Journal of Computational Mathematics
Vol.41, No.6, 2023, 1246–1280.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2203-m2021-0031

A MODIFIED WEAK GALERKIN FINITE ELEMENT METHOD FOR SINGULARLY PERTURBED PARABOLIC CONVECTION-DIFFUSION-REACTION PROBLEMS*

Suayip Toprakseven[1)]
Faculty of Engineering, Department of Computer Science, Artvin Çoruh University, Artvin, 08100, Turkey
Email: topraksp@artvin.edu.tr
Fuzheng Gao
School of Mathematics, Shandong University, Jinan 250100, China
Email: fzgao@sdu.edu.cn

Abstract

In this work, a modified weak Galerkin finite element method is proposed for solving second order linear parabolic singularly perturbed convection-diffusion equations. The key feature of the proposed method is to replace the classical gradient and divergence operators by the modified weak gradient and modified divergence operators, respectively. We apply the backward finite difference method in time and the modified weak Galerkin finite element method in space on uniform mesh. The stability analyses are presented for both semi-discrete and fully-discrete modified weak Galerkin finite element methods. Optimal order of convergences are obtained in suitable norms. We have achieved the same accuracy with the weak Galerkin method while the degrees of freedom are reduced in our method. Various numerical examples are presented to support the theoretical results. It is theoretically and numerically shown that the method is quite stable.

Mathematics subject classification: 65N15, 65N30, 35J50.
Key words: The modified weak Galerkin finite element method, Backward Euler method, Parabolic convection-diffusion problems, Error estimates.

1. Introduction

In this paper, we propose a modified weak Galerkin finite element method (MWG-FEM) for the following parabolic convection-diffusion problem:

$$
\begin{aligned}
&\partial_t u-\varepsilon\Delta u+\nabla\cdot(\mathbf{b}u)+cu=f && \text{in} \quad Q_T=\Omega\times(0,T],\\
&u=0 && \text{on} \quad \partial\Omega\times(0,T],\\
&u=u_0 && \text{in} \quad \Omega\times\{0\},
\end{aligned}
\tag{1.1}
$$

where $\varepsilon\in(0,1]$ is a small parameter and Ω is a bounded polygonal domain in $\mathbb{R}^2$ with the boundary $\partial\Omega$, $\partial_t u=\frac{\partial u}{\partial t}$ and $u_0\in L^2(\Omega)$. For the well-posedness of the problem [24], we assume that $\mathbf{b},c$ and f are smooth functions, $\mathbf{b}\in[W^{1,\infty}(\Omega)]^2$ and for some constant a_0 such that

$$
c+\frac{1}{2}\nabla\cdot\mathbf{b}\geq a_0>0. \tag{1.2}
$$

* Received January 28, 2021 / Revised version received June 11, 2021 / Accepted March 11, 2022 / Published online December 7, 2022 /
1) Corresponding author

Convection-diffusion equations are commonly used to describe a wide range of differential equations arising from the mathematical modeling of real word problems in science and engineering involving fluid, petroleum simulation, groundwater contamination and gas dynamics [2,3,30], etc. Applications generally involve time-dependent convection-dominated problems for the mathematical modeling of physical processes. It is well known that the solution of the singularly perturbed convection-diffusion problems possess boundary or interior layers. It is well known that these layers lead to unsatisfactory numerical solutions with non-physical oscillations when the conventional numerical methods such as finite difference (FD) methods and the standard finite element methods are applied. To recover these non-physical oscillations, some stabilization methods have been proposed over the last decades, including streamline-upwind Petrov-Galerkin (SUPG) methods proposed by Hughes and Brooks [4], local projection stabilization method [15, 19] and the interior penalty method [35]. However, there are some disadvantages of these methods for convection-dominated problems. For instance, the popular SUPG methods have the stabilization term which includes many terms for the time dependent problems. Moreover, they produce overshoots and undershoots near the layer region. A large of papers has been devoted to the numerical methods for convection-diffusion problems on some layer adapted meshes in the literature [18]. Unfortunately, the location of the layer must be known in prior in order to use the layer adapted meshes. The layers may move as time varies in the parabolic convection-diffusion problems. This leads to use fitted operator methods for the numerical solutions of unsteady convection-diffusion-reaction equations.

Wang and Ye [32] first introduced the weak Galerkin finite element method (WG-FEM) and analyzed for numerical solution of second order differential equations. The WG-FEMs introduce a space of weak functions, weak gradient and weak divergence on the space of completely discontinuous piecewise polynomials. The weak functions in WG-FEMs consist of the form $u = \{u_0, u_b\}$ with $u = u_0$ inside of the element and $u = u_b$ on the boundary of the element. Later on, WG finite element methods have further been presented for a large variety of PDEs including the implementation results [21], parabolic problems [16], the Helmholtz equations with high wave numbers in [22] and the time fractional reaction-diffusion-convection problems in [27]. The weak gradient and weak divergence operators have been introduced for convection-dominated problems in [5] and [17]. The WG-FEM has been studied and analyzed for time-dependent convection-diffusion equations with convection term in non-conservation form based on these newly defined operators [34]. While the formulation of WG-FEM is simple and parameter-free, it adds more degrees of freedom since it has two components for each function in the approximation space. In order to reduce the degrees of freedom in the formulation of the WG-FEM, a modified WG-FEM (MWG-FEM) introduced in [31] eliminates u_b from the space of weak functions and uses the average $\{u_0\}$ of the u_0 on the boundary of element. As a result, the weak functions in the MWG-FEM of the form $u = \{u_0, \{u_0\}\}$ and for simplicity we denote by u. As the WG-FEM, the MWG-FEM is a parameter free method and it has the same degrees of freedom as the discontinuous Galerkin (DG) methods. In other words, the MWG-FEM inherits from the properties of the WG-FEM with the reduced number of unknowns in the associated discrete systems. Compared to DG methods, the formulation of the MWG-FEM is simple, symmetric and the resulting system is positive definite while they have the same same finite element space and there is no need a large penalization parameter for the MWG-FEM. MWG-FEMs have been further developed for a variety of PDEs such as convection-diffusion problems [12], parabolic equations [11], Stokes equations [20,26], convection-diffusion problems in one dimension [28] and in higher dimension with weakly imposed boundary condition [9].

Superconvergence approximation of the MWG-FEM is also presented in [29]. This paper aims to introduce a MWG-FEM for time dependent convection-diffusion problems. This modified scheme has less the degree of freedom than WG-FEM proposed in [34] while the accuracy remains the same. In this paper, we approximate the convection term by a modified weak divergence operator and introduce a simple upwinding-type stabilizer for the convection term and we do not require extra conditions on the convection coefficient.

The rest of this paper is organized as follows. In Section 2, a MWG-FEM is introduced and semi-discrete MWG-FEM and fully discrete MWG-FEM are proposed. Stability analyses of the semi-discrete and backward Euler difference time discrete MWG finite element schemes are established. Some error equations are derived. Error analyses and an optimal convergence result in the energy norm and a suboptimal order error estimate in L^2-norm for both schemes are given in Section 3. The optimal convergence order in L^2-norm is established in Section 4. Numerical results are given in Section 5 to verify the theoretical findings.

Throughout this article, we use C or with subscript such as C_1, C_2 for a generic constant independent of ε, time step size δ and mesh step size h unless otherwise stated.

2. MWG-FEMs

We use the standard notation for Sobolev spaces $H^l(\Omega)$ for any domain $\Omega \subset \mathbb{R}^2$ with $l \geq 0$. The inner product, semi-norm and norm in $H^l(S)$ are given $(\cdot,\cdot)_{l,S}$, $\;|\cdot|_{l,S}$ and $\|\cdot\|_{l,S}$ for subset $S \subset \Omega$, respectively. We sometimes skip the subset S when $S = \Omega$ and use the notation $|\cdot|_l$ and $\|\cdot\|_l$. Moreover, we denote by $\|\cdot\|_\infty$ the norm on $L^\infty(\Omega)$ and $\|\cdot\|$ the norm on $L^2(\Omega)$.

The variational form for the problem (1.1) is to seek $u \in H_0^1(\Omega), t \in [0,T]$ such that the following equations hold:

$$\begin{aligned} &(\partial_t u, v) + A(u,v) = (f,v), \quad \forall v \in H_0^1(\Omega), \quad t > 0,\\ &u(x,0) = u_0(x), \qquad\qquad\quad x \in \Omega, \end{aligned} \tag{2.1}$$

where $A(u,v) = \varepsilon(\nabla u, \nabla v) + (\nabla\cdot(\mathbf{b}u), v) + (cu, v)$ and $(\cdot,\cdot)$ is the standard inner product in $L^2(\Omega)$.

Let $\mathcal{T}_h$ be a partition of the domain Ω consisting of polygons which are closed and simply connected elements. The set of all edges in $\mathcal{T}_h$ is denoted by $\mathcal{E}_h$ and the set of all interior edges by $\mathcal{E}_h^0 = \mathcal{E}_h \backslash \partial\Omega$. Denote by h_T the diameter of elements $T \in \mathcal{T}_h$ and $h = \max_{T\in\mathcal{T}_h} h_T$. We follow the shape regularity assumptions $A1 - A4$ for the partition $\mathcal{T}_h$ detailed as in [17].

Let T_1, T_2 be two adjacent triangles with common edge e and unit outward normal vectors $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ on e associated with T_1 and T_2, respectively. The average $\{\cdot\}$ and jump $[\cdot]$ of a scalar valued function u on e are defined by

$$\{u\}_e = \begin{cases} \frac{1}{2}(u|_{T_1} + u|_{T_2}), & e \in \mathcal{E}_h^0,\\ u, & e \in \partial\Omega, \end{cases} \qquad [u]_e = \begin{cases} u|_{T_1}\boldsymbol{n_1} + u|_{T_2}\boldsymbol{n_2}, & e \in \mathcal{E}_h^0,\\ u\boldsymbol{n}, & e \in \partial\Omega. \end{cases}$$

Similarly, we define the average and jump operators for a vector valued function $\mathbf{v}$

$$\{\boldsymbol{v}\}_e = \begin{cases} \frac{1}{2}(\boldsymbol{v}|_{T_1} + \boldsymbol{v}|_{T_2}), & e \in \mathcal{E}_h^0,\\ \boldsymbol{v}, & e \in \partial\Omega, \end{cases} \qquad [\boldsymbol{v}]_e = \begin{cases} \boldsymbol{v}|_{T_1}\cdot\boldsymbol{n_1} + \boldsymbol{v}|_{T_2}\cdot\boldsymbol{n_2}, & e \in \mathcal{E}_h^0,\\ \boldsymbol{v}\cdot\boldsymbol{n}, & e \in \partial\Omega. \end{cases}$$

For a given integer $k \geq 1$, we define the finite element space $S_h(k)$ associated with $\mathcal{T}_h$ as follows:

$$S_h(k) = \{u \in L^2(\Omega) : u|_T \in \mathbb{P}_k(T), \forall T \in \mathcal{T}_h\}, \tag{2.2}$$

and its subspace $S_h^0(k)$ as

$$S_h^0(k) = \{u \in S_h(k) : u|_e = 0, e \in \partial\Omega\}, \tag{2.3}$$

where $\mathbb{P}_k(T)$ is the set of polynomials on T of degree at most k.

For any function $u \in S_h(k)$, *the modified weak gradient* $\nabla_w u \in [\mathbb{P}_{k-1}(T)]^2$ is defined on T as the unique polynomial satisfying the following equation:

$$\big(\nabla_w u, \boldsymbol{\tau}\big)_T = -\big(u, \nabla \cdot \boldsymbol{\tau}\big)_T + \big\langle \{u\}, \boldsymbol{\tau} \cdot \mathbf{n}\big\rangle_{\partial T}, \quad \forall \boldsymbol{\tau} \in [\mathbb{P}_{k-1}(T)]^2, \tag{2.4}$$

where $\mathbf{n}$ is the unit outward normal to ∂T and $\big(\cdot,\cdot\big)_T$ and $\big\langle\cdot,\cdot\big\rangle_{\partial T}$ are the L^2 inner products on T and ∂T, respectively.

For any function $u \in S_h(k)$, *the modified weak divergence* $\nabla_w \cdot (\mathbf{b}u) \in \mathbb{P}_{k-1}(T)$ related to $\mathbf{b}$ is defined on T as the unique polynomial satisfying the following equation:

$$\big(\nabla_w \cdot (\mathbf{b}u), w\big)_T = -\big(\mathbf{b}u, \nabla w\big)_T + \langle \{u\}, \mathbf{b} \cdot \boldsymbol{n} w\rangle_{\partial T}, \quad \forall w \in \mathbb{P}_{k-1}(T). \tag{2.5}$$

Remark 2.1. This newly defined modified weak gradient is different from the weak gradient operator defined in [10]. This modified definition replaces the values u_b on the boundary by the average operator $\{u\}$ of u on the boundary of T. This reduces the degree of freedom for the problem, that is, the unknown coefficients in the system are reduced.

Remark 2.2. If u is continuous in Ω, then we have

$$\{u\} = u \text{ on } \partial T, \quad \forall T \in \mathcal{T}_h.$$

We see from the definition of weak gradient in (2.4) that

$$\begin{aligned}\int_T \nabla_w u v \, dx &= -\int_T u \nabla \cdot v \, dx + \int_{\partial T} u v \cdot \mathbf{n} \, ds \\ &= \int_T \nabla u v \, dx \quad \forall v \in [\mathbb{P}_{k-1}(T)]^2,\end{aligned}$$

which implies the modified weak gradient in fact is the L^2 projection of the classical gradient operator on the space of polynomials. Thus, we have $\nabla_w u = \nabla u$ when $u \in \mathbb{P}_k(\Omega)$.

Similarly, if u is continuous in Ω, then from the definition of modified weak divergence given by (2.5), we have

$$\begin{aligned}\int_T \nabla_w \cdot u w \, dx &= -\int_T u \nabla w \, dx + \int_{\partial T} u w \mathbf{n} \, ds \\ &= \int_T \nabla \cdot u w \, dx \quad \forall w \in \mathbb{P}_{k-1}(T),\end{aligned}$$

showing that the modified weak divergence is the L^2 projection of the classical divergence operator on the space $[\mathbb{P}_k(\Omega)]^2$. Thus we have $\nabla_w \cdot u = \nabla \cdot u$ when $u \in [\mathbb{P}_k(\Omega)]^2$.

In order to analyze and investigate the proposed method, we introduce the local L^2 projection. We first define the local projection Q_h given by

$$Q_h : L^2(T) \to \mathbb{P}_k(T), \quad (Q_h q - q, p)_T = 0, \quad \forall p \in \mathbb{P}_k(T) \tag{2.6}$$

for each element $T \in \mathcal{T}_h$. The other projection is the L^2 projection on the local weak gradient space defined by

$$\mathbb{Q}_h : [L^2(T)]^2 \to [\mathbb{P}_{k-1}(T)]^2, \quad (\mathbb{Q}_h \boldsymbol{\tau} - \boldsymbol{\tau}, \boldsymbol{\sigma})_T = 0, \quad \forall \boldsymbol{\sigma} \in [\mathbb{P}_{k-1}(T)]^2 \tag{2.7}$$

for each element T. The following error estimates are standard and the proof can be found in [11].

Lemma 2.1 ([17]). *Let $\mathcal{T}_h$ be a finite element partition of Ω satisfying the regularity requirements. Then, for any $u \in H^{1+k}(\Omega)$ with $k > 0$, we have*

$$\sum_{T \in \mathcal{T}_h} (\|u - Q_h u\|_T^2 + h_T^2 \|\nabla(u - Q_h u)\|_T^2) \leq C h^{2(k+1)} \|u\|_{1+k}^2, \tag{2.8}$$

$$\sum_{T \in \mathcal{T}_h} (\|\nabla u - \mathbb{Q}_h(\nabla u)\|_T^2 + h_T^2 |\nabla u - \mathbb{Q}_h(\nabla u)|_{1,T}^2) \leq C h^{2k} \|u\|_{1+k}^2. \tag{2.9}$$

Example 2.1 ([12]). Let T be the reference triangle element $\triangle ABC$ with nodes $A(0,0)$, $B(1,1)$ and $C(0,1)$. Suppose that $u|_T = (1,1)$ and $u = 0$ on $\Omega \backslash T$. Then we have $\{u\}|_{e_j} = (\frac{1}{2}, \frac{1}{2})$ where $e_1 = \overline{AB}$, $e_2 = \overline{BC}$ and $e_3 = \overline{CA}$.

a. If $k = 1$, then we have $\nabla_w \cdot u|_T = 0$.

b. If $k = 2$ then we have $\nabla_w \cdot u|_T = 6 - 6x - 6y$.

For the future reference, we use the following notations:

$$\begin{aligned}
(u, v) &= \sum_{T \in \mathcal{T}_h} (u, v)_T = \sum_{T \in \mathcal{T}_h} \int_T uv \, dx, \\
\langle u, v \rangle &= \sum_{T \in \mathcal{T}_h} \langle u, v \rangle_{\partial T} = \sum_{T \in \mathcal{T}_h} \int_{\partial T} uv \, ds.
\end{aligned}$$

For L^2- norm, we suppress the subscript and use the notation $\| \cdot \|$ in the sequel.

For $u_h, v_h \in S_h(k)$, we define a bilinear form on $S_h(k)$ as follows:

$$a(u_h, v_h) = \varepsilon(\nabla_w u_h, \nabla_w v_h) + (\nabla_w \cdot (\mathbf{b} u_h), v_h) + (c u_h, v_h) + s_c(u_h, v_h) + s_d(u_h, v_h), \tag{2.10}$$

where

$$\begin{aligned}
s_c(u, v) &= \sum_{T \in \mathcal{T}_h} \langle \mathbf{b} \cdot \mathbf{n}(u - \{u\}), v - \{v\} \rangle_{\partial_+ T}, \\
s_d(u, v) &= \sum_{e \in \mathcal{E}_h} \varepsilon h_e^{-1} \langle [u], [v] \rangle_e,
\end{aligned}$$

and

$$\partial_+ T = \{y \in \partial T : \mathbf{b}(y) \cdot \mathbf{n}(y) \geq 0\}.$$

Remark 2.3. The stabilization terms $s_c(\cdot, \cdot)$ and $s_d(\cdot, \cdot)$ defined in this paper is different from the stabilization terms defined in [12,17].

We then propose the following modified weak Galerkin finite scheme:

Algorithm 2.1. The Modified Weak Galerkin Scheme

A modified weak Galerkin solution to (1.1) is to find $u_h(t) \in S_h^0(k)$ such that

$$\begin{aligned} &\big(\partial_t u_h, v_h\big) + a(u_h, v_h) = \big(f, v_h\big), \quad \forall v \in S_h^0(k), \quad t > 0, \\ &u_h(x,0) = Q_h u_0(x), \qquad\qquad\qquad x \in \Omega. \end{aligned} \tag{2.11}$$

Let $dim(S_h(k)) = M$ and $dim(S_h^0(k)) = N$. If $\{\phi_n(x) : n = 1, \ldots, N, N+1, \ldots, M\}$ is basis functions of $S_h(k)$ and $\{\phi_n(x) : n = 1, \ldots, N\}$ is basis functions of $S_h^0(k)$, then the matrix form of the MWG-FEM given by (2.11) can be written as

$$\mathbf{M}\frac{d\mathbf{C}}{dt} + \mathbf{AC} = \mathbf{F}, \tag{2.12}$$

where $\mathbf{M}$ is the mass matrix given by

$$\mathbf{M} = \big(\phi_n, \phi_m\big), \quad n, m = 1, \ldots, N,$$

the stiffness matrix $\mathbf{A}$

$$\mathbf{A} = a(\phi_n, \phi_m), \quad n, m = 1, \ldots, N,$$

the forcing vector $\mathbf{F}$

$$\mathbf{F} = \left[-\sum_{n=N+1}^{M} \frac{dc_n}{dt}(\phi_n, \phi_1) + (f, \phi_1), \ldots, -\sum_{n=N+1}^{M} \frac{dc_n}{dt}(\phi_n, \phi_N) + (f, \phi_N) \right]^T$$

and the constant vector $\mathbf{C}$

$$\mathbf{C} = [c_1, c_2, \ldots, c_N]^T$$

for the numerical approximation

$$u_h(t) = \sum_{n=1}^{N} c_n(t)\phi_n(x) + \sum_{n=N+1}^{M} c_n(t)\phi_n(x).$$

2.1. Stability

The following multiplicative trace inequality will be useful in proving the error estimates.

Lemma 2.2 ([23]). *Let $T \in \mathcal{T}_h$ and $e \in \partial T$. For any $\phi \in H^1(T)$, the following trace inequality holds*

$$\|\phi\|_e^2 \le C\Big(h_T^{-1}\|\phi\|_T^2 + h_T\|\nabla\phi\|_T^2\Big). \tag{2.13}$$

We also frequently use the following identity [20]:

$$\big\langle v - \{v\}, \boldsymbol{\tau} \cdot \boldsymbol{n} \big\rangle = \sum_{e \in \mathcal{E}_h} \big\langle [v], \{\boldsymbol{\tau}\} \big\rangle_e, \tag{2.14}$$

which follows from the equality

$$\big\langle v, \boldsymbol{\tau} \cdot \boldsymbol{n} \big\rangle = \sum_{e \in \mathcal{E}_h^0} \big\langle \{v\}, [[\boldsymbol{\tau}]] \big\rangle_e + \sum_{e \in \mathcal{E}_h} \big\langle [v], \{\boldsymbol{\tau}\} \big\rangle_e.$$

Next, we define an energy norm in $S_h^0(k)$: for $v_h \in S_h^0(k)$

$$|||v_h|||^2 := \|v_h\|_w^2 + \|v_h\|_s^2, \tag{2.15}$$

where

$$\|v_h\|_w^2 = \sum_{T\in\mathcal{T}_h} \varepsilon\|\nabla_w v_h\|_T^2 + s_d^2(v_h, v_h),$$
$$\|v_h\|_s^2 = \sum_{T\in\mathcal{T}_h} \||\mathbf{b}\cdot\mathbf{n}|^{\frac{1}{2}}(v_h - \{v_h\})\|_{\partial T}^2 + \|v_h\|^2.$$

An energy-like norm $|||\cdot|||_\varepsilon$ in the space $S_h^0(k) + H^1(\Omega)$ is defined for $v_h \in S_h^0(k) + H^1(\Omega)$

$$|||v_h|||_\varepsilon^2 = \|v_h\|_{1,h}^2 + \|v_h\|_s^2, \tag{2.16}$$

where

$$\|v_h\|_{1,h}^2 = \sum_{T\in\mathcal{T}_h} \varepsilon\|\nabla v_h\|_T^2 + s_d^2(v_h, v_h).$$

Then, we will show that these two norms are equivalent in the next lemma.

Lemma 2.3. *For any $v_h \in S_h^0(k)$, we have the following*

$$C|||v_h||| \le |||v_h|||_\varepsilon \le C|||v|||.$$

Proof. Let $v_h \in S_h^0(k)$. The definition of the modified weak gradient (2.4) and integration by parts imply that

$$\big(\nabla_w v_h, \boldsymbol{\sigma}\big)_T = \big(\nabla v_h, \boldsymbol{\sigma}\big)_T + \langle\{v_h\} - v_h, \boldsymbol{\sigma}\cdot\boldsymbol{n}\rangle_{\partial T}, \quad \forall\boldsymbol{\sigma} \in [\mathbb{P}_{k-1}(T)]^2. \tag{2.17}$$

Choosing $\boldsymbol{\sigma} = \nabla_w v_h$ in (2.17), we have

$$\|\nabla_w v_h\|_T^2 = \big(\nabla v_h, \nabla_w v_h\big)_T + \langle\{v_h\} - v_h, \nabla_w v_h\cdot\boldsymbol{n}\rangle_{\partial T}.$$

Using the Cauchy-Schwartz inequality, the equality (2.14) and the multiplicative trace inequality (2.2), we arrive at

$$\begin{aligned}\|\nabla_w v_h\|_T^2 &\le \|\nabla v_h\|_T\|\nabla_w v_h\|_T + \|[v_h]\|_{\partial T}\|\nabla_w v_h\|_{\partial T}\\ &\le \Big(\|\nabla v_h\|_T + Ch_T^{-1/2}\|[v_h]\|_{\partial T}\Big)\|\nabla_w v_h\|_T.\end{aligned}$$

Therefore, we get

$$\|\nabla_w v_h\|_T \le \|\nabla v_h\|_T + Ch_T^{-\frac{1}{2}}\|[v_h]\|_{\partial T}.$$

Taking square on both sides of the above inequality and summing up over all element $T \in \mathcal{T}_h$ leads to

$$\varepsilon\|\nabla_w v_h\|^2 \le C\Big(\varepsilon\|\nabla v_h\|^2 + s_d(v_h, v_h)\Big).$$

As a result,

$$\|v_h\|_w \le C\|v_h\|_{1,h}. \tag{2.18}$$

Taking $\boldsymbol{\sigma} = \nabla v_h$ in (2.17), we have

$$\|\nabla v_h\|_T^2 = \big(\nabla v_h, \nabla_w v_h\big)_T - \langle \{v_h\} - v_h, \nabla v_h \cdot \boldsymbol{n}\rangle_{\partial T}.$$

Using the Cauchy-Schwartz inequality, the equality (2.14) and the multiplicative trace inequality (2.2), we arrive at

$$\begin{aligned}\|\nabla v_h\|_T^2 &\le \|\nabla v_h\|_T \|\nabla_w v_h\|_T + \|[v_h]\|_{\partial T}\|\nabla v_h\|_{\partial T}\\ &\le \bigg(\|\nabla_w v_h\|_T + Ch_T^{-1/2}\|[v_h]\|_{\partial T}\bigg)\|\nabla v_h\|_T.\end{aligned}$$

Hence, we obtain

$$\|\nabla v_h\|_T \le \|\nabla_w v_h\|_T + Ch_T^{-\frac{1}{2}}\|[v_h]\|_{\partial T}.$$

Taking square on both sides of the above inequality and summing up over all element $T \in \mathcal{T}_h$ leads to

$$\varepsilon\|\nabla v_h\|^2 \le C\bigg(\varepsilon\|\nabla_w v_h\|^2 + s_d(v_h, v_h)\bigg).$$

Then we have

$$\|v_h\|_{1,h} \le C\|v_h\|_w. \tag{2.19}$$

From the inequality (2.18) and inequality (2.19), we arrive at

$$C\|v_h\|_w \le \|v_h\|_{1,h} \le C\|v_h\|_w.$$

The definition of $|||\cdot|||-$ norm and $|||\cdot|||_\varepsilon-$ norm conclude that

$$C|||v_h||| \le |||v_h|||_\varepsilon \le C|||v_h|||,$$

which completes the proof. $\square$

We now prove that the bilinear form $a(\cdot,\cdot)$ is continuous and coercive with respect to the $|||\cdot|||$-norm defined by (2.15).

Lemma 2.4. *For $u_h, v_h \in S_h^0(k)$, there exist positive constants C and γ such that*

$$a(u_h, v_h) \le C|||v_h||||||v_h|||, \tag{2.20}$$

$$a(v_h, v_h) \ge \gamma|||v_h|||^2. \tag{2.21}$$

Proof. Let $u_h, v_h \in S_h^0(k)$. It follows from the definition of the bilinear form $a(\cdot,\cdot)$ and the Cauchy-Schwarz inequality that

$$\begin{aligned}a(u_h, v_h) \le &C\Big(\varepsilon\|\nabla_w u_h\|^2 + \|u_h\|^2 + s_d(u_h,u_h) + \big\langle u_h - \{u_h\}, |\mathbf{b}\cdot\boldsymbol{n}|(u_h - \{u_h\})\big\rangle\Big)^{\frac{1}{2}}\\ &\cdot\Big(\varepsilon\|\nabla_w v_h\|^2 + \|v_h\|^2 + s_d(v_h,v_h) + \big\langle v_h - \{v_h\}, |\mathbf{b}\cdot\boldsymbol{n}|(v_h - \{v_h\})\big\rangle\Big)^{\frac{1}{2}}.\end{aligned}$$

Then, we have (2.20) by the definition of $|||\cdot|||$-norm.

The definition of the modified weak divergence (2.5) and integration by parts lead to write

$$\begin{aligned}\big(\nabla_w\cdot(\mathbf{b}u_h), v_h\big) &= -\big(\mathbf{b}u_h, \nabla v_h\big) + \langle\{u_h\}, \mathbf{b}\cdot\boldsymbol{n}v_h\rangle\\ &= \big(\nabla\cdot\mathbf{b}u_h, v_h\big) + \big(\mathbf{b}v_h, \nabla u_h\big) - \langle u_h - \{u_h\}, \mathbf{b}\cdot\boldsymbol{n}v_h\rangle,\end{aligned} \tag{2.22}$$

and

$$\begin{aligned}\big(\nabla_w \cdot (\mathbf{b}v_h), u_h\big) &= -\big(\mathbf{b}v_h, \nabla u_h\big) + \langle \{v_h\}, \mathbf{b}\cdot\boldsymbol{n}u_h\rangle \\ &= -\big(\mathbf{b}v_h, \nabla u_h\big) + \langle \{v_h\}, \mathbf{b}\cdot\boldsymbol{n}(u_h - \{u_h\})\rangle,\end{aligned} \tag{2.23}$$

where we have used the fact that $\langle \mathbf{b}\cdot\boldsymbol{n}\{u_h\}, \{v_h\}\rangle = 0$ for $u_h, v_h \in S_h^0(k)$ in the last equality.

Summing up (2.22) and (2.23) and taking $u_h = v_h$, we obtain

$$\big(\nabla_w \cdot (\mathbf{b}u_h), u_h\big) = \frac{1}{2}\big(\nabla\cdot\mathbf{b}u_h, u_h\big) - \frac{1}{2}\langle u_h - \{u_h\}, \mathbf{b}\cdot\boldsymbol{n}(u_h - \{u_h\})\rangle. \tag{2.24}$$

Using (2.24), we get

$$\begin{aligned}a(u_h,u_h) &= \varepsilon\big(\nabla_w u_h, \nabla_w u_h\big) + \Big((c + \frac{1}{2}\nabla\cdot\mathbf{b})u_h, u_h\Big) \\ &\quad - \frac{1}{2}\big\langle u_h - \{u_h\}, \mathbf{b}\cdot\boldsymbol{n}(u_h - \{u_h\})\big\rangle + s_c(u_h,u_h) + s_d(u_h,u_h) \\ &\geq \varepsilon\|\nabla_w u_h\|^2 + a_0\|u_h\|^2 + \frac{1}{2}\big\langle u_h - \{u_h\}, |\mathbf{b}\cdot\boldsymbol{n}|(u_h - \{u_h\})\big\rangle + s_d(u_h,u_h) \geq \gamma|||u_h|||^2,\end{aligned}$$

with $\gamma = \min\{a_0, \frac{1}{2}\}$. This completes the proof. ☐

Lemmas 2.3 and 2.4 conclude that the bilinear form $a(\cdot,\cdot)$ is also coercive in the norm $|||\cdot|||_\varepsilon$ defined by (2.16).

Lemma 2.5. *For $v_h \in S_h^0(k)$, there exists a positive constant α such that*

$$a(v_h, v_h) \geq \alpha|||v_h|||_\varepsilon{}^2. \tag{2.25}$$

Now, we consider the semi-discrete approximation of a parabolic problem formulated in Algorithm 2.1. The following is a basic stability estimate for the continuous time semi-discrete problem for Eq. (1.1).

Theorem 2.1. *The MWG-FEM solution $u_h(t)$ to the problem stated in Algorithm* 2.1 *has the following good stability inequality:*

$$\|u_h\|^2 + 2\varepsilon\int_0^t \|\nabla_w u_h)\|^2\,dt \leq 2C\int_0^t \|f\|^2\,dt + \|u_h(0)\|^2. \tag{2.26}$$

Proof. Choosing $v = u_h$ in Eq. (2.11), we have

$$(\partial_t u_h, u_h) + a(u_h, u_h) = (f, u_h).$$

The coercivity property (2.21) of the bilinear form $a(\cdot,\cdot)$ implies that

$$a(u_h,u_h) \geq |||u_h|||^2.$$

Thus, we have

$$(\partial_t u_h, u_h) + |||u_h|||^2 \leq C\|f\|^2 + \frac{a_0}{2}\|u_h\|^2.$$

Using the definition of $|||u_h|||$, we get

$$(\partial_t u_h, u_h) + \frac{1}{2}|||u_h|||^2 \leq C\|f\|^2. \tag{2.27}$$

Then, integrating from $t = 0$ to T, the result follows. Thus we complete the proof. ☐

Let δ be a time step size and $t_m = m\delta$ with $0 \le m \le N$ and $N\delta = T$. We denote by $U^m \in S_h^0(k)$ the numerical approximation to $u_h(t_m)$. A fully discrete MWG-FEM for the problem (2.11) is to seek $U^m \in S_h^0(k)$ such that $U^0 = Q_h u_0$ and

$$\big(\hat{\partial}_t U^m, v_h\big) + a(U^m, v_h) = \big(f, v_h\big), \quad \forall v_h \in S_h^0(k), \tag{2.28}$$

where $\hat{\partial}_t U^m = \frac{U^m - U^{m-1}}{\delta}$ is the backward Euler difference at time $t = t_m$. Equivalently we can rewrite this as follows:

$$\begin{aligned} &\big(U^m, v_h\big) + \delta a(U^m, v_h) = \big(U^{m-1} + \delta f(\cdot, t_m), v_h\big), \quad \forall v_h \in S_h^0(k), \quad k = 1, 2, \ldots, \\ &U^0 = Q_h u_0. \end{aligned} \tag{2.29}$$

Again using the coercivity property (2.21), we have with $v_h = U^m$

$$\frac{1}{\delta}\big(U^m, U^m\big) + a(U^m, U^m) \ge C_\gamma |||U^m|||^2, \quad \forall U^m \in S_h^0(k),$$

where $C_\gamma = \min\{a_0 + \frac{1}{\delta}, \frac{1}{2}\}$. The existence and uniqueness of the problem (2.29) follow.

2.2. Error equations

The MWG-FEM lacks of consistency property since the exact solution does not satisfy the numerical scheme (2.11). This property is the key for the Galerkin orthogonality in the conventional finite element methods. Thus, the MWG-FEM does not have the Galerkin orthogonality. In order to establish the error estimates without the Galerkin orthogonality, we will first derive some error equations which will be useful in our later analysis.

Lemma 2.6. *Let u be the solution of the problem* (1.1). *Then for $v_h \in S_h^0(k)$,*

$$-\varepsilon\big(\Delta u, v_h\big) = \varepsilon\big(\nabla_w Q_h u, \nabla_w v_h\big) - T_1(u, v_h), \tag{2.30}$$

$$\big(\nabla \cdot (\boldsymbol{b}u), v_h\big) = \big(\nabla_w \cdot (\boldsymbol{b}Q_h u), v_h\big) - T_2(u, v_h), \tag{2.31}$$

$$\big(cu, v_h\big) = \big(cQ_h u, v_h\big) - T_3(u, v_h), \tag{2.32}$$

where

$$T_1(u, v_h) = \varepsilon\langle\{\nabla u - \mathbb{Q}_h(\nabla u)\}, [v_h]\rangle + \varepsilon\langle\{Q_h u\} - u, \nabla_w v_h \cdot \boldsymbol{n}\rangle, \tag{2.33}$$

$$T_2(u, v_h) = \big(u - Q_h u, \boldsymbol{b} \cdot \nabla v_h\big) - \langle u - \{Q_h u\}, \boldsymbol{b} \cdot \boldsymbol{n} v_h\rangle, \tag{2.34}$$

$$T_3(u, v_h) = -\big(cu - cQ_h u, v_h\big). \tag{2.35}$$

Proof. We start with proving Eq. (2.30). From the definition of the modified weak gradient (2.4) and integration by parts, for $v_h \in S_h^0(k)$, we obtain

$$\begin{aligned} \big(\nabla_w(Q_h u), \nabla_w v_h\big)_T &= -\big(Q_h u, \nabla \cdot (\nabla_w v_h)\big)_T + \langle\{Q_h u\}, \nabla_w v_h \cdot \boldsymbol{n}\rangle_{\partial T} \\ &= -\big(u, \nabla \cdot (\nabla_w v_h)\big)_T + \langle\{Q_h u\}, \nabla_w v_h \cdot \boldsymbol{n}\rangle_{\partial T} \\ &= \big(\nabla u, \nabla_w v_h\big)_T - \langle u - \{Q_h u\}, \nabla_w v_h \cdot \boldsymbol{n}\rangle_{\partial T} \\ &= \big(\mathbb{Q}_h(\nabla u), \nabla_w v_h\big)_T - \langle u - \{Q_h u\}, \nabla_w v_h \cdot \boldsymbol{n}\rangle_{\partial T}. \end{aligned}$$

Then, we have

$$\big(\nabla_w(Q_h u), \nabla_w v_h\big) = \big(\mathbb{Q}_h(\nabla u), \nabla_w v_h\big) + \langle\{Q_h u - u\}, \nabla_w v_h \cdot \boldsymbol{n}\rangle. \tag{2.36}$$

Here, we use the fact that $\{u\} = u$ as u is continuous function. It follows from the definition of modified weak gradient and integration by parts that

$$\begin{aligned}\big(\mathbb{Q}_h(\nabla u), \nabla_w v_h\big)_T &= -\big(v_h, \nabla \cdot \mathbb{Q}_h(\nabla u)\big)_T + \langle \{v_h\}, \mathbb{Q}_h(\nabla u) \cdot \boldsymbol{n}\rangle_{\partial T} \\ &= \big(\nabla v_h, \mathbb{Q}_h(\nabla u)\big)_T + \langle \{v_h\} - v_h, \mathbb{Q}_h(\nabla u) \cdot \boldsymbol{n}\rangle_{\partial T} \\ &= \big(\nabla v_h, \nabla u\big)_T + \langle \{v_h\} - v_h, \mathbb{Q}_h(\nabla u) \cdot \boldsymbol{n}\rangle_{\partial T},\end{aligned}$$

which leads to

$$\big(\mathbb{Q}_h(\nabla u), \nabla_w v_h\big) = \big(\nabla v_h, \nabla u\big) + \langle \{v_h\} - v_h, \mathbb{Q}_h(\nabla u) \cdot \boldsymbol{n}\rangle. \tag{2.37}$$

Multiplying the term $-\Delta u$ by the test function $v_h \in S_h^0(k)$ yields

$$\begin{aligned}-\big(\Delta u, v_h\big) &= \big(\nabla u, \nabla v_h\big) - \langle \nabla u \cdot \boldsymbol{n}, v_h\rangle && (2.38)\\ &= \big(\nabla u, \nabla v_h\big) - \langle \nabla u \cdot \boldsymbol{n}, v_h - \{v_h\}\rangle, && (2.39)\end{aligned}$$

where we use the fact that $\langle \nabla u \cdot \boldsymbol{n}, \{v_h\}\rangle = 0$.

Combining altogether Eqs. (2.36), (2.37) and (2.38) and making use of the identity (2.14), we get the desired result (2.30). Next, we prove Eq. (2.31). The definition of the modified weak divergence (2.5) and integration by parts lead to

$$\begin{aligned}\big(\nabla \cdot (\mathbf{b}u), v_h\big)_T =& -\big(\mathbf{b}u, \nabla v_h\big)_T + \langle u, \mathbf{b} \cdot \boldsymbol{n} v_h\rangle_{\partial T} \\ =& -\big(\mathbf{b}Q_h u, \nabla v_h\big)_T - \big(u - Q_h u, \mathbf{b} \cdot \nabla v_h\big)_T + \langle \{Q_h u\}, \mathbf{b} \cdot \boldsymbol{n} v_h\rangle_{\partial T} \\ & - \langle \{Q_h u\}, \mathbf{b} \cdot \boldsymbol{n} v_h\rangle_{\partial T} + \langle u, \mathbf{b} \cdot \boldsymbol{n} v_h\rangle_{\partial T} \\ =& \big(\nabla_w \cdot (\mathbf{b}Q_h u), v_h\big)_T - \big(u - Q_h u, \mathbf{b} \cdot \nabla v_h\big)_T + \langle u - \{Q_h u\}, \mathbf{b} \cdot \boldsymbol{n} v_h\rangle_{\partial T}.\end{aligned}$$

Summing up over all $T \in \mathcal{T}_h$ gives that

$$\big(\nabla \cdot (\mathbf{b}u), v_h\big) = \big(\nabla_w \cdot (\mathbf{b}Q_h u), v_h\big) - T_2(u, v_h),$$

which is Eq. (2.31). Eq. (2.32) is clear. Thus, we complete the proof. □

3. Error Analysis

We will present the error estimates in this section. First we will derive the a priori error estimates for the semi-discrete MWG-FEM defined by (2.11) and then for the fully discrete MWG-FEM scheme given by (2.29).

Let u be the exact solution of the problem (1.1) and u_h be the solution of the semi-discrete problem given by (2.11), respectively. Let $e := Q_h u - u_h$ be the error between the L^2 projection of the true solution u and the MWG-FEM solution u_h computed by (2.11). Then we have the following error equation for e which will be needed in the error estimates.

Lemma 3.1. *Let $e = Q_h u - u_h \in S_h(k)$. For any $v_h \in S_h^0(k)$ we have*

$$\big(\partial_t e, v_h\big) + a(e, v_h) = T_1(u, v_h) + T_2(u, v_h) + T_3(u, v_h) + s_c(Q_h u, v_h) + s_d(Q_h u, v_h). \tag{3.1}$$

Proof. Multiplying (1.1) by the test function $v_h \in S_h^0(k)$, we obtain

$$(\partial_t u, v_h) - \varepsilon(\Delta u, v_h) + (\nabla \cdot (\mathbf{b}u), v_h) + (cu, v_h) = (f, v_h). \tag{3.2}$$

By making use of Eqs. (2.30)-(2.32), we can write the above equation as

$$\begin{aligned}&(\partial_t u, v_h) + \varepsilon(\nabla_w Q_h u, \nabla_w v_h) + (\nabla_w \cdot (\mathbf{b} Q_h u), v_h) + (c Q_h u, v_h)\\=&(f, v_h) + T_1(u, v_h) + T_2(u, v_h) + T_3(u, v_h).\end{aligned} \tag{3.3}$$

We add the terms $s_c(Q_h u, v_h)$ and $s_d(Q_h u, v_h)$ to both sides of Eq. (3.3) and we get

$$\begin{aligned}&(\partial_t (Q_h u), v_h) + a(Q_h u, v_h)\\=&(f, v_h) + T_1(u, v_h) + T_2(u, v_h) + T_3(u, v_h) + s_c(Q_h u, v_h) + s_d(Q_h u, v_h),\end{aligned} \tag{3.4}$$

where we use the fact that $(Q_h \partial_t u - \partial_t u, v_h) = 0$ for $v_h \in S_h^0(k)$ and the commutative property of the L^2 projection on the time derivative, that is, $(Q_h \partial_t u - \partial_t u, v_h) = (\partial_t (Q_h u) - \partial_t u, v_h) = 0$.

Subtracting (2.11) from (3.4) gives the error equation (3.1). Thus the proof of the lemma is now complete. $\square$

In order to obtain the error estimates, we need to have the error bounds for each term $T_i(u, v_h), i = 1, 2, 3$. The following lemma gives the bounds for these terms.

Lemma 3.2. *If u is the exact solution of the problem* (1.1)*, then for any $v_h \in S_h^0(k)$ we have the following estimates:*

$$|T_1(u, v_h)| \le C\varepsilon^{\frac{1}{2}} h^k |u|_{k+1} |||v_h|||, \tag{3.5}$$

$$|T_2(u, v_h)| \le C h^{k+\frac{1}{2}} |u|_{k+1} |||v_h|||, \tag{3.6}$$

$$|T_3(u, v_h)| \le C h^{k+1} |u|_{k+1} |||v_h|||, \tag{3.7}$$

$$|s_c(Q_h u, v_h)| \le C h^{k+\frac{1}{2}} |u|_{k+1} |||v_h|||, \tag{3.8}$$

$$|s_d(Q_h u, v_h)| \le C\varepsilon^{\frac{1}{2}} h^k |u|_{k+1} |||v_h|||. \tag{3.9}$$

Proof. With the aid of the Cauchy-Schwartz inequality, the trace inequality (2.2) and Lemma 2.1, we obtain

$$\begin{aligned}&\sum_{T\in\mathcal{T}_h} \left|\langle \varepsilon\{\nabla u - \mathbb{Q}_h(\nabla u)\}, [v_h]\rangle_{\partial T}\right| \le C \sum_{T\in\mathcal{T}_h} \varepsilon \|\nabla u - \mathbb{Q}_h(\nabla u)\|_{\partial T} \|[v_h]\|_{\partial T}\\\le& C\Big(\sum_{T\in\mathcal{T}_h} h_T \varepsilon \|\nabla u - \mathbb{Q}_h(\nabla u)\|_{\partial T}^2\Big)^{\frac{1}{2}} \Big(\sum_{T\in\mathcal{T}_h} h_T^{-1} \varepsilon \|[v_h]\|_{\partial T}^2\Big)^{\frac{1}{2}}\\\le& C\varepsilon^{\frac{1}{2}} \Big(\sum_{T\in\mathcal{T}_h} (\|\nabla u - \mathbb{Q}_h(\nabla u)\|_T^2 + h_T^2 |\nabla u - \mathbb{Q}_h(\nabla u)|_{1,T}^2)^{\frac{1}{2}} s_d^{\frac{1}{2}}(v_h, v_h)\\\le& C\varepsilon^{\frac{1}{2}} h^k |u|_{k+1} |||v_h|||.\end{aligned}$$

Similarly, one can show that

$$\sum_{T\in\mathcal{T}_h} \left|\varepsilon\langle \{Q_h u\} - u, \nabla_w v_h \cdot \boldsymbol{n}\rangle_{\partial T}\right| \le C\varepsilon^{\frac{1}{2}} h^k |u|_{k+1} |||v_h|||.$$

Consequently, one has

$$|T_1(u, v_h)| \le C\varepsilon^{\frac{1}{2}} h^k |u|_{k+1} |||v_h|||.$$

We next derive the error bound for the term $T_2(u, v_h)$ in (3.6). Let $\overline{\mathbf{b}}_T$ be the constant value of the average of $\mathbf{b}$ over the element T. Using the Cauchy-Schwartz inequality, Lemmas 2.1

and 2.3, we have

$$\begin{aligned}
\big(u - Q_h u, \mathbf{b}\cdot\nabla v_h\big) &= \sum_{T\in\mathcal{T}_h} \big(u - Q_h u, (\mathbf{b} - \overline{\mathbf{b}}_T)\cdot\nabla v_h\big)_T \\
&\le \sum_{T\in\mathcal{T}_h} \|u - Q_h u\|_T \|\mathbf{b} - \overline{\mathbf{b}}_T\|_{\infty,T} \|\nabla v_h\|_T \\
&\le C h^{k+1} |u|_{k+1} |||v_h|||.
\end{aligned}$$

Moreover, using the Cauchy-Schwartz inequality, the trace inequality (2.2) and Lemma 2.1, one can obtain

$$\begin{aligned}
\langle u - \{Q_h u\}, \mathbf{b}\cdot\boldsymbol{n} v_h\rangle &= \langle u - \{Q_h u\}, \mathbf{b}\cdot\boldsymbol{n}(v_h - \{v_h\})\rangle \\
&\le C \sum_{T\in\mathcal{T}_h} \|u - Q_h u\|_{\partial T} \||\mathbf{b}\cdot\boldsymbol{n}|^{\frac{1}{2}}(v_h - \{v_h\})\|_{\partial T} \\
&\le C\Big(\sum_{T\in\mathcal{T}_h} (h_T^{-1}\|u - Q_h u\|_T^2 + h_T\|\nabla(u - Q_h u)\|_T^2)\Big)^{\frac{1}{2}} s_c^{\frac{1}{2}}(v_h, v_h) \\
&\le C h^{k+\frac{1}{2}} |u|_{k+1} |||v|||,
\end{aligned}$$

where we use the fact that $\langle u - \{Q_h u\}, \mathbf{b}\cdot\boldsymbol{n}\{v_h\}\rangle = 0$. As a result,

$$T_2(u, v_h) \le C h^{k+\frac{1}{2}} |u|_{k+1} |||v|||.$$

Similarly we can prove (3.7).

From the Cauchy-Schwartz inequality, trace inequality (2.2) and Lemma 2.1, we infer that

$$\begin{aligned}
s_c(Q_h u, v_h) &\le \sum_{T\in\mathcal{T}_h} \big|\langle \mathbf{b}\cdot\boldsymbol{n}(Q_h u - \{Q_h u\}), v_h - \{v_h\}\rangle_{\partial_+ T}\big| \\
&= \sum_{T\in\mathcal{T}_h} \big|\langle \mathbf{b}\cdot\boldsymbol{n}(Q_h u - u + u - \{Q_h u\}), v_h - \{v_h\}\rangle_{\partial_+ T}\big| \\
&\le \sum_{T\in\mathcal{T}_h} \big|\langle \mathbf{b}\cdot\boldsymbol{n}(Q_h u - u), v_h - \{v_h\}\rangle_{\partial_+ T}\big| + \sum_{T\in\mathcal{T}_h} \big|\langle \mathbf{b}\cdot\boldsymbol{n}(\{u - Q_h u\}), v_h - \{v_h\}\rangle_{\partial_+ T}\big| \\
&\le C\Big(\sum_{T\in\mathcal{T}_h} \big\||\mathbf{b}\cdot\boldsymbol{n}|^{\frac{1}{2}}(u - Q_h u)\big\|_{\partial T}^2\Big)^{\frac{1}{2}} s_c^{\frac{1}{2}}(v_h, v_h) \\
&\le C h^{k+\frac{1}{2}} |u|_{k+1} |||v_h|||.
\end{aligned}$$

Similar argument shows that the estimate (3.9) holds true. Thus we complete the proof. □

We are now ready to state and prove an error estimate for the time continuous semi-discrete modified WG-FEM approximation (2.11).

Theorem 3.1. *Let $u(x,t)$ and $u_h(x,t)$ be the exact solution of the problem* (1.1) *and the solution of the semi-discrete modified WG-FEM given by* (2.11), *respectively. Assume that $u, \partial_t u \in H^{k+1}(\Omega)$. Then there is a positive constant C independent of ε and the mesh size h such that*

$$\|u - u_h\|^2 \le C\left(\|u_0 - u_h^0\|^2 + h^{2k}(h^2\|u_0\|_{k+1}^2 + (\varepsilon + h)\int_0^t |u|_{k+1}^2\, ds)\right). \tag{3.10}$$

Proof. We aligned the error $e_h = u - u_h = u - Q_h u + Q_h u - u_h := \theta + e$. Lemma 2.1 provides the error estimate for the first term θ

$$\|\theta\| \le Ch^{k+1}|u|_{k+1}. \tag{3.11}$$

Thus, we will derive the error estimate for e. The coercivity property (2.21) of the bilinear form implies that

$$a(e,e) \ge \gamma|||e|||^2.$$

Taking $v_h = e$ in the error equation (3.1) we have

$$(\partial_t e, e) + a(e,e) = T_1(u,e) + T_2(u,e) + T_3(u,e) + s_c(Q_h u, e) + s_d(Q_h u, e).$$

Combining two expressions above yields

$$(\partial_t e, e) + \gamma|||e|||^2 \le T_1(u,e) + T_2(u,e) + T_3(u,e) + s_c(Q_h u, e) + s_d(Q_h u, e). \tag{3.12}$$

Using the Young's inequality and Lemma 3.2, we obtain

$$\begin{aligned}
|T_1(u,e)| &\le C_\gamma \varepsilon h^{2k}|u|_{k+1}^2 + \frac{\gamma}{5}|||e|||^2,\\
|T_2(u,e)| &\le C_\gamma h^{2k+1}|u|_{k+1}^2 + \frac{\gamma}{5}|||e|||^2,\\
|T_3(u,e)| &\le C_\gamma h^{2k+2}|u|_{k+1}^2 + \frac{\gamma}{5}|||e|||^2,\\
|s_c(Q_h u,e)| &\le C_\gamma h^{2k+1}|u|_{k+1}^2 + \frac{\gamma}{5}|||e|||^2,\\
|s_d(Q_h u,e)| &\le C_\gamma \varepsilon h^{2k}|u|_{k+1}^2 + \frac{\gamma}{5}|||e|||^2,
\end{aligned}$$

Substituting the results above into (3.12) reveals

$$\frac{1}{2}\frac{d}{dt}\|e\|^2 \le C(\varepsilon + h)h^{2k}|u|_{k+1}^2. \tag{3.13}$$

Integrating (3.13) with respect to t over $[0,t]$ we find

$$\|e\|^2 \le C\left(\|e(0)\|^2 + (\varepsilon + h)h^{2k}\int_0^t |u|_{k+1}^2\, ds\right). \tag{3.14}$$

With the help of Lemma 2.1, we have

$$\|e(0)\| \le \|Q_h u_0 - u_0\| + \|u_0 - u_h^0\| \le Ch^{k+1}\|u_0\|^2 + \|u_0 - u_h^0\|. \tag{3.15}$$

Substituting the estimate (3.15) into (3.14), we obtain

$$\|e\|^2 \le C\left(\|u_0 - u_h^0\|^2 + h^{2k}(h^2\|u_0\|_{k+1}^2 + (\varepsilon + h)\int_0^t |u|_{k+1}^2\, ds)\right).$$

Combining the above estimate and the estimate (3.11) give the desired result (3.10). □

Theorem 3.2. *Let $u(x,t)$ and $u_h(x,t)$ be the exact solution of the problem* (1.1) *and the solution of the semi-discrete modified WG-FEM given by* (2.11)*, respectively. Assume that $u, \partial_t u, u_0 \in H^{k+1}(\Omega)$. Then there is a positive constant C independent of ε and the mesh size h such that*

$$\begin{aligned}
&\int_0^t \|\partial_t(u-u_h)\|^2\,ds + \frac{\gamma}{4}|||u-u_h|||^2 \\
&\le C\Bigg(|||u_0-u_h^0|||^2 + h^{2k}\Bigg(\|u_0\|_{k+1}^2 + (\varepsilon+h) \\
&\qquad \times\left(\int_0^t |u|_{k+1}^2\,ds + \int_0^t |\partial_t u|_{k+1}^2\,ds + \int_0^t |u_0|_{k+1}^2\,ds\right)\Bigg)\Bigg). \qquad (3.16)
\end{aligned}$$

Proof. With the same notation as in the previous theorem, we estimate the error for e. Taking $v_h = \partial_t e$ in the error equation (3.1) we get

$$\begin{aligned}
&(\partial_t e, \partial_t e) + a(e, \partial_t e) \\
=&T_1(u,\partial_t e) + T_2(u,\partial_t e) + T_3(u,\partial_t e) + s_c(Q_h u, \partial_t e) + s_d(Q_h u, \partial_t e) \\
=&\frac{\partial}{\partial t}T_1(u,e) - T_1(\partial_t u, e) + \frac{\partial}{\partial t}T_2(u,e) - T_2(\partial_t u, e) + \frac{\partial}{\partial t}T_3(u,e) \\
&- T_3(\partial_t u, e) + \frac{\partial}{\partial t}s_c(Q_h u, e) - s_c(Q_h\partial_t u, e) + \frac{\partial}{\partial t}s(Q_h u, e) - s(Q_h\partial_t u, e).
\end{aligned}$$

Using Lemma 3.2 we have

$$\begin{aligned}
&\|\partial_t e\|^2 + \frac{1}{2}\frac{d}{dt}a(e,e) \\
\le&C(\varepsilon^{\frac12}+h^{\frac12})h^k|\partial_t u|_{k+1}|||e||| + \frac{\partial}{\partial t}T_1(u,e) \\
&+ \frac{\partial}{\partial t}T_2(u,e) + \frac{\partial}{\partial t}T_3(u,e) + \frac{\partial}{\partial t}s_c(Q_h u, e) + \frac{\partial}{\partial t}s_d(Q_h u, e).
\end{aligned}$$

Integrating the last inequality with respect to t gives

$$\begin{aligned}
&\int_0^t \|\partial_t e\|^2\,ds + \frac{\gamma}{2}|||e|||^2 \\
\le&C\Bigg(\frac{\gamma}{2}|||e(0)|||^2 + (\varepsilon^{\frac12}+h^{\frac12})h^k\int_0^t |\partial_t u|_{k+1}|||e|||\,ds \\
&+ T_1(u,e) + T_2(u,e) + T_3(u,e) + s_c(Q_h u, e) + s_d(Q_h u, e) \\
&- T_1(u(0),e(0)) - T_2(u(0),e(0)) - T_3(u(0),e(0)) \\
&- s_c(Q_h u(0), e(0)) - s(Q_h u(0), e(0))\Bigg).
\end{aligned}$$

Again using Lemma 3.2 and the Young's inequality we get

$$\begin{aligned}
&\int_0^t \|\partial_t e\|^2\,ds + \frac{\gamma}{2}|||e|||^2 \\
&\le C\Bigg(\frac{\gamma}{2}|||e(0)|||^2 + (\varepsilon+h)h^{2k}\left(\int_0^t |u|_{k+1}^2\,ds + \int_0^t |\partial_t u|_{k+1}^2\,ds + \int_0^t |u_0|_{k+1}^2\,ds\right)
\end{aligned}$$

$$+ \frac{\gamma}{4}|||e|||^2 + \frac{\gamma}{2}|||e(0)|||^2\Bigg).$$

Note that,

$$|||e(0)||| \le C(h^k\|u_0\|_{k+1} + |||u_0 - u_h^0|||).$$

Thus, we have

$$\begin{aligned}
&\int_0^t \|\partial_t e\|^2\,ds + \frac{\gamma}{4}|||e|||^2 \\
\le& C\left(|||u_0 - u_h^0|||^2 + h^{2k}\left(\|u_0\|_{k+1}^2 + (\varepsilon + h)\left(\int_0^t |u|_{k+1}^2\,ds + \int_0^t |\partial_t u|_{k+1}^2\,ds + \int_0^t |u_0|_{k+1}^2\,ds\right)\right)\right).
\end{aligned}$$

The above estimate and the estimate for θ give the desired result (3.16). Thus, the proof is completed. □

Next, we give the error estimates for the fully discrete MWG-FEM solution computed by (2.28).

Theorem 3.3. *Let u and U^m be the exact solution of the problem* (1.1) *and the solution of the fully discrete MWG-FEM given by* (2.28), *respectively. Assume that $u, \partial_t u, u_0 \in H^{k+1}(\Omega)$. Then there is a positive constant C independent of ε and h such that for $0 < m \le N$*

$$\begin{aligned}
&\|u(t_m) - U^m\|^2 \\
\le& C\Bigg(\|u_0 - U^0\|^2 + \delta^2 \int_0^{t_m} \left\|\frac{\partial^2 u}{\partial t^2}\right\|^2 ds \\
&+ h^{2k}\left(h^2\left(\|u_0\|_{k+1}^2 + \int_0^{t_m} \|\partial_t u\|_{k+1}^2\,ds\right) + \delta(\varepsilon + h)\sum_{n=1}^{m} |u_n|_{k+1}^2\right)\Bigg).
\end{aligned} \tag{3.17}$$

Proof. We aligned the error $e_h^m = u(t_m) - U^m = \theta^m + e^m$ where

$$\theta^m = u(t_m) - Q_h u(t_m), \quad e^m = Q_h u(t_m) - U^m.$$

We know from Lemma 2.1 that

$$\|\theta^m\| \le Ch^{k+1}|u(t_m)|_{k+1} \le Ch^{k+1}\left(\|u_0\|_{k+1} + \int_0^{t_m} \|\partial_t u\|_{k+1}\,dt\right). \tag{3.18}$$

From the definition of Q_h, we have for $v_h \in S_h^0(k)$

$$\begin{aligned}
\big(Q_h\partial_t u(t_m) - \hat{\partial}_t U^m, v_h\big) &= \big(Q_h\partial_t u(t_m) - \hat{\partial}_t Q_h u(t_m), v_h\big) + \big(\hat{\partial}_t(Q_h u(t_m) - U^m), v_h\big) \\
&= \big(\partial_t u(t_m) - \hat{\partial}_t u(t_m), v_h\big) + \big(\hat{\partial}_t(Q_h u(t_m) - U^m), v_h\big),
\end{aligned}$$

or, equivalently, using the definition of the projection Q_h

$$\big(\hat{\partial}_t(Q_h u(t_m) - U^m), v_h\big) - \big(Q_h\partial_t u(t_m) - \hat{\partial}_t U^m, v_h\big) = \big(\hat{\partial}_t u(t_m) - \partial_t u(t_m), v_h\big).$$

In what follows we will find an equivalent expression for the term $(\partial_t u(t_m) - \hat{\partial}_t U^m, v_h)$. From the variational formulation given in (2.1) and the fully discrete MWG-FEM scheme defined by (2.28) we obtain

$$(\partial_t u(t_m) - \hat{\partial}_t U^m, v_h) = a(U^m, v_h) - A(u(t_m), v_h).$$

Using the same argument in deriving the error equations in Lemma 2.6 for the semi-discrete case, we have the following error equation for the term $a(U^m, v_h) - A(u(t_m), v_h)$:

$$\begin{aligned} &a(U^m, v_h) - A(u(t_m), v_h) \\ =& -\Big(a(Q_h u(t_m) - U^m, v_h) - T_1(u(t_m), v_h) - T_2(u(t_m), v_h) \\ &\quad - T_3(u(t_m), v_h) - s_c(Q_h u(t_m), v_h) - s_d(Q_h u(t_m), v_h)\Big). \end{aligned} \tag{3.19}$$

Thus, we have the following error equation for the fully discrete MWG-FEM for any $v_h \in S_h^0(k)$

$$\begin{aligned} &\big(\hat{\partial}_t(Q_h u(t_m) - U^m), v_h\big) + a(Q_h u(t_m) - U^m, v_h) \\ =&\big(\hat{\partial}_t u(t_m) - \partial_t u(t_m), v_h\big) + T_1(u(t_m), v_h) + T_2(u(t_m), v_h) \\ &+ T_3(u(t_m), v_h) + s_c(Q_h u(t_m), v_h) + s_d(Q_h u(t_m), v_h). \end{aligned} \tag{3.20}$$

Letting $v_h = e^m$ in the above equation (3.20), we have

$$\begin{aligned} &(\hat{\partial}_t e^m, e^m) + a(e^m, e^m) \\ =& (w^m, e^m) + T_1(u(t_m), e^m) + T_2(u(t_m), e^m) + T_3(u(t_m), e^m) \\ &+ s_c(Q_h u(t_m), e^m) + s_d(Q_h u(t_m), e^m), \end{aligned}$$

where

$$w^m = \hat{\partial}_t u(t_m) - \partial_t u(t_m).$$

Using the similar argument in the semi-discrete case, one can show that

$$(\hat{\partial}_t e^m, e^m) + a(e^m, e^m) \le (w^m, e^m) + C(\varepsilon^{\frac{1}{2}} + h^{\frac{1}{2}})h^k |u(t_m)|_{k+1}^2 |||e^m|||.$$

Using the Cauchy-Schwartz inequality, the Young's inequality and the coercivity of the bilinear form, we estimate the two terms on the left hand side of the above inequality

$$\big|(\hat{\partial}_t e^m, e^m)\big| = \frac{1}{\delta}(e^m - e^{m-1}, e^m) \ge \frac{1}{2\delta}\Big(\|e^m\|^2 - \|e^{m-1}\|^2\Big), \tag{3.21}$$

and

$$a(e^m, e^m) \ge \gamma |||e^m|||^2. \tag{3.22}$$

Combining (3.21) and (3.22) leads to

$$\begin{aligned} \|e^m\|^2 + 2\gamma\delta |||e^m|||^2 \le& \|e^{m-1}\|^2 + \delta\|w^m\|^2 + \delta\|e^m\| \\ &+ C\frac{\delta}{8\gamma}(\varepsilon + h)h^{2k}|u(t_m)|_{k+1}^2 + 2\gamma\delta |||e^m|||^2. \end{aligned}$$

Hence,

$$\|e^m\|^2 \le \|e^{m-1}\|^2 + \delta\|w^m\|^2 + \delta\|e^m\|^2 + C\frac{\delta}{8\gamma}(\varepsilon + h)h^{2k}|u(t_m)|_{k+1}^2.$$

By induction argument, one has

$$\|e^m\|^2 \le \|e^0\|^2 + \delta\left(\sum_{n=1}^m \|w^n\|^2 + \sum_{n=1}^m \|e^n\|^2\right) + C\frac{\delta}{8\gamma}(\varepsilon + h)h^{2k}\sum_{n=1}^m |u(t_n)|_{k+1}^2. \tag{3.23}$$

We note that

$$w^n = \frac{u(t_n) - u(t_{n-1})}{\delta} - \partial_t u(t_n) = -\frac{1}{\delta}\int_{t_{n-1}}^{t_n} (s - t_{n-1})\frac{\partial^2 u}{\partial t^2}\,ds,$$

which gives

$$\sum_{n=1}^{m} \|w^n\|^2 \le \sum_{n=1}^{m} \frac{1}{\delta^2}\left(\int_{t_{n-1}}^{t_n} (s-t_{n-1})^2\,ds\right)\left(\int_{t_{n-1}}^{t_n} \left\|\frac{\partial^2 u}{\partial t^2}\right\|^2 ds\right) = \frac{\delta}{3}\int_0^{t_m} \left\|\frac{\partial^2 u}{\partial t^2}\right\|^2 ds. \tag{3.24}$$

With the aid of Lemma 2.1, we have

$$\|e^0\|^2 \le Ch^{2k+2}\|u_0\|_{k+1}^2 + \|u_0 - U^0\|^2. \tag{3.25}$$

Combining (3.18), (3.23), (3.25) and the discrete *Gronwall* inequality yield the desired result (3.17). The proof is now complete. □

Theorem 3.4. *Let u and U^m be the exact solution of the problem* (1.1) *and the solution of the fully discrete MWG-FEM given by* (2.28), *respectively. Assume that $u, \partial_t u, u_0 \in H^{k+1}(\Omega)$. Then there is a positive constant C independent of ε and h such that for $0 < m \le N$*

$$|||u(t_m) - U^m|||^2 \le C\Bigg(|||u_0 - U^0|||^2 + \delta^2\int_0^{t_m}\left\|\frac{\partial^2 u}{\partial t^2}\right\|^2 ds \tag{3.26}$$
$$+ h^{2k}\left(\|u_0\|_{k+1}^2 + (\varepsilon + h)\delta\sum_{n=1}^{m}|u_n|_{k+1}^2\right)\Bigg).$$

Proof. Taking $v_h = \hat{\partial}_t e^m$ in the error equation (3.20), we have

$$\begin{aligned}
&\big(\hat{\partial}_t e^m, \hat{\partial}_t e^m\big) + a(e^m, \hat{\partial}_t e^m)\\
=&\big(w^m, \hat{\partial}_t e^m\big) + T_1(u(t_m), \hat{\partial}_t e^m) + T_2(u(t_m), \hat{\partial}_t e^m)\\
&+ T_3(u(t_m), \hat{\partial}_t e^m) + s_c(Q_h u(t_m), \hat{\partial}_t e^m) + s(Q_h u(t_m), \hat{\partial}_t e^m).
\end{aligned}$$

Observe that

$$\begin{aligned}
\big(\hat{\partial}_t e^m, \hat{\partial}_t e^m\big) &= \|\hat{\partial}_t e^m\|^2,\\
a(e^m, \hat{\partial}_t e^m) &= \frac{1}{\delta}\left(a(e^m, e^m) - a(e^m, e^{m-1})\right)\\
&\ge \frac{1}{\delta}\left(\gamma|||e^m|||^2 - \beta|||e^m||||||e^{m-1}|||\right)\\
&\ge \frac{1}{\delta}\left(\left(\gamma - \frac{\beta}{2}\right)|||e^m|||^2 - \frac{\beta}{2}|||e^{m-1}|||^2\right)\\
&= \frac{C}{\delta}\left(|||e^m|||^2 - |||e^{m-1}|||^2\right),
\end{aligned}$$

where we use the continuity of the bilinear form $a(\cdot,\cdot)$ given by (2.20).

Using the Young's inequality and Lemma 3.2, we have

$$|T_1(u(t_m), \hat{\partial}_t e^m)| \le C_\gamma \varepsilon\delta^{-1}h^{2k}|u(t_m)|_{k+1}^2 + \frac{\gamma}{5}\delta^{-1}\left(|||e^m|||^2 + |||e^{m-1}|||^2\right),$$

$$|T_2(u(t_m), \hat{\partial}_t e^m)| \le C_\gamma \delta^{-1} h^{2k+1} |u(t_m)|_{k+1}^2 + \frac{\gamma}{5}\delta^{-1}\left(|||e^m|||^2 + |||e^{m-1}|||^2\right),$$
$$|T_3(u(t_m), \hat{\partial}_t e^m)| \le C_\gamma \varepsilon \delta^{-1} h^{2k+2} |u(t_m)|_{k+1}^2 + \frac{\gamma}{5}\delta^{-1}\left(|||e^m|||^2 + |||e^{m-1}|||^2\right),$$
$$|s_c(Q_h u(t_m), \hat{\partial}_t e^m)| \le C_\gamma \delta^{-1} h^{2k+1} |u(t_m)|_{k+1}^2 + \frac{\gamma}{5}\delta^{-1}\left(|||e^m|||^2 + |||e^{m-1}|||^2\right),$$
$$|s_d(Q_h u(t_m), \hat{\partial}_t e^m)| \le C_\gamma \varepsilon \delta^{-1} h^{2k} |u(t_m)|_{k+1}^2 + \frac{\gamma}{5}\delta^{-1}\left(|||e^m|||^2 + |||e^{m-1}|||^2\right).$$

The Cauchy-Schwartz inequality implies that

$$\delta\|\hat{\partial}_t e^m\|^2 + C|||e^m|||^2 \le C|||e^{m-1}|||^2 + \frac{\delta}{4}\|w^m\|^2 + \delta\|\hat{\partial}_t e^m\|^2 + C(\varepsilon + h)h^{2k}|u(t_m)|_{k+1}^2.$$

By induction argument along with cancellation we are led to

$$|||e^m|||^2 \le C|||e^0|||^2 + \frac{\delta}{4}\sum_{n=1}^{m}\|w^n\|^2 + C(\varepsilon + h)h^{2k}\sum_{n=1}^{m}|u(t_n)|_{k+1}^2. \tag{3.27}$$

Using Lemma 2.1 we obtain

$$|||e^0||| \le Ch^k\|u_0\| + |||u_0 - U^0|||. \tag{3.28}$$

Substituting the inequalities (3.28) and (3.24) in the above inequality (3.27) gives the desired result. □

4. Optimal Order Error Estimates in L^2-norm

We have derived the optimal order of error estimate in H^1-norm for both semi-discrete and fully discrete MWG-FEM schemes in the previous section. In this section, we present the optimal order of error estimate in L^2-norm for semi-discrete and fully discrete MWG-FEM. For this reason, similar to [33], we define an elliptic projection $R_h u(t)$ of u on the space $S_h^0(k)$ as follows. For each $t \in [0,T]$, $R_h u(t) : H_0^1(\Omega) \cap H^2(\Omega) \to S_h^0(k)$ such that

$$a(R_h u, v_h) = -\varepsilon(\Delta u, v_h) + (\nabla \cdot (\mathbf{b}u), v_h) + (cu, v_h), \quad \forall v_h \in S_h^0(k). \tag{4.1}$$

The projection $R_h u(t)$ is well-defined since the bilinear form $a(\cdot,\cdot)$ is bounded and coercive by Lemma 2.4. The following lemma will be useful in the sequel.

Lemma 4.1. *Let $u \in H^{k+1}(\Omega)$ be the exact solution of problem* (1.1)*. Then there is a positive constant C such that*

$$|||R_h u - Q_h u||| \le C(\varepsilon^{\frac{1}{2}} + h^{\frac{1}{2}})h^k\|u\|_{k+1}, \tag{4.2}$$
$$\|R_h u - Q_h u\| \le C(\varepsilon^{-\frac{1}{2}}h^{\frac{1}{2}} + \varepsilon^{-1} + 1)h^{k+1}\|u\|_{k+1}. \tag{4.3}$$

Proof. From the definition of elliptic projection we have

$$(\partial_t u, v_h) + a(R_h u, v_h) = (\partial_t u, v_h) + A(u, v_h) = (f, v_h), \quad \forall v_h \in S_h^0(k). \tag{4.4}$$

Testing (1.1) by $v_h \in S_h^0(k)$ yields

$$(\partial_t u, v_h) - \varepsilon(\Delta u, v_h) + (\nabla \cdot (\mathbf{b}u), v_h) + (cu, v_h) = (f, v_h). \tag{4.5}$$

With the help of Lemma 2.6, Eq. (4.5) becomes

$$\begin{aligned}&(\partial_t u, v_h) + \varepsilon(\nabla_w Q_h u, \nabla_w v_h) + (\nabla_w \cdot (\mathbf{b} Q_h u), v_h) + (c Q_h u, v_h)\\ =&(f, v_h) + T_1(u, v_h) + T_2(u, v_h) + T_3(u, v_h).\end{aligned}$$

We add $s_c(Q_h u, v_h)$ and $s_d(Q_h u, v_h)$ to both side of the equation above to obtain

$$\begin{aligned}(\partial_t u, v_h) + a(Q_h u, v_h) = (f, v_h) &+ T_1(u, v_h) + T_2(u, v_h) + T_3(u, v_h)\\ &+ s_c(Q_h u, v_h) + s_d(Q_h u, v_h).\end{aligned} \tag{4.6}$$

Let $E := Q_h u - R_h u$. Subtracting (4.4) from (4.6) gives

$$a(E, v_h) = T_1(u, v_h) + T_2(u, v_h) + T_3(u, v_h) + s_c(Q_h u, v_h) + s_d(Q_h u, v_h). \tag{4.7}$$

Taking $v_h = E$ in (4.7) and using the coercivity of the bilinear form $a(\cdot,\cdot)$ and Lemma 3.2 we have the desired result

$$|||E||| \leq C(\varepsilon^{\frac{1}{2}} + h^{\frac{1}{2}})h^k |u|_{k+1},$$

which proves the estimate (4.2).

Next we will prove the estimate (4.3). We consider the following dual problem : Find $\phi \in H_0^1(\Omega) \cap H^2(\Omega)$ satisfying

$$-\varepsilon\Delta\phi - \mathbf{b} \cdot \nabla\phi + c\phi = E \quad \text{in } \Omega, \tag{4.8}$$

where $c \in W^{1,\infty}(\Omega)$ and for some constant c_0 such that $c - \frac{1}{2}\nabla \cdot \mathbf{b} \geq c_0 > 0$. The following lemma is needed for the rest of the proof.

Lemma 4.2. *Let $\phi \in H_0^1(\Omega) \cap H^2(\Omega)$ be the exact solution of the problem* (4.8). *Then for any $v_h \in S_h^0(k)$ we have the following estimates:*

$$|T_1(\phi, v_h)| \leq C\varepsilon^{\frac{1}{2}} h |\phi|_2 |||v_h|||, \tag{4.9}$$

$$|T_2(\phi, v_h)| \leq C h^{\frac{3}{2}} |\phi|_2 |||v_h|||, \tag{4.10}$$

$$|T_3(\phi, v_h)| \leq C h^2 |\phi|_2 |||v_h|||, \tag{4.11}$$

$$|s_c(Q_h\phi, v_h)| \leq C h^{\frac{3}{2}} |\phi|_2 |||v_h|||, \tag{4.12}$$

$$|s_d(Q_h\phi, v_h)| \leq C\varepsilon^{\frac{1}{2}} h |\phi|_2 |||v_h|||, \tag{4.13}$$

where $T_i(u,v), i = 1, 2, 3$ are defined by (2.33), (2.34) *and* (2.35), *respectively.*

Proof. The proof is similar to the proof of Lemma 3.2 and thus we omit the proof. □.

We assume that the following H^2-regularity for the dual problem (4.8) holds

$$\varepsilon\|\phi\|_2 \leq C\|E\|. \tag{4.14}$$

To see this, an H^1-energy estimate can be found very quickly and is

$$\varepsilon|\phi|_{1,\Omega}^2 + c_0\|\phi\|_{L^2(\Omega)}^2 \leq \|E\|_{L^2(\Omega)}^2. \tag{4.15}$$

We now proceed with an H^2-regularity estimate. Multiplying the dual problem by $\Delta\phi$ and integrating gives

$$-\varepsilon\int_\Omega (\Delta\phi)^2\, dx - \int_\Omega \mathbf{b} \cdot \nabla\phi\Delta\phi\, dx + \int_\Omega c\phi\Delta\phi\, dx = \int_\Omega E\Delta\phi\, dx.$$

We use integration-by-parts on the second and third terms on the left side as well as the BC's on ϕ so this becomes

$$-\varepsilon \int_\Omega (\Delta\phi)^2 \, dx - \int_\Omega \left(c - \frac{1}{2}\nabla\cdot\mathbf{b}\right)(\nabla\phi)^2 \, dx \le \int_\Omega E\Delta\phi \, dx + \int_\Omega \nabla c\phi\nabla\phi \, dx, \tag{4.16}$$

which leads to the inequality

$$\varepsilon|\phi|_{2,\Omega}^2 + c_0|\phi|_{1,\Omega}^2 \le \|E\|_{L^2(\Omega)}|\phi|_{2,\Omega} + c_2\|\phi\|_{L^2(\Omega)}|\phi|_{1,\Omega},$$

where $\|\nabla c\|_{L^\infty(\Omega)} \le c_2$. We now use Young's inequality on each of the terms on the right side of the equation

$$\varepsilon|\phi|_{2,\Omega}^2 + c_0|\phi|_{1,\Omega}^2 \le \frac{1}{4\varepsilon}\|E\|_{L^2(\Omega)}^2 + \frac{\varepsilon}{2}|\phi|_{2,\Omega}^2 + \frac{4c_2^2}{c_0}\|\phi\|_{L^2(\Omega)}^2 + \frac{c_0}{2}|\phi|_{1,\Omega}^2,$$

or

$$\varepsilon|\phi|_{2,\Omega}^2 + c_0|\phi|_{1,\Omega}^2 \le \frac{1}{2\varepsilon}\|E\|_{L^2(\Omega)}^2 + \frac{8c_2^2}{c_0}\|\phi\|_{L^2(\Omega)}^2.$$

Now, from our H^1-regularity (4.15), we have $c_0\|\phi\|_{L^2(\Omega)}^2 \le \|E\|_{L^2(\Omega)}^2$ and thus conclude $\exists C > 0$ dependent only on c_0 and c_2 so

$$\varepsilon|\phi|_{2,\Omega}^2 + c_0\|\phi\|_{1,\Omega}^2 \le \frac{C}{\varepsilon}\|E\|_{L^2(\Omega)}^2, \tag{4.17}$$

which shows (4.14).

Multiplying the dual problem (4.8) by the test function E and using the Eqs. (2.30) and (2.32), we obtain

$$\begin{aligned}\|E\|^2 &= -\varepsilon(\Delta\phi, E) - (\mathbf{b}\cdot\nabla\phi, E) + (c\phi, E)\\ &= \varepsilon(\nabla_w Q_h\phi, \nabla_w E) - (\mathbf{b}\cdot\nabla\phi, E) + (cQ_h\phi, E) - T_1(\phi, E) - T_3(\phi, E).\end{aligned} \tag{4.18}$$

Integration by parts and the definition of the weak divergence reveal that

$$\begin{aligned}-(\mathbf{b}\cdot\nabla\phi, E)_T &= (\phi, \nabla\cdot(\mathbf{b}E))_T - \langle\phi, \mathbf{b}\cdot\mathbf{n}E\rangle_{\partial T}\\ &= (Q_h\phi, \nabla\cdot(\mathbf{b}E))_T + (\phi - Q_h\phi, \nabla\cdot(\mathbf{b}E))_T - \langle\phi, \mathbf{b}\cdot\mathbf{n}E\rangle_{\partial T}\\ &= (\nabla_w\cdot(\mathbf{b}E), Q_h\phi)_T - T_2^*(\phi, E)_T,\end{aligned}$$

where $T_2^*(\phi, E)_T = (\phi - Q_h\phi, \nabla\cdot(\mathbf{b}E))_T - \langle\{\phi - Q_h\phi\}, \mathbf{b}\cdot\mathbf{n}E\rangle_{\partial T}$. As a result, by summing all over the element $T \in \mathcal{T}_h$ we have

$$-(\mathbf{b}\cdot\nabla\phi, E) = (\nabla_w\cdot(\mathbf{b}E), Q_h\phi) - T_2^*(\phi, E). \tag{4.19}$$

Combining Eqs. (4.18) and (4.19) leads to

$$\begin{aligned}\|E\|^2 &= \varepsilon(\nabla_w E, \nabla_w Q_h\phi) + (\nabla_w\cdot(\mathbf{b}E), Q_h\phi)\\ &\quad + (cE, Q_h\phi) - T_1(\phi, E) - T_2^*(\phi, E) - T_3(\phi, E)\\ &= a(E, Q_h\phi) - s_c(E, Q_h\phi) - s_d(E, Q_h\phi)\\ &\quad - T_1(\phi, E) - T_2^*(\phi, E) - T_3(\phi, E).\end{aligned} \tag{4.20}$$

From (4.7) with v_h is replaced by $Q_h\phi$, we have

$$\begin{aligned}a(E, Q_h\phi) &= T_1(u, Q_h\phi) + T_2(u, Q_h\phi) + T_3(u, Q_h\phi)\\ &\quad + s_c(Q_h u, Q_h\phi) + s_d(Q_h u, Q_h\phi).\end{aligned} \tag{4.21}$$

Eqs. (4.20) and (4.21) imply that

$$\begin{aligned}\|E\|^2 = &T_1(u, Q_h\phi) + T_2(u, Q_h\phi) + T_3(u, Q_h\phi)\\&+ s_c(Q_hu, Q_h\phi) + s_d(Q_hu, Q_h\phi)\\&- T_1(\phi, E) - T_2^*(\phi, E) - T_3(\phi, E)\\&- s_c(E, Q_h\phi) - s_d(E, Q_h\phi).\end{aligned}\tag{4.22}$$

We will derive a bound for each term on the RHS of Eq. (4.22) separately.

We begin with the first term $T_1(u, Q_h\phi)$. It follows from the Cauchy-Schwarz inequality, the trace inequality (2.13), and Lemma 2.1 that

$$\begin{aligned}&\sum_{T\in\mathcal{T}_h}\Big|\big\langle\varepsilon\{\nabla u - \mathbb{Q}_h(\nabla u)\}, [\phi - Q_h\phi]\big\rangle_{\partial T}\Big|\\&\le C\sum_{T\in\mathcal{T}_h}\varepsilon\|\{\nabla u - \mathbb{Q}_h(\nabla u)\}\|_{\partial T}\|[\phi - Q_h\phi]\|_{\partial T}\\&\le C\Big(\sum_{T\in\mathcal{T}_h}\varepsilon h_T\|\{\nabla u - \mathbb{Q}_h(\nabla u)\}\|^2_{\partial T}\Big)^{\frac12}\Big(\sum_{T\in\mathcal{T}_h}\varepsilon h_T^{-1}\|[\phi - Q_h\phi]\|^2_{\partial T}\Big)^{\frac12}\\&\le C\varepsilon^{\frac12}\Big(\sum_{T\in\mathcal{T}_h}(\|\nabla u - \mathbb{Q}_h(\nabla u)\|^2_T + h_T^2|\nabla u - \mathbb{Q}_h(\nabla u)|^2_{1,T})\Big)^{\frac12}|||\phi - Q_h\phi|||\\&\le C\varepsilon^{\frac12}h^k|u|_{k+1}\varepsilon^{\frac12}h\|\phi\|_2 \le C\varepsilon h^{k+1}|u|_{k+1}\|\phi\|_2,\end{aligned}$$

where we use the fact that $[\phi] = 0$.

Similarly, one can show that

$$\sum_{T\in\mathcal{T}_h}\big|\varepsilon\langle\{Q_hu\} - u, \nabla_w Q_h\phi\cdot\boldsymbol{n}\rangle_{\partial T}\big| \le C\varepsilon h^{k+1}|u|_{k+1}\|\phi\|_2.$$

Consequently, one has

$$|T_1(u, Q_h\phi)| \le C\varepsilon h^{k+1}|u|_{k+1}\|\phi\|_2.\tag{4.23}$$

We next derive the error bound for the term $T_2(u, Q_h\phi)$. Let $\overline{\mathbf{b}}_T$ be the constant value of the average of $\mathbf{b}$ over the element T. Using the Cauchy-Schwartz inequality, Lemma 2.1 and the Poincare inequality, we have

$$\begin{aligned}&\big(u - Q_hu, \mathbf{b}\cdot\nabla Q_h\phi\big)\\&=\sum_{T\in\mathcal{T}_h}\big(u - Q_hu, (\mathbf{b} - \overline{\mathbf{b}}_T)\cdot\nabla\phi\big)_T\\&\quad-\sum_{T\in\mathcal{T}_h}\big(u - Q_hu, (\mathbf{b} - \overline{\mathbf{b}}_T)\cdot\nabla(\phi - Q_h\phi)\big)_T\\&\le\sum_{T\in\mathcal{T}_h}\|u - Q_hu\|_T\|\mathbf{b} - \overline{\mathbf{b}}_T\|_{\infty,T}\|\nabla\phi\|_T\\&\quad+\sum_{T\in\mathcal{T}_h}\|u - Q_hu\|_T\|\mathbf{b} - \overline{\mathbf{b}}_T\|_{\infty,T}\|\nabla(\phi - Q_h\phi)\|_T\\&\le Ch^{k+1}|u|_{k+1}\|\phi\|_1 + Ch^{k+2}|u|_{k+1}\|\phi\|_2\\&\le Ch^{k+1}|u|_{k+1}\|\phi\|_2.\end{aligned}$$

Moreover, using the Cauchy-Schwartz inequality, the trace inequality (2.13) and Lemma 2.1, one can obtain

$$\begin{aligned}
&\langle u-\{Q_hu\},\mathbf{b}\cdot\boldsymbol{n}Q_h\phi\rangle\\
=&\langle u-\{Q_hu\},\mathbf{b}\cdot\boldsymbol{n}((Q_h\phi-\phi)-\{Q_h\phi-\phi\})\rangle\\
\le&C\sum_{T\in\mathcal{T}_h}\|u-Q_hu\|_{\partial T}\||\mathbf{b}\cdot\boldsymbol{n}|^{\frac12}(\phi-Q_h\phi)\|_{\partial T}\\
\le&Ch^{k+2}|u|_{k+1}\|\phi\|_2,
\end{aligned}$$

where we use the facts that $\langle u-\{Q_hu\},\mathbf{b}\cdot\boldsymbol{n}\{Q_h\phi\}\rangle=0$ and $\{\phi\}=\phi$. As a result,

$$|T_2(u,Q_h\phi)|\le Ch^{k+1}|u|_{k+1}\|\phi\|_2. \tag{4.24}$$

It follows from the Cauchy-Schwartz inequality and Lemma 2.1 that

$$|T_3(u,Q_h\phi)|\le Ch^{k+1}|u|_{k+1}\|\phi\|_2. \tag{4.25}$$

From the Cauchy-Schwartz inequality, trace inequality (2.13) and Lemma 2.1, we infer that

$$\begin{aligned}
|s_c(Q_hu,Q_h\phi)|\le&\sum_{T\in\mathcal{T}_h}\left|\langle\mathbf{b}\cdot\boldsymbol{n}(Q_hu-\{Q_hu\}),Q_h\phi-\{Q_h\phi\}\rangle_{\partial_+T}\right|\\
=&\sum_{T\in\mathcal{T}_h}\left|\langle\mathbf{b}\cdot\boldsymbol{n}(Q_hu-u+u-\{Q_hu\}),(Q_h\phi-\phi+\phi-\{Q_h\phi\})\rangle_{\partial_+T}\right|\\
\le&\sum_{T\in\mathcal{T}_h}\left|\langle\mathbf{b}\cdot\boldsymbol{n}(Q_hu-u),Q_h\phi-\phi+\phi-\{Q_h\phi\}\rangle_{\partial_+T}\right|\\
&+\sum_{T\in\mathcal{T}_h}\left|\langle\mathbf{b}\cdot\boldsymbol{n}(\{u-Q_hu\}),Q_h\phi-\phi+\phi-\{Q_h\phi\}\rangle_{\partial_+T}\right|\\
\le&C\Big(\sum_{T\in\mathcal{T}_h}\left\||\mathbf{b}\cdot\boldsymbol{n}|^{\frac12}(u-Q_hu)\right\|^2_{\partial T}\Big)^{\frac12}\Big(\sum_{T\in\mathcal{T}_h}\left\||\mathbf{b}\cdot\boldsymbol{n}|^{\frac12}(\phi-Q_h\phi)\right\|^2_{\partial T}\Big)^{\frac12}\\
\le&Ch^{k+\frac12}|u|_{k+1}h^{\frac32}\|\phi\|_2=Ch^{k+2}|u|_{k+1}\|\phi\|_2.
\end{aligned} \tag{4.26}$$

Similarly one can show that

$$|s_d(Q_hu,Q_h\phi)|\le C\varepsilon h^{k+\frac32}|u|_{k+1}\|\phi\|_2. \tag{4.27}$$

The estimates (4.9) and (4.2) imply that

$$|T_1(\phi,E)|\le C\varepsilon^{\frac12}h\|\phi\|_2|||E|||\le C\varepsilon^{\frac12}(\varepsilon^{\frac12}+h^{\frac12})h^{k+1}|u|_{k+1}\|\phi\|_2. \tag{4.28}$$

Next we estimate the term $T_2^*(\phi,E)$. Let $\overline{\mathbf{b}}_T=\dfrac{1}{|T|}(\mathbf{b},1)_T$. Then using the Cauchy-Schwartz inequality we obtain

$$\begin{aligned}
(\phi-Q_h\phi,\nabla\cdot(\mathbf{b}E))=&(\phi-Q_h\phi,(\nabla\cdot\mathbf{b})E)+(\phi-Q_h\phi,\mathbf{b}\cdot\nabla E)\\
=&\sum_{T\in\mathcal{T}_h}\Big((\phi-Q_h\phi,(\nabla\cdot\mathbf{b})E)_T+(\phi-Q_h\phi,(\mathbf{b}-\overline{\mathbf{b}}_T)\cdot\nabla E)\Big)\\
\le&Ch^2\|\phi\|_2|||E|||\\
\le&C\varepsilon^{\frac12}(\varepsilon^{\frac12}+h^{\frac12})h^{k+2}|u|_{k+1}\|\phi\|_2.
\end{aligned}$$

The above estimate and the estimate (4.10) imply that

$$|T_2^*(\phi,E)| \le C(\varepsilon^{\frac{1}{2}}+h^{\frac{1}{2}})h^{k+\frac{3}{2}}|u|_{k+1}\|\phi\|_2. \tag{4.29}$$

Owing to (4.11), one can easily show that

$$|T_3(\phi,E)| \le C(\varepsilon^{\frac{1}{2}}+h^{\frac{1}{2}})h^{k+2}|u|_{k+1}\|\phi\|_2. \tag{4.30}$$

We also infer from the estimates (4.12) and (4.13) that

$$|s_c(E,Q_h\phi)| \le C(\varepsilon^{\frac{1}{2}}+h^{\frac{1}{2}})h^{k+\frac{3}{2}}|u|_{k+1}\|\phi\|_2, \tag{4.31}$$

and

$$|s_d(E,Q_h\phi)| \le C\varepsilon^{\frac{1}{2}}(\varepsilon^{\frac{1}{2}}+h^{\frac{1}{2}})h^{k+1}|u|_{k+1}\|\phi\|_2. \tag{4.32}$$

Combining altogether the estimates above (4.23)-(4.32) gives the desired result (4.3). Thus we complete the proof. □

We aligned the error $e_h = u - u_h = \nu + E + \rho$ where $\nu = u - Q_h u$, $E = Q_h u - R_h u$ and $\rho = R_h u - u_h$. Now we are ready to state and prove the error estimates for semi-discrete MWG-FEM in L^2-norm and $|||\cdot|||$-norm in the following theorems.

Theorem 4.1. *Let $u \in H^{k+1}(\Omega)$ and u_h be the exact solution for the problem* (1.1) *and the solution of the semi-discrete MWG-FEM given by* (2.11)*, respectively. Assume that $\partial_t u, u_0 \in H^{k+1}(\Omega)$. Then there is a constant C such that*

$$\begin{aligned}&\|u-u_h\|^2\\ \le& C\left(\|u(0)-u_h(0)\|^2 + (\varepsilon^{-\frac{1}{2}}h^{\frac{1}{2}}+\varepsilon^{-1}+1)^2h^{2k+2}\left(\|u_0\|_{k+1}^2+\int_0^t \|\partial_t u\|_{k+1}^2\,ds\right)\right).\end{aligned} \tag{4.33}$$

Proof. If we estimate the error ρ, then the desired results follows from the following observation:

$$\|u-u_h\| \le \|\nu\| + \|E\| + \|\rho\|. \tag{4.34}$$

The first and second terms in (4.34) can be estimated by Lemma 2.1 and the estimate (4.3) as follows:

$$\begin{aligned}&\|\nu\| \le Ch^{k+1}\|u\|_{k+1}, \quad \|\partial_t\nu\| \le Ch^{k+1}\|\partial_t u\|_{k+1},\\ &\|E\| \le C(\varepsilon^{-\frac{1}{2}}h^{\frac{1}{2}}+\varepsilon^{-1}+1)h^{k+1}\|u\|_{k+1},\\ &\|\partial_t E\| \le C(\varepsilon^{-\frac{1}{2}}h^{\frac{1}{2}}+\varepsilon^{-1}+1)h^{k+1}\|\partial_t u\|_{k+1}.\end{aligned} \tag{4.35}$$

In order to estimate ρ, by the definition we note that for $v_h \in S_h^0(k)$

$$\begin{aligned}(\partial_t\rho,v_h)+a(\rho,v_h) &= (R_h\partial_t u,v_h)+a(R_h u,v_h)-(\partial_t u_h,v_h)-a(u_h,v_h)\\ &= (R_h\partial_t u,v_h)+a(R_h u,v_h)-(f,v_h)\\ &= (R_h\partial_t u,v_h)-\varepsilon(\Delta u,v_h)+(\nabla\cdot(\mathbf{b}u),v_h)+(cu,v_h)-(f,v_h)\\ &= (R_h\partial_t u,v_h)+(\partial_t Q_h u,v_h)-(\partial_t Q_h u,v_h)-(\partial_t u,v_h)\\ &= -(\partial_t E,v_h)-(\partial_t\nu,v_h).\end{aligned} \tag{4.36}$$

Taking $v_h = \rho$ in (4.36) yields

$$(\partial_t \rho, \rho) + a(\rho, \rho) = -(\partial_t E, \rho) - (\partial_t \nu, \rho), \quad t > 0. \tag{4.37}$$

Using the coercivity of the bilinear form, the Cauchy-Schwartz inequality and the Young's inequality we obtain

$$\frac{1}{2}\frac{\partial \|\rho\|^2}{\partial t} + \gamma |||\rho|||^2 \leq C\Big(\|\partial_t \nu\|^2 + \|\partial_t E\|^2 + \|\rho\|^2\Big),$$

which gives after integration with respect to t

$$\|\rho\|^2 \leq \|\rho(0)\|^2 + C\left(\int_0^t \|\partial_t \nu\|^2\, ds + \int_0^t \|\partial_t E\|^2\, ds + \int_0^t \|\rho\|^2\, ds\right). \tag{4.38}$$

Using Lemma 4.1 we obtain

$$\begin{aligned}
\|\rho(0)\| &= \|u_h(0) - R_h u(0)\| \\
&\leq \|u_h(0) - Q_h u(0)\| + \|Q_h u(0) - R_h u(0)\| \\
&\leq \|u_h(0) - Q_h u(0)\| + C(\varepsilon^{-\frac{1}{2}} h^{\frac{1}{2}} + \varepsilon^{-1} + 1) h^{k+1} \|u_0\|_{k+1} \\
&\leq \|u_0 - u_h(0)\| + \|u_0 - Q_h u(0)\| + C(\varepsilon^{-\frac{1}{2}} h^{\frac{1}{2}} + \varepsilon^{-1} + 1) h^{k+1} \|u_0\|_{k+1} \\
&\leq \|u_0 - u_h(0)\| + C h^{k+1} \|u_0\|_{k+1} + C(\varepsilon^{-\frac{1}{2}} h^{\frac{1}{2}} + \varepsilon^{-1} + 1) h^{k+1} \|u_0\|_{k+1} \\
&\leq \|u_0 - u_h(0)\| + C(\varepsilon^{-\frac{1}{2}} h^{\frac{1}{2}} + \varepsilon^{-1} + 1) h^{k+1} \|u_0\|_{k+1}.
\end{aligned} \tag{4.39}$$

From (4.35) and (4.39) together with the Gronwall lemma, we have the desired result (4.33). The proof is now complete. □

We now give an error estimate for the fully discrete MWG-FEM in the next theorem.

Theorem 4.2. *Let $u \in H^{k+1}(\Omega)$ and U^m be the exact solution of the problem* (1.1) *and the solution of the fully discrete MWG-FEM approximation computed by* (2.28)*, respectively. Assume also that $\partial_t u, u_0 \in H^{k+1}(\Omega)$. Then we have the following error estimate:*

$$\begin{aligned}
\|u^m - U^m\| \leq C\Bigg(&\|u_0 - U^0\| + \delta \int_0^{t_m} \left\|\frac{\partial^2 u}{\partial t^2}\right\| ds + \left(\varepsilon^{-\frac{1}{2}} h^{\frac{1}{2}} + \varepsilon^{-1} + 1\right) h^{k+1} \\
&\times \left(\|u_0\|_{k+1} + \int_0^{t_m} \|\partial_t u\|_{k+1}\, ds\right)\Bigg).
\end{aligned} \tag{4.40}$$

Proof. Similar to the previous proof, we again write

$$\begin{aligned}
\|u(t_m) - U^m\| &= \|u(t_m) - Q_h u(t_m) + Q_h u(t_m) - R_h u(t_m) + R_h u(t_m) - U^m\| \\
&\leq \|u(t_m) - Q_h u(t_m)\| + \|Q_h u(t_m) - R_h u(t_m)\| + \|R_h u(t_m) - U^m\| \\
&=: \|\nu^m\| + \|E^m\| + \|\rho^m\|.
\end{aligned} \tag{4.41}$$

Lemmas 2.1 and 4.1 imply that

$$\|\nu^m\| \leq C h^{k+1} \|u^m\|_{k+1} \leq C h^{k+1}\left(\|u_0\|_{k+1} + \int_0^{t_m} \|\partial_t u\|_{k+1}\, ds\right), \tag{4.42}$$

$$\begin{aligned}
\|E^m\| &\leq C(\varepsilon^{-\frac{1}{2}} h^{\frac{1}{2}} + \varepsilon^{-1} + 1) h^{k+1} \|u^m\|_{k+1} \\
&\leq C(\varepsilon^{-\frac{1}{2}} h^{\frac{1}{2}} + \varepsilon^{-1} + 1) h^{k+1}\left(\|u_0\|_{k+1} + \int_0^{t_m} \|\partial_t u\|_{k+1}\, ds\right).
\end{aligned} \tag{4.43}$$

To estimate ρ^m, we note that for $\forall v_h \in S_h^0(k)$

$$\begin{aligned}(\hat{\partial}_t\rho^m, v_h) + a(\rho^m, v_h) &= (\hat{\partial}_t R_h u^m, v_h) + a(R_h u^m, v_h) - (\hat{\partial}_t U^m, v_h) - a(U^m, v_h)\\ &= (\hat{\partial}_t R_h u^m, v_h) + a(R_h u^m, v_h) - (f(t_m), v_h)\\ &= (\hat{\partial}_t R_h u^m, v_h) - \varepsilon(\Delta u^m, v_h) + (\nabla\cdot(\mathbf{b}u^m), v_h)\\ &\quad + (cu^m, v_h) - (f(t_m), v_h)\\ &= (\hat{\partial}_t R_h u^m, v_h) - (\partial_t u^m, v_h)\\ &= -(\hat{\partial}_t E^m, v_h) - (\hat{\partial}_t \nu^m, v_h) - (\partial_t u^m - \hat{\partial}_t u^m, v_h).\end{aligned}$$

Thus we have

$$(\hat{\partial}_t\rho^m, v_h) + a(\rho^m, v_h) = -(\hat{\partial}_t E^m, v_h) - (\hat{\partial}_t \nu^m, v_h) - (w^m, v_h), \tag{4.44}$$

where

$$w^m = \partial_t u^m - \hat{\partial}_t u^m.$$

Taking $v_h = \rho^m$ in (4.44), we find

$$(\hat{\partial}_t\rho^m, \rho^m) \le (\|\hat{\partial}_t E^m\| + \|\hat{\partial}_t \nu^m\| + \|w^m\|)\|\rho^m\|.$$

Thus we have

$$\|\rho^m\|^2 - (\rho^{m-1}, \rho^m) \le \delta(\|\hat{\partial}_t E^m\| + \|\hat{\partial}_t \nu^m\| + \|w^m\|)\|\rho^m\|,$$

which is equivalent to

$$\|\rho^m\| \le \|\rho^{m-1}\| + \delta(\|\hat{\partial}_t E^m\| + \|\hat{\partial}_t \nu^m\| + \|w^m\|).$$

Induction argument gives

$$\|\rho^m\| \le \|\rho^0\| + \delta\sum_{n=1}^{m}(T_1^n + T_2^n + T_3^n), \tag{4.45}$$

where

$$T_1^n = \|\hat{\partial}_t E^n\|, \quad T_2^n = \|\hat{\partial}_t \nu^n\|, \quad T_3^n = \|w^n\|.$$

As in (4.39), we find the following estimate for $\rho^0 = \rho(0)$:

$$\|\rho^0\| \le \|u_0 - u_h(0)\| + C(\varepsilon^{-\frac{1}{2}}h^{\frac{1}{2}} + \varepsilon^{-1} + 1)h^{k+1}\|u_0\|_{k+1}.$$

From (3.24), we also have

$$\delta\sum_{n=1}^{m} T_3^n \le \sum_{n=1}^{m}\left\|\int_{t_{n-1}}^{t_n}(s - t_{n-1})\frac{\partial^2 u}{\partial t^2}\,ds\right\| \le \delta\int_0^{t_m}\left\|\frac{\partial^2 u}{\partial t^2}\right\| ds. \tag{4.46}$$

Observe that by the definition

$$\begin{aligned}\hat{\partial}_t E^n &= \hat{\partial}_t Q_h u(t_n) - \hat{\partial}_t R_h u(t_n) = -\frac{1}{\delta}\int_{t_{n-1}}^{t_n}(Q_h - R_h)\partial_t u\,ds,\\ -\hat{\partial}_t \nu^n &= -(\hat{\partial}_t u(t_n) - \hat{\partial}_t Q_h u(t_n)) = -\frac{1}{\delta}\int_{t_{n-1}}^{t_n}(I - Q_h)\partial_t u\,ds.\end{aligned}$$

From Lemmas 2.1 and 4.1, we get

$$\begin{aligned} \delta\sum_{n=1}^{m} T_1^n &\le C(\varepsilon^{-\frac{1}{2}}h^{\frac{1}{2}}+\varepsilon^{-1}+1)h^{k+1}\int_0^{t_m}\|\partial_t u\|_{k+1}\,ds,\\ \delta\sum_{n=1}^{m} T_2^n &\le Ch^{k+1}\int_0^{t_m}\|\partial_t u\|_{k+1}\,ds. \end{aligned} \tag{4.47}$$

Combining (4.42), (4.43) ,(4.46) and (4.47), we have the desired result (4.40). This completes the proof. □

Although the numerical experiments in the next section show that the optimal order rate in L^2-norm of order $\mathcal{O}(h^{k+1})$, the optimal order error estimates in L^2-norm in Theorems 4.1 and 4.2 are of order $\mathcal{O}((\varepsilon^{-\frac{1}{2}}h^{\frac{1}{2}}+\varepsilon^{-1}+1)h^{k+1})$ which deteriorates for small $\varepsilon \ll h^2$. To overcome this issue, we imitate the ideas in [6, 7] to get optimal weighted error estimates if $\varepsilon \le h^2$. The improved optimal order estimates hold for for triangulations $\mathcal{T}_h$ made of simplexes T satisfying the simple flow conditions with respect to $\boldsymbol{b}$

$$\begin{aligned} &\text{Each simplex } T \text{ has a unique outflow face with respect to } \boldsymbol{b}, e_T^+.\\ &\text{Each interior face } e_T^+ \text{is included in an inflow face with respect to } \boldsymbol{b}\\ &\text{of another simplex.} \end{aligned} \tag{4.48}$$

We first introduce a weight function. For simplicity we assume $\boldsymbol{b}=(1,0)$. Accordingly, we set, for fixed x_0, y_1 and y_2,

$$\Omega_0 = ((-\infty, x_0]\times[y_1,y_2])\cap\Omega$$

and construct a function ω satisfying

$$\begin{aligned} C_1 \le \omega(x,y) &\le C_2 && \text{for} \quad (x,y)\in\Omega_0,\\ |\omega(x,y)| &\le C_2 e^{-\frac{(x-x_0)}{M\rho}} && \text{for} \quad x\ge x_0+h,\\ |\omega(x,y)| &\le C_2 e^{-\frac{(y-y_2)}{M\sigma}} && \text{for} \quad y\ge y_2+h,\\ |\omega(x,y)| &\le C_2 e^{-\frac{(y_1-y)}{M\sigma}} && \text{for} \quad y\le y_1-h. \end{aligned}$$

Here $\rho\ge 0, \sigma\ge 0$ are parameters that will depend on the mesh size h and ε. We say that ρ is the size of the upwind layer and σ is the size of the crosswind layer. The positive constants C_1, C_2 and M are fixed.

We will use two projections defined as follows:

$$(\Pi^{\pm}u-u, v)_T = 0 \quad \text{for all} \quad v\in\mathbb{P}_{k-1}(T), \tag{4.49}$$

$$\langle \Pi^{+}u-u, w\rangle_{e_T^+} = 0 \quad \text{for all} \quad w\in\mathbb{P}_{k}\left(e_T^+\right), \tag{4.50}$$

$$\langle \Pi^{-}u-u, w\rangle_{e_T^-} = 0 \quad \text{for all} \quad w\in\mathbb{P}_{k}\left(e_T^-\right). \tag{4.51}$$

Using the properties of the projections and results from [8], we have the following result.

Lemma 4.3. *If the triangulation $\mathcal{T}_h$ satisfies the flow condition* (4.48)*, the projections $\Pi^{\pm}$ given by* (4.49) *are well defined. Moreover, if the triangulation $\mathcal{T}_h$ is shape-regular, then, on each simplex $T\in\mathcal{T}_h$ we have*

$$\|\omega(\Pi^{\pm}u-u)\|_{L^2(T)} \le Ch^{k+1}|\omega u|_{H^{k+1}(T)},$$

where C only depends on k and the shape regularity constant.

Theorem 4.3. *Assume that $\epsilon \leq h$ and that the triangulation $\mathcal{T}_h$ is quasi-uniform, that is, assume that there is a parameter $\kappa > 0$ such that*

$$\max_{T\in\mathcal{T}_h}\{h_T\} \leq \kappa \min_{T\in\mathcal{T}_h}\{h_T\}.$$

Assume $\mathcal{T}_h$ satisfies the flow conditions (4.48) *with respect to* $\boldsymbol{b}$. *If $u \in H^{k+1}(\Omega)$ and u_h are the exact solution for the problem* (1.1) *and the solution of the semi-discrete MWG-FEM given by* (2.11), *respectively, then we have*

$$\|\omega\,(u-u_h)\|_{L^2(\Omega)} \leq C\int_0^t L_\varepsilon(u-\Pi^+u)\,ds.$$

Here,

$$L_\varepsilon^2(u) = \left(1+\frac{\epsilon^{\frac{1}{2}}}{h}\right)\|\omega u\|_{L^2(\mathcal{T}_h)} + \epsilon^{\frac{1}{2}}\|\omega\nabla u\|_{L^2(\mathcal{T}_h)} + h\epsilon^{\frac{1}{2}}\left\|\omega D^2 u\right\|_{L^2(\mathcal{T}_h)}, \tag{4.52}$$

where ω is given above with $\rho = \log(\frac{1}{h})h, \sigma = h^{1/2}$ and M is a sufficiently large fixed constant.

Proof. The proof is rather technical and longer, hence we only sketch the main ideas. Note that for any suitable function v, we have

$$\int_0^t (\partial_t v, \omega^2 v)_{L^2(\Omega)} = \frac{1}{2}(\|\omega v\|^2 - \|\omega v(0)\|^2). \tag{4.53}$$

Mimicking the analyses in [14], one can prove that

$$a(u_h, \omega^2 u_h) \leq CM^{-1}|||u_h|||_\omega^2 + CL_\varepsilon^2(u-\Pi^+u), \tag{4.54}$$

where

$$|||u_h|||_\omega^2 = \sum_{T\in\mathcal{T}_h}\|\omega\nabla u_h\|_T^2 + s_d^2(\omega v_h, \omega v_h) + \|\omega u_h\|_s^2.$$

First, we aligned the error as $u - u_h = u - \Pi^+u + \Pi^+u - u_h =: \eta + E_h$ and Lemma 4.3 gives the bound for the term $\omega\eta$. Thus we only need to estimate E_h.

From (4.53) and (4.54), we have, for sufficiently large M,

$$\|\omega E_h\|^2 \leq \|\omega E_h(0)\|^2 + C\int_0^t L_\varepsilon^2(u-\Pi^+u)\,ds.$$

Since $E_h(0) = 0$, we conclude the result. $\square$

The result immediately implies that optimal error estimates can be achieved when $\varepsilon \leq h^2$. This result and the numerical experiments in the next section demonstrate that the optimal error estimate is of order $\mathcal{O}(h^{k+1})$ in L^2-norm.

5. Numerical Experiments

We present various numerical examples to show the performance and efficiency of the MWG-FEM in this section. The error is computed for the backward Euler MWG-FEM solution in the following norms:

$$\|u(t_m) - U^m\|^2 = \sum_{T \in \mathcal{T}_h} \int_T |u(t_m) - U^m|^2 \, dx,$$

$$|||u - u_h|||^2 = \delta \sum_{n=0}^{N} |||u(t_n) - U^n|||^2,$$

where $\delta N = T$ and the triple-norm defined by (2.15).

Example 5.1 (Smooth Solution). We first consider the problem (1.1) in $Q_T = [0,1]^2 \times (0,1]$ and $\mathbf{b} = (1,1)^T$ with the different value of small parameter ε. We choose f and u_0 such that the true (smooth) solution is

$$u(x,y,t) = \exp(-t)\sin(2\pi x)\sin(2\pi y).$$

We first test the temporal convergence of the backward Euler MWG-FEM (2.29). The numerical experiments are carried out on the uniform triangulation mesh with 64×64 elements and the quadratic polynomials $k = 2$ for the time step $\delta = 2^{-n}, n = 3,4,5,6$. We show the history of convergence results in the L^2- norm estimates and the order of convergence (OC) for the MWG-FEM solution u_h at the final time in Table 5.1. The result shows that the method converges in L^2- norm of first order in time which confirms the theoretical analysis stated in (3.17).

Table 5.1: L^2- norm errors and OC of the backward Euler MWG-FEM for Example 5.1 at the final time for a fixed triangular mesh $h = 1/64$ using $\mathbb{P}_2$ element.

	$\varepsilon = 1$		$\varepsilon = 10^{-3}$		$\varepsilon = 10^{-9}$	
δ	$\|u(t_m) - U^m\|$	OC	$\|u(t_m) - U^m\|$	OC	$\|u(t_m) - U^m\|$	OC
1/8	3.4278e-02	-	5.5438e-02	-	6.7543e-02	-
1/16	1.9352e-02	0.8248	3.1692e-02	0.8067	3.8717e-02	0.8028
1/32	1.0249e-02	0.9169	1.6721e-02	0.9224	1.9825e-02	0.9656
1/64	5.1206e-03	1.0010	8.3011e-03	1.0102	9.8686e-03	1.0064

The uniform triangles meshes are employed with $M \times M$ elements for $M = 2,4,8,16,32,64$. The uniform meshes are used for the time discretization with the time step $\delta = 0.0001$ which is sufficiently small to ensure the convergence. We report the errors and OC in Tables 5.2 and 5.3 for the different values of the diffusion coefficient $\varepsilon = 10^{-r}, r = 0,3,9$ using polynomials of order $k = 1,2$ for the fixed time $\delta = 0.0001$. We observe that the optimal order of $\mathcal{O}(h^{k+1})$ error in L^2- norm and the optimal order of $\mathcal{O}(h^{k+1/2})$ error in $|||\cdot|||_\varepsilon$ norm which confirms the error estimation in the theoretic results.

Example 5.2 (Boundary Layer). We next consider the following BVP with boundary layers to show the efficiency of the MWG-FEM:

$$\begin{aligned}
&\partial_t u - \varepsilon \Delta u + \nabla \cdot (\mathbf{b} u) + cu = f(x,y,t) && \text{in } Q_T = \Omega \times (0,T],\\
&u = 0 && \text{on } \partial\Omega \times (0,T],\\
&u = u_0(x,y) && \text{in } \Omega \times \{0\},
\end{aligned}$$

Table 5.2: L^2- norm errors and OC of the MWG-FEM for Example 5.1 at the final time for a fixed $\delta = 0.0001$ using $\mathbb{P}_1$ and $\mathbb{P}_2$.

		$\varepsilon = 1$		$\varepsilon = 10^{-3}$		$\varepsilon = 10^{-9}$	
k	M	$\|u(t_m) - U^m\|$	OC	$\|u(t_m) - U^m\|$	OC	$\|u(t_m) - U^m\|$	OC
1	2	9.5869e-01	-	9.9742e-01	-	9.9851e-01	-
	4	2.5436e-01	1.9145	2.6562e-01	1.9088	2.6516e-01	1.9129
	8	6.5283e-02	1.9530	7.1837e-02	1.9440	7.1839e-02	1.9121
	16	1.6502e-02	1.9839	1.7456e-02	1.9835	1.8610e-02	1.9486
	32	4.1593e-03	1.9883	4.4431e-03	1.9740	4.4561e-03	2.0286
	64	1.0378e-03	2.0028	1.1094e-03	2.0017	1.1094e-03	2.0248
2	2	4.7654e-01	-	5.8564e-01	-	5.9853e-01	-
	4	5.1046e-02	3.2227	6.2451e-02	3.2292	6.3681e-02	3.2324
	8	6.1869e-03	3.0445	7.5580e-03	3.0466	7.7362e-03	3.0411
	16	7.8354e-04	2.9811	9.5438e-04	2.9853	9.7743e-04	2.9845
	32	9.8466e-05	2.9923	1.1987e-04	2.9930	1.2234e-04	2.9980
	64	1.2353e-05	2.9947	1.5001e-05	2.9983	1.5298e-05	2.9994

Table 5.3: $|||\cdot|||_\varepsilon-$ norm errors and OC of the MWG-FEM for Example 5.1 at the final time for a fixed $\delta = 0.0001$ using $\mathbb{P}_1$ and $\mathbb{P}_2$.

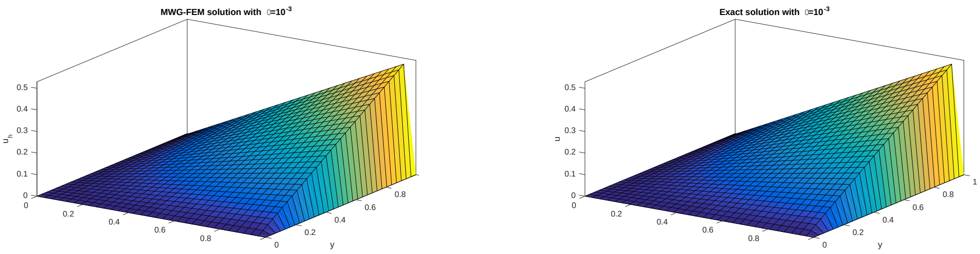
		$\varepsilon = 1$		$\varepsilon = 10^{-3}$		$\varepsilon = 10^{-9}$	
k	M	$\|\|\|u - u_h\|\|\|_\varepsilon$	OC	$\|\|\|u - u_h\|\|\|_\varepsilon$	OC	$\|\|\|u - u_h\|\|\|_\varepsilon$	OC
1	2	8.5869e+00	-	5.1672e-01	-	7.2436e-01	-
	4	5.4213e+00	0.6634	2.1241e-01	1.2825	3.1048e-01	1.2222
	8	3.0638e+00	0.8233	1.0241e-01	1.0524	1.1562e-02	1.4251
	16	1.5638e+00	0.9702	4.9879e-02	1.0378	3.8873e-03	1.5725
	32	7.6294e−01	1.0354	2.1932e-02	1.1853	1.2457e-03	1.6418
	64	3.7059e−01	1.0417	9.5848e-03	1.1942	4.0139e-04	1.6338
2	2	4.2543e+00	-	2.1241e-01	-	2.3657e-01	-
	4	1.0230e+00	2.0561	4.8256e-02	2.1380	4.1532e-02	2.5099
	8	2.2375e−01	2.1928	1.1001e-02	2.1330	6.5521e-03	2.6641
	16	5.1268e−02	2.1257	2.5282e-03	2.1214	1.1486e-03	2.5120
	32	1.1879e−02	2.1096	5.7639e-04	2.1329	2.0241e-04	2.5045
	64	2.7094e−03	2.1323	1.3255e-04	2.1205	3.5087e-05	2.5282

where $\Omega = (0,1)^2, \mathbf{b} = (1,1)^T, T = 1$ and $c = 1$. We choose the initial function $u_0(x,y)$ and the forcing function $f(x,y,t)$ such that the exact solution

$$u(x,y,t) = (1 - \exp(-t))\beta(x)\beta(y),$$

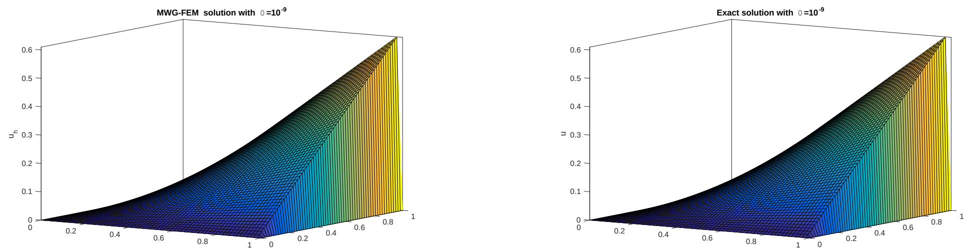
where $\beta(z) = (\exp(-1/\varepsilon) - 1)z - \exp(-1/\varepsilon) + \exp(-(1-z)/\varepsilon)$. The solution exhibits two boundary layers of width $\mathcal{O}(\varepsilon)$ along the sides $x = 1$ and $y = 1$.

We divide the domain Ω into $M \times M$ squares. Then the triangular mesh is constructed by partitioning the each square into triangles by a diagonal line. Let $h = 1/M$ be the mesh size for the number of elements M in each direction. Linear finite elements are employed on uniform meshes both in the space and in the time discretization with the time step $\delta = 1/N$ with N mesh intervals in the time direction. We plot the exact solution and MWG-FEM solution in Figs. 5.1 and 5.2 with $\varepsilon = 10^{-3}$ and $\varepsilon = 10^{-9}$.



MWG-FEM solution with $\varepsilon = 10^{-3}$ The exact solution with $\varepsilon = 10^{-3}$

Fig. 5.1. The exact solution and the MWG-FEM solution of Example 5.2 with $\varepsilon = 10^{-3}$ using linear elements over a mesh of 64×64 at $\delta = 0.01$.



MWG-FEM solution with $\varepsilon = 10^{-9}$ The exact solution with $\varepsilon = 10^{-9}$

Fig. 5.2. The exact solution and the MWG-FEM solution of Example 5.2 with $\varepsilon = 10^{-9}$ using linear elements over a mesh of 64×64 at $\delta = 0.01$.

We report the errors and the order of convergence (OC) in L^2- norm and $|||\cdot|||_\varepsilon$ norms in Table 5.5 and Table 5.6. From these results, we see that the optimal order of convergence of $\mathcal{O}(h^{k+1})$ in L^2 norm for $\varepsilon = 10^{-r}, r = 0, 3, 9$. On the other hand, we achieve numerically the optimal convergence order of $\mathcal{O}(h^k)$ and $\mathcal{O}(h^{k+1/2})$ in the energy norm $|||e_h|||_\varepsilon$ for $\varepsilon = 10^{-3}$ and $\varepsilon = 10^{-9}$, respectively, which confirms theoretical results obtained in this paper.

Table 5.4: L^2- norm errors and OC of the backward Euler MWG-FEM for Example 5.2 at the final time for a fixed triangular mesh $h = 1/64$ using $\mathbb{P}_2$ element.

	$\varepsilon = 1.0$		$\varepsilon = 10^{-9}$	
δ	$\|u(t_m) - U^m\|$	OC	$\|u(t_m) - U^m\|$	OC
$8.00e-02$	7.4384e-06	-	1.0467e-03	-
$4.00e-02$	3.6241e-06	1.0373	4.9768e-04	1.0725
$2.00e-02$	1.7380e-06	1.0601	2.3851e-04	1.0611
$1.00e-02$	8.6539e-07	1.0060	1.1923e-04	1.0003

Example 5.3 (Rotating pulse). We next consider the following variable coefficient convection diffusion equation:

$$\begin{aligned} &\partial_t u - \varepsilon \Delta u + \nabla \cdot (\mathbf{b} u) + cu = f && \text{in } \Omega \times (0, \pi/4], \\ &u = 0 && \text{on } \partial\Omega \times (0, T], \\ &u = u_0 && \text{in } \Omega \times \{0\}, \end{aligned} \tag{5.1}$$

Table 5.5: L^2- norm errors and OC of the MWG-FEM for Example 5.2 at the final time for $\delta = M^{-(k+1)}$, $k = 1, 2$ using $\mathbb{P}_1$ and $\mathbb{P}_2$.

		$\varepsilon = 1$		$\varepsilon = 10^{-3}$		$\varepsilon = 10^{-9}$	
k	M	$\Vert u(t_m) - U^m\Vert$	OC	$\Vert u(t_m) - U^m\Vert$	OC	$\Vert u(t_m) - U^m\Vert$	OC
1	2	2.9874e-04	-	5.7435e-03	-	8.9851e-03	-
	4	7.3218e-05	2.0286	1.4345e-03	2.0013	2.2926e-03	1.9705
	8	1.8056e-05	2.0197	3.5440e-04	2.0170	6.0102e-04	1.9314
	16	4.5101e-06	2.0012	8.8591e-05	2.0001	1.5214e-04	1.9820
	32	1.0672e-06	2.0793	2.1948e-05	2.0130	3.8407e-05	1.9859
	64	2.6591e-07	2.0048	5.4679e-06	2.0050	9.5546e-05	2.0071
2	2	1.5347e-05	-	1.4617e-04	-	2.5012e-04	-
	4	1.9562e-06	2.9718	1.7969e-05	3.0240	3.4935e-05	2.8398
	8	2.4101e-07	3.0208	2.2172e-06	3.0186	4.3814e-06	2.9952
	16	3.0042e-08	3.0040	2.7698e-07	3.0008	5.4520e-07	3.0065
	32	3.5189e-09	3.0937	3.3607e-08	3.0429	6.7391e-08	3.0161
	64	4.3502e-10	3.0159	4.1612e-09	3.0137	8.3072e-09	3.0201

Table 5.6: $|||\cdot|||_\varepsilon-$ norm errors and OC of the MWG-FEM for Example 5.2 at the final time for $\delta = M^{-2}$ using $\mathbb{P}_1$.

	$\varepsilon = 10^{-3}$		$\varepsilon = 10^{-9}$	
M	*Error*	OC	*Error*	OC
2	3.2650e+00	-	3.3540e+00	-
4	1.5846e+00	1.0429	1.11865e+00	1.4990
8	7.8234e−01	1.0182	4.1912e−01	1.5000
16	3.8639e−01	1.0177	1.4402e−01	1.5097
32	1.9068e−01	1.0189	5.0810e−02	1.5030
64	9.1846e−02	1.0538	1.7924e−02	1.5032

where $u = u(x, y, t)$, $f = f(x, y, t)$ and $\Omega = (-\frac{1}{2}, \frac{1}{2})^2$. Let the convection field be a pure rotation $\mathbf{b} = [-4y, 4x]$, $f = c = 0$ and the initial function u_0 be given as (see also Fig. 5.3)

$$u_0(x, y) = \exp\Big(-\frac{(x + 0.2)^2 + y^2}{2(0.1)^2}\Big).$$

In Figs. 5.4 and 5.5, we plot the exact solution u and numerical solution u_h of the convection diffusion problem (5.1) with the rotating pulse for $\varepsilon = 10^{-2}$ and $\varepsilon = 10^{-4}$ over a uniform 64×64 bilinear mesh with the time step $\delta = 0.01$.

We note that the proposed MWG-FEM presents poor convergence at the boundary layer when $\varepsilon = 10^{-3}$, however, it converges very good in the case when $\varepsilon = 10^{-9}$. Discontinuous Galerkin or SUPG converges poorly in the case when the diffusion parameter is in the intermediate regimes (for example, $\varepsilon = 10^{-3}$ see [1]). The numerical experiments show that the MWG-FEM is a stable numerical method.

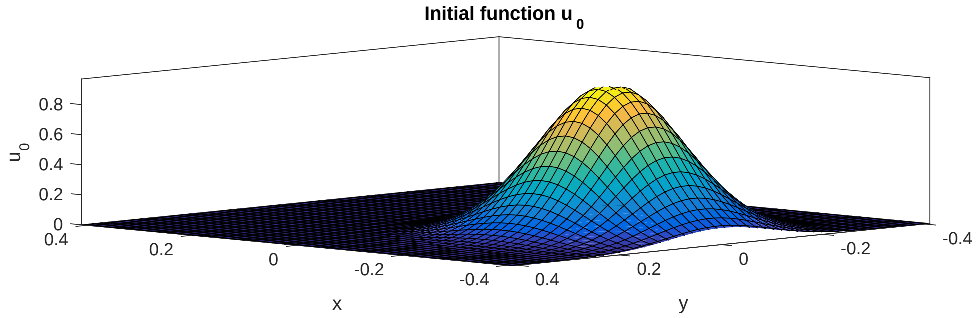
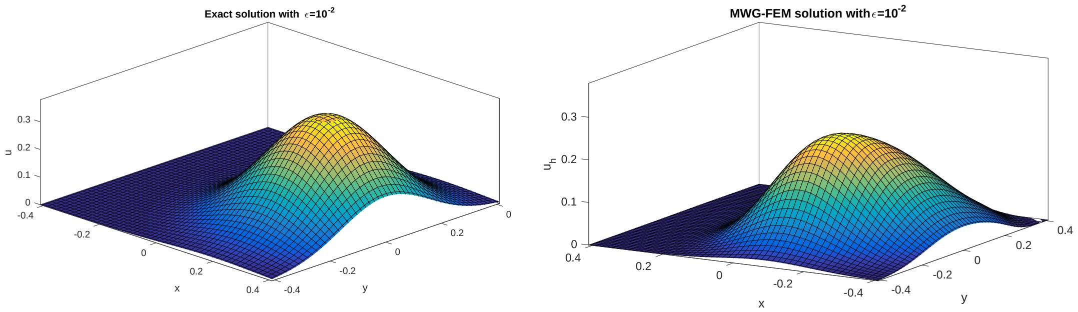
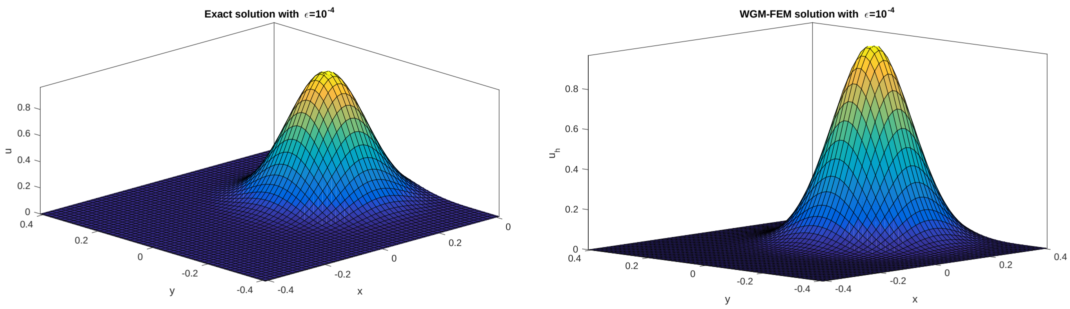


Fig. 5.3. Initial function u_0: Rotating pulse.



The exact solution with $\varepsilon = 10^{-2}$ MWG-FEM solution with $\varepsilon = 10^{-2}$

Fig. 5.4. The exact and MWG-FEM solutions of the problem (5.1) with $\varepsilon = 10^{-2}$ using linear elements over a mesh of 64×64 at $t = \pi/4$.



The exact solution with $\varepsilon = 10^{-4}$ MWG-FEM solution using linear elements over a mesh of 64×64.

Fig. 5.5. The exact and MWG-FEM solutions of the problem (5.1) with $\varepsilon = 10^{-4}$ at $t = \pi/4$.

We emphasize that we use uniform triangulation in all computations for the purpose of ease of construction of meshes. However, one can use polygonal meshes generated by PolyMesher [25]. Polygonal meshes can be defined on rectangular domains and on L-shaped domains as well. For more details, we refer the reader to [25].

Acknowledgments. The second authors research is supported in part by National Natural Science Foundation of China (No. 11871038).

References

[1] B. Ayuso de Dios and D. Marini, Discontinuous Galerkin Methods for Advection-Diffusion-Reaction Problems,*SIAM J. Numer. Anal.*, **47**:2 (2009), 1391–1420.

[2] K. Aziz and A. Settari, *Petroleum Reservoir Simulation*, Applied Science Publisher Ltd., London, 1979.

[3] J. Bear, *Dynamics of Fluids in Porous Media*, American Elsevier Publishing Company, New York, 1972.

[4] A.N. Brooks and T.J. Hughes, Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations, *Comput. Methods Appl. Mech. Engrg*, **32** (1982), 199–259.

[5] G. Chen, M. Feng and X. Xie, A robust WG finite element method for convection-diffusion-reaction equations, *J. Comput. Appl. Math*, **315** (2017), 107–125.

[6] B. Cockburn, B. Dong and J. Guzmán, Optimal convergence of the original DG method for the transport-reaction equation on special meshes , *SIAM J. Numer. Anal.*, **46(3)** (2008), 1250–1265.

[7] B. Cockburn, B. Dong, J. Guzmán and J. Qian, Optimal convergence of the original dg method on special meshes for variable transport velocity, , *SIAM J. Numer. Anal.*, **48** (2010), 133–146.

[8] B. Cockburn, B. Dong and J. Guzmán, A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems, *Math. Comp.*, **77(264)** (2008), 1887–1916.

[9] F. Gao, S. Zhang, P. Zhu, Modified weak Galerkin method with weakly imposed boundary condition for convection-dominated diffusion equations, *Appl. Numer. Math.* , **157** (2020), 490–504.

[10] F. Gao and L. Mu, On L^2-error estimate for weak Galerkin finite element methods for parabolic problems, *Journal of Computational Mathematics*, **32** (2014), 195–204.

[11] F. Gao and X. Wang, A modified weak Galerkin finite element method for a class of parabolic problems, *J. Comput. Appl. Math*, **271** (2014), 1–19.

[12] F. Gao, X. Wang and L. Mu, A modified weak Galerkin finite element methods for convection-diffusion problems in 2D, *J. Appl. Math. Comput.*,**49** (2015), 493–511.

[13] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, 1998.

[14] J. Guzman, Local analysis of discontinuous Galerkin methods applied to singularly perturbed problems, *J. Numer. Math.*, **14** (2006), 41–56.

[15] P. Knobloch and G. Lube, Local projection stabilization for advection-diffusion-reaction problems: One-level vs. two-level approach, *Appl. Numer. Math.* , **59** (2009), 2891–2907.

[16] Q.H. Li and J.Wang, Weak Galerkin finite element methods for parabolic equations:, *Numer. Methods Partial Differential Equations*, **29** (2023), 2004–2024.

[17] R. Lin, X. Ye, S. Zhang and P. Zhu, A weak Galerkin finite element method for singularly perturbed convection-diffusion-reaction problems, *SIAM J. Numer. Anal.*, **56**:3 (2018), 1482–1497.

[18] T. Linss, *Layer-Adapted Meshes for Reaction-Convection-Diffusion Problems*, Spring Heidelberg Dordrecht London, New York, 2009.

[19] G. Matthies, P. Skrzypacz and L. Tobiska, Stabilization of local projection type applied to convection-diffusion problems with mixed boundary conditions, *Electronic Transactions on Numerical Analysis ETNA*, **32** (2008), 90–105.

[20] L. Mu, X. Wang and X. Ye, A modified weak Galerkin finite element methods for the Stokes equations, *J. Comput. Appl. Math.*, **275** (2015), 79–90.

[21] L Mu, J. Wang, X. Ye and S. Zhao, A computational study of the weak Galerkin method for the second order elliptic equations, *Numer. Algorithms*, **63** (2013), 753–777.

[22] 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**:5 (2014), 1461–1479.

[23] D.A. Pietro and A, Ern, *Mathematical Aspects of Discontinuous Galerkrin Methods*, Springer-Verlag, Berlin, 2012.

[24] H.G. Roos, M. Stynes and L. Tobiska, *Robust Numerical Methods for Singularly Perturbed Differential Equations*, Springer, Berlin, 2008.

[25] C. Talischi, G.H. Paulino, A. Pereira and I.F.M. Menezes, PolyMesher: a general-purpose mesh generator for polygonal elements written in Matlab, *Struct. Multidiscip. Optim.* **45**:3 (2012), 309–328.

[26] T. Tian, Q. Zhai and R. Zhang, A new modified weak Galerkin finite element scheme for solving the stationary Stokes equations, *J. Comput. Appl. Math.*, **329** (2018), 268–279.

[27] Ş. Toprakseven, A weak Galerkin finite element method for time fractional reaction-diffusion-convection problems with variable coefficients, *Appl. Numer. Math.* , **168** (2021), 1–12.

[28] Ş. Toprakseven and P. Zhu, Uniform convergent modified weak Galerkin method for convection-dominated two-point boundary value problems. *Turkish J. Math.*, **45** (2021), 2703–2730.

[29] Ş. Toprakseven, Superconvergence of a modified weak Galerkin method for singularly perturbed two-point elliptic boundary-value problems. *Calcolo*, **59**:1 (2022), 1–35.

[30] H. Wang, D. Liang, R.E. Ewing, S.L. Lyons and G. Quin, An approximation to miscible fluid flows in porous media with point sources and sinks by an Eulerian-Lagrangian localized adjoint method and mixed finite element methods, *SIAM J. Sci. Comput.*, **22** (2000), 561–581.

[31] X. Wang, N.S. Malluwawadu, F. Gao and T.C. McMillan, A modified weak Galerkin finite element method, *J. Comput. Appl. Math.*, **271** (2014), 319–327.

[32] J. Wang and X. Ye, A weak Galerkin mixed finite element method for second order elliptic problems, *J. Comput. Appl. Math.*, **241** (2013), 103–115.

[33] M.F. Wheeler, A priori L^2 error estimates for Galerkin approximations to parabolic partial differential equations, *SIAM J. Numer. Anal.* **10** (1973), 723–759.

[34] S. Xie, P. Zhu and X. Wang, Error analysis of weak Galerkin finite element methods for time-dependent convection-diffusion equations, *Appl. Numer. Math*, **137** (2019), 19–33.

[35] H. Zarin and H. Roos, Interior penalty discontinuous approximations of convection-diffusion problems with parabolic layers, *Numerische Mathematik*, **100** (2005), 735–759.