INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 13, Number 4, Pages 525–544

WEAK GALERKIN FINITE ELEMENT METHOD FOR SECOND ORDER PARABOLIC EQUATIONS

HONGQIN ZHANG, YONGKUI ZOU, YINGXIANG XU, QILONG ZHAI, AND HUA YUE

Abstract. We apply in this paper the weak Galerkin method to the second order parabolic differential equations based on a discrete weak gradient operator. We establish both the continuous time and the discrete time weak Galerkin finite element schemes, which allow using the totally discrete functions in approximation space and the finite element partitions of arbitrary polygons with certain shape regularity. We show as well that the continuous time weak Galerkin finite element method preserves the energy conservation law. The optimal convergence order estimates in both H^1 and L^2 norms are obtained. Numerical experiments are performed to confirm the theoretical results.

Key words. Weak Galerkin finite element methods, discrete gradient, parabolic equations.

1. Introduction

The weak Galerkin (WG for short) finite element method refers to the finite element techniques for partial differential equations where the differential operators (e.g., gradient, divergence, curl, Laplacian) are approximated by weak forms. In [17], a WG method was introduced and analyzed for second order elliptic equations based on a discrete weak gradient arising from RT element [13] or BDM element [1]. When using the RT or the BDM element, the WG finite element method requires the classical finite element partitions such as triangles for 2-dimensional elements and tetrahedra for 3-dimensional elements, which limits the use of the newly-developed method. This problem was first dealt with in [16] where for the second order elliptic equation the authors using the stabilization for the flux variable established a WG mixed finite element method that is applicable for general finite element partitions consisting of shape regular polytopes (e.g., polygons in 2D and polyhedra in 3D). The idea of stabilization has been applied to the Galerkin finite element method for the second order elliptic equation, see [10]. At present, the WG method has attracted many attentions and successfully found its way to many applications, for example, see [12] for the Helmholtz equation and [11] for elliptic interface problems.

As far as the parabolic problem is concerned, there are surely many classical numerical methods applicable. For example, see [5, 3] for the classical finite element methods, [6, 9] for the discontinuous Galerkin finite element methods, [19, 2, 4, 14, 8] for the finite volume methods. We note here that the WG method is also applicable for such kind of time dependent problems. In [7], the authors discussed the WG finite element method for the parabolic equations, where again the definition of the discrete weak gradient operator proposed in [17] was applied. Comparing to the existing methods, the WG finite element method allows using discontinuous function space as the approximation space and thus it is not necessarily to require the underlying solutions to be smooth enough as in the usual sense. This property makes the WG finite element method more flexible in applications.

Received by the editors February 15, 2015 and, in revised form, September 30, 2015.
2000 *Mathematics Subject Classification.* 65M60, 65M12.

The goal of this paper is, different from the technique applied in [7], to apply the WG finite element method to the parabolic partial differential equations, by using the idea of stabilization. We consider the following initial-boundary value problem for the second order parabolic equations

$$u_t - \nabla \cdot (a\nabla u) = f, \quad \text{for} \quad x \in \Omega, t \in J, \tag{1}$$

$$u = 0, \quad \text{for} \quad x \in \partial\Omega, t \in J, \tag{2}$$

$$u(\cdot, 0) = \psi, \quad \text{for} \quad x \in \Omega, \tag{3}$$

where Ω is a polygonal domain in $\mathbb{R}^2$ with Lipschitz-continuous boundary, $J = (0, T]$ with $T > 0$, $a = a(\cdot)_{2\times 2} \in [L^\infty(\Omega)]^{2\times 2}$ is a symmetric matrix-valued function. Assume that the matrix function $a(\cdot)$ satisfies the following property: there exist two constants $0 < \bar{\alpha}_1 < \bar{\alpha}_2$ such that

$$\bar{\alpha}_1 \xi^T \xi \leq \xi^T a \xi \leq \bar{\alpha}_2 \xi^T \xi, \qquad \forall \xi \in \mathbb{R}^2. \tag{4}$$

The standard variational form for (1)-(3) seeks $u(\cdot, t) \in H_0^1(\Omega)$ such that

$$\begin{aligned} (u_t, v) + (a\nabla u, \nabla v) = (f, v), \quad \forall v \in H_0^1(\Omega), t \in J. \\ u(\cdot, 0) = \psi, \end{aligned} \tag{5}$$

where $(\cdot, \cdot)$ denotes the L^2-inner product.

In this paper, we, based on the definition of a discrete weak gradient operator proposed in [10], derive the continuous time and the discrete time WG finite element methods for problems (1)-(3) by taking into account the idea of stabilization. The new obtained methods allow the application of the more general finite element partitions satisfying certain shape regular conditions, and allow as well using totally discontinuous function space as the approximation space. In addition, the continuous time WG finite element method preserves the energy conservation law.

The rest of this paper is organized as follows. In Section 2, we introduce the notations and establish the continuous time and the discrete time WG finite element schemes for problems (1)-(3). We then prove the energy conservation law of the continuous time WG approximation in Section 3. In Section 4, the optimal error estimates in both H^1 norm and L^2 norm are proved. Finally, we present the numerical example to verify the theory.

2. The WG approximation

To introduce the WG finite element method for the parabolic equations (1)-(3), we need to consider first the weak gradient and the discrete weak gradient. The gradient ∇ is a principle differential operator involved in the variational form. Thus, it is critical to define and understand discrete weak gradients for the corresponding numerical methods. Following the idea in [17, 10], the discrete weak gradient is given by approximating the weak functions with piecewise polynomial functions, as shown in what follows.

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

$$W(K) := \{v = \{v_0, v_b\} : v_0 \in L^2(K), v_b \in H^{1/2}(\partial K)\}. \tag{6}$$

Let $(\cdot,\cdot)_K$ and $\langle\cdot,\cdot\rangle_{\partial K}$ represent the inner product in $L^2(K)$ and $H^{\frac{1}{2}}(\partial K)$, respectively,

$$(v,w)_K = \int_K vwdx, \quad \forall v,w \in L^2(K),$$
$$\langle v,w\rangle_{\partial K} = \int_{\partial K} vwds, \quad \forall v,w \in H^{\frac{1}{2}}(\partial K).$$

Define a space

$$H(\mathrm{div},K) = \{\mathbf{v} : \mathbf{v} \in [L^2(K)]^2, \nabla\cdot\mathbf{v} \in L^2(K)\}. \tag{7}$$

The standard gradient operator can be equivalently formulated as following

$$\begin{aligned} \nabla : H^1(K) &\to (H(\mathrm{div},K))^*, \\ u &\mapsto \nabla u, \end{aligned} \tag{8}$$

where the gradient ∇ is interpreted as a linear functional on the space $H(\mathrm{div},K)$ and satisfies

$$(\nabla u,\mathbf{q}) = -(u,\nabla\cdot\mathbf{q})_K + \langle u,\mathbf{q}\cdot\mathbf{n}\rangle_{\partial K}, \quad \forall \mathbf{q} \in H(\mathrm{div},K), \tag{9}$$

where $\mathbf{n}$ is the unit outward normal vector to ∂K.

The weak gradient operator ∇_ω is defined by replacing the space $H^1(K)$ by $W(K)$ and applying the expression in (9), cf. [17, 10].

Definition 2.1. *Define a weak gradient operator ∇_ω by*

$$\begin{aligned} \nabla_\omega : W(K) &\to (H(\mathrm{div},K))^*, \\ v &\mapsto \nabla_\omega v, \end{aligned}$$

where $\nabla_\omega v$ is a functional on the space $H(\mathrm{div},K)$ which is determined by

$$(\nabla_\omega v,\mathbf{q}) := -(v_0,\nabla\cdot\mathbf{q})_K + \langle v_b,\mathbf{q}\cdot\mathbf{n}\rangle_{\partial K}, \quad \forall \mathbf{q} \in H(\mathrm{div},K). \tag{10}$$

The Sobolev space $H^1(K)$ can be embedded into the weak space $W(K)$ by an inclusion map $i_W : H^1(K) \to W(K)$ defined as follows

$$i_W(\phi) = \{\phi|_K, \phi|_{\partial K}\}, \qquad \forall \phi \in H^1(K).$$

With the help of the inclusion map i_W, the Sobolev space $H^1(K)$ can be viewed as a subspace of $W(K)$ by identifying each $\phi \in H^1(K)$ with $i_W(\phi)$. Analogously, a weak function $v = \{v_0, v_b\} \in W(K)$ is said to be in $H^1(K)$ if it can be identified with a function $\phi \in H^1(K)$ through the above inclusion map. It is not hard to see that weak gradient is identical with the strong gradient (i.e.$\nabla_\omega\phi = \nabla\phi$) for each smooth function $\phi \in H^1(K)$.

Recall that the discrete weak gradient operator is defined by approximating ∇_ω in a polynomial subspace. More precisely, for any nonnegative integer r, denote by $P_r(K)$ the set of polynomials on K with degree no more than r. The discrete weak gradient operator, denoted by ∇_d, is defined as: $\nabla_d v$ is the unique polynomial in $[P_r(K)]^2$ and satisfies the following equation

$$(\nabla_d v,\mathbf{q}) = -(v_0,\nabla\cdot\mathbf{q})_K + \langle v_b,\mathbf{q}\cdot\mathbf{n}\rangle_{\partial K}, \qquad \forall \mathbf{q} \in [P_r(K)]^2. \tag{11}$$

In this paper, we shall allow a greater flexibility in the definition and computation of the discrete weak gradient operator $\nabla_d v \in [P_r(K)]^2$ by using the usual polynomial space $[P_r(K)]^2$. This will result in a new class of WG finite element schemes for parabolic equations with remarkable properties to be detailed in the following sections.

2.2. The WG method. In this section, we design a continuous time and a discretime WG finite element schemes for the initial-boundary value problems (1)-(3).

Let $\mathcal{T}_h$ be a family of partitions of the domain Ω, where h is the partition diameter. We assume throughout the paper that $\mathcal{T}_h$ is *shape regular*, namely, satisfying the shape regularity assumptions A1-A4 in [16]. Denote by T^0 its interior and by ∂T its boundary for any $T \in \mathcal{T}_h$, respectively. For each $T \in \mathcal{T}_h$, let $P_r(T^0)$ and $P_r(\partial T)$ be the sets of polynomials on T^0 and on ∂T, respectively, with degree no more than r. Denoted by V the weak function space on $\mathcal{T}_h$ given by

$$V := \{v = \{v_0, v_b\} : \{v_0, v_b\}|_T \in W(T), T \in \mathcal{T}_h\},$$

where $\{v_0, v_b\}|_T = \{v_0|_T, v_b|_{\partial T}\}$ is the restriction of v on the element T. For any given integer $r \geq 1$, let $W_r(T)$ be the weak finite element space consisting of polynomials of degree no more than r in T^0 and piecewise polynomials of degree no more than r on ∂T, i.e.,

$$W_r(T) := \{v = \{v_0, v_b\} : v_0|_T \in P_r(T^0), v_b|_{\overline{e}} \in P_r(\overline{e}), \overline{e} \subset \partial T\}.$$

We restrict the domain of the weak gradient operator ∇_d on the finite dimensional polynomial space $W_r(T)$ and obtain a linear operator from $W_r(T)$ to $G_{r-1}(T) := [P_{r-1}(T)]^2$, which is determined by

$$(\nabla_d v, \mathbf{q})_T = -(v_0, \nabla \cdot \mathbf{q})_T + \langle v_b, \mathbf{q} \cdot \mathbf{n} \rangle_{\partial T}, \quad \forall v \in W_r(T), \mathbf{q} \in G_{r-1}(T). \tag{12}$$

Now, we are ready to introduce the WG finite element method for approximating equations (1)-(3). Define WG finite element spaces

$$V_h := \{v = \{v_0, v_b\} : \{v_0, v_b\}|_T \in W_r(T), T \in \mathcal{T}_h\},$$

and

$$V_h^0 := \{v : v \in V_h, v_b = 0 \quad \text{on} \quad \partial\Omega\}. \tag{13}$$

Define two bilinear forms on V_h: for any $v, w \in V_h$

$$\bar{a}(v, w) = \sum_{T \in \mathcal{T}_h} \int_T (a \nabla_d v) \cdot \nabla_d w dT, \tag{14}$$

$$s(v, w) = \sum_{T \in \mathcal{T}_h} h_T^{-1} \langle v_0 - v_b, w_0 - w_b \rangle_{\partial T}. \tag{15}$$

Denote by $a_s(\cdot, \cdot)$ a stabilization of $\bar{a}(\cdot, \cdot)$ given by

$$a_s(v, w) = \bar{a}(v, w) + s(v, w). \tag{16}$$

In [10, Lemma7.2], it is proved that there exist two constants $\alpha, \beta > 0$ such that for $u, v \in V_h$

$$|a_s(u, v)| \leq \beta \,|||\, u \,|||\, \cdot \,|||\, v \,|||\, .$$
$$\alpha \,|||\, u|||^2 \leq a_s(u, u),$$

where

$$|||v|||^2 = \sum_{T \in \mathcal{T}_h} (\nabla_d v, \nabla_d v)_T + \sum_{T \in \mathcal{T}_h} h_T^{-1} (v_0 - v_b, v_0 - v_b)_{\partial T}, \quad v \in V_h.$$

We define two projections on each triangle: one is $Q_h u = \{Q_0 u, Q_b u\}$, the L^2 projection of $H^1(T)$ onto $P_r(T^0) \times P_r(\partial T)$ and the other is R_h, the L^2 projection of $[L^2(T)]^2$ onto $G_{r-1}(T)$.

The following identity will be frequently used in our analysis, cf. [10, Lemma 5.1]:

$$\nabla_d(Q_h u) = R_h(\nabla u), \qquad \forall u \in H^1(T). \tag{17}$$

Now we are ready to describe the WG finite element methods for problems (1). The main idea of the weak Galerkin method is to use the space V_h as testing and trial space and replace the classical gradient operator by its weak version.

We propose the continuous time WG finite element method, based on variational form (5) and the weak Galerkin operator (11). The semi-discrete WG finite element method for equations (1)-(3) is to find $u_h(t) = \{u_0(\cdot,t), u_b(\cdot,t)\} \in V_h^0$ for $t \geq 0$ such that $u_h(0) = Q_h\psi$ and the following equation holds

$$((u_h)_t, v_0) + a_s(u_h, v) = (f, v_0), \qquad \forall v = \{v_0, v_b\} \in V_h^0, \quad t > 0. \tag{18}$$

Let $k > 0$ be a time step-size. At the time level $t = t_n = nk$, with integer $0 \leq n \leq N, Nk = T$, denote by $U^n = U_h^n \in V_h$ the approximation of $u(t_n)$. We further discretize equation (18) with respect to t by the backward Euler method to obtain a full discrete WG finite element method: seek $U^n \in V_h (n = 0, 1, 2, \cdots, N)$ such that $U^0 = Q_h\psi$ and

$$(\bar{\partial} U^n, v_0) + a_s(U^n, v) = (f(t_n), v_0), \quad \forall v = \{v_0, v_b\} \in V_h^0, \tag{19}$$

where $\bar{\partial} U^n = (U^n - U^{n-1})/k$. This is equivalent to

$$(U^n, v_0) + k a_s(U^n, v) = (U^{n-1} + k f(t_n), v_0), \quad \forall v = \{v_0, v_b\} \in V_h^0. \tag{20}$$

3. Energy conservation of WG

In this section, we investigate the energy conservation property of the semi-discrete WG finite element approximation u_h. The solution u of the problem (1)-(3) has the following energy preserving property on each $T \in \mathcal{T}_h$ [7]:

$$\int_{t-\triangle t}^{t+\triangle t} \int_T u_t dx dt + \int_{t-\triangle t}^{t+\triangle t} \int_{\partial T} \mathbf{q} \cdot \mathbf{n} ds dt = \int_{t-\triangle t}^{t+\triangle t} \int_T f dx dt, \tag{21}$$

where $\mathbf{q} = -a\nabla u$ is the flow rate of heat energy. We claim that the semi-discrete WG finite element method for (1)-(3) preserves the energy conservation property (21).

Choosing in (18) the test function $v = \{v_0, v_b = 0\}$ so that $v_0 = 1$ on T and $v_0 = 0$ elsewhere. We then obtain by integrating over the time period $[t - \triangle t, t + \triangle t]$

$$\int_{t-\triangle t}^{t+\triangle t} \int_T u_t dx dt + \int_{t-\triangle t}^{t+\triangle t} a_s(u_h, v) dt = \int_{t-\triangle t}^{t+\triangle t} \int_T f dx dt, \tag{22}$$

where

$$a_s(u_h, v) = \int_T a\nabla_d u_h \cdot \nabla_d v dx + h_T^{-1} \int_{\partial T} (u_0 - u_b) ds. \tag{23}$$

Using the definitions of operators R_h and ∇_d in (11), we obtain

$$\begin{aligned} \int_T a\nabla_d u_h \cdot \nabla_d v dx &= \int_T R_h(a\nabla_d u_h) \cdot \nabla_d v dx \\ &= -\int_T \nabla \cdot R_h(a\nabla_d u_h) dx \\ &= -\int_{\partial T} R_h(a\nabla_d u_h) \cdot \mathbf{n} ds. \end{aligned} \tag{24}$$

Substituting (24) to (23), together with (22), we have

$$\begin{aligned}&\int_{t-\triangle t}^{t+\triangle t}\int_T u_t dxdt + \int_{t-\triangle t}^{t+\triangle t}\int_{\partial T}\{-R_h(a\nabla_d u_h) + h_T^{-1}(u_0-u_b)\mathbf{n}\}\cdot\mathbf{n}dsdt\\ =&\int_{t-\triangle t}^{t+\triangle t}\int_T f dxdt,\end{aligned}\tag{25}$$

which provides a numerical flux

$$\mathbf{q}_h\cdot\mathbf{n} = \{-R_h(a\nabla_d u_h) + h_T^{-1}(u_0-u_b)\mathbf{n}\}\cdot\mathbf{n}.$$

We then aim to verify that the numerical flux $\mathbf{q}_h\cdot\mathbf{n}$ crosses continuously the edge of each T. Denote by $\overline{e}$ the common edge between two elements T_1 and T_2. Choose the test function $v=\{v_0,v_b\}$ so that $v_0\equiv 0$, v_b arbitrary on $\overline{e}$ and zero elsewhere. Using $v_0\equiv 0$ and (18), we have

$$\begin{aligned}\int_{T_1\cup T_2} a\nabla_d u_h\cdot\nabla_d v dx - h_{T_1}^{-1}\int_{\partial T_1\cap\overline{e}}(u_0-u_b)\mid_{T_1} v_b ds&\\ - h_{T_2}^{-1}\int_{\partial T_2\cap\overline{e}}(u_0-u_b)\mid_{T_2} v_b ds = 0.&\end{aligned}\tag{26}$$

Using (11), we obtain

$$\begin{aligned}&\int_{T_1\cup T_2} a\nabla_d u_h\cdot\nabla_d v dx\\ =&\int_{T_1\cup T_2} R_h(a\nabla_d u_h)\cdot\nabla_d v dx\\ =&\int_{\overline{e}}(R_h(a\nabla_d u_h)\mid_{T_1}\cdot\mathbf{n_1} + R_h(a\nabla_d u_h)\mid_{T_2}\cdot\mathbf{n_2})v_b ds,\end{aligned}$$

where $\mathbf{n_1}$ and $\mathbf{n_2}$ are the outward normal vectors to T_1 and T_2 on the edge $\overline{e}$, respectively. Noting the fact $\mathbf{n_1}+\mathbf{n_2}=0$, substituting the above equation into (26) yields

$$\begin{aligned}&\int_{\overline{e}}\{-R_h(a\nabla_d u_h)\mid_{T_1} + h_{T_1}^{-1}(u_0-u_b)\mid_{T_1}\mathbf{n_1}\}\cdot\mathbf{n_1}v_b ds\\ +&\int_{\overline{e}}\{-R_h(a\nabla_d u_h)\mid_{T_2} + h_{T_2}^{-1}(u_0-u_b)\mid_{T_2}\mathbf{n_2}\}\cdot\mathbf{n_2}v_b ds = 0,\end{aligned}\tag{27}$$

which shows the numerical flux $\mathbf{q}_h\cdot\mathbf{n}$ is continuous along the normal direction.

4. Error analysis

In this section, we estimate the errors for both continuous and discrete time WG finite element methods. The difference between WG finite element approximation u_h and the L^2 projection $Q_h u$ of the exact solution u is measured in different norms.

To this end, we need the following estimates concerning the projection operators Q_h and R_h, see [16]. Note here the underlying mesh $\mathcal{T}_h$ is general enough and allows polygons.

Lemma 4.1. [16, Lemma5.1] *There hold for any* $\phi\in H^{r+1}(\Omega)$

$$\sum_{T\in\mathcal{T}_h}\|\phi-Q_0\phi\|_T^2 + \sum_{T\in\mathcal{T}_h} h_T^2\|\nabla(\phi-Q_0\phi)\|_T^2 \le Ch^{2(r+1)}\|\phi\|_{r+1}^2,\tag{28}$$

$$\sum_{T\in\mathcal{T}_h}\|a(\nabla\phi-R_h(\nabla\phi))\|_T^2 \le Ch^{2r}\|\phi\|_{r+1}^2,\tag{29}$$

where C *is a generic constant independent of the mesh size* h.

In addition, the following trace inequality holds [16, Lemma A.1]

$$\|\phi\|_{\partial T}^2 \le C(h_T^{-1}\|\phi\|_T^2 + h_T\|\nabla\phi\|_T^2), \quad \forall T \in \mathcal{T}_h, \forall \phi \in H^1(T). \tag{30}$$

4.1. Continuous time WG finite element method. We then in this subsection aim to investigate the approximation properties for the semi-discrete solution.

For $\phi \in H^1(T)$ and $v \in V_h$, based on (11), (17), we have by integration by parts

$$\begin{aligned}(a\nabla_d Q_h\phi, \nabla_d v)_T &= (aR_h(\nabla\phi), \nabla_d v)_T \\ &= -(v_0, \nabla\cdot(aR_h\nabla\phi))_T + \langle v_b, (aR_h\nabla\phi)\cdot\mathbf{n}\rangle_{\partial T} \\ &= (\nabla v_0, aR_h\nabla\phi)_T - \langle v_0 - v_b, (aR_h\nabla\phi)\cdot\mathbf{n}\rangle_{\partial T} \\ &= (a\nabla\phi, \nabla v_0)_T - \langle (aR_h\nabla\phi)\cdot\mathbf{n}, v_0 - v_b\rangle_{\partial T}.\end{aligned} \tag{31}$$

Let u and u_h be the solutions of (1)-(3) and (18), respectively. Denote by $e := u_h - Q_h u \in V_h^0$ the difference between the weak Galerkin approximation and the L^2 projection of the exact solution u.

Theorem 4.2. *Assume $u \in H^{r+1}(\Omega)$. Then there exists a constant $C > 0$ independent of the mesh size h such that the following estimates hold*

$$\|e(\cdot,t)\|^2 + \int_0^t \alpha \,|||\, e \,|||^2\, ds \le \|e(\cdot,0)\|^2 + Ch^{2r}\int_0^t \|u\|_{r+1}^2 ds,$$

and

$$\begin{aligned}&\int_0^t \|e_t\|^2 ds + \frac{\alpha}{4}\,|||\, e \,|||^2 + \|e\| \le \beta \,|||\, e(\cdot,0) \,|||^2 + \|e(\cdot,0)\|^2 \\ &+ Ch^{2r}(\|\psi\|_{r+1}^2 + \|u\|_{r+1}^2 + \int_0^t \|u\|_{r+1}^2 ds + \int_0^t \|u_t\|_{r+1}^2 ds).\end{aligned}$$

Proof. Let $v = \{v_0, v_b\} \in V_h^0$ be a test function. By testing (1.1) against v_0, together with $R_h(\nabla u) = \nabla_d(Q_h u)$ for $u \in H^1$, $(Q_h u_t, v_0) = (u_t, v_0)$, and (31), we obtain

$$\begin{aligned}&(f, v_0) \\ =&(u_t, v_0) + \sum_{T\in\mathcal{T}_h}(-\nabla\cdot a\nabla u, v_0)_T \\ =&(Q_0 u_t, v_0) + \sum_{T\in\mathcal{T}_h}(a\nabla u, \nabla v_0)_T - \sum_{T\in\mathcal{T}_h}\langle v_0, a(\nabla u)\cdot\mathbf{n}\rangle_{\partial T} \\ =&(Q_0 u_t, v_0) + \sum_{T\in\mathcal{T}_h}(a\nabla_d Q_h u, \nabla_d v)_T + \sum_{T\in\mathcal{T}_h}\langle a(R_h\nabla u - \nabla u)\cdot\mathbf{n}, v_0 - v_b\rangle_{\partial T},\end{aligned}$$

where the fact that $\sum_{T\in\mathcal{T}_h}\langle a(\nabla u)\cdot\mathbf{n}, v_b\rangle_{\partial T} = 0$ is applied. Adding $s(Q_h u, v)$ to both sides of the above equation gives

$$\begin{aligned}&(f, v_0) + s(Q_h u, v) \\ =&(Q_0 u_t, v_0) + a_s(Q_h u, v) + \sum_{T\in\mathcal{T}_h}\langle a(R_h\nabla u - \nabla u)\cdot\mathbf{n}, v_0 - v_b\rangle_{\partial T}.\end{aligned} \tag{32}$$

Using (18), we then obtain an error equation

$$\begin{aligned}&((u_0 - Q_0 u)_t, v_0) + a_s(u_h - Q_h u, v) \\ =&\sum_{T\in\mathcal{T}_h}\langle a(\nabla u - R_h\nabla u)\cdot\mathbf{n}, v_0 - v_b\rangle_{\partial T} + s(Q_h u, v).\end{aligned} \tag{33}$$

Now, we bound the terms in the right-hand side of (33) one by one. Using the Cauchy-Schwarz inequality, the trace inequality, and the Young' s inequality consecutively we have

$$\begin{aligned}
&|\sum_{T\in\mathcal{T}_h}\langle a(\nabla u - R_h\nabla u)\cdot\mathbf{n}, v_0 - v_b\rangle_{\partial T}| \\
\leq&(\sum_{T\in\mathcal{T}_h} h_T\|a(\nabla u - R_h\nabla u\|_{\partial T}^2)^{1/2}(\sum_{T\in\mathcal{T}_h} h_T^{-1}\|v_0 - v_b\|_{\partial T}^2)^{1/2} \\
\leq&\frac{1}{\alpha}\sum_{T\in\mathcal{T}_h} h_T\|a(\nabla u - R_h\nabla u)\|_{\partial T}^2 + \frac{\alpha}{4}\sum_{T\in\mathcal{T}_h} h_T^{-1}\|v_0 - v_b\|_{\partial T}^2 \\
\leq&\frac{C}{\alpha}(\sum_{T\in\mathcal{T}_h}(\|a(\nabla u - R_h\nabla u)\|_T^2 + h_T^2\|\nabla[a(\nabla u - R_h\nabla u)]\|_T^2) + \frac{\alpha}{4}|||v|||^2 \\
\leq&Ch^{2r}\|u\|_{r+1}^2 + \frac{\alpha}{4}|||v|||^2 .
\end{aligned} \tag{34}$$

Similarly,

$$\begin{aligned}
&|s(Q_hu, v)| \\
=&|\sum_{T\in\mathcal{T}_h} h_T^{-1}\langle Q_0u - Q_bu, v_0 - v_b\rangle_{\partial T}| \\
=&|\sum_{T\in\mathcal{T}_h} h_T^{-1}\langle Q_0u - u, v_0 - v_b\rangle_{\partial T}| \\
\leq&(\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0u - u\|_{\partial T}^2)^{1/2}(\sum_{T\in\mathcal{T}_h} h_T^{-1}\|v_0 - v_b\|_{\partial T}^2)^{1/2} \\
\leq&\frac{1}{\alpha}\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0u - u\|_{\partial T}^2 + \frac{\alpha}{4}\sum_{T\in\mathcal{T}_h} h_T^{-1}\|v_0 - v_b\|_{\partial T}^2 \\
\leq&\frac{C}{\alpha}\sum_{T\in\mathcal{T}_h}(h_T^{-2}\|Q_0u - u\|_T^2 + \|\nabla(Q_0u - u)\|_T^2) + \frac{\alpha}{4}|||v|||^2 \\
\leq&Ch^{2r}\|u\|_{r+1}^2 + \frac{\alpha}{4}|||v|||^2 .
\end{aligned} \tag{35}$$

Choosing $v = e$ in the error equation (33), we obtain

$$(e_t, e) + a_s(e, e) = \sum_{T\in\mathcal{T}_h}\langle a(\nabla u - R_h\nabla u)\cdot\mathbf{n}, e_0 - e_b\rangle_{\partial T} + s(Q_hu, e).$$

Using again the Cauchy-Schwarz inequality and the coercivity of the bilinear form, we have

$$\frac{1}{2}\frac{d}{dt}\|e\|^2 + \alpha|||e|||^2 \leq \frac{1}{\alpha}\sum_{T\in\mathcal{T}_h}(h_T\|a(\nabla u - R_h\nabla u)\|_{\partial T}^2 + h_T^{-1}\|Q_0u - u\|_{\partial T}^2) + \frac{\alpha}{2}|||e|||^2 .$$

By integrating over the time period $[0, t]$, we get

$$\begin{aligned}
&\|e\|^2 + \int_0^t \alpha|||e|||^2\,ds \leq \|e(\cdot, 0)\|^2 \\
&+\frac{2}{\alpha}\int_0^t\sum_{T\in\mathcal{T}_h}(h_T\|a(\nabla u - R_h\nabla u)\|_{\partial T}^2 + h_T^{-1}\|Q_0u - u\|_{\partial T}^2)ds.
\end{aligned} \tag{36}$$

Hence, we have

$$\|e\|^2 + \int_0^t \alpha|||e|||^2\,ds \leq \|e(\cdot, 0)\|^2 + Ch^{2r}\int_0^t\|u\|_{r+1}^2 ds.$$

In order to estimate $|||e|||$, we apply the error equation with $v = (u_h - Q_h u)_t = e_t$,

$$
\begin{aligned}
&(e_t, e_t) + a_s(e, e_t) \\
=&\sum_{T\in\mathcal{T}_h} \langle a(\nabla u - R_h \nabla u)\cdot \mathbf{n}, (e_t)_0 - (e_t)_b\rangle_{\partial T} + s(Q_h u, e_t) \\
=&\frac{d}{dt}\sum_{T\in\mathcal{T}_h} \langle a(\nabla u - R_h \nabla u)\cdot \mathbf{n}, e_0 - e_b\rangle_{\partial T} + \frac{d}{dt}s(Q_h u, e) \\
&-\sum_{T\in\mathcal{T}_h} \langle a(\nabla u_t - R_h \nabla u_t)\cdot \mathbf{n}, e_0 - e_b\rangle_{\partial T} - s(Q_h u_t, e).
\end{aligned}
$$

This in turn gives by Cauchy-Schwarz inequality

$$
\begin{aligned}
&\|e_t\|^2 + \frac{1}{2}\frac{d}{dt}a_s(e,e) \\
\leq&\frac{d}{dt}\sum_{T\in\mathcal{T}_h} \langle a(\nabla u - R_h \nabla u), \cdot\mathbf{n}, e_0 - e_b\rangle_{\partial T} + \frac{d}{dt}s(Q_h u, e) \\
&+\frac{1}{2\alpha}\sum_{T\in\mathcal{T}_h} h_T\|a(\nabla u_t - R_h\nabla u_t)\|_{\partial T}^2 + \frac{\alpha}{2}\, |||\, e|||^2 \\
&+\frac{1}{2\alpha}\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0 u_t - u_t\|_{\partial T}^2 + \frac{\alpha}{2}\, |||\, e\, |||^2 .
\end{aligned}
$$

Thus, integrating with respect to t and together with the coercivity and boundedness yields

$$
\begin{aligned}
&\int_0^t \|e_t\|^2 ds + \frac{\alpha}{2}\, |||\, e|||^2 \\
\leq&\|e(\cdot,0)\|^2 + \frac{\beta}{2}\, |||\, e(\cdot,0)\, |||^2 + s(Q_h u, e) - s(Q_h u(\cdot,0), e(\cdot,0)) \\
&+\sum_{T\in\mathcal{T}_h} \langle a(\nabla u - R_h\nabla u)\cdot\mathbf{n}, e_0 - e_b\rangle_{\partial T} \\
&-\sum_{T\in\mathcal{T}_h} \langle a(\nabla u(\cdot,0) - R_h\nabla u(\cdot,0))\cdot\mathbf{n}, e(\cdot,0)_0 - e(\cdot,0)_b\rangle_{\partial T} \\
&+\int_0^t \frac{1}{2\alpha}(\sum_{T\in\mathcal{T}_h} h_T\|a(\nabla u_t - R_h\nabla u_t)\|_{\partial T}^2 ds \\
&+\int_0^t \frac{1}{2\alpha}\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0 u_t - u_t\|_{\partial T}^2 ds + \int_0^t \alpha\, |||\, e\, |||^2\, ds \\
\leq&\|e(\cdot,0)\|^2 + \frac{\beta}{2}\, |||\, e(\cdot,0)|||^2 \\
&+\frac{2}{\alpha}\sum_{T\in\mathcal{T}_h} (h_T\|a(\nabla u - R_h\nabla u)\|_{\partial T}^2 + h_T^{-1}\|Q_0 u - u\|_{\partial T}^2) + \frac{\alpha}{4}\, |||\, e|||^2 \\
&+\frac{1}{\beta}\sum_{T\in\mathcal{T}_h} (h_T\|a(\nabla u(\cdot,0) - R_h\nabla u(\cdot,0))\|_{\partial T}^2 \\
&+h_T^{-1}\|Q_0 u(\cdot,0) - u(\cdot,0)\|_{\partial T}^2) + \frac{\beta}{2}\, |||\, e(\cdot,0)|||^2
\end{aligned}
$$

$$
\begin{aligned}
&+\int_0^t \frac{1}{2\alpha}(\sum_{T\in\mathcal{T}_h} h_T\|a(\nabla u_t - R_h\nabla u_t)\|^2_{\partial T}ds\\
&+\int_0^t \frac{1}{2\alpha}\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0u_t - u_t\|^2_{\partial T}ds + \int_0^t \alpha ⦀ e ⦀^2\, ds.
\end{aligned}
$$

Combining (34), (35), (36) and the above equation, using the Young' s inequality, we have

$$
\begin{aligned}
&\|e\|^2 + \int_0^t \|e_t\|^2 ds + \frac{\alpha}{4} ⦀ e⦀^2\\
\le&\|e(\cdot,0)\|^2 + \beta ⦀ e(\cdot,0)⦀^2\\
&+ Ch^{2r}(\|u(\cdot,0)\|^2_{r+1} + \|u\|^2_{r+1} + \int_0^t \|u\|^2_{r+1}ds + \int_0^t \|u_t\|^2_{r+1}ds).
\end{aligned}
$$

This completes the proof. □

4.2. Discrete time WG finite element method. We then in this subsection investigate the error analysis for full discrete weak Galerkin finite element method. To this end, we need the following Poincare inequality related to the weak gradient operator.

Lemma 4.3. [10, Lemma7.1] *There exists a constant C independent of the mesh size h such that*

$$\|v\|^2 \le C ⦀ v⦀^2, \quad \forall v = \{v_0, v_b\} \in V_h^0.$$

We estimate the error of the full discrete WG element method. Let u and U^n be the solutions of (1)-(3) and (19), respectively. Denote by $e^n := U^n - Q_hu(t_n)$ the difference between the backward Euler WG approximation and the L^2 projection of the exact solution u.

Theorem 4.4. *Assume $u \in C^2\big([0,T];H^{r+1}(\Omega)\big)$. Then there exists a constant $C > 0$ independent of the mesh size h such that for $0 < n \le N$*

$$\|e^n\|^2 + \sum_{j=1}^n \alpha k ⦀ e^j⦀^2 \le \|e^0\|^2 + C(h^{2r}\|u\|^2_{r+1,\infty} + k^2\int_0^{t_n} \|u_{tt}\|^2 ds),$$

and

$$
\begin{aligned}
&⦀ e^n⦀^2 \le C\big\{\|e^0\|^2 + ⦀e^0⦀^2\\
&+ h^{2r}\big(\|u(\cdot,0)\|^2_{r+1} + \|u\|^2_{r+1,\infty} + \|u_t\|^2_{r+1,\infty} + k^2\int_0^{t_n}\|u_{tt}\|^2_{r+1}ds\big)\\
&+ k^2\int_0^{t_n} \|u_{tt}\|^2 ds\big\},
\end{aligned}
$$

where

$$\|u\|_{r+1,\infty} = \max_{0\le t\le T}\{\|u(t)\|_{r+1}\}.$$

Proof. It is easy to see that

$$
\begin{aligned}
&(\bar\partial U^n - Q_hu_t(t_n), v_0)\\
=&(\bar\partial(U^n - Q_hu(t_n)), v_0) + (\bar\partial Q_hu(t_n) - Q_hu_t(t_n), v_0)\\
=&(\bar\partial(U^n - Q_hu(t_n)), v_0) + (\bar\partial u(t_n) - u_t(t_n), v_0).
\end{aligned}
$$

Then, we obtain the following error equation for the backward Euler WG method

$$
\begin{aligned}
&(\bar{\partial}(U^n - Q_h u(t_n)), v_0) + a_s(U^n - Q_h u(t_n)), v) \\
=&(u_t - \bar{\partial}u(t_n), v_0) + \sum_{T\in\mathcal{T}_h} \langle a(\nabla u(t_n) - R_h \nabla u(t_n)) \cdot \mathbf{n}, v_0 - v_b\rangle_{\partial T} + s(Q_h u(t_n), v),
\end{aligned}
$$

i.e.

$$
\begin{aligned}
&(\bar{\partial}e^n, v_0) + a_s(e^n, v) \\
=&(u_t - \bar{\partial}u(t_n), v_0) + \sum_{T\in\mathcal{T}_h} \langle a(\nabla u - R_h \nabla u) \cdot \mathbf{n}, v_0 - v_b\rangle_{\partial T} + s(Q_h u(t_n), v).
\end{aligned} \tag{37}
$$

Let

$$
\begin{aligned}
\omega_1^n &= u_t(t_n) - \bar{\partial}u(t_n), \\
\omega_{2T}^n &= (a(\nabla u(t_n) - R_h \nabla u(t_n)) \cdot \mathbf{n})\mid_{\partial T}, \\
\omega_{3,T}^n &= (Q_0 u(t_n) - Q_b u(t_n))\mid_{\partial T}, \\
e_T^n &= (e_0(t_n) - e_b(t_n))\mid_{\partial T}.
\end{aligned}
$$

Choosing $v = e^n$ in (37) gives

$$
(\bar{\partial}e^n, e^n) + a_s(e^n, e^n) = (\omega_1^n, e_0^n) + \sum_{T\in\mathcal{T}_h} (\langle \omega_{2,T}^n, e_T^n\rangle + h_T^{-1}\langle \omega_{3,T}^n, e_T^n\rangle). \tag{38}
$$

By the coercivity of the bilinear form (16) and the Cauchy-Schwarz inequality, we obtain

$$
\begin{aligned}
&\|e^n\|^2 + \alpha k \,|\!|\!|\, e^n |\!|\!|^2 \\
\le&(e^{n-1}, e^n) + k\|\omega_1^n\| \cdot \|e^n\| + k\Big|\sum_{T\in\mathcal{T}_h} (\langle \omega_{2,T}^n, e_T^n\rangle + h_T^{-1}\langle \omega_{3,T}^n, e_T^n\rangle)\Big| \\
\le&\frac{1}{2}\|e^{n-1}\|^2 + \frac{1}{2}\|e^n\|^2 + k\|\omega_1^n\| \cdot \|e^n\| + k\Big|\sum_{T\in\mathcal{T}_h} (\langle \omega_{2,T}^n, e_T^n\rangle + h_T^{-1}\langle \omega_{3,T}^n, e_T^n\rangle)\Big|.
\end{aligned}
$$

By the Poincare inequality in Lemma 4.3, we have

$$
\begin{aligned}
&\frac{1}{2}\|e^n\|^2 + \alpha k \,|\!|\!|\, e^n |\!|\!|^2 \\
\le&\frac{1}{2}\|e^{n-1}\|^2 + \frac{k\alpha}{4} \,|\!|\!|\, e^n |\!|\!|^2 + \frac{ck}{\alpha}\|\omega_1^n\|^2 \\
&+ \frac{2k}{\alpha} \sum_{T\in\mathcal{T}_h} h_T \|a(\nabla u(t_n) - R_h \nabla u(t_n)\|_{\partial T}^2 + \frac{k\alpha}{8} \,|\!|\!|\, e^n |\!|\!|^2 \\
&+ \frac{2k}{\alpha} \sum_{T\in\mathcal{T}_h} h_T^{-1} \|Q_0 u(t_n) - u(t_n)\|_{\partial T}^2 + \frac{k\alpha}{8} \,|\!|\!|\, e^n |\!|\!|^2 .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\frac{1}{2}\|e^n\|^2 + \frac{1}{2}\alpha k \,|\!|\!|\, e^n |\!|\!|^2 \le&\frac{1}{2}\|e^{n-1}\|^2 + \frac{ck}{\alpha}\|\omega_1^n\|^2 \\
&+ \frac{2k}{\alpha} \sum_{T\in\mathcal{T}_h} h_T \|a(\nabla u(t_n) - R_h \nabla u(t_n)\|_{\partial T}^2 \\
&+ \frac{2k}{\alpha} \sum_{T\in\mathcal{T}_h} h_T^{-1} \|Q_0 u(t_n) - u(t_n)\|_{\partial T}^2,
\end{aligned}
$$

which gives by induction

$$
\begin{aligned}
\|e^n\|^2+\sum_{j=1}^{n}\alpha k \interleave e^j\interleave^2 \le & \|e^0\|^2+\frac{ck}{\alpha}\sum_{j=1}^{n}\|\omega_1^j\|^2 \\
& +\frac{4k}{\alpha}\sum_{j=1}^{n}\sum_{T\in\mathcal{T}_h} h_T\|a(\nabla u(t_j)-R_h\nabla u(t_j)\|_{\partial T}^2 \\
& +\frac{4k}{\alpha}\sum_{j=1}^{n}\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0u(t_j)-u(t_j)\|_{\partial T}^2.
\end{aligned}
\tag{39}
$$

Noting that

$$
\omega_1^j=u_t(t_j)-\frac{u(t_j)-u(t_{j-1})}{k}=\frac{1}{k}\int_{t_{j-1}}^{t_j}(s-t_{j-1})u_{tt}ds,
$$

we then obtain

$$
\begin{aligned}
\|\omega_1^j\|^2 &= \int_\Omega\Big(\frac{1}{k}\int_{t_{j-1}}^{t_j}(s-t_{j-1})u_{tt}ds\Big)^2dx \\
&\le \frac{1}{k^2}\int_\Omega\int_{t_{j-1}}^{t_j}(s-t_{j-1})^2ds\int_{t_{j-1}}^{t_j}u_{tt}^2dsdx \\
&\le Ck\int_{t_{j-1}}^{t_j}\|u_{tt}\|^2ds.
\end{aligned}
\tag{40}
$$

Similar to the analysis in (34) and (35), we get

$$
\sum_{T\in\mathcal{T}_h} h_T\|a(\nabla u(t_j)-R_h\nabla u(t_j)\|_{\partial T}^2+\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0u(t_j)-u(t_j)\|_{\partial T}^2\le Ch^{2r}\|u(t_j)\|_{r+1}^2.
\tag{41}
$$

Thus we have proved the error estimate for $\|e^n\|$. In order to estimate $\interleave e^n\interleave$, we choose $v=\bar\partial e^n$ in error equation (37) and obtain

$$
(\bar\partial e^n,\bar\partial e^n)+a_s(e^n,\bar\partial e^n)=(\omega_1^n,\bar\partial e_0^n)+\sum_{T\in\mathcal{T}_h}(\langle\omega_{2,T}^n,\bar\partial e_T^n\rangle+h_T^{-1}\langle\omega_{3,T}^n,\bar\partial e_T^n\rangle),
$$

where $\bar\partial e_T^n=\frac{1}{k}((e_0^n-e_0^{n-1})-(e_b^n-e_b^{n-1}))\mid_{\partial T}$. Noting that $\langle\omega_{2,T}^n,\bar\partial e_T^n\rangle$ on the right-hand side can be written as

$$
\langle\omega_{2,T}^n,\bar\partial e_T^n\rangle=\bar\partial\langle\omega_{2,T}^n,e_T^n\rangle+\langle(\omega_2^n)_t-\bar\partial\omega_{2,T}^n,e_T^{n-1}\rangle-\langle(\omega_2^n)_t,e_T^{n-1}\rangle,
$$

where $(\omega_2^n)_t=a(\nabla u_t(t_n)-R_h\nabla u_t(t_n))\cdot\mathbf{n}\mid_T$. Analogously, Let $(\omega_3^n)_t=(Q_0u_t(t_n)-Q_bu_t(t_n))\mid_T$, the following form holds

$$
\langle\omega_{3,T}^n,\bar\partial e_T^n\rangle=\bar\partial\langle\omega_{3,T}^n,e_T^n\rangle+\langle(\omega_3^n)_t-\bar\partial\omega_{3,T}^n,e_T^{n-1}\rangle-\langle(\omega_3^n)_t,e_T^{n-1}\rangle.
$$

Then we obtain

$$
\begin{aligned}
& k\|\bar\partial e^n\|^2+a_s(e^n,e^n) \\
=& a_s(e^n,e^{n-1})+k(\omega_1^n,\bar\partial e^n) \\
& +k\sum_{T\in\mathcal{T}_h}\{\bar\partial\langle\omega_{2,T}^n,e_T^n\rangle+\langle(\omega_2^n)_t-\bar\partial\omega_{2,T}^n,e_T^{n-1}\rangle-\langle(\omega_2^n)_t,e_T^{n-1}\rangle \\
& +h_T^{-1}(\bar\partial\langle\omega_{3,T}^n,e_T^n\rangle+\langle(\omega_3^n)_t-\bar\partial\omega_{3,T}^n,e_T^{n-1}\rangle-\langle(\omega_3^n)_t,e_T^{n-1}\rangle)\}.
\end{aligned}
$$

As what we have done to (34) and (35), by the Cauchy-Schwarz inequality and the triangle inequality, we have

$$
\begin{aligned}
& k\|\bar{\partial} e^n\|^2 + a_s(e^n, e^n) \\
\leq & \frac{1}{2} a_s(e^n, e^n) + \frac{1}{2} a_s(e^{n-1}, e^{n-1}) + k \sum_{T\in\mathcal{T}_h} (\bar{\partial}\langle \omega_{2,T}^n, e_T^n\rangle + h_T^{-1}\bar{\partial}\langle \omega_{3,T}^n, e_T^n\rangle) \\
& + \frac{k}{4}\|\omega_1^n\|^2 + k\|\bar{\partial} e^n)\|^2 \\
& + \sum_{T\in\mathcal{T}_h} k(h_T\|(\omega_2^n)_t - \bar{\partial}\omega_{2,T}^n\|_{\partial T}^2 + h_T^{-1}\|(\omega_3^n)_t - \bar{\partial}\omega_{3,T}^n\|_{\partial T}^2) + \frac{k}{2} ||| e^{n-1} |||^2 \\
& + \sum_{T\in\mathcal{T}_h} k(h_T\|(\omega_{2,T}^n)_t\|_{\partial T}^2 + h_T^{-1}\|(\omega_{3,T}^n)_t\|_{\partial T}^2) + \frac{k}{2} ||| e^{n-1} |||^2,
\end{aligned}
$$

where

$$
\begin{aligned}
\|\omega_{2,T}^j\|_{\partial T}^2 &= \langle a(\nabla u(t_j) - R_h\nabla u(t_j)), a(\nabla u(t_j) - R_h\nabla u(t_j))\rangle_{\partial T}, \\
\|\omega_{3,T}^j\|_{\partial T}^2 &= \langle Q_0 u(t_j) - Q_b u(t_j), Q_0 u(t_j) - Q_b u(t_j)\rangle_{\partial T}.
\end{aligned}
$$

Moreover, by the triangle inequality and (28), we have

$$
\begin{aligned}
\sum_{T\in\mathcal{T}_h} h_T\|\omega_{2,T}^j\|_{\partial T}^2 &= \sum_{T\in\mathcal{T}_h} h_T\|a(\nabla u(t_j) - R_h\nabla u(t_j)\|_{\partial T}^2 \\
&\leq \sum_{T\in\mathcal{T}_h} C(\|a(\nabla u(t_j) - R_h\nabla u(t_j))\|_T^2 \\
&\quad + h_T^2\|\nabla(a(\nabla u(t_j) - R_h\nabla u(t_j)))\|_T^2 \\
&\leq Ch^{2r}\|u(t_j)\|_{r+1}^2.
\end{aligned}
$$

Using the Cauchy-Schwarz inequality, the trace inequality (30) and (28) the following form holds

$$
\begin{aligned}
& \sum_{T\in\mathcal{T}_h} h_T^{-1}\|\omega_{3,T}^j\|_{\partial T}^2 \\
= & \sum_{T\in\mathcal{T}_h} h_T^{-1}\langle Q_0 u(t_j) - Q_b u(t_j), Q_0 u(t_j) - Q_b u(t_j)\rangle_{\partial T} \\
= & \sum_{T\in\mathcal{T}_h} h_T^{-1}\langle Q_0 u(t_j) - u(t_j), Q_0 u(t_j) - Q_b u(t_j)\rangle_{\partial T} \\
\leq & (\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0 u(t_j) - u(t_j)\|_{\partial T}^2)^{1/2} (\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0 u(t_j) - Q_b u(t_j)\|_{\partial T}^2)^{1/2} \\
\leq & Ch^r\|u(t_j)\|_{r+1} (\sum_{T\in\mathcal{T}_h} h_T^{-1}\|Q_0 u(t_j) - Q_b u(t_j)\|_{\partial T}^2)^{1/2}.
\end{aligned}
$$

Thus,

$$
\sum_{T\in\mathcal{T}_h} h_T\|\omega_{2,T}^j\|_{\partial T}^2 + \sum_{T\in\mathcal{T}_h} h_T^{-1}\|\omega_{3,T}^j\|_{\partial T}^2 \leq Ch^{2r}\|u(t_j)\|_{r+1}^2. \tag{42}
$$

After cancellation and by induction, we arrive at

$$\begin{aligned}\frac{1}{2}a_s(e^n,e^n)\leq&\frac{1}{2}a_s(e^0,e^0)\\&+\sum_{T\in\mathcal{T}_h}\{(\langle\omega_{2,T}^n,e_T^n\rangle-\langle\omega_{2,T}^0,e_T^0\rangle)+h_T^{-1}(\langle\omega_{3,T}^n,e_T^n\rangle-\langle\omega_{3,T}^0,e_T^0\rangle)\}\\&+\frac{k}{4}\sum_{j=1}^n\|\omega_1^j\|^2+k\sum_{j=1}^n|||e^{j-1}|||^2\\&+k\sum_{j=1}^n\sum_{T\in\mathcal{T}_h}(h_T\|(\omega_2^j)_t-\bar{\partial}\omega_{2,T}^j\|_{\partial T}^2+h_T^{-1}\|(\omega_3^j)_t-\bar{\partial}\omega_{3,T}^j\|_{\partial T}^2)\\&+k\sum_{j=1}^n\sum_{T\in\mathcal{T}_h}(h_T\|(\omega_{2,T}^j)_t\|_{\partial T}^2+h_T^{-1}\|(\omega_{3,T}^j)_t\|_{\partial T}^2).\end{aligned}$$

By the coercivity of the bilinear form (16) and the Cauchy-Schwarz inequality again, we obtain

$$\begin{aligned}\frac{\alpha}{2}|||\,e^n|||^2\leq&\frac{\beta}{2}|||\,e^0\,|||^2+\frac{k}{4}\sum_{j=1}^n\|\omega_1^j\|^2+k\sum_{j=1}^n|||e^{j-1}|||^2\\&+\frac{2}{\alpha}\sum_{T\in\mathcal{T}_h}(h_T\|\omega_{2,T}^n\|_{\partial T}^2+h_T^{-1}\|\omega_{3,T}^n\|_{\partial T}^2)+\frac{\alpha}{4}|||\,e^n|||^2\\&+\frac{1}{\beta}\sum_{T\in\mathcal{T}_h}(h_T\|\omega_{2,T}^0\|_{\partial T}^2+h_T^{-1}\|\omega_{3,T}^0\|_{\partial T}^2)+\frac{\beta}{2}|||\,e^0|||^2\\&+k\sum_{j=1}^n\sum_{T\in\mathcal{T}_h}(h_T\|(\omega_2^j)_t-\bar{\partial}\omega_{2,T}^j\|_{\partial T}^2+h_T^{-1}\|(\omega_3^j)_t-\bar{\partial}\omega_{3,T}^j\|_{\partial T}^2)\\&+k\sum_{j=1}^n\sum_{T\in\mathcal{T}_h}(h_T\|(\omega_{2,T}^j)_t\|_{\partial T}^2+h_T^{-1}\|(\omega_{3,T}^j)_t\|_{\partial T}^2).\end{aligned}\tag{43}$$

Similar to (40), we obtain

$$\sum_{T\in\mathcal{T}_h}h_T\|(\omega_2^j)_t-\bar{\partial}\omega_2^j\|_{\partial T}^2\leq Ck\sum_{T\in\mathcal{T}_h}h_T\int_{t_{j-1}}^{t_j}\|(\omega_2)_{tt}\|_{\partial T}^2ds\leq Ckh^{2r}\int_{t_{j-1}}^{t_j}\|u_{tt}\|_{r+1}^2ds,$$

and

$$\sum_{T\in\mathcal{T}_h}h_T^{-1}\|(\omega_3^j)_t-\bar{\partial}\omega_3^j\|_{\partial T}^2\leq Ck\sum_{T\in\mathcal{T}_h}h_T^{-1}\int_{t_{j-1}}^{t_j}\|(\omega_3)_{tt}\|_{\partial T}^2ds\leq Ckh^{2r}\int_{t_{j-1}}^{t_j}\|u_{tt}\|_{r+1}^2ds.$$

Substituting the above inequality into (43), together with (39), (40) and (42), we complete the proof. □

4.3. Optimal order of error estimates in L^2 and $|||\cdot|||$ norms. The optimal order of error estimates for $\nabla_d e$ and $\nabla_d e^n$ was obtained in Section 4.2. In this section,we derive an optimal order of estimate for e in L^2-norm, the basic idea applied is to use Wheeler's projection as [18, 15]. Now, we define an elliptic projection E_h onto the discrete weak apace V_h as follows

$$a_s(E_hu,\chi)=(-\nabla\cdot(a\nabla u),\chi),\quad\forall\chi\in V_h,u\in H_0^1,\tag{44}$$

which can be viewed as that $E_h u$ is the WG finite element approximation of the solution of the corresponding elliptic problem with exact solution u

$$\begin{aligned} -\nabla\cdot(a\nabla u) &= f \quad in \quad \Omega, \\ u &= 0 \quad on \quad \Omega. \end{aligned} \tag{45}$$

The error estimate for $E_h u$, as shown in the following lemma, should be applied.

Lemma 4.5. [10, Theorem8.1,Theorem8.2] *Let $E_h u \in V_h$ be the WG finite element solution of problem (45). Assume that the exact solution of problem (45) is so regular that $u \in H^{r+1}(\Omega)$. In addition, assume that the dual problem of problem (45) has the usual H^2-regularity. Then, there exists a constant C such that*

$$|||E_h u - Q_h u||| \le Ch^r \|u\|_{r+1}, \tag{46}$$

$$\|E_h u - Q_h u\|_{L^2} \le Ch^{r+1} \|u\|_{r+1}. \tag{47}$$

We consider the error $u_h - Q_h u$, which can be separated as

$$u_h(t) - Q_h u(t) = \theta(t) + \rho(t), \tag{48}$$

where $\rho = E_h u - Q_h u$ is estimated using Lemma 4.5, $\theta = u_h - E_h u$ is thus the term we need to further bound. The error estimates for continuous time WG finite element method in L^2 and $|||\cdot|||$ norms are given in the following two theorems, respectively.

Theorem 4.6. *Let $u \in H^{r+1}(\Omega)$ and the corresponding elliptic problem (45) has the H^2-regularity. Then there exists a constant $C > 0$ independent of the mesh size h such that*

$$\|u_h(t) - Q_h u(t)\| \le \|u_h(0) - Q_h u(0)\| + Ch^{r+1}(\|\psi\|_{r+1} + \int_0^t \|u_t\|_{r+1} ds).$$

Proof. According to Lemma 4.5, we obtain

$$\|\rho\| \le Ch^{r+1}\|u\|_{r+1} \le Ch^{r+1}(\|\psi\|_{r+1} + \int_0^t \|u_t\|_{r+1} ds). \tag{49}$$

In order to estimate θ, notice that

$$\begin{aligned} (\theta_t,\chi) + a_s(\theta,\chi) &= (u_{h,t},\chi) + a_s(u_h,\chi) - (E_h u_t,\chi) - a_s(E_h u,\chi) \\ &= (f,\chi) - (E_h u_t,\chi) - a_s(E_h u,\chi) \\ &= (f,\chi) + (\nabla\cdot(a\nabla u),\chi) - (E_h u_t,\chi) \\ &= (u_t,\chi) - (E_h u_t,\chi) \\ &= (Q_h u_t,\chi) - (E_h u_t,\chi) \\ &= -(\rho_t,\chi), \end{aligned} \tag{50}$$

where the property that the operator E_h commutes with time differentiation is applied. Since θ belongs to V_h, we choose $\chi = \theta$ in (50) and conclude

$$(\theta_t,\theta) + a_s(\theta,\theta) = -(\rho_t,\chi), \qquad t > 0. \tag{51}$$

Removing the nonnegative term $a_s(\theta,\theta)$ to obtain

$$\frac{1}{2}\frac{d}{dt}\|\theta\|^2 = \|\theta\|\frac{d}{dt}\|\theta\| \le \|\rho_t\|\|\theta\|,$$

which leads to

$$\|\theta(t)\| \le \|\theta(0)\| + \int_0^t \|\rho_t\| ds. \tag{52}$$

Using Lemma 4.5, we find

$$\begin{aligned} \|\theta(0)\| = \|u_h(0) - E_h u(0)\| &= \|u_h(0) - Q_h u(0)\| + \|E_h u(0) - Q_h u(0)\| \\ &\le \|u_h(0) - Q_h u(0)\| + Ch^{r+1}\|\psi\|_{r+1}, \end{aligned} \tag{53}$$

where the following inequality is applied

$$\|\rho_t\| \le Ch^{r+1}\|u_t\|_{r+1}. \tag{54}$$

The desired bound for $\theta(t)$ now follows. □

Theorem 4.7. *Let the assumptions in Theorem 4.6 hold. Then, there exists a constant $C > 0$ independent of the mesh size h such that*

$$\begin{aligned} |||u_h(t) - Q_h u(t)|||^2 \le & \frac{2\beta}{\alpha} ||| u_h(0) - Q_h u(0) |||^2 + Ch^{2r}(\|\psi\|_{r+1}^2 + \|u\|_{r+1}^2) \\ & + Ch^{2(r+1)} \int_0^t \|u_t\|_{r+1}^2 ds. \end{aligned}$$

Proof. As in the proof of Theorem 4.6, we write the error in the form (48). Here by Lemma 4.5, there holds

$$|||\rho(t)||| \le Ch^r\|u\|_{r+1}. \tag{55}$$

In order to estimate $|||\theta|||$, we choose $\chi = \theta_t$ in (50) to obtain

$$(\theta_t, \theta_t) + a_s(\theta, \theta_t) = -(\rho_t, \theta_t), \qquad t > 0,$$

which gives

$$\|\theta_t\|^2 + \frac{1}{2}\frac{d}{dt}a_s(\theta, \theta) = -(\rho_t, \theta_t) \le \frac{1}{2}\|\rho_t\|^2 + \frac{1}{2}\|\theta_t\|^2.$$

As a consequence, we have

$$\frac{d}{dt}a_s(\theta, \theta) \le \|\rho_t\|^2.$$

And integrating with respect to time t to obtain

$$a_s(\theta, \theta) \le a_s(\theta(0), \theta(0)) + \int_0^t \|\rho_t\|^2 ds,$$

where

$$\begin{aligned} a_s(\theta(0), \theta(0)) = & \sum_{T\in\mathcal{T}_h} (a\nabla_d(u_h(0) - E_h u(0)), \nabla_d(u_h(0) - E_h u(0))) \\ & + s\big((u_h(0) - E_h u(0)), (u_h(0) - E_h u(0))\big). \end{aligned}$$

Using the coercivity and the boundedness of the bilinear form (16), we obtain

$$\begin{aligned} \alpha ||| \theta|||^2 \le & a_s(\theta, \theta) \le a_s(\theta(0), \theta(0)) + \int_0^t \|\rho_t\|^2 ds \\ \le & 2a_s(u_h(0) - Q_h u(0), u_h(0) - Q_h u(0)) \\ & + 2a_s(E_h u(0) - Q_h u(0), E_h u(0) - Q_h u(0)) + \int_0^t \|\rho_t\|^2 ds \\ \le & 2\beta(|||u_h(0) - Q_h u(0) |||^2 + ||| E_h u(0) - Q_h u(0)|||^2) + \int_0^t \|\rho_t\|^2 ds, \end{aligned}$$

which gives together with (46) and (54)

$$|||\theta||| \le \frac{2\beta}{\alpha} ||| u_h(0) - Q_h u(0) |||^2 + C(h^{2r}\|\psi\|_{r+1}^2 + h^{2(r+1)} \int_0^t \|u_t\|_{r+1}^2 ds).$$

This ends the proof. □

Theorem 4.8. *Let $u \in H^{r+1}(\Omega)$. Then there exists a constant $C > 0$ independent of the mesh size h such that*

$$\begin{aligned}&\|U^n - Q_h u(t_n)\|\\ \le&\|U^0 - Q_h u(0)\| + Ch^{r+1}\big(\|\psi\|_{r+1} + \int_0^{t_n} \|u_t\|_{r+1} ds\big) + Ck\int_0^{t_n} \|u_{tt}\| ds.\end{aligned}$$

Proof. Similarly, we write

$$U^n - Q_h u(t_n) = (U^n - E_h u(t_n)) + (E_h u(t_n) - Q_h u(t_n)) = \theta^n + \rho^n, \tag{56}$$

where $\rho^n = \rho(t_n)$ is bounded as in the following way

$$\|\rho^n\| = \|E_h u(t_n) - Q_h u(t_n)\| \le Ch^{r+1}(\|\psi\|_{r+1} + \int_0^{t_n} \|u_t\|_{r+1} ds).$$

In order to bound θ^n, we use

$$\begin{aligned}(\bar\partial\theta^n, \chi) + a_s(\theta^n, \chi) &= (\bar\partial U^n, \chi) + a_s(U^n, \chi) - (\bar\partial E_h u(t_n), \chi) - a_s(E_h u(t_n), \chi)\\ &= (f(t_n), \chi) - (\bar\partial E_h u(t_n), \chi) - a_s(E_h u(t_n), \chi)\\ &= (f(t_n), \chi) + (\nabla\cdot(a\nabla u(t_n)) - (\bar\partial E_h u(t_n), \chi)\\ &= (u_t(t_n), \chi) - (\bar\partial E_h u(t_n), \chi)\\ &= (u_t(t_n) - \bar\partial u(t_n), \chi) + (\bar\partial u(t_n) - \bar\partial E_h u(t_n), \chi),\end{aligned}$$

i.e.

$$(\bar\partial\theta, \chi) + a_s(\theta^n, \chi) = (\omega^n, \chi), \tag{57}$$

where

$$\omega^n = (u_t(t_n) - \bar\partial u(t_n)) + (\bar\partial u(t_n) - \bar\partial E_h u(t_n)) = \omega_1^n + \omega_4^n,$$

with $\omega_4^n = \bar\partial u(t_n) - \bar\partial E_h u(t_n)$. Choosing $\chi = \theta^n$ in (57), we have

$$(\bar\partial\theta^n, \theta^n) \le \|\omega^n\|\|\theta^n\|.$$

Consequently, we obtain

$$\|\theta^n\|^2 - (\theta^{n-1}, \theta^n) \le k\|\omega^n\|\|\theta^n\|,$$

or, equivalently

$$\|\theta^n\| \le \|\theta^{n-1}\| + k\|\omega^n\|.$$

It follows by induction

$$\|\theta^n\| \le \|\theta^0\| + k\sum_{j=1}^n \|\omega^j\| \le \|\theta^0\| + k\sum_{j=1}^n \|\omega_1^j\| + k\sum_{j=1}^n \|\omega_4^j\|.$$

As in (53), $\theta^0 = \theta(0)$ is bounded . According to (40), we obtain

$$k\sum_{j=1}^n \|\omega_1^j\| \le \sum_{j=1}^n \|\int_{t_{j-1}}^{t_j} (s - t_{j-1}) u_{tt}(s) ds\| \le k\int_0^{t_n} \|u_{tt}\| ds. \tag{58}$$

Noting that

$$\omega_4^j = \bar\partial u(t_j) - \bar\partial E_h u(t_j) = (I - E_h)k^{-1}\int_{t_{j-1}}^{t_j} u_t ds = k^{-1}\int_{t_{j-1}}^{t_j} (I - E_h) u_t ds, \tag{59}$$

and by (54) we have

$$k\sum_{j=1}^n \|\omega_4^j\| \le Ch^{r+1}\int_0^{t_n} \|u_t\|_{r+1} ds.$$

The proof is complete. □

Theorem 4.9. *Let $u \in H^{r+1}(\Omega)$. Then there exists a constant $C > 0$ independent of the mesh size h such that*

$$|||U^n - Q_h u(t_n)|||^2 \leq 2 \, ||| \, U^0 - Q_h u(0) \, |||^2 + C\{h^{2r}(\|\psi\|_{r+1}^2 + \|u\|_{r+1}^2 ds) + h^{2(r+1)} \int_0^{t_n} \|u\|_{r+1}^2 ds + k^2 \int_0^{t_n} \|u_{tt}\|^2 ds\}.$$

Proof. Let θ^n and ρ^n be defined as in (56). By (47), we have

$$|||\rho^n|||^2 \leq Ch^{2r}\|u\|_{r+1}^2.$$

In order to estimate $|||\theta^n|||^2$, we choose $\chi = \bar{\partial}\theta^n$ in (57), we then get

$$(\bar{\partial}\theta^n, \bar{\partial}\theta^n) + a_s(\theta^n, \bar{\partial}\theta^n) = \|\bar{\partial}\theta^n\|^2 + \frac{1}{2}\bar{\partial}a_s(\theta^n, \theta^n) + \frac{1}{2}a_s(\bar{\partial}\theta^n, \bar{\partial}\theta^n) = (\omega^n, \bar{\partial}\theta^n) \leq \frac{1}{2}\|\omega^n\|^2 + \frac{1}{2}\|\bar{\partial}\theta^n\|^2.$$

By the coercivity of the bilinear form, we have

$$\alpha\bar{\partial} \, ||| \, \theta^n |||^2 \leq \|\omega^n\|^2.$$

Thus, we get

$$|||\theta^n|||^2 \leq |||\theta^{n-1} \, |||^2 + \frac{k}{\alpha}\|\omega^n\|^2.$$

It gives by induction

$$|||\theta^n|||^2 \leq |||\theta^0 \, |||^2 + \frac{k}{\alpha}\sum_{j=1}^{n}\|\omega^j\|^2 \leq |||\theta^0 \, |||^2 + \frac{2k}{\alpha}\sum_{j=1}^{n}\|\omega_1^j\|^2 + \frac{2k}{\alpha}\sum_{j=1}^{n}\|\omega_4^j\|. \tag{60}$$

Similar to (58) and (59), we have

$$k\sum_{j=1}^{n}\|\omega_1^j\|^2 \leq \sum_{j=1}^{n}\|\int_{t_{j-1}}^{t_j}((s - t_{j-1})u_{tt}(s))^2 ds\| \leq k^2 \int_0^{t_n} \|u_{tt}\|^2 ds, \tag{61}$$

and

$$\begin{aligned} k\sum_{j=1}^{n}\|\omega_4^j\|^2 &= k\sum_{j=1}^{n}\int_\Omega (k^{-1}\int_{t_{j-1}}^{t_j} \rho_t ds)^2 dx \\ &\leq \sum_{j=1}^{n}\int_\Omega \int_{t_{j-1}}^{t_j} \rho_t^2 ds) dx \leq \sum_{j=1}^{n}\int_{t_{j-1}}^{t_j} \|\rho_t\|^2 ds) \\ &\leq \int_0^{t_n} \|\rho_t\|^2 ds) \leq Ch^{2(r+1)}\int_0^{t_n} \|u\|_{r+1}^2 ds. \end{aligned} \tag{62}$$

It is easy to verify that

$$|||\theta^0|||^2 \leq 2(|||U^0 - Q_h u(0) \, |||^2 + Ch^{2r}\|\psi\|_{r+1}^2), \tag{63}$$

which together with our estimates (60), (61) and (62) completes the proof. □

5. Numerical Experiment

In this section, we shall present some numerical results for problems (1)-(3). The domain Ω is chosen to be the unit square $[0, 1] \times [0, 1]$, and the time interval is $[0, 1]$.

In the example, the uniform triangle partition is employed, where h denotes the spatial mesh size. We also use the uniform partition for the time discretization with τ denoting the time step. The degree of polynomial r is set to be 2, i.e. the finite element space is

$$V_h = \{\{v_0, v_b\} : v_0 \in P_2(T), v_b \in P_2(e)\}.$$

The analytic solution is chosen to be

$$u = e^{-t}\sin(\pi x)\sin(\pi y),$$

the boundary condition and the source term can be calculated accordingly.

First, we fix a sufficient small mesh size $h = 1/256$ and we obtain the error order with respect to the time step in Table 1 which conforms well the theoretical analysis.

TABLE 1. Numerical results for $h = 1/256$.

τ	$\|\!\|\!\|e_h\|\!\|\!\|$	order	$\|e_0\|$	order
1/4	5.7524e-03		1.2541e-03	
1/8	2.7501e-03	1.0647	5.9954e-04	1.0647
1/16	1.3443e-03	1.0326	2.9303e-04	1.0328
1/32	6.6492e-04	1.0156	1.4487e-04	1.0163
1/64	3.3123e-04	1.0053	7.2028e-05	1.0081
1/128	1.6642e-04	0.9930	3.5913e-05	1.0040

Then, we fix a sufficient small time step $\tau = 1/8192$ and we get the error order with respect to the mesh size in Table 2 which also conforms well the theoretical analysis.

TABLE 2. Numerical results for $\tau = 1/8192$.

h	$\|\!\|\!\|e_h\|\!\|\!\|$	order	$\|e_0\|$	order
1/4	9.6144e-02		8.2838e-03	
1/8	2.4149e-02	1.9932	1.0306e-03	3.0068
1/16	6.0448e-03	1.9982	1.2862e-04	3.0022
1/32	1.5118e-03	1.9995	1.6091e-05	2.9988
1/64	3.7800e-04	1.9998	2.0924e-06	2.9431

Acknowledgements

The authors would like to show their heartfelt gratitude to the anonymous referees since their comments and suggestions improved the paper a lot. This work was supported by NSFC Grants Nos.11201061, 11271065,11371171,11271157, U1530116, the Program for New Century Excellent Talents in University of Ministry of Education of China, China Postdoctoral Science Foundation 2012M52065 and the Science and Technology Development Planning of Jilin Province 20140520058JH.

References

[1] F. Brezzi, J. Douglas, Jr. and L. D. Marini, Two families of mixed finite element for second order elloptic problem, Numer. Math., 47 (1985), pp.217-235.
[2] P. Chatzipantelidis, R. D. Lazarov and V. Thomée, Error estimates for the finite volume element method for parabolic equations in convex polygonal domains, Numer. Methods Partial Differential Equations, 20(2003), pp.650-674.
[3] P. Chatzipantelidis, R. D. Lazarov and V. Thomée, Some error estimates for the lumped mass finite element method for a parabolic problem, Math. Comp., 81(2012), pp.1-20.
[4] P. Chatzipantelidis, R. D. Lazarov and V. Thomée, Some error estimates for the finite volume element method for a parabolic problem, Comput. Meth. Appl. Math., 13(2013), pp.251-279.
[5] K. Erilksson and C. Johnson, Adaptive finite element methods for parabolic problems IV: Nonlinear Problems, SIAM J. Numer. Anal., 32(1995), pp.1729-1749.

[6] S. Larsson, V. Thomée and L. B. Wahlbin, Numerical solution of parabolic integro-differential equations by the discontinuous Galerkin method, Math. Comp., 67(1998), pp.45-71.
[7] Q. H. Li and J. Wang, Weak Galerkin finite element methods for parabolic equations, Numer. Methods Partial Differential Equations, 29(2013), pp. 2004-2024.
[8] X. Ma, S. Shu and A. Zhou, Symmetric finite volume discretizations for parabolic problems, Compt. Methods Appl. Mech. Eng., 192(2003), pp.4467-4485.
[9] Ch. G. Makridakis and I. Babuška, On the stability of the discontinuous Galerkin method for the heat equation, SIAM J. Numer. Anal., 34(1997), pp.389-401.
[10] L. Mu, J. Wang, G. Wei and X. Ye, Weak Galerkin finite element methods on polytopal meshes, arXiv:1204.3655v2, 2012 .
[11] L. Mu, J. Wang, G. Wei, X. Ye and S. Zhao, Weak Galerkin methods for second order elliptic interface problems, J. Comput. Phys., 250(2013), pp.106-125.
[12] L. Mu, J. Wang, X. Ye and S. Zhao, A numerical study on the weak Galerkin method for the Helmholtz equation, Commun. Comput. Phys., 15(2014), pp. 1461-1479.
[13] P. A. Raviart and J. M. Thomas, A mixed finite element method for 2-nd order elliptic problems, Mathematical Aspects of Finite Element Method, I. Galligani, E. Magenes, eds., Lecture Notes in Mathematics 606, Springer-Verlag, New York, 1977.
[14] R. K. Sinha, R. E. Ewing and, R. D. Lazarov, Some new error estimates of a semidiscrete finite volume element method for a parabolic integro-differential equation with nonsmooth initial data, SIAM J. Numer. Anal., 43 (2006), pp.2320-2344.
[15] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, 2nd ed., Springer-Verlag, New York, 2006.
[16] J. Wang and X. Ye, A weak Galerkin mixed finite elememt method for second order problems, arXiv:1202.3655v1, 2012.
[17] J. Wang and X. Ye, A weak Galerkin finite elememt method for second-order problems, J. Comput. Appl. Math., 241 (2013) pp.103-115.
[18] M. F. Wheeler, A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations, SIAM J. Numer. Anal., 10 (1973), pp.723-759.
[19] M. Yang and J. Liu, A quadratic finite volume element method for parabolic problems on quadrilateral meshes, IMA J. Numer. Anal., 31(2011), pp.1038-1061.

a. School of Mathematics, Jinlin University, Changchun 130012, China
b. School of Information and Technology, Jilin Agricultural University, Changchun 130118, China
E-mail: zhanghq_@126.com

School of Mathematics, Jinlin University, Changchun 130012, China
E-mail: zouyk@jlu.edu.cn

Corresponding author
School of Mathematics and Statistics, Northeast Normal University, Changchun 130024, China
E-mail: yxxu@nenu.edu.cn

School of Mathematics, Jinlin University, Changchun 130012, China
E-mail: diql@mails.jlu.edu.cn

School of Mathematics, Jinlin University, Changchun 130012, China
E-mail: yueh@jlu.edu.cn