Bull. Korean Math. Soc. **48** (2011), No. 5, pp. 897–915
http://dx.doi.org/10.4134/BKMS.2011.48.5.897

L^2-ERROR ANALYSIS OF FULLY DISCRETE DISCONTINUOUS GALERKIN APPROXIMATIONS FOR NONLINEAR SOBOLEV EQUATIONS

Mi Ray Ohm and Hyun Young Lee

Abstract. In this paper, we develop a symmetric Galerkin method with interior penalty terms to construct fully discrete approximations of the solution for nonlinear Sobolev equations. To analyze the convergence of discontinuous Galerkin approximations, we introduce an appropriate projection and derive the optimal L^2 error estimates.

1. Introduction

In this paper, we consider the following nonlinear Sobolev equation

$$(1.1)\qquad \begin{aligned} &u_t - \nabla\cdot(a(u)\nabla u + b(u)\nabla u_t) = f(u) && \text{in } \Omega\times(0,T],\\ &(a(u)\nabla u + b(u)\nabla u_t)\cdot n = 0 && \text{on } \partial\Omega\times(0,T],\\ &u(x,0) = u_0(x) && \text{on } \Omega, \end{aligned}$$

where Ω is an open bounded domain in $\mathbb{R}^d$, $d = 2$ with smooth boundary $\partial\Omega$, n denotes the unit outward normal vector to $\partial\Omega$ and a, b and f are known functions.

The problem (1.1) represents many physical phenomena such as the flow of fluids through fissured materials [5], thermodynamics [7] and others. For the existence, uniqueness and stability of the solution of the problem (1.1), we refer to [6, 8, 20] and the reference cited in [10]. By many authors the finite difference approximations to the solution of (1.1) have been investigated in the several literatures [9, 11, 12].

Arnold, Douglas and Thomee [1, 2] and Nakao [15] constructed the Galerkin approximations to the solution of (1.1) with periodic boundary conditions in

Received October 1, 2010.
2010 *Mathematics Subject Classification.* 65M15, 65N30.
Key words and phrases. nonlinear Sobolev equation, discontinuous Galerkin approximation, fully discrete approximations, optimal L^2 error estimates.
This research was supported by Kyungsung University Research Grants in 2011.
This paper was issued from the BIEN2009 conference: INWES Asian Network held from August 27 to 29 in Busan, Korea, hosted by the Association of Korean Women Scientists and Engineers(KWSE).

one dimensional space, and obtained the optimal L^2 error estimates and superconvergence results.

In [13], Lin studied the Galerkin approximation of (1.1) with $d = 2$ and derived the optimal L^2 error estimates using Crank-Nicolson method. Lin and Zhang [14] constructed the semidiscrete finite element approximations and derived the optimal L^2 error estimates with nonlinear boundary condition.

Recently in [18, 19], the authors adapted discontinuous Galerkin methods to (1.1) and obtained the optimal H^1 error estimates.

In this paper, we construct the fully discrete approximation of the problem (1.1) using the discontinuous Galerkin method for the spatial discretization and Crank-Nicolson method for the time step discretization. Compared to the classical Galerkin method, the discontinuous Galerkin method has advantages such as the potential error control, the mesh adaption and the local mass conservation.

This paper is organized as follows: In Section 2, we introduce some notations and necessary assumptions on the data and the regularity of the solution of the problem (1.1). In Section 3, we construct finite element spaces, introduce the convergence of the approximations and construct discrete some bilinear forms and auxiliary projection. In Section 4, we formulate discrete fully discrete discontinuous Galerkin approximations and prove the optimal convergence in $\ell^\infty(L^2)$ error estimates.

2. Notations and basic assumptions

For an $s \geq 0$ and a domain $E \subset \mathbb{R}^d$, we let $H^s(E)$ the Sobolev space of order s equipped with the usual Sobolev norm $\|u\|_{s,E}$.

We simply write $\|\cdot\|_s$ instead of $\|\cdot\|_{s,\Omega}$ if $E = \Omega$. And also the usual seminorm of a function defined on E is denoted by $|\cdot|_{s,E}$ and we simply denote $|\cdot|_s$ instead of $|\cdot|_{s,\Omega}$, if $E = \Omega$.

Let $\mathcal{E}_h = \{E_1, E_2, \ldots, E_{N_h}\}$ be a regular quasi-uniform subdivision of Ω where E_i is a triangle or a quadrilateral if $d = 2$ and E_i is a 3-simplex or 3-rectangle if $d = 3$. Let $h_i = \operatorname{diam}(E_i)$ be the diameter of E_i and $h = \max_{1 \leq i \leq N_h} h_i$. We assume that there exists a constant $\rho > 0$ such that each E_i contains a ball of radius ρh_i. The quasiuniformity requirement is that there is a constant $\gamma > 0$ such that

$$\frac{h}{h_i} \leq \gamma, \quad i = 1, 2, \ldots, N_h.$$

Now we assume that for $t \in [0, T]$ and $(x, p) \in \Omega \times \mathbb{R}$, the following regularity conditions on a, b, f and u are satisfied:

(A1) There exist constants a_0 and a^* such that

$$0 < a_0 \leq a(x, p), \quad b(x, p) \leq a^*.$$

(A2) $a(x,p)$, $b(x,p)$ and $f(x,p)$ are continuously differentiable with respect to each variable and are uniformly bounded such as $\left|\frac{\partial a}{\partial p}\right|$, $\left|\frac{\partial b}{\partial p}\right|$, $\left|\frac{\partial^2 a}{\partial p^2}\right|$, $\left|\frac{\partial^2 b}{\partial p^2}\right|$, $\left|\frac{\partial f}{\partial p}\right| \le K$ for some constant $K > 0$.

(A3) $u \in C^2(\Omega \times [0,T])$ is a unique solution to the problem (1.1) and u, $u_t \in L^\infty([0,T], H^s(\Omega))$ for $s \ge \frac{d}{2} + 1 + \delta$ for $\delta > 0$, $u_{tt} \in L^\infty(H^{\frac{d}{2}+1})$.

(A4) ∇u and ∇u_t are bounded in $L^\infty(\Omega \times [0,T])$.

3. Finite element spaces and an auxiliary projection

For an $s \ge 0$ and a given subdivision $\mathcal{E}_h$, we define the following space

$$H^s(\mathcal{E}_h) = \{v \in L^2(\Omega) \mid v|_{E_i} \in H^s(E_i), \quad i = 1, 2, \ldots, N_h\}.$$

Let the edges (or faces if $d = 3$) of $\mathcal{E}_h$ be denoted by $\{e_1, e_2, \ldots, e_{P_h}, e_{P_h+1}, \ldots, e_{M_h}\}$ where $e_k \subset \Omega$, $1 \le k \le P_h$ and $e_k \subset \partial\Omega$ $P_h + 1 \le k \le M_h$. With each edge (or face) e_k, $1 \le k \le P_h$, we associate a unit outward normal vector n_k to E_i if $e_k = \partial E_i \cap \partial E_j$ and $i < j$. For $P_h + 1 \le k \le M_h$, we define $n_k = n$ the unit outward normal vector to $\partial\Omega$.

To present the discontinuous Galerkin scheme, we need some notations on edges, if $d = 2$, or faces if $d = 3$, between two elements. For $\phi \in H^s(\mathcal{E}_h)$, $s > \frac{1}{2}$ we define the following average function $\{\phi\}$ and the jump function $[\phi]$ such that

$$\{\phi\} = \frac{1}{2}(\phi|_{E_i})|_{e_k} + \frac{1}{2}(\phi|_{E_j})|_{e_k}, \quad \forall x \in e_k, \quad 1 \le k \le P_h,$$
$$[\phi] = (\phi|_{E_i})|_{e_k} - (\phi|_{E_j})|_{e_k}, \quad \forall x \in e_k, \quad 1 \le k \le P_h,$$

where $e_k = \partial E_i \cap \partial E_j$, $i < j$.

We associate the following broken norms with the space $H^s(\mathcal{E}_h)$, $s \ge 1$

$$|||\phi|||_0^2 = \sum_{i=1}^{N_h} \|\phi\|_{0,E_i}^2,$$
$$|||\phi|||_1^2 = \sum_{i=1}^{N_h} (\|\phi\|_{1,E_i}^2 + h_i^2 \|\nabla^2 \phi\|_{0,E_i}^2) + J_\beta^\sigma(\phi, \phi),$$

where

$$J_\beta^\sigma(\phi, \psi) = \sum_{k=1}^{P_h} \frac{\sigma_k}{|e_k|^\beta} \int_{e_k} [\phi][\psi] ds, \quad \beta > 0$$

is an interior penalty term and σ is a discrete positive function that takes the constant value σ_k on the edge e_k and is bounded below by $\sigma_0 > 0$ and above by $\sigma^* > 0$.

Let r be a positive integer. The finite element space used in this paper is taken to be

$$D_r(\mathcal{E}_h) = \{v \in L^2(\Omega) : v|_{E_j} \in P_r(E_j), \ j = 1, 2, \ldots, N_h\},$$

where $P_r(E_j)$ denotes the set of polynomials of degree less than or equal to r on E_j.

Now we state the following hp-approximation properties whose proofs can be found in [3, 4].

Lemma 3.1. *Let $E_j \in \mathcal{E}_h$ and $\phi \in H^s(E_j)$. Then there exist a positive constant C depending on s, γ, and ρ but independent of ϕ, r and h and a sequence $z_r^h \in P_r(E_j)$, $r = 1, 2, \ldots$ such that for any $0 \le q \le s$,*

$$\|\phi - z_r^h\|_{q,E_j} \le C \frac{h_j^{\mu-q}}{r^{s-q}} \|\phi\|_{s,E_j} \quad s \ge 0,$$

$$\|\phi - z_r^h\|_{0,e_j} \le C \frac{h_j^{\mu-\frac{1}{2}}}{r^{s-\frac{1}{2}}} \|\phi\|_{s,E_j} \quad s > \frac{1}{2},$$

$$\|\phi - z_r^h\|_{1,e_j} \le C \frac{h_j^{\mu-\frac{3}{2}}}{r^{s-\frac{3}{2}}} \|\phi\|_{s,E_j} \quad s > \frac{3}{2},$$

where $\mu = \min(r+1, s)$ and e_j is an edge or a face of E_j.

Lemma 3.2. *For each $E_j \in \mathcal{E}_h$, there exists a positive constant C depending only on γ and ρ such that the two following trace inequalities hold:*

$$\|\phi\|_{0,e_j}^2 \le C \left(\frac{1}{h_j} |\phi|_{0,E_j}^2 + h_j |\phi|_{1,E_j}^2 \right), \quad \forall \phi \in H^1(E_j),$$

$$\left\| \frac{\partial \phi}{\partial n_j} \right\|_{0,e_j}^2 \le C \left(\frac{1}{h_j} |\phi|_{1,E_j}^2 + h_j |\phi|_{2,E_j}^2 \right), \quad \forall \phi \in H^2(E_j),$$

where e_j is an edge or a face of E_j and n_j is the unit outward normal vector to E_j.

Now we introduce the following bilinear mappings $A(\rho; \cdot, \cdot)$ and $B(\rho; \cdot, \cdot)$ defined on $H^s(\mathcal{E}_h) \times H^s(\mathcal{E}_h)$

$$\begin{aligned} A(\rho; \phi, \psi) = {} & (a(\rho)\nabla\phi, \nabla\psi) - \sum_{k=1}^{P_h} \int_{e_k} \{a(\rho)\nabla\phi \cdot n_k\}[\psi] - \sum_{k=1}^{P_h} \int_{e_k} \{a(\rho)\nabla\psi \cdot n_k\}[\phi] \\ & + J_\beta^\sigma(\phi, \psi), \\ B(\rho; \phi, \psi) = {} & (b(\rho)\nabla\phi, \nabla\psi) - \sum_{k=1}^{P_h} \int_{e_k} \{b(\rho)\nabla\phi \cdot n_k\}[\psi] - \sum_{k=1}^{P_h} \int_{e_k} \{b(\rho)\nabla\psi \cdot n_k\}[\phi] \\ & + J_\beta^\sigma(\phi, \psi). \end{aligned}$$

Using the bilinear mappings A and B, we construct the weak formulation of the problem (1.1) as follows

$$\begin{aligned} & (u_t(t_{j\theta}), v) + A(u(t_{j\theta}); u(t_{j\theta}), v) + B(u(t_{j\theta}); u_t(t_{j\theta}), v) = (f(u(t_{j\theta})), v), \\ & \forall\, v \in H^s(\mathcal{E}_h). \end{aligned} \tag{3.1}$$

Now for $\lambda > 0$ we define the following bilinear forms $A_\lambda(\rho;\cdot,\cdot)$ and $B_\lambda(\rho;\cdot,\cdot)$ on $H^s(\mathcal{E}_h)\times H^s(\mathcal{E}_h)$ such that

$$\begin{aligned} A_\lambda(\rho;\phi,\psi) &= A(\rho;\phi,\psi)+\lambda(\phi,\psi),\\ B_\lambda(\rho;\phi,\psi) &= B(\rho;\phi,\psi)+\lambda(\phi,\psi). \end{aligned}$$

A_λ and B_λ satisfy the following boundedness and coercivity properties, respectively. The proofs can be found in [16, 17].

Lemma 3.3. *For $\lambda > 0$, there exists a constant $C > 0$ satisfying*

$$\begin{aligned} |A_\lambda(\rho;\phi,\psi)| &\le C|||\phi|||_1|||\psi|||_1,\\ |B_\lambda(\rho;\phi,\psi)| &\le C|||\phi|||_1|||\psi|||_1, \quad \forall\phi,\psi\in H^s(\mathcal{E}_h). \end{aligned}$$

Lemma 3.4. *For $\lambda > 0$, there exists a constant $\underset{\sim}{c} > 0$ satisfying*

$$\begin{aligned} A_\lambda(\rho;\phi,\phi) &\ge \underset{\sim}{c}|||\phi|||_1^2,\\ B_\lambda(\rho;\phi,\phi) &\ge \underset{\sim}{c}|||\phi|||_1^2, \quad \forall\phi\in D_r(\mathcal{E}_h). \end{aligned}$$

Wheeler [21] introduced an elliptic projection to prove the optimal L^2 error estimates for Galerkin approximation to parabolic differential equations. Modifying this idea we define a differentiable map $\widetilde{u}(t):[0,T]\to D_r(\mathcal{E}_h)$ such that

$$A_\lambda(u;u-\widetilde{u},v)+B_\lambda(u;u_t-\widetilde{u}_t,v)=0, \quad \forall v\in D_r(\mathcal{E}_h). \tag{3.2}$$

By Lemma 3.3 and Lemma 3.4, $\widetilde{u}(t)$ is well-defined.

4. The optimal $\ell^\infty(L^2)$ error estimates

For a positive integer $N > 0$, we let $\Delta t = \frac{T}{N}$ and $t_j = j(\Delta t)$, $0 \le k \le N$. For $0\le j\le N$, we define $g_j = g(x,t_j)$ and for $0\le j\le N-1$, we define

$$\begin{aligned} &\partial_t g_j = \frac{g_{j+1}-g_j}{\Delta t}, \quad g_{j\theta} = \frac{1}{2}(1+\theta)g_{j+1}+\frac{1}{2}(1-\theta)g_j \ \text{ and}\\ &t_{j\theta} = \frac{1}{2}(1+\theta)t_{j+1}+\frac{1}{2}(1-\theta)t_j. \end{aligned}$$

Now we formulate a fully discrete discontinuous Galerkin approximation to (3.1) as follows: Find $\{U_n\}_{n=1}^N\subset D_r(\mathcal{E}_h)$ satisfying

$$\begin{aligned} &(\partial_t U_j,v)+A(U_{j\theta};U_{j\theta},v)+B(U_{j\theta};\partial_t U_j,v)=(f(U_{j\theta}),v), \quad v\in D_r(\mathcal{E}_h)\\ &(U_0,v)=(u_0,v). \end{aligned} \tag{4.1}$$

In this section we prove the optimal convergence of $u(t_j)-U_j$ in L^2 space. To analyze the error in L^2 space, we denote $\eta(x,t)=u(x,t)-\widetilde{u}(x,t)$ and $\xi(x,t_j)=\widetilde{u}(x,t_j)-U_j(x)$, $j=0,1,\ldots,N$.

Now we state the following approximations for η and η_t whose proofs can be found in [16, 17].

Theorem 4.1. *If $u_t \in L^2(H^s)$ and $u_0 \in H^s$, then there exists a constant C independent of h and Δt such that*

(i) $\|\eta_t\|_0 + h\|\eta_t\|_1 \le Ch^s(\|u_t\|_{L^2(H^s)} + \|u_0\|_s)$,

(ii) $\|\eta\|_0 + h\|\eta\|_1 \le Ch^s(\|u_t\|_{L^2(H^s)} + \|u_0\|_s)$.

Theorem 4.2. *If $u_t \in L^2(H^s)$, $u_{tt} \in H^s$, $u_{ttt} \in H^s$ and $u_0 \in H^s$, then there exists a constant C independent of h and Δt such that*

(i) $\|\eta_{tt}\|_1 \le Ch^{s-1}\{\|u_t\|_{L^2(H^s)} + \|u_{tt}\|_{H^s} + \|u_0\|_s\}$,

(ii) $\|\eta_{ttt}\|_1 \le Ch^{s-1}\{\|u_t\|_{L^2(H^s)} + \|u_{tt}\|_{H^s} + \|u_{ttt}\|_{H^s} + \|u_0\|_s\}$

provided that $\beta \le \frac{1}{d-1}$.

Proof. Differentiating the both sides of (3.2) with respect to t, we get

$$\begin{aligned} &\Big(\Big(\frac{d}{dt}a(u)\Big)\nabla\eta, \nabla v\Big) - \sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}a(u)\Big)\nabla\eta\cdot n_k\Big\}[v] \\ &-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}a(u)\Big)\nabla v\cdot n_k\Big\}[\eta] + A_\lambda(u;\eta_t,v) + \Big(\Big(\frac{d}{dt}b(u)\Big)\nabla\eta_t, \nabla v\Big) \\ &-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}b(u)\Big)\nabla\eta_t\cdot n_k\Big\}[v] - \sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}b(u)\Big)\nabla v\cdot n_k\Big\}[\eta_t] \\ &+ B_\lambda(u;\eta_{tt},v) = 0. \end{aligned} \tag{4.2}$$

Now we define the bilinear forms A_t and B_t as follows

$$\begin{aligned} A_t(u;\eta,v) &= \Big(\Big(\frac{d}{dt}a(u)\Big)\nabla\eta, \nabla v\Big) - \sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}a(u)\Big)\nabla\eta\cdot n_k\Big\}[v] \\ &\quad -\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}a(u)\Big)\nabla v\cdot n_k\Big\}[\eta], \\ B_t(u;\eta_t,v) &= \Big(\Big(\frac{d}{dt}b(u)\Big)\nabla\eta_t, \nabla v\Big) - \sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}b(u)\Big)\nabla\eta_t\cdot n_k\Big\}[v] \\ &\quad -\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}b(u)\Big)\nabla v\cdot n_k\Big\}[\eta_t]. \end{aligned}$$

From (4.2) and the definitions of A_t and B_t, we get

$$B_\lambda(u;\eta_{tt},v) = -A_\lambda(u;\eta_t,v) - A_t(u;\eta,v) - B_t(u;\eta_t,v), \quad \forall v \in D_r(\mathcal{E}_h).$$

We let

$$\theta = P_h u - \widetilde{u},$$

where $P_h u$ is the polynomial approximation of u satisfying the properties in Lemma 3.1. Then we have the following equalities

$$B_\lambda(u;\theta_{tt},\theta_{tt}) = B_\lambda(u;u_{tt} + P_h u_{tt} - u_{tt} - \widetilde{u}_{tt},\theta_{tt})$$

$$
\begin{aligned}
&= B_\lambda(u;\eta_{tt}\theta_{tt}) - B_\lambda(u;u_{tt}-P_hu_{tt},\theta_{tt})\\
&= -A_\lambda(u;\eta_t,\theta_{tt}) - A_t(u;\eta,\theta_{tt}) - B_t(u;\eta_t,\theta_{tt})\\
&\quad - B_\lambda(u;u_{tt}-P_hu_{tt},\theta_{tt}).
\end{aligned}
$$

By Lemma 3.3 and Lemma 3.4 we obtain

$$\|\theta_{tt}\|_1 \le C(\|\eta_t\|_1 + \|\eta\|_1 + \|u_{tt}-P_hu_{tt}\|_1),$$

which implies

$$
\begin{aligned}
\|\eta_{tt}\|_1 &\le \|u_{tt}-P_hu_{tt}\|_1 + \|P_hu_{tt}-\widetilde{u}_{tt}\|_1\\
&\le \|u_{tt}-P_hu_{tt}\|_1 + C(\|\eta_t\|_1 + \|\eta\|_1 + \|u_{tt}-P_hu_{tt}\|_1).
\end{aligned}
\tag{4.3}
$$

From Lemma 3.1 we have the following estimation

$$
\begin{aligned}
&\|u_{tt}-P_hu_{tt}\|_1^2\\
&= \sum_{i=1}^{N_h}\|u_{tt}-P_hu_{tt}\|_{1,E_i}^2 + \sum_{i=1}^{N_h} h_i^2\|\nabla^2(u_{tt}-P_hu_{tt})\|_{0,E_i}^2\\
&\quad + J_\beta^\sigma(u_{tt}-P_hu_{tt},u_{tt}-P_hu_{tt})\\
&\le C\Big(h^{2(s-1)}\|u_{tt}\|_s^2 + h^{-(d-1)\beta}\sum_{k=1}^{P_h}\|[u_{tt}-P_hu_{tt}]\|_{0,e_k}^2\Big)\\
&\le C[h^{2(s-1)}\|u_{tt}\|_s^2 + h^{-(d-1)\beta-1}(h^{2s}\|u_{tt}\|_s^2)]\\
&\le Ch^{2(s-1)}\|u_{tt}\|_s^2.
\end{aligned}
$$

By applying the inequality above to (4.3) we obtain the following

$$
\begin{aligned}
\|\eta_{tt}\|_1^2 &\le C(\|u_{tt}-P_hu_{tt}\|_1^2 + \|\eta\|_1^2 + \|\eta_t\|_1^2)\\
&\le Ch^{2(s-1)}(\|u_{tt}\|_{H^s}^2 + \|u_t\|_{L^2(H^s)}^2 + \|u_0\|_s^2),
\end{aligned}
$$

which implies

$$\|\eta_{tt}\|_1 \le Ch^{s-1}(\|u_{tt}\|_{H^s} + \|u_t\|_{L^2(H^s)} + \|u_0\|_s). \tag{4.4}$$

Differentiating the both sides of (4.2) with respect to t, we get

(4.5)

$$
\begin{aligned}
&\Big(\Big(\frac{d^2}{dt^2}a(u)\Big)\nabla\eta,\nabla v\Big) - \sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d^2}{dt^2}a(u)\Big)\nabla\eta\cdot n_k\Big\}[v]\\
&-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d^2}{dt^2}a(u)\Big)\nabla v\cdot n_k\Big\}[\eta] + \Big(\Big(\frac{d}{dt}a(u)\Big)\nabla\eta_t,\nabla v\Big)\\
&-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}a(u)\Big)\nabla\eta_t\cdot n_k\Big\}[v] - \sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}a(u)\Big)\nabla v\cdot n_k\Big\}[\eta_t]
\end{aligned}
$$

$$
\begin{aligned}
&+\Big(\Big(\frac{d}{dt}a(u)\Big)\nabla\eta_t,\nabla v\Big)-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}a(u)\Big)\nabla\eta_t\cdot n_k\Big\}[v]\\
&-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}a(u)\Big)\nabla v\cdot n_k\Big\}[\eta_t]+A_\lambda(u;\eta_{tt},v)+\Big(\Big(\frac{d^2}{dt^2}b(u)\Big)\nabla\eta_t,\nabla v\Big)\\
&-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d^2}{dt^2}b(u)\Big)\nabla\eta_t\cdot n_k\Big\}[v]-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d^2}{dt^2}b(u)\Big)\nabla v\cdot n_k\Big\}[\eta_t]\\
&+\Big(\Big(\frac{d}{dt}b(u)\Big)\nabla\eta_{tt},\nabla v\Big)-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}b(u)\Big)\nabla\eta_{tt}\cdot n_k\Big\}[v]\\
&-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}b(u)\Big)\nabla v\cdot n_k\Big\}[\eta_{tt}]+\Big(\Big(\frac{d}{dt}b(u)\Big)\nabla\eta_{tt},\nabla v\Big)\\
&-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}b(u)\Big)\nabla\eta_{tt}\cdot n_k\Big\}[v]-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d}{dt}b(u)\Big)\nabla v\cdot n_k\Big\}[\eta_{tt}]\\
&+B_\lambda(u;\eta_{ttt},v)=0.
\end{aligned}
$$

Now we define the bilinear mappings A_{tt} and B_{tt} as follows

$$
\begin{aligned}
A_{tt}(u;\eta,v)&=\Big(\Big(\frac{d^2}{dt^2}a(u)\Big)\nabla\eta,\nabla v\Big)-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d^2}{dt^2}a(u)\Big)\nabla\eta\cdot n_k\Big\}[v]\\
&\quad-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d^2}{dt^2}a(u)\Big)\nabla v\cdot n_k\Big\}[\eta],\\
B_{tt}(u;\eta_t,v)&=\Big(\Big(\frac{d^2}{dt^2}b(u)\Big)\nabla\eta_t,\nabla v\Big)-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d^2}{dt^2}b(u)\Big)\nabla\eta_t\cdot n_k\Big\}[v]\\
&\quad-\sum_{k=1}^{P_h}\int_{e_k}\Big\{\Big(\frac{d^2}{dt^2}b(u)\Big)\nabla v\cdot n_k\Big\}[\eta_t].
\end{aligned}
$$

From (4.5) and the definitions of A_{tt} and B_{tt}, we get

$$
\begin{aligned}
B_\lambda(u;\eta_{ttt},v)=&-A_\lambda(u;\eta_{tt},v)-A_{tt}(u;\eta,v)-2A_t(u;\eta_t,v)\\
&-B_{tt}(u;\eta_t,v)-2B_t(u;\eta_{tt},v),\quad\forall v\in D_r(\mathcal{E}_h).
\end{aligned}
\tag{4.6}
$$

By (4.6) we obtain

$$
\begin{aligned}
B_\lambda(u;\theta_{ttt},\theta_{ttt})&=B_\lambda(u;u_{ttt}+P_hu_{ttt}-u_{ttt}-\widetilde{u}_{ttt},\theta_{ttt})\\
&=B_\lambda(u;\eta_{ttt},\theta_{ttt})-B_\lambda(u;u_{ttt}-P_hu_{ttt},\theta_{ttt})\\
&=-A_\lambda(u;\eta_{tt},\theta_{ttt})-A_{tt}(u;\eta,\theta_{ttt})\\
&\quad-2A_t(u;\eta_t,\theta_{ttt})-B_{tt}(u;\eta_t,\theta_{ttt})\\
&\quad-2B_t(u;\eta_{tt},\theta_{ttt})-B_\lambda(u;u_{ttt}-P_hu_{ttt},\theta_{ttt}).
\end{aligned}
$$

By Lemma 3.3, Lemma 3.4, (4.4) and Theorem 3.1, we obtain

$$\begin{aligned}|||\theta_{ttt}|||_1 &\le C(|||u_{tt}|||_1 + |||\eta|||_1 + |||\eta_t|||_1 + |||u_{ttt} - P_h u_{ttt}|||_1)\\ &\le Ch^{s-1}(\|u_{tt}\|_{H^s} + \|u_t\|_{L^2(H^s)} + \|u_0\|_s + \|u_{ttt}\|_{H^s}),\end{aligned}$$

which implies

$$\begin{aligned}|||\eta_{ttt}|||_1 &\le |||u_{ttt} - P_h u_{ttt}|||_1 + |||P_h u_{ttt} - \widetilde{u}_{ttt}|||_1\\ &\le Ch^{s-1}(\|u_{tt}\|_{H^s} + \|u_t\|_{L^2(H^s)} + \|u_0\|_s + \|u_{ttt}\|_{H^s}).\end{aligned}$$ □

By simple computation, we obviously obtain the two following lemmas.

Lemma 4.1. *If $\rho_{j\theta}$ satisfies*

$$\partial_t \widetilde{u}_j - \widetilde{u}_t(t_{j\theta}) = (\Delta t)\rho_{j\theta},$$

then there exists a constant C independent of h and Δt such that

(i) *if $0 < \theta \le 1$*

$$\begin{aligned}|||\rho_{j\theta}|||_0 &\le C(\|u_0\|_s + \|u_t\|_{L^2(H^s)} + \|u_{tt}\|_{L^\infty(H^s)}),\\ |||\rho_{j\theta}|||_1 &\le C(\|u_0\|_s + \|u_t\|_{L^2(H^s)} + \|u_{tt}\|_{L^\infty(H^s)}),\end{aligned}$$

(ii) *if $\theta = 0$*

$$\begin{aligned}|||\rho_{j\theta}|||_0 &\le C\Delta t(\|u_0\|_s + \|u_t\|_{L^2(H^s)} + \|u_{tt}\|_{L^\infty(H^s)} + \|u_{ttt}\|_{L^\infty(H^s)}),\\ |||\rho_{j\theta}|||_1 &\le C\Delta t(\|u_0\|_s + \|u_t\|_{L^2(H^s)} + \|u_{tt}\|_{L^\infty(H^s)} + \|u_{ttt}\|_{L^\infty(H^s)}).\end{aligned}$$

Lemma 4.2. *If we let $r_{j\theta} = \widetilde{u}(t_{j\theta}) - \widetilde{u}_{j\theta}$, then there exists a constant C independent of h and Δt such that*

$$\begin{aligned}|||r_{j\theta}|||_0 &\le C(\Delta t)^2(\|u_0\|_s + \|u_t\|_{L^2(H^s)} + \|u_{tt}\|_{L^\infty(H^s)}),\\ |||r_{j\theta}|||_1 &\le C(\Delta t)^2(\|u_0\|_s + \|u_t\|_{L^2(H^s)} + \|u_{tt}\|_{L^\infty(H^s)}).\end{aligned}$$

Theorem 4.3. *For sufficiently small $\lambda > 0$ and $\delta > 0$, if $u_t \in L^\infty(H^s)$ and $u_{tt} \in L^\infty(H^s)$, then there exists a constant $C > 0$ independent on h and Δt such that*

(i) *if $\theta \in (0, 1]$ and there exists a constant C_0 satisfying $h^{-\frac{d}{2}}\Delta t \le C_0$, then*

$$|||u(t_j) - U_j|||_0 \le C(h^\mu + \Delta t)(\|u_0\|_s + \|u_t\|_{L^\infty(H^s)} + \|\nabla u_t\|_{L^\infty} + \|u_{tt}\|_{L^\infty(H^1)}),$$

(ii) *if $\theta = 0$, then*

$$\begin{aligned}|||u(t_j) - U_j|||_0 \le C(h^\mu + \Delta t^2)(&\|u_0\|_s + \|u_t\|_{L^\infty(H^s)} + \|\nabla u_t\|_{L^\infty} + \|u_{tt}\|_{L^\infty(H^1)}\\ &+ \|u_{ttt}\|_{L^\infty(H^1)})\end{aligned}$$

hold for $j = 1, 2, \ldots, N$, where $s = \frac{d}{2} + 1 + \delta$ and $\mu = \min(r + 1, s)$.

Proof. From (3.1) and (4.1) we have

$$\begin{aligned}&(u_t(t_{j\theta}) - \partial_t U_j, v) + A_\lambda(u(t_{j\theta}); u(t_{j\theta}), v) - A_\lambda(U_{j\theta}; U_{j\theta}, v)\\ &\quad + B_\lambda(u(t_{j\theta}); u_t(t_{j\theta}), v) - B_\lambda(U_{j\theta}; \partial_t U_j, v)\\ &= (f(u(t_{j\theta})) - f(U_{j\theta}), v) + \lambda(u(t_{j\theta}) - U_{j\theta}, v) + \lambda(u_t(t_{j\theta}) - \partial_t U_j, v).\end{aligned} \tag{4.7}$$

From the definition of η and ξ, we have

$$\begin{aligned} u_t(t_{j\theta}) - \partial_t U_j &= \eta_t(t_{j\theta}) + \widetilde{u}_t(t_{j\theta}) - \partial_t \widetilde{u}_j + \partial_t \xi_j \\ &= \eta_t(t_{j\theta}) + \Delta t \rho_{j\theta} + \partial_t \xi_j. \end{aligned} \tag{4.8}$$

By the definition of η, we get

$$\begin{aligned} &A_\lambda(u(t_{j\theta}); u(t_{j\theta}), v) - A_\lambda(U_{j\theta}; U_{j\theta}, v) \\ =\ &A_\lambda(U_{j\theta}; \xi_{j\theta}, v) + A_\lambda(u(t_{j\theta}); u(t_{j\theta}), v) - A_\lambda(U_{j\theta}; \widetilde{u}_{j\theta}, v) \\ =\ &A_\lambda(U_{j\theta}; \xi_{j\theta}, v) + A_\lambda(u(t_{j\theta}); \eta(t_{j\theta}), v) + A_\lambda(u(t_{j\theta}); \widetilde{u}(t_{j\theta}) - \widetilde{u}_{j\theta}, v) \\ &+ A_\lambda(u(t_{j\theta}); \widetilde{u}_{j\theta}, v) - A_\lambda(U_{j\theta}; \widetilde{u}_{j\theta}, v). \end{aligned} \tag{4.9}$$

Form the definition of η, we have

$$\begin{aligned} &B_\lambda(u(t_{j\theta}); u_t(t_{j\theta}), v) - B_\lambda(U_{j\theta}; \partial_t U_j, v) \\ =\ &B_\lambda(U_{j\theta}; \partial_t \xi_j, v) + B_\lambda(u(t_{j\theta}); u_t(t_{j\theta}), v) - B_\lambda(U_{j\theta}; \partial_t \widetilde{u}_j, v) \\ =\ &B_\lambda(U_{j\theta}; \partial_t \xi_j, v) + B_\lambda(u(t_{j\theta}); \eta_t(t_{j\theta}), v) \\ &+ B_\lambda(u(t_{j\theta}); \widetilde{u}_t(t_{j\theta}) - \partial_t \widetilde{u}_j, v) + B_\lambda(u(t_{j\theta}); \partial_t \widetilde{u}_j, v) \\ &- B_\lambda(U_{j\theta}; \partial_t \widetilde{u}_j, v). \end{aligned} \tag{4.10}$$

Substituting (4.8)-(4.10) in (4.7) implies

$$\begin{aligned} &(\partial_t \xi_j, v) + A_\lambda(U_{j\theta}; \xi_{j\theta}, v) + B_\lambda(U_{j\theta}; \partial_t \xi_j, v) \\ =\ &-(\eta_t(t_{j\theta}) + \Delta t \rho_{j\theta}, v) - A_\lambda(u(t_{j\theta}); \eta(t_{j\theta}), v) - A_\lambda(u(t_{j\theta}); \widetilde{u}(t_{j\theta}) - \widetilde{u}_{j\theta}, v) \\ &- A_\lambda(u(t_{j\theta}); \widetilde{u}_{j\theta}, v) + A_\lambda(U_{j\theta}; \widetilde{u}_{j\theta}, v) - B_\lambda(u(t_{j\theta}); \eta_t(t_{j\theta}), v) \\ &- B_\lambda(u(t_{j\theta}); \widetilde{u}_t(t_{j\theta}) - \partial_t \widetilde{u}_j, v) - B_\lambda(u(t_{j\theta}); \partial_t \widetilde{u}_j, v) + B_\lambda(U_{j\theta}; \partial_t \widetilde{u}_j, v) \\ &+ f(u(t_{j\theta}) - f(U_{j\theta}), v) + \lambda(u(t_{j\theta}) - U_{j\theta}, v) + \lambda(u_t(t_{j\theta}) - \partial_t U_j, v). \end{aligned}$$

By (3.2) and (4.8), we have

$$\begin{aligned} &(\partial_t \xi_j, v) + A_\lambda(U_{j\theta}; \xi_{j\theta}, v) + B_\lambda(U_{j\theta}; \partial_t \xi_j, v) \\ =\ &-(\eta_t(t_{j\theta}) + \Delta t \rho_{j\theta}, v) - A_\lambda(u(t_{j\theta}); \widetilde{u}(t_{j\theta}) - \widetilde{u}_{j\theta}, v) \\ &- A_\lambda(u(t_{j\theta}); \widetilde{u}_{j\theta}, v) + A_\lambda(U_{j\theta}; \widetilde{u}_{j\theta}, v) \\ &- B_\lambda(u(t_{j\theta}); \widetilde{u}_t(t_{j\theta}) - \partial_t \widetilde{u}_j, v) - B_\lambda(u(t_{j\theta}); \partial_t \widetilde{u}_j, v) \\ &+ B_\lambda(U_{j\theta}; \partial_t \widetilde{u}_j, v) + (f(u(t_{j\theta})) - f(U_{j\theta}), v) \\ &+ \lambda(\eta(t_{j\theta}) + \widetilde{u}(t_{j\theta}) - \widetilde{u}_{j\theta} + \xi_{j\theta}, v) \\ &+ \lambda(\eta_t(t_{j\theta}) + \Delta t \rho_{j\theta} + \partial_t \xi_j, v). \end{aligned} \tag{4.11}$$

If we choose $v = \xi_{j\theta} + \partial_t \xi_j$ in (4.11), we get

$$\begin{aligned} &(\partial_t \xi_j, \xi_{j\theta} + \partial_t \xi_j) + A_\lambda(U_{j\theta}; \xi_{j\theta}, \xi_{j\theta}) + B_\lambda(U_{j\theta}; \partial_t \xi_j, \xi_{j\theta}) \\ &+ A_\lambda(U_{j\theta}; \xi_{j\theta}, \partial_t \xi_j) + B_\lambda(U_{j\theta}; \partial_t \xi_j, \partial_t \xi_j) \\ =\ &-(\eta_t(t_{j\theta}) + \Delta t \rho_{j\theta}, \xi_{j\theta} + \partial_t \xi_j) - A_\lambda(u(t_{j\theta}); \widetilde{u}(t_{j\theta}) - \widetilde{u}_{j\theta}, \xi_{j\theta} + \partial_t \xi_j) \\ &- A_\lambda(u(t_{j\theta}); \widetilde{u}_{j\theta}, \xi_{j\theta} + \partial_t \xi_j) + A_\lambda(U_{j\theta}; \widetilde{u}_{j\theta}, \xi_{j\theta} + \partial_t \xi_j) \end{aligned} \tag{4.12}$$

$$
\begin{aligned}
&+ B_\lambda(u(t_{j\theta}); \Delta t\rho_{j\theta}, \xi_{j\theta} + \partial_t\xi_j) - B_\lambda(u(t_{j\theta}); \partial_t\widetilde{u}_j, \xi_{j\theta} + \partial_t\xi_j) \\
&+ B_\lambda(U_{j\theta}; \partial_t\widetilde{u}_j, \xi_{j\theta} + \partial_t\xi_j) + (f(u(t_{j\theta})) - f(U_{j\theta}), \xi_{j\theta} + \partial_t\xi_j) \\
&+ \lambda(\eta(t_{j\theta}) + \widetilde{u}(t_{j\theta}) - \widetilde{u}_{j\theta} + \xi_{j\theta}, \xi_{j\theta} + \partial_t\xi_j) \\
&+ \lambda(\eta_t(t_{j\theta}) + \Delta t\rho_{j\theta} + \partial_t\xi_j, \xi_{j\theta} + \partial_t\xi_j).
\end{aligned}
$$

By Cauchy Schwarz's inequality, we obviously have

$$
\begin{aligned}
(\xi_{j\theta}, \partial_t\xi_j) &= \left(\frac{1+\theta}{2}\xi_{j+1} + \frac{1-\theta}{2}\xi_j, \frac{\xi_{j+1}-\xi_j}{\Delta t}\right) \\
&\geq \frac{1}{2\Delta t}(\|\xi_{j+1}\|_0^2 - \|\xi_j\|_0^2).
\end{aligned} \tag{4.13}
$$

By the definition of A_λ we obtain

$$
\begin{aligned}
&A_\lambda(U_{j\theta}; \xi_{j\theta}, \partial_t\xi_j) \\
&= (a(U_{j\theta})\nabla\xi_{j\theta}, \nabla(\partial_t\xi_j)) - \sum_{k=1}^{P_h}\int_{e_k}\{a(U_{j\theta})\nabla\xi_{j\theta}\cdot n_k\}[\partial_t\xi_j] \\
&\quad - \sum_{k=1}^{P_h}\int_{e_k}\{a(U_{j\theta})\nabla(\partial_t\xi_j)\cdot n_k\}[\xi_{j\theta}] + J_\beta^\sigma(\xi_{j\theta}, \partial_t\xi_j) + \lambda(\xi_{j\theta}, \partial_t\xi_j) \\
&\geq \frac{a_0}{2\Delta t}(\|\nabla\xi_{j+1}\|_0^2 - \|\nabla\xi_j\|_0^2) - \sum_{k=1}^{P_h}\int_{e_k}\{a(U_{j\theta})\nabla\xi_{j\theta}\cdot n_k\}[\partial_t\xi_j] \\
&\quad - \sum_{k=1}^{P_h}\int_{e_k}\{a(U_{j\theta})\nabla(\partial_t\xi_j)\cdot n_k\}[\xi_{j\theta}] + \frac{1}{2\Delta t}\Big[J_\beta^\sigma(\xi_{j+1}, \xi_{j+1}) \\
&\quad - J_\beta^\sigma(\xi_j, \xi_j)\Big] + \frac{\lambda}{2\Delta t}(\|\xi_{j+1}\|_0^2 - \|\xi_j\|_0^2).
\end{aligned} \tag{4.14}
$$

From the definition of B_λ we get the following

$$
\begin{aligned}
&B_\lambda(U_{j\theta}; \partial_t\xi_j, \xi_{j\theta}) \\
&= (b(U_{j\theta})\nabla(\partial_t\xi_j), \nabla\xi_{j\theta}) - \sum_{k=1}^{P_h}\int_{e_k}\{b(U_{j\theta})\nabla(\partial_t\xi_j)\cdot n_k\}[\xi_{j\theta}] \\
&\quad - \sum_{k=1}^{P_h}\int_{e_k}\{b(U_{j\theta})\nabla\xi_{j\theta}\cdot n_k\}[\partial_t\xi_j] + J_\beta^\sigma(\partial_t\xi_j, \xi_{j\theta}) + \lambda(\partial_t\xi_j, \xi_{j\theta}) \\
&\geq \frac{a_0}{2\Delta t}(\|\nabla\xi_{j+1}\|_0^2 - \|\nabla\xi_j\|_0^2) - \sum_{k=1}^{P_h}\int_{e_k}\{b(U_{j\theta})\nabla(\partial_t\xi_j)\cdot n_k\}[\xi_{j\theta}] \\
&\quad - \sum_{k=1}^{P_h}\int_{e_k}\{b(U_{j\theta})\nabla\xi_{j\theta}\cdot n_k\}[\partial_t\xi_j] + \frac{1}{2\Delta t}\Big(J_\beta^\sigma(\xi_{j+1}, \xi_{j+1}) \\
&\quad - J_\beta^\sigma(\xi_j, \xi_j)\Big) + \frac{\lambda}{2\Delta t}(\|\xi_{j+1}\|_0^2 - \|\xi_j\|_0^2).
\end{aligned} \tag{4.15}
$$

Applying (4.13)-(4.15) in (4.12) we have the following inequality

$$
\begin{aligned}
&\frac{1}{2\Delta t}\Big[(1+2\lambda)(\|\xi_{j+1}\|_0^2 - \|\xi_j\|_0^2) + 2(J_\beta^\sigma(\xi_{j+1},\xi_{j+1}) - J_\beta^\sigma(\xi_j,\xi_j)) \\
&+ 2a_0(\|\nabla\xi_{j+1}\|_0^2 - \|\nabla\xi_j\|_0^2)\Big] + \|\partial_t\xi_j\|_0^2 + \underset{\sim}{c}(\|\xi_{j\theta}\|_1^2 + \|\partial_t\xi_j\|_1^2) \\
\le\ & -(\eta_t(t_{j\theta}) + \Delta t\rho_{j\theta}, \xi_{j\theta} + \partial_t\xi_j) \\
&- A_\lambda(u(t_{j\theta}); \widetilde{u}(t_{j\theta}) - \widetilde{u}_{j\theta}, \xi_{j\theta} + \partial_t\xi_j) \\
&- A_\lambda(u(t_{j\theta}); \widetilde{u}_{j\theta}, \xi_{j\theta} + \partial_t\xi_j) + A_\lambda(U_{j\theta}; \widetilde{u}_{j\theta}, \xi_{j\theta} + \partial_t\xi_j) \\
&+ B_\lambda(u(t_{j\theta}); \Delta t\rho_{j\theta}, \xi_{j\theta} + \partial_t\xi_j) \\
&- B_\lambda(u(t_{j\theta}); \partial_t\widetilde{u}_j, \xi_{j\theta} + \partial_t\xi_j) + B_\lambda(U_{j\theta}; \partial_t\widetilde{u}_j, \xi_{j\theta} + \partial_t\xi_j) \\
&+ (f(u(t_{j\theta})) - f(U_{j\theta}), \xi_{j\theta} + \partial_t\xi_j) \\
&+ \lambda(\eta(t_{j\theta}) + \xi_{j\theta}, \xi_{j\theta} + \partial_t\xi_j) \\
&+ \lambda(\widetilde{u}(t_{j\theta}) - \widetilde{u}_{j\theta}, \xi_{j\theta} + \partial_t\xi_j) \\
&+ \lambda(\eta_t(t_{j\theta}) + \Delta t\rho_{j\theta} + \partial_t\xi_j, \xi_{j\theta} + \partial_t\xi_j) \\
&+ \sum_{k=1}^{P_h}\int_{e_k}\{(a(U_{j\theta}) + b(U_{j\theta}))\nabla\xi_{j\theta}\cdot n_k\}[\partial_t\xi_j] \\
&+ \sum_{k=1}^{P_h}\int_{e_k}\{(a(U_{j\theta}) + b(U_{j\theta}))\nabla(\partial_t\xi_j)\cdot n_k\}[\xi_{j\theta}] \\
=\ & \sum_{i=1}^{11} I_i.
\end{aligned} \tag{4.16}
$$

For sufficiently small $\varepsilon > 0$ there exists a constant $C > 0$ such that

$$
\begin{aligned}
|I_1| &\le (\|\eta_t(t_{j\theta})\|_0 + \|\Delta t\rho_{j\theta}\|_0)(\|\xi_{j\theta}\|_0 + \|\partial_t\xi_j\|_0) \\
&\le C\Big(\|\eta_t(t_{j\theta})\|_0^2 + (\Delta t)^2\|\rho_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2\Big) + \varepsilon\|\partial_t\xi_j\|_0^2.
\end{aligned}
$$

From Lemma 4.2 I_2 can be estimated as follows

$$
|I_2| \le C\|r_{j\theta}\|_1\|\xi_{j\theta} + \partial_t\xi_j\|_1 \le C(\Delta t)^4 + \varepsilon\|\xi_{j\theta}\|_1^2 + \varepsilon\|\partial_t\xi_j\|_1^2.
$$

Now we can separate I_3 as follows

$$
\begin{aligned}
I_3 &= A_\lambda(a(U_{j\theta}); \widetilde{u}_{j\theta}; \xi_{j\theta} + \partial_t\xi_j) - A_\lambda(u(t_{j\theta}); \widetilde{u}_{j\theta}, \xi_{j\theta} + \partial_t\xi_j) \\
&= \Big((a(U_{j\theta}) - a(u(t_{j\theta})))\nabla\widetilde{u}_{j\theta}, \nabla(\xi_{j\theta} + \partial_t\xi_j)\Big) \\
&\quad - \sum_{k=1}^{P_h}\int_{e_k}\{(a(U_{j\theta}) - a(u(t_{j\theta})))\nabla\widetilde{u}_{j\theta}\cdot n_k\}[\xi_{j\theta} + \partial_t\xi_j] \\
&\quad - \sum_{k=1}^{P_h}\int_{e_k}\{(a(U_{j\theta}) - a(u(t_{j\theta})))\nabla(\xi_{j\theta} + \partial_t\xi_j)\cdot n_k\}[\widetilde{u}_{j\theta}]
\end{aligned}
$$

$$= \sum_{l=1}^{3} I_{3l}.$$

Now we estimate I_{31}:

$$\begin{aligned}
I_{31} &\le C\|\nabla\widetilde{u}_{j\theta}\|_{L^\infty}(\|\eta(t_{j\theta})\|_0 + \|r_{j\theta}\|_0 + \|\xi_{j\theta}\|_0)(\|\nabla\xi_{j\theta}\|_0 + \|\nabla(\partial_t\xi_j)\|_0)\\
&\le C\Big(\|\eta(t_{j\theta})\|_0^2 + \|r_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2 + \|\nabla\xi_{j\theta}\|_0^2\Big) + \varepsilon\|\partial_t\xi_j\|_1^2.
\end{aligned}$$

We apply Lemma 3.2 to estimate I_{32} as follows

$$\begin{aligned}
I_{32} \le& \sum_{k=1}^{P_h}\|\nabla\widetilde{u}_{j\theta}\|_{\infty,e_k}(\|\eta(t_{j\theta})\|_{0,e_k} + \|r_{j\theta}\|_{0,e_k} + \|\xi_{j\theta}\|_{0,e_k})(\|[\xi_{j\theta}]\|_{0,e_k}\\
&+ \|[\partial_t\xi_j]\|_{0,e_k})\\
\le& C\sum_{i=1}^{N_h}\|\nabla\widetilde{u}_{j\theta}\|_{\infty,E_i}\Big(h_i^{-1/2}\|\eta(t_{j\theta})\|_{0,E_i} + h_i^{1/2}\|\nabla\eta(t_{j\theta})\|_{0,E_i}\\
&+ h^{-1/2}\|r_{j\theta}\|_{0,E_i} + h^{1/2}\|\nabla r_{j\theta}\|_{0,E_i} + h^{-1/2}\|\xi_{j\theta}\|_{0,E_i}\Big)\\
&\cdot h^{\frac{\beta(d-1)}{2}}\Big(J(\xi_{j\theta},\xi_{j\theta})^{1/2} + J(\partial_t\xi_j,\partial_t\xi_j)^{1/2}\Big)\\
\le& C\Big(\|\eta(t_{j\theta})\|_0^2 + h^2\|\nabla\eta(t_{j\theta})\|_0^2 + \|r_{j\theta}\|_0^2 + h^2\|\nabla r_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2\Big)\\
&+ \varepsilon\|\xi_{j\theta}\|_1^2 + \varepsilon\|\partial_t\xi_j\|_1^2.
\end{aligned}$$

From the approximation of η in Theorem 4.1

$$\begin{aligned}
I_{33} \le& C\sum_{k=1}^{P_h}\Big(\|\nabla(\xi_{j\theta})\|_{\infty,e_k} + \|\nabla(\partial_t\xi_j)\|_{\infty,e_k}\Big)\Big(\|\eta(t_{j\theta})\|_{0,e_k} + \|r_{j\theta}\|_{0,e_k}\\
&+ \|\xi_{j\theta}\|_{0,e_k}\Big)\|[\eta_{j\theta}]\|_{0,e_k}\\
\le& C\sum_{i=1}^{N_h}\Big(\|\nabla\xi_{j\theta}\|_{\infty,E_i}+\|\nabla(\partial_t\xi_j)\|_{\infty,E_i}\Big)h^{-1/2}\Big(\|\eta(t_{j\theta})\|_{0,E_i}+h\|\nabla\eta(t_{j\theta})\|_{0,E_i}\\
&+ \|r_{j\theta}\|_{0,E_i} + h\|\nabla r_{j\theta}\|_{0,E_i} + \|\xi_{j\theta}\|_{0,E_i}\Big)h^{-1/2}\Big(\|[\eta(t_{j\theta})]\|_{0,E_i}\\
&+ h\|[\nabla\eta(t_{j\theta})]\|_{0,E_i}\Big)\\
\le& Ch^{-\frac{d}{2}-1}\sum_{i=1}^{N_h}\Big(\|\nabla\xi_{j\theta}\|_{0,E_i}+\|\nabla(\partial_t\xi_j)\|_{0,E_i}\Big)\Big(\|\eta(t_{j\theta})\|_{0,E_i}+h\|\nabla\eta(t_{j\theta})\|_{0,E_i}\\
&+ \|r_{j\theta}\|_{0,E_i} + \|\xi_{j\theta}\|_{0,E_i}\Big)\cdot h^{\frac{d}{2}+1}(\|u_t\|_{L^\infty(H^{\frac{d}{2}+1})} + \|u_0\|_{\frac{d}{2}+1})\\
\le& C\Big(\|\eta(t_{j\theta})\|_0^2 + h^2\|\nabla\eta(t_{j\theta})\|_0^2 + \|r_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2 + \|\nabla\xi_{j\theta}\|_0^2\Big) + \varepsilon\|\partial_t\xi_j\|_1^2.
\end{aligned}$$

Therefore, we get

$$|I_3| \le C\Big(\|\eta(t_{j\theta})\|_0^2 + h^2\|\nabla\eta(t_{j\theta})\|_0^2 + \|r_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2 + \|\nabla\xi_{j\theta}\|_0^2\Big) + \varepsilon\|\xi_{j\theta}\|_1^2 + 3\varepsilon\|\partial_t\xi_j\|_1^2.$$

From the definition of I_4, we obtain

$$\begin{aligned}|I_4| &\le C\|\Delta t\rho_{j\theta}\|_1(\|\xi_{j\theta}\|_1 + \|\partial_t\xi_j\|_1)\\ &\le C(\Delta t)^2\|\rho_{j\theta}\|_1^2 + \varepsilon\|\xi_{j\theta}\|_1^2 + \varepsilon\|\partial_t\xi_j\|_1^2.\end{aligned}$$

From the definition of I_5, we can separated I_5 as follows

$$\begin{aligned}I_5 &= B_\lambda(U_{j\theta};\partial_t\widetilde{u}_j, \xi_{j\theta}+\partial_t\xi_j) - B_\lambda(u(t_{j\theta});\partial_t\widetilde{u}_j, \xi_{j\theta}+\partial_t\xi_j)\\ &= \Big((b(U_{j\theta}) - b(u(t_{j\theta})))\nabla(\partial_t\widetilde{u}_j), \nabla(\xi_{j\theta}+\partial_t\xi_j)\Big)\\ &\quad - \sum_{k=1}^{P_h}\int_{e_k}\{(b(U_{j\theta}) - b(u(t_{j\theta})))\nabla(\partial_t\widetilde{u}_j)\cdot n_k\}[\xi_{j\theta}+\partial_t\xi_j]\\ &\quad - \sum_{k=1}^{P_h}\int_{e_k}\{(b(U_{j\theta}) - b(u(t_{j\theta})))\nabla(\xi_{j\theta}+\partial_t\xi_j)\cdot n_k\}[\partial_t\widetilde{u}_j]\\ &= \sum_{j=1}^{3} I_{5j}.\end{aligned}$$

I_{51} is estimated in the following way

$$\begin{aligned}I_{51} &\le C\|\nabla(\partial_t\widetilde{u}_j)\|_\infty\Big(\|\eta(t_{j\theta})\|_0 + \|r_{j\theta}\|_0 + \|\xi_{j\theta}\|_0\Big)\Big(\|\nabla\xi_{j\theta}\|_0 + \|\nabla(\partial_t\xi_j)\|_0\Big)\\ &\le C\Big(\|\eta(t_{j\theta})\|_0^2 + \|r_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2 + \|\nabla\xi_{j\theta}\|_0^2\Big) + \varepsilon\|\partial_t\xi_j\|_1^2,\end{aligned}$$

because

$$\begin{aligned}\|\nabla(\partial_t\widetilde{u}_j)\|_{L^\infty} &= \left\|\frac{\nabla\widetilde{u}_{j+1} - \nabla\widetilde{u}_j}{\Delta t}\right\|_{L^\infty} = \left\|\frac{1}{\Delta t}\int_{t_j}^{t_{j+1}}\nabla\widetilde{u}_t(\tau)d\tau\right\|_{L^\infty}\\ &\le \|\nabla\widetilde{u}_t\|_{L^\infty} \le C\|\eta_t\|_{\frac{d}{2}+1+\delta} + \|\nabla u_t\|_{L^\infty} \le C,\end{aligned}$$

provided that $u_t \in L^\infty(H^{\frac{d}{2}+1+\delta})$ and $\nabla u_t \in L^\infty$.

Similarly, there exists $C > 0$ such that

$$\begin{aligned}I_{52} &\le C\sum_{k=1}^{P_h}\|\nabla(\partial_t\widetilde{u}_j)\|_{\infty,e_k}\Big(\|\eta(t_{j\theta})\|_{0,e_k} + \|r_{j\theta}\|_{0,e_k} + \|\xi_{j\theta}\|_{0,e_k}\Big)\Big(\|[\xi_{j\theta}]\|_{0,e_k}\\ &\quad + \|[\partial_t\xi_j]\|_{0,e_k}\Big)\\ &\le C\sum_{i=1}^{N_h}\|\nabla(\partial_t\widetilde{u}_j)\|_{\infty,E_i}h_i^{-1/2}\Big(\|\eta(t_{j\theta})\|_{0,E_i} + h_i\|\nabla\eta(t_{j\theta})\|_{0,E_i} + \|r_{j\theta}\|_{0,E_i}\end{aligned}$$

$$+ \|\xi_{j\theta}\|_{0,E_i}\Big) h_i^{\frac{\beta(d-1)}{2}} \Big(J(\xi_{j\theta}, \xi_{j\theta})^{1/2} + J(\partial_t \xi_j, \partial_t \xi_j)^{1/2} \Big)$$
$$\le C\Big(\|\eta(t_{j\theta})\|_0^2 + h^2 \|\nabla \eta(t_{j\theta})\|_0^2 + \|r_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2 \Big) + \varepsilon \|\xi_{j\theta}\|_1^2 + \varepsilon \|\partial_t \xi_j\|_1^2.$$

By applying the trace inequality and Lemma 4.1, we have

$$I_{53} \le C \sum_{k=1}^{P_h} \Big(\|\nabla \xi_{j\theta}\|_{\infty,e_k} + \|\nabla(\partial_t \xi_j)\|_{\infty,e_k} \Big) \Big(\|\eta(t_{j\theta})\|_{0,e_k} + \|r_{j\theta}\|_{0,e_k} + \|\xi_{j\theta}\|_{0,e_k} \Big)$$
$$\|[\partial_t \widetilde{u}_j - u_t(t_{j\theta})]\|_{0,e_k}$$
$$\le C \sum_{i=1}^{N_h} h_i^{-\frac{d}{2}} \Big(\|\nabla \xi_{j\theta}\|_{0,E_i} + \|\nabla(\partial_t \xi_j)\|_{0,E_i} \Big) h_i^{-\frac{1}{2}} \Big(\|\eta(t_{j\theta})\|_{0,E_i} + h \|\nabla \eta(t_{j\theta})\|_{0,E_i}$$
$$+ \|r_{j\theta}\|_{0,E_i} + \|\xi_{j\theta}\|_{0,E_i} \Big) \Big[h_i^{-\frac{1}{2}} (\|\eta_t(t_{j\theta})\|_{0,E_i} + h \|\nabla \eta_t(t_{j\theta})\|_{0,E_i})$$
$$+ h_i^{\frac{\beta(d-1)}{2}} \Delta t \|\rho_{j\theta}\|_{1,E_i} \Big].$$

Since

$\|\eta_t(t_{j\theta})\|_{0,E_i} + h\|\nabla \eta_t(t_{j\theta})\|_{0,E_i} \le C h^{\frac{d}{2}+1} \|u_t\|_{L^\infty(H^{\frac{d}{2}+1})}$ and $h^{-\frac{d}{2}} \Delta t \|\rho_{j\theta}\|_1 \le C$,

$$I_{53} \le C\Big(\|\eta(t_{j\theta})\|_0^2 + h^2 \|\nabla \eta(t_{j\theta})\|_0^2 + \|r_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2 \Big) + \varepsilon \|\xi_{j\theta}\|_1^2 + \varepsilon \|\partial_t \xi_{j\theta}\|_1^2.$$

From the estimations of I_{5i}, $1 \le i \le 3$, we have

$$|I_5| \le C(\|\eta(t_{j\theta})\|_0^2 + h^2 \|\nabla \eta(t_{j\theta})\|_0^2 + \|r_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2 + \|\nabla \xi_{j\theta}\|_0^2)$$
$$+ 2\varepsilon \|\xi_{j\theta}\|_1^2 + 3\varepsilon \|\partial_t \xi_j\|_1^2.$$

We obtain the following estimates of I_i for each $6 \le i \le 11$.

$$|I_6| \le C(\|\eta(t_{j\theta})\|_0 + \|\xi_{j\theta}\|_0 + \|r_{j\theta}\|_0)(\|\xi_{j\theta}\|_0 + \|\partial_t \xi_j\|_0)$$
$$\le C\Big(\|\eta(t_{j\theta})\|_0^2 + \|r_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2 \Big) + \varepsilon \|\partial_t \xi_j\|_0^2,$$
$$|I_7| \le C\Big(\|\eta(t_{j\theta})\|_0^2 + \|\xi_{j\theta}\|_0^2 \Big) + \varepsilon \|\partial_t \xi_j\|_0^2,$$
$$|I_8| \le C\|r_{j\theta}\|_0(\|\xi_{j\theta}\|_0 + \|\partial_t \xi_j\|_0) \le C\|r_{j\theta}\|_0^2 + \varepsilon \|\xi_{j\theta}\|_0^2 + \varepsilon \|\partial_t \xi_j\|_0^2,$$
$$|I_9| \le \lambda(\|\eta_t(t_{j\theta})\|_0 + \|\Delta t \rho_{j\theta}\|_0 + \|\partial_t \xi_j\|_0)(\|\xi_{j\theta}\|_0 + \|\partial_t \xi_j\|_0)$$
$$\le \lambda\Big(\|\eta_t(t_{j\theta})\|_0^2 + (\Delta t)^2 \|\rho_{j\theta}\|_0^2 + \frac{3}{2}\|\xi_{j\theta}\|_0^2 \Big) + \frac{5}{2}\lambda \|\partial_t \xi_j\|_0^2,$$
$$|I_{10}| = \sum_{k=1}^{P_h} \int_{e_k} \{(a(U_{j\theta}) + b(U_{j\theta}))\nabla \xi_{j\theta} \cdot n_k\}[\partial_t \xi_j] \le C\|\nabla \xi_{j\theta}\|_0^2 + \varepsilon \|\partial_t \xi_j\|_1^2,$$
$$|I_{11}| \le CJ(\xi_{j\theta}, \xi_{j\theta}) + \varepsilon \|\partial_t \xi_j\|_1^2 \le C(J(\xi_{j+1}, \xi_{j+1}) + J(\xi_j, \xi_j)) + \varepsilon \|\partial_t \xi_j\|_1^2.$$

Substituting the estimates of I_i $(1 \le i \le 11)$ into (4.16), we get

$$(4.17) \qquad \frac{1}{2\Delta t}\Big[(1+2\lambda)(\|\xi_{j+1}\|_0^2 - \|\xi_j\|_0^2) + 2(J_\beta^\sigma(\xi_{j+1}, \xi_{j+1}) - J_\beta^\sigma(\xi_j, \xi_j))$$

$$
\begin{aligned}
&+ 2a_0(\|\nabla\xi_{j+1}\|_0^2 - \|\nabla\xi_j\|_0^2)\Big] + \frac{1}{2}\|\partial_t\xi_j\|_0^2 + \frac{\underset{\sim}{c}}{2}\|\xi_{j\theta}\|_1^2 + \frac{\underset{\sim}{c}}{2}\|\partial_t\xi_j\|_1^2 \\
&\le C\Big[\|\eta_t(t_{j\theta})\|_0^2 + (\Delta t)^2\|\rho_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2 + (\Delta t)^4 + \|\eta(t_{j\theta})\|_0^2 \\
&\quad + h^2\|\nabla\eta(t_{j\theta})\|_0^2 + \|r_{j\theta}\|_0^2 + \|\nabla\xi_{j\theta}\|_0^2 + (\Delta t)^2\|\rho_{j\theta}\|_1^2 \\
&\quad + J_\beta^\sigma(\xi_{j+1},\xi_{j+1}) + J_\beta^\sigma(\xi_j,\xi_j)\Big].
\end{aligned}
$$

If we sum both sides of (4.17) from $j = 0$ to $N-1$, then we obtain

$$
\begin{aligned}
&\|\xi_N\|_0^2 - \|\xi_0\|_0^2 + \frac{2}{3}J(\xi_N,\xi_N) - 2J(\xi_0,\xi_0) + a_0\Delta t\Big(\frac{4}{3}\|\nabla\xi_N\|_0^2 - 4\|\nabla\xi_0\|_0^2\Big) \\
&+ \frac{2}{3}\Delta t\sum_{j=0}^{N-1}(\|\partial_t\xi_j\|_0^2 + \underset{\sim}{c}\|\xi_{j\theta}\|_1^2 + \underset{\sim}{c}\|\partial_t\xi_j\|_1^2) \\
&\le C(\Delta t)\sum_{j=0}^{N-1}\Big(\|\eta_t(t_{j\theta})\|_0^2 + (\Delta t)^2\|\rho_{j\theta}\|_0^2 + \|\xi_{j\theta}\|_0^2 + (\Delta t)^4 + \|\eta(t_{j\theta})\|_0^2 \\
&\quad + h^2\|\nabla\eta(t_{j\theta})\|_0^2 + \|r_{j\theta}\|_0^2 + \|\nabla\xi_{j\theta}\|_0^2 + J_\beta^\sigma(\xi_j,\xi_j) + J_\beta^\sigma(\xi_{j+1},\xi_{j+1})\Big),
\end{aligned}
$$

which implies

$$
\begin{aligned}
&\|\xi_N\|_0^2 + \frac{2}{3}J(\xi_N,\xi_N) + \frac{4}{3}a_0\Delta t\|\nabla\xi_N\|_0^2 + \frac{2\underset{\sim}{c}}{3}\Delta t\sum_{j=0}^{N-1}(\|\partial_t\xi_j\|_0^2 + \|\xi_{j\theta}\|_1^2 + \|\partial_t\xi_j\|_1^2) \\
&\le \Big(\|\xi_0\|_0^2 + 2a_0J(\xi_0,\xi_0) + 4a_0\Delta t\|\nabla\xi_0\|_0^2\Big) \\
&\quad + C(\Delta t)\sum_{j=0}^{N-1}\Big(\|\eta_t(t_{j\theta})\|_0^2 + (\Delta t)^2\|\rho_{j\theta}\|_0^2 + \|\eta(t_{j\theta})\|_0^2 + h^2\|\nabla\eta(t_{j\theta})\|_0^2\Big) \\
&\quad + C\Delta t\sum_{j=0}^{N}\Big(\|\nabla\xi_j\|_0^2 + (\Delta t)^4 + \|r_j\|_0^2\Big) \\
&\quad + C(\Delta t)\sum_{j=0}^{N-1}\Big(\|\nabla\xi_{j\theta}\|_0^2 + J_\beta^\sigma(\xi_j,\xi_j) + J_\beta^\sigma(\xi_{j+1},\xi_{j+1})\Big),
\end{aligned}
$$

where Δt is sufficiently small.

By the discrete version of Gronwall's lemma, we get

$$
\begin{aligned}
&\|\xi_N\|_0^2 + J(\xi_N,\xi_N) + \Delta t\|\nabla\xi_N\|_0^2 + \Delta t\sum_{j=0}^{N-1}\Big(\|\partial_t\xi_j\|_0^2 + \|\xi_{j\theta}\|_1^2 + \|\partial_t\xi_j\|_1^2\Big) \\
&\le C\Big\{\|\xi_0\|^2 + J(\xi_0,\xi_0) + \Delta t\|\nabla\xi_0\|_0^2 + \Delta t\sum_{j=0}^{N-1}\Big(\|\eta_t(t_{j\theta})\|_0^2 + (\Delta t)^2\|\rho_{j\theta}\|_0^2
\end{aligned}
$$

$$+ \|\eta(t_{j\theta})\|_0^2 + h^2\|\nabla\eta(t_{j\theta})\|_0^2\Big) + \Delta t \sum_{j=0}^{N} \Big((\Delta t)^4 + \|r_j\|_0^2\Big)\Big\}.$$

Now we choose an appropriate initial approximation $U_0(x)$ for example $U_0(x) = \widetilde{u}(x,0)$, to get

$$\|\xi_N\|_0^2 + J(\xi_N,\xi_N) + \|\nabla\xi_N\|_0^2 + \Delta t \sum_{j=0}^{N-1} (\|\partial_t\xi_j\|_0^2 + \|\xi_{j\theta}\|_1^2 + \|\partial_t\xi_j\|_1^2)$$

$$\leq C(\Delta t)h^{2\mu} \sum_{j=0}^{N-1} (\|u_0\|_s^2 + \|u_t\|_{L^2(H^s)}^2 + (\Delta t)^2\|\rho_{j\theta}\|_0^2) + C(\Delta t)\sum_{j=0}^{N}[(\Delta t)^4 + \|r_j\|_0^2],$$

where $\mu = \min(r+1, s)$.
If $\theta \in (0,1]$,

$$\|\xi_N\|_0^2 + J(\xi_N,\xi_N) + \|\nabla\xi_N\|_0^2 + \Delta t \sum_{j=0}^{N-1} (\|\partial_t\xi_j\|_0^2 + \|\xi_{j\theta}\|_1^2 + \|\partial_t\xi_j\|_1^2)$$

$$\leq C(\Delta t)h^{2\mu}\Big(\sum_{j=0}^{N}(\|u_0\|_s^2 + \|u_t\|_{L^2(H^s)}^2 + \|u_{tt}\|_{L^\infty((t_j,t_{j+1});L_2)})\Big)$$

$$+ C(\Delta t)^2\|u_{tt}\|_{L^\infty(H^2)}^2.$$

If $\theta = 0$,

$$\|\xi_N\|_0^2 + J(\xi_N,\xi_N) + \|\nabla\xi_N\|_0^2 + \Delta t \sum_{j=0}^{N-1} (\|\partial_t\xi_j\|_0^2 + \|\xi_{j\theta}\|_1^2 + \|\partial_t\xi_j\|_1^2)$$

$$\leq C(\Delta t)\Big[h^{2\mu}\sum_{j=0}^{N}\Big(\|u_0\|_s^2 + \|u_t\|_{L^2(H^s)}^2 + (\Delta t)^4\|u_{ttt}\|_{L^\infty((t_j,t_{j+1});L_2)}$$

$$+ (\Delta t)^4\|u_{tt}\|_{(L^\infty(t_j,t_{j+1});H^2)}^2\Big)\Big].$$

Therefore if $\theta \in (0,1]$ we have

$$\|e\|_{\ell^\infty(L^2)} \leq C(h^\mu + \Delta t)(\|u_0\|_s + \|u_t\|_{L^\infty(H^s)} + \|\nabla u_t\|_{L^\infty} + \|u_{tt}\|_{L^\infty(H^1)})$$

and if $\theta = 0$,

$$\|e\|_{\ell^\infty(L^2)} \leq C(h^\mu + (\Delta t)^2)(\|u_0\|_s + \|u_t\|_{L^\infty(H^s)} + \|\nabla u_t\|_{L^\infty} + \|u_{tt}\|_{L^\infty(H^1)}$$
$$+ \|u_{ttt}\|_{L^\infty(H^1)}).$$

□

References

[1] D. N. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal. **19** (1982), no. 4, 724–760.

[2] D. N. Arnold, J. Jr. Douglas, and V. Thomée, *Superconvergence of a finite approximation to the solution of a Sobolev equation in a single space variable*, Math. Comp. **36** (1981), 53–63.
[3] I. Babuška and M. Suri, *The h-p version of the finite element method with quasi-uniform meshes*, RAIRO Modél. Math. Anal. Numer. **21** (1987), no. 2, 199–238.
[4] ______, *The optimal convergence rates of the p-version of the finite element method*, SIAM J. Numer. Anal. **24** (1987), no. 4, 750–776.
[5] G. I. Barenblatt, I. P. Zheltov, and I. N. Kochina, *Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks*, J. Appl. Math. Mech. **23** (1960), 1286–1303.
[6] R. W. Carroll and R. E. Showalter, *Singular and Degenerate Cauchy Problems*, Mathematics in Sciences and Engineering, Vol. 127, Academic Press, New York, 1976.
[7] P. J. Chen and M. E. Gurtin, *On a theory of heat conduction involving two temperatures*, Z. Angew Math. Phys. (1968), 614–627.
[8] P. L. Davis, *A quasilinear parabolic and a related third order problem*, J. Math. Anal. Appl. **40** (1972), 327–335.
[9] R. E. Ewing, *The approximation of certain parabolic equations backward in time by Sobolev equations*, SIAM J. Math. Anal. **6** (1975), 283–294.
[10] ______, *Time-stepping Galerkin methods for nonlinear Sobolev partial differential equations*, SIAM J. Numer. Anal. **15** (1978), no. 6, 1125–1150.
[11] W. H. Ford and T. W. Ting, *Stability and convergence of difference approximations to pseudo-parabolic partial differential equations*, Math. Comput. **27** (1973), 737–743.
[12] ______, *Uniform error estimates for difference approximations to nonlinear pseudo-parabolic partial differential equations*, SIAM J. Numer. Anal. **11** (1974), 155–169.
[13] Y. Lin, *Galerkin methods for nonlinear Sobolev equations*, Aequationes Math. **40** (1990), no. 1, 54–66.
[14] Y. Lin and T. Zhang, *Finite element methods for nonlinear Sobolev equations with nonlinear boundary conditions*, J. Math. Anal. Appl. **165** (1992), no. 1, 180–191.
[15] M. T. Nakao, *Error estimates of a Galerkin method for some nonlinear Sobolev equations in one space dimension*, Numer. Math. **47** (1985), no. 1, 139–157.
[16] M. R. Ohm, H. Y. Lee, and J. Y. Shin, *Error estimates for discontinuous Galerkin method for nonlinear parabolic equations*, J. Math. Anal. Appl. **315** (2006), no. 1, 132–143.
[17] ______, *L^2-error analysis of discontinuous Galerkin approximations for nonlinear Sobolev equations*, submitted.
[18] T. Sun and D. Yang, *A priori error estimates for interior penalty discontinuous Galerkin method applied to nonlinear Sobolev equations*, Appl. Math. Comput. **200** (2008), no. 1, 147–159.
[19] ______, *Error estimates for a discontinuous Galerkin method with interior penalties applied to nonlinear Sobolev equations*, Numer. Methods Partial Differential Equations 24 (2008), no. 3, 879–896.
[20] T. W. Ting, *A cooling process according to two-temperature theory of heat conduction*, J. Math. Anal. Appl. **45** (1974), 289–303.
[21] M. F. Wheeler, *A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations*, SIAM J. Numer. Anal. **10** (1973), 723–759.

Mi Ray Ohm
Division of Information Systems Engineering
Dongseo University
Busan 617-716, Korea
E-mail address: mrohm@dongseo.ac.kr

Hyun Young Lee
Department of Mathematics
Kyungsung University
Busan 608-736, Korea
E-mail address: hylee@ks.ac.kr