Journal of Scientific Computing (2021) 86:17
https://doi.org/10.1007/s10915-020-01387-7

Convergence of Adaptive Weak Galerkin Finite Element Methods for Second Order Elliptic Problems

Yingying Xie[1] · Liuqiang Zhong[1]

Received: 21 July 2019 / Revised: 25 August 2020 / Accepted: 12 December 2020 /
Published online: 5 January 2021

Abstract
We consider a standard Adaptive weak Galerkin (AWG) finite element method for second order elliptic problems. We prove that the sum of the energy error and the scaled error estimator of AWG method, between two consecutive adaptive loops, is a contraction. At last, we present some numerical experiments to support the theoretical results.

Keywords Weak Galerkin · Finite element methods · Convergence · Second-order elliptic problems · Adaptive

Mathematics Subject Classification 65N12 · 65N30 · 35J20

1 Introduction

In this paper, we consider the convergence of Adaptive weak Galerkin finite element methods (AWG) for the following model second order elliptic problems:

$$-\nabla\cdot(A\nabla u)=f \quad \text{in } \Omega, \tag{1}$$
$$u=0 \quad \text{on } \partial\Omega, \tag{2}$$

where Ω is a bounded polygonal or polyhedral domain in $\mathbb{R}^d (d=2,3)$ and is partitioned into non-overlapping subdomains Ω_i, $1\le i\le m$. Here, we need to assume that an initial partition $\mathcal{T}_0$ of Ω, which is consistent with the partition $\bar{\Omega}=\prod_{i=1}^m \Omega_i$ in the sense that each $\mathcal{T}_0\cap\Omega_i$, $1\le i\le m$, inherits a partition of Ω_i. For all $\tau\in\mathcal{T}_0$, we consider the case that the

This research was supported in part by supported by the National Natural Science Foundation of China (Nos. 11671159, 12071160), the Guangdong Basic and Applied Basic Research Foundation (No. 2019A1515010724), the Characteristic Innovation Projects of Guangdong colleges and universities, China (No. 2018KTSCX044) and the General Project topic of Science and Technology in Guangzhou, China (No. 201904010117).

✉ Liuqiang Zhong
zhong@m.scnu.edu.cn

Yingying Xie
xieyy@m.scnu.edu.cn

1 School of Mathematics Sciences, South China Normal University, Guangzhou 510631, China

coefficient A is a piece-wise constant. We assume that the coefficient A satisfies the following property: there exist constants $\alpha > 0$ and $\beta > 0$ such that $\alpha \leq A \leq \beta$.

Weak Galerkin (WG) makes use of discontinuous finite element functions for partial differential equations in which differential operators are approximated by weak forms as distributions. WG methods were first used to solve second order elliptic problem for simplicial grids in [27] and later for regular polytopal meshes in [21]. Then, WG methods in mixed form have been applied to solve the second order elliptic problem for arbitrary shapes of polygons (or polyhedra) in $2D$ (or $3D$) in [28]. WG methods were subsequently applied to other problems, such as second order elliptic interface problems [17], the Helmholtz equation [7,19,23], the biharmonic equation [18,22,26], Darcy equation [12] and so on. WG methods are closely related to the mixed finite element methods and hybridized discontinuous Galerkin(DG) methods. However, when the coefficients are general variable functions, the WG methods are different from these methods.

Computation with adaptive grid refinement has proved to be a useful and efficient tool in scientific computing over the last several decades. We consider the following standard adaptive procedure:

$$\textbf{SOLVE} \rightarrow \textbf{ESTIMATE} \rightarrow \textbf{MARK} \rightarrow \textbf{REFINE}. \tag{3}$$

The precise definition of the algorithm can be found in Sect. 3. For elliptic and Maxwell problems, the theory of convergence and computational complexity in the form of (3) have been great developments in the past few decades, such as [1,3,8,15,33,34] etc. We also refer to [24] for an introduction to the theory of adaptive finite element methods.

For adaptive WG methods, there are only few research results for a posterior error estimates. For second order elliptic problems, a residual type a posteriori error estimator is first presented and analyzed in [6]; a posteriori error estimator is considered for a modified WG method of second order elliptic problems in [31]; a residual type error estimator is proposed which provides global upper and lower bounds of the WG method for second order elliptic problems in a discrete H^1-norm in [29]; recently, a simple posteriori error estimator which can be applied to general meshes such as hybrid, polytopal and those with hanging nodes is introduced for the WG method for second order elliptic problems in [11]; a posteriori error estimate of weak Galerkin (WG) finite element methods for the second order elliptic interface problems is presented in [16]; A residual-based a posteriori error estimator is discussed for the Stokes problem in [32]. However, to our best knowledge, there exists no work in the literature which studies the convergence of adaptive WG methods.

Our work is motivated by the convergence analysis of adaptive mixed finite element methods(AMFEM) in [5,9]. In both approaches, the authors study AMFEM for second order elliptic problems with constant coefficient. In this paper, we will present the convergence of the AWG method for second order elliptic problems whose coefficient is piece-wise constant. Because the weak gradient is defined in polynomial space and the finite element spaces are different from the classical finite element spaces, the proof of the quasi-orthogonality in [5,9] cannot be used directly. The data oscillation and the error indicator are estimated separately in [5,9]. However, in WG methods, the data oscillation is one part of the corresponding error indicator and we have to estimate the data oscillation and the corresponding error indicator together. We also notice that the corresponding error estimates of WG methods are more complicated than ones for mixed element methods.

In this paper, we shall follow the state-of-the-art convergence theory [9] to prove the convergence of adaptive WG methods without extra marking for the data oscillation. We stress that the extension of the convergence theory to adaptive WG methods is not straightforward, since the data oscillation and the error indicator in [9] are estimated, separately, but in WG methods, the data oscillation is one part of the corresponding error indicator and we have to estimate the data oscillation and the corresponding error indicator together. We also notice that the convergence technique used for hybridized DG or mixed methods cannot be applied directly to the WG methods, since the corresponding error estimates of WG methods are more complicated than ones for mixed element methods. Especially, we need to establish the corresponding quasi-orthogonality.

We summarize our main result in the following theorem.

Theorem 1 *Given a parameter $\theta \in (0, 1)$ and initial mesh $\mathcal{T}_0$. Let u be the solution of (1)–(2), $\{\mathcal{T}_k, u_k, \eta(u_k, \mathcal{T}_k)\}_{k\geq 0}$ be a sequence of meshes, finite element solutions and error estimates produced by the AWG method. Then there exist constants $\rho \in (0, 1)$, $\sigma_1 > 0$, $\sigma_2 > 0$ and ϵ depending only on the shape regularity of $\mathcal{T}_0$, the polynomial order l, coefficient A, parameters θ and μ_0, such that if*

$$0 < \epsilon < \min\left(\frac{\sigma_1(1-\xi)}{C_1}, 1\right),$$

then

$$\begin{aligned}
&(1-\epsilon)\|A^{1/2}(\nabla u - \nabla_w u_{k+1})\|^2_{\mathcal{T}_{k+1}} + \sigma_1\eta^2(u_{k+1}, \mathcal{T}_{k+1}) + \sigma_2 osc^2(f, \mathcal{T}_{k+1})\\
&\leq \rho\Big((1-\epsilon)\|A^{1/2}(\nabla u - \nabla_w u_k)\|^2_{\mathcal{T}_k} + \sigma_1\eta^2(u_k, \mathcal{T}_k) + \sigma_2 osc^2(f, \mathcal{T}_k)\Big),
\end{aligned}$$

where the constants C_1 and ξ are given by Lemmas 7 and 12 , respectively.

As a consequence, the AWG method will converges in finite steps for a give tolerance.

Here is some notation used throughout the paper. The following shorthand notation will be used to avoid the repeated constants, following [33], $x \lesssim y$ means $x \leq Cy$, where C are generic positive constants independent of the variables that appear in the inequalities and especially the mesh parameters. The notation C_i, with subscript, denotes specific and important constants.

The rest of the article is organized as follows. In Sect. 2, we describe the definitions of weak gradient and discrete weak gradient, the weak Galerkin finite element spaces and the corresponding bilinear form $a(\cdot, \cdot)$. In Sect. 3, we present the adaptive algorithm and discuss each procedure of (3) in detail. We prove the convergence of the proposed adaptive algorithm in Sect. 4 and report some numerical results in support of theoretical ones in Sect. 5.

2 Prelimimaries and Notations

In this section, we recall the definitions of weak gradient and discrete weak gradient, the weak Galerkin finite element spaces and the corresponding bilinear form $a(\cdot, \cdot)$.

First, we present some notations. For any domain $D \subset \mathbb{R}^d$, $d = 2, 3$, we use standard definitions for the Sobolev spaces $H^s(D)$ and their associated norms $\|\cdot\|_{s,D}$ for $s \geq 0$. Note that the space $L^2(D)$ is $H^0(D)$, we denote its norm by $\|\cdot\|_D$. When $D = \Omega$, we shall simplify the notation as $\|\cdot\|$. More specially, we define $\mathbf{H}(\mathrm{div}, D) = \{\boldsymbol{q} : \boldsymbol{q} \in (L^2(D))^d, \nabla\cdot\boldsymbol{q} \in L^2(D)\}$, $d = 2, 3$.

2.1 Weak Gradient and Discrete Weak Gradient

Let K be any polygonal domain with boundary ∂K. Following [27], a weak function on the region K refers to a function $v = \{v_0, v_b\}$ such that $v_0 \in L^2(K)$ and $v_b \in H^{\frac{1}{2}}(\partial K)$. The first component v_0 can be understood as the value of v in K, and the second component v_b represents v on the boundary of K. Note that v_b may not necessarily be related to the trace of v_0 on ∂K even if the trace is well-defined. Denote by $W(K)$ the space of weak functions on K

$$W(K) := \{v = \{v_0, v_b\} : v_0 \in L^2(K), v_b \in H^{\frac{1}{2}}(\partial K)\}. \tag{4}$$

According to [27], we define the weak gradient as follows.

Definition 1 (*Weak Gradient*) The weak gradient of $v = \{v_0, v_b\} \in W(K)$ is defined as a linear functional $\nabla_w v$ in the dual space of $\mathbf{H}(\mathrm{div}, K)$ satisfying the following equation

$$(\nabla_w v, \boldsymbol{q})_K := -(v_0, \nabla \cdot \boldsymbol{q})_K + \langle v_b, \boldsymbol{q} \cdot \boldsymbol{n}\rangle_{\partial K} \qquad \forall \boldsymbol{q} \in \mathbf{H}(\mathrm{div}, K), \tag{5}$$

where $\boldsymbol{n}$ is the unit outward normal direction to ∂K, $(v_0, \nabla \cdot \boldsymbol{q})_K = \int_K v_0(\nabla \cdot \boldsymbol{q})\mathrm{d}\boldsymbol{x}$ is the action of v_0 on $\nabla \cdot \boldsymbol{q}$, and $\langle v_b, \boldsymbol{q} \cdot \boldsymbol{n}\rangle_{\partial K} = \int_{\partial K} v_b(\boldsymbol{q} \cdot \boldsymbol{n})\mathrm{d}s$ is the action of $\boldsymbol{q} \cdot \boldsymbol{n}$ on $v_b \in H^{\frac{1}{2}}(\partial K)$.

In WG methods, we also need discrete analogues of the weak gradient. We consider a shape-regular partition $\mathcal{T} = \cup\{\tau\}$ for the domain Ω. For each integer $l \geq 0$, let $P_l(\tau)$ be the set of polynomials on τ with degree no more than l and $\hat{P}_l(\tau)$ be the set of homogeneous polynomials of order l in the variable $\boldsymbol{x} = (x_1, \ldots, x_d)^T$. Let $\mathbf{G}_l(\tau)$ be either $(P_l(\tau))^d$ or $RT_l(\tau) = (P_l(\tau))^d + \hat{P}_l(\tau)\boldsymbol{x}$. For the weak function space $W(\tau)$, we discretize it by $W_{i,j}(\tau)$ given as follows

$$W_{i,j}(\tau) := \left\{v = \{v_0, v_b\} : v_0 \in P_i(\tau), v_b \in P_j(\partial\tau)\right\}.$$

Definition 2 (*Discrete weak gradient*) The discrete weak gradient of $v = \{v_0, v_b\} \in W_{i,j}(\tau)$ denoted by $\nabla_{w,l,\tau} v$ is defined as the unique polynomial $\nabla_{w,l,\tau} v \in \mathbf{G}_l(\tau)$ satisfying the following equation

$$(\nabla_{w,l,\tau} v, \boldsymbol{q})_\tau := -(v_0, \nabla \cdot \boldsymbol{q})_\tau + \langle v_b, \boldsymbol{q} \cdot \boldsymbol{n}\rangle_{\partial\tau}, \qquad \forall \boldsymbol{q} \in \mathbf{G}_l(\tau). \tag{6}$$

Note that if $v \in H^1(\tau)$ and $\nabla v \in \mathbf{G}_l(\tau)$, then $\nabla_{w,l,\tau} v = \nabla v$.

Different weak Galerkin finite element methods can be derived by choosing $W_{i,j}(\tau)$ and $\mathbf{G}_l(\tau)$ with various combinations of the indices i, j and l (see [20,27]). This paper shall mainly consider two pairs $W_{l,l}(\tau) - RT_l(\tau)$ and $W_{l,l+1}(\tau) - (P_{l+1}(\tau))^d$, for integers $l \geq 0$ defined on simplices τ.

In next subsection, the weak Galerkin finite element spaces and the corresponding bilinear form $a(\cdot, \cdot)$ will be presented.

2.2 Weak Galerkin Finite Element Method

Let $\mathcal{T}_h$ be a shape-regular partition of the domain Ω into a set of elements τ. We use the notation $\mathcal{E}_h$ to denote the set of all edges or faces in $\mathcal{T}_h$ and $\mathcal{E}_h^0 = \mathcal{E}_h \setminus \partial\Omega$ denote the set of all interior edges or faces. For a d-dimensional simplex S, we write $h_S = |S|^{1/d}$ to denote the size of the element S where $|S|$ is the d-dimensional Lebesgue measure of S.

Denote by $W_l(\tau) - \mathbf{G}_l(\tau)$ a local weak Galerkin element that can be either $W_{l,l}(\tau) - RT_l(\tau)$ or $W_{l,l+1}(\tau) - (P_{l+1}(\tau))^d$. Associated with $\mathcal{T}_h$ and a local element $W_l(\tau) - \mathbf{G}_l(\tau)$, we define global weak Galerkin finite element spaces,

$$\begin{aligned} V_h &:= \{v = \{v_0, v_b\} : \{v_0, v_b\}|_\tau \in W_l(\tau)\}, \\ V_h^0 &:= \{v : v \in V_h, v_b = 0 \ \text{ on } \ \partial\Omega\}. \end{aligned}$$

We would like to emphasize that any function $v = \{v_0, v_b\} \in V_h^0$ has a single value v_b on each edge $e \in \mathcal{E}_h$. We can also note that $v = \{v_0, v_b\} \in V_h^0$ is a reasonable approximation of a function in $H_0^1(\Omega)$ (see Sect. 3 in [6]).

Now, we can define the discrete weak gradient operator $\nabla_{w,l}$ on the weak finite element space V_h, which is computed element-wise by using (6); i.e., for any τ, we have $\nabla_{w,l,\tau}(v|_\tau) \in \mathbf{G}_l(\tau)$ and

$$(\nabla_{w,l} v)|_\tau := \nabla_{w,l,\tau}(v|_\tau) \quad \forall v \in V_h.$$

Here and afterwards, for simplicity of notation, we shall drop the subscript l in the notation $\nabla_{w,l}$ for the discrete weak gradient when no confusion arises.

For any $w, v \in W_l(\tau) - \mathbf{G}_l(\tau)$, we present the bilinear form as follows

$$a(w, v) = (A\nabla_w w, \nabla_w v)_{\mathcal{T}_h} := \sum_{\tau \in \mathcal{T}_h} (A\nabla_w w, \nabla_w v)_\tau.$$

The WG methods for solving for (1)–(2): find $u_h = \{u_0^h, u_b^h\} \in V_h^0$, such that

$$a(u_h, v_h) = (f, v_0^h), \quad \forall v_h = \{v_0^h, v_b^h\} \in V_h^0. \tag{7}$$

The well-posedness of variational problem (7) can be found in [27].

Remark 1 Optimal order error estimates, which are between weak Galerkin finite element solutions and the exact solution in both the discrete H^1 and L^2 norms, were also presented in [27].

3 Adaptive Weak Galerkin Finite Element Methods

In this section, we present the standard adaptive algorithm (see Sect. 5 in [6]) and discuss each step in the algorithm in detail.

$[u_J, \mathcal{T}_J] = \mathrm{AWG}(\mathcal{T}_0, f, \mathrm{tol}, \theta)$
AWG compute an approximation u_J by adaptive finite element methods.
Input: $\mathcal{T}_0$ initial triangulation; f data; tol stopping criteria; $\theta \in (0, 1)$ marking parameter.
Output: $\mathcal{T}_J$ a triangulation; u_J WG finite element approximation on $\mathcal{T}_J$.
$\eta = 1$; $k = 0$;
while $\eta \geq$ tol
    SOLVE: equation (7) 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: elements in $\mathcal{M}_k$ and necessary elements to a conforming triangulation $\mathcal{T}_{k+1}$;
    $k = k + 1$
end
$u_J = u_k$; $\mathcal{T}_J = \mathcal{T}_k$;

The goal of this paper is to prove that the algorithm AWG will terminate in finite steps for a given tolerance. In the following subsections, we shall discuss each step in detail.

3.1 Procedure SOLVE

Given a function $f \in L^2(\Omega)$ and a shape regular mesh $\mathcal{T}_k$, let u_k be the exact WG solution of (7). In this step, we suppose that the finite dimensional problems (7) will be solved efficiently and accurately.

3.2 Procedure ESTIMATE

The crucial ingredient of the AWG is the control of the error by the estimator, namely the so-called reliability. Here, we will use a similar residual-type a posteriori error estimator in [6]. Given a mesh $\mathcal{T}_h$, assume two elements τ_1 and τ_2 sharing a common edge or face e and denote $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ the unit normal vectors on e exterior to τ_1 and τ_2. In $\mathbb{R}^2$, the unit tangential vectors $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ will obtained by rotating $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ 90 degrees counterclockwise, then denote $\gamma_{t,\partial\tau_i}(\boldsymbol{v}) = \boldsymbol{v}\cdot\boldsymbol{t}_i$ the tangential trace in τ_i of a vector function $\boldsymbol{v}$. In $\mathbb{R}^3$, the tangential trace for $\boldsymbol{v}$ in τ_i is $\gamma_{t,\partial\tau_i}(\boldsymbol{v}) = \boldsymbol{v}\times\boldsymbol{n}_i$ for $i = 1, 2$. Then the normal jump across e is defined as $[\boldsymbol{w}\cdot\boldsymbol{n}]_e = \boldsymbol{w}|_{\partial\tau_1}\cdot\boldsymbol{n}_1 + \boldsymbol{w}|_{\partial\tau_2}\cdot\boldsymbol{n}_2$ and the tangential jump across e is defined as $[\gamma_t(\boldsymbol{w})]_e = \gamma_{t,\partial\tau_1}(\boldsymbol{w}) + \gamma_{t,\partial\tau_2}(\boldsymbol{w})$. For $\forall v_h \in V_h$, we define

$$\mathbf{J}_e(A\nabla_w v_h) = \begin{cases} [A\nabla_w v_h\cdot\boldsymbol{n}]_e, & \text{if } e \in \mathcal{E}_h^0 \\ 0, & \text{otherwise,} \end{cases}$$

$$\mathbf{J}_e(\gamma_t(\nabla_w v_h)) = \begin{cases} [\gamma_t(\nabla_w v_h)]_e, & \text{if } e \in \mathcal{E}_h^0 \\ 2\gamma_t(\nabla_w v_h), & \text{otherwise.} \end{cases}$$

For $e \in \mathcal{E}_h^0$, denote by $\omega_e = \tau_1 \cup \tau_2$ the macro-element associated with e, where τ_1 and τ_2 are two elements in $\mathcal{T}_h$ sharing e as a common edge/face. Similarly, we define $\omega_x = \{\tau' \in \mathcal{T}_h, x \in \tau'\}$ for a vertex x, and $\omega_\tau = \{\tau' \in \mathcal{T}_h, \tau'\cap\tau \neq \varnothing\}$ for an element $\tau \in \mathcal{T}_h$. For the piece-wise constant A, we use $|A|$ to denote its absolute value. We use the notations $A_\tau = A|_\tau$, $|A_e^{\max}| = \max_{\tau\in w_e}|A_\tau|$, and $|A_e^{\min}| = \min_{\tau\in w_e}|A_\tau|$.

Let f_h be the L^2 projection of f to the discontinuous Galerkin space

$$S_h = \{w \in L^2(\Omega) : w|_\tau \in P_l(\tau), \forall\tau \in \mathcal{T}_h\}. \tag{8}$$

Then, for $v_h \in V_h$ and $\tau \in \mathcal{T}_h$, we define

$$\eta_c^2(v_h,\tau) = h_\tau^2|A_\tau|^{-1}\|f_h + \nabla\cdot(A\nabla_w v_h)\|_\tau^2 + \frac{1}{2}\sum_{e\in\partial\tau} h_\tau|A_e^{\max}|^{-1}\int_e \mathbf{J}_e^2(A\nabla_w v_h), \tag{9}$$

$$\eta_m^2(v_h,\tau) = h_\tau^2|A_\tau|\cdot\|\nabla\times\nabla_w v_h\|_\tau^2 + \frac{1}{2}\sum_{e\in\partial\tau} h_\tau|A_e^{\min}|\int_e \mathbf{J}_e^2(\gamma_t(\nabla_w v_h)), \tag{10}$$

$$\operatorname{osc}^2(f,\tau) = h_\tau^2|A_\tau|^{-1}\|f - f_h\|_\tau^2, \tag{11}$$

and element-wise error estimator

$$\eta^2(v_h,\tau) = \operatorname{osc}^2(f,\tau) + \eta_c^2(v_h,\tau) + \eta_m^2(v_h,\tau), \tag{12}$$

Remark 2 Note that $\eta_c(v_h,\tau)$ is an analogy of the error estimator for the conforming finite element and $\eta_m(v_h,\tau)$ is an analogy of the error estimator for the mixed finite element. The $\operatorname{osc}(f,\tau)$ is an analogy of the data oscillation for conforming finite elements.

Remark 3 There is a slight difference between the error estimator given in (12) and one introduced in [6]. For the mesh size in the jump terms, we use h_τ instead of h_e. Although h_τ and h_e are comparable, the use of h_τ is crucial for the reduction of the error estimator, as we can see from the proof of Lemma 10.

For any subset $\mathcal{W}_h \subset \mathcal{T}_h$ and $v_h \in V_h$, we define

$$\eta^2(v_h, \mathcal{W}_h) = \sum_{\tau \in \mathcal{W}_h} \eta^2(v_h, \tau), \quad \text{osc}^2(f, \mathcal{W}_h) = \sum_{\tau \in \mathcal{W}_h} \text{osc}^2(f, \tau).$$

3.3 Procedure MARK

In the selection of elements, we rely on the Dörfler marking [8]. Given a mesh $\mathcal{T}_k$, a set of indicators $\{\eta^2(u_k, \tau_k)\}_{\tau_k \in \mathcal{T}_k}$, and a marking parameter $\theta \in (0, 1)$, we suppose that the procedure **MARK** outputs a subset of marked elements $\mathcal{M}_k \subset \mathcal{T}_k$ with minimal cardinality, such that

$$\eta^2(u_k, \mathcal{M}_k) \geq \theta \eta^2(u_k, \mathcal{T}_k). \tag{13}$$

3.4 Procedure REFINE

Starting from an initial triangulation $\mathcal{T}_0$, we denote by

$$\mathbb{L}(\mathcal{T}_0) = \{\mathcal{T} : \mathcal{T} \text{ is conforming and refined from } \mathcal{T}_0\}, \tag{14}$$

and $\mathcal{T}_1 \leq \mathcal{T}_2$ if $\mathcal{T}_2$ is a refinement of $\mathcal{T}_1$.

For any $\mathcal{T}_k \in \mathbb{L}(\mathcal{T}_0)$ and a subset $\mathcal{M}_k \subset \mathcal{T}_k$ of marked elements, we suppose that *Procedure* **REFINE** outputs a conforming triangulation $\mathcal{T}_{k+1} \in \mathbb{L}(\mathcal{T}_0)$, i.e.,

$$\mathcal{T}_{k+1} = \text{REFINE}(\mathcal{T}_k, \mathcal{M}_k).$$

To generate $\mathcal{T}_{k+1}$, we first subdivide the marked elements in $\mathcal{M}_k$ to get new triangulation $\mathcal{T}'_k$. In general, $\mathcal{T}'_k$ might have hanging nodes; therefore, we have to refine additional elements in $\mathcal{T}_k \backslash \mathcal{M}_k$ to obtain a conforming triangulation $\mathcal{T}_{k+1}$. Throughout this paper, we shall impose the local refinement $\mathbb{L}(\mathcal{T}_0)$ is shape regular.

4 Convergence of the AWG Method

In this section, we begin with a quasi-orthogonality result. Then, we recall the upper bound of the a posteriori error estimator (see [6]). Moreover, we present the reduction of $\text{osc}^2(f, \mathcal{T}_h)$ and $\eta_1^2(v_h, \mathcal{T}_h) = \sum_{\tau \in \mathcal{T}_h} (\eta_c^2(v_h, \tau) + \eta_m^2(v_h, \tau))$, respectively. At last, we prove that the sum of the energy error and the error estimator, between two consecutive adaptive loops, is a contraction and the adaptive algorithm will terminate in finite steps within a given tolerance.

4.1 Quasi-Orthogonality

The standard convergence of adaptive Galerkin method is based on the orthogonality or quasi-orthogonality of the error in different finite element spaces. Especially, for the case of the mixed methods, we refer to [5,9]. However, the quasi-orthogonality of WG methods are more complicated than ones for mixed element methods.

First, for $\tau \in \mathcal{T}_h$, we denote the L^2 projection onto $W_l(\tau)$ by $Q_\tau \cdot = \{Q_0^\tau \cdot, Q_b^\tau \cdot\}$ and L^2 projection onto $\mathbf{G}_l(\tau)$ by $\mathbb{Q}_\tau$. Next Lemma presents the conservation property of the WG approximation.

Lemma 1 *Let u be the solution of (1)–(2) and $u_h = \{u_0^h, u_b^h\} \in V_h^0$ be the solution of (7). Then we have $A\nabla_w u_h \in \mathbf{H}(\mathrm{div}, \Omega)$ and*

$$- \nabla \cdot (A\nabla_w u_h) = f_h, \tag{15}$$

where f_h is the L^2 projection of f to the space S_h.

Proof The proof of the Lemma 1 is similar as Lemma 3.3 in [6]. Notice that the coefficient A is piece-wise constant.

Let $v = \{0, v_b\}$ in (6),

$$\begin{aligned} a(u_h, v) &= (A\nabla_w u_h, \nabla_w v) \\ &= \sum_{\tau \in \mathcal{T}_h} (A\nabla_w u_h, \nabla_w v)_\tau \\ &= \sum_{\tau \in \mathcal{T}_h} (-(v_0, \nabla \cdot (A\nabla_w u_h))_\tau + \langle v_b, (A\nabla_w u_h) \cdot \mathbf{n} \rangle_{\partial\tau} \\ &= \sum_{\tau \in \mathcal{T}_h} \langle v_b, (A\nabla_w u_h) \cdot \mathbf{n} \rangle_{\partial\tau} \\ &= \sum_{e \in \mathcal{E}_h^0} \langle v_b, \boldsymbol{J}_e((A\nabla_w u_h) \cdot \mathbf{n}) \rangle_e, \end{aligned}$$

using (7) leads to

$$a(u_h, v) = (f, v_0) = (f, 0) = 0,$$

we have

$$\sum_{e \in \mathcal{E}_h^o} \langle v_b, \boldsymbol{J}_e((A\nabla_w u_h) \cdot \mathbf{n}) \rangle_e = 0.$$

Choose $v_b|_e = \boldsymbol{J}_e((A\nabla_w u_h) \cdot \mathbf{n})$, we have

$$\boldsymbol{J}_e((A\nabla_w u_h) \cdot \mathbf{n}) = 0,$$

such that $(A\nabla_w u_h) \cdot \mathbf{n}$ is continuous across every edge/face. Therefore, $A\nabla_w u_h \in \mathbf{H}(\mathrm{div}, \Omega)$.

When $v = \{v_0, 0\}$, we get

$$\begin{aligned} a(u_h, v) &= \sum_{\tau \in \mathcal{T}_h} (-(v_0, \nabla \cdot (A\nabla_w u_h))_\tau + \langle v_b, (A\nabla_w u_h) \cdot \mathbf{n} \rangle_{\partial\tau} \\ &= -\sum_{\tau \in \mathcal{T}_h} (v_0, \nabla \cdot (A\nabla_w u_h))_\tau \\ &= (f, v_0) = (f_h, v_0), \end{aligned}$$

which implies

$$-\nabla \cdot (A\nabla_w u_h) = f_h.$$

□

For two nested triangulations $\mathcal{T}_h, \mathcal{T}_{h_*} \in \mathbb{L}(\mathcal{T}_0)$ with $\mathcal{T}_h \leq \mathcal{T}_{h_*}$, in order to prove the quasi-orthogonality of WG methods, we also introduce an intermediate solution $\tilde{u}_{h_*} = \{\tilde{u}_0^{h_*}, \tilde{u}_b^{h_*}\} \in V_{h_*}^0$ satisfying the following equation,

$$(A\nabla_w \tilde{u}_{h_*}, \nabla_w v_{h_*})_{\mathcal{T}_{h_*}} = (f_h, v_0^{h_*}) \qquad \forall v_{h_*} = \{v_0^{h_*}, v_b^{h_*}\} \in V_{h_*}^0. \tag{16}$$

The following lemma presents the property of the intermediate solution $\tilde{u}_{h_*}$.

Lemma 2 *Given an $f \in L^2(\Omega)$ and two nested triangulations $\mathcal{T}_h, \mathcal{T}_{h_*} \in \mathbb{L}(\mathcal{T}_0)$ with $\mathcal{T}_h \leq \mathcal{T}_{h_*}$, let u be the solution of (1)–(2), $u_h = \{u_0^h, u_b^h\} \in V_h^0$ and $u_{h_*} = \{u_0^{h_*}, u_b^{h_*}\} \in V_{h_*}^0$ be the corresponding WG solutions of (7), $\tilde{u}_{h_*} = \{\tilde{u}_0^{h_*}, \tilde{u}_b^{h_*}\} \in V_{h_*}^0$ be the solution of (16). Then*

$$\sum_{\tau_* \in \mathcal{T}_{h_*}} (\nabla u - \nabla_{w,\tau_*} u_{h_*}, A(\nabla_{w,\tau_*} \tilde{u}_{h_*} - \nabla_{w,\tau} u_h))_{\tau_*} = 0, \tag{17}$$

where $\tau \in \mathcal{T}_h$, $\tau_* \in \mathcal{T}_{h_*}$ and $\tau_* \subseteq \tau$.

Proof The main idea follows from [5].

For all $\tau_* \in \mathcal{T}_{h_*}$, let $Q_{\tau_*} \cdot = \{Q_0^{\tau_*} \cdot, Q_b^{\tau_*} \cdot\}$ and the $\mathbb{Q}_{\tau_*}$ be the L^2 projection to $W_l(\tau_*)$ and $\mathbf{G}_l(\tau_*)$, respectively. Comparing the right-hand sides of (7) and (16), then using the similar proof of Lemma 1 and note that projection from V_h to V_h is the identity operator, we obtain $A\nabla_w \tilde{u}_{h_*} \in \mathbf{H}(\mathrm{div}, \Omega)$ and

$$-\nabla \cdot (A\nabla_{w,\tau_*} \tilde{u}_{h_*}) = f_h. \tag{18}$$

For all $\tau_* \in \mathcal{T}_{h_*}$, we have $\nabla_{w,\tau_*}(Q_{\tau_*} v) = \mathbb{Q}_{\tau_*} \nabla v, \forall v \in H^1(\tau_*)$.

Then Lemma 1 implies $A\nabla_w u_h \in \mathbf{H}(\mathrm{div}, \Omega)$, both $u \in H_0^1(\Omega)$ and $u_{h_*} = \{u_0^{h_*}, u_b^{h_*}\} \in V_{h_*}^0$ implies that $Q_b^{\tau_*} u$, $u_b^{h_*}$ severally have a single value on each edge $e \in \mathcal{E}_*^0$, $Q_b^{\tau_*} u|_{\partial\Omega} = u_b^{h_*}|_{\partial\Omega} = 0$, (15) and (18), we have

$$\begin{aligned}
&\sum_{\tau_* \in \mathcal{T}_{h_*}} \left(\nabla u - \nabla_{w,\tau_*} u_{h_*}, A(\nabla_{w,\tau_*} \tilde{u}_{h_*} - \nabla_{w,\tau} u_h)\right)_{\tau_*} \\
&\quad = \sum_{\tau_* \in \mathcal{T}_{h_*}} \left(\mathbb{Q}_{\tau_*}(\nabla u - \nabla_{w,\tau_*} u_{h_*}), A(\nabla_{w,\tau_*} \tilde{u}_{h_*} - \nabla_{w,\tau} u_h)\right)_{\tau_*} \\
&\quad = \sum_{\tau_* \in \mathcal{T}_{h_*}} \left(\mathbb{Q}_{\tau_*} \nabla u - \nabla_{w,\tau_*} u_{h_*}, A(\nabla_{w,\tau_*} \tilde{u}_{h_*} - \nabla_{w,\tau} u_h)\right)_{\tau_*} \\
&\quad = \sum_{\tau_* \in \mathcal{T}_{h_*}} \left(\nabla_{w,\tau_*}(Q_{\tau_*} u - u_{h_*}), A\nabla_{w,\tau_*} \tilde{u}_{h_*}\right)_{\tau_*} \\
&\qquad - \sum_{\tau_* \in \mathcal{T}_{h_*}} \left(\nabla_{w,\tau_*}(Q_{\tau_*} u - u_{h_*}), A\nabla_{w,\tau} u_h\right)_{\tau_*} \\
&\quad = -\sum_{\tau_* \in \mathcal{T}_{h_*}} \left(Q_0^{\tau_*} u - u_0^{h_*}, \nabla \cdot (A\nabla_{w,\tau_*} \tilde{u}_{h_*})\right)_{\tau_*} \\
&\qquad + \sum_{\tau_* \in \mathcal{T}_{h_*}} \langle Q_b^{\tau_*} u - u_b^{h_*}, (A\nabla_{w,\tau_*} \tilde{u}_{h_*}) \cdot \boldsymbol{n} \rangle_{\partial\tau_*} \\
&\qquad + \sum_{\tau_* \in \mathcal{T}_{h_*}} \left(Q_0^{\tau_*} u - u_0^{h_*}, \nabla \cdot (A\nabla_{w,\tau} u_h)\right)_{\tau_*}
\end{aligned}$$

$$
\begin{aligned}
&- \sum_{\tau_*\in\mathcal{T}_{h_*}} \langle Q_b^{\tau_*}u - u_b^{h_*}, (A\nabla_{w,\tau}u_h)\cdot\boldsymbol{n}\rangle_{\partial\tau_*} \\
&= \sum_{\tau_*\in\mathcal{T}_{h_*}} (Q_0^{\tau_*}u - u_0^{h_*}, f_h - f_h)_{\tau_*} = 0. \qquad (19)
\end{aligned}
$$

□

The following lemma reveals the relationship between $\tilde{u}_0^{h_*} - u_0^{h_*}$ and $\nabla_w\tilde{u}_{h_*} - \nabla_w u_{h_*}$.

Lemma 3 *Let $u_{h_*} = \{u_0^{h_*}, u_b^{h_*}\} \in V_{h_*}^0$ and $\tilde{u}_{h_*} = \{\tilde{u}_0^{h_*}, \tilde{u}_b^{h_*}\} \in V_{h_*}^0$ be the WG solutions of (7) and (16), respectively. Assume that problem (1)–(2) has the H^{1+s} regularity with $s \in (0, 1]$. Then, we have*

$$
\|\tilde{u}_0^{h_*} - u_0^{h_*}\|_{\mathcal{T}_{h_*}} \lesssim h_{\tau_*}^s \|\nabla_w\tilde{u}_{h_*} - \nabla_w u_{h_*}\|_{\mathcal{T}_{h_*}}, \qquad (20)
$$

where the constant only depends on the shape regularity of $\mathcal{T}_{h_*}$ and coefficient A.

Proof Here, we adapt the technique from [27].

Let $w \in H^1(\Omega)$ solve the following auxiliary problem

$$
\begin{cases} -\nabla\cdot(A\nabla w) = \tilde{u}_0^{h_*} - u_0^{h_*} & \text{in } \Omega, \\ w = 0 \quad \text{on } \partial\Omega. \end{cases} \qquad (21)
$$

Then the assumption of H^{1+s} regularity implies that $w \in H^{1+s}(\Omega)$ such that

$$
\|w\|_{1+s} \lesssim \|\tilde{u}_0^{h_*} - u_0^{h_*}\|_{\mathcal{T}_{h_*}}. \qquad (22)
$$

We choose the projection Π_h introduced in [2] satisfying the following two properties

$$
\sum_{\tau\in\mathcal{T}_h} (-\nabla\cdot\boldsymbol{q}, v_0)_\tau = \sum_{\tau\in\mathcal{T}_h} (\Pi_h\boldsymbol{q}, \nabla_w v)_\tau \quad \forall\boldsymbol{q}\in\mathbf{H}(\mathrm{div},\Omega), \forall v = \{v_0, v_b\}\in V_h, \qquad (23)
$$

$$
\|\Pi_h(A\nabla u) - A\nabla_w(Q_\tau u)\| \lesssim h^s\|u\|_{1+s} \qquad \forall u\in H^{1+s}(\Omega), s > 0. \qquad (24)
$$

Formulas (23) and (24) can be found in the Lemmas 7.2 and 7.3 of [27], respectively.

Using the variational problem of (21) with the test function $\tilde{u}_0^{h_*} - u_0^{h_*}$, (23) and (24), we have

$$
\begin{aligned}
\|\tilde{u}_0^{h_*} - u_0^{h_*}\|_{\mathcal{T}_{h_*}}^2 &= \sum_{\tau_*\in\mathcal{T}_{h_*}} (-\nabla\cdot(A\nabla w), \tilde{u}_0^{h_*} - u_0^{h_*})_{\tau_*} \\
&= \sum_{\tau_*\in\mathcal{T}_{h_*}} (\Pi_{h_*}(A\nabla w), \nabla_w\tilde{u}_{h_*} - \nabla_w u_{h_*})_{\tau_*} \\
&= \big(\Pi_{h_*}(A\nabla w) - A\nabla_w(Q_{h_*}w), \nabla_w\tilde{u}_{h_*} - \nabla_w u_{h_*}\big)_{\mathcal{T}_{h_*}} \\
&\lesssim h_{\tau_*}^s\|w\|_{1+s}\|\nabla_w\tilde{u}_{h_*} - \nabla_w u_{h_*}\|_{\mathcal{T}_{h_*}}, \qquad (25)
\end{aligned}
$$

where the constant only depends on the shape regularity of $\mathcal{T}_{h_*}$ and coefficient A. We also used the following equality in the last equal

$$
(A\nabla_w(Q_{h_*}w), \nabla_w\tilde{u}_{h_*} - \nabla_w u_{h_*})_{\mathcal{T}_{h_*}} = (f_h - f_{h_*}, Q_0^{h_*}w)_{\mathcal{T}_{h_*}} = 0.
$$

Substituting (22) into (25), we arrive at

$$
\|\tilde{u}_0^{h_*} - u_0^{h_*}\|_{\mathcal{T}_{h_*}} \lesssim h_{\tau_*}^s\|\nabla_w\tilde{u}_{h_*} - \nabla_w u_{h_*}\|_{\mathcal{T}_{h_*}},
$$

which completes the proof. □

Now we define $\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}$ as the set of refined elements from $\mathcal{T}_h$ to $\mathcal{T}_{h_*}$ and $\overline{\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}}$ as the set of new elements refined from $\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}$. Obviously, $\mathcal{T}_{h_*} \backslash \overline{\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}} = \mathcal{T}_h \backslash \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}$ are unchanged elements.

Lemma 4 *For $\mathcal{T}_h, \mathcal{T}_{h_*} \in \mathbb{L}(\mathcal{T}_0)$ with $\mathcal{T}_h \leq \mathcal{T}_{h_*}$, then we have*

$$\| f_{h_*} - f_h \|_{\tau_*} \begin{cases} = 0, & \forall \tau_* \in \mathcal{T}_{h_*} \backslash \overline{\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}}, \\ \leq \| f - f_h \|_{\tau_*}, & \forall \tau_* \in \overline{\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}}. \end{cases}$$

Proof Notice that the functions f_h and f_{h_*} are the L^2 projections of f to the spaces S_h and S_{h_*}, respectively. Then for any $\tau_* \in \mathcal{T}_{h_*} \backslash \overline{\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}}$, we easily get $\| f_{h_*} - f_h \|_{\tau_*} = 0$. For any $\tau_* \in \overline{\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}}$, since $(f - f_{h_*}, v_{h_*})_{\tau_*} = 0, \forall v_{h_*} \in V_{h_*}$. In particular, let

$$v_{h_*} = \begin{cases} f_h - f_{h_*}, & \text{on } \tau_* \in \overline{\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}}, \\ 0, & \text{otherwise}. \end{cases}$$

Then

$$(f - f_{h_*}, f_h - f_{h_*})_{\tau_*} = 0. \tag{26}$$

According to (26) and Cauchy–Schwarz inequality, we get

$$\begin{aligned} \| f_{h_*} - f_h \|_{\tau_*}^2 &= (f_{h_*} - f_h, f_{h_*} - f_h)_{\tau_*} \\ &= (f_{h_*} - f, f_{h_*} - f_h)_{\tau_*} + (f - f_h, f_{h_*} - f_h)_{\tau_*} \\ &= (f - f_h, f_{h_*} - f_h)_{\tau_*} \leq \| f - f_h \|_{\tau_*} \| f_{h_*} - f_h \|_{\tau_*}. \end{aligned}$$

Canceling one $\| f_{h_*} - f_h \|_{\tau_*}$, we will get $\| f_{h_*} - f_h \|_{\tau_*} \leq \| f - f_h \|_{\tau_*}$. □

In the rest of this subsection, we will prove the following discrete result, and use it to derive the quasi-orthogonality.

Lemma 5 *Given an $f \in L^2(\Omega)$ and two triangulations $\mathcal{T}_h, \mathcal{T}_{h_*} \in \mathbb{L}(\mathcal{T}_0)$ with $\mathcal{T}_h \leq \mathcal{T}_{h_*}$, let u be the solution of (1)–(2), $u_h = \{u_0^h, u_b^h\} \in V_h^0$ and $u_{h_*} = \{u_0^{h_*}, u_b^{h_*}\} \in V_{h_*}^0$ be the corresponding WG solutions of (7), $\tilde{u}_{h_*} = \{\tilde{u}_0^{h_*}, \tilde{u}_b^{h_*}\} \in V_{h_*}^0$ be the solution of the variational problem (16). Then there exists a constant C_0 which depends only on the shape regularity of $\mathcal{T}_{h_*}$, satisfying*

$$\| A^{1/2} (\nabla_w \tilde{u}_{h_*} - \nabla_w u_{h_*}) \|_{\mathcal{T}_{h_*}} \leq \sqrt{C_0} osc(f, \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}).$$

Proof Applying (7) and (16), then for any $v_{h_*} = \{v_0^{h_*}, v_b^{h_*}\} \in V_{h_*}^0$, we have

$$(A(\nabla_w \tilde{u}_{h_*} - \nabla_w u_{h_*}), \nabla_w v_{h_*})_{\mathcal{T}_{h_*}} = (f_h - f_{h_*}, v_0^{h_*}). \tag{27}$$

Noting that $\mathcal{T}_{h_*} \backslash \overline{\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}} = \mathcal{T}_h \backslash \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}$ are unchanged elements, choosing $v_{h_*} = \tilde{u}_{h_*} - u_{h_*} \in V_{h_*}^0$ in (27) and using the property of L^2 projection, Hölder inequality, (20), Cauchy–Schwarz inequality, we arrive at

$$\begin{aligned} & \| A^{1/2} (\nabla_w \tilde{u}_{h_*} - \nabla_w u_{h_*}) \|_{\mathcal{T}_{h_*}}^2 \\ & = (f_h - f_{h_*}, \tilde{u}_0^{h_*} - u_0^{h_*})_{\mathcal{T}_{h_*}} = \sum_{\tau_* \in \overline{\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}}} (f_h - f, \tilde{u}_0^{h_*} - u_0^{h_*})_{\tau_*} \\ & \leq \sum_{\tau_* \in \overline{\mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}}} \| f - f_h \|_{\tau_*} \cdot \| \tilde{u}_0^{h_*} - u_0^{h_*} \|_{\tau_*} \end{aligned}$$

$$
\begin{aligned}
&\lesssim \sum_{\tau_* \in \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}} \|f - f_h\|_{\tau_*} \cdot h_{\tau_*} \|\nabla_w \tilde{u}_{h_*} - \nabla_w u_{h_*}\|_{\tau_*} \\
&\lesssim \sum_{\tau_* \in \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}} |A|^{-\frac{1}{2}} \|f - f_h\|_{\tau_*} \cdot h_{\tau_*} \|A^{1/2}(\nabla_w \tilde{u}_{h_*} - \nabla_w u_{h_*})\|_{\tau_*} \\
&\lesssim \left(\sum_{\tau_* \in \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}} h_{\tau_*}^2 |A|^{-1} \|f - f_h\|_{\tau_*}^2 \right)^{1/2} \left(\sum_{\tau_* \in \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}} \|A^{1/2}(\nabla_w \tilde{u}_{h_*} - \nabla_w u_{h_*})\|_{\tau_*}^2 \right)^{1/2} \\
&\lesssim \left(\sum_{\tau \in \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}} h_{\tau}^2 |A|^{-1} \|f - f_h\|_{\tau}^2 \right)^{1/2} \left(\sum_{\tau_* \in \mathcal{T}_{h_*}} \|A^{1/2}(\nabla_w \tilde{u}_{h_*} - \nabla_w u_{h_*})\|_{\tau_*}^2 \right)^{1/2} \\
&\lesssim \mathrm{osc}(f, \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}) \cdot \|A^{1/2}(\nabla_w \tilde{u}_{h_*} - \nabla_w u_{h_*})\|_{\mathcal{T}_{h_*}},
\end{aligned}
$$

where the constants only depends on the shape regularity of $\mathcal{T}_{h_*}$. At last, canceling one $\|A^{1/2}(\nabla_w \tilde{u}_{h_*} - \nabla_w u_{h_*})\|_{\mathcal{T}_{h_*}}$, then there exist a constant C_0, such that

$$
\|A^{1/2}(\nabla_w \tilde{u}_{h_*} - \nabla_w u_{h_*})\|_{\mathcal{T}_{h_*}} \le \sqrt{C_0}\mathrm{osc}(f, \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}).
$$

□

Now, we use Lemmas 2 and 5 to derive a quasi-orthogonality result.

Lemma 6 *Given an $f \in L^2(\Omega)$ and two triangulations $\mathcal{T}_h, \mathcal{T}_{h_*} \in \mathbb{L}(\mathcal{T}_0)$ defined in (14) with $\mathcal{T}_h \le \mathcal{T}_{h_*}$, let u be the solution of (1)–(2), $u_h = \{u_0^h, u_b^h\} \in V_h^0$ and $u_{h_*} = \{u_0^{h_*}, u_b^{h_*}\} \in V_{h_*}^0$ be the corresponding WG solutions of (7). Then for any $\epsilon \in (0, 1)$, we have*

$$
\begin{aligned}
(1-\epsilon)\|A^{1/2}(\nabla u - \nabla_{w,\tau_*} u_{h_*})\|_{\mathcal{T}_{h_*}}^2 &\le \|A^{1/2}(\nabla u - \nabla_{w,\tau} u_h)\|_{\mathcal{T}_h}^2 \\
&\quad -\|A^{1/2}(\nabla_{w,\tau_*} u_{h_*} - \nabla_{w,\tau} u_h)\|_{\mathcal{T}_{h_*}}^2 + \frac{C_0}{\epsilon} osc^2(f, \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}),
\end{aligned}
\tag{28}
$$

where the constant C_0 is given in Lemma 5, $\tau \in \mathcal{T}_h$, $\tau_* \in \mathcal{T}_{h_*}$ and $\tau_* \subseteq \tau$.

Proof First, making use of Lemma 2, Cauchy–Schwarz inequality and Lemma 5, we obtain

$$
\begin{aligned}
&(A^{1/2}(\nabla u - \nabla_{w,\tau_*} u_{h_*}), A^{1/2}(\nabla_{w,\tau} u_h - \nabla_{w,\tau_*} u_{h_*}))_{\mathcal{T}_{h_*}} \\
&\quad = (A(\nabla u - \nabla_{w,\tau_*} u_{h_*}), \nabla_{w,\tau} u_h - \nabla_{w,\tau_*} \tilde{u}_{h_*})_{\mathcal{T}_{h_*}} \\
&\qquad + (A^{1/2}(\nabla u - \nabla_{w,\tau_*} u_{h_*}), A^{1/2}(\nabla_{w,\tau_*} \tilde{u}_{h_*} - \nabla_{w,\tau_*} u_{h_*}))_{\mathcal{T}_{h_*}} \\
&\quad = (A^{1/2}(\nabla u - \nabla_{w,\tau_*} u_{h_*}), A^{1/2}(\nabla_{w,\tau_*} \tilde{u}_{h_*} - \nabla_{w,\tau_*} u_{h_*}))_{\mathcal{T}_{h_*}} \\
&\quad \le \|A^{1/2}(\nabla u - \nabla_{w,\tau_*} u_{h_*})\|_{\mathcal{T}_{h_*}} \|A^{1/2}(\nabla_{w,\tau_*} \tilde{u}_{h_*} - \nabla_{w,\tau_*} u_{h_*})\|_{\mathcal{T}_{h_*}} \\
&\quad \le \sqrt{C_0}\|A^{1/2}(\nabla u - \nabla_{w,\tau_*} u_{h_*})\|_{\mathcal{T}_{h_*}} \mathrm{osc}(f, \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}).
\end{aligned}
\tag{29}
$$

For any $\epsilon > 0$, using the inequality $2ab \le \epsilon a^2 + \frac{1}{\epsilon} b^2$ and (29), we have

$$
\begin{aligned}
&\|A^{1/2}(\nabla u - \nabla_{w,\tau} u_h)\|_{\mathcal{T}_h}^2 \\
&\quad = \|A^{1/2}(\nabla u - \nabla_{w,\tau_*} u_{h_*})\|_{\mathcal{T}_{h_*}}^2 + \|A^{1/2}(\nabla_{w,\tau_*} u_{h_*} - \nabla_{w,\tau} u_h)\|_{\mathcal{T}_{h_*}}^2
\end{aligned}
$$

$$
\begin{aligned}
&-2\left(A^{1/2}(\nabla u-\nabla_{w,\tau_*}u_{h_*}),A^{1/2}(\nabla_{w,\tau_*}u_{h_*}-\nabla_{w,\tau}u_h)\right)_{\mathcal{T}_{h*}}\\
\geq{}& \|A^{1/2}(\nabla u-\nabla_{w,\tau_*}u_{h_*})\|^2_{\mathcal{T}_{h*}}+\|A^{1/2}(\nabla_{w,\tau_*}u_{h_*}-\nabla_{w,\tau}u_h)\|^2_{\mathcal{T}_{h*}}\\
&-2\|A^{1/2}(\nabla u-\nabla_{w,\tau_*}u_{h_*})\|_{\mathcal{T}_{h*}}\sqrt{C_0}\mathrm{osc}(f,\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h*}})\\
\geq{}&(1-\epsilon)\|A^{1/2}(\nabla u-\nabla_{w,\tau_*}u_{h_*})\|^2_{\mathcal{T}_{h*}}+\|A^{1/2}(\nabla_{w,\tau_*}u_{h_*}-\nabla_{w,\tau}u_h)\|^2_{\mathcal{T}_{h*}}\\
&-\frac{C_0}{\epsilon}\mathrm{osc}^2(f,\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h*}}).
\end{aligned}
$$

This completes the proof. □

4.2 Residual Type Error Estimate: Upper Bound

In this subsection, we will recall the upper bound, which is important to prove the convergence of the adaptive WG methods.

Lemma 7 (Theorem 4.4 in [6]) *Let u be the solution of (1)–(2) and $u_h=\{u_0^h,u_b^h\}\in V_h^0$ be the solution of (7). Then, there exists a positive constant C_1 depending on the shape regularity of $\mathcal{T}_h$ and coefficient A, such that*

$$\|A^{1/2}(\nabla u-\nabla_w u_h)\|_{\mathcal{T}_h}\leq C_1\eta(u_h,\mathcal{T}_h). \tag{30}$$

Remark 4 Although the error estimator in the above inequality is different from one introduced in [6], they can control each other. We can see from the Remark 3.

4.3 Contraction of the Error Estimator

In this subsection, we shall introduce the contraction of the error estimator. In order to prove that, we will divide the error estimator $\eta(v_h,\mathcal{T}_h)$ into two parts $\mathrm{osc}^2(f,\mathcal{T}_h)$ and $\eta_1^2(v_h,\mathcal{T}_h)$ and present separately the reduction of the two parts.

First, we prove the the reduction of oscillation $\mathrm{osc}^2(f,\mathcal{T}_h)$.

Lemma 8 *For $\mathcal{T}_h,\mathcal{T}_{h_*}\in\mathbb{L}(\mathcal{T}_0)$ with $\mathcal{T}_h\leq\mathcal{T}_{h_*}$, let $\lambda:=1-\mu\in(0,1)$, where $\mu:=2^{-1/d}\in(0,1)$. We have*

$$osc^2(f,\mathcal{T}_{h_*})\leq osc^2(f,\mathcal{T}_h)-\lambda osc^2(f,\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}). \tag{31}$$

Proof For all $\tau_*\in\overline{\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}$, applying with (26), we arrive at

$$
\begin{aligned}
\|f-f_{h_*}\|^2_{\tau_*}&=(f-f_{h_*},f-f_{h_*})_{\tau_*}\\
&=(f-f_{h_*},f-f_h)_{\tau_*}\\
&\leq\|f-f_{h_*}\|_{\tau_*}\|f-f_h\|_{\tau_*},
\end{aligned}
$$

which implies

$$\|f-f_{h_*}\|_{\tau_*}\leq\|f-f_h\|_{\tau_*}. \tag{32}$$

For all $\tau\in\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}$, we suppose that τ is bisected into $\tau_*^1,\tau_*^2\in\mathcal{T}_{h_*}$, then $h^d_{\tau_*^1}=|\tau_*^1|=|\tau_*^2|=h^d_{\tau_*^2}=\frac{1}{2}|\tau|=\frac{1}{2}h^d_\tau(d=2,3)$ together with (11) and (32), yields

$$
\begin{aligned}
&\operatorname{osc}^2(f,\tau_*^1)+\operatorname{osc}^2(f,\tau_*^2)\\
&\quad= h_{\tau_*^1}^2|A_{\tau_*^1}|^{-1}\|f-f_{h_*}\|_{\tau_*^1}^2+h_{\tau_*^2}^2|A_{\tau_*^2}|^{-1}\|f-f_{h_*}\|_{\tau_*^2}^2\\
&\quad\le h_{\tau_*^1}^2|A_{\tau_*^1}|^{-1}\|f-f_{h}\|_{\tau_*^1}^2+h_{\tau_*^2}^2|A_{\tau_*^2}|^{-1}\|f-f_{h}\|_{\tau_*^2}^2\\
&\quad= 2^{-2/d}\cdot h_\tau^2|A_\tau|^{-1}\|f-f_h\|_{\tau_*^1}^2+2^{-2/d}\cdot h_\tau^2|A_\tau|^{-1}\|f-f_h\|_{\tau_*^2}^2\\
&\quad= 2^{-2/d}h_\tau^2|A_\tau|^{-1}\|f-f_h\|_\tau^2\\
&\quad< \mu h_\tau^2|A_\tau|^{-1}\|f-f_h\|_\tau^2\\
&\quad= \mu\operatorname{osc}^2(f,\tau),
\end{aligned}\tag{33}
$$

Using the fact that $\mathcal{T}_{h_*}\backslash\overline{\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}=\mathcal{T}_h\backslash\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}$ in conjunction with (33) and (11), we arrive at

$$
\begin{aligned}
&\operatorname{osc}^2(f,\mathcal{T}_{h_*})\\
&\quad=\sum_{\tau_*\in\mathcal{T}_{h_*}\backslash\overline{\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}} h_{\tau_*}^2|A_{\tau_*}|^{-1}\|f-f_{h_*}\|_{\tau_*}^2+\sum_{\tau_*\in\overline{\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}} h_{\tau_*}^2|A_{\tau_*}|^{-1}\|f-f_{h_*}\|_{\tau_*}^2\\
&\quad\le\sum_{\tau_*\in\mathcal{T}_{h_*}\backslash\overline{\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}} h_{\tau_*}^2|A_{\tau_*}|^{-1}\|f-f_{h_*}\|_{\tau_*}^2+\mu\sum_{\tau\in\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}\operatorname{osc}^2(f,\tau)\\
&\quad=\sum_{\tau\in\mathcal{T}_h\backslash\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}} h_\tau^2|A_\tau|^{-1}\|f-f_h\|_\tau^2+\mu\sum_{\tau\in\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}\operatorname{osc}^2(f,\tau)\\
&\quad\le\sum_{\tau\in\mathcal{T}_h}\operatorname{osc}^2(f,\tau)-\sum_{\tau\in\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}\operatorname{osc}^2(f,\tau)+\mu\sum_{\tau\in\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}\operatorname{osc}^2(f,\tau)\\
&\quad\le\operatorname{osc}^2(f,\mathcal{T}_h)-\lambda\operatorname{osc}^2(f,\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}).
\end{aligned}
$$

We complete the proof. ◻

Now we are in a position to present the reduction of the second part. We first present the difference between $\eta_1^2(v_{h_*},\mathcal{T}_{h_*})$ and $\eta_1^2(v_h,\mathcal{T}_{h_*})$.

Lemma 9 *For $\mathcal{T}_h,\mathcal{T}_{h_*}\in\mathbb{L}(\mathcal{T}_0)$ with $\mathcal{T}_h\le\mathcal{T}_{h_*}$, let $v_h=\{v_0^h,v_b^h\}\in V_h^0$, $v_{h_*}=\{v_0^{h_*},v_b^{h_*}\}\in V_{h_*}$. Then for any $\zeta>0$, there exists constant σ_1 depending on the shape regularity of $\mathcal{T}_{h_*}$, the polynomial order l, coefficient A and parameter ζ, such that*

$$
\begin{aligned}
\eta_1^2(v_{h_*},\mathcal{T}_{h_*})&\le(1+\zeta)\eta_1^2(v_h,\mathcal{T}_{h_*})\\
&\quad+\frac{1}{\sigma_1}\Big(\mu osc^2(f,\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}})+\|A^{1/2}(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\mathcal{T}_{h_*}}^2\Big).
\end{aligned}\tag{34}
$$

Proof For each $\tau_*\in\mathcal{T}_{h_*}$, we will consider the four terms in $\eta_1^2(v_{h_*},\mathcal{T}_{h_*})$ one by one.

a) We first deal with the element terms $R_1(v_{h_*},f_{h_*}):=f_{h_*}+\nabla\cdot(A\nabla_w v_{h_*})$ and $R_2(v_{h_*}):=\nabla\times\nabla_w v_{h_*}$. For $R_1(v_{h_*},f_{h_*})$, using the triangle inequality, we have

$$
\begin{aligned}
&h_{\tau_*}|A_{\tau_*}|^{-1/2}\|R_1(v_{h_*},f_{h_*})\|_{\tau_*}\\
&\quad=h_{\tau_*}|A_{\tau_*}|^{-1/2}\|f_{h_*}+\nabla\cdot(A\nabla_w v_{h_*})\|_{\tau_*}\\
&\quad=h_{\tau_*}|A_{\tau_*}|^{-1/2}\|(f_h+\nabla\cdot(A\nabla_{w,\tau}v_h)+f_{h_*}-f_h+\nabla\cdot A(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*}\\
&\quad\le h_{\tau_*}|A_{\tau_*}|^{-1/2}\|R_1(v_h,f_h)\|_{\tau_*}\\
&\qquad+h_{\tau_*}|A_{\tau_*}|^{-1/2}\|f_{h_*}-f_h+\nabla\cdot A(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*},
\end{aligned}\tag{35}
$$

Applying triangle inequality, chain rule and inverse inequality, we obtain

$$\begin{aligned}
&h_{\tau_*}|A_{\tau_*}|^{-1/2}\|f_{h_*}-f_h+\nabla\cdot A(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*}\\
&\lesssim h_{\tau_*}|A_{\tau_*}|^{-1/2}\|f_{h_*}-f_h\|_{\tau_*}+h_{\tau_*}\|\nabla\cdot A(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*}\\
&\lesssim h_{\tau_*}|A_{\tau_*}|^{-1/2}\|f_{h_*}-f_h\|_{\tau_*}+\|A(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*}\\
&\lesssim h_{\tau_*}|A_{\tau_*}|^{-1/2}\|f_{h_*}-f_h\|_{\tau_*}+\|A^{1/2}(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*}.
\end{aligned}\tag{36}$$

Substituting (36) into (35) and making use of Lemma 4, for any $\tau_*\in\mathcal{T}_{h_*}\backslash\overline{\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}$, we have

$$\begin{aligned}
&h_{\tau_*}|A_{\tau_*}|^{-1/2}\|R_1(v_{h_*},f_{h_*})\|_{\tau_*}\\
&\lesssim h_{\tau_*}|A_{\tau_*}|^{-1/2}\|R_1(v_h,f_h)\|_{\tau_*}+\|A^{1/2}(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*},
\end{aligned}\tag{37}$$

and for any $\tau_*\in\overline{\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}$, we have

$$\begin{aligned}
&h_{\tau_*}|A_{\tau_*}|^{-1/2}\|R_1(v_{h_*},f_{h_*})\|_{\tau_*}\\
&\lesssim h_{\tau_*}|A_{\tau_*}|^{-1/2}\|R_1(v_h,f_h)\|_{\tau_*}+h_{\tau_*}|A_{\tau_*}|^{-1/2}\|f-f_h\|_{\tau_*}\\
&\quad+\|A^{1/2}(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*}.
\end{aligned}\tag{38}$$

For $R_2(v_{h_*})$, a similar method for proving (37), we get

$$\begin{aligned}
&h_{\tau_*}|A_{\tau_*}|^{1/2}\|R_2(v_{h_*})\|_{\tau_*}\\
&\leq h_{\tau_*}|A_{\tau_*}|^{1/2}\|\nabla\times\nabla_{w,\tau}v_h\|_{\tau_*}+h_{\tau_*}|A_{\tau_*}|^{1/2}\|\nabla\times(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*}\\
&\lesssim h_{\tau_*}|A_{\tau_*}|^{1/2}\|R_2(v_h)\|_{\tau_*}+\|A^{1/2}(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*}.
\end{aligned}\tag{39}$$

b) Now, we consider the jump terms $\mathbf{J}_{e_*}(A\nabla_w v_{h_*})$ and $\mathbf{J}_{e_*}(\gamma_t(\nabla_w v_{h_*}))$. For each $e_*\in\mathcal{E}_{h_*}^0$, we assume that $e_*=\tau_*^1\cap\tau_*^2$ with $\tau_*^1,\tau_*^2\in\mathcal{T}_{h_*}$. Let $\boldsymbol{n}_*^1$ and $\boldsymbol{n}_*^2$ be the unit normal vectors on e_* exterior to τ_*^1 and τ_*^2, respectively. Applying the triangle inequality, we obtain

$$\begin{aligned}
&h_{\tau_*}^{1/2}|A_{e_*}^{\max}|^{-1/2}\|\mathbf{J}_{e_*}(A\nabla_{w,\tau_*}v_{h_*})\|_{e_*}\\
&\leq h_{\tau_*}^{1/2}|A_{e_*}^{\max}|^{-1/2}\|\mathbf{J}_{e_*}(A\nabla_{w,\tau}v_h)\|_{e_*}\\
&\quad+h_{\tau_*}^{1/2}|A_{e_*}^{\max}|^{-1/2}\|\mathbf{J}_{e_*}(A(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{e_*}.
\end{aligned}\tag{40}$$

Using the definition of $\mathbf{J}_{e_*}(A(\nabla_{w,\tau}v_h-\nabla_{w,\tau_*}v_{h_*}))$ and trace inequality, we have

$$\begin{aligned}
&h_{\tau_*}^{1/2}|A_{e_*}^{\max}|^{-1/2}\|\mathbf{J}_{e_*}(A(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h))\|_{e_*}\\
&\leq h_{\tau_*}^{1/2}|A_{e_*}^{\max}|^{-1/2}\|A(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)|_{\tau_*^1}\cdot\boldsymbol{n}_*^1\|_{e_*}\\
&\quad+h_{\tau_*}^{1/2}|A_{e_*}^{\max}|^{-1/2}\cdot\|A(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)|_{\tau_*^2}\cdot\boldsymbol{n}_*^2\|_{e_*}\\
&\lesssim h_{\tau_*}^{1/2}\|A(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)|_{\tau_*^1}\|_{e_*}+h_{\tau_*}^{1/2}\|A(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)|_{\tau_*^2}\|_{e_*}\\
&\lesssim\|A^{1/2}(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*^1\cup\tau_*^2}.
\end{aligned}\tag{41}$$

Substituting (41) into (40), we get

$$\begin{aligned}
&h_{\tau_*}^{1/2}|A_{e_*}^{\max}|^{-1/2}\|\mathbf{J}_{e_*}(A\nabla_w v_{h_*})\|_{e_*}\\
&\leq h_{\tau_*}^{1/2}|A_{e_*}^{\max}|^{-1/2}\|\mathbf{J}_{e_*}(A\nabla_{w,\tau}v_h)\|_{e_*}\\
&\quad+\|A^{1/2}(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*^1\cup\tau_*^2}.
\end{aligned}\tag{42}$$

Similar to the proof of (42), we obtain

$$
\begin{aligned}
&h_{\tau_*}^{1/2}|A_{e_*}^{\min}|^{1/2}\|\mathbf{J}_{e_*}(\gamma_t(\nabla_{w,\tau_*}v_{h_*}))\|_{e_*}\\
&\quad\le h_{\tau_*}^{1/2}|A_{e_*}^{\min}|^{1/2}\|\mathbf{J}_{e_*}(\gamma_t(\nabla_{w,\tau}v_h))\|_{e_*}\\
&\qquad+h_{\tau_*}^{1/2}|A_{e_*}^{\min}|^{1/2}\|\mathbf{J}_{e_*}(\gamma_t(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h))\|_{e_*}\\
&\quad\lesssim h_{\tau_*}^{1/2}|A_{e_*}^{\min}|^{1/2}\|\mathbf{J}_{e_*}(\gamma_t(\nabla_{w,\tau}v_h))\|_{e_*}+h_{\tau_*}^{1/2}\|(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)|_{\tau_*^1}\|_{e_*}\\
&\qquad+h_{\tau_*}^{1/2}\|(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)|_{\tau_*^2}\|_{e_*}\\
&\quad\lesssim h_{\tau_*}^{1/2}|A_{e_*}^{\min}|^{1/2}\|\mathbf{J}_{e_*}(\gamma_t(\nabla_{w,\tau}v_h))\|_{e_*}\\
&\qquad+\|A^{1/2}(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*^1\cup\tau_*^2}.
\end{aligned}\tag{43}
$$

For each $e_*\in\mathcal{E}_{h_*}\cap\partial\Omega$, we assume that $e_*\subset\partial\tau_*$ with $\tau_*\in\mathcal{T}_{h_*}$. By the definition of $\mathbf{J}_{e_*}(A\nabla_w v_{h_*})$, we have $\mathbf{J}_{e_*}(A\nabla_w v_{h_*})=0$. Next, a similar method for proving (43), we get

$$
\begin{aligned}
&h_{\tau_*}^{1/2}|A_{e_*}^{\min}|^{1/2}\|\mathbf{J}_{e_*}(\gamma_t(\nabla_w v_{h_*}))\|_{e_*}\\
&\quad\le h_{\tau_*}^{1/2}|A_{e_*}^{\min}|^{1/2}\|\mathbf{J}_{e_*}(\gamma_t(\nabla_{w,\tau}v_h))\|_{e_*}\\
&\qquad+h_{\tau_*}^{1/2}|A_{e_*}^{\min}|^{1/2}\|\mathbf{J}_{e_*}\left(\gamma_t(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\right)\|_{e_*}\\
&\quad\lesssim h_{\tau_*}^{1/2}|A_{e_*}^{\min}|^{1/2}\|\mathbf{J}_{e_*}(\gamma_t(\nabla_{w,\tau}v_h))\|_{e_*}+\|A^{1/2}(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\tau_*}.
\end{aligned}\tag{44}
$$

From (33), we also arrive at

$$
\sum_{\tau_*\in\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}} h_{\tau_*}|A_{\tau_*}|^{-1/2}\|f-f_h\|_{\tau_*}\le\mu\,\mathrm{osc}^2(f,\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}).\tag{45}
$$

Squaring both sides of (37), (38), (39), (42), (43), (44), applying Young's inequality $2ab\le\zeta a^2+\zeta^{-1}b^2$ for $a,b>0,\zeta>0$, summing all elements $\tau_*\in\mathcal{T}_{h_*}$ and edges/faces $e_*\in\mathcal{E}_{h_*}$, observing the shape regularity of the mesh $\mathcal{T}_{h_*}$ and using (45), we arrive at

$$
\begin{aligned}
&\eta_1^2(v_{h_*},\mathcal{T}_{h_*})\\
&\quad\le(1+\zeta)\eta_1^2(v_h,\mathcal{T}_{h_*})+C_2(1+\zeta^{-1})\Big(\sum_{\tau_*\in\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}} h_{\tau_*}|A_{\tau_*}|^{-1/2}\|f-f_h\|_{\tau_*}\\
&\qquad+\|A^{1/2}(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\mathcal{T}_{h_*}}^2\Big)\\
&\quad\le(1+\zeta)\eta_1^2(v_h,\mathcal{T}_{h_*})\\
&\qquad+C_2(1+\zeta^{-1})\Big(\mu\,\mathrm{osc}^2(f,\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}})+\|A^{1/2}(\nabla_{w,\tau_*}v_{h_*}-\nabla_{w,\tau}v_h)\|_{\mathcal{T}_{h_*}}^2\Big).
\end{aligned}
$$

The constant C_2 depends on the shape regularity of $\mathcal{T}_{h_*}$, coefficient A and the polynomial order l. At last, let $1/\sigma_1=C_2(1+\zeta^{-1})$, we get the desired inequality (34). □

Next, we prove the contraction of the error estimator if the solution does not change.

Lemma 10 *Let $\mathcal{T}_{h_*}$ be a shape regular triangulation which is refined from a shape regular triangulation $\mathcal{T}_h$. Let $u_h\in V_h$ be the discrete solution of* (7). *Then*

$$
\eta_1^2(u_h,\mathcal{T}_{h_*})\le\eta_1^2(u_h,\mathcal{T}_h)-\lambda\eta_1^2(u_h,\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}).
$$

Proof We shall divide the proof into two steps. In the first step, we prove the element-wise contraction if one element is divided into at least two parts, and in the second step, we prove the global version.

Step 1. Suppose $\tau \in \mathcal{R}_{\mathcal{T}_h \to \mathcal{T}_{h_*}}$ is bisected into $\tau_*^1 \in \mathcal{T}_{h_*}$ and $\tau_*^2 \in \mathcal{T}_{h_*}$. We shall prove that

$$\eta_1^2(u_h, \tau_*^1) + \eta_1^2(u_h, \tau_*^2) \le \mu \eta_1^2(u_h, \tau), \tag{46}$$

where $\mu \in (0, 1)$ is given in Lemma 8.

In fact, similar to the proof of (33), we can obtain that the two element-wise terms are reduced, namely,

$$\begin{aligned} &h_{\tau_*^1}^2 |A_{\tau_*^1}|^{-1} \|R_1(u_h, f_h)\|_{\tau_*^1}^2 + h_{\tau_*^2}^2 |A_{\tau_*^2}|^{-1} \|R_1(u_h, f_h)\|_{\tau_*^2}^2 \\ &\quad \le \mu h_\tau^2 |A_\tau^{-1}| \cdot \|R_1(u_h, f_h)\|_\tau^2, \end{aligned} \tag{47}$$

and

$$h_{\tau_*^1}^2 |A_{\tau_*^1}| \cdot \|R_2(u_h)\|_{\tau_*^1}^2 + h_{\tau_*^2}^2 |A_{\tau_*^2}| \cdot \|R_2(u_h)\|_{\tau_*^2}^2 \le \mu h_\tau^2 |A_\tau| \cdot \|R_2(u_h)\|_\tau^2. \tag{48}$$

On the jump residual associated with edges/faces, we note that after $\tau \in \mathcal{T}_h$ is bisected, in $\tau_*^1 \in \mathcal{T}_{h_*}$ and $\tau_*^2 \in \mathcal{T}_{h_*}$, there are three types of faces.

1. For the new edge/face e_* created by the bisection, which is inside the element τ, the function $\nabla_w u_h|_\tau$ is a polynomial and its coefficients are continuous. Therefore $[A\nabla_w u_h \cdot \boldsymbol{n}]_{e_*}$ and $[\gamma_t(\nabla_w u_h)]_{e_*}$ are zero.
2. For the edges/faces divided from τ, the jump values are invariant. But the mesh size is changed.
   For each $e \in \mathcal{E}_h^0$, where $e = \tau_1 \cap \tau_2$ with $\tau_1, \tau_2 \in \mathcal{T}_h$. Let $\tau_{*,i}^1 \in \mathcal{T}_{h_*}$ and $\tau_{*,i}^2 \in \mathcal{T}_{h_*}$ be the children of $\tau_i (i = 1, 2)$, define $e_*^i = \tau_{*,1}^i \cap \tau_{*,2}^i$, then we have $e = e_*^1 \cup e_*^2$. For the first jump term, applying Lemma 1, we obtain $\mathbf{J}_{e_*^i}(A\nabla_w u_h) = 0, i = 1, 2$. For the second jump term,

$$\begin{aligned} &\frac{1}{2} h_{\tau_*} |A_{e_*^1}^{\min}| \cdot \|\mathbf{J}_{e_*^1}(\gamma_t(\nabla_{w,\tau} u_h))\|_{e_*^1}^2 + \frac{1}{2} h_{\tau_*} |A_{e_*^2}^{\min}| \cdot \|\mathbf{J}_{e_*^2}(\gamma_t(\nabla_{w,\tau} u_h))\|_{e_*^2}^2 \\ &= 2^{-1/d} \frac{|A_{e_*^1}^{\min}|}{|A_e^{\min}|} \cdot \frac{1}{2} h_\tau |A_e^{\min}| \cdot \|\mathbf{J}_{e_*^1}(\gamma_t(\nabla_{w,\tau} u_h))\|_{e_*^1}^2 \\ &\quad + 2^{-1/d} \frac{|A_{e_*^2}^{\min}|}{|A_e^{\min}|} \cdot \frac{1}{2} h_\tau |A_e^{\min}| \cdot \|\mathbf{J}_{e_*^2}(\gamma_t(\nabla_{w,\tau} u_h))\|_{e_*^2}^2 \\ &\le \mu \cdot \frac{1}{2} h_\tau |A_e^{\min}| \cdot \|\mathbf{J}_e(\gamma_t(\nabla_w u_h))\|_e^2, \end{aligned} \tag{49}$$

   in the last step, we use the fact $\dfrac{|A_{e_*^i}^{\min}|}{|A_e^{\min}|} = 1$.
   For each $e \in \mathcal{E}_h \cap \partial\Omega$, where $e = \partial\tau$ with $\tau \in \mathcal{T}_h$. Let $\tau_*^1 \in \mathcal{T}_{h_*}$ and $\tau_*^2 \in \mathcal{T}_{h_*}$ be the children of τ, define $e_*^i \in e \cap \tau_*^i (i = 1, 2)$, then $e = e_*^1 \cup e_*^2$. For the first jump term, using the definition of $\mathbf{J}_e(A\nabla_w u_h)$, we obtain $\mathbf{J}_{e_*^i}(A\nabla_w u_h) = 0, i = 1, 2$. For the second jump term, using a similar method to prove (49), we have

$$\begin{aligned} &\frac{1}{2} h_{\tau_*} |A_{e_*^1}^{\min}| \cdot \|\mathbf{J}_{e_*^1}(\gamma_t(\nabla_{w,\tau} u_h))\|_{e_*^1}^2 + \frac{1}{2} h_{\tau_*} |A_{e_*^2}^{\min}| \cdot \|\mathbf{J}_{e_*^2}(\gamma_t(\nabla_{w,\tau} u_h))\|_{e_*^2}^2 \\ &\le \mu \cdot \frac{1}{2} h_\tau |A_e^{\min}| \cdot \|\mathbf{J}_e(\gamma_t(\nabla_w u_h))\|_e^2. \end{aligned} \tag{50}$$

3. For the edges/faces unchanged or inherited from τ, also the jump values are invariant but the mesh size is decreased by $2^{-1/d}$, $d = 2, 3$. The crucial observation is that we use the mesh size h_τ in the jump residual.

Hence, using (47), (48), (49), (50) and the fact $\mu = 2^{-1/d}$, $d = 2, 3$, we get the inequality (46).

Step 2. Notice that $\mathcal{T}_{h_*}\backslash\overline{\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}} = \mathcal{T}_h\backslash\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}$, together with (46), we get

$$
\begin{aligned}
\eta_1^2(u_h, \mathcal{T}_{h_*}) &= \eta_1^2(u_h, \overline{\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}) + \eta_1^2(u_h, \mathcal{T}_{h_*}\backslash\overline{\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}})\\
&= \eta_1^2(u_h, \overline{\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}}) + \eta_1^2(u_h, \mathcal{T}_h\backslash\mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}})\\
&\le \mu\eta_1^2(u_h, \mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}) + \eta_1^2(u_h, \mathcal{T}_h) - \eta_1^2(u_h, \mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}})\\
&\le \eta_1^2(u_h, \mathcal{T}_h) - \lambda\eta_1^2(u_h, \mathcal{R}_{\mathcal{T}_h\to\mathcal{T}_{h_*}}).
\end{aligned}
$$

□

The following lemma summarizes the contraction of $\eta_1^2(\cdot,\cdot)$ by using Lemmas 9 and 10.

Lemma 11 *For any $\zeta > 0$, there exists constant σ_1 depending on the shape regularity of $\mathcal{T}_{k+1}$, the polynomial order l, coefficient A and parameter ζ, such that*

$$
\begin{aligned}
\eta_1^2(u_{k+1}, \mathcal{T}_{k+1}) \le (1+\zeta)\left(\eta_1^2(u_k, \mathcal{T}_k) - \lambda\eta_1^2(u_k, \mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}})\right)&\\
+\frac{1}{\sigma_1}\Big(\mu osc^2(f, \mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}}) + \|A^{1/2}(\nabla_{w,\tau_{k+1}}u_{k+1} - \nabla_{w,\tau_k}u_k)\|^2_{\mathcal{T}_{k+1}}\Big)&,
\end{aligned}
\tag{51}
$$

where $\tau_k \in \mathcal{T}_k$, $\tau_{k+1} \in \mathcal{T}_{k+1}$ and $\tau_{k+1} \subseteq \tau_k$.

Proof Let $\mathcal{T}_h = \mathcal{T}_k$ and $\mathcal{T}_{h_*} = \mathcal{T}_{k+1}$ in Lemmas 9 and 10, we get the desired result (51). □

At the end of this section, we present the contraction of the error estimator by using Lemmas 8 and 11 .

Lemma 12 *There exists $\xi \in (0, 1)$ depending only on the shape regularity of $\mathcal{T}_{k+1}$, the parameters θ, λ and ζ given in the marking strategy (13), Lemmas 8 and 9, respectively. There holds*

$$
\begin{aligned}
\eta^2(u_{k+1}, \mathcal{T}_{k+1}) \le \xi\eta^2(u_k, \mathcal{T}_k) - \zeta osc^2(f, \mathcal{T}_k) + \left(\zeta\lambda + \frac{\mu}{\sigma_1}\right) osc^2(f, \mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}})&\\
+\frac{1}{\sigma_1}\|A^{1/2}(\nabla_{w,\tau_{k+1}}u_{k+1} - \nabla_{w,\tau_k}u_k)\|^2_{\mathcal{T}_{k+1}}&,
\end{aligned}
$$

where μ, σ_1 are defined in Lemmas 8 and 9, respectively; $\tau_k \in \mathcal{T}_k$, $\tau_{k+1} \in \mathcal{T}_{k+1}$ and $\tau_{k+1} \subseteq \tau_k$.

Proof Making use of the definition of the error estimator $\eta^2(\cdot,\cdot)$, Lemma 11 and let $\mathcal{T}_h = \mathcal{T}_k$, $\mathcal{T}_{h_*} = \mathcal{T}_{k+1}$ in Lemma 8, we have

$$
\begin{aligned}
&\eta^2(u_{k+1}, \mathcal{T}_{k+1})\\
&= \eta_1^2(u_{k+1}, \mathcal{T}_{k+1}) + \mathrm{osc}^2(f, \mathcal{T}_{k+1})\\
&\quad\le (1+\zeta)\left(\eta_1^2(u_k, \mathcal{T}_k) - \lambda\eta_1^2(u_k, \mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}})\right) + \mathrm{osc}^2(f, \mathcal{T}_k)\\
&\qquad+\frac{1}{\sigma_1}\Big(\mu\mathrm{osc}^2(f, \mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}}) + \|A^{1/2}(\nabla_{w,\tau_{k+1}}u_{k+1} - \nabla_{w,\tau_k}u_k)\|^2_{\mathcal{T}_{k+1}}\Big)
\end{aligned}
$$

$$
\begin{aligned}
&-\lambda \mathrm{osc}^2(f, \mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}})\\
&= (1+\zeta)\left(\eta^2(u_k,\mathcal{T}_k) - \lambda\eta^2(u_k,\mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}})\right)\\
&\quad -\zeta\left(\mathrm{osc}^2(f,\mathcal{T}_k) - \lambda\mathrm{osc}^2(f,\mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}})\right)\\
&\quad +\frac{1}{\sigma_1}\left(\mu\mathrm{osc}^2(f,\mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}}) + \|A^{1/2}(\nabla_{w,\tau_{k+1}}u_{k+1} - \nabla_{w,\tau_k}u_k)\|^2_{\mathcal{T}_{k+1}}\right)\\
&= (1+\zeta)\left(\eta^2(u_k,\mathcal{T}_k) - \lambda\eta^2(u_k,\mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}})\right) - \zeta\mathrm{osc}^2(f,\mathcal{T}_k)\\
&\quad +\left(\zeta\lambda + \frac{\mu}{\sigma_1}\right)\mathrm{osc}^2(f,\mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}})\\
&\quad +\frac{1}{\sigma_1}\|A^{1/2}(\nabla_{w,\tau_{k+1}}u_{k+1} - \nabla_{w,\tau_k}u_k)\|^2_{\mathcal{T}_{k+1}}.
\end{aligned}
\tag{52}
$$

Applying the marking strategy (13) and choosing ζ small enough such that $\xi := (1+\zeta)(1-\theta\lambda) \in (0,1)$, in conjunction with (52), we obtain

$$
\begin{aligned}
\eta^2(u_{k+1},\mathcal{T}_{k+1}) \le \xi\eta^2(u_k,\mathcal{T}_k) - \zeta\mathrm{osc}^2(f,\mathcal{T}_k) + \left(\zeta\lambda + \frac{\mu}{\sigma_1}\right)\mathrm{osc}^2(f,\mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}})\\
+\frac{1}{\sigma_1}\|A^{1/2}(\nabla_{w,\tau_{k+1}}u_{k+1} - \nabla_{w,\tau_k}u_k)\|^2_{\mathcal{T}_{k+1}},
\end{aligned}
$$

which completes the proof. □

4.4 Convergence of the AWG

In this subsection, we prove the algorithm AWG will terminate in finite steps within a given tolerance. First of all, we shall prove the contraction of summation of the energy error and the scaled error indicator.

Theorem 2 *Given a marking parameter $\theta \in (0,1)$ and initial mesh $\mathcal{T}_0$. Let u be the solution of (1)–(2), $\{\mathcal{T}_k, u_k, \eta(u_k,\mathcal{T}_k)\}_{k\ge 0}$ be a sequence of meshes, finite element solutions and error estimates produced by the AWG. Then there exist constants $\rho \in (0,1), \sigma_1 > 0, \sigma_2 > 0$ depending only on the shape regularity of $\mathcal{T}_0$, the polynomial order l, coefficient A, parameters θ, μ_0 and ϵ, such that if*

$$
0 < \epsilon < \min\left(\frac{\sigma_1(1-\xi)}{C_1}, 1\right),
$$

then

$$
\begin{aligned}
&(1-\epsilon)\|A^{1/2}(\nabla u - \nabla_w u_{k+1})\|^2_{\mathcal{T}_{k+1}} + \sigma_1\eta^2(u_{k+1},\mathcal{T}_{k+1}) + \sigma_2 osc^2(f,\mathcal{T}_{k+1})\\
&\quad \le \rho\Big((1-\epsilon)\|A^{1/2}(\nabla u - \nabla_w u_k)\|^2_{\mathcal{T}_k} + \sigma_1\eta^2(u_k,\mathcal{T}_k) + \sigma_2 osc^2(f,\mathcal{T}_k)\Big),
\end{aligned}
$$

where the constants C_1 and ξ are given by Lemmas 7 and 12 , respectively.

Remark 5 Notice that the data oscillation $\mathrm{osc}^2(f,\cdot)$ is one part of the error indicator $\eta^2(\cdot,\cdot)$. If we want to get rid of the term $\sigma_2\mathrm{osc}^2(f,\cdot)$, we have to add an extra marking for the data oscillation, see [5].

Proof By adding $\sigma_1\eta^2(u_{k+1},\mathcal{T}_{k+1})$ to both sides of (28), then applying Lemma 12 , we have

$$
(1-\epsilon)\|A^{1/2}(\nabla u - \nabla_w u_{k+1})\|^2_{\mathcal{T}_{k+1}} + \sigma_1\eta^2(u_{k+1},\mathcal{T}_{k+1})
$$

$$
\begin{aligned}
&\le \|A^{1/2}(\nabla u - \nabla_w u_k)\|^2_{\mathcal{T}_k} - \|A^{1/2}(\nabla_{w,\tau_{k+1}} u_{k+1} - \nabla_{w,\tau_k} u_k)\|^2_{\mathcal{T}_{k+1}} \\
&\quad + \frac{C_0}{\epsilon}\mathrm{osc}(f, \mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}}) + \sigma_1\eta^2(u_{k+1}, \mathcal{T}_{k+1}) \\
&\le \|A^{1/2}(\nabla u - \nabla_w u_k)\|^2_{\mathcal{T}_k} - \|A^{1/2}(\nabla_{w,\tau_{k+1}} u_{k+1} - \nabla_{w,\tau_k} u_k)\|^2_{\mathcal{T}_{k+1}} \\
&\quad + \frac{C_0}{\epsilon}\mathrm{osc}^2(f, \mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}}) + \sigma_1\xi\eta^2(u_k, \mathcal{T}_k) - \sigma_1\zeta\,\mathrm{osc}^2(f, \mathcal{T}_k) \\
&\quad + (\zeta\lambda\sigma_1 + \mu)\mathrm{osc}^2(f, \mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}}) + \|A^{1/2}(\nabla_{w,\tau_{k+1}} u_{k+1} - \nabla_{w,\tau_k} u_k)\|^2_{\mathcal{T}_{k+1}} \\
&\le \|A^{1/2}(\nabla u - \nabla_w u_k)\|^2_{\mathcal{T}_k} + \sigma_1\xi\eta^2(u_k, \mathcal{T}_k) \\
&\quad + \left(\frac{C_0}{\epsilon} + \zeta\lambda\sigma_1 + \mu\right)\mathrm{osc}^2(f, \mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}}) - \sigma_1\zeta\,\mathrm{osc}^2(f, \mathcal{T}_k), \qquad (53)
\end{aligned}
$$

for any constant $\epsilon \in (0, 1)$. Suppose $\sigma_2 > 0$, which will be determined later. By adding $\sigma_2\mathrm{osc}^2(f, \mathcal{T}_{k+1})$ in the both sides of (53) and let $\mathcal{T}_h = \mathcal{T}_k, \mathcal{T}_{h_*} = \mathcal{T}_{k+1}$ in Lemma 8, we obtain

$$
\begin{aligned}
&(1-\epsilon)\|A^{1/2}(\nabla u - \nabla_w u_{k+1})\|^2_{\mathcal{T}_{k+1}} + \sigma_1\eta^2(u_{k+1}, \mathcal{T}_{k+1}) + \sigma_2\mathrm{osc}^2(f, \mathcal{T}_{k+1}) \\
&\quad \le \|A^{1/2}(\nabla u - \nabla_w u_k)\|^2_{\mathcal{T}_k} + \sigma_1\xi\eta^2(u_k, \mathcal{T}_k) + (\sigma_2 - \sigma_1\zeta)\mathrm{osc}^2(f, \mathcal{T}_k) \\
&\quad + \left(\frac{C_0}{\epsilon} + \mu - (\sigma_2 - \zeta\sigma_1)\lambda\right)\mathrm{osc}^2(f, \mathcal{R}_{\mathcal{T}_k\to\mathcal{T}_{k+1}}). \qquad (54)
\end{aligned}
$$

The above inequality (54) along with a sufficiently large σ_2 satisfying

$$
\frac{C_0}{\epsilon} + \mu - (\sigma_2 - \zeta\sigma_1)\lambda \le 0, \qquad (55)
$$

and some $\rho_1 \in (0, 1)$ to be determined later implies

$$
\begin{aligned}
&(1-\epsilon)\|A^{1/2}(\nabla u - \nabla_w u_{k+1})\|^2_{\mathcal{T}_{k+1}} + \sigma_1\eta^2(u_{k+1}, \mathcal{T}_{k+1}) + \sigma_2\mathrm{osc}^2(f, \mathcal{T}_{k+1}) \\
&\quad \le \|A^{1/2}(\nabla u - \nabla_w u_k)\|^2_{\mathcal{T}_k} + \sigma_1\xi\eta^2(u_k, \mathcal{T}_k) + (\sigma_2 - \sigma_1\zeta)\mathrm{osc}^2(f, \mathcal{T}_k) \\
&\quad \le \rho_1(1-\epsilon)\|A^{1/2}(\nabla u - \nabla_w u_k)\|^2_{\mathcal{T}_k} + (1 - \rho_1(1-\epsilon))\,\|A^{1/2}(\nabla u - \nabla_w u_k)\|^2 \\
&\quad + \sigma_1\xi\eta^2(u_k, \mathcal{T}_k) + (\sigma_2 - \sigma_1\zeta)\mathrm{osc}^2(f, \mathcal{T}_k). \qquad (56)
\end{aligned}
$$

The upper bound (30) together with (56), yields

$$
\begin{aligned}
&(1-\epsilon)\|A^{1/2}(\nabla u - \nabla_w u_{k+1})\|^2_{\mathcal{T}_{k+1}} + \sigma_1\eta^2(u_{k+1}, \mathcal{T}_{k+1}) + \sigma_2\mathrm{osc}^2(f, \mathcal{T}_{k+1}) \\
&\quad \le \rho_1(1-\epsilon)\|A^{1/2}(\nabla u - \nabla_w u_k)\|^2_{\mathcal{T}_k} + \Big(C_1 - C_1\rho_1(1-\epsilon) + \sigma_1\xi\Big)\eta^2(u_k, \mathcal{T}_k) \\
&\quad + (\sigma_2 - \sigma_1\zeta)\mathrm{osc}^2(f, \mathcal{T}_k), \qquad (57)
\end{aligned}
$$

according to

$$
\rho_1\sigma_1 = C_1 - C_1\rho_1(1-\epsilon) + \sigma_1\xi,
$$

choose

$$
\rho_1 = \frac{C_1 + \sigma_1\xi}{C_1 + \sigma_1 - C_1\epsilon},
$$

the requirement $0 < \epsilon < \min\left(\frac{\sigma_1(1-\xi)}{C_1}, 1\right)$ with $\xi \in (0, 1)$ leads to $\rho_1 \in (0, 1)$. By (55), we obtain $\sigma_2 - \sigma_1\zeta > 0$. Then let $\rho_2 = \frac{\sigma_2 - \sigma_1\zeta}{\sigma_2}$, we get $\rho_2 \in (0, 1)$ and

$$
\begin{aligned}
&(1-\epsilon)\|A^{1/2}(\nabla u-\nabla_w u_{k+1})\|^2_{\mathcal{T}_{k+1}}+\sigma_1\eta^2(u_{k+1},\mathcal{T}_{k+1})+\sigma_2\mathrm{osc}^2(f,\mathcal{T}_{k+1})\\
&\quad\le \rho_1(1-\epsilon)\|A^{1/2}(\nabla u-\nabla_w u_k)\|^2_{\mathcal{T}_k}+\rho_1\sigma_1\eta^2(u_k,\mathcal{T}_k)+(\sigma_2-\sigma_1\zeta)\mathrm{osc}^2(f,\mathcal{T}_k)\\
&\quad\le \rho_1(1-\epsilon)\|A^{1/2}(\nabla u-\nabla_w u_k)\|^2_{\mathcal{T}_k}+\rho_1\sigma_1\eta^2(u_k,\mathcal{T}_k)+\rho_2\sigma_2\mathrm{osc}^2(f,\mathcal{T}_k).
\end{aligned}
$$

We complete the proof by setting $\rho=\max\{\rho_1,\rho_2\}\in(0,1)$. □

By recursion, we get the decay of the error plus the estimator.

Corollary 1 *Under the hypotheses of Theorem 2, then we have*

$$
(1-\epsilon)\|A^{1/2}(\nabla u-\nabla_w u_k)\|^2_{\mathcal{T}_k}+\sigma_1\eta^2(u_k,\mathcal{T}_k)+\sigma_2 osc^2(f,\mathcal{T}_k)\le \hat{C}_0\rho^k,
$$

where the constant $\epsilon,\sigma_1,\sigma_2,\rho$ are given in Theorem 2, and $\hat{C}_0=(1-\epsilon)\|A^{1/2}(\nabla u-\nabla_w u_0)\|^2_{\mathcal{T}_0}+\sigma_1\eta^2(u_0,\mathcal{T}_0)+\sigma_2 osc^2(f,\mathcal{T}_0)$. Thus the algorithm AWG will terminate in finite steps.

5 Numerical Experiments

In this section, we test some experiments to show the performance of the adaptive algorithm AWG. We carry out these numerical experiments by using the MATLAB software package iFEM [4]. We choose the lowest order WG method and estimate the energy error $\|A^{1/2}(\nabla u-\nabla_w u_k)\|_{\mathcal{T}_k}$ in the following numerical experiments.

Example 1 In this example, we test 'L-shape' problem in two dimensions. We choose an L-shape domain $\Omega=(-1,1)^2/[0,1)^2$ and the coefficient $A=\mathbf{I}$. For the source $f=0$, the exact solution is $u=r^{2/3}\sin(\frac{2}{3}\theta)$ in polar coordinates. The left of Fig. 1 shows the initial mesh $\mathcal{T}_0$, and the right of Fig. 1 shows an adaptively refined mesh with marking parameter $\theta=0.5$ after $k=14$ iterative steps, which indicates the mesh is locally refined in a small vicinity of the edge singularity.

Denote $\#\mathcal{T}_k$ the number of elements and u_k the corresponding weak finite element solution associated to the mesh $\mathcal{T}_k$. The left of Fig. 2 shows the curves of $\log\#\mathcal{T}_k-\log\|A^{1/2}(\nabla u-\nabla_w u_k)\|_{\mathcal{T}_k}$ for marking parameters $\theta=0.1, 0.3, 0.5$ which indicates the convergence and the quasi-optimality of the adaptive algorithm AWG of the energy error $\|A^{1/2}(\nabla u-\nabla_w u_k)\|_{\mathcal{T}_k}$, namely

$$
\|A^{1/2}(\nabla u-\nabla_w u_k)\|_{\mathcal{T}_k}\lesssim(\#\mathcal{T}_k)^{-1/2}.
$$

And the right of Fig. 2 plots the performances of $\|A^{1/2}(\nabla u-\nabla_w u_k)\|_{\mathcal{T}_k}$ and $\eta(u_k,\mathcal{T}_k)$ which shows that the energy error $\|A^{1/2}(\nabla u-\nabla_w u_k)\|_{\mathcal{T}_k}$ can be controlled by the error estimator $\eta(u_k,\mathcal{T}_k)$ and the optimal rates of the energy error and the corresponding error estimators are approximate.

Example 2 In this example, we employ the Kellogg problem introduced in [10]. We choose a domain $\Omega=(-1,1)^2$ and for $f=0$, the exact solution in polar coordinates is $u(r,\theta)=r^\gamma\mu(\theta)$ where

$$
\mu(\theta)=\begin{cases}
\cos\left(\left(\frac{\pi}{2}-\sigma\right)\gamma\right)\cos\left(\left(\theta-\frac{\pi}{2}+\rho\right)\gamma\right) & \text{if } 0\le\theta\le\frac{\pi}{2},\\
\cos(\rho\gamma)\cos((\theta-\pi+\sigma)\gamma) & \text{if } \frac{\pi}{2}\le\theta\le\pi,\\
\cos(\sigma\gamma)\cos((\theta-\pi-\rho)\gamma) & \text{if } \pi\le\theta\le\frac{3\pi}{2},\\
\cos\left(\left(\frac{\pi}{2}-\rho\right)\gamma\right)\cos\left(\left(\theta-\frac{3\pi}{2}-\sigma\right)\gamma\right) & \text{if } \frac{3\pi}{2}\le\theta\le 2\pi,
\end{cases}
$$

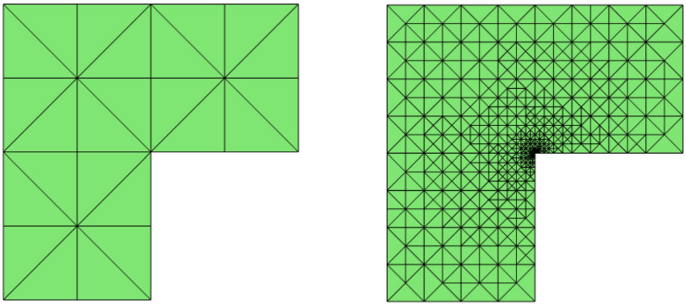

Fig. 1 The initial mesh $\mathcal{T}_0$ (left); An adaptively refined mesh after 14 adaptive iterations with marking parameter $\theta = 0.5$ (right)

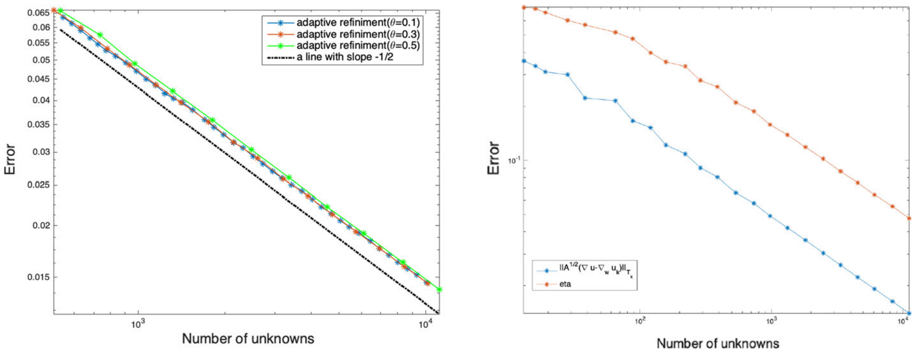

Fig. 2 Quasi optimality of the adaptive algorithm AWG of the error $\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ with different marking parameters θ(left); the performances of $\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ and $\eta(u_k, \mathcal{T}_k)$ for Example 1 with $\theta = 0.5$ (right)

the coefficient matrix A is piecewise constant: $A = 161.44764\,\mathbf{I}$ in the first and third quadrants and $A = \mathbf{I}$ in the second and fourth quadrants and the constants $\gamma = 0.1, \sigma = -14.92256, \rho = \pi/4$. Indeed, the exact solution $u \in H^{1+\gamma}(\Omega)$. The left of Fig. 3 shows the initial mesh $\mathcal{T}_0$, and the right of Fig. 3 shows an adaptively refined mesh with marking parameter $\theta = 0.5$ after $k = 130$ iterative steps. We can also see that the mesh is locally refined in a small vicinity of the edge singularity.

The left of Fig. 4 shows the curves of $\log \#\mathcal{T}_k - \log \|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ for Kellogg problem with different marking parameters $\theta = 0.1, 0.3, 0.5$ which also indicates the convergence and the next quasi-optimality of adaptive algorithm AWG, i.e.

$$\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k} \lesssim (\#\mathcal{T}_k)^{-1/2}.$$

And the right of Fig. 4 plots the performances of $\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ and $\eta(u_k, \mathcal{T}_k)$ which shows that the energy error $\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ can be controlled by the error estimator $\eta(u_k, \mathcal{T}_k)$ and the optimal rates of the energy error and the corresponding error estimators are approximate.

Example 3 In this example, we test 'L-shape' problem in three dimensions. We choose an L-shape domain $\Omega = (-1, 1)^3/[0, 1) \times [0, 1) \times (-1, 1)$. We get an initial mesh $\mathcal{T}_0$ by partitioning the given domain Ω into four subintervals in x-, y- and z-axes and then dividing

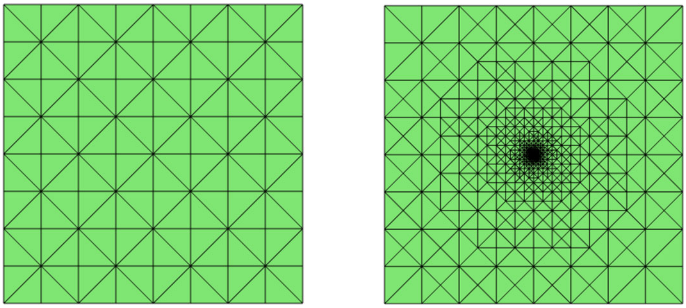

Fig. 3 The initial mesh $\mathcal{T}_0$ (left); an adaptively refined mesh for Kellogg problem with marking parameter $\theta = 0.5$ after $k = 130$ adaptive iterations (right)

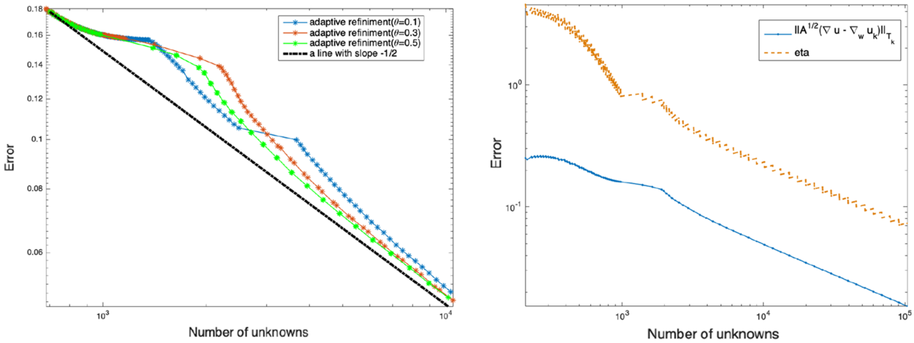

Fig. 4 Quasi optimality of the adaptive algorithm AWG of the error $\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ with different marking parameters θ(left); the performances of $\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ and $\eta(u_k, \mathcal{T}_k)$ for Example 2 (right)

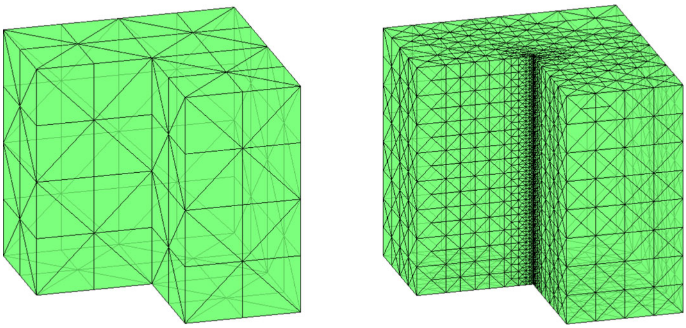

Fig. 5 The initial mesh $\mathcal{T}_0$ (left); an adaptively refined mesh for L-shape problem in three dimensions with marking parameter $\theta = 0.5$ after $k = 17$ adaptive iterations (right)

every cube into 6 tetrahedrons. We set $A = \mathbf{I}$ and the source $f = 0$ such that the exact solution in the cylindrical coordinate is $u = r^{\frac{2}{3}} \sin\left(\frac{2}{3}\theta\right)$. The left of Fig. 5 shows the initial mesh $\mathcal{T}_0$, and the right of Fig. 5 shows an adaptively refined mesh with marking parameter $\theta = 0.5$ after $k = 17$ iterative steps which also indicates the mesh is locally refined.

The left of Fig. 6 plots the curves of $\log \#\mathcal{T}_k - \log \|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ for $\theta = 0.1, 0.3, 0.5$ which indicates the convergence and the next quasi-optimality of adaptive algo-

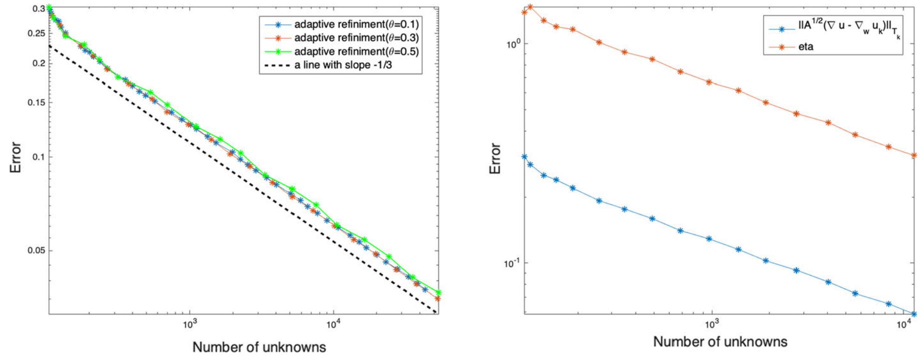

Fig. 6 Quasi optimality of the adaptive algorithm AWG of the error $\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ with different marking parameters θ(left); the performances of $\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ and $\eta(u_k, \mathcal{T}_k)$ for Example 3 (right)

rithm AWG of the energy error, i.e.

$$\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k} \lesssim (\#\mathcal{T}_k)^{-1/3}.$$

And the right of Fig. 6 plots the performances of $\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ and $\eta(u_k, \mathcal{T}_k)$ for Example 3 which shows that the energy error $\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k}$ can be controlled by the error estimator $\eta(u_k, \mathcal{T}_k)$ and the optimal rates of the energy error and the corresponding error estimators are approximate.

From above numerical examples, we know that the AWG method introduced in Sect. 3 is convergent and the numerical examples also indicate next quasi-optimality

$$\|A^{1/2}(\nabla u - \nabla_w u_k)\|_{\mathcal{T}_k} \lesssim (\#\mathcal{T}_k)^{-1/d}, d = 2, 3.$$

Acknowledgements The authors would like to thank Professor Long Chen, University of California at Irvine, for providing many constructive suggestions.

References

1. Binev, P., Dahmen, W., DeVore, R.A.: Adaptive finite element methods with convergence rates. Numer. Math. **97**(2), 219–268 (2004)
2. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer, New York (1991)
3. Cascon, J.M., Kreuzer, C., Nochetto, R.H., Siebert, K.G.: Quasi-optimal convergence rate for an adaptive finite element method. SIAM J. Numer. Anal. **46**(5), 2524–2550 (2008)
4. Chen, L.: iFEM: An Integrated Finite Element Methods Package in MATLAB. University of California, Irvine (2009)
5. Chen, L., Holst, M., Xu, J.: Convergence and optimality of adaptive mixed finite element methods. Math. Comput. **78**(265), 35–53 (2009)
6. Chen, L., Wang, J., Ye, X.: A posteriori error estimates for weak Galerkin finite element methods for second order elliptic problems. J. Sci. Comput. **59**(2), 496–511 (2014)
7. Du, Y., Zhang, Z.: A numerical analysis of the weak Galerkin method for the Helmholtz equation with high wave number. Commun. Comput. Phys. **22**(1), 133–156 (2017)
8. Dörfler, W.: A convergent adaptive algorithm for Poisson's equation. SIAM J. Numer. Anal. **33**(3), 1106–1124 (1996)
9. Huang, J., Xu, Y.: Convergence and complexity of arbitrary order adaptive mixed element methods for the Poisson equation. Sci. China Math. **55**(5), 1083–1098 (2012)
10. Kellogg, R.B.: On the Poisson equation with intersecting interfaces. Appl. Anal. **4**(2), 101–129 (1974)

11. Li, H.G., Mu, L., Ye, X.: A posteriori error estimates for the weak Galerkin finite element methods on polytopal meshes. Commun. Comput. Phys. **26**(2), 558–578 (2019)
12. Lin, G., Liu, J., Mu, L., Ye, X.: Weak Galerkin finite element methods for Darcy flow: anisotropy and heterogeneity. J. Comput. Phys. **276**, 422–437 (2014)
13. Mitchell, W.F.: A comparison of adaptive refinement techniques for elliptic problems. ACM Trans. Math. Softw. **15**(4), 326–347 (1989)
14. Morin, P., Nochetto, R.H., Siebert, K.G.: Data oscillation and convergence of adaptive FEM. SIAM J. Numer. Anal. **38**(2), 466–488 (2000)
15. Morin, P., Nochetto, R.H., Siebert, K.G.: Convergence of adaptive finite element methods. SIAM Rev. **44**(4), 631–658 (2002)
16. Mu, L.: Weak Galerkin based a posteriori error estimates for second order elliptic interface problems on polygonal meshes. J. Comput. Appl. Math. **361**, 413–425 (2019)
17. Mu, L., Wang, J., Wei, G., Zhao, S.: Weak Galerkin methods for second order elliptic interface problems. J. Comput. Phys. **250**, 106–125 (2013)
18. Mu, L., Wang, J., Ye, X.: Weak Galerkin finite element methods for the biharmonic equation on polytopal meshes. Numer. Methods Part. Differ. Equ. **30**(3), 1003–1029 (2014)
19. Mu, L., Wang, J., Ye, X.: A new weak Galerkin finite element method for the Helmholtz equation. IMA J. Numer. Anal. **35**(3), 1228–1255 (2015)
20. Mu, L., Wang, J., Ye, X.: A weak Galerkin finite element method with polynomial reduction. J. Comp. Appl. Math. **285**, 45–58 (2015)
21. Mu, L., Wang, J., Ye, X.: Weak Galerkin finite element methods on polytopal meshes. Int. J. Numer. Anal. Mod. **12**(1), 31–53 (2015)
22. Mu, L., Wang, J., Ye, X., Zhang, S.: A C^0-weak Galerkin finite element method for the biharmonic equation. J. Sci. Comput. **59**(2), 473–495 (2014)
23. Mu, L., Wang, J., Ye, X., Zhao, S.: A numerical study on the weak Galerkin method for the Helmholtz equation. Commun. Comput. Phys. **15**(5), 1461–1479 (2014)
24. Nochetto, R.H., Siebert, K.G., Veeser, A.: Theory of adaptive finite element methods: an introduction. In: Devore, R.A., Kunoth, A. (eds.) Multiscale, Nonlinear and Adaptive Approximation, pp. 409–542. Springer, Berlin (2009)
25. Stevenson, R.: The completion of locally refined simplicial partitions created by bisection. Math. Comp. **77**(261), 227–241 (2008)
26. Wang, C., Wang, J.: An efficient numerical scheme for the biharmonic equation by weak Galerkin finite element methods on polygonal or polyhedral meshes. Comput. Math. Appl. **68**(12), 2314–2330 (2013)
27. Wang, J., Ye, X.: A weak Galerkin finite element method for second-order elliptic problems. J. Comput. Appl. Math. **241**(0), 103–115 (2013)
28. Wang, J., Ye, X.: A weak Galerkin mixed finite element method for second-order elliptic problems. Math. Comput. **83**(289), 2101–2126 (2014)
29. Zhang, T., Chen, Y.: A posteriori error analysis for the weak Galerkin method for solving elliptic problems. Int. J. Comput. Methods **15**(8), 1850075 (2018)
30. Zhang, J., Zhang, K., Li, J., Wang, X.: A weak Galerkin finite element method for the Navier–Stokes equations. Commun. Comput. Phys. **23**, 706–746 (2018)
31. Zhang, T., Lin, T.: A posteriori error estimate for a modified weak Galerkin method solving elliptic problems. Numer. Methods Part. D. E **33**(1), 381–398 (2017)
32. Zheng, X., Xie, X.: A posteriori error estimator for a weak Galerkin finite element solution of the Stokes problem. E. Asian. J. Appl. Math. **7**(3), 508–529 (2017)
33. Zhong, L., Chen, L., Shu, S., Wittum, G., Xu, J.: Convergence and optimality of adaptive edge finite element methods for time-harmonic Maxwell equations. Math. Comput. **81**(278), 623–642 (2012)
34. Zhong, L., Shu, S., Chen, L., Xu, J.: Convergence of adaptive edge finite element methods for H(curl)-elliptic problems. Numer. Linear Algebra Appl. **17**(2–3), 415–432 (2010)