INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 20, Number 2, Pages 199–228

doi: 10.4208/ijnam2023-1009

WEAK GALERKIN FINITE ELEMENT METHODS FOR PARABOLIC PROBLEMS WITH L^2 INITIAL DATA

NARESH KUMAR AND BHUPEN DEKA

Abstract. We analyze the weak Galerkin finite element methods for second-order linear parabolic problems with L^2 initial data, both in a spatially semidiscrete case and in a fully discrete case based on the backward Euler method. We have established optimal L^2 error estimates of order $O(h^2/t)$ for semisdiscrete scheme. Subsequently, the results are extended for fully discrete scheme. The error analysis has been carried out on polygonal meshes for discontinuous piecewise polynomials in finite element partitions. Finally, numerical experiments confirm our theoretical convergence results and efficiency of the scheme.

Key words. Parabolic equations, weak Galerkin method, non-smooth data, polygonal mesh, optimal L^2 error estimates.

1. Introduction

There are various applications of parabolic partial differential equations (PDEs) with non-smooth data arising in sciences and engineering such as chemical diffusion, heat conduction processes, thermodynamics, and medical science [6,7,14,16]. Classical finite element methods for parabolic problems with non-smooth initial data have been studied broadly so far, an extensive literature for the same can be obtained from [19,23–25,31]. Developing numerical algorithms of parabolic equations with non-smooth initial data has been a flourishing concern. Recently, the weak Galerkin finite element method has attracted much attention in the field of numerical PDEs. As referred to in [28], the weak Galerkin finite element methods (WG-FEMs) have been established as a new finite element technique for solving PDEs, which are derived from weak formulations of problems to replace the classical differential operators (e.g., gradient, divergence, curl) by weak differential operators which is approximated in suitable polynomial spaces and adding the stabilizer term. There is no need to select the parameters of the stabilizer broadly. More precisely, the WG-FEMs have a simple and parameter-free formulation and the flexibility of using general polygonal meshes. With the new concepts of weak function and weak gradient, the WG-FEMs allow discontinuous function space as the approximation space on each element. Unlike the classical finite element method, the WG-FEM is applicable for unstructured polygonal meshes making it more suitable for complex geometry that usually appears in real-life problems. The WG approach has been developed for various types of PDEs in existing literature, such as elliptic equations [18,21,26,29], parabolic equations [9,10,17,33,34], and the hyperbolic problems [15,32]. The hybrid high-order (HHO) method is closely

Received by the editors on March 19, 2022; accepted on October 29, 2022.
2020 *Mathematics Subject Classification.* 65N15, 65N30.

related to WG finite element method as the reconstruction operator in the HHO method corresponds to the weak gradient in WG methods [5,11]. The only difference between HHO and WG methods lies in the choice of the discrete unknowns and the stabilization pattern. However, the links between HHO and WG methods are not fully explored yet; nevertheless, they share something in their roots. (cf. [3,8,11]). It is noteworthy that WG and HHO are based on different devising viewpoints and use somewhat different analysis techniques.

We know that the higher order of convergence of finite element approximations depends on the higher smoothness of the true solutions, which demands higher regularity of the initial functions. The main concern of this work is to study the convergence of weak Galerkin finite element approximations for homogeneous equations with non-smooth initial data using polygonal meshes. The error analysis is highly motivated by the fact that the solutions of parabolic problems have the so-called smoothing property (cf. [19]). The solution is smooth for positive time t, even when the initial data are not H^1 regular. Under the low regularity of solutions, convergence analysis has remained a significant part of mathematical study up to the present day. To derive optimal $\mathrm{O}(h^{r+1})$ $(r \geq 1)$ in the L^2 norm for WG-FEM, the minimum regularity assumption on the exact solution u should be $u \in H^1(0,T;H^{r+1}(\Omega))$ (for instance, see [17,34,35]). More recently, in [9], the authors have shown the convergence of WG finite element solution to the true solution at an optimal rate in $L^2(L^2)$ norm under the assumption that $u \in L^2(0,T;H^{r+1}(\Omega)) \cap H^1(0,T;H^{r-1}(\Omega))$. In the case of piecewise linear WG-FEM (i.e., $r=1$), the optimal error estimate requires the initial value to be in H^1 (see, Theorem 3.2 in [9]) and for L^2 initial data error analysis in [9] leads to sub-optimal order of convergence in $L^2(L^2)$ norm (see, Remark 3.2). In fact, optimal $L^\infty(L^2)$ error estimate in [10] for linear weak Galerkin elements demands initial data $u^0 \in H^3(\Omega)$ (see, Remark 3.4). In this work, assuming initial data in L^2, we have shown the convergence of WG finite element solution to the true solution at an optimal rate in L^2 norm on WG finite element space $(\mathcal{P}_1,\ \mathcal{P}_1,\ \mathcal{P}_0^2)$ (see, Theorem 3.2 and Theorem 4.1). The non-smooth data error analysis heavily depends on the newly derived optimal L^2 norm error estimates with smooth initial data $u^0 \in H_0^1 \cap H^2$ (see, Lemma 3.8 and Lemma 4.1). The obtained results intend to enhance the numerical analysis of linear parabolic equations on polygonal meshes with non-smooth initial data. To the best of our knowledge, the smoothing property of the WG-FEM and HHO methods for the parabolic equation has not been studied earlier.

The rest of this work is organized as follows. In Sec. 2, we have introduced some commonly used notations and reviewed the weak Galerkin discretization. Sec. 3 is concerned with the error analysis of the semidiscrete WG finite element algorithm. In Sec. 4, the backward Euler scheme is proposed, and optimal a priori error bounds in $L^\infty(L^2)$ norm is established. Sec. 5 discusses several numerical examples which demonstrate the robustness of

the WG-FEMs. Finally, in Sec. 6, results are summarized with some future extensions of this work.

2. Preliminaries and weak Galerkin discretization

2.1. Basic notation and model problem. This paper uses the standard notation for Sobolev spaces and norms [1]. Denoted $H^m(\mathcal{K})$ the typical Sobolev space for a domain $\mathcal{K} \subseteq \Omega \subset \mathbb{R}^2$ and non-negative integer m. Notations $\|\cdot\|_{m,\mathcal{K}}$ and $|\cdot|_{m,\mathcal{K}}$ are used to denote the norm and semi-norm in the Sobolev space $H^m(\mathcal{K})$, respectively. The space $L^2(\mathcal{K})$ are used instead of $H^0(\mathcal{K})$, whose inner product and norm are denoted by $(\cdot,\cdot)_{\mathcal{K}}$ and $\|\cdot\|_{\mathcal{K}}$, respectively. For simplicity of notation, we skip the subscript $\mathcal{K}$ in the norm and inner product notation when $\mathcal{K} = \Omega$. $H_0^1(\Omega)$ is a closed subspace of $H^1(\Omega)$, which is also closure of $C_0^\infty(\Omega)$ (the set of all C^∞ functions with compact support) with respect to the norm of $H^m(\Omega)$. In addition, the standard Bôchner spaces $L^2(J;\mathcal{B})$ and $L^\infty(J;\mathcal{B})$, where $\mathcal{B}$ is a real Banach space with norm $\|.\|_{\mathcal{B}}$ and $J = [0,T]$, consisting of all measurable functions $\phi : J \to \mathcal{B}$ for which,

$$\begin{aligned} \|\phi\|_{L^2(J;\mathcal{B})} &:= \Big(\int_0^T \|\phi(t)\|_{\mathcal{B}}^2 dt\Big)^{\frac{1}{2}} < \infty \text{ and} \\ \|\phi\|_{L^\infty(J;\mathcal{B})} &:= \operatorname*{ess\,sup}_{t\in[0,T]} \|\phi(t)\|_{\mathcal{B}} < \infty, \end{aligned}$$

respectively. We denote $H^1(J;\mathcal{B})$ as the space of all measurable functions $\phi : J \to \mathcal{B}$ such that ϕ', the derivative of ϕ with respect to time variable exists and belongs to $L^2(J;\mathcal{B})$, endowed with the norm

$$\|\phi\|_{H^1(J;\mathcal{B})} = \left(\|\phi\|_{L^2(J;\mathcal{B})}^2 + \|\phi'\|_{L^2(J;\mathcal{B})}^2\right)^{\frac{1}{2}}.$$

For our convenience, we use $L^2(\mathcal{B})$ for $L^2(J;\mathcal{B})$, $L^\infty(\mathcal{B})$ for $L^\infty(J;\mathcal{B})$ and $H^1(\mathcal{B})$ for $H^1(J;\mathcal{B})$.

Further, $H^{-m}(\Omega)$ denotes the space of all bounded linear functionals on $H_0^m(\Omega)$. For a functional $f \in H^{-m}(\Omega)$, its action on a function $\phi \in H_0^m(\Omega)$ is denoted by (f,ϕ), which represents the duality pairing between $H^{-m}(\Omega)$ and $H_0^m(\Omega)$. The negative Sobolev norm is defined as

$$\|f\|_{-m} = \sup_{0\neq\phi\in H_0^m(\Omega)} \frac{(f,\phi)}{\|\phi\|_m}. \tag{1}$$

Throughout our work, C denotes a positive generic constant and independent of the mesh parameters $\{h,\tau\}$.

In this article, we consider a second-order linear parabolic equation of the form

$$u_t - \nabla\cdot(\alpha\nabla u) = f \quad \text{in } \Omega\times(0,T],\ \ T<\infty, \tag{2}$$

with initial and boundary conditions

$$u(x,0) = u^0 \ \text{ in } \ \Omega, \quad u(x,t) = 0 \ \text{ on } \ \partial\Omega\times(0,T], \tag{3}$$

where Ω is a convex polygonal domain in $\mathbb{R}^2$ with boundary $\partial\Omega$. We assume that the coefficient matrix $\alpha = (\alpha_{ij}(x))_{2\times 2} \in \left[L^\infty(\Omega)\right]^{2\times 2}$ is symmetric, uniformly positive definite in Ω. Further, $f : \Omega \times (0,T] \to \mathbb{R}$ denotes the source function and $u^0 \in L^2(\Omega)$ is the initial function.

The standard weak formulation of the problem (2)-(3) is stated as follows: Find $u : [0,\infty) \to H_0^1(\Omega)$ such that

$$\begin{aligned} (u_t, v) + \mathcal{A}(u,v) &= (f,v)\ \forall v \in H_0^1(\Omega) \text{ and} \\ u(0) &= u^0 \text{ in } \Omega, \end{aligned} \tag{4}$$

where $\mathcal{A}(\cdot,\cdot)$ is a bilinear form on $H_0^1(\Omega) \times H_0^1(\Omega)$ and given by

$$\mathcal{A}(u,v) = \int_\Omega \alpha \nabla u \cdot \nabla v dx.$$

We end this section with the following regularity result for the initial boundary value problem (2)-(3) (see, [13], p. 363, Theorem 3).

Theorem 2.1. *Assume that $f \in L^2(\Omega)$ and $u^0 \in L^2(\Omega)$. Then the solution of (2)-(3) satisfies*

$$u \in L^2(0,T; H^2(\Omega)) \cap H^1(0,T; H^{-1}(\Omega)).$$

2.2. Weak Galerkin discretization. This section describes the weak Galerkin finite element discretization for the problem (2)-(3) and reviews the definition of weak gradient operator.

For some $h_0 > 0$ and $h \in (0, h_0]$, let $\mathcal{T}_h$ be a partition of the domain Ω consisting of polygons in two dimension satisfying a set of conditions specified in [29]. Denote by $\mathcal{F}_h$ the set of all edges in $\mathcal{T}_h$ and let $\mathcal{F}_h^0 = \mathcal{F}_h \backslash \partial\Omega$ be the set of all interior edges. For every element $K \in \mathcal{T}_h$, we denote by $|K|$ the measure of K and by h_K its diameter, and mesh size $h = \max_{K\in\mathcal{T}_h} h_K$ for $\mathcal{T}_h$.

Let K be any polygonal domain with interior K^0 and boundary ∂K. A weak function on the region K refers to a pair of scalar-valued functions $v = \{v_0, v_b\}$ such that $v_0 \in L^2(K)$ and $v_b \in L^2(\partial K)$. Here, v_b may not necessarily be related to the trace of v_0 on ∂K. Weak gradient operators and their discrete versions are introduced in [28, 29].

Denoted by $\mathcal{V}(K)$ the space of weak scalar-valued functions on K, i.e.,

$$\mathcal{V}(K) = \{v = \{v_0, v_b\} : v_0 \in L^2(K), v_b \in L^2(\partial K)\}. \tag{5}$$

Denoted by $\mathcal{V}$ the weak scalar-valued function space on $\mathcal{T}_h$ given by

$$\mathcal{V} = \prod_{K\in\mathcal{T}_h} \mathcal{V}(K). \tag{6}$$

Let $V \subset \mathcal{V}$ be a subspace of $\mathcal{V}$ consisting of weak valued functions which are continuous across each interior edge, i.e.,

$$V = \{v \in \mathcal{V} :\ [v]_e = 0 \quad \forall e \in \mathcal{F}_h^0\}. \tag{7}$$

Here, $[v]_e$ denotes the jump of $v \in \mathcal{V}$ across an interior edge $e \in \mathcal{F}_h^0$.

For $k \geq 1$, now, we define weak Galerkin finite element space $V_h \subset V$ as

$$V_h = \{v_h = \{v_0, v_b\} : \ v_0|_{K^0} \in \mathcal{P}_k(K), \ v_b|_e \in \mathcal{P}_k(e), \ e \in \partial K, \ K \in \mathcal{T}_h\} \tag{8}$$

and

$$V_h^0 = \{v_h = \{v_0, v_b\} \in V_h : \ v_b = 0 \ \text{ on } \partial\Omega\}. \tag{9}$$

For each $v_h = \{v_0, v_b\} \in V_h$, the weak gradient of it, denoted by ∇_w, is defined as the unique polynomial $(\nabla_w v_h) \in [\mathcal{P}_{k-1}(K)]^2$ that satisfies the following equation

$$(\nabla_w v_h, \mathbf{q})_K = -\int_K v_0(\nabla \cdot \mathbf{q})dK + \int_{\partial K} v_b(\mathbf{q} \cdot \mathbf{n})ds \ \ \forall \mathbf{q} \in [\mathcal{P}_{k-1}(K)]^2. \tag{10}$$

The usual L^2-inner product can be written locally on each element as follows

$$(\nabla_w v_h, \nabla_w w_h) = \sum_{K \in \mathcal{T}_h} (\nabla_w v_h, \nabla_w w_h)_K. \tag{11}$$

For each element $K \in \mathcal{T}_h$, Q_0 and Q_b are the usual L^2 projection operators onto $\mathcal{P}_k(K)$ and $\mathcal{P}_k(\partial K)$, respectively. We shall combine Q_0 with Q_b by writing $Q_h = \{Q_0, Q_b\}$. More precisely, for $w \in H^1(K)$, we write $Q_h w = \{Q_0 w, Q_b w\}$.

We recall the following crucial approximation properties for local projections Q_0. For details, we refer to (Lemma 4.1, [29]).

Lemma 2.1. *Let $\mathcal{T}_h$ be a finite element partition of Ω satisfying the shape regularity assumption as specified in [29]. Then, for $v \in H^{k+1}(\Omega)$, we have*

$$\sum_{K \in \mathcal{T}_h} \left(\|v - Q_0 v\|_K^2 + h_K^2 \|\nabla(v - Q_0 v)\|_K^2 \right) \leq C h^{2(k+1)} \|v\|_{k+1}^2.$$

To introduce WG finite element approximation, we introduce a bilinear map $\mathcal{A}_w(\cdot, \cdot)$ on $V_h \times V_h$, which is defined as follows

$$\mathcal{A}_w(w_h, v_h) := \sum_{K \in \mathcal{T}_h} \left(\alpha \nabla_w w_h, \nabla_w v_h \right)_K + \mathcal{S}(w_h, v_h) \ \ \forall w_h, v_h \in V_h, \tag{12}$$

where $\mathcal{S}(\cdot, \cdot)$ is known as stabilizer, which is a semi-positive definite bilinear form defined on $V_h \times V_h$. Stabilizer $\mathcal{S}(\cdot, \cdot)$ is often chosen in such a way that it fits well into the theory and implementation of the WG numerical scheme [27]. Here, for $w_h = \{w_0, w_b\}$, $v_h = \{v_0, v_b\} \in V_h$, the stabilizer $\mathcal{S}(\cdot, \cdot)$ is defined as

$$\mathcal{S}(w_h, v_h) = \sum_{K \in \mathcal{T}_h} h_K^{-1} \langle w_0 - w_b, v_0 - v_b \rangle_{\partial K}. \tag{13}$$

Here, $\langle \cdot, \cdot \rangle_{\partial K}$ denotes the L^2 inner product on ∂K and symbolically, we have

$$\langle \cdot, \cdot \rangle_{\partial K} = \sum_{e \in \partial K} \langle \cdot, \cdot \rangle_e,$$

where $\langle \cdot, \cdot \rangle_e$ denotes the L^2 inner product on $e \in \mathcal{F}_h$.

The weak finite element space V_h^0 is a normed linear space with a triple-bar norm given by (cf. [27])

$$|||v_h|||^2 = \sum_{K\in\mathcal{T}_h} \|\alpha^{\frac{1}{2}}\nabla_w v_h\|_K^2 + \sum_{K\in\mathcal{T}_h} h_K^{-1}\|v_0 - v_b\|_{\partial K}^2 = \mathcal{A}_w(v_h, v_h). \tag{14}$$

Let K be an element with e as an edge. For any function $\varphi \in H^1(K)$, the following trace inequality holds true (see, [29] for details)

$$\|\varphi\|_e^2 \le C(h_K^{-1}\|\varphi\|_K^2 + h_K\|\nabla\varphi\|_K^2). \tag{15}$$

It is easy to verify the equivalency of the triple-bar norm (14) and following discrete H^1 norm (cf. [27])

$$\|v_h\|_{1,h} = \Big(\sum_{K\in\mathcal{T}_h} (\|\nabla v_0\|_K^2 + h_K^{-1}\|v_0 - v_b\|_{\partial K}^2)\Big)^{\frac{1}{2}},\ v_h = \{v_0, v_b\} \in V_h^0. \tag{16}$$

More precisely, there exist constants $C_1,\ C_2 > 0$ such that for any $v_h \in V_h^0$, the following inequality holds true

$$C_1\|v_h\|_{1,h}^2 \le \mathcal{A}_w(v_h, v_h) \le C_2\|v_h\|_{1,h}^2. \tag{17}$$

Moreover, the following Poincaŕe-type inequality holds true (cf. [22])

$$\|v_0\| \le C|||v_h|||,\ v_h = \{v_0, v_b\} \in V_h^0. \tag{18}$$

3. Error analysis for the semidiscrete scheme

In this section, we have obtained the optimal order of convergence with L^2 initial data for the spatially discrete scheme in L^2 norm.

A time-dependent weak function $v_h : [0, T] \to V_h$ is written as $v_h(t) := \{v_0(t), v_b(t)\}$ and subsequently we define $v_{ht}(t) := \{v_0'(t), v_b'(t)\}$, where '$\prime$' denotes the time derivatives. For simplicity, we use $v_h = \{v_0, v_b\}$ for $v_h(t)$ and $v_{ht} = \{v_0', v_b'\}$ for $v_{ht}(t)$. It is easy to note from the definition of weak gradient (10) that $(\nabla_w v_h)' = \nabla_w v_h'$ and $(\nabla_w v_h)|_{t=0} = \nabla_w v_h(0)$ for all $v_h \in V_h$.

The semidiscrete weak Galerkin finite element approximation for (2)-(3) is to find $u_h(t) = \{u_0(t), u_b(t)\} \in V_h^0$ satisfying $u_h(0) = \{Q_0u^0, Q_b(Q_0u^0)\}$ such that

$$(u_{ht}, v_0) + \mathcal{A}_w(u_h, v_h) = (f, v_0) \quad \forall v_h = \{v_0, v_b\} \in V_h^0,\ t > 0, \tag{19}$$

where the bilinear map $\mathcal{A}_w(\cdot,\cdot)$ is as defined in (12).

Remark 3.1. *For $u^0 \in L^2(\Omega)$, restriction of u^0 on the boundary of a element $K \in \mathcal{T}_h$ is not defined. Therefore, we first project u^0 in the interior K^0 via the L^2 projection and then extend it to the boundary through the L^2 projection Q_b. Now, the initial approximation*

$$u_h(0) = \{Q_0u^0, Q_b(Q_0u^0)\}$$

is well defined. Next, we have to make sure that $Q_b(Q_0u^0)$ takes unique value on interior edges. For an interior edge $e \in \mathcal{F}_h$ shared by the two elements

T_1 and T_2, we define $Q_b(Q_0u^0)$ in the following way:

$$Q_b(Q_0u^0) = \begin{cases} Q_b(Q_0u^0|_{T_1\cap e}) \ \text{on} & T_1 \cap e, \\ Q_b(Q_0u^0|_{T_2\cap e}) + Q_b(\left[Q_0u^0\right]_e) \ \text{on} & T_2 \cap e, \end{cases} \tag{20}$$

where $\left[Q_0u^0\right]_e$ denotes the jump of Q_0u^0 along the edge e.

The following result deals with the existence and uniqueness of the WG solution u_h.

Theorem 3.1. *For each $h \in (0, h_0]$, there exists a function $u_h \in C^1(0,T;V_h^0)$ satisfying (19).*

Proof. For a given element $K \in \{\mathcal{T}_h\}_{0<h\le h_0}$, let $\{\phi_{0,i} : \ i = 1,2,\dots,N_0\}$ be a set of basis functions for $\mathcal{P}_k(K)$ and $\{\phi_{b,i} : \ i = 1,2,\dots,N_b\}$ be a set of basis function for $\mathcal{P}_k(e)$. Then every $v_h = \{v_0, v_b\} \in \{V_h^0\}_{0<h\le h_0}$ can be written as

$$v_h|_K = \left\{ \sum_{i=1}^{N_0} d_{0,i}(t)\phi_{0,i}, \sum_{j=1}^{N_b} d_{b,j}(t)\phi_{b,j} \right\},$$

where $d_{0,i},\ d_{b,j} : (0,T] \to \mathbb{R}$ are the coefficient functions for $1 \le i \le N_0$ and $1 \le j \le N_b$. For $1 \le i \le N_0 + N_b$, we write $\hat{\phi}_{i,h} = \{\hat{\phi}_{0,i}, \hat{\phi}_{b,i}\}$ with

$$\hat{\phi}_{0,i} = \phi_{0,i} \ \text{ for } 1 \le i \le N_0 \quad \& \quad \hat{\phi}_{0,i} = 0 \ \text{ for } \ N_0 + 1 \le i \le N_0 + N_b,$$
$$\hat{\phi}_{b,i} = 0 \ \text{ for } \ 1 \le i \le N_0 \quad \& \quad \hat{\phi}_{b,i} = \phi_{b,i-N_0} \ \text{ for } \ N_0 + 1 \le i \le N_0 + N_b,$$

and similarly to capture the unknown coefficient functions, we define

$$\hat{d}_{i,h} = d_{0,i} \ \text{ for } 1 \le i \le N_0 \ \ \& \ \ \hat{d}_{i,h} = d_{b,i-N_0} \ \text{ for } N_0 + 1 \le i \le N_0 + N_b.$$

Then, we seek our semidiscrete solution $u_h = \{u_0, u_b\} \in V_h^0$ such that

$$u_h|_K = \sum_{i=1}^{N_0+N_b} \hat{d}_{i,h}(t)\hat{\phi}_{i,h} = \left\{ \sum_{i=1}^{N_0+N_b} \hat{d}_{i,h}(t)\hat{\phi}_{0,i}, \sum_{j=1}^{N_0+N_b} \hat{d}_{j,h}(t)\hat{\phi}_{b,j} \right\}, \ \ K \in \mathcal{T}_h.$$

Now, set $v_h = \hat{\phi}_{j,h},\ j = 1,2,\dots,N_0 + N_b$ in (19) to obtain

$$\left(\sum_{i=1}^{N_0+N_b} \hat{d}'_{i,h}(t)\hat{\phi}_{0,i}, \hat{\phi}_{0,j} \right) + \mathcal{A}_w \left(\sum_{i=1}^{N_0+N_b} \hat{d}_{i,h}(t)\hat{\phi}_{i,h}, \hat{\phi}_{j,h} \right)$$
$$= (f, \hat{\phi}_{0,j}), \ \ j = 1,\dots,N_0 + N_b.$$

We can rearrange the above equations as

$$\sum_{i=1}^{N_0+N_b} \hat{d}'_{i,h}(t)\big(\hat{\phi}_{0,i}, \hat{\phi}_{0,j}\big) + \sum_{i=1}^{N_0+N_b} \hat{d}_{i,h}(t)\mathcal{A}_w\big(\hat{\phi}_{i,h}, \hat{\phi}_{j,h}\big)$$
$$= (f, \hat{\phi}_{0,j}),\ j = 1,\dots,N_0 + N_b.$$

On each element K, the local stiffness matrix $\mathcal{A}_K$ associated with the bilinear map $\mathcal{A}_w(\cdot,\cdot)$ defined by (12) can thus be written as a block matrix

$$\mathcal{A}_K = \begin{bmatrix} \mathcal{A}_{0,0} & \mathcal{A}_{0,b} \\ \mathcal{A}_{b,0} & \mathcal{A}_{b,b} \end{bmatrix}, \tag{21}$$

where $\mathcal{A}_{0,0}$ is a $N_0 \times N_0$, $\mathcal{A}_{0,b}$ is a $N_0 \times N_b$, $\mathcal{A}_{b,0}$ is a $N_b \times N_0$, and $\mathcal{A}_{b,b}$ is a $N_b \times N_b$ matrices. More precisely, these matrices are given by

$$\mathcal{A}_{0,0} = [\mathcal{A}_w(\phi_{0,j}, \phi_{0,i})_K]_{i,j}, \ \mathcal{A}_{0,b} = [\mathcal{A}_w(\phi_{0,j}, \phi_{b,i})_K]_{i,j}$$
$$\mathcal{A}_{b,0} = [\mathcal{A}_w(\phi_{b,j}, \phi_{0,i})_K]_{i,j}, \ \mathcal{A}_{b,b} = [\mathcal{A}_w(\phi_{b,j}, \phi_{b,i})_K]_{i,j},$$

where i, j are the row and column indices, respectively.

We denote by

$$\hat{d}_{h,0} = [\hat{d}_{1,h}(0), \ldots, \hat{d}_{N_0+N_b,h}(0)]^T$$

the components of given initial approximation $u_h(0)$. Then, for our semidiscrete solution, we need to find unknown vector $\hat{d}_h(t) = [\hat{d}_{1,h}(t), \ldots, \hat{d}_{N_0+N_b,h}(t)]^T$ such that

$$\begin{cases} \mathcal{C}_K \hat{d}_h'(t) + \mathcal{A}_K \hat{d}_h(t) = F_K(t), \\ \hat{d}_h(0) = \hat{d}_{h,0} \ \text{ for } \ t \in (0, T], \end{cases} \tag{22}$$

where the coefficient matrix C_K is given by

$$\mathcal{C}_K = [\mathcal{C}_{j,i}], \ \ \mathcal{C}_{j,i} = (\hat{\phi}_{0,j}, \hat{\phi}_{0,i})_K, \ \ 1 \le i, j \le N_0 + N_b$$

and the source vector is given by

$$F_K = [F_1, \ldots, F_{N_0+N_b}], \ \ F_j = (f, \hat{\phi}_{0,j})_K, \ \ 1 \le i, j \le N_0 + N_b.$$

For all $t \in J = (0, T]$, it is easy to note that

$$|(\hat{\phi}_{0,i}, \hat{\phi}_{0,j})| \le \|\hat{\phi}_{0,i}\| \|\hat{\phi}_{0,j}\|, \ \ |(f, \hat{\phi}_{0,j})| \le \|f\| \|\hat{\phi}_{0,j}\| \ \ \&$$
$$|\mathcal{A}_w(\hat{\phi}_{i,h}, \hat{\phi}_{j,h})| \le C \|\hat{\phi}_{i,h}\|_{1,h} \|\hat{\phi}_{j,h}\|_{1,h}.$$

Furthermore, for any $v \in \mathbb{R}^{N_0+N_b} \setminus \{\mathbf{0}\}$, we have

$$v^T C_K v = (\hat{v}, \hat{v})_K > 0, \ \ \hat{v} = \sum_{i=1}^{N_0+N_b} v_i \hat{\phi}_{0,i}.$$

Hence, the matrix C_K is invertible for all $t \in J$ and the equation (22) can be restated as

$$\begin{cases} \hat{d}_h'(t) + \mathcal{C}_K^{-1} \mathcal{A}_K \hat{d}_h(t) = \mathcal{C}_K^{-1} F_K(t), \\ \hat{d}_h(0) = \hat{d}_{h,0}. \end{cases}$$

Now, the existence of the solution $u_h \in C^1(0, T; V_h^0)$ follows from the standard ODE theory. The proof is completed. □

We have adopted following few results from [19] to state some a priori bounds for the solution u satisfying (4) under appropriate regularity assumptions on the initial data u^0 and source function f.

Lemma 3.1. *Let u satisfies (2). If $u^0 \in L^2(\Omega)$ and $f \in L^2(\Omega)$, then*

$$\|u(t)\|^2 + \int_0^t \|u(s)\|_1^2 ds \le C\Big(\|u^0\|^2 + \int_0^t \|f(s)\|^2 ds\Big).$$

Moreover, when $u^0 \in H_0^1(\Omega)$ and $f \in L^2(\Omega)$, we have

$$\|u(t)\|_1^2 + \int_0^t \{\|u_s(s)\|^2 + \|u(s)\|_2^2\} ds \le C\Big(\|u^0\|_1^2 + \int_0^t \|f(s)\|^2 ds\Big).$$

Lemma 3.2. *Let u satisfies (2) with $f = 0$, and let $0 \le i, j, s \le 2$. If $0 \le s + 2j - i \le 2$, then*

$$t^i \left\| \frac{\partial^j u}{\partial t^j}(t) \right\|_s^2 \le C \|u^0\|_{s+2j-i}^2.$$

Further, if $0 \le s + 2j - i - 1 \le 2$, then

$$\int_0^t \vartheta^i \left\| \frac{\partial^j u}{\partial \vartheta^j}(\vartheta) \right\|_s^2 d\vartheta \le C \|u^0\|_{s+2j-i-1}^2.$$

For the error analysis, we split our semidiscrete error $e_h = u - u_h$ into two components using an intermediate operator. We write

$$e_h = u - u_h := (u - Q_h u) + (Q_h u - u_h).$$

The following estimate for e_h is derived over the time interval $[0, t]$, which is imperative for further analysis.

Lemma 3.3. *Let $u^0 \in H_0^1(\Omega) \cap H^r(\Omega)$, $r \in \{1, 2\}$ and $f = 0$, then*

$$\int_0^t \|e_h\|^2 ds \le C t h^{2r} \|u^0\|_r^2. \tag{23}$$

Proof. To derive L^2 error estimate over the time interval $[0, t]$, we recall the following estimate from earlier work (cf. [9], Theorem 3.2)

$$\int_0^t \|e_h\|^2 ds \le C h^{2r} \int_0^t \|u\|_r^2 ds. \tag{24}$$

Now, apply Lemma 3.2 with $i = 0 = j$ and $s = r$ in the above estimate to have

$$\begin{aligned} \int_0^t \|e_h\|^2 ds &\le C h^{2r} \int_0^t \|u^0\|_r^2 ds \\ &\le C t h^{2r} \|u^0\|_r^2. \ \square \end{aligned}$$

Remark 3.2. *For $u^0 \in L^2(\Omega)$ and $f = 0$, let u solves (2)-(3). Then, from the Lemma 3.1, we have $u \in L^2(0, T; H^1(\Omega))$. As a consequence, we obtain*

$$\int_0^t \|e_h\|^2 ds \le C h^2 \int_0^t \|u\|_1^2 ds. \tag{25}$$

Here, we have used (24) with $r = 1$. Finally, a priori estimate in Lemma 3.1 yields

$$\int_0^t \|e_h\|^2 ds \le C h^2 \|u^0\|^2. \tag{26}$$

Clearly, $L^2(L^2)$ error estimate (26) is not optimal for non-smooth initial data u^0. The rest of the section is devoted to deriving optimal $L^\infty(L^2)$ norm estimate for the WG finite element approximation with $u^0 \in L^2(\Omega)$.

In the following Lemmas, we have proved basic stability results for the weak Galerkin solution u_h satisfying (19).

Lemma 3.4. *Assume that u_h is a solution of (19). Let $f \in L^2(\Omega)$ and $u^0 \in L^2(\Omega)$ with initial approximation $u_h(0) = (Q_0u^0, Q_b(Q_0u^0))$, then*

$$\|u_h(t)\|^2 + \int_0^t \|u_h\|_{1,h}^2 \leq C\Big(\|u_h(0)\|^2 + \int_0^t \|f(s)\|^2 ds\Big). \tag{27}$$

Proof. Taking $v_h = u_h$ in (19), we get

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

Now, integrate the above equation with respect to t to obtain

$$\frac{1}{2}\|u_h(t)\|^2 + \int_0^t \mathcal{A}_w(u_h, u_h)ds = \frac{1}{2}\|u_h(0)\|^2 + \int_0^t (f, u_h)ds$$

Then, Young's inequality and (17) leads to

$$\begin{aligned}\frac{1}{2}\|u_h(t)\|^2 + \int_0^t \|u_h\|_{1,h}^2 ds \quad &\leq \quad \frac{1}{2}\|u_h(0)\|^2 + C_\nu \int_0^t \|f(s)\|^2 ds \\ &\quad + C(\nu)\int_0^t \|u_h(s)\|^2 ds,\end{aligned}$$

for some $\nu > 0$. Finally, suitable $\nu > 0$ and (18) yield desired result. □

Lemma 3.5. *Assume that u_h is a solution of (19). Let $f = 0$ and $u^0 \in L^2(\Omega)$ with initial approximation $u_h(0) = (Q_0u^0, Q_b(Q_0u^0))$, then*

$$t\|u_h(t)\|_{1,h}^2 + \int_0^t s\|u_{hs}\|^2 ds \leq C\|u_h(0)\|^2. \tag{28}$$

Proof. Setting $v_h = u_{ht}$ in (19), we get

$$\|u_{ht}(t)\|^2 + \frac{1}{2}\frac{d}{dt}\mathcal{A}_w(u_h, u_h) = 0.$$

Multiplying the above equation by t, we obtain

$$t\|u_{ht}(t)\|^2 + \frac{1}{2}\frac{d}{dt}\{t\mathcal{A}_w(u_h(t), u_h(t))\} = \frac{1}{2}\mathcal{A}_w(u_h, u_h).$$

Then, integrate the above equation over $[0, t]$ to get

$$t\|u_h(t)\|_{1,h}^2 + \int_0^t s\|u_{hs}(s)\|^2 ds \leq C\int_0^t \|u_h(s)\|_{1,h}^2 ds,$$

which together with estimate (27) yields (28). □

Lemma 3.6. *Assume that u_h is a solution of (19). Let $f = 0$ and $u^0 \in H_0^1(\Omega) \cap H^2(\Omega)$ with suitable initial approximation $u_h(0) \in V_h^0$, then*

$$\|u_{ht}\|^2 + \int_0^t \|u_{hs}\|_{1,h}^2 ds \leq C\|u_{ht}(0)\|^2. \tag{29}$$

Further, if $u_0 \in L^2(\Omega)$ and $f = 0$ with $u_h(0) = (Q_0u^0, Q_b(Q_0u^0))$, then

$$t^2\|u_{ht}(t)\|^2 + \int_0^t s^2\|u_{hs}\|_{1,h}^2 ds \leq C\|u_h(0)\|^2. \tag{30}$$

Proof. We differentiate (19) with respect to time t to obtain

$$(u_{htt}, v_h) + \mathcal{A}_w(u_{ht}, v_h) = 0.$$

Now, set $v_h = u_{ht}$ in the above equation to have

$$\frac{1}{2}\frac{d}{dt}\|u_{ht}\|^2 + \mathcal{A}_w(u_{ht}, u_{ht}) = 0. \tag{31}$$

Integrating above equation from 0 to t and applying (17), we arrive at (29).

In the case $u^0 \in L^2(\Omega)$, like earlier, we multiply (31) by t^2 and obtain

$$\frac{1}{2}\frac{d}{dt}\{t^2\|u_{ht}\|^2\} + t^2\mathcal{A}_w(u_{ht}, u_{ht}) = t\|u_{ht}\|^2,$$

which upon integrating over $[0, t]$ leads to

$$\frac{t^2}{2}\|u_{ht}\|^2 + \int_0^t s^2\mathcal{A}_w(u_{hs}, u_{hs})ds = \int_0^t s\|u_{hs}\|^2 ds.$$

Now, apply the coercive property (17) to obtain

$$t^2\|u_{ht}\|^2 + \int_0^t s^2\|u_{hs}\|_{1,h}^2 ds \le C\int_0^t s\|u_{hs}\|^2 ds. \tag{32}$$

Finally, estimates (28) and (32) lead to desire result. □

Next, we applied elliptic projection to derive optimal L^2 error estimates. For $w \in H^2(\Omega) \cap H_0^1(\Omega)$, we define

$$f_w = -\nabla \cdot (\alpha \nabla w) \text{ in } \Omega.$$

Now, define $\mathcal{E}_h : H^2(\Omega) \cap H_0^1(\Omega) \to V_h^0$ by

$$\mathcal{A}_w(\mathcal{E}_h w, v_h) = (f_w, v_0) \ \forall v_h = \{v_0, v_b\} \in V_h^0. \tag{33}$$

It is easy to observe from the definition of elliptic projection and equation (19) that

$$(u_{ht}, v_h) + \mathcal{A}_w(u_h - \mathcal{E}_h u, v_h) = (f, v_h) + (\nabla \cdot (a\nabla u), v_h) = (u_t, v_h), \tag{34}$$

for all $v_h = \{v_0, v_b\} \in V_h^0$. Here, we have used equation (2).

Remark 3.3. *For $u^0 \in H_0^1(\Omega) \cap H^2(\Omega)$ and $f = 0$ with $u_h(0) = \mathcal{E}_h u^0$, identity (34) yields*

$$(u_{ht}(0), v_h) = \ (u_t(0), v_h) \ \forall v_h = \{v_0, v_b\} \in \mathcal{V}_h^0,$$

which implies

$$\|u_{ht}(0)\| \le \|u_t(0)\| \le C\|u^0\|_2. \tag{35}$$

Hence, a priori estimate (29) reduces to

$$\|u_{ht}\|^2 + \int_0^t \|u_{hs}\|_{1,h}^2 ds \le C\|u^0\|_2^2. \tag{36}$$

We can derive the following lemma from the literature (see proof of Theorem 3.1.2 in [10] and [27]) directly. We omit the details.

Lemma 3.7. *For $v \in H^1(0,T;H^{\lambda+1}(\Omega)\cap H_0^1(\Omega))$, $0 \le \lambda \le k$, we have*

$$\begin{aligned}\|Q_h v - \mathcal{E}_h v\| + h\|Q_h v - \mathcal{E}_h v\|_{1,h} &\le Ch^{\lambda+1}\|v\|_{\lambda+1} \quad \text{and}\\ \int_0^t s^2\Big(\|(v-Q_hv)_t\|^2 + \|\big(Q_hv-\mathcal{E}_hv\big)_t\|^2\Big)ds &\le Ch^{2(\lambda+1)}\int_0^t s^2\|v_t\|^2_{\lambda+1}ds.\end{aligned}$$

Corollary 3.1. *For $u^0 \in H_0^1(\Omega)\cap H^2(\Omega)$ and $f = 0$, Lemma 2.1, Lemma 3.2 and Lemma 3.7 yield*

$$\|u - Q_h u\| + \|Q_h u - \mathcal{E}_h u\| \le Ch^2\|u\|_2 \le Ch^2\|u^0\|_2$$

and

$$\begin{aligned}&\int_0^t s^2\Big(\|(u-Q_hu)_t\|^2 + \|\big(Q_hu-\mathcal{E}_hu\big)_t\|^2\Big)ds\\ \le\ & Ch^4\int_0^t s^2\|u_t\|_2^2 ds \le Ch^4 t\|u^0\|_2^2.\end{aligned}$$

Remark 3.4. *Theorem 3.1.2 in [10] and Lemma 2.1, for the weak Galerkin space $(\mathcal{P}_1(K),\ \mathcal{P}_1(\partial K),\ \big[\mathcal{P}_0(K)\big]^2)$, yields*

$$\|(u-u_h)(t)\|^2 \le Ch^4\left(\|u\|_2^2 + \int_0^t \|u_t\|_2^2 d\vartheta\right), \tag{37}$$

which demands $u \in H^1(0,T;H^2(\Omega))$. Again, due to Theorem 2.10 in [20], solution $u \in H^1(0,T;H^2(\Omega))$ provided initial data $u^0 \in H^3(\Omega)$. Thus, for $u^0 \in H_0^1(\Omega)\cap H^2(\Omega)$, we can not directly use estimate (37).

Next result illustrates optimal $L^\infty(L^2)$ norm error estimate for the semidiscrete approximation on linear weak Galerkin space $(\mathcal{P}_1(K),\ \mathcal{P}_1(\partial K),\ \big[\mathcal{P}_0(K)\big]^2)$ with $u^0 \in H_0^1(\Omega)\cap H^2(\Omega)$.

Lemma 3.8. *If $u^0 \in H_0^1(\Omega)\cap H^2(\Omega)$ and $f = 0$, then*

$$\|e_h(t)\| \le Ch^2\|u^0\|_2. \tag{38}$$

Proof. First, we split our error into three components and we write

$$e_h = u - u_h = (u - Q_h u) + (Q_h u - \mathcal{E}_h u) + (\mathcal{E}_h u - u_h) := \rho + \eta + \xi, \tag{39}$$

where $\rho := u - Q_h u$, $\eta := Q_h u - \mathcal{E}_h u$ and $\xi := \mathcal{E}_h u - u_h$.

According to the definition of projection $\mathcal{E}_h u$, we have

$$\begin{aligned}&(\xi_t, v_h) + \mathcal{A}_w(\xi, v_h) = (\mathcal{E}_h u_t, v_h) + \mathcal{A}_w(\mathcal{E}_h u, v_h) - (u_{ht}, v_h) - \mathcal{A}_w(u_h, v_h)\\ &= -(\eta_t, v_h) - (\rho_t, v_h).\end{aligned}$$

Choose $v_h = \xi$ to have

$$\frac{1}{2}\frac{d}{dt}\|\xi\|^2 + \mathcal{A}_w(\xi,\xi) = -(\eta_t,\xi) - (\rho_t,\xi). \tag{40}$$

Multiplying (40) by t, we get

$$\frac{1}{2}\frac{d}{dt}\{t\|\xi\|^2\} + t\mathcal{A}_w(\xi,\xi) = -t(\eta_t,\xi) - t(\rho_t,\xi) + \frac{1}{2}\|\xi\|^2. \tag{41}$$

Integration of (41) over $[0,t]$ yields

$$\frac{t}{2}\|\xi\|^2+\int_0^t s\mathcal{A}_w(\xi,\xi)ds=-\int_0^t s(\eta_s,\xi)ds-\int_0^t s(\rho_s,\xi)ds+\frac{1}{2}\int_0^t\|\xi\|^2ds.$$

Next, standard inequality and positivity of $\mathcal{A}_w(\cdot,\cdot)$ lead to

$$\frac{t}{2}\|\xi\|^2\le C\int_0^t s^2\{\|\eta_s\|^2+\|\rho_s\|^2\}ds+C\int_0^t\|\xi\|^2ds. \tag{42}$$

Now, from (39) and estimate (42), we have

$$\begin{aligned} t\|e_h\|^2 &\le C\big(t\|\rho\|^2+t\|\eta\|^2+t\|\xi\|^2\big)\\ &\le Ct(\|\rho\|^2+\|\eta\|^2)+C\int_0^t\|\xi\|^2ds+C\int_0^t s^2(\|\eta_t\|^2+\|\rho_t\|^2)ds.\\ &\le Ct(\|\rho\|^2+\|\eta\|^2)+C\int_0^t\big(\|\eta\|^2+s^2\|\eta_t\|^2\big)ds\\ &\quad +C\int_0^t\big(\|\rho\|^2+s^2\|\rho_t\|^2\big)ds+C\int_0^t\|e_h\|^2ds. \end{aligned} \tag{43}$$

Finally, Lemma 3.3 and Corollary 3.1 yield

$$t\|e_h(t)\|^2\le Cth^4\|u^0\|_2^2.\quad\square$$

Next, we proceed with the error analysis with non-smooth initial data.

Lemma 3.9. *If $u^0\in L^2$ and $f=0$. Then, for $t>0$, we have*

$$\|e_h(t)\|\le C\frac{h}{t^{1/2}}\|u^0\|. \tag{44}$$

Proof. We first recall the estimate (43)

$$\begin{aligned} t\|e_h\|^2 &\le Ct(\|\rho\|^2+t\|\eta\|^2)+C\int_0^t\big(\|\eta\|^2+s^2\|\eta_t\|^2\big)ds\\ &\quad +C\int_0^t\big(\|\rho\|^2+s^2\|\rho_t\|^2\big)ds+C\int_0^t\|e_h\|^2ds. \end{aligned} \tag{45}$$

Then apply Lemma 3.7 with $\lambda=0$ to arrive at

$$t\|e_h\|^2\le Ch^2t\|u\|_1^2+Ch^2\int_0^t\big(\|u\|_1^2+s^2\|u_t\|_1^2\big)ds+C\int_0^t\|e_h\|^2ds.$$

Then, using Lemmas 3.1-3.2 and estimate (26), we obtain

$$t\|e_h\|^2\le Ch^2\|u^0\|^2.$$

This completes the rest of the proof. $\square$

We now use a dual parabolic argument considering a backward problem to obtain optimal $L^\infty(L^2)$ norm error estimate. For fixed $t>0$ and $g\in L^2(\Omega)$, define $w:[0,t)\to H_0^1(\Omega)$ by

$$\begin{aligned} (\phi,w_s)-\mathcal{A}(\phi,w) &= 0\ \forall\phi\in H_0^1(\Omega),\ s\le t,\\ w(t) &= g, \end{aligned} \tag{46}$$

and define $w_h = \{w_0, w_b\} : [0,t) \to V_h^0$ by

$$\begin{aligned} (v_0, w_{hs}) - \mathcal{A}_w(v_h, w_h) &= 0 \ \ \forall v_h = \{v_0, v_b\} \in V_h^0, \ \ s \le t, \\ w_h(t) &= g_h. \end{aligned} \tag{47}$$

Then, by construction, we have

$$\frac{d}{ds}\{(u,w) - (u_h, w_h)\} = (f, w - w_h). \tag{48}$$

Integrating above from 0 to t and setting $\tilde{e}_h := w - w_h$, we obtain

$$\begin{aligned} (u(t), w(t)) - (u_h(t), w_h(t)) &= (u^0, w(0)) - (u_h(0), w_h(0)) \\ &\quad + \int_0^t (f, \tilde{e}_h) ds. \end{aligned} \tag{49}$$

Further, if $u_h(0) = (Q_0 u^0, Q_b(Q_0 u^0)$ and $g_h = (Q_0 g, Q_b(Q_0 g))$, we arrive at following crucial identity

$$(e_h(t), g) = (u^0, \tilde{e}_h(0)) + \int_0^t (f, \tilde{e}_h) ds. \tag{50}$$

Lemma 3.10. *If $u^0 \in L^2$ and $f = 0$, then*

$$\|e_h(t)\|_{-2} \le Ch^2 \|u^0\|. \tag{51}$$

Proof. From the equation (50) with $f = 0$, we have

$$(e_h(t), g) = (u^0, \tilde{e}_h(0)).$$

Thus, applying estimate (38) to the backward error $\tilde{e}_h(t)$ yields

$$|(e_h(t), g)| \le Ch^2 \|u^0\| \|g\|_2, \ \ g \in H_0^1(\Omega) \cap H^2(\Omega).$$

Then, definition (1) with $m = 2$ leads to desire estimate. □

Now, we are in a position to discuss the main result of this section.

Theorem 3.2. *Let u be the solution of (2)-(3) and $u_h \in V_h^0$ be the solution of (19). Assume that $u^0 \in L^2(\Omega)$ and $f = 0$, then there is a positive constant C independent of h such that*

$$\|e_h(t)\| \le C \frac{h^2}{t} \|u^0\|, \ \ t > 0. \tag{52}$$

Proof. Integrating the equation (48) from $t/2$ to t and using the fact $g_h = (Q_0 g, Q_b(Q_0 g))$, we get

$$\begin{aligned} (e_h(t), g) &= \left(u\left(\tfrac{t}{2}\right), w\left(\tfrac{t}{2}\right)\right) - \left(u_h\left(\tfrac{t}{2}\right), w_h\left(\tfrac{t}{2}\right)\right) \\ &= \left(e_h\left(\tfrac{t}{2}\right), w\left(\tfrac{t}{2}\right)\right) - \left(e_h\left(\tfrac{t}{2}\right), \tilde{e}_h\left(\tfrac{t}{2}\right)\right) \\ &\quad + \left(u\left(\tfrac{t}{2}\right), \tilde{e}_h\left(\tfrac{t}{2}\right)\right) \\ &\le \left\|e_h\left(\tfrac{t}{2}\right)\right\|_{-2} \left\|w\left(\tfrac{t}{2}\right)\right\|_2 + \left\|e_h\left(\tfrac{t}{2}\right)\right\| \left\|\tilde{e}_h\left(\tfrac{t}{2}\right)\right\| \\ &\quad + \left\|u\left(\tfrac{t}{2}\right)\right\|_2 \left\|\tilde{e}_h\left(\tfrac{t}{2}\right)\right\|_{-2}. \end{aligned} \tag{53}$$

Applying estimate (51) for the terms $\|e_h(\frac{t}{2})\|_{-2}$ and $\|\tilde{e}_h(\frac{t}{2})\|_{-2}$, we obtain

$$\left\|e_h\left(\frac{t}{2}\right)\right\|_{-2} \leq Ch^2\|u^0\| \quad \& \quad \left\|\tilde{e}_h\left(\frac{t}{2}\right)\right\|_{-2} \leq Ch^2\|g\|. \tag{54}$$

Further, using estimate (44) for the terms $\left\|e_h\left(\frac{t}{2}\right)\right\|$ and $\|\tilde{e}_h\left(\frac{t}{2}\right)\|$, we get

$$\left\|e_h\left(\frac{t}{2}\right)\right\| \leq C\frac{h}{t^{1/2}}\|u^0\| \ \& \ \left\|\tilde{e}_h\left(\frac{t}{2}\right)\right\| \leq C\frac{h}{t^{1/2}}\|g\|. \tag{55}$$

Again, setting $i = 2,\ s = 2$ and $j = 0$ in Lemma 3.2, we arrive at

$$\left\|u\left(\frac{t}{2}\right)\right\|_2 \leq \frac{C}{t}\|u_0\| \ \text{ and } \ \left\|w\left(\frac{t}{2}\right)\right\|_2 \leq \frac{C}{t}\|g\|. \tag{56}$$

Finally, $g = e_h(t)$ and estimates (53)-(56) lead to desired result. □

4. Fully discrete scheme

In this section, we have extended the classical finite element error analysis technique [19] to WG-FEMs with L^2 initial data for a first-order backward time fully discrete scheme. It is worth noting that the algorithms presented in [19] are only valid for finite linear elements with triangular meshes. Optimal order error estimates in $L^\infty(L^2)$ norm is shown to hold even if the initial data is in $L^2(\Omega)$.

First, we divide the time interval $J = [0, T]$ into M equally spaced subintervals $I_n = (t_{n-1}, t_n]$, $n = 1, 2, \ldots, M$ with $t_0 = 0$, and $t_M = T$ and $\tau = t_n - t_{n-1}$, the time step. For a continuous mapping $\chi : [0, T] \to L^2(\Omega)$, we define $\chi^n = \chi(\cdot, t_n)$. Then, for a sequence $\{p^n\}_{n=0}^M \subset L^2(\Omega)$, we define

$$\partial_\tau p^n = \frac{p^n - p^{n-1}}{\tau}, \quad n \geq 1.$$

We now introduce the fully discrete weak Galerkin finite element approximation to the problem (2)-(3). Let $U_h^0 = (Q_0u^0, Q_b(Q_0u^0))$ and $U_h^n = \{U_0^n, U_b^n\} \in V_h^0$ be the fully discrete solution of u at $t = t_n$, which we have defined through the following scheme

$$(\partial_\tau U_h^n, v_0) + \mathcal{A}_w(U_h^n, v_h) = (\overline{f}^{\,n}, v_0) \ \ \forall v_h = \{v_0, v_b\} \in V_h^0, \ \ n \geq 1, \tag{57}$$

where

$$\overline{f}^{\,n} = \tau^{-1}\int_{t_{n-1}}^{t_n} f(x, t)dt. \tag{58}$$

We now present the main result of this section in the following theorem.

Theorem 4.1. *Suppose u and U_h^n be the solution of (2)-(3) and (1), respectively. In addition, assume that initial data $u^0 \in L^2(\Omega)$ and $f = 0$. Then, we have*

$$\|U_h^n - u^n\| \leq C\frac{(h^2 + \tau)}{t_n}\|u^0\|. \tag{59}$$

To prove the above theorem, the following preparations are required. Similar to (1), we have used the following discrete negative norm

$$\|\chi\|_{-1,h} := \sup_{0\neq v_h\in V_h^0} \frac{(\chi, v_h)}{\|v_h\|_{1,h}}.$$

Remark 4.1. *By differentiating (19) with respect to t for $f = 0$, we obtain*

$$(u_{htt}, v_h) + \mathcal{A}_w(u_{ht}, v_h) = 0 \ \ \forall v_h = \{v_0, v_b\} \in V_h^0.$$

Now, apply continuity and coercivity properties of $A_w(\cdot,\cdot)$ to have

$$(u_{htt}, v_h) \leq C\|u_{ht}\|_{1,h}\|v_h\|_{1,h} \ \ \forall v_h = \{v_0, v_b\} \in V_h^0,$$

which yields

$$\sup_{0\neq v_h\in V_h^0} \frac{(u_{htt}, v_h)}{\|v_h\|_{1,h}} \leq C\|u_{ht}\|_{1,h},$$

so that we obtain

$$\|v_{htt}\|_{-1,h} \leq C\|u_{ht}\|_{1,h}. \tag{60}$$

Now, for the fully discrete error analysis, we split our error into two components using a semidiscrete solution as follows

$$U_h^n - u^n := (U_h^n - u_h^n) + (u_h^n - u^n). \tag{61}$$

The rest of the section analyzes error bounds for the first component in (5).

Lemma 4.1. *If $f = 0$ and $u^0 \in H_0^1(\Omega) \cap H^2(\Omega)$, then we have*

$$\|u_h^n - U_h^n\| \leq C\tau\|u_{ht}(0)\|.$$

Proof. Setting $t = t_{n+1}$ in (19), we obtain

$$(\partial_\tau u_h^{n+1}, v_0) + \mathcal{A}_w(u_h^{n+1}, v_h) = (\zeta^{n+1}, v_0) \ \ \forall v_h = \{v_0, v_b\} \in V_h^0,$$

where $\zeta^{n+1} := \partial_\tau u_h^{n+1} - u_{ht}^{n+1}$.

Let $\xi^n = u_h^n - U_h^n \in V_h^0$, then it is easy to note that

$$(\partial_\tau \xi^{n+1}, v_0) + \mathcal{A}_w(\xi^{n+1}, v_h) = (\zeta^{n+1}, v_0) \ \ \forall v_h = \{v_0, v_b\} \in V_h^0, \tag{62}$$

with $\xi^0 = 0$.

Set $v_h = \xi^{n+1}$ in (6) and apply Young's inequality with suitable $\nu > 0$ in the right hand side to obtain

$$\begin{aligned}
&\|\xi^{n+1}\|^2 + 2\tau\mathcal{A}_w(\xi^{n+1}, \xi^{n+1}) + \|\xi^{n+1} - \xi^n\|^2 \\
&= \|\xi^n\|^2 + 2\tau(\zeta^{n+1}, \xi^{n+1}) \\
&\leq \|\xi^n\|^2 + C_\nu\tau\|\zeta^{n+1}\|_{-1,h}^2 + C(\nu)\tau\|\xi^{n+1}\|_{1,h}^2 \\
&\leq \|\xi^n\|^2 + C_\nu\tau\|\zeta^{n+1}\|_{-1,h}^2 + C(\nu)\tau\mathcal{A}_w(\xi^{n+1}, \xi^{n+1}). \qquad (63)
\end{aligned}$$

Note that, in the above estimate, we have used the following fact

$$(\partial_\tau \xi^{n+1}, \xi^{n+1}) = \frac{1}{2\tau}(\|\xi^{n+1}\|^2 - \|\xi^n\|^2) + \frac{1}{2\tau}\|\xi^{n+1} - \xi^n\|^2.$$

Then, we sum (7) over n from 1 to l so that

$$\|\xi^{l+1}\|^2 \le C\tau \sum_{m=1}^{l+1} \|\zeta^m\|^2_{-1,h}, \quad 1 \le l \le M-1. \tag{64}$$

Now, by Taylor's theorem

$$\zeta^m = \frac{1}{\tau}\int_{t_{m-1}}^{t_m} (s - t_{m-1}) u_{hss}(s) ds. \tag{65}$$

Hence

$$\|\zeta^m\|^2_{-1,h} \le \frac{\tau}{3}\int_{t_{m-1}}^{t_m} \|u_{hss}(s)\|^2_{-1,h} ds. \tag{66}$$

Thus, combining (4) and (8), we get

$$\|\xi^{l+1}\|^2 \le C\tau^2 \int_0^{t_{l+1}} \|u_{hs}(s)\|^2_{1,h} ds. \tag{67}$$

Now, apply (29) leads us to

$$\|\xi^{l+1}\|^2 \quad \le \quad C\tau^2 \|u_{ht}(0)\|^2. \ \square$$

Let $\{F^j\}_{j=1}^M \subset V_h^0$, and $\{z_h^m\}_{m=1}^M \subset V_h^0$ be the solution of the discrete time backward problem

$$(v_0, \partial_\tau z_h^m) - \mathcal{A}_w(v_h, z_h^{m-1}) = (v_0, F^m) \ \ \forall v_h = \{v_0, v_b\} \in V_h^0, \tag{68}$$

with $z_h^M = 0$.

Regarding the stability of z_h^m, we have the following result.

Lemma 4.2. *For z_h^m satisfying (12), we obtain*

$$\max_{1\le m\le M} \|z_h^{m-1}\|^2_{1,h} + \sum_{m=1}^{M} \tau \|\partial_\tau z_h^m\|^2 \le C \sum_{m=1}^{M} \tau \|F^m\|^2.$$

Proof. Setting $v_h = \tau \partial_\tau z_h^m$ in (12) and applying Cauchy-Schwarz inequality, we have

$$\tau \|\partial_\tau z_h^m\|^2 + \mathcal{A}_w(z_h^{m-1} - z_h^m, z_h^{m-1}) \le \|\tau \partial_\tau z_h^m\| \|F^m\|.$$

Next, applying Young's inequality for suitable $\nu > 0$, we arrive at

$$\tau \|\partial_\tau z_h^m\|^2 + \mathcal{A}_w(z_h^{m-1} - z_h^m, z_h^{m-1}) \le C_\nu \tau \|\partial_\tau z_h^m\|^2 + C(\nu)\tau \|F^m\|^2.$$

Now, selecting $\nu > 0$ appropriately leads us

$$\tau \|\partial_\tau z_h^m\|^2 + \mathcal{A}_w(z_h^{m-1} - z_h^m, z_h^{m-1}) \le C\tau \|F^m\|^2. \tag{69}$$

It is easy to verify that

$$\begin{aligned} \mathcal{A}_w(z_h^{m-1} - z_h^m, z_h^{m-1}) \quad &= \quad \frac{\tau^2}{2}\mathcal{A}_w(\partial_\tau z_h^m, \partial_\tau z_h^m) \\ &\quad -\frac{1}{2}\mathcal{A}_w(z_h^m, z_h^m) + \frac{1}{2}\mathcal{A}_w(z_h^{m-1}, z_h^{m-1}). \end{aligned}$$

This, combined with (13) yields

$$\tau\|\partial_\tau z_h^m\|^2 - \frac{1}{2}\mathcal{A}_w(z_h^m, z_h^m) + \frac{1}{2}\mathcal{A}_w(z_h^{m-1}, z_h^{m-1}) \le C\tau\|F^m\|^2,$$

and then summing over m from $m = l$ to $m = M$ yields

$$\sum_{m=l}^{M} \tau\|\partial_\tau z_h^m\|^2 + \mathcal{A}_w(z_h^{l-1}, z_h^{l-1}) \le C\sum_{m=l}^{M}\tau\|F^m\|^2, \quad 1 \le l \le M-1.$$

This completes the rest of the proof. □

Lemma 4.3. *If $f = 0$ and $u^0 \in L^2(\Omega)$, then we have*

$$\|u_h^n - U_h^n\| \le C\frac{\tau^{1/2}}{t_n^{1/2}}\|u_h(0)\|.$$

Proof. We multiply (7) by $t_{n+1} = t_n + \tau$ to obtain

$$t_{n+1}\|\xi^{n+1}\|^2 \le t_n\|\xi^n\|^2 + \tau\|\xi^n\|^2 + C\tau t_{n+1}\|\zeta^{n+1}\|_{-1,h}^2. \tag{70}$$

Sum (14) over n from 1 to l with $1 \le l \le M-1$ to have

$$\begin{aligned} t_{l+1}\|\xi^{l+1}\|^2 &\le \sum_{n=1}^{l}\tau\|\xi^n\|^2 + C\tau\sum_{n=1}^{l} t_{n+1}\|\zeta^{n+1}\|_{-1,h}^2 \\ &= \sum_{n=1}^{l}\tau\|\xi^n\|^2 + C\tau^2\sum_{n=1}^{l}\|\zeta^{n+1}\|_{-1,h}^2 + C\tau\sum_{n=1}^{l} t_n\|\zeta^{n+1}\|_{-1,h}^2 \\ &= \sum_{n=1}^{l}\tau\|\xi^n\|^2 + C\tau^2\sum_{n=1}^{l}\|\zeta^{n+1}\|_{-1,h}^2 \\ &\quad + C\sum_{m=2}^{l+1}\frac{t_{m-1}}{m-1}t_{m-1}\|\zeta^m\|_{-1,h}^2. \end{aligned} \tag{71}$$

In order to estimate $\sum_{m=1}^{l}\tau\|\xi^m\|^2$, let $\{z_h^m\}_{m=0}^M \subset V_h^0$ be the solution of (12) with $F^m = \xi^m$. Set $v_h = \xi^m$ in (12) to obtain

$$\begin{aligned} \|\xi^m\|^2 &= (\xi^m, \partial_\tau z_h^m) - \mathcal{A}_w(\xi^m, z_h^{m-1}) \\ &= \partial_\tau(\xi, z_h)^m - (\partial_\tau\xi^m, z_h^{m-1}) - \mathcal{A}_w(\xi^m, z_h^{m-1}) \\ &= \partial_\tau(\xi, z_h)^m - (\zeta^m, z_h^{m-1}). \end{aligned} \tag{72}$$

Here, we have used (6). Since $z_h^M = 0$, we obtain by summing (16)

$$\sum_{m=1}^{M}\tau\|\xi^m\|^2 = -\sum_{m=1}^{M}\tau(\zeta^m, z_h^{m-1}).$$

Now, apply Cauchy-Schwarz inequality and then Young's inequality to have

$$\begin{aligned}\sum_{m=1}^{M}\tau\|\xi^m\|^2 &\le \sum_{m=1}^{M}\tau\|\zeta^m\|_{-1,h}\|z_h^{m-1}\|_{1,h}\\ &\le \max_{1\le m\le M}\|z_h^{m-1}\|_{1,h}\sum_{m=1}^{M}\tau\|\zeta^m\|_{-1,h}.\\ &\le C_\nu\max_{1\le m\le M}\|z_h^{m-1}\|_{1,h}^2+C(\nu)\Big(\sum_{m=1}^{M}\tau\|\zeta^m\|_{-1,h}\Big)^2,\end{aligned}$$

where ν is a suitable positive real number. It then follows from Lemma 4.2 with $F^m=\xi^m$ that

$$\begin{aligned}\sum_{m=1}^{M}\tau\|\xi^m\|^2 &\le C\Big(\sum_{m=1}^{M}\tau\|\zeta^m\|_{-1,h}\Big)^2\\ &\le C\tau^2\|\zeta^1\|_{-1,h}^2+C\Big(\sum_{m=2}^{M}\frac{\tau^2}{t_{m-1}^2}\Big)\Big(\sum_{m=2}^{M}t_{m-1}^2\|\zeta^m\|_{-1,h}^2\Big)\\ &\le C\tau^2\|\zeta^1\|_{-1,h}^2+C\sum_{m=2}^{M}t_{m-1}^2\|\zeta^m\|_{-1,h}^2.\end{aligned}$$

In the last term, we have used that

$$\sum_{m=2}^{M}\frac{\tau^2}{t_{m-1}^2}=\sum_{m=2}^{M}\frac{\tau^2}{\tau^2(m-1)^2}=\sum_{m=2}^{M}\frac{1}{(m-1)^2}\le\frac{\pi^2}{6}.$$

Hence, it follows from (15) that

$$t_{l+1}\|\xi^{l+1}\|^2\le C\tau^2\|\zeta^1\|_{-1,h}^2+C\sum_{m=2}^{M}t_{m-1}^2\|\zeta^m\|_{-1,h}^2, \tag{73}$$

for $1\le l\le M-1$. Now from (9) to get

$$\begin{aligned}\tau^2\|\zeta^1\|_{-1,h}^2 &= \Big\|\int_0^\tau su_{hss}(s)ds\Big\|_{-1,h}^2\\ &\le \tau\int_0^\tau s^2\|u_{hss}(s)\|_{-1,h}^2ds\\ &\le C\tau\int_0^\tau s^2\|u_{hs}(s)\|_{1,h}^2ds\\ &\le C\tau\|u_h(0)\|^2. \qquad (74)\end{aligned}$$

Here, we have used estimate (30) and estimate (4). Also, from (10),

$$\begin{aligned} t_{m-1}^2\|\zeta^m\|_{-1,h}^2 &\leq \frac{\tau}{3}t_{m-1}^2\int_{t_{m-1}}^{t_m}\|u_{hss}(s)\|_{-1,h}^2 ds \\ &\leq \frac{\tau}{3}\int_{t_{m-1}}^{t_m} s^2\|u_{hss}(s)\|_{-1,h}^2 ds \\ &\leq C\tau\int_{t_{m-1}}^{t_m} s^2\|u_{hs}(s)\|_{1,h}^2 ds. \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{m=2}^{M} t_{m-1}^2\|\zeta^m\|_{-1,h}^2 &\leq C\tau\int_0^{t_M} s^2\|u_{hs}(s)\|_{1,h}^2 ds \\ &\leq C\tau\|u_h(0)\|^2. \end{aligned} \tag{75}$$

Now, combining estimates (17)-(19), we obtain

$$t_{l+1}\|\xi^{l+1}\|^2 \leq C\tau\|u_h(0)\|^2, \quad 1\leq l\leq M-1, \tag{76}$$

, which completes the rest of the proof. □

Like earlier, we use the parabolic duality argument for the optimal L^2 norm error estimate with non-smooth data. For any fixed $t_n > 0$ ($1 \leq n \leq M$) and any function $\psi_h \in V_h^0$, define $w_h(s) \in V_h^0$ to be the continuous-time weak Galerkin solution of the backward problem

$$(v_0, w_{hs}) - \mathcal{A}_w(v_h, w_h) = 0 \quad \forall v_h = \{v_0, v_b\} \in V_h^0, \quad s \leq t_n, \tag{77}$$

with $w_h(t_n) = \psi_h$.

Let $\{W_h^m\}_{m=0}^n \subset V_h^0$ be the solution of discrete time backward problem

$$(v_0, \partial_\tau W_h^m) - \mathcal{A}_w(v_h, W_h^{m-1}) = 0 \ \forall v_h = \{v_0, v_b\} \in V_h^0, \ 1 \leq m \leq n, \tag{78}$$

with $W_h^n = \psi_h$.

Now, from the discrete analog of (48), we have

$$\partial_\tau\{(u_h, w_h) - (U_h, W_h)\}^m = 0. \tag{79}$$

Sum (23) over m from $1 \leq m \leq n$, we obtain

$$(\xi^n, \psi_h) = (u_h(0), w_h(0) - W_h^0). \tag{80}$$

Also, by Lemma 4.1 applied to $w_h(t_m) - W_h^m$ with time reversed, we get

$$(\xi^n, \psi_h) \leq \|u_h(0)\|\|w_h(0) - W_h^0\| \leq C\tau\|u_h(0)\|\|w_{ht}(t_n)\|. \tag{81}$$

Lemma 4.4. *If $f = 0$ and $u^0 \in L^2(\Omega)$, then*

$$\|u_h^n - U_h^n\| \leq \frac{C\tau}{t_n}\|u_h(0)\|. \tag{82}$$

Proof. Summing (23) over $m = q+1, \ldots, n$ with $q = [n/2]$ and setting $\psi_h = \xi^n$, we obtain

$$\begin{aligned} \|\xi^n\|^2 &= (\xi^q, w_h^q) + (U_h^q, (w_h - W_h)^q) \\ &= -(\xi^q, (w_h - W_h)^q) + (\xi^q, w_h^q) + (u_h^q, (w_h - W_h)^q). \end{aligned}$$

Next, apply the Cauchy-Schwarz inequality to obtain

$$\|\xi^n\|^2 \le \|\xi^q\|\|(w_h - W_h)^q\| + \|\xi^q\|\|w_h^q\| + \|u_h^q\|\|(w_h - W_h)^q\|. \tag{83}$$

Now by Lemma 4.3, we obtain

$$\|(w_h - W_h)^q\| \le \frac{C\tau^{1/2}}{t_q^{1/2}}\|w_h(t_n)\|.$$

In the above relation, multiplying and dividing by $(t_n - t_q)^{1/2}$, we get

$$\|(w_h - W_h)^q\| \le \frac{C\tau^{1/2}}{(t_n - t_q)^{1/2}}\|w_h(t_n)\|.$$

Here, we have used the fact that

$$\frac{(t_n - t_q)^{1/2}}{t_q^{1/2}} = \Big(\frac{t_n}{t_q} - 1\Big)^{1/2} = (\frac{n}{[n/2]} - 1)^{1/2} \le \sqrt{2}.$$

Now, we multiply and divide $t_n^{1/2}$, we have

$$\|(w_h - W_h)^q\| \le \frac{Ct_n^{1/2}\tau^{1/2}}{t_n^{1/2}\Big(t_n - t_q\Big)^{1/2}}\|w_h(t_n)\| \le \frac{C\tau^{1/2}}{t_n^{1/2}}\|\xi^n\|. \tag{84}$$

Here, we have used the fact that

$$\frac{t_n^{1/2}}{(t_n - t_q)^{1/2}} = \Big(\frac{n}{n - [n/2]}\Big)^{1/2} \le \sqrt{2}.$$

It follows by (25) with time-reversed and (30) that

$$(\xi^q, w_h^q) \le C\tau\|u_h(0)\|\|w_{ht}(t_q)\| \le \frac{C\tau}{t_n}\|u_h(0)\|\|\xi^n\|, \tag{85}$$

and

$$(u_h^q, (w_h - W_h)^q) \le C\tau\|w_h(t_n)\|\|u_{ht}(t_q)\| \le \frac{C\tau}{t_n}\|u_h(0)\|\|\xi^n\|. \tag{86}$$

Now, estimates (28)-(30) together with (27) leads to

$$\|\xi^n\| \le \frac{C\tau^{1/2}}{t_n^{1/2}}\|\xi^q\| + \frac{C\tau}{t_n}\|u_h(0)\| \le \frac{C\tau}{t_n}\|u_h(0)\|. \ \square \tag{87}$$

5. Numerical Section

In this section, we have tested various numerical examples for the parabolic problem (2)-(3) in $\Omega \times J$, where $\Omega \subset \mathbb{R}^2$ and $J = (0, T]$. Finite element partitions with different configurations like triangular, rectangular, and polygonal meshes are used for solving parabolic problems to confirm the flexibility of the WG method. These numerical results demonstrate that the scheme is robust and accurate. All computations are carried out using MATLAB software.

For a given finite number of successive iterations (indexed by i), let e_i be the error corresponding to the $L^\infty(L^2)$-norm on the i-th iteration, and h_i

is the corresponding mesh size. Then expected order of convergence (EOC) can be defined by

$$EOC(e_i) = \log\Big(\frac{e_{i+1}}{e_i}\Big)/\log\Big(\frac{h_{i+1}}{h_i}\Big).$$

Let U_h^n be the weak Galerkin solution defined by (1). Then, we have calculated the following error

$$e_h^n := Q_h u(x, t_n) - U_h^n$$

at final time $t_n = T$ with respect to L^2-norm for the linear WG space of the form $(\mathcal{P}_1(K),\ \mathcal{P}_1(\partial K),\ [\mathcal{P}_0(K)]^2)$.

Example 5.1. Non-smooth data with triangular mesh: This example is derived from (cf. Exercise 3.7, page 171, [2]). Consider the problem (2)-(3) in $\Omega \times J$, where $\Omega = (0, 2) \times (0, 2)$. The exact solution of the given problem is defined as

$$u(x,y,t) = \frac{200}{\pi^2}\sum_{i,j=1}^{\infty}\frac{1}{ij}\{1+(-1)^{i+1}\}\{1-\cos(\frac{j\pi}{2})\}\exp\{-\pi^2 t(\frac{i^2+j^2}{36})\} \times \sin(\frac{i\pi x}{2})\sin(\frac{j\pi y}{2}),$$

with initial data

$$u^0 = \begin{cases} 50 \ \text{ if } \ y \le 1, \\ 0 \ \text{otherwise.} \end{cases}$$

The right-hand side f can be evaluated from the exact solution u and the diffusion-coefficient

$$\alpha = \begin{bmatrix} 1/9 & 0 \\ 0 & 1/9 \end{bmatrix}.$$

The triangular mesh is used in this example, depicted in Figure 5.1. The domain is uniformly partitioned into $n \times n$ sub rectangles in such a way that each rectangular element is further divided into two triangles using a diagonal line with a negative slope, where the mesh size is $h = 1/n$. The errors with respect to $L^\infty(L^2)$ norm for linear WG space is reported in Table 5.1 at final time $T = 1$.

Example 5.2. Non-smooth data with polygonal mesh: We consider a second-order parabolic equation on a two-dimensional domain $\Omega \times J$, where $\Omega = (0, 1) \times (0, 1)$ for which the exact solution possesses non-smooth initial data. The example discussed here is extracted from [4].

$$\begin{aligned} u_t - \frac{1}{12}\Delta u &= 0 \ \text{ in } \ \Omega \times J, \\ u &= 0 \ \text{ on } \ \partial\Omega \times J, \\ u(x,y,0) &= u^0 \ \text{ in } \ \Omega. \end{aligned} \tag{88}$$

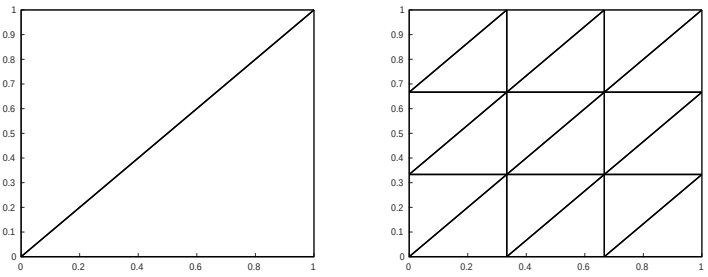

FIGURE 5.1. An initial triangular mesh with $h = 1/2$ (left), and its refinement with $h = 1/8$ (right).

We select the data appearing in (1) by setting the exact solution,

$$u(x,y,t) = \frac{8}{\pi^2}\sum_{i,j=0}^{\infty} c_i c_j \exp\{-\pi^2 t\frac{(2i+1)^2+(2j+1)^2}{12}\} \times \sin(\pi x(2i+1))\sin(\pi y(2j+1)),$$

where

$$c_i = \begin{cases} (-1)^{(i/2)}(2i+1)^{-1} \text{ if } i \text{ is even,} \\ (-1)^{(i+1)/2}(2i+1)^{-1} \text{ otherwise,} \end{cases}$$

with initial data

$$u^0 = \begin{cases} 1 \text{ if } 1/4 \le x, y \le 3/4, \\ 0 \text{ otherwise.} \end{cases}$$

In this example, we have used polygonal meshes. A typical polygonal mesh and its refinement are depicted in Figure 5.2. The errors with respect to $L^\infty(L^2)$-norm for linear WG space is represented in Table 5.1 at final time $T = 1$. We have achieved optimal order of convergence in $L^\infty(L^2)$-norm, which confirms the theoretical prediction as proved in Theorem 4.1. This can be observed from Table 5.1.

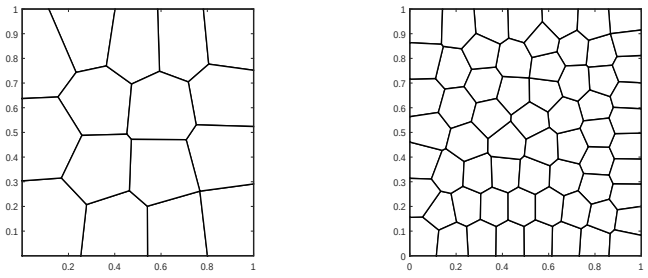

FIGURE 5.2. An initial polygonal mesh (left) and its refinement (right).

TABLE 5.1. The history of $L^\infty(L^2)$ error convergence with time step $\tau = h^2$.

	Example 5.1 (Triangular mesh)		*Example 5.2 (Polygonal mesh)*	
h	$\|e_h^n\|$	*EOC*	$\|e_h^n\|$	*EOC*
1/4	*7.127644e-01*	-	*3.303126e-03*	-
1/8	*1.762327e-01*	*2.015943e+00*	*8.109071e-04*	*2.026224e+00*
1/16	*4.407500e-02*	*1.999450e+00*	*2.021031e-04*	*2.004445e+00*
1/32	*1.102763e-02*	*1.998837e+00*	*5.049565e-05*	*2.000861e+00*
1/64	*2.757641e-03*	*1.999617e+00*	*1.262222e-05*	*2.000194e+00*
1/128	*6.894597e-04*	*1.999896e+00*	*3.155453e-06*	*2.000047e+00*

Example 5.3. Non-smooth data with rectangular mesh: In the following example, we choose $\Omega = (0,1) \times (0,1)$. We select the data appearing in (2)-(3) by setting the exact solution, which we have taken from [2] as

$$u(x,y,t) = \frac{400}{\pi^2} \sum_{i,j=1}^{\infty} \frac{1}{ij} \{1 - \cos(\frac{i\pi}{2})\}\{1 - \cos(\frac{j\pi}{2})\} \exp(-\pi^2 t(\frac{i^2+j^2}{12})) \times \sin(i\pi x) \sin(j\pi y),$$

with initial data

$$u^0 = \begin{cases} 100 \text{ if } x, y \leq 1/2, \\ 0 \text{ otherwise,} \end{cases}$$

and diffusion-coefficient

$$\alpha = \begin{bmatrix} 1/12 & 0 \\ 0 & 1/12 \end{bmatrix}.$$

In this example, we have used rectangular mesh, as shown in Figure 5.3. The errors with respect to $L^\infty(L^2)$-norm for linear WG space is reported in Table 5.2 at final time $T = 1$.

Example 5.4. Comparison among smooth data and non-smooth data: We consider a two-dimensional heat transfer equation with homogeneous boundary conditions. The solution to the model problem represents the temperature distribution in a thin rectangular plate. The temperature distribution of the insulated edges and faces of the plate is kept at zero.

$$(\mathbf{P}^*) \begin{cases} u_t - \alpha \Delta u = 0 \ \text{ in } \ \Omega \times J, \\ u(x,0) = u^0, \ \text{ in } \Omega, \\ u(x,t) = 0 \text{ on } \partial\Omega \times J, \end{cases}$$

where α denotes the thermal diffusivity and $\Omega = (0,1) \times (0,1)$.

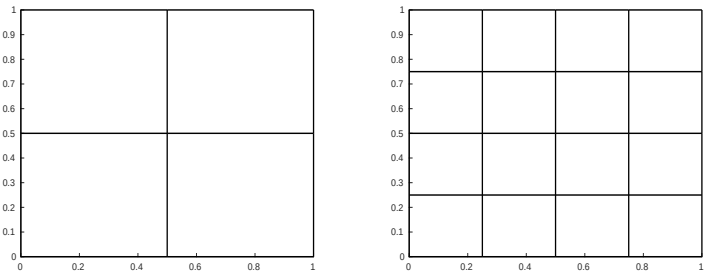

FIGURE 5.3. An initial rectangular mesh with $h = 1/2$ (left), and its refinement with $h = 1/4$ (right).

TABLE 5.2. $L^\infty(L^2)$ error convergence with time step $\tau = h^2$.

h	$\|e_h^n\|$	EOC
1/4	3.915363e+00	-
1/8	8.092984e-01	2.274402e+00
1/16	1.893695e-01	2.095468e+00
1/32	4.647979e-02	2.026528e+00
1/64	1.156529e-02	2.006802e+00
1/128	2.887895e-03	2.001711e+00

Here, we would check the behavior of the given solution with smooth and non-smooth data. For non-smooth data, we have opted the same exact solution as given in example 5.3, whereas, for the smooth data, we have taken the exact solution from (cf. Exercise 3.7, page 171, [2]) as

$$u(x,y,t) = \frac{1600}{\pi^2} \sum_{i,j=\text{odd}}^{\infty} \frac{\sin(i\pi x)\sin(j\pi y)}{ij} \times \exp(-\pi^2 t(\frac{i^2+j^2}{3}))$$

with initial data $u^0 = 100$, and the thermal diffusion-coefficient

$$\alpha = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix}.$$

The uniform triangular mesh is used as taken in example 5.1. The error with respect to $L^\infty(L^2)$ norm for linear WG space is represented in Table 5.3 at final time $T = 1$. We have obtained the optimal order of convergence in $L^\infty(L^2)$ norm as shown in Table 5.3. Figure 5.4 show the temperature distribution in the plat at $t = 1$ for smooth initial data. Further, Figure 5.5, and Figure 5.6 shows the temperature distribution in the plat at various time level $t = 0.01,\ 1,\ 5,\ 10$ with a fix time step $\tau = 0.002$, when initial data is non-smooth. It can be demonstrated from the Figures 5.5, 5.6 that

when time t increases, then solution decays to zero. Here, we can observe from Table 5.3 that when non-smooth initial data u^0 appeared in $(\mathbf{P}^*)$, its convergence behavior is the same as smooth initial data, and we have achieved an optimal rate of convergence in $L^\infty(L^2)$-norm.

TABLE 5.3. The history of $L^\infty(L^2)$ error convergence with time step $\tau = h^2$.

	Smooth data		Non-smooth data	
h	$\|e_h^n\|$	EOC	$\|e_h^n\|$	EOC
1/2	6.929241e-01	-	1.435194e+00	-
1/4	2.222611e-01	1.640442e+00	3.429445e-01	2.065199e+00
1/8	4.778263e-02	2.217698e+00	8.387519e-02	2.031659e+00
1/16	1.124189e-02	2.087601e+00	2.085701e-02	2.007712e+00
1/32	2.763636e-03	2.024246e+00	5.207374e-03	2.001905e+00
1/64	6.879407e-04	2.006212e+00	1.301416e-03	2.000474e+00
1/128	1.717990e-04	2.001562e+00	3.155453e-04	2.044163e+00

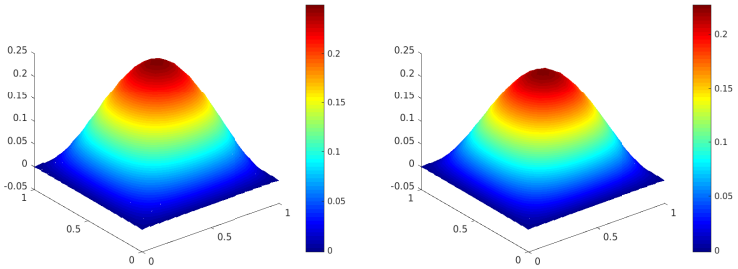

FIGURE 5.4. WG solution plot (left), exact solution plot (right) at $t = 1$ in example 5.4 for smooth initial data.

Example 5.5. Non-smooth data on rectangular mesh with Hanging nodes: In this example, we solve the same problem as in example 5.2 on rectangular mesh with hanging nodes in the finite element partition. The initial mesh is shown in Figure 5.7 (Left). A uniform refinement procedure generates the mesh on the right in Figure 5.7. It should be pointed out that the initial mesh has hanging nodes A, B, C, and D. For the finite element partition $\mathcal{T}_h$ with hanging nodes, we notice that the WG algorithm is still

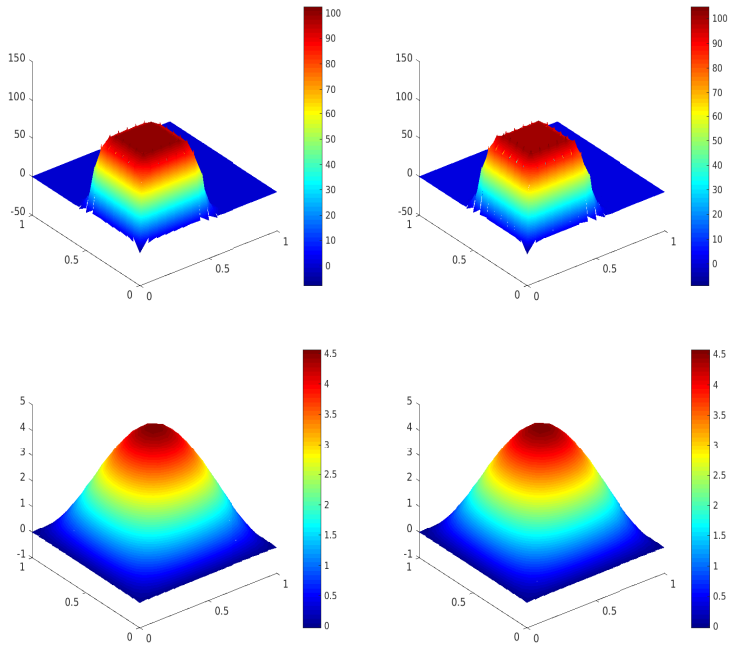

FIGURE 5.5. WG solution plot (top left), exact solution plot (top right) at $t = 0.01$ and WG solution plot (bottom left), exact solution plot (bottom right) at $t = 1$ in example 5.4 for non-smooth initial data

holding on to these refinements. The errors with respect to $L^\infty(L^2)$ norm for linear WG space is reported in Table 5.4 at final time $T = 1$.

TABLE 5.4. $L^\infty(L^2)$ error convergence with time step $\tau = h^2$.

h	$\|e_h^n\|$	EOC
1/4	*3.918887e-01*	-
1/8	*1.554516e-01*	*1.333979e+00*
1/16	*4.764673e-02*	*1.706016e+00*
1/32	*1.271481e-02*	*1.905868e+00*
1/64	*3.234341e-03*	*1.974966e+00*
1/128	*7.590261e-04*	*2.091250e+00*

6. Conclusion

In this work, we have developed a fully discrete scheme based on the weak Galerkin method in space and backward-Euler in time for a linear parabolic

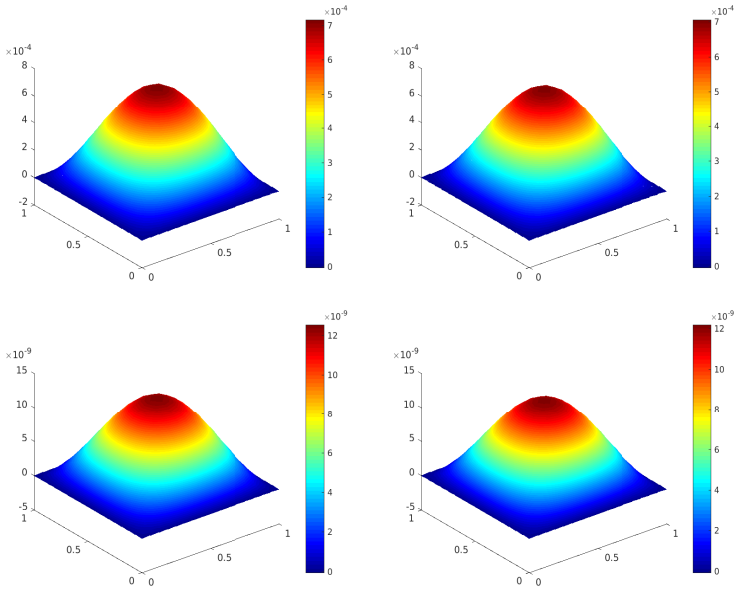

FIGURE 5.6. WG solution plot (top left), exact solution plot (top right) at $t = 5$ and WG solution plot (bottom left), exact solution plot (bottom right) at $t = 10$ in example 5.4 for non-smooth initial data.

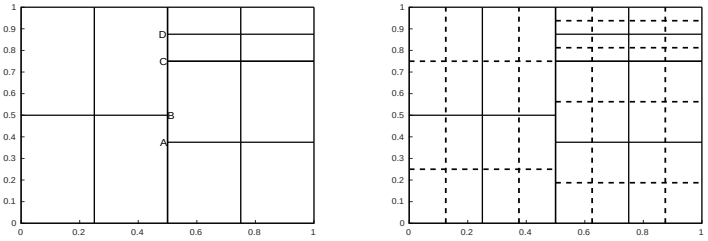

FIGURE 5.7. An initial rectangular mesh with Hanging nodes (left) and its refinement (right).

equation with non-smooth initial data on polygonal meshes. The optimal order of convergence in the $L^\infty(L^2)$ norm is proved for both semidiscrete and full-discrete WG schemes for linear WG space. The above-discussed analysis can also be used for proving non-homogeneous data as well. The obtained results and numerical scheme can help to solve a wide variety of heat conduction models with non-homogeneous inner structures. Currently, we are working on the stabilizer-free WG-FEMs for the following Cahn-Hilliard equation with non-smooth initial data (cf. [12, 30])

$$u_t - \Delta(-\Delta u + \phi(u)) = 0 \quad x \in \Omega \subset \mathbb{R}^3, \quad t > 0,$$

together with appropriate boundary conditions.

Acknowledgments

The authors are grateful to the anonymous referees for their valuable comments and suggestions which greatly improved the presentation of this paper.

Disclosure Statement

The authors declare that they have no conflict of interest.

References

[1] R. Adams and J. Fournier. Sobolev Spaces, sec. ed. Academic Press, Amsterdam, 2003.
[2] N. Asmar. Partial Differential Equations with Fourier Series and Boundary Value Problems: Third Edition. Dover Books on Mathematics. Dover Publications, 2017.
[3] G. R. Barrenechea, F. Brezzi, A. Cangiani, and E. H. Georgoulis. Building bridges: connections and challenges in modern approaches to numerical partial differential equations, volume 114. Springer, 2016.
[4] J. Bramble, J. Pasciak, P. Sammon, and V. Thomee. Incomplete iterations in multistep backward difference methods for parabolic problems with smooth and nonsmooth data. Math. Comp., 52, 05 1989.
[5] E. Burman, O. Duran, A. Ern, and M. Steins. Convergence analysis of hybrid high-order methods for the wave equation. J. Sci. Comput., 87(3):1–30, 2021.
[6] J. R. Cannon. The one-dimensional heat equation. Number 23. Cambridge University Press, 1984.
[7] V. Capasso and K. Kunisch. A reaction-diffusion system arising in modelling man-environment diseases. Quart. Appl. Math., 46(3):431–450, 1988.
[8] B. Cockburn, D. A. Di Pietro, and A. Ern. Bridging the hybrid high-order and hybridizable discontinuous Galerkin methods. ESAIM Math. Model. Numer. Anal., 50(3):635–650, 2016.
[9] B. Deka and N. Kumar. Error estimates in weak Galerkin finite element methods for parabolic equations under low regularity assumptions. Appl. Numer. Math., 162:81–105, 2021.
[10] B. Deka and N. Kumar. A systematic study on weak Galerkin finite element method for second order parabolic problems. arXiv preprint arXiv:2103.13669, 2021.
[11] Z. Dong and A. Ern. Hybrid high-order and weak Galerkin methods for the biharmonic problem. arXiv preprint arXiv:2103.16404, 2021.
[12] C. M. Elliott and S. Larsson. Error estimates with smooth and nonsmooth data for a finite element method for the cahn-hilliard equation. Math. Comp., 58(198):603–630, 1992.
[13] L. C. Evans. Partial differential equations. Amer. Math. Soc., Providence, R.I., 2010.
[14] G. Fichera. Existence theorems in elasticity. In Linear theories of elasticity and thermoelasticity, pages 347–389. Springer, 1973.
[15] Y. Huang, J. Li, and D. Li. Developing weak Galerkin finite element methods for the wave equation. Numer. Methods Partial Differential Equations, 33(3):868–884, 2017.
[16] V. Kosmopoulos and T. Keller. Damage-based finite-element vertebroplasty simulations. Eur. Spine J., 13(7):617–625, 2004.
[17] Q. H. Li and J. Wang. Weak Galerkin finite element methods for parabolic equations. Numer. Methods Partial Differential Equations, 29(6):2004–2024, 2013.
[18] 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):1482–1497, 2018.
[19] M. Luskin and R. Rannacher. On the smoothing property of the Galerkin method for parabolic equations. SIAM J. Numer. Anal., 19(1):93–113, 1982.

[20] J. Mélek, J. Nečas, M. Rokyta, and M. Råužička. Weak and measure-valued solutions to evolutionary PDEs, volume 13. Chapman & Hall, London, UK, 1996.
[21] L. Mu, J. Wang, and X. Ye. A least-squares-based weak Galerkin finite element method for second order elliptic equations. SIAM J. Sci. Comput., 39(4):A1531–A1557, 2017.
[22] L. Mu and J. W. . X. Ye. Weak Galerkin finite element methods on polytopal meshes. Int. J. Numer. Anal. Model., 12(1):31–53, 2015.
[23] R. Rannacher. Finite element solution of diffusion problems with irregular data. Numer. Math., 43(2):309–327, 1984.
[24] P. Sammon. Fully discrete approximation methods for parabolic poblems with non-smooth initial data. SIAM J. Numer. Anal., 20(3):437–470, 1983.
[25] V. Thomée. The finite element method for parabolic problems. In Mathematical Theory of Finite and Boundary Element Methods, pages 135–218. Springer, 1990.
[26] C. Wang and J. Wang. A primal-dual weak Galerkin finite element method for second order elliptic equations in non-divergence form. Math. Comp., 87(310):515–545, 2018.
[27] J. Wang, R. Wang, Q. Zhai, and R. Zhang. A systematic study on weak Galerkin finite element methods for second order elliptic problems. J. Sci. Comput., 74(3):1369–1396, 2018.
[28] J. Wang and X. Ye. A weak Galerkin finite element method for second-order elliptic problems. J. Comput. Appl. Math., 241:103–115, 2013.
[29] J. Wang and X. Ye. A weak Galerkin mixed finite element method for second order elliptic problems. Math. Comp., 83(289):2101–2126, 2014.
[30] J. Wang, Q. Zhai, R. Zhang, and S. Zhang. A weak galerkin finite element scheme for the cahn-hilliard equation. Math. Comp., 88(315):211–235, 2019.
[31] K. Wang and N. Wang. Analysis of a fully discrete finite element method for parabolic interface problems with nonsmooth initial data. J. Comput. Math., 2021.
[32] X. Wang, F. Gao, and Z. Sun. Weak Galerkin finite element method for viscoelastic wave equations. J. Comput. Appl. Math., 375:112816, 2020.
[33] X. Wang, Y. Liu, and Q. Zhai. The weak Galerkin finite element method for solving the time-dependent Stokes flow. Int. J. Numer. Anal. Model., 17(5):732–745, 2020.
[34] H. Zhang, Y. Zou, Y. Xu, Q. Zhai, and H. Yue. Weak Galerkin finite element method for second order parabolic equations. Int. J. Numer. Anal. Model, 13(4):525–544, 2016.
[35] S. Zhou, F. Gao, B. Li, and Z. Sun. Weak Galerkin finite element method with second-order accuracy in time for parabolic problems. Appl. Math. Lett., 90:118–123, 2019.

Department of Mathematics, Indian Institute of Technology Guwahati, Guwahati - 781039, India.
E-mail: nares176123101@iitg.ac.in
E-mail: bdeka@iitg.ac.in