DISCONTINUOUS GALERKIN DISCRETIZATION IN TIME OF SYSTEMS OF SECOND-ORDER NONLINEAR HYPERBOLIC EQUATIONS

AILI SHAO*

Abstract. In this paper we study the finite element approximation of systems of second-order nonlinear hyperbolic equations. The proposed numerical method combines a $h$-$p$-version discontinuous Galerkin finite element approximation in the time direction with an $H^1(\Omega)$-conforming finite element approximation in the spatial variables. Error bounds at the temporal nodal points are derived under a weak restriction on the temporal step size in terms of the spatial mesh size. Numerical experiments are presented to verify the theoretical results.

Mathematics Subject Classification. 65M60, 65M12, 35L72, 35L53.

Received December 17, 2021. Accepted September 5, 2022.

1. Introduction

This paper aims to show how a discontinuous Galerkin time-stepping method can be used to approximate solutions of second-order quasilinear hyperbolic systems, which arise in a range of relevant applications, namely elastodynamics and general relativity. There has been a substantial body of research devoted to both the theoretical and numerical analysis of solutions of second-order hyperbolic equations. In particular, Kato [1] established the existence of solutions to the initial-boundary-value problem for quasilinear hyperbolic equations using semigroup theory. Building on Kato’s work [1], Hughes et al. [2] analyzed the existence, uniqueness and well-posedness for a more general class of quasilinear second-order hyperbolic systems on a short time interval. They also applied these results to elastodynamics and Einstein’s equations for the Lorentz metric $g_{\alpha,\beta}$ on $\mathbb{R}^4$, $0 \leq \alpha, \beta \leq 3$. In contrast to the semigroup approach, Dafermos and Hrusa [3] used energy methods to establish local in time existence of smooth solutions to initial-boundary-value problems for such hyperbolic systems on a bounded domain $\Omega \subseteq \mathbb{R}^d$ where $d = 1, 2, 3$. In the case of $d = 3$, Chen and Von Wahl proved an existence theorem for similar initial-boundary-value problems in [4]. Concerning numerical approximations of second-order hyperbolic equations, fully discrete schemes based on Galerkin finite element approximations in space for the linear case can be found in [5–7]. These time-discrete schemes are generated from the rational approximations to either the cosine or the exponential and were later generalized by Bales and Dougalis [8, 9] to approximate nonlinear hyperbolic problems. In [9], Bales considered a scalar nonlinear wave equation and introduced a class of single-step fully discrete schemes, which have temporal accuracy up to fourth-order. Dupont [10] and Dendy
For hyperelastic materials, $S$ is the gradient of a scalar-valued ‘stored energy function’. Hence, if
\[ A_{i\alpha j\beta} := \frac{\partial}{\partial \eta_{j\beta}} S_{i\alpha} (\eta), \quad \eta \in \mathbb{R}^{d\times d}, \]
the elasticities $A_{i\alpha j\beta}$ satisfy
\[ A_{i\alpha j\beta} = A_{j\beta i\alpha}, 1 \leq i, \alpha, j, \beta \leq d. \quad \text{(S1a)} \]
We assume that $A_{i\alpha j\beta}$ satisfy the strict Legendre–Hadamard condition
\[ \sum_{i,\alpha,\beta,j=1}^{d} A_{i\alpha j\beta}(\eta) \zeta_{\alpha} \zeta_{\beta} \xi_{i} \xi_{j} \geq M_0 |\zeta|^2 |\xi|^2 \quad \text{for all } \eta \in \mathcal{O} \text{ and } \zeta, \xi \in \mathbb{R}^{d}, \quad \text{(S1b)} \]
for some real number $M_0 > 0$, where $\mathcal{O}$ is the domain of definition of $A_{i\alpha j\beta}$ and here $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^{d}$. This condition (S1b) is indeed satisfied by the constitutive relations of the standard material models on sizeable portions of the displacement gradient space [3].

The initial-boundary-value problem (1)–(3) does not have a global smooth solution as a result of breaking waves and shocks no matter how smooth $u_0$, $u_1$ and $\mathbf{f}$ are. It was proved by Dafermos and Hrusa [3] that there exists a unique local solution to the problem (1)–(3) provided that (S1a,b) are satisfied. We summarize this existence result in the following theorem.

\begin{theorem}
Let $\Omega$ be a bounded domain in $\mathbb{R}^{d}$ with smooth boundary $\partial \Omega$. Assume that (S1a,b) hold, that $A_{i\alpha j\beta}$ and $\mathbf{f}$ are sufficiently smooth, and that $u_0 \in [H^{m}(\Omega)]^{d}$ and $u_1 \in [H^{m-1}(\Omega)]^{d}$ for some integer $m \geq \left\lceil \frac{d}{2} \right\rceil + 3$.
\end{theorem}
Assume further that the initial values of the time derivatives of $u$ up to order $m - 1$ vanish on $\partial \Omega$ and that $\nabla u_0|_{\partial \Omega} \subset \mathcal{O}$. Then, there exists a finite time $T > 0$ for which (1)–(3) has a unique solution $u$ such that

$$u \in \bigcap_{s=0}^{m} C^{m-s}\left([0, T]; [H^s(\Omega)]^d\right).$$

By the Sobolev embedding theorem, (4) implies that

$$u \in \left[C^\beta([0, T] \times \overline{\Omega})^d\right] := \bigcap_{s=0}^{\beta} C^{\beta-s}\left([0, T]; C^s(\overline{\Omega})^d\right),$$

where $\beta = m - \left[\frac{d}{2}\right] - 1$. Note that the assumption on $m$ implies that $\beta \geq 2$. We shall assume throughout the paper that the above assumptions are satisfied for $m$ sufficiently large so that a unique solution of (1) exists. In fact, by Theorem 1.1, this unique solution satisfies

$$u \in \left.C^1\left([0, T], [H^{m-1}(\Omega)]^d\right) \cap C\left([0, T], [H^m(\Omega)]^d\right)\right).$$

There have been a few contributions to the literature discussing the numerical approximations of (1)–(3). In particular, Makridakis [12] proved an optimal $L^2$ error estimate for the classical conforming method under the assumption that $p > \frac{d}{2}$, where $p$ is the polynomial degree in space, and for appropriate initial approximations. In [13], Suli and Ortner showed an optimal error estimate based on the broken $H^1$ norm for the semi-discrete discontinuous Galerkin finite element approximations of a similar problem, but with mixed Dirichlet-Neumann boundary conditions and weaker Lipschitz assumptions on the nonlinear term. Following [8, 9], Makridakis also proposed and analyzed two different fully discrete schemes to approximate (1)–(3). The first scheme is based on second-order accurate approximations of the cosine while the second group of fully discrete methods that have temporal order of accuracy up to fourth-order are based on rational approximations of the exponential function. Discontinuous Galerkin in time methods, which are the focus of this paper are, however, arbitrarily high-order accurate. In contrast with the above-mentioned finite difference time integration schemes, for which the solution at the current time step depends on the previous steps, this discontinuous-in-time scheme on the time interval $(t_n, t_{n+1}]$ only depends on the solution at $t_n$. Since the local polynomial degree is free to vary between time steps, this method is also naturally suited for an adaptive choice of the time discretization parameters. To the best of our knowledge, the analysis of a fully discrete scheme based on such discontinuous-in-time discretization of second-order quasilinear hyperbolic systems has not been previously considered in the literature.

Discontinuous Galerkin methods [17, 18] have been widely and successfully used for the numerical approximations of PDEs. They have been first introduced by Reed and Hill [17] to solve the hyperbolic neutron transport equations. Simultaneously, but independently, they were proposed as non-standard numerical schemes for solving elliptic and parabolic problems by Babuška & Zlámal [19], Baker [20], Wheeler [21], Arnold [22] and Rivière [23] etc. In recent years, there has been considerable interest in applying discontinuous Galerkin finite element methods to nonlinear hyperbolic PDEs. In particular, Antonietti et al. [24] developed a high-order discontinuous Galerkin scheme for the spatial discretization of nonlinear acoustic waves. Muhr et al. [25] also proposed and analyzed a hybrid discontinuous Galerkin coupling approach for the semi-discrete nonlinear elasto-acoustic problem. However, there has been little work on the construction and mathematical analysis of fully discrete discontinuous Galerkin schemes for second-order nonlinear hyperbolic PDEs. In this article, a high-order discontinuous Galerkin finite element method for the time integration will be proposed and analyzed. To construct such a scheme, we first discretize with respect to the spatial variables by the means of a Galerkin finite element method, which results in a system of ordinary differential equations (ODEs) in time; then we discretize the resulting ODE system using discontinuous Galerkin method in time (e.g. see [26]). The resulting weak formulation in time is based on weakly imposing the continuity of the approximate displacements and velocities between time steps by penalizing jumps in these quantities in the definition of the numerical method.
The paper is structured as follows. The next section sets up the assumptions required for the numerical approximation. Section 3 discusses the construction of a fully discrete scheme for the approximations of (1)–(3) using a time-discontinuous Galerkin method. In Section 4, we perform the convergence analysis of the discontinuous-in-time scheme under the hypotheses (S2a,b). Building on the work of Makridakis [12], this convergence proof is based on Banach’s fixed point theorem and a nonlinear elliptic projection operator whose approximation properties will be analyzed in Appendix B. Finally, numerical experiments are presented in Section 5 to verify the theoretical results.

2. Definition and assumptions

In order to find a numerical approximation to the solution of the hyperbolic system (1)–(3), we discretize it in space using a continuous Galerkin method, and then apply a discontinuous Galerkin method in time. For the sake of showing the well-posedness of the resulting numerical method, we consider the substitution $u = e^{\gamma t}v$, with $\gamma > 0$, resulting in the equivalent equation:

$$
\tilde{v}_i(x, t) + 2\gamma \tilde{v}_i(x, t) + \gamma^2 v_i(x, t) - e^{-\gamma t} \sum_{\alpha=1}^d \partial_{\alpha} S_{i\alpha}(e^{\gamma t} \nabla v(x, t)) = \tilde{f}_i(x, t) \text{ in } \Omega \times (0, T],
$$

for each $i = 1, \ldots, d$, where $\tilde{f} = e^{-\gamma t}f$, $\gamma > 0$ is a fixed constant,

$$
v(x, t) = 0 \text{ on } \partial \Omega \times (0, T],
$$

$$
v(x, 0) = v_0(x) \in [H^m(\Omega)]^d \cap [H_0^1(\Omega)]^d, \quad \tilde{v}(x, 0) = v_1(x) \in [H^{m-1}(\Omega)]^d,
$$

where $v_0(x) = u_0(x)$ and $v_1(x) = u_1(x) - \gamma u_0(x)$. From now on, we focus on the system of equations (5)–(7) only. By Theorem 1.1, we have

$$
v = u e^{-\gamma t} \in C^1\left([0, T], [H^{m-1}(\Omega)]^d \right) \cap C\left([0, T], [H^m(\Omega)]^d\right).
$$

This shows that the initial conditions stated as (7) in the above definition are meaningful.

Before describing its discretization, we first fix the notation. We use the symbol $:\equiv$ to indicate an equality in which the left hand side is defined by the right hand side. We denote by $\langle \cdot, \cdot \rangle_{L^2}$ the inner product in $L^2(\Omega)$ and $[L^2(\Omega)]^d$. Following standard notational conventions, we shall write $W^{s,p} := [W^{s,p}(\Omega)]^d$ for $s \in \mathbb{Z}, p \in \mathbb{R}^+ \cup \{\infty\}$, and put $H^s := W^{s,2}$. Similarly, $H_0^1 := [H_0^1(\Omega)]^d$ and $L^2 := [L^2(\Omega)]^d$.

We define the following time-dependent semilinear form

$$
a(v(t), \varphi) := \sum_{i,\alpha=1}^d e^{-\gamma t} (S_{i\alpha}(e^{\gamma t} \nabla v(t)), \partial_{\alpha} \varphi_i)_{L^2}, \quad \text{for } \varphi \in H_0^1.
$$

Since we also need to approximate the gradient of the solution $\nabla u = e^{\gamma t} \nabla v$, we assume that there exists an open convex set $\mathcal{M}$ with $\overline{\mathcal{M}} \subset \mathcal{O}$ such that $\nabla u([\overline{\Omega} \times [0, T]]) \subset \mathcal{M}$. If the distance of $\overline{\mathcal{M}}$ from $\partial \mathcal{O}$ is $\delta$, we consider the set

$$
\mathcal{M}_\delta := \left\{ \eta \in \mathbb{R}^{d \times d} : \inf_{\sigma \in \mathcal{M}} |\eta - \sigma| \leq \delta \right\},
$$

where $|\cdot|$ denotes the Frobenius norm on $\mathbb{R}^{d \times d}$, defined, for $\eta \in \mathbb{R}^{d \times d}$, by $|\eta| = (\eta : \eta)^{\frac{1}{2}}$. Notice that the set $\mathcal{M}_\delta$ is convex (cf. [27], Lem. 1). Since we only require $S$ to be locally Lipschitz continuous in $\mathcal{M}_\delta$, we define the local Lipschitz constant of $S$ in $\mathcal{M}_\delta$ by

$$
K_\delta := \sup_{\eta \in \mathcal{M}_\delta} \left( \sum_{i,\alpha,j,\beta=1}^d |A_{i\alpha j\beta}(\eta)|^2 \right)^{\frac{1}{2}},
$$
and the local Lipschitz constant of the fourth-order elasticity tensor $A = \nabla S$ by

$$L_\delta := \sup_{\eta_1, \eta_2 \in \mathcal{M}_\delta, \eta_1 \neq \eta_2} |\eta_1 - \eta_2|^{-1} \left( \sum_{i,\alpha,j,\beta=1}^d |A_{i\alpha j\beta}(\eta_1) - A_{i\alpha j\beta}(\eta_2)|^2 \right)^{\frac{1}{2}}. \tag{10}$$

Since the set $\mathcal{M}_\delta$ is a compact subset of $\mathbb{R}^{d \times d}$ for every $\delta > 0$ and $A_{i\alpha j\beta}$ is sufficiently smooth (and in particular continuously differentiable on $\mathcal{M}_\delta$), it follows that $K_\delta$ and $L_\delta$ are finite. We also define

$$\mathcal{Z}_\delta := \left\{ \Phi \in L^\infty(\Omega)^{d \times d} : \Phi(x) \in \mathcal{M}_\delta, x \in \Omega \right\}. \tag{11}$$

This set $\mathcal{Z}_\delta$ is expected to contain the gradients of approximations of $u$. We define

$$\tilde{a}(\varphi; \phi, \psi) := \sum_{i,\alpha,j,\beta=1}^d (A_{i\alpha j\beta}(\nabla \varphi) \partial_\beta \phi_j, \partial_\alpha \psi_i)_{L^2}, \quad \text{for } \varphi, \phi, \psi \in H^1_0. \tag{12}$$

By the definition of $\mathcal{Z}_\delta$ and (S1a), we have

$$\tilde{a}(\varphi; \phi, \psi) = \hat{a}(\varphi; \psi, \phi), \quad \text{for } \varphi, \phi, \psi \in H^1_0, \nabla \varphi \in \mathcal{Z}_\delta. \tag{S2a}$$

We also assume that there exists a real number $M_1 > 0$ such that

$$\hat{a}(\varphi; \phi, \phi) \geq M_1 \|\nabla \phi\|^2_{L^2}, \quad \text{for } \varphi, \phi \in H^1_0, \nabla \varphi \in \mathcal{Z}_\delta. \tag{S2b}$$

Note that (S2b) is a stronger assumption than (S1b). In general, (S1b) does not imply (S2b) for $d > 1$. We refer the reader to [28–30] for counterexamples. In fact, (S1b) only implies the following Garding's inequality:

$$\hat{a}(\varphi; \phi, \phi) \geq \frac{1}{2} M_0 \|\nabla \phi\|^2_{L^2} - \mu \|\phi\|^2_{L^2} \quad \text{for } \mu \geq 0, \varphi, \phi \in H^1_0, \nabla \varphi \in \mathcal{Z}_\delta, \tag{13}$$

cf. Theorem 6.5.1 in [31] and Lemma 5 in [27]. We note that the techniques of this paper can be extended so that our results are still valid under this weak condition.

### 3. Numerical Scheme

#### 3.1. Semi-discrete approximation

We shall discretize the problem (5)–(7) in space using a continuous Galerkin method. For the spatial discretization parameter $h \in (0, 1)$, we define $\mathcal{V}_h$ to be a given family of finite-dimensional subspaces of $H^1_0 \cap H^m$ with polynomial degree $p \geq 1$. We shall assume that the triangulation $\{T_h\}_{h>0}$ of $\Omega$ into $d$-dimensional simplices, which are possibly curved along the boundary $\partial \Omega$, is shape-regular and quasi-uniform. It follows from Bernardi’s work [32] that

$$\inf_{v_h \in \mathcal{V}_h} \{\|v - v_h\|_{L^2} + h\|v - v_h\|_{H^1} \leq Ch^{r+1}\|v\|_{H^{r+1}}, \quad 1 \leq r \leq \min(p, m - 1), \quad v \in H^m \cap H^1_0. \tag{i}$$

Further, the following inverse inequalities follow directly from the quasi-uniformity of the triangulation.

There exists a positive constant $C_0$ such that, for every $v_h \in \mathcal{V}_h$,

$$\|\nabla v_h\|_{L^2} \leq C_0 h^{-1}\|v_h\|_{L^2} \quad \text{and} \quad \|\nabla v_h\|_{L^\infty} \leq C_0 h^{-1}\|v_h\|_{L^\infty}. \tag{ii,a}$$

There exists a positive constant $C_1$ such that, for every $v_h \in \mathcal{V}_h$,

$$\|\nabla v_h\|_{L^\infty} \leq C_1 h^{-\frac{d}{2}}\|\nabla v_h\|_{L^2} \quad \text{and} \quad \|v_h\|_{L^\infty} \leq C_1 h^{-\frac{d}{2}}\|v_h\|_{L^2}. \tag{ii,b}$$
With the above assumptions, we are ready to construct the continuous-in-time finite element approximation $v_h$ of $v$. The semi-discrete approximation $v_h : [0, T] \rightarrow \mathcal{V}_h$ of the solution of (5)–(7) satisfies the following initial-value problem in $\mathcal{V}_h$:

$$\left( \tilde{v}_h(t), \varphi \right)_{L^2} + a(v_h(t), \varphi) + 2\gamma(\tilde{v}_h(t), \varphi)_{L^2} + \gamma^2(v_h(t), \varphi)_{L^2} = \left( \tilde{f}(t), \varphi \right)_{L^2}$$

for all $\varphi \in \mathcal{V}_h$, $0 \leq t \leq T$,

$$v_h(0) = v_{0,h} \in \mathcal{V}_h, \quad \tilde{v}_h(0) = v_{1,h} \in \mathcal{V}_h,$$

where $v_{0,h}$ and $v_{1,h}$ are specially chosen initial values. It was proved by Makridakis [12] that the semi-discrete problem (14), (15) with $\gamma = 0$ (the semi-discrete form based on a continuous finite element approximation of the original problem (1)–(3)) has a locally unique solution and that the optimal-order $L^2$ error estimate

$$\max_{0 \leq t \leq T} \|v(t) - v_h(t)\|_{L^2} \leq C(v)h^{p+1}$$

holds for sufficiently smooth initial data. Here $p$ is the polynomial degree of the elements of the finite-dimensional space $\mathcal{V}_h$, which satisfies $p > \frac{d}{2}$. The proofs of these assertions for $\gamma > 0$ are completely analogous and are therefore omitted.

### 3.2. Discontinuous-in-time fully discrete scheme

In this section we shall construct a fully discrete approximation of the solution of (5)–(7) by applying a discontinuous Galerkin method in time. For this purpose, we partition the time interval $I = (0, T]$ into $N$ sub-intervals $I_n = (t_{n-1}, t_n]$ having length $k_n = t_n - t_{n-1}$ for $n = 1, 2, \ldots, N$, with $t_0 = 0$ and $t_N = T$. To deal with the discontinuity at each $t_n$ in the numerical approximation to $v$, we introduce the jump operator

$$[v_h]_n := v_h(t^+_n) - v_h(t^-_n) \quad \text{for } n = 0, 1, \ldots, N - 1,$$

where

$$v_h(t^+_n) := \lim_{\varepsilon \rightarrow 0^+} v_h(t_n + \varepsilon) \quad \text{for } n = 0, 1, \ldots, N - 1.$$

By convention, we assume that $v_h(0^-) = v_{0,h}$ and $\tilde{v}_h(0^-) = v_{1,h}$. Moreover, we define $v_{h,n}^+ := v_h(t^+_n)$ and $v_{h,n}^- := v_h(t^-_n).$ To deal with the nonlinear term, we apply Taylor’s theorem with an integral remainder to have

$$S_{\alpha}(\nabla v_h(t)e^\gamma t) = S_{\alpha}(0) + \sum_{j, \beta=1}^d \partial_\beta v_{h,j}(t)e^\gamma t \int_0^1 A_{i\alpha j\beta}(\tau \nabla v_h(t)e^\gamma t) d\tau.$$

By assuming that $S(0) = 0$, we can write the semilinear term as

$$\int_{t_{n-1}}^{t_n} \sum_{i, \alpha=1}^d e^{-\gamma t} \left( S_{\alpha}(\nabla v_h(t)e^\gamma t), \partial_\alpha v_{h,i}(t) \right)_{L^2} dt$$

$$= \int_{t_{n-1}}^{t_n} \sum_{i, \alpha, j, \beta=1}^d \left( \partial_\beta v_{h,j}(t) \int_0^1 A_{i\alpha j\beta}(\tau \nabla v_h(t)e^\gamma t) d\tau, \partial_\alpha v_{h,i}(t) \right)_{L^2} dt$$

$$= \int_{t_{n-1}}^{t_n} \int_0^1 \sum_{i, \alpha, j, \beta=1}^d \left( A_{i\alpha j\beta}(\tau \nabla v_h(t)e^\gamma t) \partial_\beta v_{h,j}(t), \partial_\alpha v_{h,i}(t) \right)_{L^2} d\tau dt$$

$$= \int_{t_{n-1}}^{t_n} \left\{ \int_0^1 \tilde{a}(\tau \nabla v_h(t)e^\gamma t; v_h(t), \dot{v}_h(t)) d\tau \right\} dt,$$
where
\[ \tilde{a}(\tau \nabla v_h(t)e^{\gamma t}; v_h(t), \dot{v}_h(t)) := \sum_{i,\alpha;j,\beta=1}^d (A_{i\alpha j\beta}(\tau \nabla v_h(t)e^{\gamma t}) \partial_{\beta} v_h(t), \partial_{\alpha} \dot{v}_h(t))_{L^2}. \]

We focus on the generic time interval \(I_n\) and assume that the solution on \(I_{n-1}\) is known. Following the discontinuous-in-time numerical scheme introduced in [26], we first test the equation (5) against \(\tilde{\varphi}\) for \(\varphi \in H^1(I_n; H^1_0)\) and integrate on \(I_n\) to obtain the following weak formulation:
\[
\int_{t_{n-1}}^{t_n} (\tilde{\nabla}(t), \tilde{\varphi}(t))_{L^2} dt = \int_{I_{n-1}}^{t_n} \left( \sum_{i,\alpha=1}^d e^{-\gamma \tau} (S_{i\alpha} (\nabla v(t)e^{\gamma \tau}), \partial_{\alpha} \tilde{\varphi}(t))_{L^2} + 2 \gamma \int_{t_{n-1}}^{t_n} (\tilde{\varphi}(t))_{L^2} dt \right)
\]
with \( q_n \geq 2 \) for each \( 1 \leq n \leq N \). For \( \mathbf{q} := [q_1, q_2, \ldots, q_N]^T \in \mathbb{N}^N \), we then define the space

\[
\mathcal{V}^{\mathbf{q}} := \{ \mathbf{v} : [0, T] \to V_h; \mathbf{v}|_{I_n} \in \mathcal{V}^{\mathbf{q}}_{V,h} \text{ for } n = 1, 2, \ldots, N \}.
\]

Then, the discontinuous-in-time fully discrete approximation of the problem reads as follows: find \( \mathbf{v}_{DG} \in \mathcal{V}^{\mathbf{q}} \) such that

\[
\mathcal{A}(\mathbf{v}_{DG}, \varphi) = \tilde{F}(\varphi) \quad \text{for all } \varphi \in \mathcal{V}_{V,h},
\]

where \( \tilde{F} \) is a modified version of \( F \) defined as

\[
\tilde{F}(\varphi) := \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left( \tilde{f}(t), \varphi(t) \right)_L dt + (\mathbf{v}_{1,h}, \varphi^+_L) + \gamma^2(\mathbf{v}_{0h}, \varphi^+_L)_L + \frac{1}{2} \int_0^1 \tilde{a}(\tau \mathbf{v}_{0h}; \mathbf{v}_{0h}, \varphi^-) d\tau.
\]

### 4. Convergence Analysis

By using the ideas introduced in [12] based on Banach’s fixed point theorem, we will show the existence and uniqueness of \( \mathbf{v}_{DG} \). We shall also prove a priori error estimates as summarized in the following theorem.

**Theorem 4.1.** Let \( \mathbf{v} \in W^{s,\infty}([0,T];H^m \cap H_0^1) \), with \( \frac{d}{2} + 1 < r \leq \min(p, m - 1) \), \( s \geq q_i + 1 \) for each \( i = 1, 2, \ldots, N \), be the solution of (5)–(7). Assume that \( k_i^{q_i-\frac{1}{2}} = o\left(h^{2+q_i}\right) \) and there exist positive constants \( \mu_i, \nu_i \) such that \( \mu_i k_i \leq h^2 \leq \nu_i k_i \) for each \( i = 1, 2, \ldots, N \). Suppose that we choose the initial data \( \mathbf{v}_{0h}, \mathbf{v}_{1,h} \in V_h \) to be

\[
\mathbf{v}_{1,h} = \mathbf{W}(0), \quad \mathbf{v}_{1,h} = \tilde{\mathbf{W}}(0),
\]

where \( \mathbf{W}(t) \in V_h \) is the nonlinear elliptic projection of \( \mathbf{v}(t) \) such that

\[
a(\mathbf{W}(t), \varphi) = a(\mathbf{v}(t), \varphi) \quad \text{for all } \varphi \in V_h.
\]

Then we have for the solutions of (19) that

\[
\|\mathbf{v}_{DG}(t_j^+ \mathbf{v}_{DG}(t_j^-) \|^2 \leq C(\mathbf{v}) \left(h^{2r+2} + \sum_{n=1}^{N} \frac{k_{2n+1}}{q_n} \right)^{\frac{1}{2}}
\]

for each \( j = 1, \ldots, N \), where \( C(\mathbf{v}) \) is a positive constant depending on the solution \( \mathbf{v} \).

**Remark 4.2.** If we use uniform time intervals \( k_n = k = h^2 \), and uniform polynomial degrees \( q_n = q \geq 2 \), for \( n = 1, \ldots, N \), then the error bound at the end nodal point becomes

\[
\|\mathbf{v}_{DG}(t_n^-) - \mathbf{v}(t_n^-) \|^2 \leq C(\mathbf{v}) \left(k^{q\frac{1}{2} + q} \right).
\]

**Remark 4.3.** The assumptions that \( k_i^{q_i-\frac{1}{2}} = o\left(h^{1+q_i}\right) \) and \( \mu_i k_i \leq h^2 \leq \nu_i k_i \) for each \( i = 1, \ldots, N \) require that \( q_i > 1 + \frac{d}{4} \) for each \( i = 1, \ldots, N \). That is, we need the polynomial degree in time satisfies \( q_i \geq 2 \) for \( d = 1, 2, 3 \) on each time interval \( I_i \), with \( i = 1, \ldots, N \).

**Remark 4.4.** By the Sobolev embedding theorem, \( \mathbf{v} \in W^{s,\infty}([0,T];H^m) \) for \( m > \frac{d}{2} + 2 \) implies that \( \mathbf{v} \in W^{s,\infty}([0,T];C^{2,\alpha}(\Omega)^{d}) \) for some \( \alpha \in (0,1) \). Note that the assumption \( m > \frac{d}{2} + 2 \) is consistent with the assumption \( m \geq \left[ \frac{d}{2} + 3 \right] \) in Theorem 1.1. That is, we need \( m \geq 3 \) for \( d = 1 \) and \( m \geq 4 \) for \( d = 2, 3 \).

It will be assumed throughout the convergence analysis that

\[
\mathbf{v} \in W^{s,\infty}([0,T];H^m \cap H_0^1).
\]
4.1. Definition of the fixed point map

It is known, see [33, 34], that (21) has, for \( h \) sufficiently small, a locally unique solution \( W(t) \in \mathcal{V}_h \) for \( 0 \leq t \leq T \). Furthermore, \( W \) satisfies the following properties, which are established in Appendix B.

There exists a constant \( C_r(v) \) depending on \( v \) such that, for \( \frac{d}{2} + 1 < r \leq \min(p, m - 1) \),

\[
\left\| \nabla v^{(j)}(t) - \nabla W^{(j)}(t) \right\|_{L^2} \leq C_r(v) h^r, \quad 0 \leq t \leq T,
\]

for \( j = 0, 1 \), where \( v^{(j)} := \frac{d^j}{dt^j} v \). In addition, we shall prove for the time-derivatives of \( W \) there holds

\[
\left\| v^{(j)}(t) - W^{(j)}(t) \right\|_{L^2} \leq \hat{C}_r(v) h^{r+1}, \quad 0 \leq t \leq T,
\]

for \( j = 0, 1, 2 \). We can also show that there exist constants \( c_0 \) and \( c_1 \), independent of \( h \), such that

\[
\left\| \nabla W(t) \right\|_{L^\infty} \leq c_0 \quad \text{and} \quad \left\| \nabla W(t) \right\|_{L^\infty} \leq c_1, \quad 0 \leq t \leq T.
\]

Let \( \Pi_k = \Pi_k^{q_n} \) denote the modified \( L^2 \)-projector in time direction. That is, for each \( n = 1, 2, \ldots, N \),

\[
(P_k W - W)(x, t_n^+) = 0; \quad (P_k W - W)(x, t_n^-) = 0; \quad \partial_t (P_k W - W)(x, t_n^) = 0; \quad \int_{t_n^-}^{t_n^+} (\partial_t (P_k W - W), \chi)_{L^2} \, dt = 0 \quad \text{for } \chi \in \mathcal{V}_{kh}^{q_n}.
\]

It was first proved in [35] and further studied in [36] that for each \( W \in H^s(I_n; L^2) \), there exists a positive constant \( C \) such that,

\[
\int_{t_n^-}^{t_n^+} \left\| \partial_t (W(\cdot, t) - P_k W(\cdot, t)) \right\|_{L^2}^2 \, dt \leq C \frac{k_n^{2(\mu-2)}}{q_n^{2(\mu-3)}} \| W \|_{H^s(I_n; L^2)}^2,
\]

\[
\int_{t_n^-}^{t_n^+} \left\| \partial_t (W(\cdot, t) - P_k W(\cdot, t)) \right\|_{L^2}^2 \, dt \leq C \frac{k_n^{2(\mu-1)}}{q_n^{2(\mu-3)}} \| W \|_{H^s(I_n; L^2)}^2.
\]

\[
\int_{t_n^-}^{t_n^+} \left\| W(\cdot, t) - P_k W(\cdot, t) \right\|_{L^2}^2 \, dt \leq C \frac{k_n^{2\mu}}{q_n^{2(s-1)}} \| W \|_{H^s(I_n; L^2)}^2,
\]

where \( \mu = \min(q_n + 1, s) \) and \( q_n \) is the polynomial degree with respect to the variable \( t \). If we change the spatial function space from \( L^2 \) to \( H^1_0 \) in (27)–(29), analogous estimates follow. Note that we can also get inverse inequalities with respect to the time derivatives in an analogous manner as (ii,a). That is, there exists a positive constant \( C_2 \) such that, for each fixed \( x \in \Omega \), for every \( \varphi \in \mathcal{V}_{kh}^{q} \),

\[
\left\| \partial_t \varphi(x, \cdot) \right\|_{L^2(I_n)} \leq C_2 k_n^{-1} \| \varphi(x, \cdot) \|_{L^2(I_n)}
\]

\[
\left\| \partial_t \varphi(x, \cdot) \right\|_{L^\infty(I_n)} \leq C_2 k_n^{-1} \| \varphi(x, \cdot) \|_{L^\infty(I_n)},
\]

for each \( n = 1, 2, \ldots, N \). We now decompose the error as

\[
v_{DG}(t) - v(t) = (v_{DG}(t) - P_k W(t)) + (P_k W(t) - W(t)) + (W(t) - v(t))
\]

\[
:= \theta(t) + \rho_1(t) + \rho_2(t)
\]
for $t \in I_n, n = 1, 2, \ldots, N$. First note that
\[
\int_{t_{n-1}}^{t_n} \left\| \rho_k^{(j)}(t) \right\|_{L^2}^2 \, dt = \int_{t_{n-1}}^{t_n} \left\| \Pi_k \mathbf{W}^{(j)}(t) - \mathbf{W}^{(j)}(t) \right\|_{L^2}^2 \, dt \leq C k_n^{2(\mu-j)} q_n^{2(s-1)} \int_{t_{n-1}}^{t_n} \sum_{\alpha=0}^{s} \left\| \mathbf{W}^{(\alpha)}(t) \right\|_{L^2}^2 \, dt
\]
\[
\leq C_1(v) k_n^{2(\mu-j)+1} q_n^{2(s-1)} \quad \text{for } j = 0, 1, \quad (32)
\]
where we have applied (28) and (29). Here $\mu = \min(s, q_n + 1)$ and $\Pi_k \mathbf{W}^{(j)} = \frac{d^j}{dt^j} \Pi_k \mathbf{W}$, $\mathbf{W}^{(j)} = \frac{d^j}{dt^j} \mathbf{W}$, with $j = 0, 1$. If we assume that the solution $\mathbf{v} \in W^{s,\infty}([0,T]; H^m \cap H_0^1)$ of (5)–(7) is sufficiently smooth (i.e. $s > q_n + 1$), then we can write
\[
\int_{t_{n-1}}^{t_n} \left\| \rho_2^{(j)}(t) \right\|_{L^2}^2 \, dt = \int_{t_{n-1}}^{t_n} \left\| \mathbf{W}^{(j)}(t) - \mathbf{v}^{(j)}(t) \right\|_{L^2}^2 \, dt \leq C_2(v) k_n h^{2r+2} \quad (33)
\]
for $j = 0, 1$. By the property (iii,b) of the elliptic projection, we know that
\[
\int_{t_{n-1}}^{t_n} \left\| \rho_2^{(j)}(t) \right\|_{L^2}^2 \, dt = \int_{t_{n-1}}^{t_n} \left\| \mathbf{W}^{(j)}(t) - \mathbf{v}^{(j)}(t) \right\|_{L^2}^2 \, dt \leq C_2(v) k_n h^{2r+2} \quad (34)
\]
for $j = 0, 1, 2$. Here $C_i(v)$ for $i = 1, 2$ are constants depending on the exact solution $\mathbf{v}$. Recall that the fully discrete scheme is
\[
\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left( \tilde{\mathbf{v}}^{DG}(t), \tilde{\phi}(t) \right)_{L^2} \, dt + \sum_{n=1}^{N-1} \left( [\tilde{\mathbf{v}}^{DG}(t_n), \tilde{\phi}(t_n^+)]_{L^2} + ([\tilde{\mathbf{v}}^{DG}(t_0^+), \tilde{\phi}(t_0^+)]_{L^2}
\]
\[
+ \gamma^2 \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\mathbf{v}^{DG}(t), \tilde{\phi}(t))_{L^2} \, dt + \gamma^2 \sum_{n=1}^{N-1} ([\mathbf{v}^{DG}(t), \tilde{\phi}(t)])_{L^2} + \gamma^2 ([\tilde{\mathbf{v}}^{DG}(t_0), \tilde{\phi}(t_0)])_{L^2}
\]
\[
+ \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \sum_{i=1}^{d} e^{-\gamma t} (S_{i,\alpha}(\nabla \mathbf{v}^{DG}(t)e^{\gamma t}), \partial_{\alpha} \tilde{\phi}(t))_{L^2} \, dt + 2\gamma \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\tilde{\mathbf{v}}^{DG}(t), \tilde{\phi}(t))_{L^2} \, dt
\]
\[
+ \sum_{n=0}^{N-1} \left\{ \int_{0}^{1} \tilde{a}(\tau \mathbf{v}^{DG}(t_n)e^{\gamma \tau n}; \mathbf{v}^{DG}(t_n), \mathbf{v}^{DG}(t_n)) \, d\tau - \int_{0}^{1} \tilde{a}(\tau \mathbf{v}^{DG}(t_n)e^{\gamma \tau n}; \mathbf{v}^{DG}(t_n), \tilde{\phi}(t_n)) \, d\tau \right\}
\]
\[
= \left( \mathbf{v}_{1,h}, \tilde{\phi}(t_0^+) \right)_{L^2} + \gamma^2 \left( \mathbf{v}_{0,h}, \tilde{\phi}(t_0^+) \right)_{L^2} + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left( \tilde{\mathbf{f}}(t), \tilde{\phi}(t) \right)_{L^2} \, dt \quad \text{for } \tilde{\phi} \in \mathcal{V}_{kh}. \quad (35)
\]

The variational form of the original problem is written as
\[
\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left( \mathbf{v}(t), \tilde{\phi}(t) \right)_{L^2} \, dt + \sum_{n=1}^{N-1} \left( [\mathbf{v}(t), \tilde{\phi}(t)]_{L^2} + ([\mathbf{v}(t_0), \tilde{\phi}(t_0)])_{L^2}
\]
\[
+ \gamma^2 \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\mathbf{v}(t), \tilde{\phi}(t))_{L^2} \, dt + \gamma^2 \sum_{n=1}^{N-1} ([\mathbf{v}(t), \tilde{\phi}(t)])_{L^2} + \gamma^2 ([\mathbf{v}(t_0), \tilde{\phi}(t_0)])_{L^2}
\]
\[
+ \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \sum_{i=1}^{d} e^{-\gamma t} (S_{i,\alpha}(\nabla \mathbf{v}(t)e^{\gamma t}), \partial_{\alpha} \tilde{\phi}(t))_{L^2} \, dt + 2\gamma \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\tilde{\mathbf{v}}(t), \tilde{\phi}(t))_{L^2} \, dt
\]
\[
+ \sum_{n=0}^{N-1} \left\{ \int_{0}^{1} \tilde{a}(\tau \mathbf{v}(t_n)e^{\gamma \tau n}; \mathbf{v}(t_n), \mathbf{v}(t_n)) \, d\tau - \int_{0}^{1} \tilde{a}(\tau \mathbf{v}(t_n)e^{\gamma \tau n}; \mathbf{v}(t_n), \phi(t_n)) \, d\tau \right\}
\]
\[
= (v_1, \dot{\varphi}(t_n^+))_{L^2} + \gamma^2 (v_0, \varphi(t_n^+))_{L^2} + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (\ddot{f}(t), \dot{\varphi}(t))_{L^2} \, dt, \quad \text{for } \varphi \in V_{kh}^q.
\] (36)

By considering the nonlinear elliptic projection of \(v(t)\) (cf. equality (21)), we can replace
\[
\sum_{i, \alpha = 1}^d e^{-\gamma t}(S_{i\alpha}(\nabla v(t)e^{\gamma t}), \partial_{\alpha} \varphi_i(t))_{L^2} \, dt \quad \text{by} \quad \sum_{i, \alpha = 1}^d e^{-\gamma t}(S_{i\alpha}(\nabla W(t)e^{\gamma t}), \partial_{\alpha} \varphi_i(t))_{L^2} \, dt
\]
in (36). Using the continuity of \(v(t)\) and \(W(t)\) in time, we can also replace
\[
\sum_{n=0}^{N-1} \left\{ \int_0^1 \tilde{a}(\tau v(t_n^+)e^{\gamma \tau n}; v(t_n^+), \varphi(t_n^+)) \, d\tau - \int_0^1 \tilde{a}(\tau v(t_n^-)e^{\gamma \tau n}; v(t_n^-), \varphi(t_n^-)) \, d\tau \right\}
\]
by
\[
\sum_{n=0}^{N-1} \left\{ \int_0^1 \tilde{a}(\tau W(t_n^+)e^{\gamma \tau n}; W(t_n^+), \varphi(t_n^+)) \, d\tau - \int_0^1 \tilde{a}(\tau W(t_n^-)e^{\gamma \tau n}; W(t_n^-), \varphi(t_n^-)) \, d\tau \right\}
\]
in (36). Subtracting the resulting equality from (35), we have
\[
\sum_{n=1}^N \int_{t_{n-1}}^{t_n} \left( \ddot{\theta}(t) + \ddot{\rho}_1(t) + \ddot{\rho}_2(t), \dot{\varphi}(t) \right)_{L^2} \, dt + \sum_{n=1}^{N-1} \left( \left[ \ddot{\theta}(t) + \ddot{\rho}_1(t) + \ddot{\rho}_2(t) \right](t_n^+), \dot{\varphi}(t_n^+) \right)_{L^2} \, dt
\]
\[
+ (\dot{\varphi}_{DG}(t_n^+), \dot{\varphi}(t_n^+))_{L^2} + 2\gamma \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \left( \ddot{\theta}(t) + \ddot{\rho}_1(t) + \ddot{\rho}_2(t), \dot{\varphi}(t) \right)_{L^2} \, dt
\]
\[
+ \gamma^2 \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (\theta(t) + \rho_1(t) + \rho_2(t), \varphi(t))_{L^2} \, dt + \gamma^2 \sum_{n=1}^{N-1} (\left[ \theta(t) + \rho_1(t) + \rho_2(t) \right](t_n^-), \varphi(t_n^-))_{L^2} \, dt
\]
\[
+ \gamma^2 (v_{DG}(t_n^+), \varphi(t_n^+))_{L^2} + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \sum_{i, \alpha = 1}^d e^{-\gamma t}(S_{i\alpha}(\nabla v_{DG}(t)e^{\gamma t}), \partial_{\alpha} \varphi_i)_{L^2} \, dt
\]
\[
+ \sum_{n=0}^{N-1} \frac{1}{2} \left\{ \int_0^1 \tilde{a}(\tau v_{DG}(t_n^+)e^{\gamma \tau n}; v_{DG}(t_n^+), \varphi(t_n^+)) \, d\tau - \int_0^1 \tilde{a}(\tau W(t_n^+)e^{\gamma \tau n}; W(t_n^+), \varphi(t_n^+)) \, d\tau \right\}
\]
\[
+ \sum_{n=0}^{N-1} \frac{1}{2} \left\{ \int_0^1 \tilde{a}(\tau W(t_n^+)e^{\gamma \tau n}; W(t_n^+), \varphi(t_n^+)) \, d\tau - \int_0^1 \tilde{a}(\tau v_{DG}(t_n^+)e^{\gamma \tau n}; v_{DG}(t_n^+), \varphi(t_n^+)) \, d\tau \right\}
\]
\[
= (v_{1,h} - v_1, \dot{\varphi}(t_n^+))_{L^2} + \gamma^2 (v_{0,h} - v, \varphi(t_n^+))_{L^2}, \quad \text{for } \varphi \in V_{kh}^q.
\] (37)

Now we consider the integral on \(I_n = (t_{n-1}, t_n)\) only,
\[
\int_{t_{n-1}}^{t_n} \left( \ddot{\theta}(t), \dot{\varphi}(t) \right)_{L^2} \, dt + \left( \left[ \theta(t) \right]_{n-1}^{t_n}, \dot{\varphi}(t_n^-) \right)_{L^2} + 2\gamma \int_{t_{n-1}}^{t_n} \left( \ddot{\theta}(t), \dot{\varphi}(t) \right)_{L^2} \, dt + \gamma^2 \int_{t_{n-1}}^{t_n} (\theta(t), \dot{\varphi}(t))_{L^2} \, dt
\]
\[
+ \gamma^2 (\theta(t)|_{n-1}, \varphi(t_{n-1}^+))_{L^2} + \int_{t_{n-1}}^{t_n} \sum_{i, \alpha = 1}^d e^{-\gamma t}(S_{i\alpha}(\nabla v_{DG}(t)e^{\gamma t}), \partial_{\alpha} \varphi_i)(t)_{L^2} \, dt
\]
\[
+ \frac{1}{2} \left\{ \int_0^1 \tilde{a}(\tau v_{DG}(t_n^+)e^{\gamma \tau n}; v_{DG}(t_n^+), \varphi(t_n^+)) \, d\tau - \int_0^1 \tilde{a}(\tau W(t_n^+)e^{\gamma \tau n}; W(t_n^+), \varphi(t_n^+)) \, d\tau \right\}
\]
If we can write \(0 \leq \rho\) where we have used the fact that theorem with an integral remainder, we have
\[
\int_{t_{n-1}}^{t_n} (\tilde{\rho}_1(t), \tilde{\varphi}(t)) dt - 2\gamma \int_{t_{n-1}}^{t_n} (\tilde{\rho}_1(t), \tilde{\varphi}(t)) dt
\]
\[
= \int_{t_{n-1}}^{t_n} (\tilde{\rho}_2(t), \tilde{\varphi}(t)) dt - 2\gamma \int_{t_{n-1}}^{t_n} (\tilde{\rho}_2(t), \tilde{\varphi}(t)) dt
\]
\[
= -\gamma^2 \int_{t_{n-1}}^{t_n} (\rho_1(t), \varphi(t)) dt - 2\gamma \int_{t_{n-1}}^{t_n} (\tilde{\rho}_1(t), \tilde{\varphi}(t)) dt
\]
\[
= -\gamma^2 \int_{t_{n-1}}^{t_n} (\rho_2(t), \varphi(t)) dt - 2\gamma \int_{t_{n-1}}^{t_n} (\tilde{\rho}_2(t), \tilde{\varphi}(t)) dt
\]
where we have used the fact that \(\rho_2(t)\) and \(\tilde{\rho}_2(t)\) are continuous in time and properties (23)–(26). By Taylor’s theorem with an integral remainder, we have
\[
S_{ta}(\nabla \nu_{DG}(t)e^{\gamma t}) = S_{ta}(\nabla W(t)e^{\gamma t}) + \sum_{j=1}^{d} e^{\gamma t} \partial_j (\nu_{DG}(t) - W(t)) \int_{0}^{1} \frac{\partial}{\partial \eta_{j\beta}} S_{ta}(\nabla W(t)e^{\gamma t} + \tau(\nabla \nu_{DG}(t) - \nabla W(t))e^{\gamma t}) d\tau.
\]
If \(\nabla \nu_{DG}(t)e^{\gamma t} \in Z_3, \nabla W(t)e^{\gamma t} \in Z_3\) for each \(t \in [0, T]\), we have \(\nabla W(t)e^{\gamma t} + \tau(\nabla \nu_{DG}(t) - \nabla W(t))e^{\gamma t} \in Z_3\) for \(0 \leq \tau \leq 1\) by the convexity of \(Z_3\). This implies that the term in the integral remainder is well-defined. Thus, we can write
\[
\sum_{i, \alpha=1}^{d} e^{-\gamma t} (S_{ta}(\nabla \nu_{DG}(t)e^{\gamma t}) - S_{ta}(\nabla W(t)e^{\gamma t}), \partial_\alpha \tilde{\varphi}(t)) \int_{0}^{1} \frac{\partial}{\partial \eta_{j}\beta} S_{ta}(\nabla W(t)e^{\gamma t} + \tau(\nabla \phi(t) - \nabla W(t))e^{\gamma t}) \partial_\beta (\nu_{DG}(t) - W(t)) \partial_\alpha \tilde{\varphi}(t) \int_{0}^{1} d\tau
\]
\[
= \int_{0}^{1} \sum_{i, \alpha, j, \beta=1}^{d} \left( A_{i\alpha,j\beta}(\nabla W(t)e^{\gamma t} + \tau(\nabla \phi(t) - \nabla W(t))e^{\gamma t}) \partial_\beta (\nu_{DG}(t) - W(t)) \partial_\alpha \tilde{\varphi}(t) \right) \int_{0}^{1} d\tau
\]
\[
= \int_{0}^{1} \left( \tilde{a}(W(t)e^{\gamma t} + \tau(\nu_{DG}(t) - W(t))e^{\gamma t}) ; \nu_{DG}(t) - W(t), \tilde{\varphi}(t) \right) d\tau.
\]
For simplicity of notation, we write
\[
A(\nu_{DG}(t)e^{\gamma t} ; \nu_{DG}(t) - W(t), \tilde{\varphi}(t)) := \int_{0}^{1} \left( \tilde{a}(W(t)e^{\gamma t} + \tau(\nu_{DG}(t) - W(t))e^{\gamma t}) ; \nu_{DG}(t) - W(t), \tilde{\varphi}(t) \right) d\tau.
\]
Analogously,
\[
\int_{0}^{1} \left( \tilde{a}(\nu_{DG}(t)e^{\gamma t} - \nu_{DG}(t) ; \nu_{DG}(t) - W(t), \tilde{\varphi}(t) \right) d\tau
\]
\[
= \int_{0}^{1} \left( \tilde{a}(W(t)e^{\gamma t} + \tau(\nu_{DG}(t) - W(t))e^{\gamma t}) ; \nu_{DG}(t) - W(t), \tilde{\varphi}(t) \right) d\tau.
\]
\[
\begin{align*}
\gamma^n & = \sum_{i=1}^d e^{-\gamma^n t_{n-1}} \left\{ \left( S_{i\alpha} (\nabla_{DG} (t_{n-1}^+) e^{\gamma^n t_{n-1}}) \right) (x) \right\} L^2 \\
& = \sum_{i=1}^d e^{-\gamma^n t_{n-1}} \left( S_{i\alpha} (\nabla_{DG} (t_{n-1}^+) e^{\gamma^n t_{n-1}}) \right) (x) L^2 \\
& = \sum_{i=1}^d e^{-\gamma^n t_{n-1}} \left( S_{i\alpha} (\nabla_{DG} (t_{n-1}^+) e^{\gamma^n t_{n-1}} - S_{i\alpha} (\nabla_{DG} (t_{n-1}^+) e^{\gamma^n t_{n-1}}) \right) L^2 \\
& = A(v_{DG}(t_{n-1}^+) e^{\gamma^n t_{n-1}}; v_{DG}(t_{n-1}^+) - W(t_{n-1}^+), \varphi(t_{n-1}^+)).
\end{align*}
\]

Since \( W(t_{n-1}^+) - \Pi_k W(t_{n}^+) = 0 \) for each \( n = 1, 2, \ldots, N \), our equation (38) becomes

\[
\begin{align*}
\int_{t_{n-1}}^{t_n} \left( \dot{\theta}(t), \dot{\varphi}(t) \right)_{L^2} dt + \left( \left( \dot{\theta}(t) \right)_{n-1}, \dot{\varphi}(t_{n-1}^+) \right)_{L^2} + 2\gamma \int_{t_{n-1}}^{t_n} \left( \dot{\theta}(t), \dot{\varphi}(t) \right)_{L^2} dt \\
+ \gamma^2 \int_{t_{n-1}}^{t_n} A(v_{DG}(t) e^{\gamma^n t_{n-1}}; v_{DG}(t), \varphi(t)) dt + \frac{1}{2} A(v_{DG}(t_{n-1}^+) e^{\gamma^n t_{n-1}}; \theta(t_{n-1}^+), \varphi(t_{n-1}^+)) \\
= -\gamma^2 \int_{t_{n-1}}^{t_n} A(\rho_1(t), \varphi(t))_{L^2} dt - 2\gamma \int_{t_{n-1}}^{t_n} \left( \dot{\rho}_1(t), \dot{\varphi}(t) \right)_{L^2} dt - \int_{t_{n-1}}^{t_n} (\dot{\varphi}(t), \dot{\varphi}(t))_{L^2} dt \\
+ \gamma^2 \int_{t_{n-1}}^{t_n} A(v_{DG}(t) e^{\gamma^n t_{n-1}}; W(t) - \Pi_k W(t), \dot{\varphi}(t)) dt + \frac{1}{2} A(v_{DG}(t_{n-1}^-) e^{\gamma^n t_{n-1}}; \theta(t_{n-1}^-), \varphi(t_{n-1}^-))
\end{align*}
\]

for \( \varphi \in V_{h\theta}^q \). Consider the following subset of \( V_{h\theta}^q \) defined by

\[
\begin{align*}
\mathcal{F} & := \left\{ \psi \in V_{h\theta}^q \mid \text{for each } j = 1, 2, \ldots, N, \| \psi(t_j^-) - \Pi_k W(t_j^-) \|_{H^1}^2, + \| \partial_t (\psi(t_j^-) - \Pi_k W(t_j^-)) \|_{L^2}^2 \\
& + \sum_{n=1}^j \int_{t_{n-1}}^{t_n} \| \partial_t (\psi(t) - \Pi_k W(t)) \|_{L^2}^2 dt \leq C_*(v) \left( \sum_{n=1}^j k_n h^{2r+2} + \frac{k_n^{2n+1}}{q^{(s-1)}} \right) e^{\gamma^n t_n} \psi(t) \in Z_\delta, \right\}
\end{align*}
\]

where \( C_*(v) \) is a positive constant depending on the solution \( v \), which will be specified later. First note that \( \mathcal{F} \) is non-empty since \( \Pi_k W \in \mathcal{F} \). In addition, \( \mathcal{F} \) is a closed and convex subset of \( V_{h\theta}^q \) in the topology induced by the norm \( \| \cdot \|_\mathcal{F} \), which is defined by

\[
\| \varphi \|_\mathcal{F} = \max_{t \in \mathcal{T}_{n,1} \leq n \leq N} (\| \varphi(t) \|_{H^1} + \| \varphi(t) \|_{L^2})
\]

for \( \varphi \in V_{h\theta}^q \). With this notation, we are ready to define a fixed point mapping \( \mathcal{N} \) on \( \mathcal{F} \) as follows: if \( \phi \in \mathcal{F} \), the image \( v_\phi := \mathcal{N}(\phi) \) is given by the relation

\[
\begin{align*}
v_\phi(0) = v_{0,h}, \quad \dot{v}_\phi(0) = v_{1,h},
\end{align*}
\]

\[
\begin{align*}
\int_{t_{n-1}}^{t_n} \left( \dot{\theta}_\phi(t), \dot{\varphi}(t) \right)_{L^2} dt + \left( \left( \dot{\theta}_\phi(t) \right)_{n-1}, \dot{\varphi}(t_{n-1}^+) \right)_{L^2} + 2\gamma \int_{t_{n-1}}^{t_n} \left( \dot{\theta}_\phi(t), \dot{\varphi}(t) \right)_{L^2} dt \\
+ \gamma^2 \int_{t_{n-1}}^{t_n} A(v_{DG}(t) e^{\gamma^n t_{n-1}}; \theta(t), \varphi(t)) dt + \gamma^2 \left( \left( \theta(t) \right)_{n-1}, \dot{\varphi}(t_{n-1}^+) \right)_{L^2}
\end{align*}
\]
+ \int_{t_{n-1}}^{t_n} A(\phi(t)e^{\gamma t}; \theta_\phi(t), \dot{\varphi}(t)) \, dt + \frac{1}{2} A(\phi(t_1)e^{\gamma t_1}; \theta_\phi(t_1), \varphi(t_1)) \nonumber \\
= - \gamma^2 \int_{t_{n-1}}^{t_n} (p_1(t), \dot{\varphi}(t))_{L^2} \, dt - 2\gamma \int_{t_{n-1}}^{t_n} \int_{t_{n-1}}^{t_n} (\tilde{p}_1(t), \dot{\varphi}(t))_{L^2} \, dt 
- \gamma^2 \int_{t_{n-1}}^{t_n} (p_2(t), \dot{\varphi}(t))_{L^2} \, dt - 2\gamma \int_{t_{n-1}}^{t_n} (\tilde{p}_2(t), \dot{\varphi}(t))_{L^2} \, dt 
+ \int_{t_{n-1}}^{t_n} A(\phi(t)e^{\gamma t}; W(t) - \Pi_k W(t), \varphi(t)) \, dt + \frac{1}{2} A(\phi(t_1)e^{\gamma t_1}; \theta_\phi(t_1), \varphi(t_1)) \tag{40} 

where \( \theta_\phi = \mathbf{v}_\phi - \Pi_k W \).

In order to complete the proof of the theorem, it suffices to show that, for each \( n = 1, \ldots, N \), the map \( \mathcal{N} \) defined by (40) has a unique fixed point in \( \mathcal{F} \). If \( \mathbf{v}_{DG} \in \mathcal{F} \) is this fixed point, then \( \mathbf{v}_{DG} \) is a solution to (19).

### 4.2. Auxiliary results

If we take \( \varphi = \theta_\phi \) in (40), then the nonlinear term inside the integral becomes

\[
A(\phi(t)e^{\gamma t}; \theta_\phi(t), \dot{\varphi}(t)) = A\left(\phi(t)e^{\gamma t}; \theta_\phi(t), \dot{\theta}_\phi(t)\right).
\]

Following the proof in [12], it is crucial to replace the expression \( A\left(\phi(t)e^{\gamma t}; \theta_\phi(t), \dot{\theta}_\phi(t)\right) \) by

\[
\frac{1}{2} \frac{d}{dt} A(\phi(t)e^{\gamma t}; \theta_\phi(t), \theta_\phi(t)) - \frac{1}{2} \sum_{i, \alpha, j, \beta = 1}^d \int_0^1 \left( \partial_i A^-_{\alpha j \beta} \partial_j \theta_\phi(t), \partial_i \theta_\phi(t) \right)_{L^2} \, d\tau,
\]

where \( A^-_{\alpha j \beta} = A_{\alpha j \beta}(\nabla W(t)e^{\gamma t} + \tau(\nabla \phi(t) - \nabla W(t))e^{\gamma t}) \) and \( t \in I_n, n = 1, 2, \ldots, N \). We shall need an estimate on the expression

\[
A_i(\phi(t)e^{\gamma t}; \varphi(t), \psi(t)) := \frac{1}{2} \sum_{i, \alpha, j, \beta = 1}^d \int_0^1 \left( \partial_i A^-_{\alpha j \beta} \partial_j \varphi(t), \partial_i \psi(t) \right)_{L^2} \, d\tau
\]

for \( \varphi, \psi \in \mathcal{V}_h^d, t \in I_n, n = 1, 2, \ldots, N \).

**Lemma 4.5.** Under the assumptions stated in Theorem 4.1, there exists a constant \( C_\tau > 0 \) such that, for \( t \in I_n, n = 1, 2, \ldots, N \),

\[
|A_i(\phi(t)e^{\gamma t}; \varphi(t), \psi(t))| \leq C_\tau \| \nabla \varphi(t) \|_{L^2} \| \nabla \psi(t) \|_{L^2}. \tag{41}
\]

**Proof.** See Appendix A. \( \square \)

### 4.3. Convergence proof

We will establish the existence of a unique fixed point in \( \mathcal{F} \) by showing that the pair \( \mathcal{F} \) and \( \mathcal{N} \) satisfies the assumptions of Banach’s fixed point theorem, namely that

(a) \( \mathcal{N}(\mathcal{F}) \subset \mathcal{F} \).

(b) \( \mathcal{N} \) is a contraction with respect to \( d(\cdot, \cdot) \) where for \( \phi, \varphi \in \mathcal{F} \),

\[
d(\phi, \varphi) := \max_{t \in I_n, 1 \leq n \leq N} \left( \| \phi(t) - \varphi(t) \|_{H^1} + \| \dot{\phi}(t) - \dot{\varphi}(t) \|_{L^2} \right).
\]
4.3.1. Existence of a fixed point of $\mathcal{N}$ in $\mathcal{F}$

For (a), we first observe that $\mathcal{N}$ is well-defined. Indeed, if $\phi \in \mathcal{F}$, since $\nabla W e^{\gamma t} \in Z_\delta$, $\nabla W e^{\gamma t} + \tau (\nabla \phi - \nabla W) e^{\gamma t} \in Z_\delta$ for $0 \leq \tau \leq 1$, and the bilinear form $A(\phi(t)e^{\gamma t}; \cdot; \cdot)$ is symmetric and positive definite. Taking $\varphi = \theta_\phi$ in (40) and replacing $A \left( \phi(t)e^{\gamma t}; \theta_\phi(t), \dot{\theta}_\phi(t) \right)$ by

$$
\frac{1}{2} \frac{d}{dt} A(\phi(t)e^{\gamma t}; \theta_\phi(t), \dot{\theta}_\phi(t)) - \frac{1}{2} \sum_{i,\alpha,\beta=1}^d \int_{t_{n-1}}^{t_n} \left( \partial_i A_{\alpha j \beta} \partial_j \theta_{\phi, j}(t), \partial_\alpha \theta_{\phi, i}(t) \right)_L L_2 \, dt,
$$

we obtain

$$
\left\| \dot{\theta}_\phi(t_n) \right\|_L^2 + \left\| \dot{\theta}_\phi(t_{n-1}) \right\|_L^2 + \gamma \left\| \theta_\phi(t_n) \right\|_L^2 + \gamma \left\| \theta_\phi(t_{n-1}) \right\|_L^2 + 4 \gamma \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_L^2 \, dt + A(\phi(t_n)e^{\gamma t}; \theta_\phi(t_n), \dot{\theta}_\phi(t_n)) - A(\phi(t_{n-1})e^{\gamma t}; \theta_\phi(t_{n-1}), \dot{\theta}_\phi(t_{n-1}))
$$

$$= -2 \gamma^2 \int_{t_{n-1}}^{t_n} \left( \rho_1(t), \dot{\theta}_\phi(t) \right)_L L_2 \, dt - 2 \gamma^2 \int_{t_{n-1}}^{t_n} \left( \rho_2(t), \dot{\theta}_\phi(t) \right)_L L_2 \, dt - 2 \gamma^2 \int_{t_{n-1}}^{t_n} \left( \rho_2(t), \dot{\theta}_\phi(t) \right)_L L_2 \, dt - 4 \gamma \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_L^2 \, dt + 2 \left( \rho_2(t_{n-1}), \dot{\theta}_\phi(t_{n-1}) \right)_L L_2^2 + 2 \gamma^2 \left( \theta_\phi(t_{n-1}), \dot{\theta}_\phi(t_{n-1}) \right)_L L_2^2
$$

$$+ 2 \int_{t_{n-1}}^{t_n} A(\phi(t)e^{\gamma t}; W(t) - \Pi_\delta W(t), \dot{\theta}_\phi(t)) \, dt + \sum_{i,\alpha,\beta=1}^d \int_{t_{n-1}}^{t_n} \int_0^1 \left( \partial_i A_{\alpha j \beta} \partial_j \theta_{\phi, j}(t), \partial_\alpha \theta_{\phi, i}(t) \right)_L L_2 \, dt \, dt.
$$

Now we need to bound the terms on the right-hand side of equation (42). By using (33), (34) and Young’s inequality, we have

$$\left| -2 \gamma^2 \int_{t_{n-1}}^{t_n} \left( \rho_1(t), \dot{\theta}_\phi(t) \right)_L L_2 \, dt - 2 \gamma^2 \int_{t_{n-1}}^{t_n} \left( \rho_2(t), \dot{\theta}_\phi(t) \right)_L L_2 \, dt \right| \leq 3 \gamma \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_L^2 dt + C_1(\gamma) \int_{t_{n-1}}^{t_n} \left( \left\| \rho_1(t) \right\|_L^2 + \left\| \rho_2(t) \right\|_L^2 \right) dt
$$

$$+ C_2(\gamma) \int_{t_{n-1}}^{t_n} \left( \left\| \dot{\rho}_2(t) \right\|_L^2 + \left\| \dot{\rho}_2(t) \right\|_L^2 \right) dt \leq 3 \gamma \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_L^2 dt + C(\gamma, \nu) \left( k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(\nu-1)}} \right),
$$

where $C_i(\gamma)$ for $i = 1, 2$ are constants depending only on $\gamma$ and all $C(\gamma, \nu)$ for $i = 1, 2$ are constants depending on both $\gamma$ and the exact solution $\nu$. By Cauchy–Schwarz inequality, we obtain

$$2 \left( \dot{\theta}_\phi(t_{n-1}), \dot{\theta}_\phi(t_{n-1}) \right)_L \leq \left\| \dot{\theta}_\phi(t_{n-1}) \right\|_L^2 + \left\| \dot{\theta}_\phi(t_{n-1}) \right\|_L^2,
$$

(44)
and

$$2\gamma^2 (\theta_\phi (t_{n-1}), \theta_\phi (t_{n-1}))_{L^2} \leq \gamma^2 \| \theta_\phi (t_{n-1}) \|_{L^2}^2 + \gamma^2 \| \theta_\phi (t_{n-1}) \|_{L^2}^2.$$  

(45)

Note that

$$A(\phi(t)e^{\gamma t}; W(t) - \Pi_k W(t), \dot{\theta}_\phi(t)) = \frac{d}{dt} A(\phi(t)e^{\gamma t}; W(t) - \Pi_k W(t), \theta_\phi(t)) - \dot{A}(\phi(t)e^{\gamma t}; \dot{\theta}_\phi(t)) - \sum_{i,\alpha,\beta = 1} d \int_0^1 \left( \partial_t A_{i\alpha j\beta} \partial_\beta (W(t) - \Pi_k W(t)_j), \partial_\alpha \theta_{\phi,i}(t) \right)_{L^2} d\tau.$$  

Using the fact that \((W - \Pi_k W)(t_{n-1}^+) = (W - \Pi_k W)(t_{n-1}^-) = 0\) for \(n = 1, 2, \ldots, N\), we have

$$\int_{t_{n-1}}^{t_n} A(\phi(t)e^{\gamma t}; W(t) - \Pi_k W(t), \dot{\theta}_\phi(t)) \, dt = -\int_{t_{n-1}}^{t_n} A(\phi(t)e^{\gamma t}; \dot{\theta}_\phi(t)) \, dt$$

$$+ \sum_{i,\alpha,\beta = 1} d \int_0^1 \left( \partial_t A_{i\alpha j\beta} \partial_\beta (W(t) - \Pi_k W(t)_j), \partial_\alpha \theta_{\phi,i}(t) \right)_{L^2} d\tau dt.$$  

(46)

Then

$$\left| \int_{t_{n-1}}^{t_n} A(\phi(t)e^{\gamma t}; W(t) - \Pi_k W(t), \dot{\theta}_\phi(t)) \, dt \right| \leq \left| -\int_{t_{n-1}}^{t_n} A(\phi(t)e^{\gamma t}; \dot{\theta}_\phi(t)) \, dt \right|$$

$$+ \sum_{i,\alpha,\beta = 1} d \int_0^1 \left( \partial_t A_{i\alpha j\beta} \partial_\beta (W(t) - \Pi_k W(t)_j), \partial_\alpha \theta_{\phi,i}(t) \right)_{L^2} d\tau dt$$

$$\leq \int_{t_{n-1}}^{t_n} K_{\delta} \| \nabla \dot{\theta}_\phi(t) \|_{L^2} \| \nabla \theta_\phi(t) \|_{L^2} \, dt + C_r \int_{t_{n-1}}^{t_n} \| \nabla (W(t) - \Pi_k W(t)) \|_{L^2} \| \nabla \theta_\phi(t) \|_{L^2} \, dt$$

$$\leq \left( \frac{K_{\delta}}{2} + \frac{C_r}{2} \right) \int_{t_{n-1}}^{t_n} \| \nabla \theta_\phi(t) \|_{L^2}^2 \, dt + \frac{C_r}{2} \int_{t_{n-1}}^{t_n} \| \nabla (W(t) - \Pi_k W(t)) \|_{L^2}^2 \, dt$$

$$+ C_r \int_{t_{n-1}}^{t_n} \| \nabla (W(t) - \Pi_k W(t)) \|_{L^2}^2 \, dt$$

$$\leq \left( \frac{K_{\delta}}{2} + \frac{C_r}{2} \right) \int_{t_{n-1}}^{t_n} \| \nabla \theta_\phi(t) \|_{L^2}^2 \, dt + C_r \left( \frac{2^{q_n+1}}{q_n 2^{(s-1)}} \right)$$

for \(W \in W^{s,\infty}([0, T]; H_0^1)\).  

(47)

To bound the terms involving \(\partial_t A_{i\alpha j\beta}^r\), we apply Lemma 4.5 to get

$$\int_{t_{n-1}}^{t_n} \sum_{i,\alpha,\beta = 1} d \int_0^1 \left( \partial_t A_{i\alpha j\beta} \partial_\beta \theta_{\phi,j}(t), \partial_\alpha \theta_{\phi,i}(t) \right)_{L^2} d\tau \, dt \leq C_r \int_{t_{n-1}}^{t_n} \| \nabla \theta_\phi(t) \|_{L^2}^2 \, dt.$$  

(48)

Combining (42)–(48), we obtain

$$\left\| \dot{\theta}_\phi (t_{n-1}) \right\|_{L^2}^2 + \gamma^2 \left\| \theta_\phi (t_{n-1}) \right\|_{L^2}^2 + \gamma \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi (t) \right\|_{L^2}^2 \, dt + A(\phi(t_{n-1})e^{\gamma t}; \theta_\phi (t_{n-1}), \theta_\phi (t_{n-1}))$$

$$\leq A(\phi(t_{n-1})\theta_\phi (t_{n-1}), \theta_\phi (t_{n-1})) + \tilde{C} \int_{t_{n-1}}^{t_n} \left\| \nabla \theta_\phi (t) \right\|_{L^2}^2 \, dt + \left\| \dot{\theta}_\phi (t_{n-1}) \right\|_{L^2}^2 + \gamma^2 \left\| \theta_\phi (t_{n-1}) \right\|_{L^2}^2.$$
Summing up over \(n = 1, \ldots, j\), we obtain
\[
\left\| \dot{\theta}_\phi(t_j) \right\|_{L^2}^2 + \gamma^2 \left\| \theta_\phi(t_j) \right\|_{L^2}^2 + \gamma \sum_{n=1}^j \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_{L^2}^2 \, dt + A(\phi(t_j)e^{\gamma t_j}; \theta_\phi(t_j), \theta_\phi(t_j)) \leq \tilde{C} \sum_{n=1}^j \int_{t_{n-1}}^{t_n} \left\| \nabla \theta_\phi(t) \right\|_{L^2}^2 \, dt + C(v) \sum_{n=1}^j \left( k_n h^{2r+2} + \frac{k_{2q_n+1}}{q_n^{2(s-1)}} \right).
\]

Using the coercivity of \(A(\phi(t_j)e^{\gamma t_j}; \theta_\phi(t_j), \theta_\phi(t_j))\) (i.e., assumption (S2b)), we have
\[
\left\| \dot{\theta}_\phi(t_j) \right\|_{L^2}^2 + \gamma^2 \left\| \theta_\phi(t_j) \right\|_{L^2}^2 + \gamma \sum_{n=1}^j \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_{L^2}^2 \, dt + M_1 \left\| \nabla \theta_\phi(t_j) \right\|_{L^2}^2 \leq \tilde{C} \sum_{n=1}^j \int_{t_{n-1}}^{t_n} \left\| \nabla \theta_\phi(t) \right\|_{L^2}^2 + 2C_0^2 h^{-2} k_n \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_{L^2}^2 \, dt.
\]

By the fundamental theorem of calculus and the triangle inequality, we have for each \(t \in I_n\), with \(n = 1, \ldots, N\),
\[
\left\| \nabla \theta_\phi(t) \right\|_{L^2}^2 \leq \left( \left\| \nabla \theta_\phi(t_n) \right\|_{L^2} + \int_{t_{n-1}}^{t_n} \left\| \nabla \theta_\phi(t) \right\|_{L^2} \, dt \right)^2 \geq 2 \left\| \nabla \theta_\phi(t_n) \right\|_{L^2}^2 + 2 C_0^2 h^{-2} k_n \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_{L^2}^2 \, dt.
\]

Substituting (62) into (51), we have
\[
\left\| \dot{\theta}_\phi(t_j) \right\|_{L^2}^2 + \gamma^2 \left\| \theta_\phi(t_j) \right\|_{L^2}^2 + \gamma \sum_{n=1}^j \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_{L^2}^2 \, dt + M_1 \left\| \nabla \theta_\phi(t_j) \right\|_{L^2}^2 \leq 2 \tilde{C} \sum_{n=1}^{j-1} k_n \left\| \nabla \theta_\phi(t_n) \right\|_{L^2}^2 + 2 \tilde{C} k_j \left\| \nabla \theta_\phi(t_j) \right\|_{L^2}^2 + 2 \tilde{C} C_0^2 \sum_{n=1}^j h^{-2} k_n \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_{L^2}^2 \, dt + C(v) \sum_{n=1}^j \left( k_n h^{2r+2} + \frac{k_{2q_n+1}}{q_n^{2(s-1)}} \right)
\]
\[
\leq 2 \tilde{C} \sum_{n=1}^{j-1} k_n \left\| \nabla \theta_\phi(t_n) \right\|_{L^2}^2 + 2 \tilde{C} k_j \left\| \nabla \theta_\phi(t_j) \right\|_{L^2}^2 + \tilde{C} \sum_{n=1}^j k_n \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_{L^2}^2 \, dt + C(v) \sum_{n=1}^j \left( k_n h^{2r+2} + \frac{k_{2q_n+1}}{q_n^{2(s-1)}} \right),
\]

where the last inequality follows from the assumption that \(\mu_i k_i \leq h^2\) for each \(i = 1, \ldots, N\), with \(\tilde{C} = 2\tilde{C} C_0^2 \max_{1 \leq i \leq j} \frac{1}{\mu_i}\). The term \(2\tilde{C} k_j \left\| \nabla \theta_\phi(t_j) \right\|_{L^2}^2\) and the sum of integrals on the right-hand side of (53) can be absorbed into the third and fourth terms of the left-hand side of (53) if we choose each time step \(k_n\) is sufficiently small. That is,
\[
\left\| \dot{\theta}_\phi(t_j) \right\|_{L^2}^2 + \gamma^2 \left\| \theta_\phi(t_j) \right\|_{L^2}^2 + \sum_{n=1}^j \left( \gamma - \tilde{C} k_n \right) \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_\phi(t) \right\|_{L^2}^2 \, dt + \left( M_1 - 2\tilde{C} k_n \right) \left\| \nabla \theta_\phi(t_n) \right\|_{L^2}^2.
\]
\[
\begin{align*}
\leq 2\tilde{C} \sum_{n=1}^{j-1} k_n \|
abla \theta_\phi(t_n^-)\|_{L^2} + C(v) \sum_{n=1}^{j} \left( k_n h^{2r+2} + \frac{k_{2q_n+1}}{q_n^{2(s-1)}} \right),
\end{align*}
\]

By choosing \( k_n \leq \min\left\{ \frac{2}{\tilde{C}}, \frac{M_n}{4\tilde{C}} \right\} \) for each \( n = 1, \ldots, N \) and applying the discrete Grönwall lemma, we have

\[
\begin{align*}
\| \dot{\theta}_\phi(t_j^-) \|_{L^2}^2 + \| \theta_\phi(t_j^-) \|_{L^2}^2 + \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \| \dot{\theta}_\phi(t) \|_{L^2}^2 \, dt + \| \nabla \theta_\phi(t_j^-) \|_{L^2}^2 \leq C_j(v) \exp\left( C \sum_{n=1}^{j} \left( k_n h^{2r+2} + \frac{k_{2q_n+1}}{q_n^{2(s-1)}} \right) \right) \leq C_{\max}(v) \sum_{n=1}^{j} \left( k_n h^{2r+2} + \frac{k_{2q_n+1}}{q_n^{2(s-1)}} \right),
\end{align*}
\]

where \( C_{\max}(v) = \max_{1 \leq n \leq N} C_n(v) \exp(CT) \). Now tracing back constants through the previous estimates, we notice that \( C_{\max}(v) \) does not depend on \( C_n(v) \), so we can define \( C_*(v) := C_{\max}(v) \). Note that

\[
\| \nabla v_\phi(t) - \nabla v(t) \|_{L^\infty} \leq \| \nabla v_\phi(t) - \nabla W(t) \|_{L^\infty} + \| \nabla v(t) - \nabla W(t) \|_{L^\infty}.
\]

By the inverse estimate (ii,b), the error bound (iii,a), and the approximation properties of \( P_h \) in the \( W^{1,\infty} \) and \( H^1 \) semi-norms, we can find an \( h_1 > 0 \) such that, for \( h < h_1 \),

\[
\| \nabla W(t) - \nabla v(t) \|_{L^\infty} \leq C_1 h^{-\frac{d}{2}} \| \nabla (W(t) - P_h v(t)) \|_{L^\infty} + \| \nabla P_h v(t) - v(t) \|_{L^\infty} \leq C_1 h^{-\frac{d}{2}} \| \nabla (W(t) - v(t)) \|_{L^2} + C(v) h^{r-\frac{d}{2}} \]

\[
\leq C(v) h^{r-\frac{d}{2}} \leq \frac{\delta}{2} e^{-\gamma T}.
\]

Since \( r > \frac{d}{2} + 1, k_i^{q_i-\frac{1}{2}} = o\left( h^{1+\frac{d}{2}} \right) \) and \( \mu_i k_i \leq h^2 \) for each \( i = 1, \ldots, N \), we can also choose \( h_2 > 0 \) such that, for \( t \in I_n, n = 1, 2, \ldots, N, \) and \( h < h_2 \),

\[
\begin{align*}
\| \nabla v_\phi(t) - \nabla W(t) \|_{L^\infty} &\leq C_1 h^{-\frac{d}{2}} \left( \| \nabla \theta_\phi(t) \|_{L^2} + \| \nabla P_h W(t) - \nabla W(t) \|_{L^2} \right) \\
&\leq C_1 h^{-\frac{d}{2}} \left( \| \nabla \theta_\phi(t_n^-) \|_{L^2} + \int_{t_{n-1}}^{t_n} \| \partial_t \nabla \theta_\phi(t) \|_{L^2} \, dt \right) \\
&\quad + \| \nabla P_h W(t_n^-) - \nabla W(t_n^-) \|_{L^2} + \int_{t_{n-1}}^{t_n} \| \partial_t (\nabla P_h W - \nabla W)(t) \|_{L^2} \, dt \\
&\leq C_1 h^{-\frac{d}{2}} \left( \| \nabla \theta_\phi(t_n^-) \|_{L^2} + \int_{t_{n-1}}^{t_n} \| \partial_t \nabla \theta_\phi(t) \|_{L^2} \, dt \right) \\
&\quad + C_1 h^{-\frac{d}{2}} \sqrt{k_n} \left( \int_{t_{n-1}}^{t_n} \| \partial_t (\nabla P_h W - \nabla W)(t) \|_{L^2}^2 \, dt \right)^{\frac{1}{2}} \\
&\leq C_1 h^{-\frac{d}{2}} \left( \| \nabla \theta_\phi(t_n^-) \|_{L^2} + C_0 h^{-1} \sqrt{k_n} \int_{t_{n-1}}^{t_n} \| \partial_t \theta_\phi(t) \|_{L^2}^2 \, dt \right)^{\frac{1}{2}} + C(v) h^{-\frac{d}{2}} \frac{k_{2q_n+1}}{q_n^{2(s-1)}} \\
&\leq C(v) h^{-\frac{d}{2}} \left( \sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_{2q_n+1}}{q_n^{2(s-1)}} \right)^{\frac{1}{2}} \leq \frac{\delta}{2} e^{-\gamma T}.
\end{align*}
\]

By choosing \( h < h_* = \min\{ h_0, h_1, h_2 \} \), we obtain \( \nabla v_\phi e^{\gamma t} \in Z_\delta \).
4.3.2. Verification that $\mathcal{N}$ is a contraction mapping

To show the contraction property (b), we consider $\mathbf{R} = \phi - \phi'$ and $\Theta = \mathbf{v}_\phi - \mathbf{v}_{\phi'}$ where $\phi, \phi' \in \mathcal{F}$. Replacing $\phi$ in (40) by $\phi'$ and subtracting the new equation from (40), we have

$$
\int_{t_{n-1}}^{t_n} \left( \dot{\Theta}(t), \dot{\phi}(t) \right)_{L^2} dt + 2 \gamma \int_{t_{n-1}}^{t_n} \left( \dot{\Theta}(t), \dot{\phi}(t) \right)_{L^2} dt + \gamma^2 \int_{t_{n-1}}^{t_n} \left( \Theta(t), \dot{\phi}(t) \right)_{L^2} dt + \gamma^2 \int_{t_{n-1}}^{t_n} \left( \Theta(t), \dot{\phi}(t) \right)_{L^2} dt
$$

\begin{align*}
&+ \gamma^2 \left( \left[ \Theta(t) \right]_{n-1}, \dot{\phi}(t) \right)_{L^2} + \int_{t_{n-1}}^{t_n} \left( A(\phi(t)e^{\gamma t}; \theta(t), \dot{\phi}(t)) - A(\phi'(t)e^{\gamma t}; \theta'(t), \dot{\phi}'(t)) \right) dt \\
&+ \frac{1}{2} \left( A(\phi(t_n^{-1})e^{\gamma t_{n-1}}; \theta(t_n^{-1}), \dot{\phi}(t_n^{-1})) - A(\phi'(t_n^{-1})e^{\gamma t_{n-1}}; \theta'(t_n^{-1}), \dot{\phi}'(t_n^{-1})) \right) \\
&+ \frac{1}{2} \left( A(\phi(t_{n-1})e^{\gamma t_{n-1}}; \theta(t_{n-1}), \dot{\phi}(t_{n-1})) - A(\phi'(t_{n-1})e^{\gamma t_{n-1}}; \theta'(t_{n-1}), \dot{\phi}'(t_{n-1})) \right) \\
&= \int_{t_{n-1}}^{t_n} \left( A(\phi(t)e^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\phi}(t)) - A(\phi'(t)e^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\phi}'(t)) \right) dt \\
&+ \frac{1}{2} \left( A(\phi(t_{n-1})e^{\gamma t_{n-1}}; \theta(t_{n-1}), \dot{\phi}(t_{n-1})) - A(\phi'(t_{n-1})e^{\gamma t_{n-1}}; \theta'(t_{n-1}), \dot{\phi}'(t_{n-1})) \right). \\
\end{align*}

(56)

Taking $\phi = \Theta$ in equation (56), replacing

$$
A\left( \phi(t)e^{\gamma t}; \theta(t), \dot{\Theta}(t) \right) - A\left( \phi'(t)e^{\gamma t}; \theta'(t), \dot{\Theta}(t) \right)
$$

by

$$
A\left( \phi(t)e^{\gamma t}; \Theta(t), \dot{\Theta}(t) \right) + A\left( \phi(t)e^{\gamma t}; \theta(t), \dot{\Theta}(t) \right) - A\left( \phi'(t)e^{\gamma t}; \theta'(t), \dot{\Theta}(t) \right),
$$

and writing

$$
\int_{t_{n-1}}^{t_n} A\left( \phi(t)e^{\gamma t}; \Theta(t), \dot{\Theta}(t) \right) dt = \frac{1}{2} \int_{t_{n-1}}^{t_n} \frac{d}{dt} A(\phi(t)e^{\gamma t}; \Theta(t), \dot{\Theta}(t)) dt
$$

\begin{align*}
&- \frac{1}{2} \int_{t_{n-1}}^{t_n} \int_{0}^{1} \sum_{i, \alpha, \beta = 1}^{d} \left( \partial_t A_{\alpha \beta \gamma \delta} \partial_{\gamma \delta} \Theta(t), \partial_{\alpha \beta} \Theta(t) \right)_{L^2} dt, \\
\end{align*}

we have

\begin{align*}
&\left\| \dot{\Theta}(t_n) \right\|_{L^2}^2 + \left\| \dot{\Theta}(t_{n-1}^+) \right\|_{L^2}^2 + \gamma^2 \left\| \Theta(t_n) \right\|_{L^2}^2 + \gamma^2 \left\| \Theta(t_{n-1}^+) \right\|_{L^2}^2 + 4 \gamma \int_{t_{n-1}}^{t_n} \left\| \dot{\Theta}(t) \right\|_{L^2}^2 dt \\
&\quad + A\left( \phi(t_n^{-1})e^{\gamma t_n^{-1}}; \Theta(t_n^{-1}), \Theta(t_{n-1}) \right) - A\left( \phi'(t_n^{-1})e^{\gamma t_{n-1}}; \Theta(t_{n-1}), \Theta(t_{n-1}) \right) \\
&\quad = 2 \left( \dot{\Theta}(t_{n-1}^{-1}), \Theta(t_{n-1}) \right)_{L^2} + 2 \gamma \left( \Theta(t_{n-1}^{-1}), \Theta(t_{n-1}) \right)_{L^2} \\
&\quad + \int_{t_{n-1}}^{t_n} \int_{0}^{1} \sum_{i, \alpha, \beta = 1}^{d} \left( \partial_t A_{\alpha \beta \gamma \delta} \partial_{\gamma \delta} \Theta(t), \partial_{\alpha \beta} \Theta(t) \right)_{L^2} dt \quad \text{dt} \\
&\quad + 2 \int_{t_{n-1}}^{t_n} \left( A\left( \phi(t)e^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\Theta}(t) \right) - A\left( \phi'(t)e^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\Theta}(t) \right) \right) dt \\
&\quad + 2 \int_{t_{n-1}}^{t_n} \left( A\left( \phi'(t)e^{\gamma t}; \theta(t), \dot{\Theta}(t) \right) - A\left( \phi'(t)e^{\gamma t}; \theta'(t), \dot{\Theta}(t) \right) \right) dt \\
&\quad + \left( A\left( \phi(t_n^{-1})e^{\gamma t_{n-1}}; \theta(t_n^{-1}), \dot{\Theta}(t_{n-1}) \right) - A\left( \phi(t_n^{-1})e^{\gamma t_{n-1}}; \theta(t_n^{-1}), \dot{\Theta}(t_{n-1}) \right) \right) \\
&\quad + \left( A\left( \phi'(t_{n-1})e^{\gamma t_{n-1}}; \theta'(t_{n-1}), \dot{\Theta}(t_{n-1}) \right) - A\left( \phi'(t_{n-1})e^{\gamma t_{n-1}}; \theta'(t_{n-1}), \dot{\Theta}(t_{n-1}) \right) \right) \\
&\quad + \left( A\left( \phi(t_{n-1})e^{\gamma t_{n-1}}; \theta(t_{n-1}), \dot{\Theta}(t_{n-1}) \right) - A\left( \phi(t_{n-1})e^{\gamma t_{n-1}}; \theta(t_{n-1}), \dot{\Theta}(t_{n-1}) \right) \right).
\end{align*}
Note that
\[
\int_{t_{n-1}}^{t_n} A\left(\phi'(t)e^{\gamma t}; \theta'_\varphi(t), \Theta(t)\right) dt + A\left(\phi'(t_{n-1}^+)e^{\gamma t_{n-1}}; \theta'_0(t_{n-1}^+), \Theta(t_{n-1}^+)\right)
- A\left(\phi'(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \theta'_0(t_{n-1}^-), \Theta(t_{n-1}^-)\right)
= -\int_{t_{n-1}}^{t_n} A\left(\phi'(t)e^{\gamma t}; \theta'_\varphi(t), \Theta(t)\right) dt - \int_{t_{n-1}}^{t_n} \int_0^1 \sum_{i,\alpha,\beta=1}^d \left( \partial_i \tilde{A}_{i\alpha j\beta}^\gamma \partial_\beta \theta'_{\varphi,j}(t), \partial_\alpha \Theta_i(t) \right) L^2 dt
+ A\left(\phi'(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \theta'_\varphi(t_{n-1}^-), \Theta(t_{n-1}^-)\right) - A\left(\phi'(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \theta'_\varphi(t_{n-1}^-), \Theta(t_{n-1}^-)\right),
\]
where \( \tilde{A}_{i\alpha j\beta}^\gamma := A_{i\alpha j\beta}(\nabla \Theta(t) e^{\gamma t} + \tau(\nabla \phi'(t) - \nabla \Theta(t)) e^{\gamma t}) \). Analogously, we have
\[
-\int_{t_{n-1}}^{t_n} A\left(\phi(t)e^{\gamma t}; \theta'_\varphi(t), \Theta(t)\right) dt - A\left(\phi(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \theta'_\varphi(t_{n-1}^-), \Theta(t_{n-1}^-)\right)
+ A\left(\phi(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \theta'_\varphi(t_{n-1}^-), \Theta(t_{n-1}^-)\right)
= \int_{t_{n-1}}^{t_n} A\left(\phi(t)e^{\gamma t}; \theta'_\varphi(t), \Theta(t)\right) dt + \int_{t_{n-1}}^{t_n} \int_0^1 \sum_{i,\alpha,\beta=1}^d \left( \partial_i \tilde{A}_{i\alpha j\beta}^\gamma \partial_\beta \theta'_{\varphi,j}(t), \partial_\alpha \Theta_i(t) \right) L^2 dt
- A\left(\phi(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \theta'_\varphi(t_{n-1}^-), \Theta(t_{n-1}^-)\right) + A\left(\phi(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \theta'_\varphi(t_{n-1}^-), \Theta(t_{n-1}^-)\right).
\]

This implies that
\[
\left\|\dot{\Theta}(t_{n-1})\right\|^2_{L^2} + \left\|\dot{\Theta}(t_{n-1}^+)\right\|^2_{L^2} + \gamma^2 \left\|\Theta(t_{n-1})\right\|^2_{L^2} + \gamma^2 \left\|\Theta(t_{n-1}^-)\right\|^2_{L^2} + 4\gamma \int_{t_{n-1}}^{t_n} \left\|\dot{\Theta}(t)\right\|^2_{L^2} dt
+ A\left(\phi(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \Theta(t_{n-1}^-), \Theta(t_{n-1}^-)\right) - A\left(\phi(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \Theta(t_{n-1}^-), \Theta(t_{n-1}^-)\right)
= 2\left(\dot{\Theta}(t_{n-1}^-), \dot{\Theta}(t_{n-1}^+)\right)_{L^2} + 2\gamma^2 \left(\Theta(t_{n-1}^-), \Theta(t_{n-1}^-)\right)_{L^2}
\quad + \int_{t_{n-1}}^{t_n} \int_0^1 \sum_{i,\alpha,\beta=1}^d \left( \partial_i \tilde{A}_{i\alpha j\beta}^\gamma \partial_\beta \Theta_j(t), \partial_\alpha \Theta_i(t) \right) L^2 dt
+ \int_{t_{n-1}}^{t_n} \left( A\left(\phi'(t)e^{\gamma t}; \theta'_\varphi(t), \Theta(t)\right) - A\left(\phi(t)e^{\gamma t}; \theta'_\varphi(t), \Theta(t)\right) \right) dt
+ \int_{t_{n-1}}^{t_n} \left( A\left(\phi(t)e^{\gamma t}; \theta'_\varphi(t), \Theta(t)\right) - A\left(\phi'(t)e^{\gamma t}; \theta'_\varphi(t), \Theta(t)\right) \right) dt
+ 2 \int_{t_{n-1}}^{t_n} \left( A\left(\phi(t)e^{\gamma t}; W(t) - \Pi_k W(t), \dot{\Theta}(t)\right) - A\left(\phi'(t)e^{\gamma t}; W(t) - \Pi_k W(t), \dot{\Theta}(t)\right) \right) dt
\quad + \int_{t_{n-1}}^{t_n} \int_0^1 \sum_{i,\alpha,\beta=1}^d \left( \left( \partial_i \tilde{A}_{i\alpha j\beta}^\gamma - \partial_i \tilde{A}_{i\alpha j\beta}^\gamma \right) \partial_\beta \Theta_j(t), \partial_\alpha \Theta_i(t) \right) L^2 dt
+ A\left(\phi'(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \theta'_\varphi(t_{n-1}^-), \Theta(t_{n-1}^-)\right) - A\left(\phi'(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \theta'_\varphi(t_{n-1}^-), \Theta(t_{n-1}^-)\right)
- A\left(\phi(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \theta'_\varphi(t_{n-1}^-), \Theta(t_{n-1}^-)\right) + A\left(\phi(t_{n-1}^-)e^{\gamma t_{n-1}^-}; \theta'_\varphi(t_{n-1}^-), \Theta(t_{n-1}^-)\right).
\]

Again, we need to bound the terms on the right-hand side of the equation (57). By the Cauchy–Schwarz inequality, we have
\[
2\left(\dot{\Theta}(t_{n-1}^-), \dot{\Theta}(t_{n-1}^-)\right)_{L^2} \leq \left\|\dot{\Theta}(t_{n-1}^-)\right\|^2_{L^2} + \left\|\dot{\Theta}(t_{n-1}^-)\right\|^2_{L^2},
\]
(58)
By Lemma 4.5, for \( \phi \in F \), we have

\[
\int_{t^+_{n-1}}^{t_n} \int_0^1 \sum_{i,\alpha,j,\beta=1}^{d} (\partial_t A_{i\alpha j\beta} \partial_\beta \Theta_j(t), \partial_n \Theta_i(t))_{L^2} \, d\tau \, dt \leq C_\tau \int_{t^+_{n-1}}^{t_n} \| \nabla \Theta(t) \|^2_{L^2} \, dt.
\]

Recalling that the values of \( \nabla W(t)e^{\tau t} + \tau \nabla (\phi(t) - W(t))e^{\tau t} \) and \( \nabla W(t)e^{\tau t} + \tau \nabla (\phi'(t) - W(t))e^{\tau t} \) belong to the convex set \( M_\delta \), and that \( A_{i\alpha j\beta} \) is Lipschitz continuous on \( M_\delta \), we have

\[
\int_{t^+_{n-1}}^{t_n} \mathcal{A}\left(\phi'(t)e^{\tau t}; \theta_{\phi'}(t), \hat{\Theta}(t)\right) - \mathcal{A}\left(\phi(t)e^{\tau t}; \theta_{\phi'}(t), \hat{\Theta}(t)\right) \, dt
\]

\[
\leq \int_{t^+_{n-1}}^{t_n} \left| \mathcal{A}\left(\phi'(t)e^{\tau t}; \theta_{\phi'}(t), \hat{\Theta}(t)\right) - \mathcal{A}\left(\phi(t)e^{\tau t}; \theta_{\phi'}(t), \hat{\Theta}(t)\right) \right| \, dt
\]

\[
\leq L_\delta \int_{t^+_{n-1}}^{t_n} \| \nabla \phi(t) - \nabla \phi'(t) \|_{L^2} \| \nabla \theta_{\phi'}(t) \|_{L^\infty} \left\| \nabla \hat{\Theta}(t) \right\|_{L^2} \, dt
\]

\[
\leq L_\delta C_0 C_1 \int_{t^+_{n-1}}^{t_n} \| \nabla R(t) \|_{L^2} h^{-1} \| \nabla \theta_{\phi'}(t) \|_{L^2} \left\| \nabla \hat{\Theta}(t) \right\|_{L^2} \, dt
\]

\[
\leq 2\gamma \int_{t^+_{n-1}}^{t_n} \left\| \hat{\Theta}(t) \right\|^2_{L^2} \, dt + C(\gamma) \max_{t \in I_n} \| \nabla R(t) \|_{L^2} h^{-2-d} \int_{t^+_{n-1}}^{t_n} \| \nabla \theta_{\phi'}(t) \|_{L^2}^2 \, dt,
\]

where we have used the inverse inequalities (ii,a) and (ii,b), Young’s inequality. To approximate the \( \nabla \theta_{\phi'} \) term, we apply the fundamental theorem of calculus as in the proof of (a) to have

\[
\int_{t^+_{n-1}}^{t_n} \| \nabla \theta_{\phi'}(t) \|_{L^2}^2 \, dt \leq 2k_n \| \nabla \theta_{\phi'}(t^+_{n-1}) \|^2_{L^2} + 2C_0^2 k_n^2 h^{-2} \int_{t^+_{n-1}}^{t_n} \left\| \hat{\theta}_{\phi'}(t) \right\|^2_{L^2} \, dt.
\]
Applying the Lipschitz continuity of \( A_{ioj\beta} \) again, we have

\[
2 \int_{t_{n-1}}^{t_n} \left| \left( A\left( \phi(t)e^{\gamma t}; W(t) - \Pi_k W(t), \dot{\Theta}(t) \right) - A\left( \phi'(t)e^{\gamma t}; W(t) - \Pi_k W(t), \dot{\Theta}(t) \right) \right) \right| dt \\
\leq 2 \int_{t_{n-1}}^{t_n} \| \nabla R(t) \|_{L^2} \| \nabla W(t) - \nabla (\Pi_k W)(t) \|_{L^\infty} \| \dot{\Theta}(t) \|_{L^2} dt \\
\leq 2L\delta \int_{t_{n-1}}^{t_n} \| \nabla R(t) \|_{L^2} \| \nabla W(t) - \nabla (\Pi_k W)(t) \|_{L^\infty} \| \dot{\Theta}(t) \|_{L^2} dt \\
\leq 2L\delta C_0 C_1 h^{-\frac{d}{2} - 1} \int_{t_{n-1}}^{t_n} \| \nabla R(t) \|_{L^2} \| \nabla W(t) - \nabla (\Pi_k W)(t) \|_{L^\infty} \| \dot{\Theta}(t) \|_{L^2} dt \\
\leq \gamma \int_{t_{n-1}}^{t_n} \| \dot{\Theta}(t) \|_{L^2}^2 dt + C(\gamma, W) h^{-d-2} \max_{t \in [0, T]} \| \nabla R(t) \|_{L^2} \|

\text{for } W \in W^{s, \infty}([0, T]; H_0^s),
\]

where \( C(\gamma, W) \) is a positive constant depending on both \( \gamma \) and the nonlinear projection \( W \). Next, we need to bound the term involving \( \partial_t A_{ioj\beta} - \partial_t \bar{A}_{ioj\beta} \). Recall that

\[
\partial_t A_{ioj\beta} := \partial_t A_{ioj\beta} (\nabla W(t)e^{\gamma t} + \tau (\nabla \phi(t) - \nabla W(t)e^{\gamma t})
\]

and

\[
\partial_t \bar{A}_{ioj\beta} := \partial_t \bar{A}_{ioj\beta} (\nabla W(t)e^{\gamma t} + \tau (\nabla \phi'(t) - \nabla W(t)e^{\gamma t}).
\]

By Taylor’s theorem with an integral remainder, we have

\[
\partial_t A_{ioj\beta} - \partial_t \bar{A}_{ioj\beta} = \int_0^1 \sum_{k, l, \delta = 1}^d \frac{\partial^2}{\partial \eta_k \partial \eta_l \partial \delta} A_{ioj\beta} (\nabla W(t)e^{\gamma t} + \tau (\nabla \phi(t) - \nabla W(t)e^{\gamma t} + \tau (\nabla \phi(t) - \nabla W(t)e^{\gamma t})
\]

\[
\times \partial_t \partial_\delta (W_1(t)e^{\gamma t} + \tau (\phi'(t) - W_1(t)e^{\gamma t} + \tau (\phi(t) - W_1(t)e^{\gamma t}) \tau \partial_\gamma (\phi_k(t) - \phi'_k(t)) d\tau.
\]

Since \( A_{ioj\beta} \) is sufficiently smooth (in particular, twice continuously differentiable), we can estimate the above difference term by

\[
\left| \partial_t A_{ioj\beta} - \partial_t \bar{A}_{ioj\beta} \right| \leq \tilde{C} \| \nabla R(t) \|_{L^{\infty}} \left( \| \nabla W(t) \|_{L^\infty} + \| \nabla \dot{W}(t) \|_{L^\infty} \right) \\
+ \tilde{C} \| \nabla R(t) \|_{L^{\infty}} \left( \| \nabla \phi(t) - \nabla \phi'(t) \|_{L^\infty} + \| \nabla \dot{\phi}(t) - \nabla \dot{\phi}'(t) \|_{L^\infty} \right) \\
+ \tilde{C} \| \nabla R(t) \|_{L^{\infty}} \left( \| \nabla \phi' - \nabla W(t) \|_{L^\infty} + \| \nabla \dot{\phi}'(t) - \nabla \dot{W}(t) \|_{L^\infty} \right).
\]

Similarly to the proof of Lemma 4.5, we can show that

\[
\| \nabla \phi' - \nabla W(t) \|_{L^{\infty}} + \| \nabla \dot{\phi}'(t) - \nabla \dot{W}(t) \|_{L^\infty} \leq C_\tau.
\]

Property (iii,c) of the nonlinear projection \( W \) implies that

\[
\| \nabla W(t) \|_{L^\infty} + \| \nabla \dot{W}(t) \|_{L^\infty} \leq c_0 + c_1.
\]
In the view of the triangle inequality, we have
\[
\| \nabla \phi(t) - \nabla \phi'(t) \|_{L^\infty} + \| \nabla \phi(t) - \nabla \phi'(t) \|_{L^\infty} \\
\leq \| \nabla \phi(t) - \nabla \hat{W}(t) \|_{L^\infty} + \| \nabla \phi'(t) - \nabla \hat{W}(t) \|_{L^\infty} + \| \nabla \phi(t) - \nabla \hat{W}(t) \|_{L^\infty} \\
+ \| \nabla \phi'(t) - \nabla \hat{W}(t) \|_{L^\infty} \\
\leq 2C_\tau.
\]

Combining (66)–(68) and applying the inverse inequality (ii,b), Young’s inequality, we obtain
\[
\int_{t_{n-1}}^{t_n} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d \left( (\partial_t A_{\alpha j}^\beta - \partial_t A_{\alpha j}^\beta) \partial_\beta \theta_{\phi', j}(t), \partial_\alpha \Theta_i(t) \right) L^2 d\tau dt \\
\leq \hat{C}_{\text{Lip}} \int_{t_{n-1}}^{t_n} \| \nabla \hat{R}(t) \|_{L^\infty} \| \nabla \theta_{\phi', j}(t) \|_{L^2} \| \nabla \Theta(t) \|_{L^2} dt \\
\leq \hat{C}_{\text{Lip}} C_1 h^{-\frac{2}{3}} \int_{t_{n-1}}^{t_n} \| \nabla \Theta(t) \|_{L^2} \| \nabla \theta_{\phi', j}(t) \|_{L^2} \| \nabla \Theta(t) \|_{L^2} dt \\
\leq \hat{C} \int_{t_{n-1}}^{t_n} \| \nabla \Theta(t) \|_{L^2}^2 dt + C h^{-d} \int_{t_{n-1}}^{t_n} \| \nabla \Theta(t) \|_{L^2}^2 \| \nabla \theta_{\phi', j}(t) \|_{L^2}^2 dt \\
\leq \hat{C} \int_{t_{n-1}}^{t_n} \| \nabla \Theta(t) \|_{L^2}^2 dt + 2C \max_{t \in I_n} \| \nabla \hat{R}(t) \|_{L^2}^2 h^{-d} k_n \left( \| \nabla \theta_{\phi', j}(t) \|_{L^2}^2 + \int_{t_{n-1}}^{t_n} \| \dot{\theta}_{\phi', j}(t) \|_{L^2}^2 dt \right),
\]
where \( \hat{C}_{\text{Lip}} = \hat{C}(3C_\tau + c_0 + c_1) \) and \( C \) is a generic positive constant. Combining the estimates (58)–(65) and (69), we obtain
\[
\| \dot{\Theta}(t_n) \|_{L^2}^2 + \gamma_2 \| \Theta(t_n) \|_{L^2}^2 + A(\phi(t_n^+) e^{\gamma t_n}; \Theta(t_n), \Theta(t_n)) + \gamma \int_{t_{n-1}}^{t_n} \| \dot{\Theta}(t) \|_{L^2}^2 dt \\
\leq A(\phi(t_n^+) e^{\gamma t_n}; \Theta(t_n), \Theta(t_n)) + \| \dot{\Theta}(t_{n-1}) \|_{L^2}^2 + \gamma_2 \| \Theta(t_{n-1}) \|_{L^2}^2 + C \int_{t_{n-1}}^{t_n} \| \nabla \Theta(t) \|_{L^2}^2 dt \\
+ C(h^{-d-2} + h^{-d}) k_n \left( \| \nabla \theta_{\phi', j}(t_n) \|_{L^2}^2 + \int_{t_{n-1}}^{t_n} \| \dot{\theta}_{\phi', j}(t) \|_{L^2}^2 dt \right) \max_{t \in I_n} \| \nabla \hat{R}(t) \|_{L^2}^2 \\
+ Ch^{-d-2} \left( \int_{t_{n-1}}^{t_n} \| \dot{\theta}_{\phi', j}(t) \|_{L^2}^2 dt \right) \max_{t \in I_n} \| \nabla \hat{R}(t) \|_{L^2}^2 + C(\gamma, W) h^{-d-2} \frac{k_n^{2+3}}{q_{2n-1}^n} \max_{t \in I_n} \| \nabla \hat{R}(t) \|_{L^2}^2 \\
+ A(\phi'(t_n^+) e^{\gamma t_n}; \Theta'(t_n), \Theta(t_n)) - A(\phi(t_n^+) e^{\gamma t_n}; \Theta(t_n), \Theta(t_n)) \\
- A(\phi(t_{n-1}^+) e^{\gamma t_{n-1}}; \Theta(t_{n-1}), \Theta(t_{n-1})) + A(\phi(t_{n-1}^+) e^{\gamma t_{n-1}}; \Theta(t_{n-1}), \Theta(t_{n-1})).
\]

Summing up over \( n = 1, \ldots, j \), we have
\[
\| \dot{\Theta}(t_j^-) \|_{L^2}^2 + \gamma_2 \| \Theta(t_j^-) \|_{L^2}^2 + A(\phi(t_j^-) e^{\gamma t_j}; \Theta(t_j^-), \Theta(t_j^-)) + \gamma \sum_{n=1}^j \int_{t_{n-1}}^{t_n} \| \dot{\Theta}(t) \|_{L^2}^2 dt \\
\leq C \sum_{n=1}^j \int_{t_{n-1}}^{t_n} \| \nabla \Theta(t) \|_{L^2}^2 dt + C(h^{-d-2} + h^{-d}) \left( \sum_{n=1}^j k_n \| \nabla \theta_{\phi', j}(t_n) \|_{L^2}^2 \right) \max_{t \in I_n, 1 \leq n \leq j} \| \nabla \hat{R}(t) \|_{L^2}^2
\]
By applying Young's inequality on the right-hand side of (73), we have

\[ C \left( h^{-d-2} + h^{-d} \right) \left( \sum_{n=1}^{\gamma} k_n \int_{t_{n-1}}^{t_n} \left\| \dot{\theta}_{\phi}^i(t) \right\|^2 L^2 dt \right) + \max_{t \in I_n, 1 \leq n \leq j} \left\| \nabla R(t) \right\|^2 L^2.

\]

Combining the estimates (71) and (72) and applying the assumption (S2b) to inequality (ii,b), we obtain

\[ C \left( \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \left\| \nabla \Theta(t) \right\|^2 L^2 dt \right) + C \left( \nu^{-d-2} + \nu^{-d} \right) \left( \sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_{2q_n}^{2q_n+1}}{q_n^{2(2s-1)}} \right) \max_{t \in I_n, 1 \leq n \leq j} \left\| \nabla R(t) \right\|^2 L^2.

\]

\[ + A \left( \phi'(t_j) \right) e^{\gamma t_j}; \theta_{\phi}^i(t_j), \Theta(t_j) \right) - A \left( \phi(t_j) \right) e^{\gamma t_j}; \theta_{\phi}^i(t_j), \Theta(t_j) \right)

\]

\[ \leq C \left( \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \left\| \nabla \Theta(t) \right\|^2 L^2 dt \right) + C \left( \nu^{-d-2} + \nu^{-d} \right) \left( \sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_{2q_n}^{2q_n+1}}{q_n^{2(2s-1)}} \right) \max_{t \in I_n, 1 \leq n \leq j} \left\| \nabla R(t) \right\|^2 L^2

\]

where \( C \) is a generic positive constant and \( C(\nu) \) is a positive constant depending on the exact solution \( \nu \). These constants may change from line to line. Using the Lipschitz continuity of \( A(\cdot; \theta_{\phi}^i(t_j), \Theta(t_j)) \) and the inverse inequality (ii,b), we obtain

\[ \left| A \left( \phi'(t_j) \right) e^{\gamma t_j}; \theta_{\phi}^i(t_j), \Theta(t_j) \right) - A \left( \phi(t_j) \right) e^{\gamma t_j}; \theta_{\phi}^i(t_j), \Theta(t_j) \right) \right|

\[ \leq L_8 \left\| \nabla R(t_j) \right\| L^2 \left\| \nabla \Theta(t_j) \right\| L^\infty \left\| \nabla \Theta(t_j) \right\| L^2

\]

\[ \leq C \nu^{-d} \left( \sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_{2q_n}^{2q_n+1}}{q_n^{2(2s-1)}} \right) \left( \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \left\| \nabla \Theta(t) \right\|^2 L^2 dt \right) + C \left( \nu^{-d-2} + \nu^{-d} \right) \left( \sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_{2q_n}^{2q_n+1}}{q_n^{2(2s-1)}} \right) \max_{t \in I_n, 1 \leq n \leq j} \left\| \nabla R(t) \right\|^2 L^2

\]

Combining the estimates (71) and (72) and applying the assumption (S2b) to \( A \left( \phi(t_j) \right) e^{\gamma t_j}; \theta_{\phi}^i(t_j), \Theta(t_j) \right) \) on the left-hand-side of the resulting inequality yield

\[ \left\| \dot{\Theta}(t_j) \right\|^2 L^2 + \frac{1}{2} \left\| \Theta(t_j) \right\|^2 L^2 + M_1 \left\| \nabla \Theta(t_j) \right\| L^2 \left( \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \left\| \nabla \Theta(t) \right\|^2 L^2 dt \right)

\]

\[ \leq C \left( \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \left\| \nabla \Theta(t) \right\|^2 L^2 dt \right) + C \left( \nu^{-d-2} + \nu^{-d} \right) \left( \sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_{2q_n}^{2q_n+1}}{q_n^{2(2s-1)}} \right) \max_{t \in I_n, 1 \leq n \leq j} \left\| \nabla R(t) \right\|^2 L^2

\]

By applying Young’s inequality on the right-hand side of (73), we have

\[ \left\| \dot{\Theta}(t_j) \right\|^2 L^2 + \frac{1}{2} \left\| \Theta(t_j) \right\|^2 L^2 + M_1 \left\| \nabla \Theta(t_j) \right\| L^2 \left( \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \left\| \nabla \Theta(t) \right\|^2 L^2 dt \right)

\]
\[
\frac{d}{dt} \theta(t) + \sum_{i=1}^{N} \frac{\theta_i(t)^2}{2} \leq \frac{M_1}{2} \frac{\|\nabla \theta(t)\|_{L^2}^2}{\epsilon} + \frac{1}{\epsilon} \max_{t \in I_n, 1 \leq n \leq j} \|\nabla R(t)\|_{L^2},
\]

where \( C(M_1, v) \) is a constant depending on \( M_1 \) and the exact solution \( v \). This implies that

\[
\left\| \frac{d}{dt} \theta_j(t) \right\|_{L^2}^2 + \left\| \theta_j(t) \right\|_{L^2}^2 + \left\| \nabla \theta_j(t) \right\|_{L^2}^2 + \left\| \nabla \theta_j(t) \right\|_{L^2}^2 \leq \frac{C_1(v)}{t} \left( \sum_{i=1}^{N} h_i^2 \right) + \frac{1}{\epsilon} \max_{t \in I_n, 1 \leq n \leq j} \|\nabla R(t)\|_{L^2}.
\]

By an analogous application of the discrete Grönwall lemma as in the proof of (a), we can deduce that, for \( k_n \) sufficiently small for each \( n = 1, \ldots, j \),

\[
\left\| \frac{d}{dt} \theta_j(t) \right\|_{L^2}^2 + \left\| \theta_j(t) \right\|_{L^2}^2 + \left\| \nabla \theta_j(t) \right\|_{L^2}^2 + \left\| \nabla \theta_j(t) \right\|_{L^2}^2 \leq \frac{C_1(v)}{t} \left( \sum_{i=1}^{N} h_i^2 \right) + \frac{1}{\epsilon} \max_{t \in I_n, 1 \leq n \leq j} \|\nabla R(t)\|_{L^2},
\]

where \( \tilde{C}(v) \) is a positive constant depending on \( v \) which may vary from line to line. By the fundamental theorem of calculus and the triangle inequality, we have for each \( t \in I_n \), with \( n = 1, \ldots, N \),

\[
\|\nabla \theta(t)\|_{L^2} \leq \|\nabla \theta(t_n)\|_{L^2} + \int_{t_{n-1}}^{t_n} \left\| \nabla \frac{d}{dt} \theta(t) \right\|_{L^2} dt
\]

\[
\leq \|\nabla \theta(t_n)\|_{L^2} + C_0 h^{-1} \int_{t_{n-1}}^{t_n} \left\| \frac{d}{dt} \theta(t) \right\|_{L^2} dt
\]

\[
\leq \frac{1}{2} \left( t_n - t_{n-1} \right) \left( \frac{\|\nabla \theta(t)\|_{L^2}}{\epsilon} \right)^2 dt
\]

This implies that, for each \( t \in I_n \) with \( 1 \leq n \leq N \),

\[
\|\nabla \theta(t)\|_{L^2} \leq \left( 1 + C_0 h^{-1} k_n^{-\frac{1}{2}} \right) \tilde{C}(v)^{\frac{1}{2}} h^{-\frac{1}{2}} \left( \sum_{i=1}^{N} h_i \right)^{\frac{1}{2}} \max_{t \in I_n, 1 \leq n \leq N} \|\nabla R(t)\|_{L^2}
\]

\[
\leq \tilde{C}_1(v)^{\frac{1}{2}} h^{-\frac{1}{2}} \left( \sum_{i=1}^{N} h_i \right)^{\frac{1}{2}} \max_{t \in I_n, 1 \leq n \leq N} \|\nabla R(t)\|_{L^2}.
\]
where the last inequality follows from the assumption that \( \mu, k \leq h^2 \) for each \( 1 \leq i \leq N \). Analogously, we have

\[
\| \dot{\Theta}(t) \|_{L^2} \leq \| \Theta(t_n) \|_{L^2} + k_n^{1/2} \left( \int_{t_{n-1}}^{t_n} \| \Theta(t) \|_{L^2}^2 \, dt \right)^{1/2} \\
\leq \tilde{C}_2(v) h^{-\frac{d}{2} - 1} \left( \sum_{i=1}^{N} k_i h^{2r+2} + \frac{k^{2q_i+1}}{q_i^{2(s-1)}} \right) \frac{1}{2} \max_{t \in I_n, 1 \leq n \leq N} \| \nabla R(t) \|_{L^2},
\]

and

\[
\| \dot{\Theta}(t) \|_{L^2} \leq \| \Theta(t_n) \|_{L^2} + \int_{t_{n-1}}^{t_n} \| \dot{\Theta}(t) \|_{L^2} \, dt \\
\leq \| \Theta(t_n) \|_{L^2} + C_2 k_n^{1/2} \left( \int_{t_{n-1}}^{t_n} \| \dot{\Theta}(t) \|_{L^2}^2 \, dt \right)^{1/2} \\
\leq \tilde{C}_3(v) h^{-\frac{d}{2} - 1} k_n^{1/2} \left( \sum_{i=1}^{N} k_i h^{2r+2} + \frac{k^{2q_i+1}}{q_i^{2(s-1)}} \right) \frac{1}{2} \max_{t \in I_n, 1 \leq n \leq N} \| \nabla R(t) \|_{L^2}.
\]

Summing up (77)–(79) and taking maximum on the left-hand for \( t \in I_n \), with \( 1 \leq n \leq N \), we have

\[
\max_{t \in I_n, 1 \leq n \leq N} \left( \| \Theta(t) \|_{H^1} + \| \dot{\Theta}(t) \|_{L^2} \right) \leq \tilde{C}(v) h^{-\frac{d}{2} - 1} \\max_{1 \leq n \leq N} k_n^{1/2} \left( \sum_{i=1}^{N} k_i h^{2r+2} + \frac{k^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} \max_{t \in I_n, 1 \leq n \leq N} \| \nabla R(t) \|_{L^2}.
\]

By choosing the mesh size \( h \) and time steps \( \{k_i\}_{i=1}^{N} \) \( < 1 \) for each \( i = 1, 2, \ldots, N \) small enough, and \( r \) and \( \{q_i\}_{i=1}^{N} \) for each \( i = 1, 2, \ldots, N \) large enough such that

\[
\tilde{C}(v) h^{-\frac{d}{2} - 1} \\max_{1 \leq n \leq N} k_n^{1/2} \left( \sum_{i=1}^{N} k_i h^{2r+2} + \frac{k^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} < 1,
\]

we obtain

\[
\max_{t \in I_n, 1 \leq n \leq N} \left( \| \Theta(t) \|_{H^1} + \| \dot{\Theta}(t) \|_{L^2} \right) \leq \max_{t \in I_n, 1 \leq n \leq N} \left( \| R(t) \|_{H^1} + \| R(t) \|_{L^2} \right).
\]

Indeed, the inequality (81) follows from our assumptions that \( r > \frac{d}{2} + 1, \) \( k_i^{\frac{d}{2} - \frac{1}{2}} = o(h^{1/2}) \) for each \( i = 1, 2, \ldots, N \). Therefore, by Banach’s fixed point theorem, \( v_{\text{DG}} = v_{\phi} \) is the unique solution to (19). By the triangle inequality, equations (23), (26) and property (iii,b) of the nonlinear projection \( W \), we have

\[
\| v_{\text{DG}}(t_j) - v(t_j) \|_{L^2} + \| \dot{v}_{\text{DG}}(t_j) - \dot{v}(t_j) \|_{L^2} \\
\leq \| \Theta(t_j) \|_{L^2} + \| \dot{\Theta}(t_j) \|_{L^2} + \| W(t_j) - v(t_j) \|_{L^2} + \| \dot{W}(t_j) - \dot{v}(t_j) \|_{L^2} \\
\leq C(v) \left( \sum_{n=1}^{j} k_n h^{2r+2} + \frac{k^{2q_n+1}}{q_n^{2(s-1)}} \right)^{\frac{1}{2}} + 2\tilde{C}_r(v) h^{r+1} \quad \text{(by (iii,b))} \\
\leq C(v) \left( h^{2r+2} + \sum_{n=1}^{j} \frac{k^{2q_n+1}}{q_n^{2(s-1)}} \right)^{\frac{1}{2}}.
\]
5. Numerical experiments

In this section, we show some numerical experiments on a simple version of the nonlinear elastodynamics equation to verify the error estimates proved in Section 4.

5.1. Numerical results for a one-dimensional nonlinear elastodynamics problem

We consider the one-dimensional nonlinear equation

\begin{align}
\ddot{u}(x, t) + 2\gamma \dot{u}(x, t) + \gamma^2 u(x, t) - \partial_x[S(\partial_x u(x, t))] &= f(x, t) \quad \text{in } (0, 1) \times (0, T], \\
u(0, t) &= u(1, t) = 0 \quad \text{for all } t \in (0, T], \\
u(x, 0) &= u_0(x), \quad \dot{u}(x, 0) = u_1(x).
\end{align}

We take \( S(\partial_x u) = \frac{1}{3}\partial_x u^3 \) and the time interval to be \( I = (0, T] \) with \( T = 1 \). Let \( u_0, u_1 \) and \( f \) be chosen such that the exact solution is

\[ u(x, t) = \sin\left(\sqrt{2\pi t}\right) \sin(\pi x). \]

That is, \( u_0(x) \equiv 0, u_1(x) = \sqrt{2\pi} \sin(\pi x), \) and

\[ f(x, t) = \left[(-2\pi^2 + \gamma^2) \sin\left(\sqrt{2\pi t}\right) + 2\sqrt{2\pi} \cos\left(\sqrt{2\pi t}\right)\right] \sin(\pi x) + \pi^4 \sin^3\left(\sqrt{2\pi t}\right) \cos^2(\pi x) \sin(\pi x). \]

We first discretize the problem in the spatial direction using continuous piecewise polynomials of degree \( p \geq 1 \). Let \( \mathcal{V}_h \) be the finite element function space with \( h \) being the spatial discretization parameter. The numerical approximation of the nonlinear wave-type equation following a Picard-type linearization in the nonlinear term is the following: find \( u_h \in \mathcal{V}_h \) such that

\[ \int_{\Omega} \dddot{u}_h \cdot v_h \, dx + \int_{\Omega} 2\gamma \ddot{u}_h \cdot v_h \, dx + \int_{\Omega} \gamma^2 u_h \cdot v_h \, dx + \frac{1}{3} \int_{\Omega} (\partial_x u_h^*)^2 \partial_x u_h \cdot \partial_x v_h \, dx = \int_{\Omega} f \cdot v_h \, dx, \]

for all \( v_h \in \mathcal{V}_h \). Here we assume that \( \partial_x u_h^* \) is known at each time step \( I_n \) either as an initial guess by using \( u_h \) over the previous time interval, or as a previous iterate in the Picard iteration. Now the problem results in the following second-order differential system for the nodal displacement \( \mathbf{U}(t) \):

\[ \begin{cases}
\dddot{\mathbf{U}}(t) + 2\gamma \ddot{\mathbf{U}}(t) + \gamma^2 \dot{\mathbf{U}}(t) + \frac{1}{3} \mathbf{K}(t) \mathbf{U}(t) = \mathbf{F}(t), & t \in (0, T], \\
\mathbf{U}(0) = \mathbf{U}_1, \quad \dot{\mathbf{U}}(0) = \mathbf{0},
\end{cases} \]

where \( \mathbf{U}_0 = [0, \ldots, 0]^T \in \mathbb{R}^d \) and \( \mathbf{U}_1 \) is the \( d \)-vector corresponding to \( u_1 \) at the grid points, \( \dddot{\mathbf{U}}(t) \) (respectively \( \ddot{\mathbf{U}}(t) \)) represents the vector of nodal acceleration (respectively velocity) and \( \mathbf{F}(t) \) is the vector of externally applied loads. \( \mathbf{M} \) is the mass matrix which is defined as

\[ \mathbf{M}_{ij} := \int_0^1 \psi_i(x) \psi_j(x) \, dx. \]

The time-dependent stiffness matrix \( \mathbf{K}(t) \) is defined as

\[ \mathbf{K}_{ij}(t) := \int_0^1 (\partial_x u_h^*(t))^2 \partial_x \psi_i(x) \partial_x \psi_j(x) \, dx, \]

where \( \{\psi_i\}_{i=1}^d \) are the basis functions in the spatial direction.

Multiplying the above algebraic formulation by \( \mathbf{M}^{-\frac{1}{2}} \) and setting \( \mathbf{Z}(t) = \mathbf{M}^{\frac{1}{2}} \mathbf{U}(t) \), we obtain

\[ \dddot{\mathbf{Z}}(t) + L \ddot{\mathbf{Z}}(t) + K(t) \mathbf{Z}(t) = \mathbf{G}(t), \quad t \in (0, T], \]

(86)
\[ \dot{Z}(0) = \tilde{M}^2 U_1, \quad Z(0) = \tilde{M}^2 U_0. \]  

Here
\[ L = 2\gamma \text{Id}, \quad K_0 = \gamma^2 \text{Id}, \]
\[ K(t) = \frac{1}{3} \tilde{M}^{-\frac{3}{2}} \tilde{K}(t) \tilde{M}^{-\frac{3}{2}}, \quad G(t) = \tilde{M}^{-\frac{3}{2}} F(t). \]

Note that both \( \tilde{K}(t) \) and \( K(t) \) are time-dependent. We subdivided \([0,T]\) into \( N \) subintervals \( I_n \), for \( n = 1, \ldots, N \), of uniform length \( k \). We assume that the polynomial degree in time is constant at each time step. That is, \( q_1 = \cdots = q_N \geq 2 \). If we consider the time integration on a generic time interval \( I_n \) for each \( n = 1, \ldots, N \), our discontinuous-in-time formulation reads as: find \( Z \in \mathcal{V}_{kh}^m \) such that
\[
\begin{aligned}
&\left( \dot{Z}(t), \dot{v} \right)_{L^2(I_n)} + \left( L \dot{Z}(t), \dot{v} \right)_{L^2(I_n)} + \left( K(t) Z(t), \dot{v} \right)_{L^2(I_n)} + \left( K_0 Z(t), \dot{v} \right)_{L^2(I_n)} \\
&\quad + \dot{Z}(t_{n-1}^+) \cdot \dot{v}(t_{n-1}^+) + K_0 Z(t_{n-1}^+) \cdot v(t_{n-1}^-) + K(t_{n-1}^+) Z(t_{n-1}^-) \cdot v(t_{n-1}^-) \\
&= \left( G(t), \dot{v} \right)_{L^2(I_n)} + \dot{Z}(t_{n-1}^-) \cdot \dot{v}(t_{n-1}^-) + K_0 Z(t_{n-1}^-) \cdot v(t_{n-1}^-) + K(t_{n-1}^-) Z(t_{n-1}^-) \cdot v(t_{n-1}^-),
\end{aligned}
\]
for all \( v \in \mathcal{V}_{kh}^m \), where on the right-hand side the values \( \dot{Z}(t_{n-1}^-) \) and \( Z(t_{n-1}^-) \) computed for \( I_{n-1} \) are used as initial conditions for the current time interval. For \( I_1 \), we set \( \dot{Z}(t_0^-) = \dot{Z}(0) \) and \( Z(t_0^-) = Z(0) \). Focusing on the generic time interval \( I_n \), we introduce the basis functions in the time direction \( \{ \phi_j(t) \}_{j=1}^{q_n+1} \) for the polynomial space \( \mathbb{P}^{q_n}(I_n) \) and define \( D = \hat{d}(q_n + 1) \), the dimension of the local finite element space \( \mathcal{V}_{kh}^m \). We also introduce the vectorial basis \( \{ \Phi_m^j(t) \}_{j=1}^{\hat{d}} \), where \( \Phi_m^j(t) \) is the \( \hat{d} \)-dimensional vector whose \( m \)-th component is \( \phi_j(t) \) and the other components are zero. We write
\[
Z(t) = \sum_{m=1}^{\hat{d}} \sum_{j=1}^{q_n+1} \alpha_m^j \Phi_m^j(t),
\]
where \( \alpha_m^j \in \mathbb{R} \) for \( m = 1, \ldots, \hat{d}, \ j = 1, \ldots, q_n + 1 \). By choosing \( v(t) = \Phi_m^j(t) \) for each \( m = 1, \ldots, \hat{d}, \ j = 1, \ldots, q_n + 1 \), we obtain the following algebraic system
\[
A Z = b,
\]
where \( z \in \mathbb{R}^D = \mathbb{R}^{(q_n+1)d} \) is the solution vector (whose entries are the values of \( \alpha_m^j \)); \( b \in \mathbb{R}^D \) corresponds to the right-hand side, which is given componentwise as
\[
b_m^j = \left( G(t), \Phi_m^j \right)_{L^2(I_n)} + \dot{Z}(t_{n-1}^-) \cdot \Phi_m^j(t_{n-1}^-) + K_0 Z(t_{n-1}^-) \cdot \Phi_m^j(t_{n-1}^-) + K(t_{n-1}^-) Z(t_{n-1}^-) \cdot \Phi_m^j(t_{n-1}^-),
\]
for \( m = 1, \ldots, \hat{d}, j = 1, \ldots, q_n + 1 \). \( A \) is the local stiffness matrix defined with its structure being discussed below. For \( l, j = 1, \ldots, q_n + 1 \),
\[
\begin{align*}
M_l^1 &= \left( \dot{\phi}_l^j, \dot{\phi}_l^j \right)_{L^2(I_n)}, & M_l^2 &= \left( \dot{\phi}_l^j, \dot{\phi}_l^j \right)_{L^2(I_n)}, & M_l^3 &= \left( \phi_l^j, \phi_l^j \right)_{L^2(I_n)}, \\
\tilde{M}_l^1 &= \left( K(t) \phi_l^j, \phi_l^j \right)_{L^2(I_n)}, & M_l^4 &= \dot{\phi}_l^j(t_{n-1}^-) \cdot \phi_l^j(t_{n-1}^-), \\
\tilde{M}_l^2 &= \dot{\phi}_l^j(t_{n-1}^-) \cdot \phi_l^j(t_{n-1}^-), & \tilde{M}_l^3 &= K(t_{n-1}^-) \phi_l^j(t_{n-1}^-) \cdot \phi_l^j(t_{n-1}^-).
\end{align*}
\]
Setting
\[
M = M^1 + M^4, \quad B_{ij} = L_{ij} M^2 + K_{0ij} (M^3 + M^5) + \left( \tilde{M}^3 + \tilde{M}^5 \right),
\]

A. SHAO
Algorithm 1. Iterative Algorithm (Multiple Picard iterations at each time interval).

Initialization: \( \partial_x u_0^n = \partial_x u_0 \) and
\[
\left[ K^0 \right]_{ij} = \int_0^1 (\partial_x u_0^0(x) \partial_x \psi_i(x)) \partial_x \psi_j(x) \, dx.
\]

Iteration: On each interval \( I_n = (t_{n-1}, t_n] \) for \( n = 1, 2, \ldots, N \), we solve
\[
\tilde{M} \tilde{U}(t) + 2\gamma \tilde{M} \tilde{U}(t) + \gamma^2 \tilde{M} \tilde{U}(t) + \frac{1}{3} \tilde{K}^n(t) \tilde{U}(t) = F(t)
\]
iteratively (using Picard iterations) by applying the discontinuous-in-time integration. Here
\[
\left[ K^n \right]_{ij} = \int_0^1 (\partial_x u_{DG}^{n-1}(t))^2 \partial_x \psi_i(x) \partial_x \psi_j(x) \, dx,
\]
where \( u_{DG}^{n-1}(t) \) is the solution obtained from the previous time interval \( I_{n-1} \).
\[
\left[ K^k \right]_{ij} = \int_0^1 (\partial_x u_{DG}^{n-k-1}(t))^2 \partial_x \psi_i(x) \partial_x \psi_j(x) \, dx,
\]
for \( k = 1, 2, \ldots \), where \( u_{DG}^{n-k-1}(t) \) is computed from the previous Picard iteration by using the stiffness matrix \( \left[ K^{k-1} \right]_{ij} \).

Update:
\[
\left[ \tilde{K}^{n+1} \right]_{ij} = \int_0^1 (\partial_x u_{DG}^n(t))^2 \partial_x \psi_i(x) \partial_x \psi_j(x) \, dx,
\]
where \( \partial_x u_{DG}^n(t) \) is computed using \( \left[ K_{k_{\text{end}}}^n \right] \). Here \( k_{\text{end}} \) is either the maximal (final) Picard iteration number or the iteration at which a certain tolerance is achieved.

Now move to the next time interval \( I_{n+1} \).
implies that we can formulate $\tilde{\phi}$ via $\frac{1}{\tilde{M}} = \frac{1}{M} + \frac{1}{2} \kappa \phi^2(t) + \frac{1}{3} \kappa \phi^3(t)$.

We use CG–$\Gamma_1$ and Table 1.

$$
\begin{array}{cccc}
 q & h & k = h^2 & L^2\text{-error} \\
 \hline
 2 & 2.50000e-1 & 6.25000e-2 & 1.2123e-2 \\
 & 2.00000e-1 & 4.00000e-2 & 4.9774e-3 \\
 & 1.25000e-1 & 1.56250e-2 & 1.1643e-3 \\
 & 6.25000e-2 & 3.90625e-3 & 1.2454e-4 \\
 3 & 2.50000e-1 & 6.25000e-2 & 4.1533e-4 \\
 & 2.00000e-1 & 4.00000e-2 & 1.8590e-4 \\
 & 1.25000e-1 & 1.56250e-2 & 2.5283e-5 \\
 & 6.25000e-2 & 3.90625e-3 & 1.5609e-6 \\
 4 & 2.50000e-1 & 6.25000e-2 & 2.5498e-5 \\
 & 2.00000e-1 & 4.00000e-2 & 8.3999e-6 \\
 & 1.25000e-1 & 1.56250e-2 & 9.7790e-7 \\
 & 6.25000e-2 & 3.90625e-3 & 3.1115e-8 \\
\end{array}
$$

Take $q = 2$ as an example; we note that on each time interval $I_n$ the solution $U(t)$ can be defined using the Legendre basis via $U(t)^n = \tilde{M}^{-\frac{1}{2}} Z^n(t) = \tilde{M}^{-\frac{1}{2}} (\alpha_1^n \phi_1(t) + \alpha_2^n \phi_2(t) + \alpha_3^n \phi_3(t))$.

Here $\alpha_i^n$ for $i = 1, 2$ and 3 are the coefficient vectors computed by extracting appropriate entries from $z$. This implies that we can formulate $\dot{K}^{n+1}(t)$ by considering

$$
\dot{K}^{n+1}(t) = \int_0^1 |\nabla U^n(t)|^2 \partial_x \psi_i \partial_x \psi_j dx
= \int_0^1 |\tilde{M}^{-\frac{1}{2}} \nabla \phi_1(t) + \tilde{M}^{-\frac{1}{2}} \nabla \phi_2(t) + \tilde{M}^{-\frac{1}{2}} \nabla \phi_3(t)|^2 \partial_x \psi_i \partial_x \psi_j dx
= \int_0^1 |\tilde{M}^{-\frac{1}{2}} \nabla \phi_1(t)|^2 \partial_x \psi_i \partial_x \psi_j dx + \int_0^1 |\tilde{M}^{-\frac{1}{2}} \nabla \phi_2(t)|^2 \partial_x \psi_i \partial_x \psi_j dx
+ \int_0^1 |\tilde{M}^{-\frac{1}{2}} \nabla \phi_3(t)|^2 \partial_x \psi_i \partial_x \psi_j dx
+ \int_0^1 2(\tilde{M}^{-\frac{1}{2}} \nabla \phi_1(t) \cdot \tilde{M}^{-\frac{1}{2}} \nabla \phi_2(t)) \partial_x \psi_i \partial_x \psi_j dx
+ \int_0^1 2(\tilde{M}^{-\frac{1}{2}} \nabla \phi_1(t) \cdot \tilde{M}^{-\frac{1}{2}} \nabla \phi_3(t)) \partial_x \psi_i \partial_x \psi_j dx
+ \int_0^1 2(\tilde{M}^{-\frac{1}{2}} \nabla \phi_2(t) \cdot \tilde{M}^{-\frac{1}{2}} \nabla \phi_3(t)) \partial_x \psi_i \partial_x \psi_j dx
:= K_{11}^{n+1}(t) + K_{22}^{n+1}(t) + K_{33}^{n+1}(t) + K_{12}^{n+1}(t) + K_{13}^{n+1}(t) + K_{23}^{n+1}(t).
$$

We use CG–$p$ elements where $p = q$ in space with $k = h^2$, $T = 1$ and $\gamma = 1$, and compute the errors $\|u(T) - u_{DG}(t^-_N)\|_{L^2} + \|\dot{u}(T) - \dot{u}_{DG}(t^-_N)\|_{L^2}$ versus $k$ for $h^2 = 6.25000e-2, 4.00000e-2, 1.56250e-2$ and 3.90625e-3 with respect to polynomial degrees 2, 3, 4 in Table 1. Note that here we use $h = 2.50000e-1, 2.00000e-2, 1.25000e-1$ and 6.25000e-2 instead of the conventional halving procedure; this is to avoid the accumulation of any unnecessary floating point errors resulting from a large number of time steps while still having sufficient data to compute the convergence rates. The computed errors are shown in Figure 1 in a log–log scale. As expected, the error decreases as we increase the polynomial degree $q$ or decrease the time step $k$. By
Remark 4.2, we expect convergence rates of order 1.5, 2.0 and 2.5 for \( q = p = 2, 3 \) and 4 respectively, which are consistent with the numerical results shown in Table 1.

**Appendix A. Proof of the auxiliary lemma**

**Lemma A.1.** Under the assumptions stated in Theorem 4.1, there exists a constant \( C_\tau > 0 \) such that, for \( t \in I_n, n = 1, 2, \ldots, N \),

\[
| A_i (\phi(t)e^{\gamma t}, \varphi(t), \psi(t)) | \leq C_\tau \| \nabla \varphi(t) \|_{L^2} \| \nabla \psi(t) \|_{L^2}. \tag{A.1}
\]

**Proof.** Note that for \( t \in I_n, n = 1, 2, \ldots, N \),

\[
\partial_t [ A_{i\alpha j\beta} (\nabla \Phi(t)e^{\gamma t} + \tau (\nabla \Phi(t) - \nabla W(t))e^{\gamma t}) ]
= \sum_{k,m=1}^{d} \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{km}} (\nabla \Phi(t)e^{\gamma t} + \tau (\nabla \Phi(t) - \nabla W(t))e^{\gamma t}) \partial_m (\partial_t (W_k(t)e^{\gamma t}) + \tau \partial_t ((\phi_k(t) - W_k(t))e^{\gamma t})).
\]

Since the values of the function \( \nabla \Phi(t)e^{\gamma t} + \tau (\nabla \Phi(t) - \nabla W(t))e^{\gamma t} \) for \( t \in [0, T], \tau \in (0, 1) \), belong to the compact convex subset \( \mathcal{M}_\delta \) of \( \mathbb{R}^{d^2} \), and because \( A_{i\alpha j\beta} \) is sufficiently smooth (and in particular continuously differentiable on \( \mathcal{M}_\delta \)), we have

\[
\left| \sum_{i,\alpha, j,\beta=1}^{d} \frac{1}{2} \int_0^1 \left( \partial_t [ A_{i\alpha j\beta} (\nabla \Phi(t)e^{\gamma t} + \tau (\nabla \Phi(t) - \nabla W(t))e^{\gamma t}) ] \partial_j \varphi(t), \partial_i \psi(t) \right)_{L^2} d\tau \right|
\leq c \left( \| \nabla \Phi(t) \|_{L^\infty} + \| \nabla W(t) \|_{L^\infty} \right) \| \nabla \varphi(t) \|_{L^2} \| \nabla \psi(t) \|_{L^2}
+ c \left( \| \nabla (\dot{\Phi}(t) - \dot{W}(t)) \|_{L^\infty} + \| \nabla (\dot{\Phi}(t) - \dot{W}(t)) \|_{L^\infty} \right) \| \nabla \varphi(t) \|_{L^2} \| \nabla \psi(t) \|_{L^2}
\leq \tilde{C}h^{-\frac{\gamma}{2}} \left( \| \nabla (\dot{\Phi}(t) - \dot{W}(t)) \|_{L^2} + \| \nabla (\dot{\Phi}(t) - \dot{W}(t)) \|_{L^2} \right) \| \nabla \varphi(t) \|_{L^2} \| \nabla \psi(t) \|_{L^2}
\]
\[ + \tilde{C} \| \nabla \varphi(t) \|_{L^2} \| \nabla \psi(t) \|_{L^2}, \]

where we have applied the inverse inequality (ii,b) and property (iii,c) of the nonlinear projection \( W \). We shall bound \( \| \nabla (\phi(t) - \tilde{W}(t)) \|_{L^2} \) and \( \| \nabla (\phi(t) - W(t)) \|_{L^2} \) for \( t \in I_n, n = 1, 2, \ldots, N \). Applying the triangle inequality, we have

\[ \| \nabla (\phi(t) - W(t)) \|_{L^2} \leq \| \nabla (\phi(t) - \Pi_k W(t)) \|_{L^2} + \| \nabla (\Pi_k W(t) - W(t)) \|_{L^2}. \]

Note that for \( t \in I_n, n = 1, 2, \ldots, N \),

\[ \| \nabla (\phi(t) - \Pi_k W(t)) \|_{L^2} \leq \| \nabla (\phi(t_n) - \Pi_k W(t_n)) \|_{L^2} + \int_{t_n}^{t_{n+1}} \| \partial_s (\nabla \phi(s) - \nabla \Pi_k W(s)) \|_{L^2} ds \]

\[ \leq \| \nabla (\phi(t_n) - \Pi_k W(t_n)) \|_{L^2} + \int_{t_n}^{t_{n+1}} \| \partial_t (\nabla \phi(t) - \nabla \Pi_k W(t)) \|_{L^2} dt \]

\[ \leq \| \nabla (\phi(t_n) - \Pi_k W(t_n)) \|_{L^2} + C_0 h^{-1} \sqrt{k_n} \left( \int_{t_n}^{t_{n+1}} \| \partial_t (\phi(t) - \Pi_k W(t)) \|_{L^2}^2 dt \right)^{\frac{1}{2}} \]

\[ \leq C(v) \left( \sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_i^{2q+1}}{\gamma_i^{2(s-1)}} \right)^{\frac{1}{2}}, \]

where we have used the inverse inequality (ii,a), Hölder’s inequality, the fact that \( \phi \in \mathcal{F} \) and the assumption that \( \mu, k_i \leq h^2 \) for each \( i = 1, \ldots, N \). Here \( C(v) \) denotes a constant depending on \( v \), which may vary throughout this proof. On the other hand,

\[ \| \nabla (\Pi_k W(t) - W(t)) \|_{L^2} \leq \| \nabla (\Pi_k W(t_n) - W(t_n)) \|_{L^2} + \int_{t_n}^{t_{n+1}} \| \partial_t (\nabla \Pi_k W(t) - \nabla W(t)) \|_{L^2} ds \]

\[ \leq \int_{t_n}^{t_{n+1}} \| \partial_t (\nabla \Pi_k W(t) - \nabla W(t)) \|_{L^2} dt \quad \text{(since } \Pi_k W(t_n) = W(t_n)) \]

\[ \leq \sqrt{k_n} \left( \int_{t_n}^{t_{n+1}} \| \partial_t (\nabla \Pi_k W(t) - \nabla W(t)) \|_{L^2}^2 dt \right)^{\frac{1}{2}} \]

\[ \leq C_{\frac{k_n^{q+n+\frac{1}{2}}}{q_n^{-1}}} \| W \|_{H^s(I_n); H^1)}, \quad \text{for } W \in H^s([0,T]; H^1), \]

where we have used inequality (28) with the \( L^2 \) norm in space replaced by the \( H^1 \) semi-norm. Thus,

\[ \max_{t \in I_n, 1 \leq n \leq N} \| \nabla (\phi(t) - W(t)) \|_{L^2} \leq C(v) \left( \sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_i^{2q+1}}{\gamma_i^{2(s-1)}} \right)^{\frac{1}{2}} + C_{\frac{k_n^{q+n+\frac{1}{2}}}{q_n^{-1}}} \| W \|_{H^s(I_n); H^1)}. \quad (A.2) \]

Applying the triangle inequality to the time derivative term, we have

\[ \| \nabla (\phi(t) - \tilde{W}(t)) \|_{L^2} \leq \| \partial_t (\nabla \phi(t) - \nabla \Pi_k W(t)) \|_{L^2} + \| \partial_t (\nabla \Pi_k W(t) - \nabla W(t)) \|_{L^2}. \]

Note that for \( t \in I_n, n = 1, 2, \ldots, N \),

\[ \| \partial_t (\nabla \phi(t) - \nabla \Pi_k W(t)) \|_{L^2} \leq C_0 h^{-1} \| \partial_t (\phi(t) - \Pi_k W(t)) \|_{L^2} \]

\[ \leq C_0 h^{-1} \| \partial_t (\phi(t_n) - \Pi_k W(t_n)) \|_{L^2} + C_0 h^{-1} \int_{t_n}^{t_{n+1}} \| \partial_s (\phi(s) - \Pi_k W(s)) \|_{L^2} ds \]
Combining (A.2) and (A.3), we have

\[
\nu_t (\nabla \Phi(t) - \Pi_k W(t)) \big|_{L^2} \leq C_0 h^{-1} \left\| \partial_t (\Phi(t_n) - \Pi_k W(t_n)) \right\|_{L^2} + C_0 h^{-1} \int_{t_{n-1}}^{t_n} \left\| \partial_{tt} (\Phi(t) - \Pi_k W(t)) \right\|_{L^2} dt.
\]

Since \( \Phi \in \mathcal{F} \), we have

\[
\left\| \partial_t (\Phi(t_n) - \Pi_k W(t_n)) \right\|_{L^2} \leq C_* (v) \left( \sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_i^{2q+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}}.
\]

Using the inverse inequality in time (30), we obtain

\[
\int_{t_{n-1}}^{t_n} \left\| \partial_{tt} (\Phi(t) - \Pi_k W(t)) \right\|_{L^2} dt \leq \sqrt{k_n} \left( \int_{t_{n-1}}^{t_n} \left\| \partial_{tt} (\Phi(t) - \Pi_k W(t)) \right\|_{L^2}^2 dt \right)^{\frac{1}{2}}
\]

\[
\leq C_2 \frac{1}{\sqrt{k_n}} \left( \int_{t_{n-1}}^{t_n} \left\| \partial_t (\Phi(t) - \Pi_k W(t)) \right\|_{L^2}^2 dt \right)^{\frac{1}{2}}
\]

\[
\leq C (v) \frac{1}{\sqrt{k_n}} \left( \sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_i^{2q+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}}.
\]

On the other hand, for \( t \in I_n, n = 1, 2, \ldots, N \), we have

\[
\left\| \partial_t (\nabla \Pi_k W(t) - \nabla W(t)) \right\|_{L^2} \leq \left\| \partial_t (\nabla \Pi_k W(t_n) - \nabla W(t_n)) \right\|_{L^2} + \int_{t}^{t_n} \left\| \partial_{tt} (\nabla \Pi_k W(s) - \nabla W(s)) \right\|_{L^2} ds
\]

\[
\leq \int_{t_{n-1}}^{t_n} \left\| \partial_{tt} (\nabla \Pi_k W(t) - \nabla W(t)) \right\|_{L^2} dt \quad \text{(since } \Pi_k W(t_n) = W(t_n) \text{)}
\]

\[
\leq \sqrt{k_n} \left( \int_{t_{n-1}}^{t_n} \left\| \partial_{tt} (\nabla \Pi_k W(t) - \nabla W(t)) \right\|_{L^2}^2 dt \right)^{\frac{1}{2}}
\]

\[
\leq C \frac{k_n^{2q+1}}{q_n^{a-3}} \| W \|_{H^r ([0,T]; H^1)} \quad \text{for } W \in H^r ([0,T]; H^1).
\]

Thus

\[
\max_{t \in I_n, 1 \leq n \leq N} \left\| \nabla \Phi (t) - \nabla W(t) \right\|_{L^2} \leq C(v) \left( h^{-1} + h^{-1} k_n^{-\frac{1}{2}} \right) \left( \sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_i^{2q+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}}
\]

\[
+ C \frac{k_n^{2q+1}}{q_n^{a-3}} \| W \|_{H^r ([0,T]; H^1)}. \quad \text{(A.3)}
\]

Combining (A.2) and (A.3), we have

\[
\tilde{C} h^{-\frac{d}{2}} \max_{t \in I_n, 1 \leq n \leq N} \left( \| \nabla \Phi (t) - \nabla W(t) \|_{L^2} + \left\| \nabla \Phi (t) - \nabla W(t) \right\|_{L^2} \right)
\]

\[
\leq \tilde{C} h^{-\frac{d}{2}} C(v) \left( 1 + h^{-1} + h^{-1} k_n^{-\frac{1}{2}} \right) \left( \sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_i^{2q+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}}
\]

\[
+ \tilde{C} h^{-\frac{d}{2}} C \left( \frac{k_n^{2q+1}}{q_n^{a-1}} + \frac{k_n^{2q-1}}{q_n^{a-3}} \right) \| W \|_{H^r ([0,T]; H^1)}. \quad \text{(A.4)}
\]
Since \( r > \frac{d}{2} + 1 \), \( k_i^{\theta - \frac{d}{2}} = o\left(h^{1+\frac{d}{2}}\right) \) and \( \mu k_i \leq h^2 \leq \nu_k \) for each \( i = 1, 2, \ldots, N \), we can choose \( h_0 > 0 \) such that for \( h \leq h_0 \), the right-hand side of \((A.4)\) is bounded by 1. Thus, equation \((41)\) follows by taking \( C_\tau = \tilde{C} + 1 \). The constant \( C_\tau \) defined in this way does not depend on \( C_\circ(v) \).

\[ \square \]

**APPENDIX B. APPROXIMATION PROPERTIES OF THE ELLIPTIC PROJECTION**

Here we derive the properties \((\text{iii.a})–(\text{iii.c})\) of the nonlinear projection \( W \). We write \( a \lesssim b \) if there exists a universal constant \( C > 0 \) independent of the spatial discretization parameter \( h \) such that \( a \leq C b \).

**B.1. \( L^2 \) bound on \( \nabla(v - W) \) and \( L^\infty \) bound on \( \nabla W \)**

Recall that for each \( t \in [0, T] \), \( a(W(t), \varphi) = a(v(t), \varphi) \) for all \( \varphi \in \mathcal{V}_h \). Let \( \mathcal{P}_h : L^2 \to \mathcal{V}_h \) denote the standard \( L^2 \)-projection operator in the spatial direction. Then we have

\[
a(W(t), \varphi) - a(\mathcal{P}_h v(t), \varphi) = a(v(t), \varphi) - a(\mathcal{P}_h v(t), \varphi) \quad \text{for all} \quad \varphi \in \mathcal{V}_h.
\]

That is,

\[
\int_0^1 \tilde{a}(\mathcal{P}_h v(t)e^{\gamma t} + \tau(W(t) - \mathcal{P}_h v(t))e^{\gamma t}; W(t) - \mathcal{P}_h v(t), \varphi) \, d\tau = 0.
\]

Define the following subset of \( H^1_0 \),

\[
\mathcal{F} = \left\{ \phi \in \mathcal{V}_h : \|\nabla(\phi - \mathcal{P}_h v)\|_{L^2} \leq C_* h^r \|v\|_{H^{r+1}} \quad \text{for} \quad \frac{d}{2} < r \leq \min(p, m - 1) \right\}
\]

where \( C_* \) is a constant independent of \( h \). The set \( \mathcal{F} \) is non-empty since for each fixed \( t \in [0, T] \), \( \mathcal{P}_h v(t) \in \mathcal{F} \). Furthermore, \( \mathcal{F} \) is a closed and convex subset of \( H^1_0 \). We define the fixed point mapping \( \mathcal{N} \) on \( \mathcal{F} \) as follows. Given \( \phi \in \mathcal{F} \), we denote by \( W_\phi \in \mathcal{V}_h \), the solution to the following linear variational problem: find \( W_\phi \in \mathcal{V}_h \) such that

\[
\int_0^1 \tilde{a}(\mathcal{P}_h v(t)e^{\gamma t} + \tau(\phi(t) - \mathcal{P}_h v(t))e^{\gamma t}; W_\phi(t) - \mathcal{P}_h v(t), \varphi) \, d\tau = 0.
\]

Since \( \mathcal{V}_h \) is a finite dimensional linear space, the existence and uniqueness of \( W_\phi(t) \in \mathcal{V}_h \) for each \( t \in [0, T] \) follows if we can show that \( \int_0^1 \tilde{a}(\mathcal{P}_h v(t)e^{\gamma t} + \tau(\phi(t) - \mathcal{P}_h v(t))e^{\gamma t}; W_\phi(t) - \mathcal{P}_h v(t), \varphi) \, d\tau \) is coercive on \( \mathcal{V}_h \times \mathcal{V}_h \) in the \( \| \cdot \|_{H^1} \) semi-norm. This is indeed true in view of the assumption \((S2b)\). For each \( t \in [0, T] \), if we take \( W(t) = W_\phi(t) \), we have

\[
\|\nabla(W(t) - \mathcal{P}_h v(t))\|_{L^2} \leq C_* h^r \|v\|_{H^{r+1}}, \quad \frac{d}{2} < r \leq \min(p, m - 1).
\]

By the approximation properties of \( \mathcal{P}_h \) in the \( \| \cdot \|_{H^1} \) semi-norm, we have

\[
\|\nabla(v(t) - \mathcal{P}_h v(t))\|_{L^2} \leq h^r \|v(t)\|_{H^{r+1}}, \quad \frac{d}{2} < r \leq \min(p, m - 1).
\]

It follows from the triangle inequality that

\[
\|\nabla(W(t) - v(t))\|_{L^2} \leq h^r \|v(t)\|_{H^{r+1}}, \quad \frac{d}{2} < r \leq \min(p, m - 1).
\]
By the approximation properties of $\mathcal{P}_h$ in the $|\cdot|_{W^{1, \infty}}$ semi-norm, we have
\[
\|\nabla (v(t) - \mathcal{P}_h v(t))\|_{L^\infty} \leq h^{r - \frac{d}{2}} \|v(t)\|_{H^{r+1}}, \quad \frac{d}{2} < r \leq \min(p, m - 1).
\]  
(B.9)
Combining (B.8) and (B.9), we obtain
\[
\|\nabla W(t)\|_{L^\infty} \leq \|\nabla v(t)\|_{L^\infty} + \|\nabla (W(t) - v(t))\|_{L^\infty}
\leq \|\nabla v(t)\|_{L^\infty} + \|\nabla (W(t) - \mathcal{P}_h v(t))\|_{L^\infty} + \|\nabla (\mathcal{P}_h v(t) - v(t))\|_{L^\infty}
\leq \|\nabla v(t)\|_{L^\infty} + C_1 h^{\frac{d}{2}} \|\nabla (W(t) - \mathcal{P}_h v(t))\|_{L^2} + C(v) h^{r - \frac{d}{2}}
\leq c_0,
\]
for some constant $c_0$. The last inequality follows from the boundedness of $\nabla v$ and the fact that $r > \frac{d}{2}$, while the second last line follows from (ii,b) and (B.9).

B.2. $L^2$ bound on $\nabla \left( \dot{W} - \dot{v} \right)$ and $L^\infty$ bound on $\nabla \dot{W}$

For the estimate of the $L^2$ bound on $\nabla \left( \dot{W} - \dot{v} \right)$, we follow the proof from Section 6 in [27]. We need to show that $t \mapsto W(t)$ is differentiable with respect to $t$. For $U \in \mathcal{V}_h$ and $t \in [0, T]$, we notice that the mapping $\varphi \mapsto a(U, \varphi) - a(v(t), \varphi)$ is a bounded linear functional on $\mathcal{V}_h$; hence by Riesz representation theorem, there exists a unique $\mathcal{A}(t, U) \in \mathcal{V}_h$ such that
\[
(\mathcal{A}(t, U), \varphi) = a(U, \varphi) - a(v(t), \varphi).
\]
It follows from the linearization process that the derivative of the nonlinear mapping $(t, U) \mapsto \mathcal{A}(t, U)$ with respect to $U$, evaluated at $U = W(t)$, exists and is invertible for any $t \in [0, T]$. We also have $\mathcal{A}(t, W(t)) = 0$. Since $v(t)$ is differentiable with respect to $t$, it follows that $\mathcal{A}(t, U)$ is differentiable in a neighbourhood of $(t_0, W(t_0))$ for any $t_0 \in (0, T)$. We then deduce from the implicit function theorem that $t \mapsto W(t)$ is differentiable in $(0, T)$. Next, we derive the error bound of $\|\nabla \left( \dot{W}(t) - \dot{v}(t) \right)\|_{L^2}$. By definition of $W(t)$, we have
\[
\int_0^1 \tilde{a}(\mathcal{P}_h v(t)e^{\gamma t} + \tau(W(t) - \mathcal{P}_h v(t))e^{\gamma t}; W(t) - \mathcal{P}_h v(t), \varphi) \, d\tau
= \int_0^1 \tilde{a}(\mathcal{P}_h v(t)e^{\gamma t} + \tau(v(t) - \mathcal{P}_h v(t))e^{\gamma t}; v(t) - \mathcal{P}_h v(t), \varphi) \, d\tau \quad \text{for all } \varphi \in \mathcal{V}_h.
\]
After differentiation with respect to $t$, we have
\[
\int_0^1 \tilde{a}(\mathcal{P}_h v(t)e^{\gamma t} + \tau(W(t) - \mathcal{P}_h v(t))e^{\gamma t}; \dot{W}(t) - \mathcal{P}_h \dot{v}(t), \varphi) \, d\tau
+ \int_0^1 \int_{\Omega} \sum_{i, a, j, k, m=1}^d \left( \frac{\partial A_{iaj\beta}}{\partial \eta_{km}}(\nabla \mathcal{P}_h v(t)e^{\gamma t} + \tau(\nabla W(t) - \nabla \mathcal{P}_h v(t))e^{\gamma t}) \right) \left( \partial_m(\delta_t[\mathcal{P}_h v_k(t)e^{\gamma t}] + \tau \partial_t[(W_k(t) - \mathcal{P}_h v_k(t))e^{\gamma t}]) \right) \partial_j(W - \mathcal{P}_h v)_{\beta}, \partial_j \varphi_{\alpha} \right) \, dx \, d\tau
= \int_0^1 \tilde{a}(\mathcal{P}_h v(t)e^{\gamma t} + \tau(v(t) - \mathcal{P}_h v(t))e^{\gamma t}; \dot{v}(t) - \mathcal{P}_h \dot{v}(t), \varphi) \, d\tau
+ \int_0^1 \int_{\Omega} \sum_{i, a, j, k, m=1}^d \left( \frac{\partial A_{iaj\beta}}{\partial \eta_{km}}(\nabla \mathcal{P}_h v(t)e^{\gamma t} + \tau(\nabla v(t) - \nabla \mathcal{P}_h v(t))e^{\gamma t}) \right) \left( \partial_m(\delta_t[v_k(t)e^{\gamma t}] + \tau \partial_t[v_k(t)e^{\gamma t}]) \right) \partial_j(v - \mathcal{P}_h v)_{\beta}, \partial_j \varphi_{\alpha} \right) \, dx \, d\tau.
\]
\[
\times \partial_m \left( \partial_t \left[ \mathcal{P}_h \mathbf{v}_k(t)e^{\gamma t} \right] + \tau \partial_t \left[ (\mathbf{v}_k(t) - \mathcal{P}_h \mathbf{v}_k(t))e^{\gamma t} \right] \right) \left\{ \partial_j (\mathbf{v} - \mathcal{P}_h \mathbf{v})_\beta, \partial_j \varphi_\alpha \right\}_{L^2} \, dx \, d\tau,
\]

for all \( \varphi \in \mathcal{V}_h \). Rearranging gives
\[
\int_0^1 \tilde{a} \left( \mathcal{P}_h \mathbf{v}(t)e^{\gamma t} + \tau (\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t))e^{\gamma t}; \dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t), \varphi \right) \, d\tau
\]
\[
= \int_0^1 \tilde{a} (\mathcal{P}_h \mathbf{v}(t)e^{\gamma t} + \tau (\mathbf{v}(t) - \mathcal{P}_h \mathbf{v}(t))e^{\gamma t}; \dot{\mathbf{v}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t), \varphi) \, d\tau
\]
\[
+ \int_0^1 \int_{\Omega} \sum_{i, \alpha, j, \beta, k, m=1}^d \left( \left\{ \frac{\partial A_{iaj\beta}}{\partial \eta_{k,m}} (\nabla \mathcal{P}_h \mathbf{v}(t)e^{\gamma t} + \tau (\nabla \mathbf{v}(t) - \nabla \mathcal{P}_h \mathbf{v}(t))e^{\gamma t}) \times \partial_m (\partial_t \left[ \mathcal{P}_h \mathbf{v}_k(t)e^{\gamma t} \right] + \tau \partial_t \left[ (\mathbf{v}_k(t) - \mathcal{P}_h \mathbf{v}_k(t))e^{\gamma t} \right] \right\} \partial_j (\mathbf{v} - \mathcal{P}_h \mathbf{v})_\beta, \partial_j \varphi_\alpha \right\}_{L^2} \, dx \, d\tau
\]
\[
- \int_0^1 \int_{\Omega} \sum_{i, \alpha, j, \beta, k, m=1}^d \left( \left\{ \frac{\partial A_{iaj\beta}}{\partial \eta_{k,m}} (\nabla \mathcal{P}_h \mathbf{v}(t)e^{\gamma t} + \tau (\nabla \mathbf{W}(t) - \nabla \mathcal{P}_h \mathbf{v}(t))e^{\gamma t}) \times \partial_m (\partial_t \left[ \mathcal{P}_h \mathbf{v}_k(t)e^{\gamma t} \right] + \tau \partial_t \left[ (\mathbf{W}_k(t) - \mathcal{P}_h \mathbf{v}_k(t))e^{\gamma t} \right] \right\} \partial_j (\mathbf{W} - \mathcal{P}_h \mathbf{v})_\beta, \partial_j \varphi_\alpha \right\}_{L^2} \, dx \, d\tau
\]
\[
:= T_1 + T_2 + T_3.
\]

Taking \( \varphi = \dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t) \), we have
\[
T_1 \lesssim h^r \| \dot{\mathbf{v}}(t) \|_{H^{r+1}} \left\| \nabla \left( \dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t) \right) \right\|_{L^2}
\]
\[
T_2 \lesssim \left( \| \nabla \mathcal{P}_h \dot{\mathbf{v}}(t) \|_{L^\infty} + \| \nabla (\dot{\mathbf{v}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \|_{L^\infty} + \| \nabla \mathcal{P}_h \mathbf{v}(t) \|_{L^\infty} + \| \nabla (\mathbf{v}(t) - \mathcal{P}_h \mathbf{v}(t)) \|_{L^\infty} \right)
\]
\[
\times \| \nabla \mathbf{v}(t) - \nabla \mathcal{P}_h \mathbf{v}(t) \|_{L^2} \left\| \nabla \left( \dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t) \right) \right\|_{L^2}
\]
\[
\lesssim h^r \| \mathbf{v}(t) \|_{H^{r+1}} \left( \| \nabla \mathcal{P}_h \dot{\mathbf{v}}(t) \|_{L^\infty} + \| \nabla (\dot{\mathbf{v}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \|_{L^\infty} + \| \nabla \mathcal{P}_h \mathbf{v}(t) \|_{L^\infty} + \| \nabla (\mathbf{v}(t) - \mathcal{P}_h \mathbf{v}(t)) \|_{L^\infty} \right)
\]
\[
\times \| \nabla (\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t)) \|_{L^2} \left\| \nabla (\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \right\|_{L^2}
\]
\[
\lesssim h^r \| \mathbf{v}(t) \|_{H^{r+1}} \left( \| \nabla \dot{\mathbf{v}}(t) \|_{L^\infty} + \| \nabla (\dot{\mathbf{v}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \|_{L^\infty} + \| \nabla (\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t)) \|_{L^\infty} \right)
\]
\[
\times \left\| \nabla \left( \dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t) \right) \right\|_{L^2} + h^{r-\frac{3}{2}} \| \mathbf{v}(t) \|_{H^{r+1}} \left\| \nabla (\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t)) \right\|_{L^2} \| \nabla (\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \|_{L^2}
\]
\[
\lesssim h^r \| \mathbf{v}(t) \|_{H^{r+1}} \left\| \nabla (\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \right\|_{L^2}
\]
\[
+ h^{r-\frac{3}{2}} \| \mathbf{v}(t) \|_{H^{r+1}} \left\| \nabla (\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t)) \right\|_{L^2} \left\| \nabla (\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \right\|_{L^2}
\]
\[
\lesssim h^r \| \mathbf{v}(t) \|_{H^{r+1}} \left\| \nabla (\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \right\|_{L^2}
\]
\[
+ h^{r-\frac{3}{2}} \| \mathbf{v}(t) \|_{H^{r+1}} \left\| \nabla (\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t)) \right\|_{L^2} \left\| \nabla (\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \right\|_{L^2}
\]
\[
\lesssim h^r \| \mathbf{v}(t) \|_{H^{r+1}} \left\| \nabla (\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \right\|_{L^2}
\]
\[
+ h^{r-\frac{3}{2}} \| \mathbf{v}(t) \|_{H^{r+1}} \left\| \nabla (\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t)) \right\|_{L^2} \left\| \nabla (\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \right\|_{L^2}
\]
Applying the strong ellipticity condition (S2b) on the left-hand side of (B.10), we have

\[
\begin{aligned}
+ h^{r-\frac{d}{2}} \|v(t)\|_{H^{r+1}} \|\nabla \left( \hat{W}(t) - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2}^2,
\end{aligned}
\]

Combining the estimates for \( T_1, T_2 \) and \( T_3 \), we have

\[
\int_0^1 a \left( \mathcal{P}_h v(t) e^{\tau t} + \tau (W(t) - \mathcal{P}_h v(t)) e^{\tau t}; \hat{W}(t) - \mathcal{P}_h \hat{v}(t), \varphi \right) \, d\tau
\leq h^r (\|v(t)\|_{H^{r+1}} + \|\hat{v}\|_{H^{r+1}}) \|\nabla \left( \hat{W} - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2}^2
\]

\[
+ h^{r-\frac{d}{2}} \|v\|_{H^{r+1}} \|\nabla (W(t) - \mathcal{P}_h v(t))\|_{L^2} \|\nabla \left( \hat{W}(t) - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2}^2
\]

\[
+ h^{r-\frac{d}{2}} \|v\|_{H^{r+1}} \|\nabla \left( \hat{W} - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2}^2.
\]

Applying the strong ellipticity condition (S2b) on the left-hand side of (B.10), we have

\[
M_1 \|\nabla \left( \hat{W}(t) - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2}^2 \lesssim h^r (\|v(t)\|_{H^{r+1}} + \|\hat{v}\|_{H^{r+1}}) \|\nabla \left( \hat{W} - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2}^2
\]

\[
+ h^{r-\frac{d}{2}} \|v\|_{H^{r+1}} \|\nabla (W(t) - \mathcal{P}_h v(t))\|_{L^2} \|\nabla \left( \hat{W}(t) - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2}^2
\]

\[
+ h^{r-\frac{d}{2}} \|v\|_{H^{r+1}} \|\nabla \left( \hat{W} - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2}^2.
\]

Dividing by \( \|\nabla \left( \hat{W} - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2} \) on both sides yields

\[
M_1 \|\nabla \left( \hat{W}(t) - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2} \lesssim h^r (\|v(t)\|_{H^{r+1}} + \|\hat{v}\|_{H^{r+1}}) + h^{r-\frac{d}{2}} \|v\|_{H^{r+1}} \|\nabla (W(t) - \mathcal{P}_h v(t))\|_{L^2}
\]

\[
+ h^{r-\frac{d}{2}} \|v\|_{H^{r+1}} \|\nabla \left( \hat{W}(t) - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2}.
\]

Since \( r > \frac{d}{2} \), we can choose \( h \) sufficiently small such that the last term on the right-hand side can be absorbed into the term on the left-hand side. This yields

\[
\|\nabla \left( \hat{W}(t) - \mathcal{P}_h \hat{v}(t) \right) \|_{L^2} \lesssim h^r (\|v(t)\|_{H^{r+1}} + \|\hat{v}(t)\|_{H^{r+1}}).
\]

(B.11)

Again, by the approximation property of \( \mathcal{P}_h \), we have, for each \( t \in [0, T] \),

\[
\|\nabla \hat{v}(t) - \mathcal{P}_h \hat{v}(t)\|_{L^2} \lesssim h^r \|\hat{v}(t)\|_{H^{r+1}}, \quad \frac{d}{2} < r \leq \min(p, m - 1).
\]

(B.12)

It follows from the triangle inequality that, for each \( t \in [0, T] \),

\[
\|\nabla \hat{W}(t) - \nabla \hat{v}(t)\|_{L^2} \lesssim h^r (\|v(t)\|_{H^{r+1}} + \|\hat{v}(t)\|_{H^{r+1}}), \quad \frac{d}{2} < r \leq \min(p, m - 1).
\]

(B.13)

By a similar argument as in the previous section, we can show that there exists a constant \( c_1 > 0 \) such that

\[
\|\nabla \hat{W}(t)\|_{L^\infty} \leq c_1.
\]

(B.14)

**B.3.** \( L^2 \) bounds on \( (v - W) \), \( (\hat{v} - \hat{W}) \) and \( (\bar{v} - \bar{W}) \)

It was proved by Dobrowolski and Rannacher in [33] that for each \( t \in [0, T] \),

\[
\|v(t) - W(t)\|_{L^2} \leq C_r(v) h^{r+1}, \quad \frac{d}{2} < r \leq \min(p, m - 1).
\]

(B.15)
We shall focus on proving the $L^2$ error bound of the time derivative using a duality argument in this section. Consider the following boundary value problem: for a given $g \in L^2$, solve $\psi \in H^1_0$ such that
\[
\tilde{a}(v; \psi; \phi) = (g, \phi)_{L^2} \quad \text{for all } \phi \in H^1_0,
\]
(B.16)
where $v$ is the solution of (5)-(7) and
\[
\tilde{a}(v; \psi; \phi) = \sum_{i,\alpha,j,\beta=1}^d (A_{i\alpha j\beta}(\nabla v) \partial_\beta \psi, \partial_\alpha \phi_i)_{L^2}.
\]
(B.17)

Since $A_{i\alpha j\beta}(\nabla v) \in W^{1,\infty}$ provided that $A_{i\alpha j\beta}$ is sufficiently smooth and $\nabla v \in C^{2,\alpha}$ (cf. Rem. 4.4), the adjoint problem (B.16) has a unique solution which satisfies the following elliptic regularity conditions, cf. Theorems 1.1 and 2.6 of Chapter 8 in [37],
\[
\|\psi\|_{H^2} \leq \hat{c}(\|\psi\|_{L^2} + \|g\|_{L^2})
\]
(B.18)
for some positive constant $\hat{c}$. Taking $\phi = \psi \in H^1_0$ in (B.17) and applying the coercive condition (S2b), we have
\[
M_1 \|\nabla \psi\|^2_{L^2} \leq \|g\|_{L^2} \|\psi\|_{L^2}.
\]
(B.19)
Applying Poincaré’s inequality in (B.19), we deduce that
\[
\|\psi\|_{L^2} \leq M_1^{-1} C_{\text{poin}} \|g\|_{L^2}.
\]

Thus
\[
\|\psi\|_{H^2} \leq c \|g\|_{L^2},
\]
(B.20)
for some positive constant $c$. The corresponding discrete problem is formulated as: find $\psi_h \in \mathcal{V}_h$ such that
\[
\tilde{a}(v; \psi_h, \phi) = (g, \phi)_{L^2}, \quad \text{for all } \phi \in \mathcal{V}_h.
\]
(B.21)

It is known that we have, cf., e.g. [33],
\[
\|\psi - \psi_h\|_{L^2} + h \|\psi - \psi_h\|_{H^1} \leq C h^{r+1} \|\psi\|_{H^{r+1}},
\]
(B.22)
for some constant $C$. Let $g = \dot{v} - \dot{W}$, then (B.16) becomes
\[
\tilde{a}(v, \psi, \phi) = \left(\dot{v} - \dot{W}, \phi\right)_{L^2}.
\]
(B.23)
Plugging $\phi = \dot{v} - \dot{W}$ into (B.23), we obtain
\[
\left\|\dot{v} - \dot{W}\right\|_{L^2}^2 = \tilde{a}(v, \psi, \dot{v} - \dot{W}).
\]
(B.24)
Using (S2a) and the definition of the elliptic projection, we have,
\[
\tilde{a}(v, \psi, \dot{v} - \dot{W}) = \tilde{a}(v, \psi - \psi_h, \dot{v} - \dot{W}) + \tilde{a}(v, \psi_h, \dot{v} - \dot{W})
= \tilde{a}(v, \psi - \psi_h, \dot{v} - \dot{W}) + \tilde{a}(v, \dot{v} - \dot{W}, \psi_h)
= \tilde{a}(v, \psi - \psi_h, \dot{v} - \dot{W}) + \tilde{a}(W, \psi_h) - \tilde{a}(v, \dot{W}, \psi_h).
\]
(B.25)
By (iii,a), (B.20) and (B.22), we have
\[
\begin{align*}
|\tilde{a}(v; \psi - \psi_h, \hat{v} - \hat{W})| & \leq K_\delta \|\nabla (\hat{v} - \hat{W})\|_{L^2} \|\nabla(\psi - \psi_h)\|_{L^2} \\
& \leq K_\delta C_r(v) h^r \cdot Ch \|\psi\|_{H^2} \quad \text{(by (iii,a) and (B.22))} \\
& \leq K_\delta c C_r(v) C h^{r+1} \|\hat{v} - \hat{W}\|_{L^2} \quad \text{(by (B.20))}.
\end{align*}
\]

For the remaining terms in (B.25), we observe that
\[
\begin{align*}
|\tilde{a}(W; \hat{W}, \psi_h) - \tilde{a}(v; \hat{W}, \psi_h)| & \leq |\tilde{a}(W; \hat{W}, \psi_h - \psi) - \tilde{a}(v; \hat{W}, \psi_h - \psi)| \\
& \quad + |\tilde{a}(W; \hat{W}, \hat{v}) - \tilde{a}(v; \hat{W}, \hat{v})| \\
& \quad + |\tilde{a}(W, \hat{v}, \hat{v}) - \tilde{a}(v, \hat{v}, \hat{v})| \\
& := T_4 + T_5 + T_6.
\end{align*}
\]

By Lipschitz continuity of $A_{\alpha j \beta}$, we have
\[
\begin{align*}
T_4 & \leq L_\delta \|\nabla(W - v)\|_{L^2} \|\nabla \hat{W}\|_{L^\infty} \|\nabla(\psi_h - \psi)\|_{L^2} \\
& \leq C_r(v) h^r c_1 \|\nabla(\psi_h - \psi)\|_{L^2} \quad \text{(by (iii,a) and (iii,c))} \\
& \leq L_\delta c C_r(v) c_1 h^r \cdot Ch \|\psi\|_{H^2} \quad \text{(by (B.22))} \\
& \leq c C C_r(v) c_1 h^{r+1} \|\hat{W} - \hat{v}\|_{L^2} \quad \text{(by (B.20))}.
\end{align*}
\]

Similarly, we have
\[
T_5 \leq L_\delta \|\nabla(W - v)\|_{L^\infty} \|\nabla(\hat{W} - \hat{v})\|_{L^2} \|\nabla \psi\|_{L^2}.
\]

Following the analysis in [38] and Chapter 8 of [39], it can be shown that
\[
\|v - W\|_{W^{1,\infty}} \leq c(v) h^r,
\]
where $c(v)$ is a positive constant depending on the exact solution $v$. Therefore, we can bound $T_5$ by
\[
T_5 \leq L_\delta c(v) h^r C_r(v) h^r \|\psi\|_{H^2} \leq L_\delta c(v) C_r(v) C h^{r+1} \|\hat{W} - \hat{v}\|_{L^2},
\]
for any $r \geq 1$ provided that $h$ is sufficiently small. We bound $T_6$ by
\[
\begin{align*}
T_6 & = |\tilde{a}(W; \hat{v}, \psi) - \tilde{a}(v; \hat{v}, \psi)| \\
& \leq \left| \sum_{i, \alpha, j, \beta, k, \gamma = 1}^d \left( \partial_\gamma(W_k - v_k) \partial_\beta \hat{v}_j \frac{\partial A_{\alpha j \beta}}{\partial \eta_{k \gamma}} (\nabla v), \partial_\alpha \psi_i \right) \right| \\
& \quad + \left| \sum_{i, \alpha, j, \beta, k, \gamma, l, \delta = 1}^d \left( \partial_\delta(W_l - v_l) \partial_\gamma(W_k - v_k) \partial_\beta \hat{v}_j \int_0^1 \frac{\partial^2 A_{\alpha j \beta}}{\partial \eta_{l \delta} \partial \eta_{k \gamma}} (\nabla v + \tau \nabla(W - v)) d\tau, \partial_\alpha \psi_i \right) \right| \\
& := b(W; v; \hat{v}, \psi) + d(W; v; \hat{v}, \psi).
\end{align*}
\]
To ensure that \( \nabla W \in \mathcal{Z}_\delta \), we take \( h \) sufficiently small. By the convexity of \( \mathcal{Z}_\delta \) and \( \mathcal{M}_\delta \), we know that 
\[
\nabla v(x) + \tau \nabla(W(x) - v(x)) \in \mathcal{M}_\delta \quad \text{for } x \in \Omega.
\]
Since \( A_{i\alpha j\beta} \) is sufficiently smooth (in particular twice continuously differentiable on \( \mathcal{M}_\delta \)), we have 
\[
d(W, v; \dot{v}, \psi) \leq C_A \| \nabla(W - v) \|_{L^\infty} \| \nabla \dot{v} \|_{L^2} \| \nabla \psi \|_{L^2} \\
\leq C_A C(v)^2 h^{2r} \| \nabla \dot{v} \|_{L^2} \| \nabla \psi \|_{L^2} \quad \text{(by (B.26))}
\]
\[
\leq C_A C(v)^2 h^{2r} \| \nabla \dot{v} \|_{H^2} \| \dot{W} - \dot{v} \|_{L^2} \quad \text{(by (B.20))}
\]
\[
\leq Ch^{r+1} \| \dot{W} - \dot{v} \|_{L^2} \text{ for } r \geq 1.
\]

For the estimation of \( b(W, v; \dot{v}, \psi) \), we apply integration by parts and the fact that \( \mathcal{V}_h \subset H_0^1 \) to obtain 
\[
b(W, v; \dot{v}, \psi) = \left| \sum_{i, \alpha, j, \beta, \gamma, \kappa = 1}^{d} \left( \partial_i \left[ \partial_j A_{i\alpha j\beta}(\nabla v), \partial_\alpha \psi \right] (W_k - v_k) \right) \right|
\]
\[
\leq C \| W - v \|_{L^2} \| \dot{v} \|_{H^2} \| W \|_{W^2, \infty} \| v \|_{W^2, \infty} \| \psi \|_{H^2}
\]
\[
\leq \tilde{C}h^{r+1} \| \dot{W} - \dot{v} \|_{L^2} \quad \text{(by (B.20) and (B.15)).}
\]

Combining the above estimates for \( T_4, T_5 \) and \( T_6 \), we have 
\[
\| \dot{W} - \dot{v} \|_{L^2} \leq \tilde{C}_r(v) h^{r+1},
\]
(B.28)

for some positive constant \( \tilde{C}_r(v) \).

By a similar argument, we can easily show that \( \dot{W}(t) \) is differentiable with respect to \( t \) and a similar \( L^2 \) error estimate for \( \dot{W} - \dot{v} \). The proof of this estimate can be found in \([40, 41]\). We omit the details here.

**Acknowledgements.** The author would like to thank Prof. Endre Suli for his helpful discussions and suggestions.

**References**

[1] T. Kato, Quasi-linear equations of evolution, with applications to partial differential equations, in Spectral Theory and Differential Equations. Springer (1975) 25–70.

[2] T.J.R. Hughes, T. Kato and J.E. Marsden, Well-posed quasi-linear second-order hyperbolic systems with applications to nonlinear elastodynamics and general relativity. *Arch. Ration. Mech. Anal.* **63** (1976) 273–294.

[3] C.M. Dafermos and W.J. Hrusa, Energy methods for quasilinear hyperbolic initial-boundary value problems: applications to elastodynamics. *Arch. Ration. Mech. Anal.* **87** (1985) 267–292.

[4] C.P. Chen and W. von Wahl, Das Rand-Anfangswertproblem für quasilineare Wellengleichungen in Sobolevräumen niedriger Ordnung. *J. Reine Angew. Math.* **337** (1982) 77–112.

[5] G.A. Baker, V.A. Dougalis and S.M. Serbin, High order accurate two-step approximations for hyperbolic equations. *RAIRO Anal. Numér.* **13** (1979) 201–226.

[6] L.A. Bales, Semidiscrete and single step fully discrete approximations for second order hyperbolic equations with time-dependent coefficients. *Math. Comp.* **43** (1984) 383–414.

[7] G.A. Baker and J.H. Bramble, Semidiscrete and single step fully discrete approximations for second order hyperbolic equations. *RAIRO Anal. Numér.* **13** (1979) 75–100.

[8] L.A. Bales and V.A. Dougalis, Cosine methods for nonlinear second-order hyperbolic equations. *Math. Comp.* **52** (1989) 299–319, S15–S33.

[9] L.A. Bales, Higher order single step fully discrete approximations for second order hyperbolic equations with time dependent coefficients. *SIAM J. Numer. Anal.* **23** (1986) 27–43.

[10] T. Dupont, \( L^2 \)-estimates for Galerkin methods for second order hyperbolic equations. *SIAM J. Numer. Anal.* **10** (1973) 880–889.

[11] J. E. Dendy, Jr., Galerkin’s method for some highly nonlinear problems. *SIAM J. Numer. Anal.* **14** (1977) 327–347.
Subscribe to Open (S2O)
A fair and sustainable open access model

This journal is currently published in open access under a Subscribe-to-Open model (S2O). S2O is a transformative model that aims to move subscription journals to open access. Open access is the free, immediate, online availability of research articles combined with the rights to use these articles fully in the digital environment. We are thankful to our subscribers and sponsors for making it possible to publish this journal in open access, free of charge for authors.

Please help to maintain this journal in open access!

Check that your library subscribes to the journal, or make a personal donation to the S2O programme, by contacting subscribers@edpsciences.org

More information, including a list of sponsors and a financial transparency report, available at: https://www.edpsciences.org/en/maths-s2o-programme