Electronic Research Archive

http://www.aimspress.com/journal/era

ERA, 32(8): 5033–5066.
DOI: 10.3934/era.2024232
Received: 17 July 2024
Revised: 17 August 2024
Accepted: 19 August 2024
Published: 22 August 2024

Research article

A weak Galerkin finite element method for parabolic singularly perturbed convection-diffusion equations on layer-adapted meshes

Suayip Toprakseven[1,*] **and Seza Dinibutun**[2]

[1] Department of Mathematics,Yozgat Bozok University, Yozgat, Turkey

[2] Department of Mathematics and Natural Sciences, International University of Kuwait, Ardiya, Kuwait

* **Correspondence:** Email: suayip.toprakseven@yobu.edu.tr; Tel: +904662121000.

Abstract: In this paper, we designed and analyzed a weak Galerkin finite element method on layer adapted meshes for solving the time-dependent convection-dominated problems. Error estimates for semi-discrete and fully-discrete schemes were presented, and the optimal order of uniform convergence has been obtained. A special interpolation was delicately designed based on the structures of the designed method and layer-adapted meshes. We provided various numerical examples to confirm the theoretical findings.

Keywords: time dependent singularly perturbed problem; weak Galerkin method; semi-discrete and fully discrete schemes; layer-adapted-meshes

1. Introduction

We will present a weak Galerkin finite element method for the following parabolic singularly perturbed convection-reaction-diffusion problem:

$$\begin{cases} \partial_t u - \varepsilon\Delta u + \mathbf{b}\cdot\nabla u + cu = f(x,y,t) & \text{in } \ \Omega\times(0,T], \\ u = 0 & \text{on } \ \partial\Omega\times(0,T], \\ u(x,0) = u^0 & \text{in } \ \overline{\Omega}, \end{cases} \tag{1.1}$$

where $\Omega = (0,1)^2$, $0 < \varepsilon \ll 1$, and $T > 0$ is a constant. Assume $\mathbf{b} = \mathbf{b}(x,y)$, $c = c(x,y)$, and $u^0 = u^0(x,y)$ are sufficiently smooth functions on Ω, and

$$\mathbf{b} = (b_1(x,y), b_2(x,y)) \geq (\beta_1, \beta_2), \quad c - \frac{1}{2}\nabla\cdot\mathbf{b} \geq c_0 > 0 \ \text{ on } \Omega, \tag{1.2}$$

for some constants β_1, β_2, and c_0. The parabolic convection-dominated problem (1.1) has been utilized in a broad range of applied mathematics and engineering including fluid dynamics, electrical engineering, and the transport problem [1, 2].

In general, the solution of the problem (1.1) will have abrupt changes along the boundary. In other words, the solution exhibits boundary/interior layers near the boundary of Ω. We are only interested in the boundary layers by excluding the interior layers which can be accomplished by assuming some extra compatible conditions on the data; see, e.g., [1, 3]. The standard numerical schemes including the finite element method give unsatisfactory numerical results due to the boundary layers. Some nonphysical oscillations in the numerical solution can occur even on adapted meshes, and it is not easy to solve efficiently the resulting discrete system [4]. There are many numerical schemes for solving convection-dominated problems efficiently and accurately in the literature. These methods include Galerkin finite element methods [5–7], weak Galerkin finite element methods (WG-FEMs) [8–10], the streamline upwind Petrov-Galerkin (SUPG) [11, 12], and the discontinuous Galerkin (DG) methods [13–15]. Among these numerical methods, the standard WG-FEM introduced in [16] is also an effective and flexible numerical algorithm for solving PDEs. Recently, the WG methods demonstrate robust and stable discretizations for singularly perturbed problems. In fact, while the WG-FEM and the hybridizable discontinuous Galerkin share something in common, the WG-FEM seems more appropriate for solving the time dependent singularly perturbed problems since the inclusion of the convective term in the context of hybridized methods is not straightforward and makes the analysis more subtle. Errors estimates of arbitrary-order methods, including the virtual element method (VEM), are typically limited by the regularity of the exact solution. A distinctive feature of the WG-FEM lies in its use of weak function spaces. Moreover, hybrid high-order (HHO) methods have similar features with WG-FEMs. In fact, the reconstruction operator in the HHO method and the weak gradient operator in WG methods are closely related, and that the main difference between HHO and WG methods lies in the choice of the discrete unknowns and the design of the stabilization operator [17]. Notably, in weak Galerkin methods, weak derivatives are used instead of strong derivatives in variational form for underlying PDEs and adding parameter free stabilization terms. Considering the application of the WG method, various PDEs arising from the mathematical modeling of practical problems in science are solved numerically via WG-FEMs using the concept of weak derivatives. There exist many papers on such PDEs including elliptic equations in [16, 18, 19], parabolic equations [20–22], hyperbolic equations [23, 24], etc.

However, to the best of the author's knowledge, there is no work regarding the uniform convergence results of the fully-discrete WG-FEM for singularly perturbed parabolic problems on layer-adapted meshes. This paper uses three layer-adapted meshes defined through mesh generating functions, namely, Shishkin-type meshes, Bakhvalov-Shishkin type meshes and Bakhvalov-type meshes given in [25]. The error estimates in this work show that one has optimal order of convergence for Bakhvalov-Shishkin type meshes and Bakhvalov-type meshes while almost optimal convergence for Shishkin-type meshes. The main ingredient of the error analysis is the use of the vertices-edges interpolation of Lin [26]. The main advantage of this interpolation operator is that we have sharper error bounds compared with the classical interpolation operators. For the sake of simplicity, the Crank–Nicolson method is used for time discretization. This scheme yields optimal order estimates for fully-discrete WG-FEM. As an alternative, one can apply a discontinuous Galerkin method in time and present optimal order estimates for the fully-discrete scheme [27].

The rest of the paper is organized as follows. In Section 2, we introduce some notations and recall some definitions. The regularity of the solution is also summarized and three layer-adapted meshes have been introduced in Section 2. Also, we define the weak gradient and weak convection operators, and using these concepts we define our bilinear forms. In Section 3, the semi-discrete WG-FEM and its stability results have been presented. Section 4 introduces a special interpolation operator and analyses interpolation error estimates. Section 5 presents error analysis of the semi-discrete WG-FEM for the problem (1.1) on the layer-adapted meshes. In Section 6, we apply the Crank-Nicolson scheme on uniform time mesh in time to obtain the fully-discrete WG-FEM, and prove uniform error estimates on the layer-adapted meshes. In Section 7, we conduct some numerical examples to validate the robustness of the WG-FEM for the problem (1.1). Summary on the contributions of this work are presented in Section 8.

2. Preliminaries and weak Galerkin finite element method

Let S be a measurable subset of Ω. We shall use the classical Sobolev spaces $W^{r,q}(S), H^r(S) = W^{r,2}(S), H_0^r(S), L^q(S) = W^{0,q}(S)$ for negative integers $r > 0$ and $1 \leq q \leq \infty$, and $(\cdot,\cdot)_S$ for the L^2 inner product on S. The semi-norm and norm on $H^r(S)$ are denoted by $|\cdot|_{r,S}$ and $\|\cdot\|_{r,S}$, respectively. If $S = \Omega$, we do not write S in the subscript. Throughout the study, we shall use C as a positive generic constant, which is independent of the mesh parameters h and ε.

2.1. Decomposition of the solution

This section deals with the introduction of a decomposition of the solution which provides a priori information on the exact solution and its derivatives. Based on this solution decomposition, we construct layer-adapted meshes. As we noted in the introduction, the solution of (1.1) exhibits typically two exponential boundary layers at $x = 1$ and $y = 1$. The following lemma gives some information on the solution decomposition and bounds on the solution of (1.1) and its derivatives.

Lemma 2.1. *Let k be positive integer and $l \in (0,1)$. Assume that the solution u of the problem (1.1) belongs to the space $C^{k+l}(Q_T)$ where $Q_T := \Omega \times (0,T]$. Assume further that the solution u can be decomposed into a smooth part u_R and layer components u_{L_0}, u_{L_1}, and u_{L_1} with*

$$u = u_R + u_L, \quad u_L = u_{L_0} + u_{L_1} + u_{L_2}, \quad \forall (x,y) \in \overline{\Omega}, \tag{2.1}$$

where the smooth and layer parts satisfy

$$\left| \frac{\partial^{i+j+r} u_R}{\partial^i x \partial^j y \partial t^r}(x,y) \right| \leq C \tag{2.2}$$

$$\left| \frac{\partial^{i+j+r} u_{L_0}}{\partial^i x \partial^j y \partial t^r}(x,y) \right| \leq C\varepsilon^{-i} e^{-\beta_1(1-x)/\varepsilon,} \tag{2.3}$$

$$\left| \frac{\partial^{i+j+r} u_{L_1}}{\partial^i x \partial^j y \partial t^r}(x,y) \right| \leq C\varepsilon^{-j} e^{-\beta_2(1-y)/\varepsilon}, \tag{2.4}$$

$$\left| \frac{\partial^{i+j+r} u_{L_2}}{\partial^i x \partial^j y \partial t^r}(x,y) \right| \leq C\varepsilon^{-(i+j)} e^{-\beta_1(1-x)/\varepsilon} e^{-\beta_2(1-y)/\varepsilon}, \tag{2.5}$$

for any $(x,y) \in \overline{\Omega}$, $t \in [0,T]$, *and positive integers* i, j, r *with* $i + j + 2r \leq k$, *and* C *only depends on* $\boldsymbol{b}, c$, *and* f *and is independent of* ε. *Here,* u_R *is the regular part of* u, u_{L_0} *is an exponential boundary layer along the side* $x = 1$ *of* Ω, u_{L_1} *is an exponential boundary layer along the side* $y = 1$, *while* u_{L_2} *is an exponential corner layer at* $(1,1)$.

Proof. Under some smoothness conditions and strong imposed compatibility requirements, Shishkin proved this solution decomposition; see, [1]. □

2.2. Layer-adapted meshes

Let N_x and N_y be positive integers. For the sake of simplicity, we assume that $N_x = N_y = N$ is an even integer number. We shall construct the tensor product mesh $\mathcal{T}_N = \{T_{ij}\}_{i,j=1,2...,N}$ in $\overline{\Omega}$ consisting of elements $T_{ij} = I_i \times K_j$ with the intervals $I_i = (x_{i-1}, x_i)$ and $K_j = (y_{j-1}, y_j)$, where the mesh points are defined by

$$0 = x_0 < x_1 < \ldots, x_{N_x} = 1, \; 0 = y_0 < y_1 < \ldots, y_{N_y} = 1.$$

Since the construction of the meshes in both directions is similar, the mesh construction in x-variable is given here.

We define the transition parameter as

$$\tau = \min\left(\frac{1}{2}, \frac{\sigma\varepsilon}{\beta_1}\varphi(1/2)\right),$$

where $\sigma \geqslant p + 1$ is a positive constant. Here, p is the degree of the polynomials used in the approximation space. The function φ obeys the conditions

$$\varphi(0) = 0, \; \varphi' > 0, \; \; \varphi'' \geq 0. \tag{2.6}$$

Assumption 1. Throughout this article, we assume that $\varepsilon \leq CN^{-1}$ such that $\tau = \frac{(p+1)\varepsilon}{\beta_1}\varphi(1/2)$, since otherwise the analysis can be carried out using uniform mesh.

Let the mesh points x_i be equally distributed in $[0, x_{N/2}]$ with $N/2$ intervals and distributed $[x_{N/2}, 1]$ with $N/2$ intervals using the mesh generating function defined by

$$x_i = \lambda(i/N) = \begin{cases} 2(1-\tau)i/N, & \text{for } i = 0, 1, \ldots, N/2, \\ 1 - \frac{(k+1)\varepsilon}{\beta_1}\varphi(1 - i/N), & \text{for } i = N/2, N/2+1, \ldots, N. \end{cases} \tag{2.7}$$

For example, as in [25], the Shishkin-type (S-type) meshes can be deduced by $\varphi(1/2) = \ln N$ while Bakhvalov-type meshes (B-type) can be recovered by taking $\varphi(1/2) = \ln(1/\varepsilon)$.

We will use the mesh characterizing function ψ defined by $\psi = e^{-\varphi}$, which is an essential tool in our analysis below.

Following [2], we list some famous adaptive meshes including S-type, Bakhvalov-Shishkin meshes (BS-mesh), and B-type in Table 1.

Table 1. Frequently used layer-adapted meshes.

	S-type	BS	B-type
$\varphi(t)$	$2t\ln N$	$-\ln\left[1-2\left(1-N^{-1}\right)t\right]$	$-\ln[1-2(1-\varepsilon)t]$
$\psi(t)$	N^{-2t}	$1-2\left(1-N^{-1}\right)t$	$1-2(1-\varepsilon)t$
$\max\lvert\psi'\rvert$	$C\ln N$	C	C

Similarly, we define the transition point in the y-direction as

$$\tau_y := \min\left(\frac{1}{2}, \frac{\sigma\varepsilon}{\beta_2}\varphi(1/2)\right).$$

We first split the domain Ω into four subdomains as in Figure 1:

$$\Omega_r := [0, 1-\tau]\times[0, 1-\tau_y], \qquad \Omega_x := [1-\tau, 1]\times[0, 1-\tau_y],$$
$$\Omega_y := [0, 1-\tau]\times[1-\tau_y, 1], \qquad \Omega_{xy} := [1-\tau, 1]\times[1-\tau_y, 1].$$

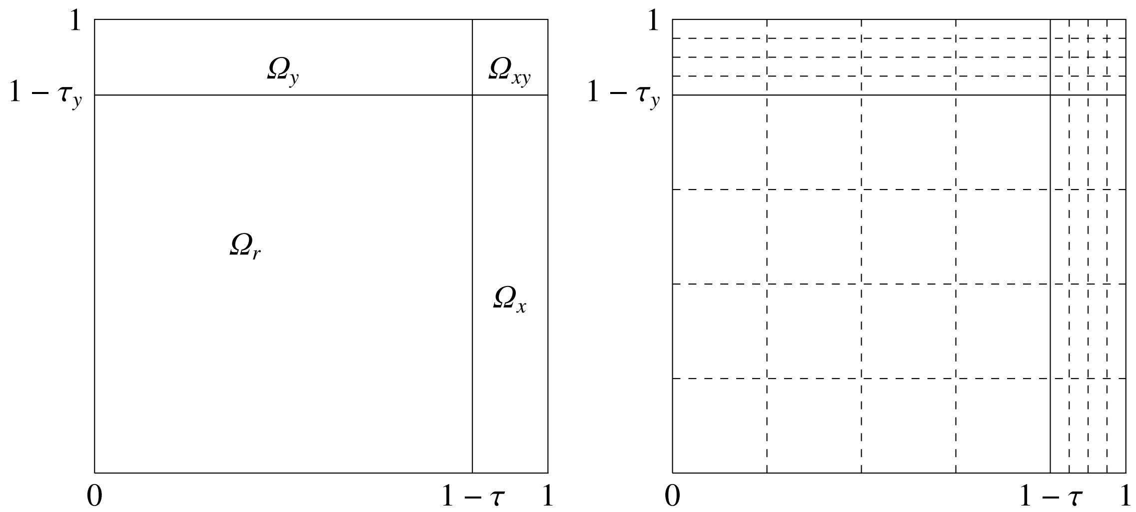


Figure 1. Tensor product Shishkin mesh for $N = 8$.

Clearly, the mesh is uniform in Ω_r with a mesh size of $O(N^{-1})$, highly anisotropic in Ω_x and Ω_y, while it is very fine in Ω_{xy}.

Let $h_i^x := x_i - x_{i-1},\ i = 1,\dots,N \quad h_j^y := y_j - y_{j-1},\ j = 1,\dots,N$ be the mesh sizes of the subintervals. For the sake of simplicity, we assume that $\beta_1 = \beta_2 = \beta$. Then, one has $h_i^x = h_j^y$, and we simply write $h_i,\ i = 1,\dots,N$ for simplicity. The following technical lemmas show the smallness of the boundary layer-functions and the basic properties of the mesh sizes of the layer-adapted meshes.

Lemma 2.2. *[28] Denote by* $\Theta_i = \min\{h_i/\varepsilon, 1\}\, e^{-\alpha(1-x_i)/\sigma\varepsilon}$ *for* $i = N/2+1,\dots,N$*. There exists a constant* $C > 0$ *independent of* ε *and* N *such that*

$$\max_{N/2+1\le i\le N}\Theta_i \le CN^{-1}\max|\psi'|$$

$$\sum_{i=N/2+1}^{N} \Theta_i \leq C$$

Lemma 2.3. *[28] For the layer-adapted meshes we considered here, we have*

$$h_1 = h_2 = \ldots = h_{N/2} \text{ and } \min_{i=1,\ldots,N} h_i \geq C\varepsilon N^{-1} \max |\psi'| .$$

Moreover, for $i = N/2 + 1, \cdots, N$,

$$\begin{aligned} h_i &= 2\tau/N, && \text{for S-type} \\ 1 \geq \tfrac{h_{i+1}}{h_i} &\geq C, && \text{for BS- mesh} \end{aligned}$$

and, for the B-type mesh,

$$\begin{aligned} i &= N/2 + 2, \cdots, N, \quad 1 \geq \tfrac{h_{i+1}}{h_i} \geq C, \\ i &= 1, 2, \cdots, N/2, \quad h_{N/2+i} \geq \tfrac{\sigma\varepsilon}{\beta} \tfrac{1}{i+1}, \end{aligned}$$

where $C > 0$ *is a constant independent of* ε *and* N.

2.3. The numerical method

A weak formulation of the problems (1.1) and (1.2) is to look for $u \in H_0^1(\Omega)$ such that

$$(u_t, v) + A(u, v) = (f, v) \quad \forall v \in H_0^1(\Omega), \tag{2.8}$$

where the bilinear from $A(\cdot, \cdot)$ is defined by

$$A(u, v) := \varepsilon(\nabla u, \nabla v) + (\mathbf{b} \cdot \nabla u, v) + (cu, v).$$

Based on the weak formulation (2.8), we define the WG-FEM on the layer-adapted mesh. Let p be a positive integer. We define a local WG-FE space $V(p, K)$ on each $K \in \mathcal{T}_N$ given by

$$V(p, K) := \left\{ v_N = \{v_0, v_b\} : v_0|_K \in \mathbb{Q}_p(K), v_b|_e \in \mathbb{P}_p(e), e \subset \partial K \right\},$$

where $\mathbb{Q}_p(K)$ is the polynomials of degree p on K in both variables, and $\mathbb{P}_p(e)$ denotes the polynomials of degree p on the edge e.

Defining the WG finite element space V_N globally on $\mathcal{T}_N$ as

$$V_N = \left\{ v_N = \{v_0, v_b\} : v_N|_K \in V(p, K), v_b|_{e \cap \partial K_1} = v_b|_{e \cap \partial K_2}, \partial K_1 \cap \partial K_2 = \{e\} \right\}, \tag{2.9}$$

and its subspace

$$V_N^0 = \{v : v \in V_N, v_b = 0 \text{ on } \partial\Omega\} .$$

The weak gradient operator $\nabla_w u_N \in [\mathbb{Q}_{p-1}(K)]^2$ can be defined on K as

$$(\nabla_w u_N, \psi)_K = -(u_0, \nabla \cdot \psi)_K + \langle u_b, \psi \cdot \mathbf{n} \rangle_{\partial K} \qquad \forall \psi \in [\mathbb{Q}_{p-1}(K)]^2, \tag{2.10}$$

where $\mathbf{n}$ represents the outward unit normal ∂K, $(w, v)_K$ denotes the inner product on K for functions w and v, and $\langle w, v \rangle_{\partial K}$ is the L^2-inner product on ∂K.

The weak convection operator $\mathbf{b} \cdot \nabla_w u_N \in \mathbb{Q}_p(K)$ can be defined on K as

$$(\mathbf{b} \cdot \nabla_w u_N, \xi)_K = -(u_0, \nabla \cdot (\mathbf{b}\xi))_K + \langle u_b, \mathbf{b} \cdot \mathbf{n}\xi \rangle_{\partial K} \qquad \forall \xi \in \mathbb{Q}_p(K). \tag{2.11}$$

For simplicity, we adapt

$$(\phi, \psi) = \sum_{K \in \mathcal{T}_N} (\phi, \psi)_K, \quad \|\phi\|^2 = (\phi, \phi), \quad \langle \phi, \psi \rangle = \sum_{K \in \mathcal{T}_N} \langle \phi, \psi \rangle_{\partial K}.$$

For $u_N = \{u_0, u_b\} \in V_N$ and $v_N = \{v_0, v_b\} \in V_N$, the bilinear form $\mathcal{A}_w(\cdot, \cdot)$ is given by

$$\begin{aligned} \mathcal{A}_w(u_N, v_N) = \varepsilon(\nabla_w u_N, \nabla_w v_N) + (\mathbf{b} \cdot \nabla_w u_N, v_0) + (cu_0, v_o) \\ + \mathcal{S}_d(u_N, v_N) + \mathcal{S}_c(u_N, v_N), \end{aligned} \tag{2.12}$$

where $s_d(\cdot, \cdot)$ and $s_c(\cdot, \cdot)$ are bilinear forms defined by

$$\begin{aligned} \mathcal{S}_d(u_N, v_N) &= \sum_{K \in \mathcal{T}_N} \rho_K \langle u_0 - u_b, v_0 - v_b \rangle_{\partial K}, \\ \mathcal{S}_c(u_N, v_N) &= \sum_{K \in \mathcal{T}_N} \langle \mathbf{b} \cdot \mathbf{n}(u_0 - u_b), v_0 - v_b \rangle_{\partial K^+} \end{aligned}$$

with $\partial K^+ = \{z \in \partial K : \mathbf{b}(z) \cdot \mathbf{n}(z) \geq 0\}$ and ρ_K is the penalty term given by

$$\rho_K := \begin{cases} 1, & \text{if } K \subset \Omega_r, \\ N(\max |\psi'|)^{-1}, & \text{if } K \subset \Omega \setminus \Omega_r. \end{cases} \tag{2.13}$$

Given a mesh rectangle K, its dimensions parallel to the x and y-axes are written as $h_{x,K}$ and $h_{y,K}$, respectively.

Lemma 2.4. *[29] For all $K \in \mathcal{T}_N$ with $h_K = \min\{h_{x,K}, h_{y,K}\}$, there exists a constant C depending only on p such that*

$$\|u_N\|^2_{L^2(\partial K)} \leq C h_K^{-1} \|u_N\|^2_{L^2(K)}, \quad \forall u_N \in \mathbb{P}_p(K). \tag{2.14}$$

We next formulate our semi-discrete WG scheme as follows (Algorithm 1).

Algorithm 1 The semi-discrete WG-FEM for problem (1.1)

Find $u_N = (u_0, u_b) \in V_N^0$ such that

$$\begin{aligned} &(u_0', v_0) + \mathcal{A}_w(u_N, v_N) = (f, v_0) \quad \forall v_N = (v_0, v_b) \in V_N^0, \\ &u_N(0) = u^0(0), \end{aligned} \tag{2.15}$$

where $u^0(0) \in V_N^0$ is an approximation of $u(0)$.

3. The semi-discrete scheme and stability analysis

This section is devoted to establishing the stability results of the WG-FEM defined by (2.15). Define the energy norm $\|\cdot\|_E$ on the weak function space V_N for $v_N = \{v_0, v_b\} \in V_N$,

$$\|v_N\|_E^2 := \varepsilon \sum_{K\in\mathcal{T}_N} \|\nabla_w v_N\|_K^2 + \sum_{K\in\mathcal{T}_N} \||\mathbf{b}\cdot\mathbf{n}|^{1/2}(v_0 - v_b)\|_{\partial K}^2 + \|v_0\|^2 + \mathcal{S}_d(v_N, v_N). \tag{3.1}$$

Define also an H^1 equivalent norm on V_N by

$$\|v_N\|_\varepsilon^2 := \varepsilon \sum_{K\in\mathcal{T}_N} \|\nabla v_0\|_K^2 + \sum_{K\in\mathcal{T}_N} \||\mathbf{b}\cdot\mathbf{n}|^{1/2}(v_0 - v_b)\|_{\partial K}^2 + \|v_0\|^2 + \mathcal{S}_d(v_N, v_N). \tag{3.2}$$

The equivalence of these two norms on V_N^0 is given in the next lemma.

Lemma 3.1. *For $v_N \in V_N^0$, one has*

$$C\|v_N\|_\varepsilon \le \|v_N\|_E \le C\|v_N\|_\varepsilon.$$

Proof. For $v_N = \{v_0, v_b\} \in V_N^0$, it follows from the weak gradient operator (2.10) and integration by parts that;

$$(\nabla_w v_N, w)_K = (\nabla v_0, w)_K - \langle v_0 - v_b, \mathbf{n}\cdot w\rangle_{\partial K}, \quad \forall\, w \in [\mathbb{Q}_{p-1}(K)]^2, \forall K \in \mathcal{T}_N. \tag{3.3}$$

Choosing $w = \nabla_w v_N$ in (3.3) and using the Cauchy-Schwartz inequality and the trace inequality (2.4), we arrive at

$$\begin{aligned}(\nabla_w v_N, \nabla_w v_N)_K &= (\nabla v_0, \nabla_w v_N)_K - \langle v_0 - v_b, \mathbf{n}\cdot\nabla_w v_N\rangle_{\partial K}\\ &\le \|\nabla v_0\|_{L^2(K)}\|\nabla_w v_N\|_{L^2(K)} + \|v_0 - v_b\|_{L^2(\partial K}\|\nabla_w v_N\|_{L^2(\partial K)}\\ &\le (\|\nabla v_0\|_{L^2(K)} + Ch_K^{-1/2}\|v_0 - v_b\|_{L^2(\partial K)})\|\nabla_w v_N\|_{L^2(K)}.\end{aligned}$$

Hence, we get

$$\|\nabla_w v_N\|_{L^2(K)} \le \|\nabla v_0\|_{L^2(K)} + Ch_K^{-1/2}\|v_0 - v_b\|_{L^2(\partial K)}.$$

Summing over all $K \in \mathcal{T}_N$ yields

$$\varepsilon\|\nabla_w v_N\|^2 \le 2\Big(\varepsilon\|\nabla v_0\|^2 + C\sum_{K\in\mathcal{T}_N} \varepsilon h_K^{-1}\|v_0 - v_b\|_{L^2(\partial K)}^2\Big).$$

From the penalty term (2.13), we get

$$\frac{\varepsilon h_K^{-1}}{\rho_K} \le C, \quad \forall K \in \mathcal{T}_N,$$

and

$$\sum_{K\in\mathcal{T}_N} \varepsilon h_K^{-1}\|v_0 - v_b\|_{L^2(\partial K)} = \sum_{K\in\mathcal{T}_N} \frac{\varepsilon h_K^{-1}}{\rho_K}\rho_K\|v_0 - v_b\|_{L^2(\partial K)} \le Cs_d(v_N, v_N).$$

As a result, for $v_N \in V_N^0$, we have

$$\|v_N\|_E \le C\|v_N\|_\varepsilon. \tag{3.4}$$

Taking $w = \nabla v_0$ in (3.3) and using the Cauchy-Schwartz inequality, we get

$$\begin{aligned}(\nabla v_0, \nabla v_0)_K &= (\nabla_w v_N, \nabla v_0)_K + \langle v_0 - v_b, \mathbf{n} \cdot \nabla v_0\rangle_{\partial K}\\ &\le \|\nabla_w v_N\|_{L^2(K)}\|\nabla v_0\|_{L^2(K)} + \|v_0 - v_b\|_{L^2(\partial K}\|\nabla v_0\|_{L^2(\partial K)}\\ &\le (\|\nabla_w v_N\|_{L^2(K)} + Ch_K^{-1/2}\|v_0 - v_b\|_{L^2(\partial K)})\|\nabla v_0\|_{L^2(K)},\end{aligned}$$

where we have again used the trace inequality (2.4).

Consequently,

$$\|\nabla v_0\|_{L^2(K)} \le \|\nabla_w v_N\|_{L^2(K)} + Ch_K^{-1/2}\|v_0 - v_b\|_{L^2(\partial K)}.$$

Summing over all $K \in \mathcal{T}_N$ yields

$$\varepsilon\|\nabla v_0\|^2 \le 2\Big(\varepsilon\|\nabla_w v_N\|^2 + C\sum_{K\in\mathcal{T}_N}\varepsilon h_K^{-1}\|v_0 - v_b\|^2_{L^2(\partial K)}\Big).$$

Therefore, we have

$$\varepsilon\|\nabla v_0\|^2 \le 2(\varepsilon\|\nabla_w v_N\|^2 + Cs_d(v_N, v_N)),$$

which implies

$$\|v_N\|_\varepsilon \le C\|v_N\|_E. \tag{3.5}$$

From (3.4) and (3.5), we obtain the desired conclusion, which completes the proof. □

We shall show the coercivity of the bilinear form $\mathcal{A}_w(\cdot,\cdot)$ in $\|\cdot\|_E$ norm on V_N^0.

Lemma 3.2. *For any $v_N \in V_N^0$, the following inequality holds:*

$$\mathcal{A}_w(v_N, v_N) \ge C\|v_N\|_E^2, \quad \forall v_N \in V_N^0. \tag{3.6}$$

Proof. For $v_N = \{v_0, v_b\}, w_N = \{w_0, w_b\} \in V_N^0$, using the weak convection derivative (2.11) and integration by parts gives

$$\begin{aligned}(\boldsymbol{b}\cdot\nabla_w v_N, w_0) &= -(v_0, \nabla\cdot(\boldsymbol{b}w_0)) + \langle v_b, \boldsymbol{b}\cdot\boldsymbol{n}w_0\rangle\\ &= (\boldsymbol{b}\cdot\nabla v_0, w_0) - \langle\boldsymbol{b}\cdot\boldsymbol{n}(v_0 - v_b), w_0\rangle\end{aligned} \tag{3.7}$$

and

$$\begin{aligned}(\boldsymbol{b}\cdot\nabla_w w_N, v_0) &= -(w_0, \nabla\cdot(\boldsymbol{b}v_0)) + \langle w_b, \boldsymbol{b}\cdot\boldsymbol{n}v_0\rangle\\ &= -(w_0, \nabla\cdot(\boldsymbol{b}v_0)) + \langle w_b, \boldsymbol{b}\cdot\boldsymbol{n}(v_0 - v_b)\rangle\end{aligned} \tag{3.8}$$

where we use the facts that $v_N, w_N \in V_N^0$, and $\langle\boldsymbol{b}\cdot\boldsymbol{n}v_b, w_b\rangle = 0$ in the last inequality.

Combining (3.7) and (3.8), and taking $v_N = w_N$, we obtain

$$(\boldsymbol{b}\cdot\nabla_w v_N, v_0) = -\frac{1}{2}(\nabla\cdot\boldsymbol{b}v_0, v_0) - \frac{1}{2}\langle\boldsymbol{b}\cdot\boldsymbol{n}(v_0 - v_b), v_0 - v_b\rangle. \tag{3.9}$$

From (3.9), we have

$$\begin{aligned}\mathcal{A}_w(v_N, v_N) =&\varepsilon\left(\nabla_w v_N, \nabla_w v_N\right) + \left(\left(-\frac{1}{2}\nabla\cdot\boldsymbol{b}\right)v_0, v_0\right) + \mathcal{S}_d(v_N, v_N) + \mathcal{S}_c(v_N, v_N)\\ &-\frac{1}{2}\langle \boldsymbol{b}\cdot\boldsymbol{n}(v_0 - v_b), v_0 - v_b\rangle\\ \geq&\varepsilon\|\nabla_w v_N\|^2 + c_0\|v_0\|^2 + \frac{1}{2}\sum_{T\in\mathcal{T}_N}\left\||\mathbf{b}\cdot\mathbf{n}|^2(v_0 - v_b)\right\|^2_{\partial K} + \mathcal{S}_d(v_N, v_N)\\ \geq& C\|v_N\|^2_E,\end{aligned}$$

which yields (3.6) with $C = \min\left\{c_0, \frac{1}{2}\right\}$. The proof is completed. □

Therefore, the existence of a unique solution to (2.15) follows from the coercivity property of the bilinear form (3.6). As a result of the two lemmas above, the bilinear form is also coercive in the $\|\cdot\|_\varepsilon$-norm in the sense that for any $v_N \in V_N^0$, there holds

$$\mathcal{A}_w(v_N, v_N) \geq C\|v_N\|^2_\varepsilon. \tag{3.10}$$

Lemma 3.3. *If $f \in L^2(\Omega)$ for each $t \in (0, T)$, then there is a constant $C > 0$ independent of t and mesh size h such that the solution $u_N(t) = \{u_0(t), u_b(t)\}$ defined in* (2.15) *satisfies*

$$\|u_0(t)\|^2 \leq C\|u^0\|^2 + \int_0^t \|f(s)\|^2 ds, \ \forall\ t \in (0, T]. \tag{3.11}$$

Proof. Choosing $v = u_N(t)$ in (2.15) gives that

$$\frac{1}{2}\frac{d}{dt}\|u_0(t)\|^2 + \mathcal{A}_w(u_N, u_N) = (f, u_0).$$

Using the Cauchy-Schwarz inequality and the coercivity (3.6) of the bilinear form $\mathcal{A}_w(\cdot,\cdot)$,

$$\begin{aligned}\frac{d}{dt}\|u_0(t)\|^2 &\leq 2(f, u_0)\\ &\leq \|f\|^2 + \|u_0\|^2.\end{aligned}$$

Integrating the above inequality with respect to the time variable t, we arrive at

$$\|u_0(t)\|^2 \leq \|u^0\|^2 + \int_0^t \|f(s)\|^2 ds + \int_0^t \|u_0(s)\|^2 ds. \tag{3.12}$$

Using the Gronwall's inequality gives the desired conclusion. We are done. □

4. Interpolation estimates

First, we define "vertices-edges' interpolation $\mathcal{P}_K v$ of a function v on an element K as follows. Let $\hat{K} := (-1, 1) \times (-1, 1)$ be the reference element with the vertices $\hat{a}_i$ and the edges $\hat{e}_i$ for $i = 1, \ldots, 4$. For $\hat{v}(\cdot,\cdot) \in C(\hat{K})$, the approximation $\hat{\mathcal{P}} : C(\hat{K}) \to \mathbb{Q}_p(\hat{K})$ is given by

$$(\hat{\mathcal{P}}\hat{v})(\hat{a}_i) = \hat{v}(\hat{a}_i) \quad \text{ for } i = 1, \ldots, 4, \tag{4.1}$$

$$\int_{\hat{e}_i} (\hat{\mathcal{P}}\hat{v}) q ds = \int_{\hat{e}_i} \hat{v} q ds \quad \forall q \in \mathbb{P}_{p-2}(\hat{e}_i) \text{ for } i = 1, \ldots, 4, \tag{4.2}$$

$$\iint_{\hat{K}} (\hat{\mathcal{P}}\hat{v}) q d\xi d\eta = \iint_{\hat{K}} \hat{v} q d\xi d\eta \quad \forall\, q \in \mathbb{Q}_{p-2}(\hat{K}). \tag{4.3}$$

The approximation operator $\hat{\mathcal{P}}$ is well-defined [30]. Thus, we can define a continuous global interpolation operator $\mathcal{P}_N : C(\overline{\Omega}) \to V_N$ by writing

$$(\mathcal{P}_N u)|_K := (\hat{\mathcal{P}}(u \circ G_K)) \circ G_K^{-1} \qquad \forall\, K \in \mathcal{T}_N, u \in C(\overline{\Omega}), \tag{4.4}$$

where the bijective mapping $G_K : \hat{K} \to K$ is given by $G_K(\xi, \eta) = (x, y) = (x_K + h_K^x \xi/2, y_K + h_K^y \eta/2)$.

This interpolation operator has the following stability estimate [30]

$$\|\mathcal{P}_N \phi\|_{L^\infty(K)} \le C\|\phi\|_{L^\infty(K)}, \quad \forall \phi \in C(K). \tag{4.5}$$

Since our approximation operator $(\mathcal{P}_N v)|_K$ is continuous on K, we have $\{(\mathcal{P}_N v)|_K, (\mathcal{P}_N v)|_e\} \in V_N$ for $e \in \partial K$. By the trace theorem, we will denote this by $\mathcal{P}_N v$.

Lemma 4.1. *[31] For any $\phi \in H^1(K)$,*

$$(\nabla_w(\mathcal{P}_N \phi), \psi) = (\nabla(\mathcal{P}_N \phi), \psi) \quad \forall \psi \in \mathbb{Q}_{p-1}(K), K \in \mathcal{T}_N.$$

We recall some technical results from [30].

Lemma 4.2. *For any $K \in \mathcal{T}_N$ and $q \in [1, \infty]$, there exists a constant C such that the vertices-edges-element approximant $\mathcal{P}_N \phi$ satisfies*

$$\|v - \mathcal{P}_N \phi\|_{L^q(K)} \le C \sum_{i+j=p+1} h_{x,K}^i h_{y,K}^j \left\| \frac{\partial^{p+1}\phi}{\partial x^i \partial y^j} \right\|_{L^q(K)}$$

for all $\phi \in W^{p+1,q}(K)$.

Lemma 4.3. *Let $s \in [1, p]$. The following estimates hold for any $K \in \mathcal{T}_N$:*

$$\left\|(\psi - \mathcal{P}_N \psi)_x\right\|_{L^2(K)} \le C \sum_{i=0}^{s} h_{x,K}^i h_{y,K}^{s-i} \left\| \frac{\partial^{s+1}\psi}{\partial x^{i+1} \partial y^{s-i}} \right\|_{L^2(K)},$$

$$\left\|(\psi - \mathcal{P}_N \psi)_y\right\|_{L^2(K)} \le C \sum_{i=0}^{s} h_{x,K}^i h_{y,K}^{s-i} \left\| \frac{\partial^{s+1}\psi}{\partial x^{i} \partial y^{s-i+1}} \right\|_{L^2(K)}$$

for all $\psi \in H^{r+1}(K)$.

A careful inspection of the proof of Lemma 4.3 in [30] reveals that the following results also hold true.

Lemma 4.4. *For $K \in \mathcal{T}_N$ and $s \in [1, p+1-\ell]$ with $\ell = 1, 2, \ldots, p+1$, there exists a constant C such that the vertices-edges-element approximant $\mathcal{P}_N \phi$ satisfies*

$$\left\| \frac{\partial^\ell(\phi - \mathcal{P}_N \phi)}{\partial x^\ell} \right\|_{L^2(K)} \le C \sum_{i=0}^{s} h_{x,K}^i h_{y,K}^{s-i} \left\| \frac{\partial^{r+\ell}\phi}{\partial x^{i+\ell} \partial y^{s-i}} \right\|_{L^2(K)},$$

$$\left\| \frac{\partial^\ell(\phi - \mathcal{P}_N \phi)}{\partial y^\ell} \right\|_{L^2(K)} \le C \sum_{i=0}^{s} h_{x,K}^i h_{y,K}^{s-i} \left\| \frac{\partial^{r+\ell}\phi}{\partial x^{i} \partial y^{s-i+\ell}} \right\|_{L^2(K)}$$

for all $\phi \in H^{r+1}(K)$.

Lemma 4.5. *Let the assumption of Lemma 2.1 hold such that $u = u_R + u_L,\ u_L = u_{L_0} + u_{L_1} + u_{L_2}$. Then there is a constant $C > 0$ such that the following interpolation error estimates are satisfied:*

$$||u_R - \mathcal{P}_N u_R||_{L^2(\Omega)} \leq CN^{-(p+1)}, \tag{4.6}$$
$$||u_L||_{L^2(\Omega_r)} \leq C\varepsilon^{1/2}N^{-(p+1)}, \tag{4.7}$$
$$||\nabla u_L||_{L^2(\Omega_r)} \leq C\varepsilon^{-1/2}N^{-(p+1)}, \tag{4.8}$$
$$||\Delta u_L||_{L^2(\Omega_r)} \leq C\varepsilon^{-3/2}N^{-(p+1)}, \tag{4.9}$$
$$N^{-1}||\nabla\mathcal{P}_N u_L||_{L^2(\Omega_r)} + ||\mathcal{P}_N u_L||_{L^2(\Omega_r)} \leq CN^{-(p+3/2)}, \tag{4.10}$$
$$||u_L - \mathcal{P}_N u_L||_{L^2(\Omega)} \leq C(N^{-1}\max|\psi'|)^{p+1}. \tag{4.11}$$

Proof. The proof of (4.6) follows from Lemma 4.2 and the solution decomposition (2.2) of Lemma 2.1. Using the decay bound (2.3) of u_{L_0} and the fact that $\varphi(1/2) \geq \ln N$, we have

$$\begin{aligned}||u_{L_0}||^2_{L^2(\Omega_r)} &= \int_0^{1-\tau}\int_0^{1-\tau}|u_{L_0}|^2 dxdy\\ &\leq C\int_0^{1-\tau}\int_0^{1-\tau} e^{-2\beta_1(1-x)/\varepsilon}dxdy \leq C\varepsilon N^{-2(p+1)},\end{aligned}$$

which shows (4.7) for u_{L_0}. Similar arguments can be applied to the layer functions u_{L_1} and u_{L_2}. Thus, we complete the proof of (4.7). For (4.8) and (4.9), we will prove for u_{L_0} since the other two parts follow similarly. One can use the decay bound (2.3) of u_{L_0} to obtain

$$\begin{aligned}\left\|\nabla u_{L_0}\right\|^2_{L^2(\Omega_r)} &\leq C\varepsilon^{-2}\int_0^{1-\tau}\int_0^{1-\tau} e^{-2\beta_1(1-x)/\varepsilon}dxdy\\ &\leq C\varepsilon^{-1}e^{-2\beta_1\tau/\varepsilon}\\ &\leq C\varepsilon^{-1}N^{-2(p+1)}.\end{aligned}$$

We now shall prove (4.9). Appealing (2.3), one gets

$$\begin{aligned}\left\|\Delta u_{L_0}\right\|^2_{L^2(\Omega_r)} &\leq C\varepsilon^{-4}\int_0^{1-\tau}\int_0^{1-\tau} e^{-2\beta_1(1-x)/\varepsilon}dxdy\\ &\leq C\varepsilon^{-3}e^{-2\beta_1\tau/\varepsilon}\\ &\leq C\varepsilon^{-3}N^{-2(p+1)}.\end{aligned}$$

The proof of (4.10) is a little longer. Using an inverse estimate yields

$$N^{-1}||\nabla\mathcal{P}_N u_L||_{L^2(\Omega_r)} + ||\mathcal{P}_N u_L||_{L^2(\Omega_r)} \leq C||\mathcal{P}_N u_L||_{L^2(\Omega_r)}.$$

Hence, we shall estimate $||\mathcal{P}_N u_L||_{L^2(\Omega_r)}$. With the help of the stability estimate (4.5) and the decay bound (2.3) of u_{L_0}, we have

$$\left\|\mathcal{P}_N u_{L_0}\right\|^2_{L^2(\Omega_r)} \leq C\int_0^{1-\tau}\sum_{i=1}^{N/2}\int_{x_{i-1}}^{x_i} e^{-2\beta_1(1-x_i)/\varepsilon}dxdy.$$

If $i < N/2$, then the sum can be small as a function of ε but not small if $i = N/2$. For $i = 1, \ldots, N/2 - 1$ and $x \in [x_{i-1}, x_i]$, we have

$$e^{-2\beta_1(1-x_i)/\varepsilon} = e^{2\beta_1\left(x_{N/2}-x_{N/2-1}\right)/\varepsilon} e^{-2\beta_1(1-x_{i-1})/\varepsilon} \le e^{2\beta_1\left(x_{N/2}-x_{N/2-1}\right)/\varepsilon} e^{-2\beta_1(1-x)/\varepsilon}$$

and when $i = N/2$, again using the fact $\varphi(1/2) \ge \ln N$,

$$e^{-2\beta_1\left(1-x_{N/2}\right)/\varepsilon} = e^{-2\beta_1\tau/\varepsilon} \le N^{-2(p+1)}.$$

Thus,

$$\begin{aligned} \left\|\mathcal{P}_N u_{L_0}\right\|^2_{L^2(\Omega_r)} &\le Ce^{2\beta_1\left(x_{N/2}-x_{N/2-1}\right)/\varepsilon} \int_0^{x_{N/2-1}} e^{-2\beta_1(1-x)/\varepsilon} dx + CN^{-2p-3} \\ &\le C\varepsilon e^{-2\beta_1\left(1-x_{N/2}\right)/\varepsilon} + CN^{-2p-3} \\ &\le C\left(\varepsilon N^{-2(p+1)} + N^{-2p-3}\right) \end{aligned}$$

which proves (4.10). To prove (4.11), we use (4.7) and (4.10) to obtain

$$\|u_L - \mathcal{P}_N u_L\|_{L^2(\Omega_r)} \le C\left(\varepsilon^{1/2} N^{-(p+1)} + N^{-p-3/2}\right) \le C(N^{-1} \max|\psi'|)^{p+1}.$$

On the set $\Omega_r \cup \Omega_y$, from the triangle inequality, one obtains

$$\left\|u_{L_0} - \mathcal{P}_N u_{L_0}\right\|_{L^2(\Omega_r\cup\Omega_y)} \le C(\left\|u_{L_0}\right\|_{L^2(\Omega_r\cup\Omega_y)} + \left\|\mathcal{P}_N u_{L_0}\right\|_{L^2(\Omega_r\cup\Omega_y)}) =: I_1 + I_2$$

For I_1, we have

$$\left\|u_{L_0}\right\|_{L^2(\Omega_r\cup\Omega_y)} \le \Big(\int_0^{1-\tau} e^{-2\beta_1(1-x)/\varepsilon} dx\Big)^{1/2} \le C\varepsilon^{1/2} N^{-(p+1)}.$$

For I_2, using (4.5) and the decay property (2.3) of u_{L_0},

$$\begin{aligned} \left\|\mathcal{P}_N u_{L_0}\right\|_{L^2(\Omega_r\cup\Omega_y)} &= \left\{\sum_{i=1}^{N/2} \int_{x=x_{i-1}}^{x_i} \int_{y=0}^{1} [\mathcal{P}_N u_{L_0}(x,y)]^2\, dydx\right\}^{1/2} \\ &\le C\left\{\sum_{i=1}^{N/2} \int_{x=x_{i-1}}^{x_i} \int_{y=0}^{1} e^{-2\beta_1(1-x_i)/\varepsilon} dydx\right\}^{1/2} \\ &\le C\left\{\sum_{i=1}^{N/2-1} \int_{x=x_i}^{x_{i+1}} \int_{y=0}^{1} e^{-2\beta_1(1-x)/\varepsilon} dydx + \int_{x=x_{N/2-1}}^{x_{N/2}} \int_{y=0}^{1} e^{-2\beta_1\tau/\varepsilon} dydx\right\}^{1/2} \\ &\le C\left\{\left(\varepsilon + N^{-1}\right) N^{-2(p+1)}\right\}^{1/2}, \end{aligned} \tag{4.12}$$

where we have used that $\varphi(1/2) \ge \ln N$. Applying Lemma 4.2 and the decay property (2.3), we have for any $K \subset \Omega_x \cup \Omega_{xy}$

$$\begin{aligned} \left\|u_{L_0} - \mathcal{P}_N u_{L_0}\right\|_{L^2(K)} &\le C(h_{x,K}^{p+1} \left\|\frac{\partial^{p+1} u_{L_0}}{\partial x^{p+1}}\right\|_{L^2(K)} + h_{y,K}^{p+1} \left\|\frac{\partial^{p+1} u_{L_0}}{\partial y^{p+1}}\right\|_{L^2(K)}) \\ &\le C\Big((\frac{h_{x,K}}{\varepsilon})^{p+1} + h_{y,K}^{p+1}\Big)\Big(\int_{1-\tau}^{1} \int_0^1 e^{-2\beta_1(1-x)/\varepsilon} dydx\Big)^{1/2} \end{aligned}$$

$$\leq C\Theta_i^{p+1} \leq C(N^{-1}\max|\psi'|)^{p+1},$$

where we have used $\dfrac{h_{x,K}}{\varepsilon} \geq CN^{-1} \geq Ch_{y,K}$ and Lemma 2.2. Similarly, one can show that (4.11) holds for u_{L_1} as well.

For $K \subset \Omega_r \cup \Omega_x \cup \Omega_y$, one can prove as above $\sum_{K\subset\Omega_r\cup\Omega_x\cup\Omega_y} \left\|u_{L_2} - \mathcal{P}_N u_{L_2}\right\|_{L^2(K)} \leq CN^{-(p+1)}$. For $K \subset \Omega_{xy}$, we obtain

$$\begin{aligned}
\left\|u_{L_2} - \mathcal{P}_N u_{L_2}\right\|_{L^2(K)} &\leq C(h_{x,K}^{p+1}\left\|\frac{\partial^{p+1}u_{L_2}}{\partial x^{p+1}}\right\|_{L^2(K)} + h_{y,K}^{p+1}\left\|\frac{\partial^{p+1}u_{L_2}}{\partial y^{p+1}}\right\|_{L^2(K)}) \\
&\leq C\Big((\frac{h_{x,K}}{\varepsilon})^{p+1} + (\frac{h_{y,K}}{\varepsilon})^{p+1}\Big)\Big(\int_{1-\tau}^{1}\int_{1-\tau_y}^{1} e^{-(2\beta_1(1-x)+2\beta_2(1-y))/\varepsilon}dydx\Big)^{1/2} \\
&\leq C(\Theta_i^{p+1} + \Theta_j^{p+1}) \leq C(N^{-1}\max|\psi'|)^{p+1}.
\end{aligned}$$

Thus, we complete the proof of (4.11). The proof of the lemma is now completed. □

Lemma 4.6. *Let $u \in H^{p+1}(\Omega)$. Assume that the conditions of Lemma 4.5 hold. Then, we have*

$$\begin{aligned}
&\|\nabla(u_R - \mathcal{P}_N u_R)\|_{L^2(\Omega)} \leq CN^{-p}, \\
&\|\Delta(u_R - \mathcal{P}_N u_R)\|_{L^2(\Omega)} \leq CN^{1-p}, \\
&\left\|\frac{\partial^l(u_L - \mathcal{P}_N u_L)}{\partial x^l}\right\|_{L^2(\Omega r)} \leq C\varepsilon^{1/2-l}N^{-(p+1)}, \\
&\left\|\frac{\partial^l(u_L - \mathcal{P}_N u_L)}{\partial x^l}\right\|_{L^2(\Omega\setminus\Omega_r)} \leq C\varepsilon^{1/2-l}(N^{-1}\max|\psi'|)^{p+1-l}, \\
&\left\|\frac{\partial^l(u_L - \mathcal{P}_N u_L)}{\partial y^l}\right\|_{L^2(\Omega r)} \leq C\varepsilon^{1/2-l}N^{-(p+1)}, \\
&\left\|\frac{\partial^l(u_L - \mathcal{P}_N u_L)}{\partial y^l}\right\|_{L^2(\Omega\setminus\Omega_r)} \leq C\varepsilon^{1/2-l}(N^{-1}\max|\psi'|)^{p+1-l}
\end{aligned}$$

for $l = 1, 2$, where u_L denotes u_{L_0}, u_{L_1}, or u_{L_2}.

Proof. The first and second estimates follow from Lemma 4.3, Lemma 4.4, and the fact that $\max\{h_{x,K}, h_{y,k}\} \leq CN^{-1}$. From the triangle inequality and (4.8) and (4.10) of Lemma 4.5, we have

$$\|\nabla(u_L - \mathcal{P}_N u_L)\|_{L^2(\Omega_r)} \leq \|\nabla u_L\|_{L^2(\Omega_r)} + \|\nabla\mathcal{P}_N u_L\|_{L^2(\Omega_r)} \leq C\varepsilon^{-1/2}N^{-(p+1)},$$

where we have used that $\varepsilon \leq CN^{-1}$. This completes the proof of the third and fifth inequalities for $l = 1$. An inverse inequality and (4.9) and (4.10) of Lemma 4.5 lead to

$$\begin{aligned}
\|\Delta(u_L - \mathcal{P}_N u_L)\|_{L^2(\Omega_r)} &\leq \|\Delta u_L\|_{L^2(\Omega_r)} + CN\|\nabla\mathcal{P}_N u_L\|_{L^2(\Omega_r)} \\
&\leq C\varepsilon^{-3/2}[1 + (\varepsilon N)^{3/2} + (\varepsilon N)^2]N^{-(p+1)} \\
&\leq C\varepsilon^{-3/2}N^{-(p+1)},
\end{aligned}$$

where again we have used $\varepsilon N \leq C$. This proves the third and fifth inequalities for $l = 2$. Using Lemma 4.4 with $r = p + 1 - \ell$ for $\ell = 1, 2$ and the decay bound (2.3) of u_{L_0}, one can show that for any $K \subset \Omega_x \cup \Omega_{xy}$,

$$\begin{aligned}
\left\|\frac{\partial^\ell (u_{L_0} - \mathcal{P}_N u_{L_0})}{\partial x^\ell}\right\|_{L^2(K)} &\leq C(h_{x,K}^{p+1-\ell}\left\|\frac{\partial^{p+1} u_{L_0}}{\partial x^{p+1}}\right\|_{L^2(K)} + h_{y,K}^{p+1-\ell}\left\|\frac{\partial^{p+1} u_{L_0}}{\partial x^\ell \partial y^{p+1-\ell}}\right\|_{L^2(K)}) \\
&\leq C\varepsilon^{-\ell}\Big((\frac{h_{x,K}}{\varepsilon})^{p+1-\ell} + h_{y,K}^{p+1-\ell}\Big)\Big(\int_{1-\tau}^1 \int_0^1 e^{-2\beta_1(1-x)/\varepsilon} dydx\Big)^{1/2} \\
&\leq C\varepsilon^{1/2-\ell}\Theta_i^{p+1-\ell} \leq C\varepsilon^{1/2-\ell}(N^{-1}\max|\psi'|)^{p+1-\ell}.
\end{aligned}$$

Similarly, one can prove that the result holds for u_{L_1}, too.

For $K \subset \Omega_r \cup \Omega_x \cup \Omega_y$, we get $\sum_{K\subset\Omega_r\cup\Omega_x\cup\Omega_y}\left\|\frac{\partial^\ell (u_{L_2} - \mathcal{P}_N u_{L_2})}{\partial x^\ell}\right\|_{L^2(K)} \leq C\varepsilon^{1/2-\ell}N^{-(p+1)}$. For $K \subset \Omega_{xy}$, we obtain

$$\begin{aligned}
\left\|\frac{\partial^\ell (u_{L_2} - \mathcal{P}_N u_{L_2})}{\partial x^\ell}\right\|_{L^2(K)} &\leq C(h_{x,K}^{p+1-\ell}\left\|\frac{\partial^{p+1} u_{L_2}}{\partial x^{p+1}}\right\|_{L^2(K)} + h_{y,K}^{p+1-\ell}\left\|\frac{\partial^{p+1} u_{L_2}}{\partial x^\ell y^{p+1-\ell}}\right\|_{L^2(K)}) \\
&\leq C\varepsilon^{-\ell}\Big((\frac{h_{x,K}}{\varepsilon})^{p+1-\ell} + (\frac{h_{y,K}}{\varepsilon})^{p+1-\ell}\Big)\times \\
&\quad \Big(\int_{1-\tau}^1 \int_{1-\tau_y}^1 e^{-2(\beta_1(1-x)+\beta_2(1-y))/\varepsilon} dydx\Big)^{1/2} \\
&\leq C\varepsilon^{1/2-\ell}(\Theta_i^{p+1-\ell} + \Theta_j^{p+1-\ell}) \leq C\varepsilon^{1/2-\ell}(N^{-1}\max|\psi'|)^{p+1-\ell},
\end{aligned}$$

which completes the proof of the fourth inequality. Similarly, one can prove that the last inequality holds true. The proof is now finished.

□

5. Error analysis of the semi-discrete WG-FEM

Unlike the classical numerical methods such as FEM and the SUPG, the proposed WG-FEM does not have Galerkin orthogonality property. This results in some inconsistency errors in the error bounds. We first give a useful error equation in the following lemma.

Lemma 5.1. *[31] Let u solve the problem (1.1). For $v_N = \{v_0, v_b\} \in V_N^0$,*

$$-\varepsilon(\Delta u, v_0) = \varepsilon(\nabla_w(\mathcal{P}_N u), \nabla_w v_N) - E_d(u, v_N), \tag{5.1}$$

$$(\boldsymbol{b} \cdot \nabla u, v_0) = (\boldsymbol{b} \cdot \nabla_w(\mathcal{P}_N u), v_0) - E_c(u, v_N) \tag{5.2}$$

$$(cu, v_0) = (c(\mathcal{P}_N u), v_0) - E_r(u, v_N) \tag{5.3}$$

where

$$E_d(u, v_N) = -\varepsilon(\nabla(u - \mathcal{P}_N u), \nabla v_0) + \varepsilon\langle\nabla(u - \mathcal{P}_N u)\cdot \boldsymbol{n}, v_0 - v_b\rangle, \tag{5.4}$$

$$E_c(u, v_N) = (u - \mathcal{P}_N u, \nabla\cdot(\boldsymbol{b}v_0)) - \langle u - \mathcal{P}_N u, \boldsymbol{b}\cdot\boldsymbol{n}v_0\rangle, \tag{5.5}$$

$$E_r(u, v_N) = (c(\mathcal{P}_N u) - cu, v_0). \tag{5.6}$$

The following error equation $e_N = \{e_0, e_b\} := \{\mathcal{P}_N u - u_0, \mathcal{P}_N u - u_b\}$ will be needed in the error analysis.

Lemma 5.2. *Let u and u_N be the solutions of (1.1) and (2.15), respectively. For $v_N = \{v_0, v_b\} \in V_N^0$, one has*

$$(e_0', v_0) + \mathcal{A}_w(e_N, v_N) = E(u, v_N), \tag{5.7}$$

where $E(u, v_N) := E_d(u, v_N) + E_c(u, v_N) + E_r(u, v_N)$, which are defined by (5.4), (5.5), and (5.6), respectively.

Proof. Multiplying (1.1) by a test function $v_N = \{v_0, v_b\} \in V_N^0$, we arrive at

$$(u_t, v_0) - \varepsilon (\Delta u, v_0) + (\mathbf{b} \cdot \nabla u, v_0) + (cu, v_0) = (f, v_0).$$

With the help of (5.1), (5.2), and (5.3), the above equation becomes

$$(u_t, v_0) + \varepsilon (\nabla_w \mathcal{P}_N u, \nabla_w v_N) + (\mathbf{b} \cdot \nabla_w \cdot (\mathcal{P}_N u), v_0) + (c\mathcal{P}_N u, v_0) = (f, v_0) + E(u, v_N).$$

Since $\mathcal{P}_N u$ is continuous in Ω, we get

$$\mathcal{S}_c(\mathcal{P}_N u, v_N) = \mathcal{S}_d(\mathcal{P}_N u, v_N) = 0.$$

Therefore, we have

$$(u_t, v_0) + \mathcal{A}_w(\mathcal{P}_N u, v_N) = (f, v_0) + E(u, v_N). \tag{5.8}$$

Subtracting (2.15) from (5.8) gives (5.7), which completes completed. □

Lemma 5.3. *Let $\mathcal{P}_N u$ be the vertex-edge-cell interpolation of the solution u of the problem (1.1). Then, there holds*

$$\left|(u - \mathcal{P}_N u)(x, y)\right| \le \begin{cases} CN^{-(p+1)}, & \text{if } (x, y) \in \Omega_r, \\ C(N^{-1} \max |\psi'|)^{p+1}, & \text{otherwise}. \end{cases}$$

Proof. The solution decomposition (2.1) implies that

$$u - \mathcal{P}_N u = (u_R - \mathcal{P}_N u_R) + (u_{L_0} - \mathcal{P}_N u_{L_0}) + (u_{L_1} - \mathcal{P}_N u_{L_1}) + (u_{L_2} - \mathcal{P}_N u_{L_2}).$$

Using Lemma 4.2 with $q = \infty$ and Lemma 2.1,

$$|(u_R - \mathcal{P}_N u_R)(x, y)| \le CN^{-(p+1)} \sum_{i+j=p+1} \|\frac{\partial^{p+1} u_R}{\partial x^i \partial y^j}\|_{L^\infty(\Omega)} \le CN^{-(p+1)} \quad \forall (x, y) \in \Omega.$$

Next, we examine the layer parts one by one. Let $(x, y) \in K \subset \Omega_r \cup \Omega_y$. From the L^∞ stability property (4.5) of the interpolation operator, one has

$$\left|(u_{L_0} - \mathcal{P}_N u_{L_0})(x, y)\right| \le C\|u_{L_0}\|_{L^\infty(K)} \le Ce^{-\beta_1(1-x_i)/\varepsilon} \le Ce^{-\beta_1\tau/\varepsilon} \le CN^{-(p+1)}.$$

Let $(x,y) \in K \subset \Omega_x \cup \Omega_{xy}$. The stability property (4.5) and Lemma 4.2 with $q = \infty$ yield

$$\begin{aligned}\|u_{L_0} - \mathcal{P}_N u_{L_0}\|_{L^\infty(K)} &\leq C \min\left\{\|u_{L_0}\|_{L^\infty(K)}, h_{x,K}^{p+1}\|\frac{\partial^{p+1}u_{L_0}}{\partial x^{p+1}}\|_{L^\infty(K)} + h_{x,K}^{p+1}\|\frac{\partial^{p+1}u_{L_0}}{\partial x^{p+1}}\|_{L^\infty(K)}\right\}\\ &\leq \min\left\{1, (\frac{h_{x,K}}{\varepsilon})^{p+1} + (h_{y,K})^{p+1}\right\}e^{-\beta_1(1-x_i)/\varepsilon}\\ &\leq C\Theta_i^{p+1} \leq C(N^{-1}\max|\psi'|)^{p+1}.\end{aligned}$$

Similarly, we can derive the estimates on the other layer components u_{L_1} and u_{L_2}. Combining the above estimates gives the desired conclusion.

Thus, we complete the proof. □

We recall the following trace inequality. For any $u \in H^1(K)$, one has

$$\|u\|_e^2 \leq C\left(h_K^{-1}\|u\|_{L^2(K)}^2 + \|u\|_{L^2(K)}\|\nabla u\|_{L^2(K)}\right). \tag{5.9}$$

Lemma 5.4. *Let $u \in H^{p+1}(\Omega)$ and ρ_K be given by (2.13). Assume that the conditions of Lemma 4.5 hold. Then, one has*

$$\left\{\sum_{K\in\mathcal{T}_N}\frac{\varepsilon^2}{\rho_K}\|\nabla(u-\mathcal{P}_N u)\|_{L^2(\partial K)}^2\right\}^{1/2} \leq C(N^{-1}\max|\psi'|)^p.$$

Proof. For the sake of simplicity, we use the following notations. Let $\zeta_R := u_R - \mathcal{P}_N u_R$ and $\zeta_L := u_L - \mathcal{P}_N u_L$ represent the interpolation errors of the regular and layer components of the solution. Hence, by the triangle inequality, we have

$$\sum_{K\in\mathcal{T}_N}\frac{\varepsilon^2}{\rho_K}\|\nabla\zeta\|_{L^2(\partial K)}^2 \leq \sum_{K\in\mathcal{T}_N}\frac{\varepsilon^2}{\rho_K}(\|\nabla\zeta_R\|_{L^2(\partial K)}^2 + \|\nabla\zeta_L\|_{L^2(\partial K)}^2). \tag{5.10}$$

With the help of the trace inequality (5.9), we have

$$\|\nabla\zeta_R\|_{L^2(\partial K)}^2 \leq C(h_K^{-1}\|\nabla\zeta_R\|_{L^2(K)}^2 + \|\nabla\zeta_R\|_{L^2(K)}\|\Delta\zeta_R\|_{L^2(K)}).$$

Now, appealing the definition (2.13) of stabilization parameter ρ_K and Lemma 4.6 gives

$$\begin{aligned}\sum_{K\in\mathcal{T}_N}\frac{\varepsilon^2}{\rho_K}\|\nabla\zeta_R\|_{L^2(\partial K)}^2 &\leq C\sum_{K\in\mathcal{T}_N}\frac{\varepsilon^2}{\rho_K}(h_K^{-1}\|\nabla\zeta_R\|_{L^2(K)}^2 + \|\nabla\zeta_R\|_{L^2(K)}\|\Delta\zeta_R\|_{L^2(K)})\\ &\leq C(\varepsilon^2 N\|\nabla\zeta_R\|_{L^2(\Omega_r)}^2 + \varepsilon\|\nabla\zeta_R\|_{L^2(\Omega\backslash\Omega_r)}^2\\ &\quad + \varepsilon^2\|\nabla\zeta_R\|_{L^2(\Omega_r)}\|\Delta\zeta_R\|_{L^2(\Omega_r)} + \varepsilon^2 N^{-1}\ln N\|\nabla\zeta_R\|_{L^2(\Omega\backslash\Omega_r)}\|\Delta\zeta_R\|_{L^2(\Omega\backslash\Omega_r)})\\ &\leq C\varepsilon N^{-2p},\end{aligned} \tag{5.11}$$

where we have used that $\varepsilon N < 1$.

Using once again the trace inequality (5.9), we have

$$\|\nabla\zeta_L\|_{L^2(\partial K)}^2 \leq C(h_K^{-1}\|\nabla\zeta_L\|_{L^2(K)}^2 + \|\nabla\zeta_L\|_{L^2(K)}\|\nabla\zeta_L\|_{L^2(K)}).$$

Now, appealing the definition (2.13) of stabilization parameter ρ_K and Lemma 4.6 again reveals that

$$\begin{aligned}\sum_{K\in\mathcal{T}_N}\frac{\varepsilon^2}{\rho_K}\|\nabla\zeta_L\|^2_{L^2(\partial K)} &\le C\sum_{K\in\mathcal{T}_N}\frac{\varepsilon^2}{\rho_K}(h_K^{-1}\|\nabla\zeta_L\|^2_{L^2(K)}+\|\nabla\zeta_L\|_{L^2(K)}\|\Delta\zeta_L\|_{L^2(K)})\\ &\le C(\varepsilon^2N\|\nabla\zeta_L\|^2_{L^2(\Omega_r)}+\varepsilon\|\nabla\zeta_L\|^2_{L^2(\Omega\setminus\Omega_r)}\\ &\quad+\ \varepsilon^2\|\nabla\zeta_L\|_{L^2(\Omega_r)}\|\Delta\zeta_L\|_{L^2(\Omega_r)}+\varepsilon^2N^{-1}\ln N\|\nabla\zeta_L\|_{L^2(\Omega\setminus\Omega_r)}\|\Delta\zeta_L\|_{L^2(\Omega\setminus\Omega_r)})\\ &\le C[(\varepsilon+N^{-1})N^{-(2p+1)}+(N^{-1}\max|\psi'|)^{2p}].\end{aligned}\tag{5.12}$$

Plugging (5.12) and (5.11) into (5.10) yields

$$\sum_{K\in\mathcal{T}_N}\frac{\varepsilon^2}{\rho_K}\|\nabla\zeta\|^2_{L^2(\partial K)}\le C[(\varepsilon+N^{-2})N^{-2p}+(N^{-1}\max|\psi'|)^{2p}].$$

Consequently, we have

$$\left\{\sum_{K\in\mathcal{T}_N}\frac{\varepsilon^2}{\rho_K}\|\nabla(u-\mathcal{P}_Nu)\|^2_{L^2(\partial K)}\right\}^{1/2}\le C(N^{-1}\max|\psi'|)^p,$$

which completes the proof.

□

Now, we shall prove the error bounds for the consistency errors.

Lemma 5.5. (A priori bounds) *Assume that $\mathcal{T}_N$ is the tensor product mesh as defined in Section 2. Then, for $u\in H^{k+1}(\Omega)$ and $v_N\in V_N^0$, we have*

$$|E_d(u,v_N)|\ \le\ C(N^{-1}\max|\psi'|)^p\|v_N\|_\varepsilon,\tag{5.13}$$

$$|E_c(u,v_N)+E_r(u,v_N)|\ \le\ C(N^{-1}\max|\psi'|)^p\|v_N\|_\varepsilon,\tag{5.14}$$

Proof. It follows from the Cauchy-Schwarz and Holder inequalities that

$$\begin{aligned}|E_d(u,v_N)| &\le \sum_{K\in\mathcal{T}_N}\varepsilon\|\nabla(u-\mathcal{P}_Nu)\|_{L^2(K)}\|\nabla v_0\|_{L^2(K)}\\ &\quad+\sum_{K\in\mathcal{T}_N}\varepsilon\|\nabla(u-\mathcal{P}_Nu)\|_{L^2(\partial K)}\|v_0-v_b\|_{L^2(\partial K)},\\ &:=\mathcal{S}_1+\mathcal{S}_2.\end{aligned}\tag{5.15}$$

Now, it then follows from Lemma 4.6 that

$$\begin{aligned}\mathcal{S}_1 &= \sum_{K\in\mathcal{T}_N}\varepsilon\|\nabla(u-\mathcal{P}_Nu)\|_{L^2(K)}\|\nabla v_0\|_{L^2(K)}\\ &= \sum_{K\in\mathcal{T}_N}\varepsilon^{1/2}\|\nabla(u-\mathcal{P}_Nu)\|_{L^2(K)}\varepsilon^{1/2}\|\nabla v_0\|_{L^2(K)}\\ &\le \Big(\sum_{K\in\mathcal{T}_N}\varepsilon^{1/2}\|\nabla(u_R-\mathcal{P}_Nu_R)\|_{L^2(K)}+\sum_{K\subset\Omega_r}\varepsilon^{1/2}\|\nabla(u_L-\mathcal{P}_Nu_L)\|_{L^2(K)}\end{aligned}$$

$$
\begin{aligned}
&+\sum_{K\subset\Omega\backslash\Omega_r}\varepsilon^{1/2}\|\nabla(u_L-\mathcal{P}_Nu_L)\|_{L^2(K)}\Big)\|\nabla v_N\|_\varepsilon\\
\leq\ & C(\varepsilon^{1/2}N^{-p}+N^{-(p+1)}+(N^{-1}\max|\psi'|)^p)\|v_N\|_\varepsilon\\
\leq\ & C(N^{-1}\max|\psi'|)^p\|v_N\|_\varepsilon. \qquad (5.16)
\end{aligned}
$$

Next, we consider the term $\mathcal{S}_2$. From the Cauchy-Schwarz inequality and Lemma 5.4, we have

$$
\begin{aligned}
\mathcal{S}_2&=\varepsilon\sum_{K\in\mathcal{T}_N}\|\nabla(u-\mathcal{P}_Nu)\|_{L^2(\partial K)}\|v_0-v_b\|_{L^2(\partial K)}\\
&\leq\left\{\sum_{K\in\mathcal{T}_N}\frac{\varepsilon^2}{\rho_K}\|\nabla(u-\mathcal{P}_Nu)\|^2_{L^2(\partial K)}\right\}^{1/2}\left\{\sum_{K\in\mathcal{T}_N}\rho_K\|v_0-v_b\|^2_{L^2(\partial K)}\right\}^{1/2}\\
&\leq C(N^{-1}\max|\psi'|)^p\|v_N\|_\varepsilon.
\end{aligned}
\tag{5.17}
$$

Combining (5.16) and (5.17), we get

$$
|E_d(u,v_N)|\leq C(N^{-1}\max|\psi'|)^p\|v_N\|_\varepsilon. \tag{5.18}
$$

From (5.5) and (5.6) and using $\langle u-I_Nu,\mathbf{b}\cdot\mathbf{n}v_b\rangle=0$, we arrive at

$$
\begin{aligned}
E_c(u,v_N)+E_r(u,v_N)&=(u-\mathcal{P}_Nu,\mathbf{b}\cdot\nabla v_0)+\langle u-\mathcal{P}_Nu,\mathbf{b}\cdot\mathbf{n}(v_0-v_b)\rangle\\
&\quad+(u-\mathcal{P}_Nu,(\nabla\cdot\mathbf{b}-c)v_0)\\
&=:\mathcal{R}_1+\mathcal{R}_1^b+\mathcal{R}_2.
\end{aligned}
$$

Now, the Hölder inequality and Lemma 5.3 lead us to write

$$
\begin{aligned}
\mathcal{R}_1\ &\leq\ C\Big(\sum_{K\subset\Omega_r}\|u-\mathcal{P}_Nu\|_{L^\infty(K)}\|\nabla v_0\|_{L^1(K)}+\sum_{K\subset\Omega\backslash\Omega_r}\|u-\mathcal{P}_Nu\|_{L^\infty(K)}\|\nabla v_0\|_{L^1(K)}\Big)\\
&\leq\ C\Big(N^{-(p+1)}\sum_{K\subset\Omega_r}\|\nabla v_0\|_{L^1(K)}+C(N^{-1}\max|\psi'|)^{p+1}\sum_{K\subset\Omega\backslash\Omega_r}\|\nabla v_0\|_{L^1(K)}\Big).
\end{aligned}
\tag{5.19}
$$

The Cauchy Schwartz and inverse inequalities give

$$
\sum_{K\subset\Omega_r}\|\nabla v_0\|_{L^1(K)}\leq CN\sum_{K\subset\Omega_r}\|v_0\|_{L^1(K)}\leq CN|\Omega_r|^{1/2}\Big(\sum_{K\subset\Omega_r}\|v_0\|^2_{L^2(K)}\Big)^{1/2}\leq CN\|v_N\|_\varepsilon. \tag{5.20}
$$

Appealing the Cauchy Schwartz inequality on $\Omega\setminus\Omega_r$, we have

$$
\begin{aligned}
\sum_{K\subset\Omega\backslash\Omega_r}\|\nabla v_0\|_{L^1(K)}&\leq\sum_{K\subset\Omega_x}\|\nabla v_0\|_{L^1(K)}+\sum_{K\subset\Omega_y}\|\nabla v_0\|_{L^1(K)}+\sum_{K\subset\Omega_{xy}}\|\nabla v_0\|_{L^1(K)}\\
&\leq\ \sqrt{\tau(1-\tau_y)}\|\nabla v_0\|_{L^2(\Omega_x)}+\ \sqrt{\tau_y(1-\tau)}\|\nabla v_0\|_{L^2(\Omega_y)}\\
&\qquad+\ \sqrt{\tau\tau_y}\|\nabla v_0\|_{L^2(\Omega_{xy})}\\
&\leq C\ln^{1/2}N\sum_{K\subset\Omega\backslash\Omega_r}\varepsilon^{1/2}\|\nabla v_0\|_{L^2(K)}\\
&\leq C(\ln N)^{1/2}\|v_N\|_\varepsilon.
\end{aligned}
\tag{5.21}
$$

Using the error bounds (5.20) and (5.21) in (5.19), we obtain

$$\begin{aligned} |\mathcal{R}_1| &\le C[N^{-p} + (N^{-1}\max|\psi'|)(\ln N)^{1/2}(N^{-1}\max|\psi'|)^p]\|v_N\|_\varepsilon \\ &\le C(N^{-1}\max|\psi'|)^p\|v_N\|_\varepsilon, \end{aligned} \tag{5.22}$$

where we have used the fact that $N^{-1}(\ln N)^{1/2}\max|\psi'| = N^{-1}(\ln N)^{1/2}|\psi'(0)| < C$.

Since u and $\mathcal{P}_N u$ are continuous, we conclude that $\|u-\mathcal{P}_N u\|_{L^\infty(e)} \le \|u-\mathcal{P}_N u\|_{L^\infty(K)}$ for any $e \subset \partial K \in \mathcal{T}_N$. Then, the Hölder inequality and Lemma 5.3 imply that

$$\begin{aligned} |\mathcal{R}_1^b| &\le \sum_{K\in\mathcal{T}_N} \|u-\mathcal{P}_N u\|_{L^\infty(\partial K)}\|v_0 - v_b\|_{L^1(\partial K)} \\ &\le \sum_{K\in\mathcal{T}_N} \|u-\mathcal{P}_N u\|_{L^\infty(\partial K)}\|v_0 - v_b\|_{L^2(\partial K)}|\partial K|^{1/2} \\ &\le \Big(\sum_{K\in\mathcal{T}_N} \|u-\mathcal{P}_N u\|^2_{L^\infty(\partial K)}|\partial K|\rho_K^{-1}\Big)^{1/2}\big(\sum_{K\in\mathcal{T}_N}\rho_K\|v_0 - v_b\|^2_{L^2(\partial K)}\big)^{1/2} \\ &\le C\Big(N^{-2p-1} + (N^{-1}\max|\psi'|)^{2(p+1)}(\varepsilon\ln N)\ln N\Big)^{1/2}\big(\sum_{K\in\mathcal{T}_N}\rho_K\|v_0 - v_b\|^2_{L^2(\partial K)}\big)^{1/2} \\ &\le CN^{-(p+1/2)}(\max|\psi'|)^{p+1}\|v_N\|_\varepsilon, \end{aligned}$$

where we have used that $\varepsilon\ln N < C$ and $(N^{-1}\max|\psi'|)^{2(p+1)}\ln N \le N^{-(2p+1)}(\max|\psi'|)^{2(p+1)}$.

Using the Cauchy Schwartz inequality and (4.6) and (4.11) of Lemma 4.5, we obtain

$$\begin{aligned} |\mathcal{R}_2| &\le C\|u-\mathcal{P}_N u\|\|v_0\| \\ &\le C(N^{-1}\max|\psi'|)^{p+1}\|v_N\|_\varepsilon. \end{aligned}$$

The proof is completed. □

By letting $v_N = e_N$ in (5.7), we obtain

$$\frac{1}{2}\frac{d}{dt}\|e_0(t)\|^2 + \mathcal{A}_w(e_N, e_N) \le |E_d(u, e_N)| + |E_c(u, e_N)| + |E_r(u, e_N)|.$$

It then follows from the estimates (5.13) and (5.14), together with Young's inequality and 3.10, that

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|e_0(t)\|^2 + C\|e_N\|_\varepsilon^2 &\le C(N^{-1}\max|\psi'|)^p\|e_N\|_\varepsilon \\ &\le C(N^{-1}\max|\psi'|)^{2p} + \frac{C}{2}\|e_N\|_\varepsilon^2. \end{aligned}$$

As a result,

$$\frac{1}{2}\frac{d}{dt}\|e_0(t)\|^2 + C\|e_N\|_\varepsilon^2 \le C(N^{-1}\max|\psi'|)^{2p}. \tag{5.23}$$

Then, by integrating from 0 to t , we have

$$\|e_0(t)\|^2 + C\int_0^t \|e_N\|_\varepsilon^2 ds \le C\Big\{\|e_0(0)\|^2 + (N^{-1}\max|\psi'|)^{2p}\Big\}.$$

This result is collected in the following theorem.

Theorem 5.1. **(Semi-discrete estimate)** *Let $u \in H^{k+1}(\Omega)$ be the solution of (1.1)-(1.2) and $u_N \in V_N^0$ be the solution of (2.15). Then, we have*

$$\|e_0(t)\|^2 + \int_0^t \|e_N\|_\varepsilon^2 ds \le C\Big\{\|e_0(0)\|^2 + (N^{-1}\max|\psi'|)^{2p}\Big\}.$$

6. Fully-discrete WG finite element scheme

In this section, we shall use the Crank-Nicolson scheme on uniform time mesh in time to derive the fully discrete approximation of the problem (1.1) and (1.2). For a given partition $0 = t_0 < t_1 < \cdots < t_M = T$ of the time interval $J = [0, T]$ for some positive integer M and step length $\tau = \frac{T}{M}$, we define

$$\partial_\tau \omega^n = \frac{\omega^{n+1} - \omega^n}{\tau} \quad \text{and} \quad \omega^{n+\frac{1}{2}} = \frac{1}{2}(\omega^{n+1} + \omega^n), \quad 0 \le n \le M-1,$$

where the sequence $\{\omega^n\}_{n=0}^M \subset L^2(\mathcal{D})$. For simplicity, we denote $\xi(\cdot, t_n)$ by ξ^n for a function $\xi : [0, T] \to L^2(\Omega)$. We now state our fully discrete weak Galerkin finite element approximation. Find $U_N^n = \{U_0^n, U_b^n\} \in V_N^0$ such that

$$(\partial_\tau U_0^n, \phi_0) + \mathcal{A}_w(U^{n+\frac{1}{2}}, \phi_N) = (f^{n+\frac{1}{2}}, \phi_0) \ \ \forall \phi_N \in V_N^0, \tag{6.1}$$

with $U_N^0 = \mathcal{P}_N u(0)$ and $f^{n+\frac{1}{2}} = \frac{1}{2}(f^{n+1} + f^n)$.

The following lemma shows that the Crank-Nicolson scheme is unconditionally stable in the L^2 norm.

Lemma 6.1. *Let $f \in C(0, T; L^2(\Omega))$. Then, we have the following stability estimate for the fully-discrete scheme* (6.1)*:*

$$\|U_0^n\| \le C\Big(\|u^0\| + T \max_{1\le n\le M} \|f(t_n)\|\Big), \ n = 0, 1, 2, \ldots, M. \tag{6.2}$$

Proof. Choosing $v = U_N^{n+1/2}$ in (6.1), and using the Cauchy-Schwarz inequality, we get

$$\frac{1}{2\tau}\left(\|U_0^{n+1}\|^2 - \|U_0^n\|^2\right) + C\|U_N^{n+\frac{1}{2}}\|_\varepsilon^2 \le C\|f^{n+\frac{1}{2}}\|\|U_0^{n+1/2}\|,$$

where we have used that $\mathcal{A}_w(U_N^{n+\frac{1}{2}}, U_N^{n+\frac{1}{2}}) \ge C\|U_N^{n+\frac{1}{2}}\|_\varepsilon^2$. Using the fact that $a^2 - b^2 = (a-b)(a+b)$ and the coercivity of the bilinear form, we have

$$\frac{1}{2\tau}\left(\|U_0^{n+1}\| - \|U_0^n\|\right) \le (C/2)\|f^{n+\frac{1}{2}}\|,$$

Let $1 \le j \le M$ be an integer. We sum the above inequality from $n = 1$ to $n = j$:

$$\begin{aligned}\|U_0^j\| &\le \|U_0^0\| + C\tau\sum_{n=1}^{j}\|f^{n+\frac{1}{2}}\| \\ &\le C(\|u^0\| + M\tau \max_{1\le n\le M}\|f(t_n)\|).\end{aligned}$$

Recalling that $M\tau = T$, the result follows. The proof is now completed.

□

Next, we shall present the convergence analysis. To begin, we prove the error estimate of the discretization error $\mathcal{P}_N u(t_n) - U_N^n$. To this end, we need to derive an error equation involving the error $e_N^n := U_N^n - \mathcal{P}_N u^n$.

We formulate the error equation for e_N^n in the following lemma.

Lemma 6.2. *For $v_N = \{v_0, v_b\} \in V_N^0$, we have*

$$(\partial_\tau e_0^n, v_0) + \mathcal{A}_w(e_N^{n+\frac{1}{2}}, v_N) \quad = \quad (\xi^n, v_0) + E(u^{n+\frac{1}{2}}, v_N), \tag{6.3}$$

where $\xi^n = \frac{1}{\tau}(\mathcal{P}_N u^{n+1} - \mathcal{P}_N u^n) - \partial_t u^{n+\frac{1}{2}}$; and $E(u, v_N) = E_d(u, v_N) + E_c(u, v_N) + E_r(u, v_N)$.

Proof. From (1.1), one obtains the following equation:

$$\partial_t u^{n+\frac{1}{2}} - \varepsilon \Delta u^{n+\frac{1}{2}} + \mathbf{b} \cdot \nabla u^{n+\frac{1}{2}} + c u^{n+\frac{1}{2}} = f^{n+\frac{1}{2}}. \tag{6.4}$$

On each element $T \in \mathcal{T}_N$, for $v_N = \{v_0, v_b\} \in V_N^0$, we test equation (6.4) against v_0 to arrive at

$$\begin{aligned}(f^{n+\frac{1}{2}}, v_0) \quad = \quad & (\partial_t u^{n+\frac{1}{2}}, v_0)_{\mathcal{T}_N} - \sum_{T \in \mathcal{T}_N} (\varepsilon \Delta u^{n+\frac{1}{2}}, v_0)_T \\ & + \sum_{T \in \mathcal{T}_N} (\mathbf{b} \cdot \nabla u^{n+\frac{1}{2}}, v_0)_T + (c u^{n+\frac{1}{2}}, v_0)_{\mathcal{T}_N}.\end{aligned} \tag{6.5}$$

Using a similar argument in deriving (5.8), one can show that

$$(\partial_t u^{n+\frac{1}{2}}, v_0)_{\mathcal{T}_N} + \mathcal{A}_w(\mathcal{P}_N u^{n+\frac{1}{2}}, v_N) = (f^{n+\frac{1}{2}}, v_0) + E_d(u^{n+\frac{1}{2}}, v_N) + E_c(u^{n+\frac{1}{2}}, v_N) + E_r(u^{n+\frac{1}{2}}, v_N),$$

where we have used that $\mathcal{S}_d(\mathcal{P}_N u^{n+\frac{1}{2}}, v_N) = 0$ and $\mathcal{S}_c(\mathcal{P}_N u^{n+\frac{1}{2}}, v_N) = 0$ since $\mathcal{P}_N u$ is continuous in Ω. Thus, we get

$$\frac{1}{\tau}(\mathcal{P}_N u^{n+1} - \mathcal{P}_N u^n, v_0)_{\mathcal{T}_N} + \mathcal{A}_w(\mathcal{P}_N u, v_N) = (f^{n+\frac{1}{2}}, v_0) + E(u, v_N) + (\xi^n, v_0)_{\mathcal{T}_N}, \tag{6.6}$$

Subtracting (6.1) from (6.6) gives the conclusion. We complete the proof. □

Lemma 6.3. *Let $u \in H^{k+1}(\Omega)$. Assume that u and U_N^n are the solutions (1.1), (1.2), and (6.1), respectively. One has for $n = 1, 2, \ldots, M$,*

$$\|e_0^n\|^2 + C\tau \sum_{j=0}^{n-1} \|e_N^{j+\frac{1}{2}}\|_\varepsilon^2 \le C\Big(\tau^4 \int_0^{t_n} \|u_{ttt}(s)\|^2 ds + (N^{-1} \max |\psi'|)^{2p}\Big). \tag{6.7}$$

Proof. Choosing $v_N = e_N^{n+\frac{1}{2}}$ in (6.3) and by the coercivity property (3.6), we find

$$\frac{1}{2\tau}(\|e_0^{n+1}\|^2 - \|e_0^n\|^2) + C\|e_N^{n+\frac{1}{2}}\|_\varepsilon^2 \quad \le \quad (\xi^n, e_0^{n+\frac{1}{2}}) + E(u^{n+\frac{1}{2}}, e_N^{n+\frac{1}{2}}),$$

or, equivalently,

$$\|e_0^{n+1}\|^2 - \|e_0^n\|^2 + 2C\tau\|e_N^{n+\frac{1}{2}}\|_\varepsilon^2 \quad \le \quad 2\tau(\xi^n, e_0^{n+\frac{1}{2}}) + 2\tau E(u^{n+\frac{1}{2}}, e_N^{n+\frac{1}{2}})$$

$$:= \quad \mathcal{W}_1 + \mathcal{W}_2. \tag{6.8}$$

We can express the term $\xi^n = (\partial_\tau \mathcal{P}_N u^n - \partial_\tau u^n) + (\partial_\tau u^n - \partial_t u^{n+\frac{1}{2}}) =: T_1 + T_2$. We write

$$T_1 = \frac{1}{\tau}\int_{t_n}^{t_{n+1}} \frac{\partial}{\partial t}(\mathcal{P}_N u(\cdot, s) - u(\cdot, s))\, ds \leq \frac{1}{\tau}\int_{t_n}^{t_{n+1}} |\mathcal{P}_N u_t(\cdot, s) - u_t(\cdot, s)|\, ds, \tag{6.9}$$

and

$$\begin{aligned} T_2 &= \frac{1}{\tau}\Big(u(t_{n+1}) - u(t_n)\Big) - \frac{1}{2}\Big(u_t(t_{n+1}) + u_t(t_n)\Big) \\ &= \frac{1}{2\tau}\Big(\int_{t_n}^{t_{n+1}} (t_{n+1} - s)(t_n - s) u_{ttt}(s) ds\Big). \end{aligned} \tag{6.10}$$

From (6.9) and (6.10), we obtain

$$\begin{aligned} \|\xi^n\|^2 &\leq \int_\Omega \left[\frac{1}{2\tau}\left[\int_{t_n}^{t_{n+1}} (t_{n+1} - s)(t_n - s) u_{ttt}(s) ds\right]^2 dx\right. \\ &\quad + \int_\Omega \left[\frac{1}{\tau}\int_{t_n}^{t_{n+1}} |\mathcal{P}_N u_t(\cdot, s) - u_t(\cdot, s)|\, ds\right]^2 dx \\ &\leq \frac{1}{4\tau^2}\int_\Omega \left[\int_{t_n}^{t_{n+1}} (t_{n+1} - s)^2 (t_n - s)^2 ds \int_{t_n}^{t_{n+1}} u_{ttt}^2(s) ds\right] dx \\ &\quad + \frac{1}{\tau^2}\int_\Omega \left[\int_{t_n}^{t_{n+1}} |\mathcal{P}_N u_t(\cdot, s) - u_t(\cdot, s)|^2\, ds\right] dx \\ &\leq \frac{\tau^3}{120}\int_{t_n}^{t_{n+1}} \|u_{ttt}(s)\|^2 dt + \frac{1}{\tau}\|\mathcal{P}_N u_t - u_t\|^2_{L^\infty(0,T;L^2(\Omega))}. \end{aligned} \tag{6.11}$$

Hence, with the aid of the Cauchy-Schwarz and the Poincare inequality, $\mathcal{W}_1$ in (6.8) can be estimated as follows.

$$\begin{aligned} |\mathcal{W}_1| &= |2\tau(\xi^n, e_0^{n+\frac{1}{2}})| \leq 2\tau\|\xi^n\|\|e_0^{n+\frac{1}{2}}\| \leq 2\tau\|\xi^n\|\|e_N^{n+\frac{1}{2}}\|_\varepsilon \\ &\leq \frac{\tau}{C}\|\xi^n\|^2 + \tau C\|e_N^{n+\frac{1}{2}}\|_\varepsilon^2 \\ &\leq C(\tau^4 \int_{t_n}^{t_{n+1}} \|u_{ttt}(s)\|^2 ds + \|\mathcal{P}_N u_t - u_t\|^2_{L^\infty(0,T;L^2(\Omega))}) + C\tau\|e_N^{n+\frac{1}{2}}\|_\varepsilon^2 \\ &\leq C(\tau^4 \int_{t_n}^{t_{n+1}} \|u_{ttt}(s)\|^2 ds + (N^{-1}\max|\psi'|)^{2p+2} + C\tau\|e_N^{n+\frac{1}{2}}\|_\varepsilon^2, \end{aligned} \tag{6.12}$$

where we have used the Young's inequities in the second inequality, and the estimate (6.11) and Lemma 4.5 in the second estimates of the righthand side. Applying Lemma 5.5 and Young's inequality, we obtain the estimate of the term I_2 in the righthand side of (6.8) as follows:

$$\begin{aligned} |I_2| &\leq C(2\tau(N^{-1}\max|\psi'|)^p)\|e_N^{n+\frac{1}{2}}\|_\varepsilon \\ &\leq C\tau(N^{-1}\max|\psi'|)^{2p} + C\tau\|e_N^{n+\frac{1}{2}}\|_\varepsilon^2. \end{aligned} \tag{6.13}$$

Combining (6.8)–(6.13) yields

$$\|e_0^{n+1}\|^2 - \|e_0^n\|^2 + C\tau\|e_N^{n+\frac{1}{2}}\|_\varepsilon^2 \leq C\tau^4 \int_{t_n}^{t_{n+1}} \|u_{ttt}(s)\|^2 ds + C\tau(N^{-1}\max|\psi'|)^{2p}.$$

Let $1 \leq j \leq M$. Using the fact that $e_N^0 = 0$, we sum the above expression from $n = 0$ to $n = j-1$ to obtain

$$\begin{aligned}\|e_0^j\|^2 + C\tau\sum_{n=0}^{j-1}\|e_N^{n+\frac{1}{2}}\|_\varepsilon^2 &\leq C\tau^4\int_0^{t_j}\|u_{ttt}(s)\|^2 ds + C\sum_{n=0}^{j-1}\tau(N^{-1}\max|\psi'|)^{2p}\\ &\leq C(\tau^4\int_0^{t_j}\|u_{ttt}(s)\|^2 ds + (N^{-1}\max|\psi'|)^{2p}).\end{aligned}$$

We complete the proof. □

Theorem 6.1. *Let $u \in H^{k+1}(\Omega)$. Assume that u and U_N^n are the solutions (1.1), (1.2), and (6.1), respectively. One has for $n = 1, 2, \ldots, M$,*

$$\|e_N^{n+1}\|_\varepsilon^2 \leq C(\tau^4\int_0^{t_m}\|u_{ttt}(s)\|^2 ds + C(N^{-1}\max|\psi'|)^{2p}).$$

Proof. Choosing $v = \partial_\tau e_N^n$ in (6.3) and by coercivity (3.6), we find

$$\|\partial_\tau e_0^n\|^2 + \mathcal{A}_w(e_N^{n+\frac{1}{2}}, \partial_\tau e_N^n) = (\xi^n, \partial_\tau e_0^n) + E(u^{n+\frac{1}{2}}, \partial_\tau e_N^n).$$

or, equivalently,

$$\begin{aligned}2\tau\|\partial_\tau e_0^n\|^2 + \mathcal{A}_w(e_N^{n+1}, e_N^{n+1}) - \mathcal{A}_w(e_N^n, e_N^n) &= 2\tau(\xi^n, \partial_\tau e_0^n) + 2\tau E(u^{n+\frac{1}{2}}, \partial_\tau e_N^n)\\ &\leq \tau\|\xi^n\|^2 + \tau\|\partial_\tau e_0^n)\|^2 + 2\tau E(u^{n+\frac{1}{2}}, \partial_\tau e_N^n).\end{aligned}$$

Thus, we have

$$\tau\|\partial_\tau e_0^n\|^2 + \mathcal{A}_w(e_N^{n+1}, e_N^{n+1}) - \mathcal{A}_w(e_N^n, e_N^n) \leq \tau\|\xi^n\|^2 + 2\tau E(u^{n+\frac{1}{2}}, \partial_\tau e_N^n).$$

Because $e_0^0 = 0$, we sum up the above term from $n = 0$ to $n = m-1$ for any fixed m to get

$$\sum_{n=0}^{m-1}\tau\left\|\partial_\tau e_0^n\right\|^2 + C\left\|e_N^m\right\|_\varepsilon^2 \leq \sum_{n=0}^{m-1}\tau\|\xi^n\|^2 + \sum_{n=0}^{m-1}2\tau E\left(u^{n+1/2}, \partial_\tau e_N^n\right). \tag{6.14}$$

From (6.11), we have

$$\sum_{n=0}^{m-1}\tau\|\xi^n\|^2 \leq \frac{\tau^4}{120}\int_0^{t_m}\|u_{ttt}(s)\|^2\ \mathrm{d}s + \|\mathcal{P}_N u_t - u_t\|_{L^\infty(0,T;L^2(\Omega))}^2. \tag{6.15}$$

Observe that

$$\sum_{n=0}^{m-1} 2\tau E\left(u^{n+1/2}, \partial_\tau e_N^n\right) = \sum_{n=0}^{m-1} \tau E\left(-\partial_\tau u^n, e_N^{n+1/2}\right) + 2E\left(u^m, e_N^m\right)$$
$$:= J_1 + J_2. \tag{6.16}$$

Similar to (6.13), one has

$$|J_1| \leq C\sum_{n=0}^{m-1} \tau(N^{-1}\max|\psi'|)^{2p} + C\tau\sum_{n=0}^{m-1} \|e_N^{n+\frac{1}{2}}\|_\varepsilon^2. \tag{6.17}$$

It follows from Lemma 5.5, the Cauchy-Schwarz inequality, and Young's inequality that

$$|J_2| \leq C(N^{-1}\max|\psi'|)^{2p} + C\|e_N^m\|_\varepsilon^2. \tag{6.18}$$

From (6.16), (6.17), and (6.18) together with $\tau M = T$, we have

$$\sum_{n=0}^{m-1} 2\tau E\left(u^{n+1/2}, \partial_\tau e_N^n\right) \leq C(N^{-1}\max|\psi'|)^{2p} + C\tau\sum_{n=0}^{m-1} \|e_N^{n+\frac{1}{2}}\|_\varepsilon^2 + C\|e_N^m\|_\varepsilon^2. \tag{6.19}$$

Combining (6.14), (6.15), and (6.19) yields that

$$\sum_{n=0}^{m-1} \tau\left\|\partial_\tau e_0^n\right\|^2 + C\left\|e_N^m\right\|_\varepsilon^2 \leq C(\tau^4\int_0^{t_m} \|u_{ttt}(s)\|^2 \, \mathrm{d}s + (N^{-1}\max|\psi'|)^{2p} + C\tau\sum_{n=0}^{m-1} \|e_N^{n+\frac{1}{2}}\|_\varepsilon^2).$$

Finally, using (6.7), we obtain

$$\sum_{n=0}^{m-1} \tau\left\|\partial_\tau e_0^n\right\|^2 + C\left\|e_N^m\right\|_\varepsilon^2 \leq C(\tau^4\int_0^{t_m} \|u_{ttt}(s)\|^2 \, \mathrm{d}s + (N^{-1}\max|\psi'|)^{2p})$$

which completes the proof.

□

7. Numerical Experiments

This section presents various numerical examples for the fully-discrete Crank-Nicolson WG finite element method. We used MATLAB R2020A in our the calculations. We also used the 5-point Gauss-Legendre quadrature rule for evaluating of all integrals. All the calculations were calculated using MATLAB R2016a. The systems of linear equations resulting from the discrete problems were solved by lower-upper (LU) decomposition.

We apply the fully-discrete WG-FEM on the adaptive meshes shown in Table 1. We choose $\sigma = p + 1$ and calculate the energy-norm $\|e_N^n\|_E$ and the L^2-norm error $\|e_0^n\|$, where $e_N = \{e_0^n, e_b^n\} = \{u - U_0^n, u - U_b^n\}$ is the error using N intervals in each direction. The order of convergence (OC) is computed by the formula

$$OC(2) = \log_2\left(\|e_N\| / \|e_{2N}\|\right), \quad OC(S) = \frac{\log\left(\|e_N\| / \|e_{2N}\|\right)}{\log(2\log N/\log(2N))}.$$

The numerical errors and the order of convergences in space are also tested. In order for the space error to dominate the errors, we take $\tau = N^{-2}$ for N element in each direction. We list the errors in the energy norm and L^2-norm and the order of convergence in Tables 2 and 3, respectively. These numerical results show that the order of convergence is of order p and of order $p+1$ in the energy and L^2 norms, respectively, which support the stated error estimates in Theorem 6.1.

Table 2. The energy-error and the order of convergence in space for Example 7.1 $\varepsilon = 10^{-5}$.

		Shishkin		Bakhvalov- Shishkin		Bakhvalov-type	
	N	$\|\|e_N^n\|\|_E$	$OC(S)$	$\|\|e_N^n\|\|_E$	$OC(2)$	$\|\|e_N^n\|\|_E$	$OC(2)$
	16	1.256×10^{-1}	–	6.311×10^{-2}	–	6.604×10^{-2}	–
	32	8.192×10^{-2}	0.90	3.381×10^{-2}	0.90	3.466×10^{-2}	0.93
	64	5.102×10^{-2}	0.93	1.774×10^{-2}	0.93	1.819×10^{-2}	0.93
$\mathbb{P}_1$	128	3.063×10^{-2}	0.95	9.182×10^{-3}	0.95	9.547×10^{-3}	0.93
	256	1.781×10^{-2}	0.97	4.687×10^{-3}	0.97	4.873×10^{-3}	0.97
	512	1.008×10^{-2}	0.99	2.359×10^{-3}	1.00	2.453×10^{-3}	0.99
	16	2.406×10^{-2}	–	3.344×10^{-3}	–	3.9238×10^{-3}	–
	32	9.828×10^{-3}	1.90	9.603×10^{-4}	1.80	1.001×10^{-3}	1.97
	64	3.644×10^{-3}	1.94	2.573×10^{-4}	1.90	2.626×10^{-4}	1.93
$\mathbb{P}_2$	128	1.267×10^{-3}	1.96	6.705×10^{-5}	1.94	6.891×10^{-5}	1.93
	256	4.189×10^{-4}	1.97	1.687×10^{-5}	1.99	1.758×10^{-6}	1.97
	512	1.325×10^{-4}	2.00	4.246×10^{-6}	1.99	4.425×10^{-6}	2.00
	16	4.603×10^{-3}	–	5.866×10^{-4}	–	6.402×10^{-4}	–
	32	1.180×10^{-3}	2.90	8.422×10^{-5}	2.80	8.576×10^{-5}	2.90
	64	2.633×10^{-4}	2.94	1.112×10^{-5}	2.92	1.125×10^{-5}	2.93
$\mathbb{P}_3$	128	5.299×10^{-5}	2.97	1.419×10^{-6}	2.97	1.435×10^{-6}	2.97
	256	1.001×10^{-5}	2.98	1.798×10^{-7}	2.98	1.806×10^{-7}	2.99
	512	1.782×10^{-6}	3.00	2.263×10^{-8}	3.00	2.273×10^{-8}	3.00

Example 7.1. *Let $\boldsymbol{b} = (1, 1)$ and $T = 1$ in the problem (1.1). We choose f and u^0 such that the exact solution is*

$$u(x, y, t) = e^{-t}xy(1 - x)(1 - y)\kappa(x)\kappa(y),$$

where $\kappa(z) = 1 - e^{-(1-z)/\varepsilon}$.

In Figure 2, we plot the numerical solutions of the WG-FEM using the $\mathbb{P}_1$ element on the three layer-adapted meshes given in Figure 1 for $\varepsilon = 10^{-5}$ and $N = 32$.

Table 3. The L^2 error and the order of convergence in space for Example 7.1 $\varepsilon = 10^{-5}$.

		Shishkin		Bakhvalov- Shishkin		Bakhvalov-type	
	N	$\|\|e_0^n\|\|$	$OC(2)$	$\|\|e_0^n\|\|$	$OC(2)$	$\|\|e_0^n\|\|$	$OC(2)$
	16	1.045×10^{-2}	–	1.021×10^{-2}	–	1.023×10^{-2}	–
	32	2.723×10^{-3}	1.94	2.671×10^{-3}	1.94	2.675×10^{-3}	1.94
	64	6.950×10^{-4}	1.97	6.944×10^{-4}	1.97	6.946×10^{-4}	1.97
$\mathbb{P}_1$	128	1.761×10^{-5}	1.98	1.760×10^{-5}	1.98	1.760×10^{-5}	1.98
	256	4.433×10^{-6}	1.99	4.430×10^{-6}	1.99	4.432×10^{-6}	1.99
	512	1.109×10^{-7}	1.99	1.105×10^{-7}	1.99	1.108×10^{-7}	1.99
	16	2.297×10^{-4}	–	2.286×10^{-4}	–	2.297×10^{-4}	–
	32	3.013×10^{-5}	2.93	3.010×10^{-5}	2.93	3.011×10^{-5}	2.93
	64	3.845×10^{-6}	2.97	3.832×10^{-6}	2.97	3.844×10^{-6}	2.97
$\mathbb{P}_2$	128	4.873×10^{-7}	2.98	4.862×10^{-7}	2.98	4.868×10^{-7}	2.98
	256	6.133×10^{-8}	2.99	6.130×10^{-8}	2.99	6.132×10^{-8}	2.99
	512	7.666×10^{-9}	3.00	7.662×10^{-9}	3.00	7.664×10^{-9}	3.00
	16	3.783×10^{-5}	–	3.780×10^{-5}	–	3.782×10^{-5}	–
	32	2.481×10^{-6}	3.93	2.479×10^{-6}	3.93	2.480×10^{-6}	3.93
	64	1.616×10^{-7}	3.94	1.613×10^{-7}	3.94	1.614×10^{-7}	3.94
$\mathbb{P}_3$	128	1.031×10^{-8}	3.97	1.028×10^{-8}	3.97	1.030×10^{-8}	3.97
	256	6.488×10^{-10}	3.99	6.485×10^{-10}	3.99	6.487×10^{-10}	3.99
	512	4.083×10^{-11}	4.00	4.077×10^{-11}	4.00	4.079×10^{-11}	4.00

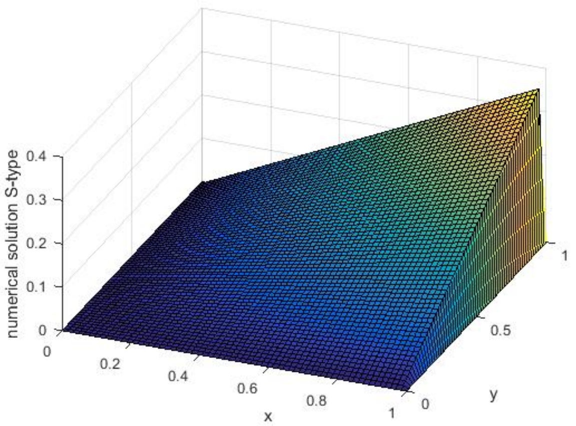
We next present the temporal convergence rate for Example 7.1. In order for the temporal error to dominate the error, we take $N = 256$ and $\varepsilon = 10^{-5}$, and use the $\mathbb{P}_3$ element. We report the results in the L^2-norm and the energy norm in Tables 4 and 5, respectively. We see that the order of convergence in time is of order $O(\tau^2)$, which verifies the theoretical estimate claimed in Theorem 6.1.

Table 4. The L^2 error and the order of convergence in time for Example 7.1 $\varepsilon = 10^{-5}$.

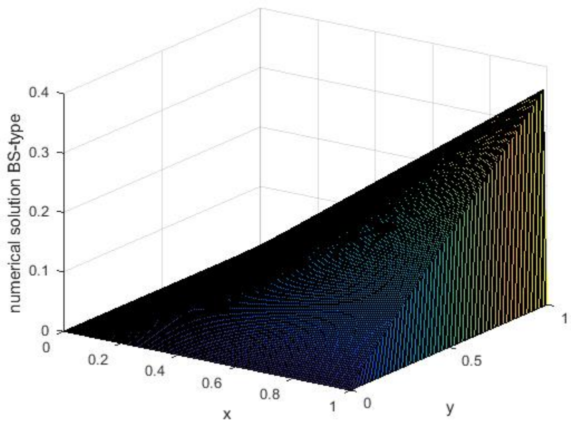
	Shishkin		Bakhvalov- Shishkin		Bakhvalov-type	
τ	$\|\|e_0^n\|\|$	$OC(2)$	$\|\|e_0^n\|\|$	$OC(2)$	$\|\|e_0^n\|\|$	$OC(2)$
1/2	7.425×10^{-3}	–	7.424×10^{-3}	–	7.425×10^{-3}	–
1/4	1.818×10^{-3}	2.03	1.816×10^{-3}	2.03	1.818×10^{-3}	2.03
1/8	4.545×10^{-4}	1.99	4.545×10^{-4}	1.99	4.545×10^{-4}	1.99
1/16	1.136×10^{-4}	2.00	1.135×10^{-4}	2.00	1.136×10^{-4}	2.00

Table 5. The energy error and the order of convergence in time for Example 7.1 $\varepsilon = 10^{-5}$.

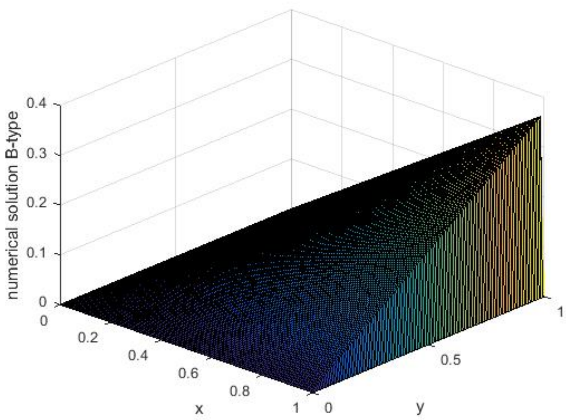
	Shishkin		Bakhvalov- Shishkin		Bakhvalov-type	
τ	$\|\|e_N^n\|\|_E$	$OC(2)$	$\|\|e_N^n\|\|_E$	$OC(2)$	$\|\|e_N^n\|\|_E$	$OC(2)$
1/2	7.425×10^{-3}	–	7.424×10^{-3}	–	7.425×10^{-3}	–
1/4	1.856×10^{-3}	2.00	1.852×10^{-3}	2.00	1.855×10^{-3}	2.00
1/8	4.672×10^{-4}	1.99	4.668×10^{-4}	1.99	4.669×10^{-4}	1.99
1/16	1.168×10^{-4}	2.00	1.104×10^{-4}	2.00	1.106×10^{-4}	2.00



(a) S-type mesh



(b) BS-type mesh



(c) B-type mesh

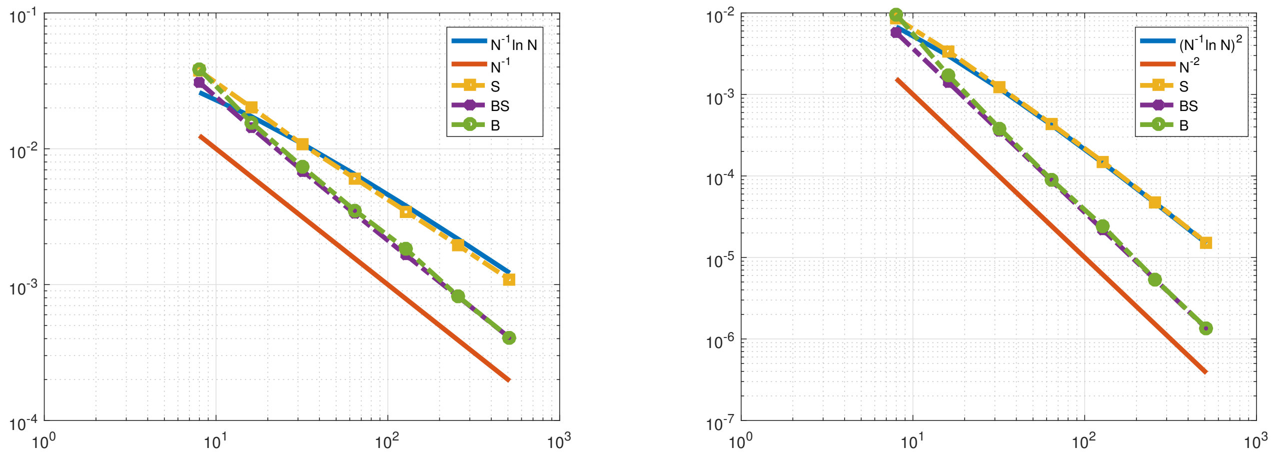
Figure 2. Numerical solution of Example 7.1 for $\varepsilon = 10^{-5}$ using $\mathbb{P}_1$.

Lastly, we test the robustness of the WG-FEM method with respect to the small parameter ε for Example 7.1. We take $N = 256$ and use the $\mathbb{P}_1$ element for the values of $\varepsilon = 10^{-r}$, $r = 5, 6, \ldots, 10$. The results are reported in Table 6. These results show that the WG-FEM is robust with respect to the perturbation parameter ε.

Table 6. The energy-error in space for Example 7.1 for the values of ε.

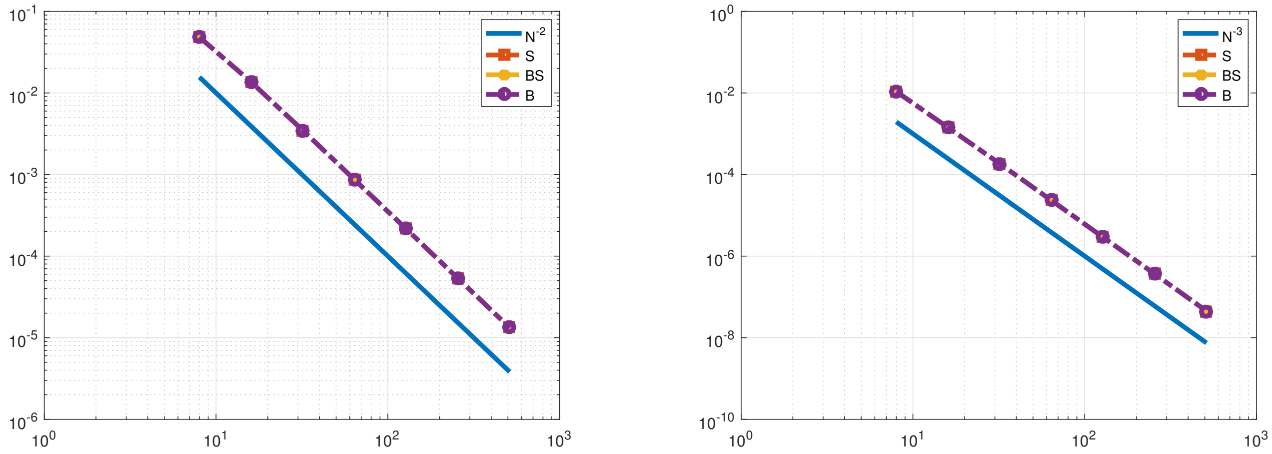
	Shishkin	Bakhvalov- Shishkin	Bakhvalov-type
ε	$\|\|e_N^n\|\|_E$	$\|\|e_N^n\|\|_E$	$\|\|e_N^n\|\|_E$
10^{-5}	1.781×10^{-2}	4.687×10^{-3}	4.873×10^{-3}
10^{-6}	1.780×10^{-2}	4.685×10^{-3}	4.877×10^{-3}
10^{-7}	1.780×10^{-2}	4.685×10^{-3}	4.877×10^{-3}
10^{-8}	1.780×10^{-2}	4.685×10^{-3}	4.877×10^{-3}
10^{-9}	1.780×10^{-2}	4.685×10^{-3}	4.686×10^{-3}
10^{-10}	1.780×10^{-2}	4.685×10^{-3}	4.879×10^{-3}

The order of convergence via loglog plot in the energy norm and L^2 norm are plotted in Figures 3 and 4, respectively, for Example 7.2. We observe that the order of convergence of order p and of order $p+1$ in the energy and L^2 norms, respectively, which support the stated error estimates in Theorem 6.1 as in Example 7.1. To test the temporal error, we choose $N = 256$ and $\varepsilon = 10^{-5}$, and use the $\mathbb{P}_1$ element. We present the results in the L^2-norm and the energy norm in Tables 7 and 8, respectively. We see that the order of convergence in time is of order $O(\tau^2)$ as claimed in Theorem 6.1.



(a) $\mathbb{P}_1$ element (b) $\mathbb{P}_2$ element.

Figure 3. The order of convergence in energy norm via loglog plot for Example 7.2 for $\varepsilon = 10^{-5}$ using $\mathbb{P}_1$ and $\mathbb{P}_2$ elements.



(a) $\mathbb{P}_1$ element (b) $\mathbb{P}_2$ element.

Figure 4. The order of convergence in L^2-norm via loglog plot for Example 7.2 for $\varepsilon = 10^{-5}$ using $\mathbb{P}_1$ and $\mathbb{P}_2$ elements.

Example 7.2. *Let* $\boldsymbol{b} = (1 + x, 2 - y)$*,* $c = (1 + x^2 + y^2)$*, and* $T = 1$ *in the problem (1.1). We take* f *and* u^0 *such that the exact solution is*

$$u(x, y, t) = e^t xy(1 - e^{-(3-2x-x^2)/(2\varepsilon)})(1 - e^{-(3-4y+y^2)/(2\varepsilon)}).$$

Table 7. The L^2 error and the order of convergence in time for Example 7.2 $\varepsilon = 10^{-5}$.

	Shishkin		Bakhvalov- Shishkin		Bakhvalov-type	
τ	$\|\|e_0^n\|\|$	$OC(2)$	$\|\|e_0^n\|\|$	$OC(2)$	$\|\|e_0^n\|\|$	$OC(2)$
1/2	7.391×10^{-2}	–	7.390×10^{-2}	–	7.391×10^{-2}	–
1/4	3.217×10^{-2}	1.20	3.215×10^{-2}	1.20	3.216×10^{-2}	1.20
1/8	9.901×10^{-3}	1.70	9.895×10^{-3}	1.70	9.900×10^{-3}	1.70
1/16	6.449×10^{-4}	2.00	6.445×10^{-4}	2.00	6.448×10^{-4}	2.00

Table 8. The energy error and the order of convergence in time for Example 7.2 $\varepsilon = 10^{-5}$.

	Shishkin		Bakhvalov- Shishkin		Bakhvalov-type	
τ	$\|\|e_N^n\|\|_E$	$OC(2)$	$\|\|e_N^n\|\|_E$	$OC(2)$	$\|\|e_N^n\|\|_E$	$OC(2)$
1/2	6.521×10^{-2}	–	6.518×10^{-2}	–	6.520×10^{-2}	–
1/4	1.723×10^{-2}	1.92	1.720×10^{-2}	1.92	1.722×10^{-2}	1.92
1/8	4.277×10^{-3}	2.01	4.274×10^{-3}	2.01	4.276×10^{-3}	2.01
1/16	1.061×10^{-3}	2.02	1.057×10^{-3}	2.02	1.060×10^{-3}	2.02

We also test the WG-FEM for Example 7.2 for the robustness against ε. The results are presented in Table 9 for $N = 256$ and the $\mathbb{P}_1$ element for the values of $\varepsilon = 10^{-r},\ r = 5, 6, \ldots, 10$. Again, one sees that the WG-FEM is the parameter-uniform method.

Table 9. The energy-error in space for Example 7.2 for the values of ε.

	Shishkin	Bakhvalov- Shishkin	Bakhvalov-type
ε	$\|\|e_N^n\|\|_E$	$\|\|e_N^n\|\|_E$	$\|\|e_N^n\|\|_E$
10^{-5}	7.856×10^{-4}	1.432×10^{-4}	1.441×10^{-4}
10^{-6}	7.852×10^{-4}	1.430×10^{-4}	1.440×10^{-4}
10^{-7}	7.852×10^{-4}	1.430×10^{-4}	1.440×10^{-4}
10^{-8}	7.852×10^{-4}	1.430×10^{-4}	1.440×10^{-4}
10^{-9}	7.852×10^{-4}	1.430×10^{-4}	1.440×10^{-4}
10^{-10}	7.843×10^{-4}	1.439×10^{-4}	1.440×10^{-4}

8. Conclusions

In this paper, we present the Crack-Nicolson- WG-FEM applied to the singularly perturbed parabolic convection-dominated problems in 2D. We use the Crack-Nicolson scheme in time on uniform mesh and the WG-FEM in space on three layer-adapted meshes: Shishkin, Bakhvalov-Shishkin, and Bakhvalov meshes. We prove (almost) uniform error estimates of order p in the energy norm and second order estimate in time. With the use of a special interpolation operator, the error analysis of the semi-discrete WG-FEM and the fully discrete WG-FEM have been carried out. Various numerical examples are conducted to validate the convergence rate of the proposed method.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The authors declare there is no conflicts of interest.

References

1. H. Roos, M. Stynes, L. Tobiska, *Robust Numerical Methods for Singularly Perturbed Differential Equations*, Springer, Berlin, 2008. https://doi.org/10.1007/978-3-540-34467-4
2. T. Linß, *Layer-adapted Meshes for Reaction-Convection-Diffusion Problems*, Springer, Berlin, 2009. https://doi.org/10.1007/978-3-642-05134-0

3. C. Clavero, J. C. Jorge, F. Lisbona, G. I. Shishkin, A fractional step method on a special mesh for the resolution of multidimensional evolutionary convection-diffusion problems, *Appl. Numer. Math.*, **27** (1998), 211–231 https://doi.org/10.1016/S0168-9274(98)00014-2

4. T. Linß, M. Stynes, Numerical methods on Shishkin meshes for linear convection-diffusion problems, *Comput. Methods Appl. Mech. Eng.*, **190** (2001), 3527–3542. https://doi.org/10.1016/S0045-7825(00)00271-1

5. T. E. Tezduyar, Y. J. Park, H. A. Deans, Finite element procedures for time-dependent convection-diffusion-reaction systems, *Int. J. Numer. Methods Fluids*. **7** (1987), 1013–1033. https://doi.org/10.1002/fld.1650071003

6. V. John, E. Schmeyer, Finite element methods for time-dependent convection-diffusion-reaction equations with small diffusion, *Comput. Methods Appl. Mech. Eng.*, **198** (2008), 475–494. https://doi.org/10.1016/j.cma.2008.08.016

7. G. Richter, A finite element method for time-dependent convection-diffusion equations, *Math. Comp.*, **54** (1990), 81–106. https://doi.org/10.1090/S0025-5718-1990-0993932-9

8. S. Toprakseven, S. Dinibutun, A high-order stabilizer-free weak Galerkin finite element method on nonuniform time meshes for subdiffusion problems, *AIMS Math.*, **8** (2023), 31022–31049. https://doi.org/10.3934/math.20231588

9. S. Toprakseven, S. Dinibutun, Error estimations of a weak Galerkin finite element method for a linear system of coupled singularly perturbed reaction-diffusion equations in the energy and balanced norms, *AIMS Math.*, **8** (2023), 15427–15465. https://doi.org/10.3934/math.2023788

10. S. Toprakseven, F. Gao, A modified weak Galerkin finite element method for singularly perturbed parabolic convection-diffusion-reaction problems, *J. Comput. Math.*, **41** (2023), 1246–1280. https://doi.org/10.4208/jcm.2203-m2021-0031

11. V. John, J. Novo, Error analysis of the SUPG finite element discretization of evolutionary convection-diffusion-reaction equations, *SIAM J. Numer. Anal.*, **49** (2011), 1149–1176. https://doi.org/10.1137/100789002

12. E. Burman, Consistent SUPG-method for transient transport problems: Stability and convergence, *Comput. Methods Appl. Mech. Eng.*, **199** (2010), 1114–1123. https://doi.org/10.1016/j.cma.2009.11.023

13. P. Houston, C. Schwab, E. Süli, Discontinuous hp-finite element methods for advection-diffusion-reaction problems, *SIAM J. Numer. Anal.*, **39** (2002), 2133–2163. https://doi.org/10.1137/S0036142900374111

14. B. Cockburn, C. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, *SIAM J. Numer. Anal.*, **35** (1998), 2440–2463. https://doi.org/10.1137/S0036142997316712

15. Y. Li, W. Zhao, W. Zhao, Spatio-temporal scalar auxiliary variable approach for the nonlinear convection-diffusion equation with discontinuous Galerkin method, *Numer. Methods Partial Differ. Equations*, **40** (2024), e23061. https://doi.org/10.1002/num.23061

16. J. Wang, X. Ye, A weak Galerkin finite element method for second-order elliptic problems, *J. Comput. Appl. Math.*, **241** (2013), 103–115. https://doi.org/10.1016/j.cam.2012.10.003

17. Z. Dong, A. Ern, Hybrid high-order method for singularly perturbed fourth-order problems on curved domains, *ESAIM: Math. Modell. Numer. Anal.*, **55** (2021), 3091–3114. https://doi.org/10.1051/m2an/2021081

18. J. Wang, R. Wang, Q. Zhai, R. Zhang, A systematic study on weak Galerkin finite element methods for second order elliptic problems, *J. Sci. Comput.*, **74** (2018), 1369–1396. https://doi.org/10.1007/s10915-017-0496-6

19. A. Al-Taweel, S. Hussain, X. Wang, B. Jones, A $P_0 - P_0$ weak Galerkin finite element method for solving singularly perturbed reaction-diffusion problems, *Numer. Methods Partial Differ. Equations*, **36** (2020), 213–227. https://doi.org/10.1002/num.22415

20. B. Deka, N. Kumar, Error estimates in weak Galerkin finite element methods for parabolic equations under low regularity assumptions, *Appl. Numer. Math.*, **162** (2021), 81–105. https://doi.org/10.1016/j.apnum.2020.12.003

21. B. Deka, N. Kumar, A systematic study on weak Galerkin finite element method for second-order parabolic problems, *Numer. Methods Partial Differ. Equations*, **39** (2023), 2444–2474. https://doi.org/10.1002/num.22973

22. H. Zhang, Y. Zou, Y. Xu, Q. Zhai, H. Yue, Weak Galerkin finite element method for second order parabolic equations, *Int. J. Numer. Anal. Model.*, **13** (2016), 525–544.

23. Y. Huang, J. Li, D. Li, Developing weak Galerkin finite element methods for the wave equation, *Numer. Methods Partial Differ. Equations*, **33** (2017), 868–884. https://doi.org/10.1002/num.22127

24. A. AL-Taweel, L. Mu, A new upwind weak Galerkin finite element method for linear hyperbolic equations, *J. Comput. Appl. Math.*, **390** (2021), 113376. https://doi.org/10.1016/j.cam.2020.113376

25. H. Roos, T. Linß, Sufficient conditions for uniform convergence on layer-adapted grids, *Computing*, **63** (1999), 27–45. https://doi.org/10.1007/s006070050049

26. Q. Lin, N. Yan, A. Zhou, A rectangle test for interpolated finite elements, *Proc. Syst. Sci. Eng.*, (1991), 217–229.

27. S. Franz, G. Matthies, A unified framework for time-dependent singularly perturbed problems with discontinuous Galerkin methods in time, *Math. Comp.*, **87** (2018), 2113–2132. https://doi.org/10.1090/mcom/3326

28. Y. Cheng, Y. Mei, H. Roos, The local discontinuous Galerkin method on layer-adapted meshes for time-dependent singularly perturbed convection-diffusion problems, *Comput. Math. Appl.*, **117** (2022), 245–256. https://doi.org/10.1016/j.camwa.2022.05.004

29. D. A. Di Pietro, A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, Springer Berlin, 2011. https://doi.org/10.1007/978-3-642-22980-0

30. M. Stynes, L. Tobiska, Using rectangular Q_p elements in the SDFEM for a convection-diffusion problem with a boundary layer, *Appl. Numer. Math.*, **58** (2008), 1789–1802. https://doi.org/10.1016/j.apnum.2007.11.004

31. N. Kumar, S. Toprakseven, N. S. Yadav, J. Yuan, A Crank-Nicolson WG-FEM for unsteady 2D convection-diffusion equation with nonlinear reaction term on layer adapted mesh, *Appl. Numer. Math.*, **201** (2024), 322–346. https://doi.org/10.1016/j.apnum.2024.03.013

AIMS Press