mathcal{T}}$ across the edge. In the continuous level, $\Delta u \in L^2(\Omega)$ implies $\nabla u \in H(\mathrm{div};\Omega)$, meaning $\nabla u \cdot n_e$ is continuous at an edge e where n_e is a unit normal vector of e. For the finite element approximation $u_{\mathcal{T}} \in \mathcal{V}_{\mathcal{T}}$, the normal derivative $\nabla u_{\mathcal{T}} \cdot n_e$ is not continuous, although the tangential derivative $\nabla u_{\mathcal{T}} \cdot t_e$ is. The discontinuity of the normal derivative across edges can be used to measure the error $\nabla u - \nabla u_{\mathcal{T}}$.

In the following, we provide a rigorous justification and a posteriori error estimate for the Poisson equation with homogeneous Dirichlet boundary condition below and refer to [14] for general elliptic equations and mixed boundary conditions.

Before delving into technical details, we highlight the orthogonality arising from the Galerkin projection.

Lemma 1.5. *Let u be the solution of (1) and $u_{\mathcal{T}} \in \mathcal{V}_{\mathcal{T}}$ be the solution of (2). Then we have the orthogonality*

$$(\nabla u - \nabla u_{\mathcal{T}}, \nabla v_{\mathcal{T}}) = 0 \quad \forall v_{\mathcal{T}} \in \mathcal{V}_{\mathcal{T}}. \tag{5}$$

Let $\mathcal{E}_{\mathcal{T}}$ denote the set of all interior edges. For each interior edge $e \in \mathcal{E}_{\mathcal{T}}$, we fix a unit normal vector n_e. Let τ_1 and τ_2 be two triangles sharing the edge e. The jump of flux across e is defined as

$$[\nabla u_{\mathcal{T}} \cdot n_e] = \nabla u_{\mathcal{T}} \cdot n_e|_{\tau_1} - \nabla u_{\mathcal{T}} \cdot n_e|_{\tau_2}.$$

We define h as a piecewise constant function on $\mathcal{T}$: for each element $\tau \in \mathcal{T}$,

$$h|_\tau = h_\tau := |\tau|^{1/2}. \tag{6}$$

We also define a piecewise constant function on $\mathcal{E}_{\mathcal{T}}$ as

$$h|_e = h_e := (h_{\tau_1} + h_{\tau_2})/2, \tag{7}$$

where $e = \tau_1 \cap \tau_2$ is the common edge of two triangles τ_1 and τ_2.

We will use the trace theorem $\|v\|_{0,\partial\tau} \lesssim \|v\|_{1,\tau}$ and apply a scaling argument to obtain

$$\|v\|_{0,e} \lesssim h_\tau^{-1/2}\|v\|_{0,\tau} + h_\tau^{1/2}|v|_{1,\tau}. \tag{8}$$

The correct scaling of h can be obtained by choosing $v = 1$ and $v = x$ in (8).

Theorem 1.6. *For a given triangulation $\mathcal{T}$, let $u_{\mathcal{T}}$ be the linear finite element approximation of the solution u of the Poisson equation. Then there exists a constant $C_1 > 0$ depending only on the shape regularity of $\mathcal{T}$ such that*

$$|u - u_{\mathcal{T}}|_1 \le C_1\Big(\sum_{\tau\in\mathcal{T}} \|hf\|_{0,\tau}^2 + \sum_{e\in\mathcal{E}_{\mathcal{T}}} \|h^{1/2}[\nabla u_{\mathcal{T}} \cdot n_e]\|_{0,e}^2\Big)^{1/2}. \tag{9}$$

Proof. For any $w \in H_0^1(\Omega)$ and any $w_{\mathcal{T}} \in \mathcal{V}_{\mathcal{T}}$, we have

$$
\begin{aligned}
& a(u-u_{\mathcal{T}}, w) \\
= \; & a(u-u_{\mathcal{T}}, w-w_{\mathcal{T}}) \\
= \; & \sum_{\tau\in\mathcal{T}} \int_\tau \nabla(u-u_{\mathcal{T}})\cdot\nabla(w-w_{\mathcal{T}})\,\mathrm{dx} \\
= \; & \sum_{\tau\in\mathcal{T}} \int_\tau -\Delta(u-u_{\mathcal{T}})(w-w_{\mathcal{T}})\,\mathrm{dx} + \sum_{\tau\in\mathcal{T}} \int_{\partial\tau} \nabla(u-u_{\mathcal{T}})\cdot n(w-w_{\mathcal{T}})\,\mathrm{dS} \\
= \; & \sum_{\tau\in\mathcal{T}} \int_\tau f(w-w_{\mathcal{T}})\,\mathrm{dx} + \sum_{e\in\mathcal{E}_h} \int_e [\nabla u_{\mathcal{T}}\cdot n_e](w-w_{\mathcal{T}})\,\mathrm{dS} \\
\le \; & \sum_{\tau\in\mathcal{T}} \|hf\|_{0,\tau}\|h^{-1}(w-w_{\mathcal{T}})\|_{0,\tau} + \sum_{e\in\mathcal{E}_h} \|h^{1/2}[\nabla u_{\mathcal{T}}\cdot n_e]\|_{0,e}\|h^{-1/2}(w-w_{\mathcal{T}})\|_{0,e} \\
\lesssim \; & \left(\sum_{\tau\in\mathcal{T}} \|hf\|_{0,\tau}^2 + \sum_{e\in\mathcal{E}_{\mathcal{T}}} \|h^{-1/2}[\nabla u_{\mathcal{T}}\cdot n_e]\|_{0,e}^2\right)^{1/2} \\
& \left(\sum_{\tau\in\mathcal{T}} \|h^{-1}(w-w_{\mathcal{T}})\|_{0,\tau}^2 + \|\nabla(w-w_{\mathcal{T}})\|_{0,\tau}^2\right)^{1/2}.
\end{aligned}
$$

In the last step, we have used the scaled trace theorem (8).

Now chose $w_{\mathcal{T}} = \Pi_{\mathcal{T}} w$ using the quasi-interpolation operator introduced in Lemma 1.4, we have

$$
\left(\sum_{\tau\in\mathcal{T}} \|h^{-1}(w-w_{\mathcal{T}})\|_{0,\tau}^2 + \|\nabla(w-w_{\mathcal{T}})\|_{0,\tau}^2\right)^{1/2} \lesssim |w|_{1,\Omega}. \tag{10}
$$

Then we end with

$$
|u-u_{\mathcal{T}}|_1 = \sup_{w\in H_0^1(\Omega)} \frac{a(u-u_{\mathcal{T}}, w)}{|w|_1} \lesssim \Big(\sum_{\tau\in\mathcal{T}} \|hf\|_{0,\tau}^2 + \sum_{e\in\mathcal{E}_{\mathcal{T}}} \|h^{1/2}[\nabla u_{\mathcal{T}}\cdot n_e]\|_{0,e}^2\Big)^{1/2}.
$$

□

To guide the local refinement, we need an element-wise error indicator. For any $\tau \in \mathcal{T}$ and any $v_{\mathcal{T}} \in \mathcal{V}_{\mathcal{T}}$, we define

$$
\eta(v_{\mathcal{T}}, \tau) = \left(\|hf\|_{0,\tau}^2 + \sum_{e\in\partial\tau} \|h^{1/2}[\nabla v_{\mathcal{T}}\cdot n_e]\|_{0,e}^2\right)^{1/2}. \tag{11}
$$

For a subset $\mathcal{M}_{\mathcal{T}} \subseteq \mathcal{T}$, we define

$$
\eta(v_{\mathcal{T}}, \mathcal{M}_{\mathcal{T}}) = \Big[\sum_{\tau} \eta^2(v_{\mathcal{T}}, \tau)\Big]^{1/2}.
$$

With these notation, the upper bound (9) can be simply written as

$$
|u-u_{\mathcal{T}}|_{1,\Omega} \le C_1 \eta(u_{\mathcal{T}}, \mathcal{T}). \tag{12}
$$

Remark 1.7. The local version of the upper bound (12)

$$
|u-u_{\mathcal{T}}|_{1,\tau} \le C_1 \eta(u_{\mathcal{T}}, \Omega_\tau)
$$

does not hold in general as the orthogonality (5) only holds globally.

1.3. **Lower Bound.** We shall derive a lower bound of the error estimator η through the following exercises. The technique is developed by Verfürth [14] and is known as bubble functions. Namely using bubble functions to embed constants into $H_0^1(\Omega)$.

Let u be the solution of the Poisson equation $-\Delta u = f$ with homogeneous Dirichlet boundary condition, and let $u_{\mathcal{T}}$ be the linear finite element approximation of u based on a shape-regular and conforming triangulation $\mathcal{T}$.

Exercise 1.8. (1) For a triangle τ, we denote $V_\tau = \{f_\tau \in L_2(\tau) \,|\, f_\tau = \text{constant}\}$ equipped with the L^2 inner product. Let $\lambda_i, i = 1, 2, 3$ be the barycentric coordinates of $x \in \tau$, and let $b_\tau = \lambda_1\lambda_2\lambda_3$ be the bubble function on τ. We define $B_\tau f_\tau = f_\tau b_\tau$.

Prove that $B_\tau : V_\tau \mapsto V = H_0^1(\Omega)$ is bounded in L^2 and H^1 norm:

$$\|B_\tau f_\tau\|_{0,\tau} = C\|f_\tau\|_{0,\tau}, \text{ and } \|\nabla(B_\tau f_\tau)\|_{0,\tau} \lesssim h_\tau^{-1}\|f_\tau\|_{0,\tau}.$$

(2) Use (1) to prove that

$$\|hf_\tau\|_{0,\tau} \lesssim |u - u_{\mathcal{T}}|_{1,\tau} + \|h(f - f_\tau)\|_{0,\tau}.$$

(3) For an interior edge e, we define $V_e = \{g_e \in L^2(e) \,|\, g_e = \text{constant}\}$. Suppose e has endpoints x_i and x_j, we define $b_e = \lambda_i\lambda_j$ and $B_e : V_e \mapsto V$ by $B_e g_e = g_e b_e$.

Let Ω_e denote the domaine formed by two triangles sharing e. Prove that

(a) $\|g_e\|_{0,e} = C\|B_e g_e\|_{0,e}$,

(b) $\|B_e g_e\|_{0,\Omega_e} \lesssim h_e^{1/2}\|g_e\|_{0,e}$ and,

(c) $\|\nabla(B_e g_e)\|_{0,\Omega_e} \lesssim h_e^{-1/2}\|g_e\|_{0,e}$.

(4) Use (3) to prove that

$$\|h^{1/2}[\nabla u_{\mathcal{T}} \cdot n_e]\|_{0,e} \lesssim \|hf\|_{0,\Omega_e} + |u - u_{\mathcal{T}}|_{1,\Omega_e}.$$

(5) Use (1) and (4) to prove the lower bound of the error estimator. There exists a constant C_2 depending only on the shape regularity of the triangulation such that for any piecewise constant approximation f_τ of $f \in L^2$,

$$C_2\eta^2(u_{\mathcal{T}}, \mathcal{T}) \le |u - u_{\mathcal{T}}|_{1,\Omega}^2 + \sum_{\tau\in\mathcal{T}_h} \|h(f - f_\tau)\|_\tau^2.$$

2. Convergence

Standard adaptive finite element methods (AFEM) based on local mesh refinement can be written as loops of the form

$$\textbf{SOLVE} \to \textbf{ESTIMATE} \to \textbf{MARK} \to \textbf{REFINE}. \tag{13}$$

Starting from an initial triangulation $\mathcal{T}_0$, to obtain $\mathcal{T}_{k+1}$ from $\mathcal{T}_k$, we first solve the equation to obtain u_k based on $\mathcal{T}_k$. The error is then estimated using u_k and $\mathcal{T}_k$, and this error estimation is used to mark a set of triangles in $\mathcal{T}_k$. Marked triangles, and possibly more neighboring triangles, are then refined in such a way that the triangulation remains shape-regular and conforming; see Section 3 for details.

2.1. **Algorithm.** In the **SOLVE** step, we assume that the solutions of the finite-dimensional problems can be efficiently solved to any desired accuracy. For multigrid methods on graded bisection grids, we refer to [5].

The *a posteriori* error estimators play a crucial role in the **ESTIMATE** step. We have provided one in the previous section and will discuss more in the next section.

The *a posteriori* error estimator is divided into local error indicators, which are then used to make local modifications by refining the elements with large error indicators and possibly coarsening the elements with small error indicators. The way we mark these triangles influences the efficiency of the adaptive algorithm. The traditional maximum marking strategy, proposed in the pioneering work of Babuška and Vogelius [1], is to mark triangles τ^* such that

$$\eta(u_{\mathcal{T}}, \tau^*) \geq \theta \max_{\tau \in \mathcal{T}} \eta(u_{\mathcal{T}}, \tau), \quad \text{for some } \theta \in (0, 1).$$

Such a marking strategy is designed to evenly distribute the error. Based on our relaxation of the equidistribution principle, we may leave some exceptional elements and focus on the overall amount of error. This leads to the bulk criterion first proposed by Dörfler [6] in order to prove the convergence of the local refinement strategy. With such a strategy, one defines the marking set $\mathcal{M}_{\mathcal{T}} \subset \mathcal{T}$ such that

$$\eta^2(u_{\mathcal{T}}, \mathcal{M}_{\mathcal{T}}) \geq \theta\, \eta^2(u_{\mathcal{T}}, \mathcal{T}), \quad \text{for some } \theta \in (0, 1). \tag{14}$$

We shall use the Dörfler marking strategy in the convergence proof.

After choosing a set of marked elements, we need to carefully design the rule for dividing the marked triangles such that the mesh obtained by this division rule remains conforming and shape-regular. We may need to refine more triangles to recover the conformity of the triangulation and thus denote the set of refined triangles by $\overline{\mathcal{M}}_k$. Additionally, we aim to control the number of elements added to ensure the optimality of the refinement. To this end, we shall use the newest vertex bisection detailed in Section 3.

Let us summarize AFEM in the following subroutine:

```
[$u_J, \mathcal{T}_J$] = AFEM ($\mathcal{T}_1, f, tol, \theta$)
% AFEM compute an approximation $u_J$ by adaptive finite element methods
% Input: $\mathcal{T}_1$ an initial triangulation; $f$ data; $tol$ <<1 tolerance; $\theta \in (0,1)$
% Output: $u_J$ linear finite element approximation; $\mathcal{T}_J$ the finest mesh
$\eta = 1, k = 0$;
while $\eta \geq tol$
    $k = k + 1$;
    SOLVE Poisson equation on $\mathcal{T}_k$ to get the solution $u_k$;
    ESTIMATE the error by $\eta = \eta(u_k, \mathcal{T}_k)$;
    MARK a set $\mathcal{M}_k \subset \mathcal{T}_k$ with minimum number such that
                    $\eta^2(u_k, \mathcal{M}_k) \geq \theta\, \eta^2(u_k, \mathcal{T}_k)$;
    REFINE $\tau \in \overline{\mathcal{M}}_k$ to get a conforming triangulation $\mathcal{T}_{k+1}$;
end
$u_J = u_k; \mathcal{T}_J = \mathcal{T}_k$;
```

2.2. **Contraction of the error estimator.** By the orthogonality (5), one can easily conclude that the error is non-increasing, i.e.,

$$|u - u_{k+1}|_1 \leq |u - u_k|_1.$$

Equality could hold, i.e., $u_{k+1} = u_k$, if the refinement did not introduce interior nodes for triangles and edges; see Examples 3.6 and 3.7 in [9]. A closer look reveals that when the solution does not change, the error estimator η will be reduced due to the change in mesh size and the Dörfler marking strategy.

Lemma 2.1. *Given $\theta \in (0,1)$, let $\mathcal{T}_{k+1}$ be a conforming and shape-regular triangulation refined from a conforming and shape-regular triangulation $\mathcal{T}_k$ using the Dörfler marking strategy (14). Let u_k be the solution of (2) in $\mathcal{V}_k$. Then*

$$\eta^2(u_k,\mathcal{T}_{k+1}) \le \rho\,\eta^2(u_k,\mathcal{T}_k) \tag{15}$$

for some $\rho \in (0,1)$ depending only on the shape regularity of $\mathcal{T}_k$ and the parameter θ used in the Dörfler marking strategy.

Proof. We study in detail the change in the error estimator due to the bisection of a triangle. Suppose τ is bisected into τ_1 and τ_2. We first prove an element-wise contraction of the error indicator: there exists a number $\bar\rho \in (0,1)$ depending only on the shape regularity of $\mathcal{T}_k$ such that

$$\eta^2(u_k,\tau_1) + \eta^2(u_k,\tau_2) \le \bar\rho\,\eta^2(u_k,\tau). \tag{16}$$

To distinguish between the different mesh size functions, we use h_{k+1} and h_k to denote the mesh size function defined on $\mathcal{T}_k$ and $\mathcal{T}_{k+1}$, respectively. Thanks to our definition, $h^2_{k+1,\tau_i} = |\tau_i| = 1/2|\tau| = 1/2\,h^2_{k,\tau}$. The part involving the element residual is reduced by one half:

$$\|h_{k+1}f\|^2_{\tau_1} + \|h_{k+1}f\|^2_{\tau_2} = \frac{1}{2}\|h_k f\|^2_\tau.$$

Regarding the jump of gradient on the edges, an important observation is that $[\nabla u_k \cdot n_e] = 0$ for the newly created edge inside τ. For other edges on the boundary of τ, h_e is reduced by a factor due to the definition of h_e, while the jump $[\nabla u_k \cdot n_e]$ remains unchanged as a constant on the coarse mesh. So $\sum_{e\in\partial\tau} \|h^{1/2}[\nabla v_{\mathcal{T}} \cdot n_e]\|^2_{0,e}$ is also reduced by a factor strictly less than one.

Since not all elements are refined, Dörfler marking ensures that a portion of the error estimator is reduced, which is sufficient. Recall that $\mathcal{M}_k \subseteq \mathcal{T}_k$ is the marked set. We may need to refine more triangles to recover the conformity of the triangulation, and thus denote the set of refined triangles by $\overline{\mathcal{M}}_k$. Since $\mathcal{M}_k \subseteq \overline{\mathcal{M}}_k$, we have

$$\eta^2(u_k,\overline{\mathcal{M}}_k) \ge \eta^2(u_k,\mathcal{M}_k) \ge \theta\eta^2(u_k,\mathcal{T}_k).$$

We use the notation $\overline{\mathcal{M}}_{k+1} \subseteq \mathcal{T}_{k+1}$ to denote the set of triangles obtained by refinement of those in $\overline{\mathcal{M}}_k$. Then $\mathcal{T}_k\backslash\overline{\mathcal{M}}_k = \mathcal{T}_{k+1}\backslash\overline{\mathcal{M}}_{k+1}$ are the untouched triangles. We then have

$$\begin{aligned}
\eta^2(u_k,\mathcal{T}_{k+1}) &= \eta^2(u_k,\mathcal{T}_{k+1}\backslash\overline{\mathcal{M}}_{k+1}) + \eta^2(u_k,\overline{\mathcal{M}}_{k+1})\\
&\le \eta^2(u_k,\mathcal{T}_k\backslash\overline{\mathcal{M}}_k) + \bar\rho\,\eta^2(u_k,\overline{\mathcal{M}}_k)\\
&= \eta^2(u_k,\mathcal{T}_k) - (1-\bar\rho)\,\eta^2(u_k,\overline{\mathcal{M}}_k)\\
&\le \eta^2(u_k,\mathcal{T}_k) - \theta(1-\bar\rho)\,\eta^2(u_k,\mathcal{T}_k)\\
&= \big[1-\theta(1-\bar\rho)\big]\eta^2(u_k,\mathcal{T}_k).
\end{aligned}$$

We obtain (15) with $\rho = 1-\theta(1-\bar\rho) \in (0,1)$. □

2.3. **Contraction of the sum of error and error estimator.** We shall prove the convergence of `AFEM` by showing the contraction of the total error between two levels. There exists a positive constant α and a constant $\delta \in (0,1)$ such that for all $k \ge 0$,

$$|u-u_{k+1}|_1^2 + \alpha\,\eta^2(u_{k+1},\mathcal{T}_{k+1}) \le \delta\left[\,|u-u_k|_1^2 + \alpha\,\eta^2(u_k,\mathcal{T}_k)\right]. \tag{17}$$

Recall that we have

$$|u - u_{k+1}|_1^2 \le |u - u_k|_1^2, \quad \eta^2(u_k, \mathcal{T}_{k+1}) \le \rho\, \eta^2(u_k, \mathcal{T}_k).$$

To prove (17), we shall explore more relation between the error and the error estimator in consecutive levels $\mathcal{T}_k$ and $\mathcal{T}_{k+1}$.

Lemma 2.2. *Given a $\theta \in (0,1)$, let $\mathcal{T}_{k+1}$ be a conforming and shape regular triangulation which is refined from a conforming and shape regular triangulation $\mathcal{T}_k$ using Dörfler marking strategy (14). Let u_{k+1} and u_k be solutions of (2) in $\mathcal{V}_{k+1}$ and $\mathcal{V}_k$, respectively. Then we have*

(1) orthogonality:

$$|u - u_{k+1}|_1^2 = |u - u_k|_1^2 - |u_{k+1} - u_k|_1^2;$$

(2) upper bound:

$$|u - u_k|_1^2 \le C_1 \eta^2(u_k, \mathcal{T}_k)$$

for some constant C_1 depending only on the shape regularity of $\mathcal{T}$;

(3) continuity of the error estimator: for any $\epsilon > 0$, there exists a constant C_ϵ such that

$$\eta^2(u_{k+1}, \mathcal{T}_{k+1}) \le (1+\epsilon)\eta^2(u_k, \mathcal{T}_{k+1}) + C_\epsilon |u_{k+1} - u_k|_1^2;$$

(4) contraction of the error estimator:

$$\eta^2(u_{k+1}, \mathcal{T}_{k+1}) \le \rho(1+\epsilon)\, \eta^2(u_k, \mathcal{T}_k) + C_\epsilon |u_{k+1} - u_k|_1^2$$

for $\rho \in (0,1)$ in Lemma 2.1.

Proof. (1) is straightforward since u_{k+1} is the H^1 projection and $u_{k+1} - u_k \in \mathcal{V}_{k+1}$ due to the nestedness of $\mathcal{T}_k$ and $\mathcal{T}_{k+1}$. (2) has been proven in the previous section.

Now let's prove (3). The part containing the element-wise residual $\|hf\|$ remains unchanged since we do not alter the triangulation. For each $e \in \mathcal{E}_{\mathcal{T}}$, let $\tau \in \mathcal{T}$ such that $e \in \partial\tau$. From the triangle inequality and the fact that $\nabla(u_{k+1} - u_k)$ is piecewise constant, we have:

$$\begin{aligned}\|h^{1/2}[\nabla u_{k+1} \cdot n_e]\|_{0,e} &\le \|h^{1/2}[\nabla u_k \cdot n_e]\|_{0,e} + \|h^{1/2}[\nabla(u_{k+1} - u_k) \cdot n_e]\|_{0,e} \\ &\le \|h^{1/2}[\nabla u_k \cdot n_e]\|_{0,e} + C|u_{k+1} - u_k|_{1,\tau}.\end{aligned}$$

Squaring both sides, applying Young's inequality $2ab \le \epsilon a^2 + \epsilon^{-1} b^2$, and summing over all edges yields the desired inequality. (4) is a combination of (3) and Lemma 2.1. □

We are now ready to prove the contraction result. We exploit the negative term arising from the orthogonality of the error to offset the positive term resulting from the reduction of the error estimator.

Theorem 2.3. *Given a $\theta \in (0,1)$, let $\mathcal{T}_{k+1}$ be a conforming and shape regular triangulation which is refined from a conforming and shape regular triangulation $\mathcal{T}_k$ using Dörfler marking strategy (14). Let u_{k+1} and u_k be solutions of (2) in $\mathcal{V}_{k+1}$ and $\mathcal{V}_k$, respectively. Then there exist constants $\delta \in (0,1)$ and α depending only on θ and the shape regularity of $\mathcal{T}_k$ such that*

$$|u - u_{k+1}|_1^2 + \alpha\, \eta^2(u_{k+1}, \mathcal{T}_{k+1}) \le \delta \left[\, |u - u_k|_1^2 + \alpha\, \eta^2(u_k, \mathcal{T}_k) \right]. \tag{18}$$

Proof. Let $\rho \in (0,1)$ be the constant in Lemma 2.1. Since $\rho \in (0,1)$, we can choose $\epsilon \in (0,1)$ small enough such that $\rho(1+\epsilon) < 1$. Let $\alpha = C_{\epsilon}^{-1}$. Adding the two inequalities in Lemma 2.2 (1) and (4) with weight α will imply

$$|u - u_{k+1}|_1^2 + \alpha\, \eta^2(u_{k+1}, \mathcal{T}_{k+1}) \leq |u - u_k|_1^2 + \alpha(1+\epsilon)\rho\, \eta^2(u_k, \mathcal{T}_k).$$

Let δ be a number in $(0,1)$ whose value will be clear in a moment. We then have

$$\begin{aligned}
&|u - u_{k+1}|_1^2 + \alpha\, \eta^2(u_{k+1}, \mathcal{T}_{k+1}) \\
\leq\, & \delta |u - u_k|_1^2 + (1-\delta)|u - u_k|_1^2 + \alpha\rho(1+\epsilon)\, \eta^2(u_k, \mathcal{T}_k) \\
\leq\, & \delta |u - u_k|_1^2 + (1-\delta) C_1 \eta^2(u_k, \mathcal{T}_k) + \alpha\rho(1+\epsilon)\, \eta^2(u_k, \mathcal{T}_k) \\
\leq\, & \delta \left[|u - u_k|_1^2 + \frac{(1-\delta)C_1 + \alpha\rho(1+\epsilon)}{\delta} \eta^2(u_k, \mathcal{T}_k) \right].
\end{aligned}$$

This suggests us to choose δ such that

$$\alpha = \frac{(1-\delta)C_1 + \alpha\rho(1+\epsilon)}{\delta}.$$

Namely

$$\delta = \frac{C_1 + \alpha\rho(1+\epsilon)}{C_1 + \alpha}. \tag{19}$$

Recall that we choose ϵ such that $\rho(1+\epsilon) < 1$, so $\delta \in (0,1)$. The desired result (18) then follows. □

As a consequence of the contraction of the total error between two levels, we can prove that `AFEM` will terminate in a finite number of steps for a given tolerance *tol* and yield a convergent approximation u_J based on an adaptive grid $\mathcal{T}_J$. For a deeper analysis of complexity, readers are encouraged to refer to [13, 4].

Theorem 2.4. *Let u_k and $\mathcal{T}_k$ be the solution and triangulation obtained in the k-th loop in the algorithm `AFEM`, then there exist constants $\delta \in (0,1)$ and α depending only on θ and the shape regularity of $\mathcal{T}_0$ such that*

$$|u - u_k|_1^2 + \alpha\, \eta^2(u_k, \mathcal{T}_k) \leq C_0 \delta^k, \tag{20}$$

and thus the algorithm `AFEM` will terminate in finite steps.

2.4. Alternative convergence proof.

We follow [7] to present an alternative convergence proof of the error estimator.

Theorem 2.5. *Let u_k and $\mathcal{T}_k$ be the solution and triangulation obtained in the k-th loop in the algorithm `AFEM`. Then there exist constants $0 < \varrho < 1$ and $C > 0$ such that: for all positive integers ℓ, m*

$$\eta^2(u_{\ell+m}, \mathcal{T}_{\ell+m}) \leq C \varrho^m \eta^2(u_\ell, \mathcal{T}_\ell). \tag{21}$$

Proof. We recall the contraction of the error estimator

$$\eta^2(u_{i+1}, \mathcal{T}_{i+1}) \leq \rho \eta^2(u_i, \mathcal{T}_i) + C_\theta |u_{i+1} - u_i|_1^2. \tag{22}$$

Therefore, for any $N \geq l+1$, it holds

$$\begin{aligned}\sum_{i=\ell+1}^{N} \eta^2(u_i, \mathcal{T}_i) &\leq \sum_{i=\ell+1}^{N} \left[\rho\eta^2(u_{i-1}, \mathcal{T}_{i-1}) + C_\theta |u_i - u_{i-1}|_1^2\right] \\ &\leq \rho \sum_{i=\ell}^{N-1} \eta^2(u_i, \mathcal{T}_i) + C_\theta |u - u_\ell|_1^2 \\ &\leq \rho \sum_{i=\ell}^{N-1} \eta^2(u_i, \mathcal{T}_i) + C_\theta C_1^2 \eta^2(u_\ell, \mathcal{T}_\ell).\end{aligned}$$

Here, in the second inequality, we have used the orthogonality to get

$$\sum_{i=\ell+1}^{N} |u_i - u_{i-1}|_1^2 = |u_N - u_\ell|_1^2 = |u - u_\ell|_1^2 - |u - u_N|_1^2 \leq |u - u_\ell|_1^2.$$

Then, rearranging the terms and with the arbitrary choice of N, we obtain

$$\sum_{i=\ell+1}^{\infty} \eta^2(u_i, \mathcal{T}_i) \leq \tilde{C}\eta^2(u_\ell, \mathcal{T}_\ell) \qquad \text{for all positive integer } l,$$

where $\tilde{C} = (\rho + C_\theta C_1^2)/(1-\rho)$. Intuitively we have a positive sequence $\{a_i\}$ with property $\sum_{i=\ell+1}^{\infty} a_i \leq \tilde{C} a_\ell$, then a_i is geometric decay.

To prove that, we first show the contraction

$$(1 + \tilde{C}^{-1}) \sum_{i=\ell+1}^{\infty} \eta^2(u_i, \mathcal{T}_i) \leq \sum_{i=\ell+1}^{\infty} \eta^2(u_i, \mathcal{T}_i) + \eta^2(u_\ell, \mathcal{T}_\ell) = \sum_{i=\ell}^{\infty} \eta^2(u_i, \mathcal{T}_i).$$

Repeat m times, we have

$$\begin{aligned}\eta^2(u_{\ell+m}, \mathcal{T}_{\ell+m}) &\leq \sum_{i=\ell+m}^{\infty} \eta^2(u_i, \mathcal{T}_i) \leq (1 + \tilde{C}^{-1})^{-m} \sum_{i=\ell}^{\infty} \eta^2(u_i, \mathcal{T}_i) \\ &\leq (1 + \tilde{C})(1 + \tilde{C}^{-1})^{-m} \eta^2(u_\ell, \mathcal{T}_\ell).\end{aligned}$$

Let $C_5^2 = 1 + \tilde{C}$ and $\varrho = (1 + \tilde{C}^{-1})^{-1}$, then the desired result follows. □

3. Newest vertex bisection

In this section we shall give a brief introduction of the newest vertex bisection. We refer to [8, 14] for detailed description of the newest vertex bisection refinement procedure and especially [3] for the control of the number of elements added by the completion process.

We first recall two important properties of triangulations. A triangulation $\mathcal{T}_h$ (also indicated by mesh or grid) of $\Omega \subset \mathbb{R}^2$ is a decomposition of Ω into a set of triangles. It is called *conforming* if the intersection of any two triangles τ and τ' in $\mathcal{T}_h$ either consists of a common vertex x_i, edge E or empty. An edge of a triangle is called *non-conforming* if there is a vertex in the interior of that edge and that interior vertex is called *hanging node*. See Fig. 3 (b) for an example of non-conforming triangles and hanging nodes. We would like to keep the conformity of the triangulations.

A triangulation $\mathcal{T}_h$ is *shape regular* if

$$\max_{\tau \in \mathcal{T}_h} \frac{\operatorname{diam}(\tau)^2}{|\tau|} \leq \sigma \tag{23}$$

where $\operatorname{diam}(\tau)$ is the diameter of τ and $|\tau|$ is the area of τ. A sequence of triangulation $\{\mathcal{T}_k, k = 0, 1, \cdots\}$ is called *uniform shape regular* if σ in (23) is independent with k.

The shape regularity of triangulations assures that angles of the triangulation remains bounded away from 0 and π which is important to control the interpolation error in H^1 norm and the condition number of the stiffness matrix. We also want to keep this property of the triangulations.

After we marked a set of triangles to be refined, we need to carefully design the rule for dividing the marked triangles such that the refined mesh is still conforming and shape regular. Such refinement rules include red and green refinement [2], longest edge bisection [11] and newest vertex bisection [12]. We shall restrict ourself to the newest vertex bisection method since it will produce nested finite element spaces and relatively easier to generalize to high dimensions.

Given an initial shape regular triangulation $\mathcal{T}_0$ of Ω, we assign to each $\tau \in \mathcal{T}_0$ exactly one vertex called *the newest vertex*. The opposite edge of the newest vertex is called *refinement edge*. One such initial labeling is to use the longest edge of each triangle (with a tie breaking scheme for edges of equal length). The rule of the newest vertex bisection includes:

(1) a triangle is divided to two new children triangles by connecting the newest vertex to the midpoint of the refinement edge;
(2) the new vertex created at a midpoint of a refinement edge is assigned to be the newest vertex of the children.

It is easy to verify that all the descendants of an original triangle fall into four similarity classes (see Figure 1) and hence the angles are bounded away from 0 and π and all triangulations refined from $\mathcal{T}_0$ using newest vertex bisection forms a shape regular class of triangulations.

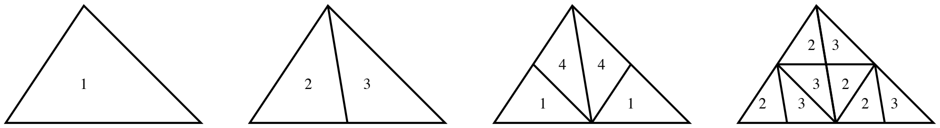


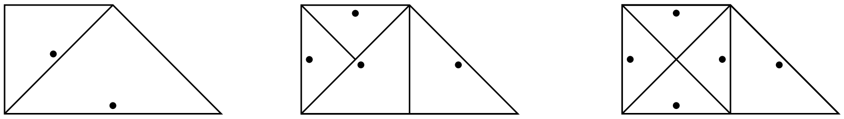
FIGURE 1. Four similarity classes of triangles generated by the newest vertex bisection

The triangulation obtained by the newest vertex might have hanging nodes. We have to make additional subdivisions to eliminate the hanging nodes, i.e., complete the new partition. The completion should also follow the bisection rules to keep the shape regularity; see Figures below for an illustration of the completion procedure.

Let M denotes the set of triangles to be refined. A standard iterative algorithm of the completion is the following.

```
function T = completion(T,M)
while M is not empty
    Update T by bisecting each triangle in M;
    Let now M be the set of non-conforming triangles.
end
```

We need to show the `while` loop will terminate. For two dimensional triangulation, this is easy. Let us denote the uniform bisection of $\mathcal{T}$ as $D(\mathcal{T})$, i.e., every triangle is bisected into



(a) An initial triangulation (b) Refine two triangles producing a non-conforming edge (c) Refine one triangle to obtain a conforming triangulation

FIGURE 2. An illustration of the completion procedure. The dot indicates the refinement edge of each triangle.

two. Note that $D(\mathcal{T})$ may not be conforming; see Fig 3 (b). But $D^2(\mathcal{T})$, which corresponds to bisecting every triangle twice, is always conforming since middle points of all edges are added from $\mathcal{T}$ to $D^2(\mathcal{T})$. We consider the completion procedure as a procedure of splitting edges. The edges split during the completion procedure is a subset of the edge set of $\mathcal{T}$ which is finite and thus the completion will terminate.

If we ask more than the termination of the completion process and want to control the number of elements refined due to the completion, we have to carefully assign the newest vertex for the initial partition $\mathcal{T}_0$. Binev, Dahmen and DeVore [3] show that if $\mathcal{M}$ is the collection of all triangles marked in going from a conforming divisible triangulation $\mathcal{T}_0$ to $\mathcal{T}_k$ then

$$\#\mathcal{T}_k \leq \#\mathcal{T}_0 + C\#\mathcal{M}, \tag{24}$$

where $\#A$ denotes the cardinality of the set A. That is the number of addition triangles refined in the completion procedure is bounded by the number of marked triangles in the l^1 sense. The inequality (24) cannot be true in the l^∞ sense. Refine one marked triangle could trigger a sequence of triangles with length equals to its generation in the completion procedure; see the following figure. The inequality (24) is crucial for the optimality of adaptive finite element methods; see [3, 4].

We conclude this section by a remark that the bisection or the regular refinement in three and higher dimensions is much more involved. The theoretical proof of the shape regularity, the termination of completion, and the control of number of elements added in the completion requires more careful combinatory study; see [10].

REFERENCES

[1] I. Babuška and M. Vogelius. Feeback and adaptive finite element solution of one-dimensional boundary value problems. *Numer. Math.*, 44:75–102, 1984. 6

[2] R. E. Bank, A. H. Sherman, and A. Weiser. Refinement algorithms and data structures for regular local mesh refinement. In *Scientific Computing*, pages 3–17. IMACS/North-Holland Publishing Company, Amsterdam, 1983. 11

[3] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004. 10, 12

[4] J. M. Cascón, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008. 9, 12

[5] L. Chen, R. H. Nochetto, and J. Xu. Optimal multilevel methods for graded bisection grids. *Numer. Math.*, 120(1):1–34, July 21, 2012. 6

[6] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33:1106–1124, 1996. 6

[7] M. Feischl, T. Führer, and D. Praetorius, Adaptive FEM with optimal convergence rates for a certain class of nonsymmetric and possibly nonlinear problems, *SIAM J. Numer. Anal.*, 52 (2014), pp 601-625. 9

[8] W. F. Mitchell. A comparison of adaptive refinement techniques for elliptic problems. *ACM Trans. Math. Softw. (TOMS) archive*, 15(4):326 – 347, 1989. 10

[9] P. Morin, R. H. Nochetto, and K. G. Siebert. Convergence of adaptive finite element methods. *SIAM Rev.*, 44(4):631–658, 2002. 6

[10] R. H. Nochetto, K. G. Siebert, and A. Veeser. Theory of adaptive finite element methods: an introduction. In R. A. DeVore and A. Kunoth, editors, *Multiscale, Nonlinear and Adaptive Approximation*. Springer, 2009. 1, 12

[11] M. C. Rivara. Mesh refinement processes based on the generalized bisection of simplices. *SIAM J. Numer. Anal.*, 21:604–613, 1984. 11

[12] E. G. Sewell. *Automatic Generation of Triangulations for Piecewise Polynomial Approximation*. PhD thesis, Purdue Univ., West Lafayette, Ind., 1972. 11

[13] R. Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, 7(2):245–269, 2007. 9

[14] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement Tecniques*. B. G. Teubner, 1996. 3, 5, 10