Discontinuous Galerkin discretisation in time of the second-order hyperbolic PDEs

Aili Shao

Magdalen College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2022

Abstract

A discontinuous Galerkin (DG) time-stepping method is presented for solving second-order hyperbolic partial differential equations (PDEs). The proposed numerical method combines the *hp*-version discontinuous Galerkin finite element method (hp-DGFEM) in the time direction with an $H^1(\Omega)$-conforming finite element approximation for the spatial variables. We start with the construction and analysis of the discontinuous-in-time scheme to linear hyperbolic PDEs of second order in Chapter 2. Our analysis shows that this method is consistent and stable, with arbitrarily high-order convergence rates in the temporal domain for sufficiently smooth solutions. Error bounds in both energy and L^2 norms are derived. These error estimates show that this method allows for a broad range of *hp*-refinement strategies with varying time step sizes and polynomial degrees, thus having the potential to give exponential rates of convergence. Numerical experiments on linear wave equations and elastodynamics systems show significant gains in accuracy over existing time integration schemes.

We then extend this DG time-stepping method to approximate solutions of second-order quasilinear hyperbolic systems in Chapter 3. In particular, we study the nonlinear elastodynamics problem. The resulting numerical scheme is stable and convergent, and we derive *a priori* error bounds at nodal points in the L^2-norm for sufficiently regular solutions. Numerical experiments on a nonlinear elastodynamics problem with smooth solutions demonstrate the convergence rates.

Chapter 4 further applies this discontinuous-in-time scheme to a nonlinear damped

equation, which is derived from Maxwell's equations by assuming a linearly polarised wave propagating on an infinite cylindrical domain. Again, we show *a priori* error estimates of the solutions at the nodal points in this chapter. Numerical experiments on nonlinear wave equations illustrate and confirm these theoretical findings.

Acknowledgement

First and foremost, my deepest thanks go to my supervisor, Professor Endre Süli, whose kind support and clear guidance have shaped me into the researcher I am today. It has been my great pleasure to work with you for these years. Whenever I need your help, you have always been generous with your time.

I would also like to express my sincere gratitude to Dr Kathryn Gillow, Professor Qian Wang and Professor Gui-Qiang Chen for acting as examiners for my Transfer of Status and Confirmation of Status. Your constructive comments and invaluable suggestions have helped shape the thesis into today's form.

Moreover, I wish to thank Professor Jan Kristensen, my undergraduate tutor from Magdalen College, for inspiring me with many advanced topics in Functional Analysis and Calculus of Variations.

The OxPDE Group has provided the best research atmosphere during my studies. Special thanks go to Dr Kaibo Hu and Fabian Laakmann for your helpful discussions about the FEniCs implementation of my numerical schemes. I would also like to thank Christopher Irving and Zhuolin Li for interesting mathematical conversations regarding topics in elliptic regularity.

Finally, my endless gratitude and love go to my partner, Greg J. Henderson, for all his inspiration, love and support.

Contents

List of Figures

List of Tables

List of Symbols

d	space dimension
$\mathbb{R}, \mathbb{N}$	real numbers, non-negative integers
Ω	an open and bounded subset of $\mathbb{R}^d$
$\overline{\Omega}$	the closure of Ω
$\partial\Omega$	the boundary of the domain Ω
$(\cdot,\cdot)_{L^2}$	the inner product in $L^2(\Omega)$ and $[L^2(\Omega)]^d$
$\langle\cdot,\cdot\rangle$	the dual pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$ or $[H^{-1}(\Omega)]^d$ and $[H_0^1(\Omega)]^d$
$C^{k,\gamma}(\overline{\Omega})$	the Hölder space consists of all functions $u \in C^k(\overline{\Omega})$ for which $\Vert u\Vert_{C^{k,\gamma}(\overline{\Omega})} := \sum_{\vert\alpha\vert\le k} \Vert D^\alpha u\Vert_{C(\overline{\Omega})} + \sum_{\vert\alpha\vert=k} [D^\alpha u]_{C^{0,\gamma}(\overline{\Omega})} < \infty$ with $[D^\alpha u]_{C^{0,\gamma}(\overline{\Omega})} := \sup\limits_{x,y\in\Omega, x\neq y} \left\{ \frac{\vert D^\alpha u(x) - D^\alpha u(y)\vert}{\vert x-y\vert^\gamma} \right\}$
$\vert\vert\vert \cdot \vert\vert\vert$	energy norm cf. Proposition 2.4
$\Vert\cdot\Vert_{L^p(\Omega)}$	Lebesgue-norms with $\Vert u\Vert_{L^p(\Omega)} := \begin{cases} \left(\int_\Omega \vert u\vert^p \,\mathrm{d}x\right)^{\frac{1}{p}}, & \text{if } p \in [1,\infty) \\ \operatorname*{ess.sup}_{\Omega} \vert u\vert, & \text{if } p = \infty \end{cases}$
$\Vert\cdot\Vert_{H_0^1(\Omega)}$	Sobolev-norm with $\Vert u\Vert_{H_0^1(\Omega)} := \left(\int_\Omega \vert\nabla u\vert^2 \,\mathrm{d}x\right)^{\frac{1}{2}}$
$\Vert\cdot\Vert_{H^{-1}(\Omega)}$	Sobolev-norm with $\Vert u\Vert_{H^{-1}(\Omega)} := \sup\left\{ \langle u,\varphi\rangle \mid \varphi \in H_0^1(\Omega), \Vert\varphi\Vert_{H_0^1(\Omega)} \le 1 \right\}$
$L^p(0,T;X)$	space of measurable functions $u\colon (0,T] \to X$ with $\Vert u\Vert_{L^p(0,T;X)} := \begin{cases} \left(\int_0^T \Vert u(t)\Vert_X^p \,\mathrm{d}t\right)^{\frac{1}{p}} < \infty, & \text{if } p \in [1,\infty), \\ \operatorname*{ess.sup}_{0\le t\le T} \Vert u(t)\Vert_X < \infty, & \text{if } p = \infty, \end{cases}$

	e.g., $X = L^2(\Omega), H_0^1(\Omega)$, $H^{-1}(\Omega)$ and $H^m(\Omega)$ for $m \in \mathbb{N}$
	or $X = [L^2(\Omega)]^d, [H_0^1(\Omega)]^d$, $[H^{-1}(\Omega)]^d$ and $[H^m(\Omega)]^d$
$H^k(0,T;X)$	space of measurable functions $u\colon (0,T] \to X$ with
	$\|u\|_{H^k(0,T;X)} := \left(\int_0^T \sum_{\alpha=0}^k \|\frac{\partial^\alpha u}{\partial t^\alpha}\|_X^2 \,\mathrm{d}t\right)^{\frac{1}{2}}$
$W^{s,\infty}(0,T;X)$	space of measurable functions $u\colon (0,T] \to X$ with
	$\|u\|_{W^{s,\infty}(0,T;X)} := \sum_{\alpha=0}^s \operatorname*{ess.sup}_{0\le t\le T} \|\frac{\partial^\alpha u(t)}{\partial t^\alpha}\|_X < \infty$ for $s \in \mathbb{N}$
$C^k([0,T];X)$	space of measurable functions $u\colon [0,T] \to X$ with
	$\|u\|_{C^k([0,T];X)} := \sum_{\alpha=0}^k \max_{0\le t\le T} \|\frac{\partial^\alpha u(t)}{\partial t^\alpha}\|_X < \infty$ for $k \in \mathbb{N}$
$\mathcal{P}_h$	Ritz projection operator in the spatial direction
$\mathcal{P}^q$	Boundary value preserving L^2-projection operator in the time direction
Π^q	the modified L^2-projection operator in the time direction, cf. Definition 2.9
∂_α	the partial derivative with respect to x_α
$[x]$	the greatest integer less than or equal to x
$\lvert \cdot \rvert$	the Euclidean l_2-norm of a vector or the Frobenius norm of a matrix
$\mathcal{T}_h$	a quasi-uniform finite element triangulation
$\mathcal{N}$	the fixed point map, cf. Eq. 3.3.22 and Eq. 4.3.19
$\mathcal{F}$	the fixed point subset, cf. Eq. 3.3.20 and Eq. 4.3.17

Chapter 1

Introduction

1.1. Motivation

Second-order hyperbolic equations arise in a wide range of relevant applications, including acoustic wave equations, nonlinear elastodynamics systems, electromagnetic wave propagation problems, etc. For nonlinear wave equations, the superposition principle does not generally apply. This means that such equations are more difficult to analyse mathematically and no general analytical method for solving them exists. Thus, developing numerical methods for wave-type problems, especially for nonlinear equations, has been a constant interest in the field of partial differential equations (PDEs). For a certain class of nonlinear second-order hyperbolic equations, no global smooth solutions exist due to the breaking of waves and the formation of shocks, and this feature adds more difficulty in solving such problems numerically. For such time-dependent problems, the crucial step in numerical discretisation is to choose an appropriate time-stepping scheme. Traditional approaches for numerical integration include explicit and implicit finite difference, Runge–Kutta [50, 67] and Newmark–β [58] methods. Though implicit schemes are typically unconditionally stable, explicit schemes are usually preferred in engineering and physical applications because of their computational convenience. The main drawback of explicit methods is the limitation on the time step imposed by the Courant–

Friedrichs–Lewy (CFL) [22] condition. Thus, a flexible and efficient time discretisation method has become a topic of great interest in numerical analysis and computational mathematics. Different from the above-mentioned finite difference schemes, we introduce an implicit arbitrarily high-order accurate time-stepping scheme based on an *hp*–version of the discontinuous Galerkin (DG) method [6, 75] to discretise second-order hyperbolic PDEs.

1.2. What is a discontinuous Galerkin time-stepping method?

Discontinuous Galerkin (DG) methods [51, 64] were first introduced by Reed and Hill [64] to solve the hyperbolic neutron transport equations, and then generalised to elliptic and parabolic problems by Babuška & Zlámal [12], Baker [13], Wheeler [78], Arnold [9] and Rivière [66] etc. Relevant analysis and applications of DG methods to first-order hyperbolic problems can be found in [20, 21, 37, 48, 51, 64, 65, 73]. Several discontinuous Galerkin finite element methods (DGFEM) for solving wave-type equations have appeared in the literature [3, 39, 47, 63].

To construct this DG time-stepping scheme, we first discretise the spatial variables using a Galerkin finite element method, which results in a system of ordinary differential equations (ODEs) with respect to the temporal variable, then discretise the resulting ODE system using the DG method. The resulting weak formulation in time is based on weakly imposing the continuity of the approximate displacements and velocities between time steps by penalising jumps in these quantities in the definition of the numerical method. In contrast with traditional 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 discretisation parameter.

1.3. A note on the history of the DG time-stepping method

Here we give a review of the development of the time-discontinuous Galerkin method.

The DG time-stepping method was first introduced for first order linear non-stiff ODEs by Delfour, Hager, and Trochu [25], and a nodal convergence of order $O(k^{2q+1})$, where q is the polynomial degree of the finite element space, was proved in this work. Then this discontinuous-in-time scheme was applied to linear parabolic problems in a time-dependent domain by Jamet [46] who proved unconditional stability of the scheme and general error estimates of order $O(k^{q+1})$ where $k = h$ is both the time step size and the spatial mesh size, and q is the minimum of the polynomial degrees in the spatial and time variables. Later on, the DGFEM for parabolic problems was studied in a series of papers by Eriksson, Johnson, Thomée *et al.* [28–34]. In these works, an h-version of the DGFEM was first introduced and analysed. That is, the convergence of the discrete solution to the exact solution is achieved by reducing the mesh size h and the time step k respectively. The rates of convergence in both time and space are only algebraic. To achieve an exponential convergence in time and space, the hp-version of the DGFEM was developed by Schötzau and Schwab [68, 69]. In particular, the following *a priori* error estimates have been proved in [69]:

Theorem 1.1. *Let $u \in H^{s+1}(J; X)$ for $s \geq 0$ be the exact solution of the following linear parabolic problem*

$$\begin{aligned} u'(t) + Lu(t) &= g(t), \quad t \in J = (0, T) \\ u(0) &= u_0, \end{aligned} \tag{1.3.1}$$

where L is a general elliptic operator, X is a general Hilbert space, u_0 is the initial data and g is the forcing term. Let $q_m = q$ be the polynomial degrees in time on each time interval I_m and set $k = \max\{k_m\}$ where k_m is the time step size on each time interval

I_m. For the DGFEM approximation $u_{\mathrm{DG}} \in \mathcal{V}^q(\mathcal{M}; X)$, where

$$\mathcal{V}^q(\mathcal{M}; X) := \{u\colon J \to X : u\,|_{I_m} \in \mathcal{P}^q(I_m, X) \text{ for } 1 \leq m \leq M\}, \tag{1.3.2}$$

we have the error bound

$$\|u - u_{\mathrm{DG}}\|_{L^2(J;X)} \leq C k^{\min(q,s)+1} q^{-(s+1)} \|u\|_{H^{s+1}(J;X)} \tag{1.3.3}$$

with a constant C depending on s only.

Let N be the total number of time degrees of freedom. If we consider the h-version of this DGFEM scheme, where the convergence is achieved by reducing the time step k at a fixed approximation order q, such that $N \sim \frac{1}{k}$, we have

$$\|u - u_{\mathrm{DG}}\|_{L^2(J;X)} \leq C N^{-\min(q,s)-1}. \tag{1.3.4}$$

If we consider the p-version of this scheme where convergence is achieved by increasing the approximation order q on a fixed time partition, such that $N \sim q$, we have

$$\|u - u_{\mathrm{DG}}\|_{L^2(J;X)} \leq C N^{-s-1}. \tag{1.3.5}$$

The extension to p- or hp-version of the original h-version DGFEM also relies on the solution behaviour of parabolic problems. It has been shown in Schötzau and Schwab's work [68, 69] that the hp-version can resolve time singularities at an exponential convergence rate, independent of the spatial discretisation. Other than analysing the error estimates of the hp-DGFEM time-stepping for parabolic problems, the algorithmic features of this scheme were also explored by Schötzau, Schwab and their coauthors in [77]. The focus of [77] is on the efficiency aspects of the algorithm. For instance, the decoupling of the spatial systems within every time step was considered, and several parallelisation strategies were discussed. It has been verified that the decoupling process is of great value both in terms of computational time and memory requirements. They also showed that the exponential convergence rates for the hp-DGFEM resulted directly in a significant CPU time reduction as compared to the original h-version time-stepping scheme.

Based on the success of the hp-DGFEM for parabolic problems, Johnson [47] extended this time integration scheme to a second-order hyperbolic problem by converting the wave equation to a bigger first-order in time PDE system. Then *a priori* and *a posteriori* error estimates for linear wave equations were proved in this paper. The approximate solution u_{DG} corresponding to the displacement u in the linear wave equation is of accuracy orders of $O(h^2k^{-\frac{1}{2}} + k^3)$ and $O(h^2 + k^3)$, respectively, in the L^2-norm, and of order $O(hk^{-\frac{1}{2}} + k^3)$ and $O(h + k^3)$, respectively, in the energy norm, where h and k are the spatial and time steps as usual. Combining with previous work cited, Johnson also argued that the DGFEM is a good time discretisation scheme for parabolic or hyperbolic (both first and second-order) problems because of the following advantages:

(a) high order implicit A-stable time-stepping schemes are generated;

(b) optimal (for parabolic problems) or nearly optimal (for hyperbolic problems) *a priori* and *a posteriori* error estimates can be proved using general variational techniques including duality;

(c) the space-time mesh on each slab S_n can be chosen independently giving the possibility of (adaptively) changing the mesh from one slab to the next orienting the mesh along with characteristics or according to features of the exact solution, which increases the precision and permits larger time steps;

(d) reliable and efficient adaptive versions based on the *a posteriori* error estimates can be constructed;

(e) moving or free boundaries can be handled without unnecessary technical difficulties.

Before Johnson's work, the DG methods were first applied to elastodynamics equations by Hulbert and Hughes [42, 45] and Hulbert [44] using the space-time finite element discretisation. That is, discontinuous Galerkin approximations were implemented on space-time 'slabs' $S_n = \Omega \times I_n$, where Ω is the underlying spatial domain and $I_n =$

(t_n, t_{n+1}), for $n = 1, 2, \ldots, N$, are time intervals. The formulation of this fully discrete scheme employs additional least-squares (also known as Petrov-Galerkin) terms, which enhance stability. The elastodynamics equations Hulbert and Hughes considered are:

$$\rho\ddot{\mathbf{u}} = \nabla \cdot \boldsymbol{\sigma}(\nabla \mathbf{u}) + \mathbf{f} \quad \text{on } Q := \Omega \times (0, T), \tag{1.3.6}$$

$$\mathbf{u} = \mathbf{g} \quad \text{on } P_g := \Gamma_g \times (0, T), \tag{1.3.7}$$

$$\mathbf{n} \cdot \boldsymbol{\sigma}(\nabla \mathbf{u}) = \mathbf{h} \quad \text{on } P_h := \Gamma_h \times (0, T), \tag{1.3.8}$$

$$\mathbf{u}(x, 0) = \mathbf{u}_0, \quad x \in \Omega, \tag{1.3.9}$$

$$\dot{\mathbf{u}}(x, 0) = \mathbf{v}_0, \quad x \in \Omega, \tag{1.3.10}$$

where $\rho = \rho(x) > 0$ is the density, $\mathbf{f}$ is the body force, $\mathbf{g}$ is the prescribed boundary displacement, $\mathbf{h}$ is the prescribed boundary traction, and $\mathbf{n}$ is the unit outward normal to Γ. Define $\hat{h} = \max(ck, h)$, where c is the dilatational wave speed, h and k are maximum element diameters in space and time respectively. Assuming that the exact solution to (1.3.6) is regular enough, that is, $\mathbf{u} \in [H^{q+1}(Q)]^d$, then it was proved by Hughes and Hulbert [45] that

$$|||\mathbf{u} - \mathbf{u}_{\mathrm{DG}}|||^2 \leq C(\mathbf{u})\hat{h}^{2q-1}, \tag{1.3.11}$$

where $\mathbf{u}_{\mathrm{DG}}$ is the DG approximation to the exact solution based on the least square Galerkin formulation, $C(\mathbf{u})$ is a constant dependent on $\mathbf{u}$, but not on $\hat{h}$, and $|||\cdot|||$ is a norm that is stronger than the total energy norm. To be more precise,

$$\begin{aligned} |||\mathbf{e}|||^2 :&= \mathcal{E}(\mathbf{e}(T^-)) + \mathcal{E}(\mathbf{e}(0^+)) + \sum_{n=1}^{N-1} \mathcal{E}([\![\mathbf{e}(t_n)]\!]) + \sum_{n=1}^{N} (\mathscr{L}\mathbf{e}, \rho^{-1}\boldsymbol{\tau}\mathscr{L}\mathbf{e})_{L^2(Q_n^\Sigma)} \\ &+ \sum_{n=1}^{N} \left(\mathbf{n} \cdot [\![\boldsymbol{\sigma}(\nabla \mathbf{e})]\!], \rho^{-1}\mathbf{s}\mathbf{n} \cdot [\![\boldsymbol{\sigma}(\nabla \mathbf{e})]\!]\right)_{L^2(Y_n^\Sigma)} \\ &+ \sum_{n=1}^{N} (\mathbf{n} \cdot \boldsymbol{\sigma}(\nabla \mathbf{e}), \rho^{-1}\mathbf{s}\mathbf{n} \cdot \boldsymbol{\sigma}(\nabla \mathbf{e}))_{L^2((Y_h)_n)}, \end{aligned} \tag{1.3.12}$$

where

$$\mathcal{E}(\mathbf{w}) := \frac{1}{2}\int_\Omega \dot{\mathbf{w}} \cdot \rho\dot{\mathbf{w}} \, \mathrm{d}\Omega + \frac{1}{2}\int_\Omega \nabla \mathbf{w} \cdot \boldsymbol{\sigma}(\nabla \mathbf{w}) \, \mathrm{d}\Omega$$

and $\mathscr{L}$ is the differential operator, defined as

$$\mathscr{L}\mathbf{w} := \nabla \cdot \boldsymbol{\sigma}(\nabla \mathbf{w}) + \mathbf{f}.$$

Here Q_n^Σ and Y_n^Σ represent the interior boundaries of the corresponding domains. The matrix operators $\boldsymbol{\tau}$ and $\mathbf{s}$ in the stability terms are the so-called time-scale and slowness matrices. Hulbert and Hughes' numerical results not only confirmed the theoretical convergence rates in [45] but also hinted that some simplifications of the DG formulation may be possible. Moreover, though Hughes [44] considered the fully coupled space-time finite element method for both linear and nonlinear elastodynamics, convergence and error estimates were only proved for the linear case. For the nonlinear elastodynamics problem, only a stability result was presented. Hence, this motivates us to design a simplified version of the discontinuous time-stepping method for nonlinear elastodynamics, or say, nonlinear hyperbolic PDEs of second-order in general, with a convergence proof included.

In recent years, different versions of discontinuous time-stepping methods have been proposed for second-order in time ODEs or linear wave equations. For instance, Adjerid and Temimi [3] introduced a new DGFEM that combines the pth-degree standard continuous finite element method in space and the qth-degree discontinuous Galerkin method in time, which is directly applied to the resulting second-order ODE system [2]. In [3], Adjerid and Temimi proved that the DG solution converges to the exact solution at $O(h^{p+1-s}) + O(k^{q+1})$ rate in the $L^2([0,T], H^s(\Omega)), s = 0, 1,$ norm provided that the exact solution is smooth enough. They also established that the DG solution is superconvergent at the tensor product of Lobatto points in space and Jacobi points in time, e.g., the DG solution at Lobatto points and t_n^- are $O(h^{p+2}) + O(k^{2q})$ super-convergent. Based on Schwab and Schötzau's hp-DGFEM scheme for linear parabolic problems, Antonietti *et al.* [6] applied this hp-version of discontinuous-in-time integration scheme to systems of second-order ODEs. Simultaneously as this thesis is written, Antonietti *et* *al.* [7] presented a new high-order space-time discretisation method for the elastic wave equation. This scheme is the result of a combination between the discontinuous-in-time discretisation presented in [6] and the spatial DG formulation introduced in [8].

Though discontinuous-in-time Galerkin schemes have been widely used to solve time-dependent problems in the past forty years, there has been little work on the construction and mathematical analysis of the DG time-stepping methods for second-order nonlinear hyperbolic PDEs, which has been the main focus of the research presented herein.

1.4. Contributions of this thesis

In this thesis, we generalise this high-order discontinuous-in-time method [6] to second-order hyperbolic-type PDEs, which arise in a wide range of relevant applications, including the acoustic wave equation (cf. Chapter 2), nonlinear elastodynamics systems (cf. Chapter 3), and an electromagnetic wave propagation problem (cf. Chapter 4).

In Chapter 2, discontinuous Galerkin time-stepping methods are presented for linear second-order hyperbolic equations. The semi-discrete scheme provides a clear demonstration of how we apply the DG scheme over the temporal domain. To achieve stability, we impose the continuity of the approximate solution and its derivative between time steps by penalising jumps in these quantities in the weak formulation. This gives us an implicit, unconditionally stable and arbitrarily high-order accurate time integration scheme. We then fully discretise the problem by combining this *hp*–version of DGFEM for temporal discretisation with an H^1-conforming finite element approximation for the spatial variables. *A priori* error estimates are derived both in the energy norm and the L^2-norm. In particular, we proved a convergence rate of $O(h^r + k_n^{q_n-\frac{1}{2}})$ for approximation of the displacement in the energy norm while a convergence rate of $O(h^{r+1} + k_n^{q_n})$ is proved for the approximation of the velocity in the L^2-norm on each time interval $I_n = (t_{n-1}, t_n]$. Here k_n, for $n = 1, \ldots, N$, and h are the temporal and spatial discretisation parameters, respectively, and q_n, for $n = 1, \ldots, N$, and r are the corresponding polynomial degrees.

Numerical experiments on both a scalar wave equation and a linearised elastodynamics system (two-dimensional) are presented to verify the theoretical results. The main advantages of this new proposed discontinuous-in-time scheme over those in the existing literature [3, 7, 44, 47] are the following:

(a) In contrast with Johnson's method [47] which converts the wave equation into a first-order system, this direct discretisation leads to a smaller fully discrete problem (i.e., the discretisation matrices are smaller in size) for the solution while maintaining arbitrarily high-order convergence rates [2].

(b) Less regularity is required for the first-order derivative of the solution.

(c) The numerical formulation is simplified as compared to Hughes' [44] space-time finite element method while the stability and convergence results are preserved.

(d) The extension to the hp–version of the DGFEM shows a better convergence result than the h–version [3].

(e) A more structured proof of the convergence results for linear hyperbolic PDEs of second-order, not limited to elastic-wave propagation [7], is presented.

(f) This flexible and efficient discontinuous-in-time scheme can be naturally applied to nonlinear hyperbolic equations of second-order.

In Chapter 3, we extend this DG time-stepping method to approximate solutions of second-order quasilinear hyperbolic systems. In contrast to Hughes' [44] space-time finite element method where the convergence result was only proved for linearised elastodynamics problems, we present a detailed convergence proof for the nonlinear elastodynamics systems based on a *Banach fixed point* argument. *A priori* error bounds in the L^2-norm for sufficiently regular solutions are proved. In particular, we have shown that

$$\|\mathbf{u}(t_j^-) - \mathbf{u}_{\mathrm{DG}}(t_j^-)\|_{L^2(\Omega)} + \|\dot{\mathbf{u}}(t_j^-) - \dot{\mathbf{u}}_{\mathrm{DG}}(t_j^-)\|_{L^2(\Omega)} \le C(\mathbf{u}) \left(h^{2r+2} + \sum_{n=1}^{j} \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} \right)^{\frac{1}{2}}, \tag{1.4.1}$$

where k_n, for $n = 1, \dots, N$, and h are the temporal and spatial discretisation parameters, respectively, such that $k_n^{q_n - \frac{1}{2}} = o(h^{\frac{d}{2}+1})$ and there exist positive constants μ_n and ν_n such that $\mu_n k_n \le h^2 \le \nu_n k_n$ for $n = 1, \dots, N$. Here q_n for each $n = 1, \dots, N$ is the polynomial degree in time integration and r satisfies $\frac{d}{2} + 1 < r \le \min(p, m-1)$, where p is the polynomial degree of the spatial finite element space while m is the regularity index of the spatial variable (i.e., $\mathbf{u} \in W^{s,\infty}(0, T; [H^m(\Omega)]^d \cap [H_0^1(\Omega)]^d)$ with $s \ge q_n + 1$). This *a priori* error estimate shows that the method is arbitrarily high-order accurate with respect to the time-step size and temporal polynomial degree. Numerical experiments on a quasilinear elastodynamics problem illustrate and confirm these theoretical findings.

Chapter 4 further applies this discontinuous-in-time scheme to a nonlinear damped wave equation, which is derived from Maxwell's equations by assuming a linearly polarised wave propagating on an infinite cylindrical domain. The existence of at least one weak solution is proved for a bounded domain with a smooth boundary under certain assumptions on the nonlinear terms, and the uniqueness follows if the exact solution is sufficiently regular. Again, we show *a priori* error estimates in the L^2-norm in this chapter. Slightly different from the error estimates for the quasilinear elastodynamics problem presented in Chapter 3, we consider the H^1-norm instead of the L^2-norm for the displacement error at nodal points. That is,

$$\|u(t_j^-) - u_{\mathrm{DG}}(t_j^-)\|_{H^1(\Omega)} + \|\dot{u}(t_j^-) - \dot{u}_{\mathrm{DG}}(t_j^-)\|_{L^2(\Omega)} \le C(u) \left(h^{2r} + \sum_{n=1}^{j} \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} \right)^{\frac{1}{2}}, \tag{1.4.2}$$

where k_n, q_n for $n = 1, \dots, N$, h and r are defined as above. Here, we only require $k_n^{q_n - \frac{1}{2}} = o(h^{\frac{d}{2}})$ instead of $k_n^{q_n - \frac{1}{2}} = o(h^{\frac{d}{2}+1})$, for each $n = 1, \dots, N$, for this case.

Conclusions and possible projects for future research are discussed in Chapter 5.

1.5. Preliminary results

We close this introductory chapter with a collection of results from Analysis, which are used repeatedly throughout the thesis.

1.5.1. Banach fixed point theorem The *Banach fixed point theorem*, introduced by Stefan Banach [14] in 1992, is an important tool for proving the existence and uniqueness of solutions for nonlinear problems. A key assumption for *Banach fixed point theorem* is the contraction, which is defined by the following.

Definition 1.2. *Let (X, d) be a metric space and $\mathcal{N}\colon \mathcal{F} \subset X \to X$. $\mathcal{N}$ is a contraction if there exists $\eta \in (0,1)$ such that*

$$\mathrm{d}(\mathcal{N}(x), \mathcal{N}(y)) \leq \eta\, \mathrm{d}(x, y). \tag{1.5.1}$$

The following version of the *Banach fixed point theorem* is used throughout the thesis (cf. Chapter 3 and Chapter 4) to prove the existence and uniqueness of discrete solutions.

Theorem 1.3. *Let (X, d) be a complete metric space. Assume that $\mathcal{F} \subset X$ is non-empty and closed, and $\mathcal{N}\colon \mathcal{F} \subset X \to \mathcal{F}$ is a contraction mapping. Then $\mathcal{N}$ admits a unique fixed point x^* in $\mathcal{F}$,(i.e., $\mathcal{N}(x^*) = x^*$). Furthermore, the sequence x_n defined via $x_{n+1} = \mathcal{N}(x_n)$ converges for every $x_0 \in \mathcal{F}$ to x^*.*

Proof. First note that since X is complete, and $\mathcal{F} \subset X$ is non-empty and closed, $(\mathcal{F}, \mathrm{d})$ is a complete metric space as well.

- Uniqueness: Assume for contradiction that there are $x, y \in \mathcal{F}$, $x \neq y$ such that $\mathcal{N}(x) = x$ and $\mathcal{N}(y) = y$, then $\mathrm{d}(\mathcal{N}(x), \mathcal{N}(y)) = \mathrm{d}(x, y)$. Since $\mathcal{N}$ is a contraction, i.e., $\mathrm{d}(\mathcal{N}(x), \mathcal{N}(y)) < \mathrm{d}(x, y)$, this yields a contradiction.

- Existence: Given $x_0 \in \mathcal{F}$, consider the iterated sequence $x_{n+1} = \mathcal{N}(x_n)$. The definition of the iterated sequences and the contraction assumption imply that

$$\mathrm{d}(x_{n+1}, x_n) = \mathrm{d}(\mathcal{N}(x_n), \mathcal{N}(x_{n-1})) \leq \eta \mathrm{d}(x_n, x_{n-1}),$$

and by induction $\mathrm{d}(x_{n+1}, x_n) \leq \eta^n \mathrm{d}(x_1, x_0)$. By the triangle inequality, we have

$$\mathrm{d}(x_{n+m}, x_n) \leq \sum_{p=0}^{m-1} \mathrm{d}(x_{n+p+1}, x_{n+p}) \leq \sum_{p=0}^{m-1} \eta^{n+p} \mathrm{d}(x_1, x_0) \leq \frac{\eta^n}{1-\eta} \mathrm{d}(x_1, x_0).$$

Since $\eta \in (0,1)$, we have $\mathrm{d}(x_{n+m}, x_n) \to 0$ as $n \to \infty$. This shows that (x_n) is a Cauchy sequence in $\mathcal{F}$. By completeness of $\mathcal{F}$, we deduce that the sequence (x_n) has a limit x^* in $\mathcal{F}$. Next, note that a contractive map is continuous, thus,

$$\mathcal{N}(x^*) = \mathcal{N}(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} \mathcal{N}(x_n) = \lim_{n\to\infty} x_{n+1} = x^*.$$

□

1.5.2. The Poincaré inequality We will make use of the following version of Poincaré inequality.

Lemma 1.4. *For every $v \in H_0^1(\Omega)$, there exists a positive constant C_{poin} depending only on the diameter of the domain Ω such that*

$$\|v\|_{H^1(\Omega)}^2 \leq C_{\mathrm{poin}} \|\nabla v\|_{L^2(\Omega)}^2. \tag{1.5.2}$$

Proof. See Adams [1]. □

1.5.3. Grönwall's inequality We need both the continuous and discrete versions of Grönwall's inequality.

Lemma 1.5 (Grönwall's inequality)**.** *Suppose that A is a bounded non-negative function on $[0,T]$ and that E and C are non-negative constants such that*

$$A(t) \leq E + C \int_0^t A(s)\,\mathrm{d}s, \quad 0 \leq t \leq T; \tag{1.5.3}$$

then we have that

$$A(t) \leq E \exp(Ct). \tag{1.5.4}$$

Proof. If we set $R(t) = E + C\int_0^t A(s)\,\mathrm{d}s$, it follows from (1.5.3) that $R'(t) \leq CR(t)$. Thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(R(t)\exp(-Ct)\Big) = \exp(-Ct)(R'(t) - CR(t)) \leq 0.$$

Then (1.5.4) follows immediately from the fact that

$$A(t)\exp(-Ct) \leq R(t)\exp(-Ct) \leq R(0)\exp(-C\cdot 0) = E.$$

□

Lemma 1.6 (Discrete Grönwall's inequality)**.** *Suppose that* $\{A_n\}$ *and* $\{g_n\}$ *are non-negative sequences and* E *is a non-negative constant such that*

$$A_n \leq E + \sum_{k=0}^{n} g_k A_k, \qquad \textit{for } n \geq 0, \tag{1.5.5}$$

then we have

$$A_n \leq E\exp\left(\sum_{k=0}^{n} g_k\right). \tag{1.5.6}$$

Proof. See Holte [41]. □

1.5.4. Other important results We also need the following results from Seregin's Lecture notes [71] and Lions and Magenes' book [53].

Theorem 1.7. *Assume that* $v \in L^2(0,T,H_0^1(\Omega))$ *and* $\partial_t v \in L^2(0,T;H^{-1}(\Omega))$*. Then* $v \in C([0,T];L^2(\Omega))$ *and*

$$\int_{t_1}^{t_2}\int_\Omega \partial_t v\cdot v\,\mathrm{d}x\,\mathrm{d}t = \frac{1}{2}\|v(\cdot,t_2)\|^2_{L^2(\Omega)} - \frac{1}{2}\|v(\cdot,t_1)\|^2_{L^2(\Omega)}$$

for all $t_1, t_2 \in [0,T]$.

Theorem 1.8. *Let* H *be a Hilbert space,* V *be a reflexive Banach space, and* V *is continuously embedded into* H*. Let* V *contain a countable set* S *which is dense in* V *and in* H*. Let* V^* *be a dual space to* V *with respect to scalar product in* H *with the norm*

$$\|v^*\|_{V^*} = \sup\{(v^*,v)_H : v \in V, \quad \|v\|_V = 1\}.$$

Assume that $v \in L^p(0,T;V) \cap L^2(0,T;H)$ *and* $\partial_t v \in L^{p'}(0,T;V^*)$ *with* $p' = \frac{p}{p-1}$ *and* $p > 1$. *Then,* $v \in C([0.T];H)$ *and*

$$\|v(\cdot,t_2)\|_H^2 - \|v(\cdot,t_1)\|_H^2 = 2\int_{t_1}^{t_2} (\partial_t v(\cdot,t), v(\cdot,t))_H \,\mathrm{d}t$$

for all $t_1, t_2 \in [0,T]$.

Chapter 2

Linear hyperbolic equations of second order

In this chapter, we apply a high-order discontinuous-in-time scheme to numerically solve linear hyperbolic PDEs of second order. We first discretise the PDEs in time while keeping the spatial differential operators undiscretised. The resulting weak formulation is based on weakly imposing the continuity of the approximate solution and its time derivative between time steps by penalising jumps in these quantities in the definition of the numerical method. We then fully discretise the problem using an H^1-conforming finite element approximation for the spatial variables.

Let $\Omega \subset \mathbb{R}^d$, for $d = 1, 2, 3$, be an open and bounded domain with sufficiently smooth boundary $\partial\Omega$ (i.e., $\partial\Omega \in C^2$). For $T > 0$, we consider the following initial boundary value problem:

$$\ddot{u}(x,t) + L\dot{u}(x,t) + Ku(x,t) = f(x,t) \text{ in } \Omega \times (0,T], \tag{2.0.1}$$

where f lies in $L^2(0,T;L^2(\Omega))$, and

$$u(0) = u_0(x) \in H_0^1(\Omega),\ \dot{u}(0) = u_1(x) \in L^2(\Omega), \tag{2.0.2}$$

$$u = 0 \text{ on } \partial\Omega \times (0,T], \tag{2.0.3}$$

are prescribed initial and boundary conditions respectively. Here the dots over u denote differentiation with respect to time t. Throughout this thesis, $\dot{u}$ (respectively $\ddot{u}$) and $\partial_t u$ (respectively $\partial_{tt} u$) are interchangeable. We also write $u(\cdot, t)$ as $u(t)$ when appropriate for the sake of having neater presentation. The differential operators K and L are defined as $K := -\Delta$ and $L := \mathrm{Id}$ respectively, where Id is the identity operator. We note that the techniques presented in this chapter can be extended to linear hyperbolic PDEs of second order with more general positive definite differential operators K and L. In this chapter, we focus on the numerical discretisation of the case $K = -\Delta$ and $L = \mathrm{Id}$ only. The numerical schemes for more general positive definite differential operators K and L are completely analogous and are therefore omitted.

2.1. Weak solution

We define the time-dependent bilinear form

$$B[u, \varphi; t] := (\dot{u}(t), \varphi)_{L^2} + (\nabla u(t), \nabla \varphi)_{L^2}$$

for $u \in L^2(0, T; H_0^1(\Omega))$ with $\dot{u} \in L^2(0, T; L^2(\Omega))$ and $\varphi \in H_0^1(\Omega)$.

Definition 2.1. *We say that a function $u \in L^2(0, T; H_0^1(\Omega))$, with $\dot{u} \in L^2(0, T; L^2(\Omega))$, $\ddot{u} \in L^2(0, T; H^{-1}(\Omega))$, is a weak solution of the hyperbolic initial boundary value problem (2.0.1)–(2.0.3) if*

(a) $\langle \ddot{u}(t), \varphi \rangle + B[u, \varphi; t] = (f(t), \varphi)_{L^2}$ *for all* $\varphi \in H_0^1(\Omega)$ *and* $0 \leq t \leq T$ *and*

(b) $u(0) = u_0$, $\dot{u}(0) = u_1$.

Remark 2.2. *It is well known that problem (2.0.1)–(2.0.3) is well-posed and admits a unique weak solution (see Theorem 29.1 in [79]). By Theorem* 8.1 *and Theorem* 8.2 *in Lions and Magenes' work [53], we know that $u \in C([0, T]; H_0^1(\Omega))$ and $\dot{u} \in C([0, T]; L^2(\Omega))$. This shows that the initial conditions are meaningfully attained in $H_0^1(\Omega)$ and $L^2(\Omega)$ respectively.*

Remark 2.3. *Note that if $K = -\Delta$ and $L = 0$, then we have the wave equation*

$$\ddot{u}(x,t) - \Delta u(x,t) = f(x,t) \text{ in } \Omega \times (0,T]$$

with $f \in L^2(0,T;L^2(\Omega))$, and initial/boundary conditions satisfying

$$u(0) = u_0 \in H_0^1(\Omega),\ \dot{u}(0) = u_1 \in L^2(\Omega),\ u = 0 \text{ on } \partial\Omega \times (0,T].$$

We can easily transform this problem into an equivalent form with L being a positive definite differential operator as discussed in this section. Consider $u = \mathrm{e}^{t\gamma}v$ for a positive real constant γ; then v satisfies

$$\mathrm{e}^{\gamma t}\ddot{v}(x,t) + 2\gamma\mathrm{e}^{\gamma t}\dot{v}(x,t) + \gamma^2\mathrm{e}^{\gamma t}v(x,t) - \mathrm{e}^{\gamma t}\Delta v = f(x,t).$$

Dividing by $\mathrm{e}^{\gamma t}$ on both sides, we have

$$\ddot{v}(x,t) + 2\gamma\dot{v}(x,t) + \gamma^2 v(x,t) - \Delta v(x,t) = f(x,t)\mathrm{e}^{-\gamma t}.$$

Then we can show the existence and uniqueness of a solution of this problem similarly as before but with $K = -\Delta + \gamma^2\mathrm{Id}$ and $L = 2\gamma\mathrm{Id}$.

2.2. Semi-discrete numerical scheme

In this section, we introduce a high-order DG method of lines approach to discretise (2.0.1)–(2.0.2) in the time direction. We will show that the resulting variational formulation is well-posed and thus admits a unique solution in the linear space, which we will define in this section.

2.2.1. Motivation We partition the interval $I = (0,T]$ into N subintervals $I_n = (t_{n-1}, t_n]$ having length $k_n = t_n - t_{n-1}$ for $n = 1, 2, \ldots, N$ with $t_0 = 0$ and $t_N = T$, as shown below. To deal with the discontinuity at each t_n in the numerical approximation to u, we introduce the jump operator

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

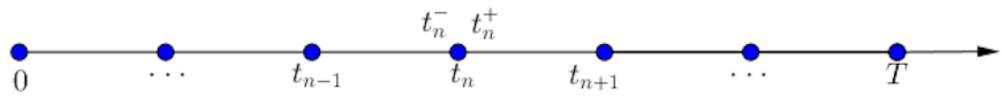

Figure 2.1: Partition of the time domain

where

$$v(t_n^{\pm}) = \lim_{\varepsilon \to 0^{\pm}} v(t_n + \varepsilon) \text{ for } n = 0, 1, \ldots, N-1,$$

for a function v. By convention, we assume that $v(0^-) = u_0$ and $\dot{v}(0^-) = u_1$. Moreover, we shall write v_n^+ for $v(t_n^+)$ and v_n^- for $v(t_n^-)$.

We first derive an energy identity for the exact solution u. On each time interval $I_n = (t_{n-1}, t_n]$, we assume that the exact solution $u \in L^2(I_n; H_0^1(\Omega))$, $\dot{u} \in L^2(I_n; H_0^1(\Omega))$, and $\ddot{u} \in L^2(I_n; H^{-1}(\Omega))$. Note that we assume $\dot{u} \in L^2(I_n; H_0^1(\Omega))$, which is a stronger assumption than that in Definition 2.1 because we need to ensure that $\nabla \dot{u}$ in the following equation is well defined. Taking $\varphi = \dot{u}(t)$ in Definition 2.1 (a) and integrating on I_n, we have

$$\int_{t_{n-1}}^{t_n} \langle \ddot{u}, \dot{u} \rangle \,\mathrm{d}t + \int_{t_{n-1}}^{t_n} (\dot{u}, \dot{u})_{L^2} \,\mathrm{d}t + \int_{t_{n-1}}^{t_n} (\nabla u, \nabla \dot{u})_{L^2} \,\mathrm{d}t = \int_{t_{n-1}}^{t_n} (f, \dot{u})_{L^2} \,\mathrm{d}t.$$

By taking $v = \dot{u}$ in Theorem 1.7 and $v = \nabla u$ with $H = V = L^2(\Omega)$ in Theorem 1.8, we obtain

$$\begin{aligned}
&\frac{1}{2}\left(\|\dot{u}(t_n^-)\|_{L^2(\Omega)}^2 - \|\dot{u}(t_{n-1}^+)\|_{L^2(\Omega)}^2\right) + \frac{1}{2}\left(\|\nabla u(t_n^-)\|_{L^2(\Omega)}^2 - \|\nabla u(t_{n-1}^+)\|_{L^2(\Omega)}^2\right) \\
&\quad + \int_{t_{n-1}}^{t_n} \|\dot{u}\|_{L^2(\Omega)}^2 \,\mathrm{d}t = \int_{t_{n-1}}^{t_n} (f, \dot{u})_{L^2} \,\mathrm{d}t,
\end{aligned}$$

for $n = 1, 2, \ldots, N$. Summing up over $n = 1, \ldots, N$, we have

$$\begin{aligned}
&\frac{1}{2}\left(\|\dot{u}(t_N^-)\|_{L^2(\Omega)}^2 - \|\dot{u}(t_0^+)\|_{L^2(\Omega)}^2\right) + \frac{1}{2}\left(\|\nabla u(t_N^-)\|_{L^2(\Omega)}^2 - \|\nabla u(t_0^+)\|_{L^2(\Omega)}^2\right) \\
&\quad - \frac{1}{2}\left\{\|\dot{u}(t_{N-1}^+)\|_{L^2(\Omega)}^2 - \|\dot{u}(t_{N-1}^-)\|_{L^2(\Omega)}^2 + \cdots + \|\dot{u}(t_1^+)\|_{L^2(\Omega)}^2 - \|\dot{u}(t_1^-)\|_{L^2(\Omega)}^2\right\}
\end{aligned}$$

$$
\begin{aligned}
&-\frac{1}{2}\left\{\|\nabla u(t_{N-1}^{+})\|_{L^2(\Omega)}^2-\|\nabla u(t_{N-1}^{-})\|_{L^2(\Omega)}^2+\cdots+\|\nabla u(t_1^{+})\|_{L^2(\Omega)}^2-\|\nabla u(t_1^{-})\|_{L^2(\Omega)}^2\right\}\\
&+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot u\|_{L^2(\Omega)}^2\,\mathrm{d}t=\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(f,\dot u)_{L^2}\,\mathrm{d}t. \qquad (2.2.1)
\end{aligned}
$$

The terms in the middle lines motivate us to add extra terms to achieve a well-posed variational form. From Remark 2.2, we know that $u \in C([0,T];H_0^1(\Omega))$ and $\dot u \in C([0,T];L^2(\Omega))$, that is, $[u]_n = 0$ and $[\dot u]_n = 0$ for all $n = 0,1,2,\dots,N-1$. Adding

$$
\sum_{n=0}^{N-1}([\dot u]_n,\dot u_n^{+})_{L^2}=\left(\dot u_{N-1}^{+}-\dot u_{N-1}^{-},\dot u_{N-1}^{+}\right)_{L^2}+\cdots+\left(\dot u_0^{+}-\dot u_0^{-},\dot u_0^{+}\right)_{L^2}
$$

and

$$
\sum_{n=0}^{N-1}(\nabla[u]_n,\nabla u_n^{+})_{L^2}=\left(\nabla u_{N-1}^{+}-\nabla u_{N-1}^{-},\nabla u_{N-1}^{+}\right)_{L^2}+\cdots+\left(\nabla u_0^{+}-\nabla u_0^{-},\nabla u_0^{+})\right)_{L^2}
$$

to the left side of the equality (2.2.1), we have

$$
\begin{aligned}
&\frac{1}{2}\|\dot u(t_N^{-})\|_{L^2(\Omega)}^2+\frac{1}{2}\|\dot u(t_0^{+})\|_{L^2(\Omega)}^2+\frac{1}{2}\sum_{n=1}^{N-1}\left(\|\dot u(t_n^{+})\|_{L^2(\Omega)}^2-2\left(\dot u(t_n^{+}),\dot u(t_n^{-})\right)_{L^2}+\|\dot u(t_n^{-})\|_{L^2(\Omega)}^2\right)\\
&\quad+\frac{1}{2}\|\nabla u(t_N^{-})\|_{L^2(\Omega)}^2+\frac{1}{2}\sum_{n=1}^{N-1}\left(\|\nabla u(t_n^{+})\|_{L^2(\Omega)}^2-2\left(\nabla u(t_n^{+}),\nabla u(t_n^{-})\right)_{L^2}+\|\nabla u(t_n^{-})\|_{L^2(\Omega)}^2\right)\\
&\quad+\frac{1}{2}\|\nabla u(t_0^{+})\|_{L^2(\Omega)}^2-\left(\nabla u(t_0^{-}),\nabla u(t_0^{+})\right)_{L^2}-\left(\dot u(t_0^{-}),\dot u(t_0^{+})\right)_{L^2}+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot u\|_{L^2(\Omega)}^2\,\mathrm{d}t\\
&=\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(f,\dot u)_{L^2}\mathrm{d}t.
\end{aligned}
$$

That is,

$$
\begin{aligned}
&\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot u\|_{L^2(\Omega)}^2\,\mathrm{d}t+\frac{1}{2}\|\dot u(t_0^{+})\|_{L^2(\Omega)}^2+\frac{1}{2}\sum_{n=1}^{N-1}\|[\dot u]_n\|_{L^2(\Omega)}^2+\frac{1}{2}\|\dot u(t_N^{-})\|_{L^2(\Omega)}^2\\
&\quad+\frac{1}{2}\|\nabla u(t_0^{+})\|_{L^2(\Omega)}^2+\frac{1}{2}\sum_{n=1}^{N-1}\|\nabla[u]_n\|_{L^2(\Omega)}^2+\frac{1}{2}\|\nabla u(t_N^{-})\|_{L^2(\Omega)}^2\\
&=\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(f,\dot u)_{L^2}\,\mathrm{d}t+\left(\dot u(t_0^{-}),\dot u(t_0^{+})\right)_{L^2}+\left(\nabla u(t_0^{-}),\nabla u(t_0^{+})\right)_{L^2}.
\end{aligned}
$$

This equality motivates us to define an energy norm for the variational formulation.

2.2.2. Discontinuous-in-time discretisation Now we follow a time integration approach to incrementally build an approximation of the exact solution u on each time interval I_n. We focus on the generic time interval I_n and assume that the solution on I_{n-1} is known. Testing the equation (2.0.1) against $\dot{v}$ where $v \in H^1(I_n; H_0^1(\Omega))$, we have

$$\int_{t_{n-1}}^{t_n} \langle \ddot{u}, \dot{v} \rangle \, \mathrm{d}t + \int_{t_{n-1}}^{t_n} (\dot{u}, \dot{v})_{L^2} \, \mathrm{d}t + \int_{t_{n-1}}^{t_n} (\nabla u, \nabla \dot{v})_{L^2} \, \mathrm{d}t = \int_{t_{n-1}}^{t_n} (f, \dot{v})_{L^2} \, \mathrm{d}t. \qquad (2.2.2)$$

We observe that $[u]_n = [\dot{u}]_n = 0$ for $n = 0, 1, \ldots, N-1$. Now rewrite (2.2.2) by adding suitable (strongly consistent) terms:

$$\begin{aligned} &\int_{t_{n-1}}^{t_n} \langle \ddot{u}, \dot{v} \rangle \, \mathrm{d}t + \int_{t_{n-1}}^{t_n} (\dot{u}, \dot{v})_{L^2} \, \mathrm{d}t + \int_{t_{n-1}}^{t_n} (\nabla u, \nabla \dot{v})_{L^2} \, \mathrm{d}t + ([\dot{u}]_{n-1}, \dot{v}_{n-1}^+)_{L^2} \\ &\quad + (\nabla [u]_{n-1}, \nabla v_{n-1}^+)_{L^2} = \int_{t_{n-1}}^{t_n} (f, \dot{v})_{L^2} \, \mathrm{d}t \text{ for } n = 1, \ldots, N. \end{aligned} \qquad (2.2.3)$$

Note that we define by convention that $u_0^- = u_0$ and $\dot{u}_0^- = u_1$. Summing over all time intervals in (2.2.3), we are able to define the bilinear form $\mathcal{A} \colon \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ with

$$\mathcal{H} := H^2(0, T; H^{-1}(\Omega)) \cap H^1(0, T; H_0^1(\Omega)) \cap L^2(0, T; H_0^1(\Omega))$$

by

$$\begin{aligned} \mathcal{A}(u, v) := &\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \langle \ddot{u}, \dot{v} \rangle \, \mathrm{d}t + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\dot{u}, \dot{v})_{L^2} \, \mathrm{d}t + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\nabla u, \nabla \dot{v})_{L^2} \, \mathrm{d}t \\ &+ \sum_{n=1}^{N-1} ([\dot{u}]_n, \dot{v}_n^+)_{L^2} + \sum_{n=1}^{N-1} (\nabla [u]_n, \nabla v_n^+)_{L^2} + (\dot{u}_0^+, \dot{v}_0^+)_{L^2} + (\nabla u_0^+, \nabla v_0^+)_{L^2} \end{aligned} \qquad (2.2.4)$$

for all $v \in \mathcal{H}$. The linear functional $F \colon \mathcal{H} \to \mathbb{R}$ is defined as

$$F(v) := \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (f, \dot{v})_{L^2} \, \mathrm{d}t + (u_1, \dot{v}_0^+)_{L^2} + (\nabla u_0, \nabla v_0^+)_{L^2}. \qquad (2.2.5)$$

Even though the bilinear form $\mathcal{A}(\cdot, \cdot)$ is defined for the continuous in time function space $\mathcal{H} \times \mathcal{H}$, it also works for piecewise in time version of $\mathcal{H} \times \mathcal{H}$. Next, we introduce the local semi-discrete space

$$\mathcal{V}^{q_n} := \{v \in L^2(0,T; H_0^1(\Omega))\colon v\,|_{I_n} \in \mathbb{P}^{q_n}(I_n; H_0^1(\Omega))\}$$

where $\mathbb{P}^{q_n}(I_n; H_0^1(\Omega))$ is the space of polynomials in t of degree less than or equal to $q_n \geq 2$ on I_n with coefficients in $H_0^1(\Omega)$. Then, introducing $\mathbf{q} := (q_1, q_2, \ldots, q_N) \in \mathbb{N}^N$ the N-dimensional polynomial degree vector, we can define the discontinuous Galerkin finite element space as

$$\mathcal{V}^{\mathbf{q}} := \{v \in L^2(0,T; H_0^1(\Omega))\colon v\,|_{I_n} \in \mathcal{V}^{q_n} \text{ for all } n = 1, 2, \ldots, N\}.$$

The discontinuous-in-time formulation of problem (2.0.1)–(2.0.3) reads as follows: find $u_{\mathrm{DG}} \in \mathcal{V}^{\mathbf{q}}$ such that

$$\mathcal{A}(u_{\mathrm{DG}}, v) = F(v), \text{ for all } v \in \mathcal{V}^{\mathbf{q}}. \tag{2.2.6}$$

2.2.3. Stability analysis We begin by defining an energy norm, which will be used in the stability and convergence analysis.

Proposition 2.4. *The function* $|||\cdot|||\colon \mathcal{V}^{\mathbf{q}} \to \mathbb{R}^+$ *defined by*

$$\begin{aligned} |||v|||^2 := {}& \frac{1}{2}\|\dot{v}_0^+\|_{L^2(\Omega)}^2 + \frac{1}{2}\sum_{n=1}^{N-1}\|[\dot{v}]_n\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\dot{v}_N^-\|_{L^2(\Omega)}^2 + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{v}\|_{L^2(\Omega)}^2\,\mathrm{d}t \\ &+ \frac{1}{2}\|\nabla v_0^+\|_{L^2(\Omega)}^2 + \frac{1}{2}\sum_{n=1}^{N-1}\|\nabla[v]_n\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\nabla v_N^-\|_{L^2(\Omega)}^2 \end{aligned} \tag{2.2.7}$$

is a norm on $\mathcal{V}^{\mathbf{q}}$.

Proof. The homogeneity and triangle inequality follow from the properties of the L^2 and H_0^1 norms. It is sufficient to show that $|||v||| = 0 \Leftrightarrow v \equiv 0$. If $v \equiv 0$, it is trivially true that $|||v||| = 0$. Now we need to show that $|||v||| = 0 \Rightarrow v \equiv 0$. If $|||v||| = 0$, then $\|v_0^+\|_{H_0^1(\Omega)} := \|\nabla v_0^+\|_{L^2(\Omega)} = 0$ and $\int_{t_0}^{t_1}\|\dot{v}\|_{L^2(\Omega)}^2\,\mathrm{d}t = 0$. We have

$$\begin{cases} \dot{v}(t) = 0 & \text{for } t \in I_1, \\ v_0^+ = 0. \end{cases}$$

This implies that $v(t) \equiv 0$ on I_1. We now proceed by induction. Assume that $v(t) \equiv 0$ on I_{n-1} and consider the time interval I_n. From $\|[v]_{n-1}\|_{H_0^1(\Omega)} := \|\nabla [v]_{n-1}\|_{L^2(\Omega)} = 0$, we know that $v_{n-1}^+ = v_{n-1}^- = 0$. Then

$$\begin{cases} \dot{v}(t) = 0 & \text{for } t \in I_n, \\ v_{n-1}^+ = 0. \end{cases}$$

Thus, we conclude that $v \equiv 0$ on each interval I_n for $n = 1, \ldots, N$. That is, $v \equiv 0$ on $(0, T]$. □

Remark 2.5. *If $L = 0$ in our model problem (2.0.1), the corresponding definition for $|||\cdot|||\colon \mathcal{V}^{\mathbf{q}} \to \mathbb{R}^+$ will be*

$$\begin{aligned} |||v|||^2 := \frac{1}{2}\|\dot{v}_0^+\|_{L^2(\Omega)}^2 + \frac{1}{2}\sum_{n=1}^{N-1}\|[\dot{v}]_n\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\dot{v}_N^-\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\nabla v_0^+\|_{L^2(\Omega)}^2 \\ + \frac{1}{2}\sum_{n=1}^{N-1}\|\nabla [v]_n\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\nabla v_N^-\|_{L^2(\Omega)}^2, \end{aligned} \tag{2.2.8}$$

which is not a well-defined norm.

Remark 2.6. *If we take $u = v$ in (2.2.4), we have*

$$\mathcal{A}(v, v) = |||v|||^2 \text{ for all } v \in \mathcal{V}^{\mathbf{q}}. \tag{2.2.9}$$

This shows that the bilinear form $\mathcal{A}(\cdot,\cdot)$ is coercive with respect to the $|||\cdot|||$ norm with coercivity constant $\alpha = 1$.

Theorem 2.7. *Let $f \in L^2(0, T; L^2(\Omega))$, $u_0 \in H_0^1(\Omega)$ and $u_1 \in L^2(\Omega)$. Then the solution of (2.2.6), $u_{\mathrm{DG}} \in \mathcal{V}^{\mathbf{q}}$ satisfies*

$$|||u_{\mathrm{DG}}||| \le \left(\|f\|_{L^2(0,T;L^2(\Omega))}^2 + 2\|\nabla u_0\|_{L^2(\Omega)}^2 + 2\|u_1\|_{L^2(\Omega)}^2\right)^{\frac{1}{2}}.$$

Proof. Using the definition of $|||\cdot|||$ in (2.2.7) and Young's inequality, we have

$$|||u_{\mathrm{DG}}|||^2 = \mathcal{A}(u_{\mathrm{DG}}, u_{\mathrm{DG}}) = F(u_{\mathrm{DG}})$$

$$
\begin{aligned}
&= \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (f, \dot{u}_{\mathrm{DG}})_{L^2}\, \mathrm{d}t + (u_1, \dot{u}_{DG}(\cdot, 0^+))_{L^2} + (\nabla u_0, \nabla u_{\mathrm{DG}}(\cdot, 0^+))_{L^2} \\
&\leq \frac{1}{2} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|f(\cdot, t)\|_{L^2(\Omega)}^2\, \mathrm{d}t + \frac{1}{2} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|\dot{u}_{\mathrm{DG}}(\cdot, t)\|_{L^2(\Omega)}^2\, \mathrm{d}t \\
&\quad + \|u_1\|_{L^2(\Omega)}^2 + \frac{1}{4} \|\dot{u}_{\mathrm{DG}}(\cdot, 0^+)\|_{L^2(\Omega)}^2 + \|\nabla u_0\|_{L^2(\Omega)}^2 + \frac{1}{4} \|\nabla u_{\mathrm{DG}}(\cdot, 0^+)\|_{L^2(\Omega)}^2 \\
&\leq \frac{1}{2} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|f(\cdot, t)\|_{L^2(\Omega)}^2\, \mathrm{d}t + \frac{1}{2} |||u_{\mathrm{DG}}|||^2 + \|u_1\|_{L^2(\Omega)}^2 + \|\nabla u_0\|_{L^2(\Omega)}^2.
\end{aligned}
$$

Then the required inequality follows. □

2.2.4. Convergence analysis In this section, we prove *a priori* error estimates with respect to the energy norm $|||\cdot|||$ defined in (2.2.7).

Before we proceed to the proof of the main convergence result, let us first define the L^2-projection operator based on Definition 3.1 in [69] (Schötzau & Schwab, 2000), Definition 3.2 in [6] (Antonietti *et al.*, 2018) and Thomée's book [75].

Definition 2.8. *Let $I = (-1, 1)$. For a function $w \in C((-1, 1]; H_0^1(\Omega)) \cap L^2(I; H_0^1(\Omega))$, we define the (boundary value preserving) L^2-projection $\mathcal{P}^q w \in \mathbb{P}^q(I; H_0^1(\Omega))$ with $q \geq 0$ by the following conditions*

$$
\mathcal{P}^q w(x, 1) = w(x, 1) \in H_0^1(\Omega), \tag{2.2.10}
$$

$$
\int_I ((w - \mathcal{P}^q w), \varphi)_{L^2}\, \mathrm{d}t = 0 \textit{ for all } \varphi \in \mathbb{P}^{q-1}(I; H_0^1(\Omega)). \tag{2.2.11}
$$

When $q = 0$, only the first condition is necessary.

Here 'boundary value preserving' means that the interpolant evaluated at the right end point is equal to the original function evaluated at the right end point, which is stated as (2.2.10) in Definition 2.8.

Definition 2.9. *Let $I = (-1, 1)$. For a function $u \in H^1(I; H_0^1(\Omega))$ such that $\dot{u} \in C((-1, 1]; H_0^1(\Omega))$, we define $\Pi^q u \in \mathbb{P}^q(I; H_0^1(\Omega))$ with $q \geq 1$ by*

$$
\Pi^q u(x, t) = u(x, -1) + \int_{-1}^{t} \mathcal{P}^{q-1} \dot{u}(x, s)\, \mathrm{d}s \textit{ for all } t \in (-1, 1],
$$

where $\mathcal{P}^{q-1}\dot{u}$ is given in Definition 2.8.

Now we need to show that $\Pi^q u(x,t)$ in Definition 2.9 is well-defined. Consider the sequence of Legendre polynomials $\{L_i\}_{i\geq 0}$, $L_i \in \mathbb{P}^i(I)$ on $I = (-1,1)$ defined by the following recurrence relation

$$(i+1)L_{i+1}(t) = (2i+1)tL_i(t) - iL_{i-1}(t)$$

with $L_0(t) = 1$ and $L_1(t) = t$.

The relevant properties of Legendre polynomials for our purpose are the following:

(a) $L_i(1) = 1$ for $i \in \mathbb{N} \cup \{0\}$;

(b) $\int_{-1}^t L_i(s)\,\mathrm{d}s = \frac{1}{2i+1}(L_{i+1}(t) - L_{i-1}(t))$ for $i \in \mathbb{N}$;

(c) $\int_{-1}^1 L_i(t)^2\,\mathrm{d}t = \frac{2}{2i+1}$ for $i \in \mathbb{N} \cup \{0\}$.

Lemma 2.10. $\Pi^q u(x,t)$ *defined in Definition 2.9 is well-defined.*

Proof. Existence: Note that $\dot{u} \in L^2(I; H_0^1(\Omega))$, so we can write $\dot{u}$ as a Legendre series in the following form

$$\dot{u}(x,t) = \sum_{i=0}^{\infty} b_i(x) L_i(t)$$

with $b_i \in H_0^1(\Omega)$ for each $i \in \mathbb{N} \cup \{0\}$. From Lemma 3.5 in [69], we find that

$$\mathcal{P}^{q-1}\dot{u}(x,t) = \sum_{i=0}^{q-2} b_i(x)L_i(t) + \left(\sum_{i=q-1}^{\infty} b_i(x)\right) L_{q-1}(t),$$

and consequently,

$$\begin{aligned}
\Pi^q u(x,t) &= u(x,-1) + \sum_{i=0}^{q-2} b_i(x) \int_{-1}^t L_i(s)\,\mathrm{d}s + \left(\sum_{i=q-1}^{\infty} b_i(x)\right) \int_{-1}^t L_{q-1}(s)\,\mathrm{d}s \\
&= u(x,-1) + b_0(x)(t+1) + \sum_{i=1}^{q-2} b_i(x) \frac{1}{2i+1}(L_{i+1}(t) - L_{i-1}(t)) \\
&\quad + \left(\sum_{i=q-1}^{\infty} b_i(x)\right) \left(\frac{1}{2q-1}(L_q(t) - L_{q-2}(t))\right)
\end{aligned}$$

$$
\begin{aligned}
&= \left[u(x,-1) + b_0(x) - \frac{b_1(x)}{3}\right] L_0(t) + \left[b_0(x) - \frac{b_2(x)}{5}\right] L_1(t) \\
&\quad + \sum_{i=2}^{q-3} \left[\frac{b_{i-1}(x)}{2i-1} - \frac{b_{i+1}(x)}{2i+3}\right] L_i(t) + \left[\frac{b_{q-3}(x)}{2q-5} - \sum_{i=q-1}^{\infty} \frac{b_i(x)}{2q-1}\right] L_{q-2}(t) \\
&\quad + \frac{b_{q-2}(x)}{2q-3} L_{q-1}(t) + \sum_{i=q-1}^{\infty} \frac{b_i(x)}{2q-1} L_q(t),
\end{aligned}
$$

or equivalently

$$
\Pi^q u(x,t) = \sum_{i=0}^{q} u_i^*(x) L_i(t), \tag{2.2.12}
$$

where

$$
\begin{aligned}
u_0^*(x) &= u(x,-1) + b_0(x) - \frac{b_1(x)}{3}, \\
u_1^*(x) &= b_0(x) - \frac{b_2(x)}{5}, \\
u_i^*(x) &= \frac{b_{i-1}(x)}{2i-1} - \frac{b_{i+1}(x)}{2i+3} \text{ for } i = 2, \dots, q-3, \\
u_{q-2}^*(x) &= \frac{b_{q-3}(x)}{2q-5} - \sum_{i=q-1}^{\infty} \frac{b_i(x)}{2q-1}, \\
u_{q-1}^*(x) &= \frac{b_{q-2}(x)}{2q-3}, \\
u_q^*(x) &= \sum_{i=q-1}^{\infty} \frac{b_i(x)}{2q-1}.
\end{aligned}
$$

Uniqueness: Assume for contradiction that both $u_1, u_2 \in \mathbb{P}^q(I, H_0^1(\Omega))$ satisfy the conditions in the definition. In particular, we have $u_1(x,1) = u_2(x,1)$. Now considering the difference $u_1 - u_2$, we can write it as a Legendre series $(u_1 - u_2)(x,t) = \sum_{i=0}^{q} a_i(x) L_i(t)$ with $a_i = \int_I (u_1 - u_2) L_i \, \mathrm{d}t \in H_0^1(\Omega)$. It follows from the orthogonality condition that

$$
\int_I (u_1 - u_2, a L_k)_{L^2} \, \mathrm{d}t = 0 \text{ for all } a \in H_0^1(\Omega), \quad 0 \leq k \leq q-1.
$$

Using the orthogonality properties of the Legendre polynomials we get $(a_k, a)_{L^2} = 0$ for all $a \in H_0^1(\Omega)$. Since $H_0^1(\Omega)$ is dense in $L^2(\Omega)$, we have that $a_k = 0$ in $L^2(\Omega)$ and thus $a_k = 0$ in $H_0^1(\Omega)$ for $0 \leq k \leq q-1$. This implies that $(u_1 - u_2)(x,t) = a_q(x) L_q(t)$. Since $u_1(x,1) = u_2(x,1)$, we have that $a_q \equiv 0$, which proves the uniqueness of a projection polynomial satisfying the conditions in Definition 2.9. $\square$

Lemma 2.11. *For any $u \in H^1(I; H_0^1(\Omega))$ such that $\dot{u} \in C((-1,1]; H_0^1(\Omega))$, we have for $q \geq 2$,*

(a) $(\Pi^q u - u)(x,-1) = 0$;

(b) $(\Pi^q u - u)(x,1) = 0$;

(c) $\partial_t(\Pi^q u - u)(x,1) = 0$;

(d) $\int_I \left(\partial_t \left(u - \Pi^q u\right), \varphi\right)_{L^2} \, \mathrm{d}t = 0$ *for all* $\varphi \in \mathbb{P}^{q-2}(I; H_0^1(\Omega))$.

Proof. (a) follows directly from Definition 2.9. For (b), first note that we can expand $\dot{u}$ as

$$\dot{u}(x,t) = \sum_{i=0}^{\infty} b_i(x) L_i(t)$$

with coefficients $b_i \in H_0^1(\Omega)$. Then u can be written as

$$u(x,t) = u(x,-1) + \sum_{i=0}^{\infty} b_i(x) \int_{-1}^{t} L_i(s) \, \mathrm{d}s. \tag{2.2.13}$$

By orthogonality of the Legendre polynomials, we have

$$u(x,1) = u(x,-1) + \sum_{i=0}^{\infty} b_i(x) \int_{-1}^{1} L_i(s) \, \mathrm{d}s = u(x,-1) + 2b_0. \tag{2.2.14}$$

Then,

$$\begin{aligned}
\Pi^q u(x,1) &= u(x,-1) + \int_{-1}^{1} \mathcal{P}^{q-1} \dot{u}(x,s) \, \mathrm{d}s \\
&= u(x,-1) + \sum_{i=0}^{q-2} b_i(x) \int_{-1}^{1} L_i(s) \, \mathrm{d}s + \left(\sum_{i=q-1}^{\infty} b_i(x) \right) \int_{-1}^{1} L_{q-1}(s) \, \mathrm{d}s \\
&= u(x,-1) + 2b_0 \\
&= u(x,1) \text{ (by (2.2.14))}.
\end{aligned}$$

The equality (c) follows from taking the derivative with respect to t in the equation

$$\Pi^q u(x,t) = u(x,-1) + \sum_{i=0}^{q-2} b_i(x) \int_{-1}^{t} L_i(s)\,\mathrm{d}s + \left(\sum_{i=q-1}^{\infty} b_i(x)\right) \int_{-1}^{t} L_{q-1}(s)\,\mathrm{d}s.$$

Note that

$$\dot{u}(x,t) = \sum_{i=0}^{\infty} b_i(x) L_i(t)$$

and

$$\partial_t \Pi^q u(x,t) = \sum_{i=0}^{q-2} b_i(x) L_i(t) + \left(\sum_{i=q-1}^{\infty} b_i(x)\right) L_{q-1}(t).$$

Evaluating at $t=1$, we have

$$\partial_t \Pi^q u(x,1) = \sum_{i=0}^{\infty} b_i(x) = \dot{u}(x,1).$$

For (d), we can write any $q \in \mathbb{P}^{q-2}(I; H_0^1(\Omega))$ as $q(x,t) = \sum_{k=0}^{q-2} a_k(x) L_k(t)$. Then

$$\begin{aligned}
&\int_I (\partial_t(u - \Pi^q u), \varphi)_{L^2}\ \mathrm{d}t = \int_I \sum_{k=0}^{q-2} (\partial_t(u - \Pi^q u), a_k L_k)_{L^2}\ \mathrm{d}t \\
&= \sum_{k=0}^{q-2} \int_I \left(\partial_t u - \mathcal{P}^{q-1}\partial_t u), a_k L_k\right)_{L^2}\ \mathrm{d}t \\
&= \sum_{k=0}^{q-2} \int_I \left(\left(\sum_{i=q}^{\infty} b_i(x) L_i(t) - \left(\sum_{i=q}^{\infty} b_i(x)\right) L_{q-1}(t)\right), a_k L_k(t)\right)_{L^2}\ \mathrm{d}t = 0
\end{aligned}$$

by the orthogonality of Legendre polynomials. □

On an arbitrary time interval $I = (a,b)$, we define Π_I^q via the linear map

$$T \colon \Omega \times (-1,1) \to \Omega \times (a,b)$$

$$(x,\xi) \mapsto (x, \frac{1}{2}(a+b+\xi(b-a)))$$

as

$$\Pi_I^q u = [\Pi^q(u \circ T)] \circ T^{-1}, \tag{2.2.15}$$

from which it follows that

(a) $(\Pi_I^q u - u)(x,a) = 0$;

(b) $(\Pi_I^q u - u)(x,b) = 0$;

(c) $\partial_t(\Pi_I^q u - u)(x,b) = 0$;

(d) $\int_I \left(\partial_t(u - \Pi_I^q u), \varphi\right)_{L^2} \,\mathrm{d}t = 0$ for all $\varphi \in \mathbb{P}^{q-2}(I; H_0^1(\Omega))$.

Similarly as before, on the generic interval $I = (a,b)$, $\Pi_I^q u$ can be written as

$$\Pi_I^q u(x,t) = \sum_{i=0}^{q} \tilde{u}_i^* \tilde{L}_i(t) = u(x,a) + \sum_{i=0}^{q-2} \tilde{b}_i(x) \int_a^t \tilde{L}_i(s)\,\mathrm{d}s + \left(\sum_{i=q-1}^{\infty} \tilde{b}_i\right) \int_a^t \tilde{L}_{q-1}(s)\,\mathrm{d}s,$$

where $\tilde{L}_i$ represents the ith mapped Legendre polynomial so that $u = \sum_{i=0}^{\infty} \tilde{u}_i \tilde{L}_i$ and the coefficients are defined as before but with b_i replaced by $\tilde{b}_i$ for each $i \in \mathbb{N} \cup \{0\}$.

Now we generalise a standard approximation result stated in [11] by Babuška and Suri to functions in Bochner spaces.

Proposition 2.12. *Let $I = (a,b)$ with $k = b - a > 0$. For every $u \in H^s(I; H_0^1(\Omega))$, there exists a sequence $\{\mathcal{Q}^q u\}_{q \geq 0}$ with $\mathcal{Q}^q u \in \mathbb{P}^q(I; H_0^1(\Omega))$ such that, for any $0 \leq m \leq s$,*

$$\|u - \mathcal{Q}^q u\|_{H^m(I;H_0^1(\Omega))} \leq C \frac{k^{\mu - m}}{q^{s-m}} \|u\|_{H^s(I;H_0^1(\Omega))} \text{ for } s \geq 0, \tag{2.2.16}$$

where $\mu = \min(q+1, s)$ and C is a constant independent of u, q and k.

Next, we define the projection operator $\mathcal{P}_I^q u$ on an arbitrary interval $I = (a,b)$ via the same linear map $T\colon \Omega \times (-1,1) \to \Omega \times (a,b)$ as before. That is,

$$\mathcal{P}_I^q u = [\mathcal{P}^q(u \circ T)] \circ T^{-1}.$$

Then the following approximation result holds.

Proposition 2.13. *Let $I = (a,b)$ with $k = b - a > 0$. For every $u \in H^s(I; H_0^1(\Omega))$, $s \geq 2$, we have*

$$\|\partial_t u - \mathcal{P}_I^{q-1}(\partial_t u)\|_{L^2(I;H_0^1(\Omega))} \leq C \frac{k^{\mu-1}}{q^{s-1}} \|u\|_{H^s(I;H_0^1(\Omega))}, \tag{2.2.17}$$

where $\mu = \min(q+1, s)$ *and* C *is a universal constant.*

Proof. We give a sketch of the proof based on several results from [69] and [70]. For $u \in H^s(I; H_0^1(\Omega))$, define $\hat{u}(x,\xi)$ for $\xi \in \hat{I} := (-1,1)$ and $t = \frac{1}{2}(a+b+\xi k)$ by

$$u(x,t) = u(x, \frac{1}{2}(a+b+\xi k)) := \hat{u}(x,\xi).$$

Then we have $\partial_\xi \hat{u}(x,\xi) = \frac{k}{2}\partial_t u(x,t)$. Similarly to the existence proof of Lemma 2.10, we have

$$\partial_\xi \hat{u}(x,\xi) = \sum_{i=0}^{\infty} b_i(x) L_i(\xi)$$

and

$$\mathcal{P}^{q-1}\partial_\xi \hat{u}(x,\xi) = \sum_{i=0}^{q-2} b_i(x) L_i(\xi) + \left(\sum_{i=q-1}^{\infty} b_i(x)\right) L_{q-1}(\xi)$$

where $\{L_i\}_{i\geq 0}$ are Legendre polynomials defined on $\hat{I} := (-1,1)$. This implies that

$$\partial_\xi \hat{u}(x,\xi) - \mathcal{P}^{q-1}\partial_\xi \hat{u}(x,\xi) = \sum_{i=q}^{\infty} b_i(x) L_i(\xi) - \left(\sum_{i=q}^{\infty} b_i(x)\right) L_{q-1}(\xi).$$

Using the triangle inequality, we obtain

$$\|\partial_\xi \hat{u} - \mathcal{P}^{q-1}\partial_\xi \hat{u}\|_{L^2(\hat{I};H_0^1(\Omega))} \leq \|\sum_{i=q}^{\infty} b_i(x) L_i(\xi)\|_{L^2(\hat{I};H_0^1(\Omega))} + \|\left(\sum_{i=q}^{\infty} b_i(x)\right) L_{q-1}(\xi)\|_{L^2(\hat{I};H_0^1(\Omega))}.$$

Note that by Lemma 3.9 from [70], we know that

$$\|\partial_\xi \hat{u} - \mathbf{P}^{q-1}\partial_\xi \hat{u}\|_{L^2(\hat{I};H_0^1(\Omega))} = \|\sum_{i=q}^{\infty} b_i(x) L_i(\xi)\|_{L^2(\hat{I};H_0^1(\Omega))},$$

where $\mathbf{P}^{q-1}$ is the standard L^2-projection operator such that

$$\|\partial_\xi \hat{u} - \mathbf{P}^{q-1}\partial_\xi \hat{u}\|_{L^2(\hat{I};H_0^1(\Omega))} = \inf_{\varphi \in \mathbb{P}^{q-1}(\hat{I};H_0^1(\Omega))} \|\partial_\xi \hat{u} - \varphi\|_{L^2(\hat{I};H_0^1(\Omega))}.$$

Using the orthogonality of Legendre polynomials on $\hat{I}$, we have

$$\|\sum_{i=q}^{\infty} b_i(x) L_i(\xi)\|^2_{L^2(\hat{I};H_0^1(\Omega))} = \sum_{i=q}^{\infty} \|b_i\|^2_{H_0^1(\Omega)} \frac{2}{2i+2}.$$

By Lemma 3.10 from [70], we know that for $1 \leq s_0 \leq s$ we have

$$\int_{-1}^{1} |\hat{u}^{(s_0)}(x,\xi)|^2 (1-\xi^2)^{s_0-1}\,\mathrm{d}\xi = \sum_{i=s_0-1}^{\infty} |b_i(x)|^2 \frac{2}{2i+1} \frac{(i+(s_0-1))!}{(i-(s_0-1))!} \tag{2.2.18}$$

where $\hat{u}^{(s_0)}(x,\xi)$ represents the s_0-th partial derivative of $\hat{u}(x,\xi)$ with respect to ξ. Thus

$$\begin{aligned}\|\sum_{i=q}^{\infty} b_i(x)L_i(\xi)\|^2_{L^2(\hat{I};H_0^1(\Omega))} &\leq \frac{(q-(s_0-1))!}{(q+(s_0-1))!} \sum_{i=q}^{\infty} \|b_i\|^2_{H_0^1(\Omega)} \frac{2}{2i+2} \frac{(i+(s_0-1))!}{(i-(s_0-1))!} \\ &\leq \frac{(q-(s_0-1))!}{(q+(s_0-1))!} |\hat{u}|^2_{H^{s_0}(\hat{I};H_0^1(\Omega))} \text{ (by (2.2.18))}\end{aligned}$$

for any $1 \leq s_0 \leq \min(q+1, s)$. Thus, we have

$$\inf_{\varphi\in\mathbb{P}^{q-1}(\hat{I};H_0^1(\Omega))} \|\partial_\xi \hat{u} - \varphi\|^2_{L^2(\hat{I};H_0^1(\Omega))} = \|\sum_{i=q}^{\infty} b_i(x)L_i(\xi)\|^2_{L^2(\hat{I};H_0^1(\Omega))} \leq \frac{(q-(\mu-1))!}{(q+(\mu-1))!} |\hat{u}|^2_{H^{\mu}(\hat{I};H_0^1(\Omega))}$$

where $\mu = \min(q+1, s)$. Note also that

$$\|\left(\sum_{i=q}^{\infty} b_i(x)\right) L_{q-1}(\xi)\|^2_{L^2(\hat{I};H_0^1(\Omega))} = \frac{2}{2q-1} \|\sum_{i=q}^{\infty} b_i(x)\|^2_{H_0^1(\Omega)} \leq \frac{4}{q} \|\sum_{i=q}^{\infty} b_i(x)\|^2_{H_0^1(\Omega)}$$

for $q \geq 2$. From Lemma 3.6 in [69], we know that

$$\|\sum_{i=q}^{\infty} b_i(x)\|^2_{H_0^1(\Omega)} \leq \frac{C_2}{q} \|\partial_{\xi\xi}\hat{u}\|^2_{L^2(\hat{I};H_0^1(\Omega))},$$

where $\partial_\xi \hat{u}(x,\xi) = \sum_{i=0}^{\infty} b_i(x)L_i(\xi)$ and C_2 is a positive constant. Combining with the previous estimate, we have

$$\|\partial_\xi \hat{u} - \mathcal{P}^{q-1}\partial_\xi \hat{u}\|^2_{L^2(\hat{I};H_0^1(\Omega))} \leq C \left\{ \|\partial_\xi \hat{u} - \mathbf{P}^{q-1}\partial_\xi \hat{u}\|^2_{L^2(\hat{I};H_0^1(\Omega))} + \frac{1}{q^2} \|\partial_{\xi\xi}\hat{u}\|^2_{L^2(\hat{I};H_0^1(\Omega))} \right\} \tag{2.2.19}$$

for some constant $C > 0$. If we replace $\partial_\xi \hat{u}$ by $\partial_\xi \hat{u} - \varphi$ in (2.2.19) for an arbitrary $\varphi \in \mathbb{P}^{q-1}(\hat{I}; H_0^1(\Omega))$, then we have

$$\begin{aligned}&\|\partial_\xi \hat{u} - \mathcal{P}^{q-1}\partial_\xi \hat{u}\|^2_{L^2(\hat{I};H_0^1(\Omega))} \\ &\leq C \left\{ \inf_{\varphi\in\mathbb{P}^{q-1}(\hat{I};H_0^1(\Omega))} \|\partial_\xi \hat{u} - \varphi\|^2_{L^2(\hat{I};H_0^1(\Omega))} + \frac{1}{q^2} \|\partial_{\xi\xi}\hat{u} - \partial_\xi\varphi\|^2_{L^2(\hat{I};H_0^1(\Omega))} \right\}\end{aligned} \tag{2.2.20}$$

using the fact that

$$\|\partial_\xi \hat{u} - \mathbf{P}^{q-1}\partial_\xi \hat{u}\|_{L^2(\hat{I};H_0^1(\Omega))} \leq \|\partial_\xi \hat{u} - \varphi\|_{L^2(I;H_0^1(\Omega))} \text{ for any } \varphi \in \mathbb{P}^{q-1}(I;H_0^1(\Omega)).$$

Similarly, we can show that

$$\|\partial_{\xi\xi}\hat{u} - \partial_\xi \varphi\|^2_{L^2(\hat{I};H_0^1(\Omega))} \leq \frac{(q-(s_0-1))!}{(q+(s_0-1))!}|\partial_\xi \hat{u}|^2_{H^{s_0}(\hat{I};H_0^1(\Omega))} = \frac{(q-(s_0-1))!}{(q+(s_0-1))!}|\hat{u}|^2_{H^{s_0+1}(\hat{I};H_0^1(\Omega))}$$

for $1 \leq s_0 \leq \min(q+1, s-1)$. Thus,

$$\|\partial_{\xi\xi}\hat{u} - \partial_\xi \varphi\|^2_{L^2(\hat{I};H_0^1(\Omega))} \leq \frac{(q-(\mu-2))!}{(q+(\mu-2))!}|\hat{u}|^2_{H^{\mu}(\hat{I};H_0^1(\Omega))}$$

for $\mu = \min(q+2, s)$. Therefore,

$$\begin{aligned}
&\|\partial_\xi \hat{u} - \mathcal{P}^{q-1}\partial_\xi \hat{u}\|^2_{L^2(\hat{I};H_0^1(\Omega))} \\
&\leq C\left\{\frac{(q-(\mu-1))!}{(q+(\mu-1))!}|\hat{u}|^2_{H^{\mu}(\hat{I};H_0^1(\Omega))} + \frac{1}{q^2}\frac{(q-(\mu-2))!}{(q+(\mu-2))!}|\hat{u}|^2_{H^{\mu}(\hat{I};H_0^1(\Omega))}\right\} \\
&\leq \tilde{C}\frac{1}{q^2}\frac{(q-(\mu-2))!}{(q+(\mu-2))!}|\hat{u}|^2_{H^{\mu}(\hat{I};H_0^1(\Omega))}
\end{aligned}$$

where $\mu = \min(q+1, s)$. Scaling back to $I = (a,b)$, we have

$$\begin{aligned}
\|\partial_t u - \mathcal{P}_I^{q-1}\partial_t u\|^2_{L^2(I;H_0^1(\Omega))} &= \int_a^b \|\partial_t u(\cdot,t) - \mathcal{P}_I^{q-1}\partial_t u(\cdot,t)\|^2_{H_0^1(\Omega)}\,\mathrm{d}t \\
&= \left(\frac{k}{2}\right)^{-1}\int_{-1}^{1}\|\partial_\xi \hat{u}(\cdot,\xi) - \mathcal{P}^{q-1}\partial_\xi u(\cdot,\xi)\|^2_{H_0^1(\Omega)}\,\mathrm{d}\xi \\
&= \left(\frac{k}{2}\right)^{-1}\|\partial_\xi \hat{u} - \mathcal{P}^{q-1}\partial_\xi \hat{u}\|^2_{L^2(\hat{I};H_0^1(\Omega))} \\
&\leq \tilde{C}\frac{1}{q^2}\frac{(q-(\mu-2))!}{(q+(\mu-2))!}\left(\frac{k}{2}\right)^{-1}|\hat{u}|^2_{H^{\mu}(\hat{I};H_0^1(\Omega))} \\
&= \tilde{C}\frac{1}{q^2}\frac{(q-(\mu-2))!}{(q+(\mu-2))!}\left(\frac{k}{2}\right)^{-1}\int_{-1}^{1}\|\hat{u}^{(\mu)}(\cdot,\xi)\|^2_{H_0^1(\Omega)}\,\mathrm{d}\xi \\
&= \tilde{C}\frac{1}{q^2}\frac{(q-(\mu-2))!}{(q+(\mu-2))!}\left(\frac{k}{2}\right)^{-1}\int_a^b\left(\frac{k}{2}\right)^{2\mu-1}\|u^{(\mu)}(\cdot,t)\|^2_{H_0^1(\Omega)}\,\mathrm{d}t \\
&= \tilde{C}\frac{1}{q^2}\frac{(q-(\mu-2))!}{(q+(\mu-2))!}\left(\frac{k}{2}\right)^{2\mu-2}\int_a^b\|u^{(\mu)}(\cdot,t)\|^2_{H_0^1(\Omega)}\,\mathrm{d}t \\
&= \tilde{C}\frac{1}{q^2}\frac{(q-(\mu-2))!}{(q+(\mu-2))!}\left(\frac{k}{2}\right)^{2\mu-2}|u|^2_{H^{\mu}(I;H_0^1(\Omega))}
\end{aligned}$$

where $\mu = \min(q+1, s)$. Selecting $\mu = s$, we have

$$\begin{aligned}\|\partial_t u - \mathcal{P}_I^{q-1}\partial_t u\|^2_{L^2(I;H_0^1(\Omega))} &\le \tilde{C}\frac{1}{q^2}\frac{(q-(s-2))!}{(q+(s-2))!}\left(\frac{k}{2}\right)^{2(s-1)}|u|^2_{H^s(I;H_0^1(\Omega))}\\ &\le C\frac{k^{2(s-1)}}{q^{2(s-1)}}|u|^2_{H^s(I;H_0^1(\Omega))}\end{aligned}$$

where the second inequality follows from Stirling's formula as $q \to \infty$. Therefore,

$$\|\partial_t u - \mathcal{P}_I^{q-1}\partial_t u\|_{L^2(I;H_0^1(\Omega))} \le C\frac{k^{\mu-1}}{q^{s-1}}\|u\|_{H^s(I;H_0^1(\Omega))}$$

for $\mu = \min(q+1, s)$ since the semi-norm $|\cdot|_{H^s(I;H_0^1(\Omega))}$ is bounded by $\|\cdot\|_{H^s(I;H_0^1(\Omega))}$. □

As a result, the following estimates also hold.

Lemma 2.14. *Let $I = (a,b)$ with $k = b-a > 0$. For every $u \in H^s(I;H_0^1(\Omega))$, $s \ge 2$, we have*

$$\|\partial_t(u - \Pi_I^q u)\|_{L^2(I;H_0^1(\Omega))} \le C\frac{k^{\mu-1}}{q^{s-1}}\|u\|_{H^s(I;H_0^1(\Omega))}, \tag{2.2.21}$$

$$\|u - \Pi_I^q u\|_{L^2(I;H_0^1(\Omega))} \le C\frac{k^{\mu}}{q^{s-1}}\|u\|_{H^s(I;H_0^1(\Omega))}, \tag{2.2.22}$$

where $\mu = \min(q+1, s)$, Π_I^q is the projection operator as defined in (2.2.15), and C is a universal constant.

Proof. To prove (2.2.21), we first note that $\partial_t(\Pi_I^q u) = \mathcal{P}_I^{q-1}\partial_t u$ by definition.

$$\begin{aligned}\|\partial_t(u - \Pi_I^q u)\|_{L^2(I;H_0^1(\Omega))} &= \|\partial_t u - \mathcal{P}_I^{q-1}\partial_t u\|_{L^2(I;H_0^1(\Omega))}\\ &\le C\frac{k^{\mu-1}}{q^{s-1}}\|u\|_{H^s(I;H_0^1(\Omega))} \text{ (by Proposition 2.13).}\end{aligned}$$

For (b), we have

$$\begin{aligned}\|u - \Pi_I^q u\|^2_{L^2(I;H_0^1(\Omega))} &= \left\|u(x,a) + \int_a^t \partial_s u(x,s)\,\mathrm{d}s - u(x,a) - \int_a^t \mathcal{P}_I^{q-1}\partial_s u(x,s)\,\mathrm{d}s\right\|^2_{L^2(I;H_0^1(\Omega))}\\ &= \left\|\int_a^t \partial_s u(x,s)\,\mathrm{d}s - \int_a^t \mathcal{P}_I^{q-1}\partial_s u(x,s)\,\mathrm{d}s\right\|^2_{L^2(I;H_0^1(\Omega))}\\ &= \int_a^b \left\|\int_a^t \partial_s u(\cdot,s) - \mathcal{P}_I^{q-1}\partial_s u(\cdot,s)\,\mathrm{d}s\right\|^2_{H_0^1(\Omega)}\mathrm{d}t\end{aligned}$$

$$
\begin{aligned}
&\le \int_a^b \left\| \left(\int_a^t 1\,\mathrm{d}s\right)^{\frac{1}{2}} \left(\int_a^t |\partial_s u(\cdot,s) - \mathcal{P}_I^{q-1}\partial_s u(\cdot,s)|^2\,\mathrm{d}s\right)^{\frac{1}{2}} \right\|_{H_0^1(\Omega)}^2 \mathrm{d}t \\
&\le \int_a^b k \left\| \left(\int_a^t |\partial_s u(\cdot,s) - \mathcal{P}_I^{q-1}\partial_s u(\cdot,s)|^2\,\mathrm{d}s\right)^{\frac{1}{2}} \right\|_{H_0^1(\Omega)}^2 \mathrm{d}t \\
&\le k^2 \left\| \left(\int_a^b |\partial_s u(\cdot,s) - \mathcal{P}_I^{q-1}\partial_s u(\cdot,s)|^2\,\mathrm{d}s\right)^{\frac{1}{2}} \right\|_{H_0^1(\Omega)}^2 \\
&= k^2 \|\partial_s u - \mathcal{P}_I^{q-1}\partial_s u\|_{L^2(I;H_0^1(\Omega))}^2 \\
&\le C^2 k^2 \frac{k^{2\mu-2}}{q^{2s-2}} \|u\|_{H^s(I;H_0^1(\Omega))}^2.
\end{aligned}
$$

Thus

$$
\|u - \Pi_I^q u\|_{L^2(I;H_0^1(\Omega))} \le C \frac{k^\mu}{q^{s-1}} \|u\|_{H^s(I;H_0^1(\Omega))}.
$$

□

We can also derive the following result using Lemma 2.14.

Corollary 2.15. *Let $I = (a,b)$ with $k = b - a$. Then, for every $u \in H^s(I; H_0^1(\Omega))$ with $s \ge 2$, we have*

$$
\|\partial_{tt}(u - \Pi_I^q u)\|_{L^2(I;H_0^1(\Omega))} \le C \frac{k^{\mu-2}}{q^{s-3}} \|u\|_{H^s(I;H_0^1(\Omega))} \tag{2.2.23}
$$

where $\mu = \min(q+1, s)$ and C is a universal constant, which may vary from line to line.

Proof.

$$
\begin{aligned}
\|\partial_{tt}(u - \Pi_I^q u)\|_{L^2(I;H_0^1(\Omega))} &\le \|\partial_{tt}(u - \mathcal{Q}^q u)\|_{L^2(I;H_0^1(\Omega))} + \|\partial_{tt}(\mathcal{Q}^q u - \Pi_I^q u)\|_{L^2(I;H_0^1(\Omega))} \\
&\le C \frac{k^{\mu-2}}{q^{s-2}} \|u\|_{H^s(I;H_0^1(\Omega))} + C_{\mathrm{inv}} \frac{q^2}{k} \|\partial_t(\mathcal{Q}^q u - \Pi_I^q u)\|_{L^2(I;H_0^1(\Omega))} \\
&= C \frac{k^{\mu-2}}{q^{s-2}} \|u\|_{H^s(I;H_0^1(\Omega))} + C_{\mathrm{inv}} \frac{q^2}{k} \|\partial_t \mathcal{Q}^q u - \mathcal{P}_I^{q-1}\partial_t u\|_{L^2(I;H_0^1(\Omega))} \\
&\le C \frac{k^{\mu-2}}{q^{s-2}} \|u\|_{H^s(I;H_0^1(\Omega))} + C_{\mathrm{inv}} \frac{q^2}{k} \|\partial_t(u - \mathcal{Q}^q u)\|_{L^2(I;H_0^1(\Omega))} \\
&\quad + C_{\mathrm{inv}} \frac{q^2}{k} \|\partial_t u - \mathcal{P}_I^{q-1}\partial_t u\|_{L^2(I;H_0^1(\Omega))}
\end{aligned}
$$

$$
\begin{aligned}
&\leq C\frac{k^{\mu-2}}{q^{s-2}}\|u\|_{H^s(I;H_0^1(\Omega))} + C\frac{q^2}{k}\frac{k^{\mu-1}}{q^{s-1}}\|u\|_{H^s(I;H_0^1(\Omega))} \\
&\quad + C\frac{q^2}{k}\frac{k^{\mu-1}}{q^{s-1}}\|u\|_{H^s(I;H_0^1(\Omega))} \\
&\leq C\frac{k^{\mu-2}}{q^{s-3}}\|u\|_{H^s(I;H_0^1(\Omega))},
\end{aligned}
$$

where we have used Proposition 2.12 and the inequality

$$
\|\varphi\|_{H^1(I;H_0^1(\Omega))} \leq C_{\text{inv}}\frac{q^2}{k}\|\varphi\|_{L^2(I;H_0^1(\Omega))}
$$

for some constant C_{inv} with $\varphi \in \mathbb{P}^{q-1}(I;H_0^1(\Omega))$ for $q \geq 1$ [70] in the second step. Note that for the second last step, we have used the estimates from Propositions 2.12 and 2.13. □

Remark 2.16. *If we change the spatial function space from $H_0^1(\Omega)$ to $L^2(\Omega)$, the same estimate follows. That is, for every $u \in H^s(I;H_0^1(\Omega))$ with $s \geq 2$, we have*

$$
\|\partial_{tt}(u - \Pi_I^q u)\|_{L^2(I;L^2(\Omega))} \leq C\frac{k^{\mu-2}}{q^{s-3}}\|u\|_{H^s(I;L^2(\Omega))} \tag{2.2.24}
$$

where $\mu = \min(q+1, s)$ and C is a universal constant.

Before we begin the proof of the main convergence theorem, we state some useful properties of the projection error first. Assume that $u \in H^s(0,T;H_0^1(\Omega))$ with $s \geq 2$ is the weak solution of (2.0.1)–(2.0.3) and let $\Pi_I^q u \in \mathcal{V}^{\mathrm{q}}$ be the projection of u such that $\Pi_I^q u_{|I_n} = \Pi_{I_n}^{q_n} u \in \mathcal{V}^{q_n}$ is defined according to (2.2.15) for $n = 1, \ldots, N$. Then the following properties hold for $q_n \geq 2$:

(a) $(u - \Pi_I^q u)(x, t_n^{\pm}) = 0$ for $n = 0, 1, \ldots, N$;

(b) $\partial_t(u - \Pi_I^q u)(x, t_n^-) = 0$ for $n = 1, \ldots, N$;

(c) $\int_{t_{n-1}}^{t_n} \big(\partial_t(u - \Pi_I^q), \varphi\big)\, \mathrm{d}t$ for all $\varphi \in \mathbb{P}^{q_n-2}(I;H_0^1(\Omega))$.

Now we are ready to prove the following convergence theorem.

Theorem 2.17. *Let u be the solution to (2.0.1)–(2.0.3), and let $u_{\mathrm{DG}} \in \mathcal{V}^{\mathbf{q}}$ be the discontinuous Galerkin approximation of u. That is,*

$$\mathcal{A}(u_{\mathrm{DG}}, v) = F(v) \text{ for all } v \in \mathcal{V}^{\mathbf{q}}.$$

If $u_{|I_n} \in H^{s_n}(I_n; H^2(\Omega) \cap H_0^1(\Omega))$ for any $n = 1, \ldots, N$ with $s_n \geq 2$, then

$$|||u - u_{\mathrm{DG}}||| \leq C \left(\sum_{n=1}^{N} \frac{k_n^{2\mu_n - 3}}{q_n^{2s_n - 6}} \|u\|^2_{H^{s_n}(I_n; L^2(\Omega))} + \sum_{n=1}^{N} \frac{k_n^{2\mu_n}}{q_n^{2s_n - 2}} \|u\|^2_{H^{s_n}(I_n; H^2(\Omega))} \right)^{\frac{1}{2}},$$

where $\mu_n = \min(q_n + 1, s_n)$ for any $n = 1, \ldots, N$ and C is a constant independent of u, q_n and k_n.

Proof. Denote the global error by $e := u - u_{\mathrm{DG}}$, and decompose it as $e = e^\pi + e^h$ with $e^\pi = u - \Pi_I^q u$ and $e^h = \Pi_I^q u - u_{\mathrm{DG}}$. First note that $e^\pi(x, t_n^\pm) = (u - \Pi_I^q u)(x, t_n^\pm) = 0$ for $n = 0, \ldots, N$ and $\dot{e}^\pi(x, t_n^-) = \partial_t (u - \Pi_I^q u)(x, t_n^-) = 0$ for $n = 1, \ldots, N$. Therefore,

$$\begin{aligned}
|||e^\pi|||^2 &= \frac{1}{2}\|\dot{e}^\pi(t_0^+)\|^2_{L^2(\Omega)} + \frac{1}{2}\sum_{n=1}^{N-1} \|[\dot{e}^\pi]_n\|^2_{L^2(\Omega)} + \frac{1}{2}\|\dot{e}^\pi(t_N^-)\|^2_{L^2(\Omega)} + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|\dot{e}^\pi\|^2_{L^2(\Omega)}\, \mathrm{d}t \\
&\quad + \frac{1}{2}\|\nabla e^\pi(t_0^+)\|^2_{L^2(\Omega)} + \frac{1}{2}\sum_{n=1}^{N-1} \|\nabla [e^\pi]_n\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla e^\pi(t_N^-)\|^2_{L^2(\Omega)} \\
&= \frac{1}{2}\|\dot{e}^\pi(t_0^+)\|^2_{L^2(\Omega)} + \frac{1}{2}\sum_{n=1}^{N-1} \|\dot{e}^\pi(t_n^+) - \dot{e}(t_n^-)\|^2_{L^2(\Omega)} + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|\dot{e}^\pi\|^2_{L^2(\Omega)}\, \mathrm{d}t \\
&= \frac{1}{2}\sum_{n=1}^{N} \|\dot{e}^\pi(t_{n-1}^+)\|^2_{L^2(\Omega)} + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|\dot{e}^\pi\|^2_{L^2(\Omega)}\, \mathrm{d}t.
\end{aligned}$$

By considering $\dot{e}^\pi(t_{n-1}^+) = -\int_{t_{n-1}^+}^{t_n^-} \ddot{e}^\pi(s)\, \mathrm{d}s$, we have

$$\begin{aligned}
|||e^\pi|||^2 &= \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|\dot{e}^\pi\|^2_{L^2(\Omega)}\, \mathrm{d}t + \frac{1}{2}\sum_{n=1}^{N} \| - \int_{t_{n-1}}^{t_n} \ddot{e}^\pi(s)\, \mathrm{d}s\|^2_{L^2(\Omega)} \\
&\leq \sum_{n=1}^{N} \left(\|\dot{e}^\pi\|^2_{L^2(I_n; L^2(\Omega))} + \frac{k_n}{2}\|\ddot{e}^\pi\|^2_{L^2(I_n; L^2(\Omega))} \right) \text{ (by Hölder's inequality)} \\
&= \sum_{n=1}^{N} \left(\|\partial_t (u - \Pi_{I_n}^{q_n} u)\|^2_{L^2(I_n; L^2(\Omega))} + \frac{k_n}{2}\|\partial_{tt}(u - \Pi_{I_n}^{q_n} u)\|^2_{L^2(I_n; L^2(\Omega))} \right)
\end{aligned}$$

$$\leq C \sum_{n=1}^{N} \frac{k_n^{2\mu_n - 2}}{q_n^{2s_n - 2}} \|u\|^2_{H^{s_n}(I_n;L^2(\Omega))} + C \sum_{n=1}^{N} \frac{k_n^{2\mu_n - 3}}{q^{2s_n - 6}} \|u\|^2_{H^{s_n}(I_n;L^2(\Omega))} \text{ (by (2.2.21) and (2.2.24))}$$
$$\leq C \sum_{n=1}^{N} \frac{k_n^{2\mu_n - 3}}{q_n^{2s_n - 6}} \|u\|^2_{H^{s_n}(I_n;L^2(\Omega))},$$

where $\mu_n = \min(q_n + 1, s_n)$ for any $n = 1, \ldots, N$ and C is a universal constant, which may vary from line to line.

By Galerkin orthogonality, we obtain $\mathcal{A}(e^\pi + e^h, e^h) = \mathcal{A}(u - u_{\mathrm{DG}}, e^h) = 0$ as $e^h = \Pi_I^q u - u_{\mathrm{DG}} \in \mathcal{V}^{\mathbf{q}}$. Then,

$$\begin{aligned}
|||e^h|||^2 &= \mathcal{A}(e^h, e^h) = -\mathcal{A}(e^\pi, e^h) \\
&= -\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left\langle \ddot{e}^\pi, \dot{e}^h \right\rangle \mathrm{d}t - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\dot{e}^\pi, \dot{e}^h)_{L^2} \, \mathrm{d}t - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\nabla e^\pi, \nabla \dot{e}^h)_{L^2} \, \mathrm{d}t \\
&\quad - \sum_{n=1}^{N-1} ([\dot{e}^\pi]_n, \dot{e}^h(t_n^+))_{L^2} - \sum_{n=1}^{N-1} (\nabla [e^\pi]_n, \nabla e^h(t_n^+))_{L^2} - (\dot{e}^\pi(t_0^+), \dot{e}^h(t_0^+))_{L^2} \\
&\quad - (\nabla e^\pi(t_0^+), \nabla e^h(t_0^+)).
\end{aligned}$$

Integrating by parts the term $\left\langle \ddot{e}^\pi, \dot{e}^h \right\rangle$ and rearranging the addends, we have

$$\begin{aligned}
|||e^h|||^2 &= \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\dot{e}^\pi, \ddot{e}^h)_{L^2} \, \mathrm{d}t - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\dot{e}^\pi, \dot{e}^h)_{L^2} \, \mathrm{d}t - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\nabla e^\pi, \nabla \dot{e}^h)_{L^2} \, \mathrm{d}t \\
&\quad + \sum_{n=1}^{N-1} (\dot{e}^\pi(t_n^-), [\dot{e}^h]_n)_{L^2} - \sum_{n=1}^{N-1} (\nabla [e^\pi]_n, \nabla e^h(t_n^+))_{L^2} - (\dot{e}^\pi(t_N^-), \dot{e}^h(t_N^-))_{L^2} \\
&\quad - (\nabla e^\pi(t_0^+), \nabla e^h(t_0^+))_{L^2}.
\end{aligned}$$

Note that the first term vanishes by the orthogonality condition that

$$\int_{t_{n-1}}^{t_n} (\dot{e}^\pi, \varphi)_{L^2} \ \mathrm{d}t = 0 \text{ for all } \varphi \in \mathbb{P}^{q_n - 2}(I_n; H_0^1(\Omega)),$$

and the terms on the second and third lines vanish by the properties of the projection error. Integrating by parts the term $\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\nabla e^\pi, \nabla \dot{e}^h)_{L^2} \, \mathrm{d}t$ in the x variable, we have

$$|||e^h|||^2 = -\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\dot{e}^\pi, \dot{e}^h)_{L^2} \, \mathrm{d}t + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\Delta e^\pi, \dot{e}^h)_{L^2} \, \mathrm{d}t$$

$$
\begin{aligned}
&\leq \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\int_\Omega |\dot{e}^\pi||\dot{e}^h|\,\mathrm{d}x\,\mathrm{d}t + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\int_\Omega |\Delta e^\pi||\dot{e}^h|\,\mathrm{d}x\,\mathrm{d}t\\
&\leq \sum_{n=1}^{N}\|\dot{e}^\pi\|_{L^2(I_n;L^2(\Omega))}\|\dot{e}^h\|_{L^2(I_n;L^2(\Omega))} + \sum_{n=1}^{N}\|\Delta e^\pi\|_{L^2(I_n;L^2(\Omega))}\|\dot{e}^h\|_{L^2(I_n;L^2(\Omega))}\\
&\leq \sum_{n=1}^{N}\left(\|\dot{e}^\pi\|^2_{L^2(I_n;L^2(\Omega))} + \|\Delta e^\pi\|^2_{L^2(I_n;L^2(\Omega))}\right) + \frac{1}{2}\sum_{n=1}^{N}\|\dot{e}^h\|^2_{L^2(I_n;L^2(\Omega))}
\end{aligned}
$$

by Young's inequality. Note that $\sum_{n=1}^{N}\|\dot{e}^h\|_{L^2(I_n;L^2(\Omega))} \leq |||e^h|||$, so we can absorb the $\frac{1}{2}\sum_{n=1}^{N}\|\dot{e}^h\|^2_{L^2(I_n;L^2(\Omega))}$ term into the left-hand side of the inequality to get

$$
\begin{aligned}
|||e^h|||^2 &\leq 2\sum_{n=1}^{N}\left(\|\dot{e}^\pi\|^2_{L^2(I_n;L^2(\Omega))} + \|\Delta e^\pi\|^2_{L^2(I_n;L^2(\Omega))}\right)\\
&\leq C\sum_{n=1}^{N}\frac{k_n^{2\mu_n-2}}{q_n^{2s_n-2}}\|u\|^2_{H^{s_n}(I_n;L^2(\Omega))} + C\sum_{n=1}^{N}\frac{k_n^{2\mu_n}}{q_n^{2s_n-2}}\|\Delta u\|^2_{H^{s_n}(I_n;L^2(\Omega))} \text{ (by Lemma 2.14)},
\end{aligned}
$$

where $\mu_n = \min(q_n+1, s_n)$ for any $n = 1, \ldots, N$. By the definition of the H^2–norm, we know that $\|\Delta u\|_{L^2(\Omega)} \leq \|u\|_{H^2(\Omega)}$. Therefore,

$$
\begin{aligned}
|||e||| &\leq |||e^\pi||| + |||e^h|||\\
&\leq \left(C\sum_{n=1}^{N}\frac{k_n^{2\mu_n-3}}{q_n^{2s_n-6}}\|u\|^2_{H^{s_n}(I_n;L^2(\Omega))}\right)^{\frac{1}{2}}\\
&\quad + \left(C\sum_{n=1}^{N}\frac{k_n^{2\mu_n-2}}{q_n^{2s_n-2}}\|u\|^2_{H^{s_n}(I_n;L^2(\Omega))} + C\sum_{n=1}^{N}\frac{k_n^{2\mu_n}}{q_n^{2s_n-2}}\|u\|^2_{H^{s_n}(I_n;H^2(\Omega))}\right)^{\frac{1}{2}}\\
&\leq C\left(\sum_{n=1}^{N}\frac{k_n^{2\mu_n-3}}{q_n^{2s_n-6}}\|u\|^2_{H^{s_n}(I_n;L^2(\Omega))} + \sum_{n=1}^{N}\frac{k_n^{2\mu_n}}{q_n^{2s_n-2}}\|u\|^2_{H^{s_n}(I_n;H^2(\Omega))}\right)^{\frac{1}{2}},
\end{aligned}
$$

where C is a universal constant, which may vary from line to line. □

Remark 2.18. *If we use uniform time intervals $k_n = k$, and uniform polynomial degrees $q_n = q \geq 2, s_n = s \geq 2$ for $n = 1, \ldots, N$, then the error bound becomes*

$$
|||u - u_{\mathrm{DG}}||| \leq C\left(\frac{k^{2\mu-3}}{q^{2s-6}}\|u\|^2_{H^s(0,T;L^2(\Omega))} + \frac{k^{2\mu}}{q^{2s-2}}\|u\|^2_{H^s(0,T;H^2(\Omega))}\right)^{\frac{1}{2}} \leq C(u)\frac{k^{\mu-\frac{3}{2}}}{q^{s-3}}
$$

where $\mu = \min(q+1, s)$ and $C(u)$ is a constant depending on the exact solution u, but *independent of q and k.*

2.3. Fully discrete numerical scheme

In this section, we shall construct a fully discrete scheme for the approximation of the solution of (2.0.1)–(2.0.3). This numerical scheme combines a hp–DGFEM in the time direction with an $H^1(\Omega)$-conforming finite element approximation in the spatial variables. It is well-known that discontinuous Galerkin finite element methods offer certain advantages over standard continuous Galerkin methods when applied to the spatial discretisation of the acoustic wave equation [4]. For instance, the mass matrix has a desirable block diagonal structure, which gives a more efficient computation when using an explicit time scheme. However, for the sake of having a neat and concise convergence analysis, we stick to a continuous Galerkin method in the spatial direction for this thesis; applying the DGFEM in both the time and spatial directions will be considered in our future work.

2.3.1. Construction of the fully discrete scheme For the spatial discretisation parameter $h \in (0,1)$, we define $\mathcal{V}_h$ to be a given family of finite element subspaces of $H_0^1(\Omega)$ with polynomial degree $p \geq 1$. We shall assume that the triangulation $\{\mathcal{T}_h\}$ of Ω 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 [15] that

$$\inf_{v\in\mathcal{V}_h}\{\|u-v\|_{L^2(\Omega)}+h\|u-v\|_{H^1(\Omega)}\} \leq Ch^{r+1}\|u\|_{H^{r+1}(\Omega)}, \quad 1 \leq r \leq \min(p, m-1), \tag{2.3.1}$$

for $u \in H^m(\Omega) \cap H_0^1(\Omega)$. Now we introduce the space-time finite element space by

$$\mathcal{V}_{kh}^{q_n} := \{u\colon [0,T] \to \mathcal{V}_h;\ u\,|_{I_n} = \sum_{j=1}^{q_n} u_j t^j,\ u_j \in \mathcal{V}_h\},$$

with $q_n \geq 2$ for each $1 \leq n \leq N$. For $\mathbf{q} := [q_1, \ldots, q_N]^T \in \mathbb{N}^N$, we then define the space

$$\mathcal{V}_{kh}^{\mathbf{q}} := \{u\colon [0,T] \to \mathcal{V}_h;\ u\,|_{I_n} \in \mathcal{V}_{kh}^{q_n} \text{ for } n = 1, \ldots, N\}.$$

Following the same procedure as in Section 2.2, we can now define the fully discrete discontinuous-in-time scheme. We seek a solution $u_{\mathrm{DG}} \in \mathcal{V}_{kh}^{\mathbf{q}}$ such that

$$\mathcal{A}(u_{\mathrm{DG}}, v) = \tilde{F}(v) \quad \text{ for all } v \in \mathcal{V}_{kh}^{\mathbf{q}}, \tag{2.3.2}$$

where the bilinear form $\mathcal{A}(\cdot,\cdot)\colon \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ is defined by (2.2.4) and $\tilde{F}$ is a modified version of F in (2.2.5) defined as

$$\tilde{F}(v) := \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (f(t), \dot{v}(t))_{L^2} \,\mathrm{d}t + \big(\mathcal{P}_h u_1, \dot{v}(t_0^+)\big)_{L^2} + \big(\nabla \mathcal{P}_h u_0, \nabla v(t_0^+)\big)_{L^2}, \tag{2.3.3}$$

where $\mathcal{P}_h\colon H_0^1(\Omega) \to \mathcal{V}_h$ is the Ritz projection such that

$$(\nabla \mathcal{P}_h u, \nabla \varphi)_{L^2} = (\nabla u, \nabla \varphi)_{L^2} \quad \text{ for all } \varphi \in \mathcal{V}_h. \tag{2.3.4}$$

The existence and uniqueness of the fully discrete solution $u_{\mathrm{DG}} \in \mathcal{V}_{kh}^{\mathbf{q}}$ of (2.3.2) follow from the following proposition.

Proposition 2.19. *For each $1 \le n \le N$, the local problem*

$$\begin{aligned}
&\int_{t_{n-1}}^{t_n} \langle \ddot{u}, \dot{v}\rangle \,\mathrm{d}t + \int_{t_{n-1}}^{t_n} (\dot{u}, \dot{v})_{L^2} \,\mathrm{d}t + \int_{t_{n-1}}^{t_n} (\nabla u, \nabla \dot{v})_{L^2} \,\mathrm{d}t + (\dot{u}_{n-1}^+, \dot{v}_{n-1}^+)_{L^2} + (\nabla u_{n-1}^+, \nabla v_{n-1}^+)_{L^2} \\
&= \int_{t_{n-1}}^{t_n} (f, \dot{v})_{L^2} \,\mathrm{d}t + (\dot{u}_{n-1}^-, \dot{v}_{n-1}^+)_{L^2} + (\nabla u_{n-1}^-, \nabla v_{n-1}^+)_{L^2}
\end{aligned} \tag{2.3.5}$$

admits a unique solution $u \in \mathcal{V}_{kh}^{q_n}$ on I_n provided that u_{n-1} and $f\,|_{I_n}$ are given.

Proof. We first note that to show uniqueness, it suffices to see that the corresponding homogeneous equation

$$\begin{aligned}
&\int_{t_{n-1}}^{t_n} \langle \ddot{u}, \dot{v}\rangle \,\mathrm{d}t + \int_{t_{n-1}}^{t_n} (\dot{u}, \dot{v})_{L^2} \,\mathrm{d}t + \int_{t_{n-1}}^{t_n} (\nabla u, \nabla \dot{v})_{L^2} \,\mathrm{d}t + (\dot{u}_{n-1}^+, \dot{v}_{n-1}^+)_{L^2} \\
&\quad + (\nabla u_{n-1}^+, \nabla v_{n-1}^+)_{L^2} = 0
\end{aligned}$$

for all $v \in \mathcal{V}_{kh}^{q_n}$ only has the trivial solution $u \equiv 0$. For this purpose, we assume that u is a solution and choose $v = u$ on I_n; then we have

$$\|\dot{u}_n^-\|_{L^2(\Omega)}^2 - \|\dot{u}_{n-1}^+\|_{L^2(\Omega)}^2 + 2\int_{t_{n-1}}^{t_n} \|\dot{u}\|_{L^2(\Omega)}^2 \,\mathrm{d}t + \|\nabla u_n^-\|_{L^2(\Omega)}^2 - \|\nabla u_{n-1}^+\|_{L^2(\Omega)}^2$$

$$+ 2\|\dot{u}_{n-1}^{+}\|_{L^2(\Omega)}^2 + 2\|\nabla u_{n-1}^{+}\|_{L^2(\Omega)}^2 = 0.$$

That is,

$$\|\dot{u}_n^{-}\|_{L^2(\Omega)}^2 + \|\dot{u}_{n-1}^{+}\|_{L^2(\Omega)}^2 + 2\int_{t_{n-1}}^{t_n} \|\dot{u}\|_{L^2(\Omega)}^2 \, \mathrm{d}t + \|\nabla u_n^{-}\|_{L^2(\Omega)}^2 + \|\nabla u_{n-1}^{+}\|_{L^2(\Omega)}^2 = 0.$$

Note that $\|u_{n-1}^{+}\|_{H_0^1(\Omega)} = \|\nabla u_{n-1}^{+}\|_{L^2(\Omega)} = 0$ implies that $u_{n-1}^{+} \equiv 0$. Then we have

$$\begin{cases} \dot{u}(t) = 0 & \text{for } t \in I_n, \\ u_{n-1}^{+} = 0. \end{cases}$$

This implies that $u(t) \equiv 0$ on I_n. The existence of a solution to the local problem (2.3.5) follows from the uniqueness since this is a finite–dimensional problem. □

2.3.2. Convergence analysis In this section we will prove a space-time *a priori error* estimate in the energy norm (2.2.7). This error estimate is optimal with respect to both the spatial and temporal parameters.

Theorem 2.20. *Let u be the solution to (2.0.1)–(2.0.3) such that $u_{|I_n} \in H^{s_n}(I_n; H^m(\Omega) \cap H_0^1(\Omega))$ for any $n = 1, \ldots, N$ with $s_n \geq 2$, and let $u_{\mathrm{DG}} \in \mathcal{V}_{kh}^{\mathbf{q}}$ be the discontinuous Galerkin approximation of u. That is,*

$$\mathcal{A}(u_{\mathrm{DG}}, v) = \tilde{F}(v) \text{ for all } v \in \mathcal{V}_{kh}^{\mathbf{q}}.$$

Assume that $1 \leq r \leq \min(p, m-1)$, $s_n \geq q_n + 1$ for each $n = 1, \ldots, N$; then we have

$$\begin{aligned} |||u - u_{\mathrm{DG}}||| \leq C\Bigg\{ \sum_{n=1}^{N} \frac{k_n^{2q_n-1}}{q_n^{2s_n-6}} \|u\|_{H^{s_n}(I_n;H^{r+1}(\Omega))}^2 + h^{2r+2} \sum_{n=1}^{N} \|u\|_{H^{s_n}(I_n;H^{r+1}(\Omega))}^2 \Bigg\}^{\frac{1}{2}} \\ + \left\{ h^{2r+2} \left(\|\dot{u}(t_0^{+})\|_{H^{r+1}(\Omega)}^2 + \|\dot{u}(t_N^{-})\|_{H^{r+1}(\Omega)}^2 \right) + h^{2r} \left(\|u(t_0^{+})\|_{H^{r+1}(\Omega)}^2 + \|u(t_N^{-})\|_{H^{r+1}(\Omega)}^2 \right) \right\}^{\frac{1}{2}}, \end{aligned} \tag{2.3.6}$$

where C is a positive constant, which may vary from line to line.

Proof. Let $\Pi_{I_n}^{q_n}$ denote the modified L^2-projection defined by Definition 2.9 on I_n for each

$n = 1, \dots, N$. We decompose the error as

$$u_{\mathrm{DG}}(t) - u(t) = \big(u_{\mathrm{DG}}(t) - \Pi_{I_n}^{q_n}\mathcal{P}_h u(t)\big) + \big(\Pi_{I_n}^{q_n}\mathcal{P}_h u(t) - \mathcal{P}_h u(t)\big) + (\mathcal{P}_h u(t) - u(t))$$
$$:= \theta(t) + \rho_1(t) + \rho_2(t),$$

for $t \in I_n$, with $n = 1, \dots, N$. For simplicity of notation, we define $(\Pi_{I_n}^{q_n} w)^{(j)} := \frac{\mathrm{d}^j \Pi_{I_n}^{q_n} w}{\mathrm{d}t^j}$ and $u^{(j)} := \frac{\mathrm{d}^j u}{\mathrm{d}t^j}$, for $j = 0, 1$ and 2. By Definition 2.9, Lemma 2.14 and Corollary 2.15, we have

$$\begin{aligned}\int_{t_{n-1}}^{t_n} \|\rho_1^{(j)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t &= \int_{t_{n-1}}^{t_n} \|(\Pi_{I_n}^{q_n}\mathcal{P}_h u)^{(j)}(t) - (\mathcal{P}_h u)^{(j)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\ &\le C \frac{k_n^{2(\mu-j)}}{q_n^{2(s_n-1)}} \|u\|_{H^{s_n}(I_n;L^2(\Omega))}^2 \quad \text{for } j = 0, 1, \end{aligned} \tag{2.3.7}$$

and

$$\begin{aligned}\int_{t_{n-1}}^{t_n} \|\rho_1^{(2)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t &= \int_{t_{n-1}}^{t_n} \|(\Pi_{I_n}^{q_n}\mathcal{P}_h u)^{(2)}(t) - (\mathcal{P}_h u)^{(2)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\ &\le C \frac{k_n^{2(\mu-2)}}{q_n^{2(s_n-3)}} \|u\|_{H^{s_n}(I_n;L^2(\Omega))}^2, \end{aligned} \tag{2.3.8}$$

where $\mu = \min(s_n, q_n+1)$ and C is a universal constant, which may vary from line to line. If we assume that the solution u of (2.0.1)–(2.0.3) is sufficiently smooth (i.e., $s_n > q_n+1$), then we can write

$$\int_{t_{n-1}}^{t_n} \|\rho_1^{(j)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \le C \frac{k_n^{2(q_n+1-j)}}{q_n^{2(s_n-1)}} \|u\|_{H^{s_n}(I_n;L^2(\Omega))}^2, \tag{2.3.9}$$

for $j = 0, 1$, and

$$\int_{t_{n-1}}^{t_n} \|\rho_1^{(2)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \le C \frac{k_n^{2q_n-2}}{q_n^{2(s_n-3)}} \|u\|_{H^{s_n}(I_n;L^2(\Omega))}^2. \tag{2.3.10}$$

By Céa's lemma and the interpolation property (2.3.1), we know that

$$\|\nabla(\mathcal{P}_h u - u)\|_{L^2(\Omega)} \le \min_{v \in \mathcal{V}_h} \|\nabla u - \nabla v\|_{L^2(\Omega)} \le C h^r \|u\|_{H^{r+1}(\Omega)}. \tag{2.3.11}$$

Applying a duality argument (cf. Theorem 5.4.8 in [17]), we have

$$\|\mathcal{P}_h u - u\|_{L^2(\Omega)} \le h\|\nabla(\mathcal{P}_h u - u)\|_{L^2(\Omega)} \le Ch^{r+1}\|u\|_{H^{r+1}(\Omega)}. \tag{2.3.12}$$

Thus, we can approximate the terms involving ρ_2 by the following:

$$\begin{aligned}\int_{t_{n-1}}^{t_n} \|\rho_2^{(j)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t &= \int_{t_{n-1}}^{t_n} \|\mathcal{P}_h u^{(j)}(t) - u^{(j)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\ &\le Ch^{2r+2}\|u\|_{H^j(I_n;H^{r+1}(\Omega))}^2 \qquad \text{for } j = 0,1,2. \end{aligned} \tag{2.3.13}$$

Subtracting the variational formulation from the fully discrete scheme, we have

$$\begin{aligned}&\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \left\langle \ddot\theta(t) + \ddot\rho_1(t) + \ddot\rho_2(t), \dot v(t)\right\rangle \,\mathrm{d}t + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \left(\dot\theta(t) + \dot\rho_1(t) + \dot\rho_2(t), \dot v(t)\right)_{L^2} \mathrm{d}t \\ &\quad + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} (\nabla\theta(t) + \nabla\rho_1(t) + \nabla\rho_2(t), \nabla\dot v(t))_{L^2} \,\mathrm{d}t + \sum_{n=1}^{N-1}([\dot\theta + \dot\rho_1 + \dot\rho_2]_n, \dot v(t_n^+))_{L^2} \\ &\quad + \sum_{n=1}^{N-1}([\nabla\theta + \nabla\rho_1 + \nabla\rho_2]_n, \nabla v(t_n^+))_{L^2} + (\dot u_{\mathrm{DG}}(t_0^+) - \dot u(t_0^+), \dot v(t_0^+))_{L^2} \\ &\quad + (\nabla u_{\mathrm{DG}}(t_0^+) - \nabla u(t_0^+), \nabla v(t_0^+))_{L^2} \\ &= \left(\mathcal{P}_h u_1 - u_1, \dot v(t_0^+)\right)_{L^2} + \left(\nabla\mathcal{P}_h u_0 - \nabla u_0, \nabla v(t_0^+)\right)_{L^2}. \end{aligned} \tag{2.3.14}$$

By taking $v = \theta$ in (2.3.14), we have

$$\begin{aligned}&\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \left\langle \ddot\theta(t), \dot\theta(t)\right\rangle \,\mathrm{d}t + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \left(\dot\theta(t), \dot\theta(t)\right)_{L^2} \mathrm{d}t + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \left(\nabla\theta(t), \nabla\dot\theta(t)\right)_{L^2} \mathrm{d}t \\ &\quad + \sum_{n=1}^{N-1}([\dot\theta]_n, \dot\theta(t_n^+))_{L^2} + \sum_{n=1}^{N-1}([\nabla\theta]_n, \nabla\theta(t_n^+))_{L^2} + (\dot\theta(t_0^+), \dot\theta(t_0^+))_{L^2} + (\nabla\theta(t_0^+), \nabla\theta(t_0^+))_{L^2} \\ &= -\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \left\langle \ddot\rho_1(t), \dot\theta(t)\right\rangle \,\mathrm{d}t - \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \left(\dot\rho_1(t), \dot\theta(t)\right)_{L^2} \mathrm{d}t - \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \left(\nabla\rho_1(t), \nabla\dot\theta(t)\right)_{L^2} \\ &\quad - \sum_{n=1}^{N-1}\left([\dot\rho_1]_n, \dot\theta(t_n^+)\right)_{L^2} \mathrm{d}t - \sum_{n=1}^{N-1}\left([\nabla\rho_1]_n, \nabla\theta(t_n^+)\right)_{L^2} \\ &\quad - \left(\dot\rho_1(t_0^+), \dot\theta(t_0^+)\right)_{L^2} - \left(\nabla\rho_1(t_0^+), \nabla\theta(t_0^+)\right)_{L^2} - \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \left\langle \ddot\rho_2(t), \dot\theta(t)\right\rangle \,\mathrm{d}t \end{aligned}$$

$$
\begin{aligned}
&- \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big(\dot{\rho}_2(t), \dot{\theta}(t)\Big)_{L^2} \,\mathrm{d}t - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big(\nabla\rho_2(t), \nabla\dot{\theta}(t)\Big)_{L^2} \,\mathrm{d}t \\
&- \sum_{n=1}^{N-1} \Big([\dot{\rho}_2]_n, \dot{\theta}(t_n^+)\Big)_{L^2} - \sum_{n=1}^{N-1} \big([\nabla\rho_2]_n, \nabla\theta(t_n^+)\big)_{L^2} - \Big(\dot{\rho}_2(t_0^+), \dot{\theta}(t_0^+)\Big)_{L^2} \\
&- \big(\nabla\rho_2(t_0^+), \nabla\theta(t_0^+)\big)_{L^2} + \Big(\mathcal{P}_h u_1 - u_1, \dot{\theta}(t_0^+)\Big)_{L^2} + \big(\nabla\mathcal{P}_h u_0 - \nabla u_0, \nabla\theta(t_0^+)\big)_{L^2}.
\end{aligned}
\tag{2.3.15}
$$

Since $\rho_2(t) := \mathcal{P}_h u(t) - u(t)$, both $\rho_2(t)$ and $\dot{\rho}_2(t)$ are continuous in t. By the property of the Ritz operator, we also have that

$$
\Big(\nabla\rho_2(t), \nabla\dot{\theta}(t)\Big)_{L^2} = 0 \quad \text{for all } t \in I_n,\ n = 1, \ldots, N.
$$

Thus, we can write (2.3.15) as follows:

$$
\begin{aligned}
&\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big\langle \ddot{\theta}(t), \dot{\theta}(t)\Big\rangle \,\mathrm{d}t + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big(\dot{\theta}(t), \dot{\theta}(t)\Big)_{L^2} \,\mathrm{d}t + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big(\nabla\theta(t), \nabla\dot{\theta}(t)\Big)_{L^2} \,\mathrm{d}t \\
&\quad + \sum_{n=1}^{N-1} ([\dot{\theta}]_n, \dot{\theta}(t_n^+))_{L^2} + \sum_{n=1}^{N-1} ([\nabla\theta]_n, \nabla\theta(t_n^+))_{L^2} + (\dot{\theta}(t_0^+), \dot{\theta}(t_0^+))_{L^2} + (\nabla\theta(t_0^+), \nabla\theta(t_0^+))_{L^2} \\
&= - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big\langle \ddot{\rho}_1(t), \dot{\theta}(t)\Big\rangle \,\mathrm{d}t - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big(\dot{\rho}_1(t), \dot{\theta}(t)\Big)_{L^2} \,\mathrm{d}t \\
&\quad - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big(\nabla\rho_1(t), \nabla\dot{\theta}(t)\Big)_{L^2} \,\mathrm{d}t - \sum_{n=1}^{N-1} \Big([\dot{\rho}_1]_n, \dot{\theta}(t_n^+)\Big)_{L^2} \\
&\quad - \sum_{n=1}^{N-1} \big([\nabla\rho_1]_n, \nabla\theta(t_n^+)\big)_{L^2} - \Big(\dot{\rho}_1(t_0^+), \dot{\theta}(t_0^+)\Big)_{L^2} - \big(\nabla\rho_1(t_0^+), \nabla\theta(t_0^+)\big)_{L^2} \\
&\quad - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big\langle \ddot{\rho}_2(t), \dot{\theta}(t)\Big\rangle \,\mathrm{d}t - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big(\dot{\rho}_2(t), \dot{\theta}(t)\Big)_{L^2} \,\mathrm{d}t.
\end{aligned}
\tag{2.3.16}
$$

Note that

$$
\begin{aligned}
\int_{t_{n-1}}^{t_n} \Big\langle \ddot{\rho}_1(t), \dot{\theta}(t)\Big\rangle \,\mathrm{d}t &= \int_{t_{n-1}}^{t_n} \frac{\mathrm{d}}{\mathrm{d}t} \Big(\dot{\rho}_1(t), \dot{\theta}(t)\Big) \,\mathrm{d}t - \int_{t_{n-1}}^{t_n} \Big(\dot{\rho}_1(t), \ddot{\theta}(t)\Big)_{L^2} \,\mathrm{d}t \\
&= \Big(\dot{\rho}_1(t_n^-), \dot{\theta}(t_n^-)\Big)_{L^2} - \Big(\dot{\rho}_1(t_{n-1}^+), \dot{\theta}(t_{n-1}^+)\Big)_{L^2},
\end{aligned}
\tag{2.3.17}
$$

where we have used (d) of Lemma 2.11. Substituting (2.3.17) into (2.3.16) and rearranging the addends yield

$$\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left\langle\ddot{\theta}(t),\dot{\theta}(t)\right\rangle\,\mathrm{d}t+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left(\dot{\theta}(t),\dot{\theta}(t)\right)_{L^2}\,\mathrm{d}t+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left(\nabla\theta(t),\nabla\dot{\theta}(t)\right)_{L^2}\,\mathrm{d}t$$
$$+\sum_{n=1}^{N-1}([\dot{\theta}]_n,\dot{\theta}(t_n^+))_{L^2}+\sum_{n=1}^{N-1}(\nabla[\theta]_n,\nabla\theta(t_n^+))_{L^2}+(\dot{\theta}(t_0^+),\dot{\theta}(t_0^+))_{L^2}+(\nabla\theta(t_0^+),\nabla\theta(t_0^+))_{L^2}$$
$$=-\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left(\dot{\rho}_1(t),\dot{\theta}(t)\right)_{L^2}\,\mathrm{d}t-\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left(\nabla\rho_1(t),\nabla\dot{\theta}(t)\right)_{L^2}\,\mathrm{d}t$$
$$-\sum_{n=1}^{N-1}\left(\dot{\rho}_1(t_n^-),[\dot{\theta}]_n\right)_{L^2}-\sum_{n=1}^{N-1}([\nabla\rho_1]_n,\nabla\theta(t_n^+))_{L^2}-\left(\dot{\rho}_1(t_N^-),\dot{\theta}(t_N^-)\right)_{L^2}$$
$$-\left(\nabla\rho_1(t_0^+),\nabla\theta(t_0^+)\right)_{L^2}-\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left\langle\ddot{\rho}_2(t),\dot{\theta}(t)\right\rangle\,\mathrm{d}t-\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left(\dot{\rho}_2(t),\dot{\theta}(t)\right)_{L^2}\,\mathrm{d}t. \tag{2.3.18}$$

By (a)–(c) in Lemma 2.11, we have $\rho_1(t_n^{\pm})=0$ and $\dot{\rho}_1(t_n^-)=0$ for $n=1,\ldots,N$. Thus,

$$\frac{1}{2}\|\dot{\theta}(t_0^+)\|_{L^2(\Omega)}^2+\frac{1}{2}\sum_{n=1}^{N-1}\|[\dot{\theta}]_n\|_{L^2(\Omega)}^2+\frac{1}{2}\|\dot{\theta}(t_N^-)\|_{L^2(\Omega)}^2+\frac{1}{2}\|\nabla\theta(t_0^+)\|_{L^2(\Omega)}^2+\frac{1}{2}\sum_{n=1}^{N-1}\|[\nabla\theta]_n\|_{L^2(\Omega)}^2$$
$$+\frac{1}{2}\|\nabla\theta(t_N^-)\|_{L^2(\Omega)}^2+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\theta}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t$$
$$=-\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left(\dot{\rho}_1(t),\dot{\theta}(t)\right)_{L^2}\,\mathrm{d}t-\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left(\nabla\rho_1(t),\nabla\dot{\theta}(t)\right)_{L^2}\,\mathrm{d}t$$
$$-\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left\langle\ddot{\rho}_2(t),\dot{\theta}(t)\right\rangle\,\mathrm{d}t-\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left(\dot{\rho}_2(t),\dot{\theta}(t)\right)_{L^2}\,\mathrm{d}t$$
$$\le 2\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\rho}_1(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t+2\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\ddot{\rho}_2(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t+2\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\rho}_2(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t$$
$$+3\times\frac{1}{8}\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\theta}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\Delta\rho_1(t)\|_{L^2(\Omega)}\|\dot{\theta}(t)\|_{L^2(\Omega)}\,\mathrm{d}t$$
$$\le 2\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\rho}_1(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t+2\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\rho}_2(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t+2\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\ddot{\rho}_2(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t$$
$$+2\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\Delta\rho_1(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t+\frac{1}{2}\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\theta}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t$$

$$
\begin{aligned}
&\leq \frac{1}{2}\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \|\dot{\theta}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t + 2C\sum_{n=1}^{N}\frac{k_n^{2q_n}}{q_n^{2(s_n-1)}}\|u\|_{H^{s_n}(I_n;L^2(\Omega))}^2 + 2C\sum_{n=1}^{N} h^{2r+2}\|u\|_{H^1(I_n;H^{r+1}(\Omega))}^2 \\
&\quad + 2C\sum_{n=1}^{N} h^{2r+2}\|u\|_{H^2(I_n;H^{r+1}(\Omega))}^2 + 2C\sum_{n=1}^{N}\frac{k_n^{2q_n+2}}{q_n^{2(s_n-1)}}\|u\|_{H^{s_n}(I_n;H^2(\Omega))}^2 \\
&\leq C\sum_{n=1}^{N}\frac{k_n^{2q_n}}{q_n^{2(s_n-1)}}\|u\|_{H^{s_n}(I_n;H^{r+1}(\Omega))}^2 + C\sum_{n=1}^{N} h^{2r+2}\|u\|_{H^2(I_n;H^{r+1}(\Omega))}^2 + \frac{1}{2}\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \|\dot{\theta}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t.
\end{aligned}
\tag{2.3.19}
$$

We can absorb the last term on the right-hand side of (2.3.19) into the left-hand side to obtain that

$$
|||\theta|||^2 \leq 2C\sum_{n=1}^{N}\frac{k_n^{2q_n}}{q_n^{2s_n-2}}\|u\|_{H^{s_n}(I_n;H^{r+1}(\Omega))}^2 + 2Ch^{2r+2}\sum_{n=1}^{N}\|u\|_{H^2(I_n;H^{r+1}(\Omega))}^2. \tag{2.3.20}
$$

Note that

$$
\begin{aligned}
|||\rho_2|||^2 &= \frac{1}{2}\|\dot{\rho}_2(t_0^+)\|_{L^2(\Omega)}^2 + \frac{1}{2}\sum_{n=1}^{N-1}\|[\dot{\rho}_2]_n\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\dot{\rho}_2(t_N^-)\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\nabla\rho_2(t_0^+)\|_{L^2(\Omega)}^2 \\
&\quad + \frac{1}{2}\sum_{n=1}^{N-1}\|[\nabla\rho_2]_n\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\nabla\rho_2(t_N^-)\|_{L^2(\Omega)}^2 + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\rho}_2(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t \\
&= \frac{1}{2}\|\dot{\rho}_2(t_0^+)\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\dot{\rho}_2(t_N^-)\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\nabla\rho_2(t_0^+)\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\nabla\rho_2(t_N^-)\|_{L^2(\Omega)}^2 \\
&\quad + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\rho}_2(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t \\
&\leq Ch^{2r+2}\left(\|\dot{u}(t_0^+)\|_{H^{r+1}(\Omega)}^2 + \|\dot{u}(t_N^-)\|_{H^{r+1}(\Omega)}^2\right) \\
&\quad + Ch^{2r}\left(\|u(t_0^+)\|_{H^{r+1}(\Omega)}^2 + \|u(t_N^-)\|_{H^{r+1}(\Omega)}^2\right) + Ch^{2r+2}\|u\|_{H^1(0,T;H^{r+1}(\Omega))}^2.
\end{aligned}
$$

We also have that

$$
\begin{aligned}
|||\rho_1|||^2 &= \frac{1}{2}\|\dot{\rho}_1(t_0^+)\|_{L^2(\Omega)}^2 + \frac{1}{2}\sum_{n=1}^{N-1}\|[\dot{\rho}_1]_n\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\dot{\rho}_1(t_N^-)\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\nabla\rho_1(t_0^+)\|_{L^2(\Omega)}^2 \\
&\quad + \frac{1}{2}\sum_{n=1}^{N-1}\|[\nabla\rho_1]_n\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\nabla\rho_1(t_N^-)\|_{L^2(\Omega)}^2 + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\rho}_1(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t \\
&= \frac{1}{2}\sum_{n=1}^{N}\|\dot{\rho}_1(t_{n-1}^+)\|_{L^2(\Omega)}^2 + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\rho}_1(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t
\end{aligned}
$$

$$= \frac{1}{2}\sum_{n=1}^{N} \| - \int_{t_{n-1}}^{t_n} \ddot{\rho}_1(s)\,\mathrm{d}s\|^2_{L^2(\Omega)} + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \|\dot{\rho}_1(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$
$$\le \sum_{n=1}^{N} \frac{k_n}{2}\|\partial_{tt}(\Pi_{I_n}^{q_n}\mathcal{P}_h u - \mathcal{P}_h u)\|^2_{L^2(I_n;L^2(\Omega))} + \sum_{n=1}^{N}\|\partial_t(\Pi_{I_n}^{q_n}\mathcal{P}_h u - \mathcal{P}_h u)\|^2_{L^2(I_n;L^2(\Omega))}$$
$$\le C\sum_{n=1}^{N}\frac{k_n^{2q_n-1}}{q_n^{2(s_n-3)}}\|\mathcal{P}_h u\|^2_{H^{s_n}(I_n;L^2(\Omega))} + C\sum_{n=1}^{N}\frac{k_n^{2q_n}}{q_n^{2(s_n-1)}}\|\mathcal{P}_h u\|^2_{H^{s_n}(I_n;L^2(\Omega))}$$
$$\le C\sum_{n=1}^{N}\frac{k_n^{2q_n-1}}{q_n^{2s_n-6}}\|u\|^2_{H^{s_n}(I_n;L^2(\Omega))}.$$

By the triangle inequality, we deduce that

$$\begin{aligned}|||u - u_{\mathrm{DG}}||| &\le |||\theta||| + |||\rho_1||| + |||\rho_2|||\\
&\le C\Bigg\{\sum_{n=1}^{N}\frac{k_n^{2q_n-1}}{q_n^{2s_n-6}}\|u\|^2_{H^{s_n}(I_n;H^{r+1}(\Omega))} + h^{2r+2}\sum_{n=1}^{N}\|u\|^2_{H^2(I_n;H^{r+1}(\Omega))}\Bigg\}^{\frac{1}{2}}\\
&\quad + C\Bigg\{h^{2r+2}\left(\|\dot{u}(t_0^+)\|^2_{H^{r+1}(\Omega)} + \|\dot{u}(t_N^-)\|^2_{H^{r+1}(\Omega)}\right)\\
&\quad + h^{2r}\left(\|u(t_0^+)\|^2_{H^{r+1}(\Omega)} + \|u(t_N^-)\|^2_{H^{r+1}(\Omega)}\right)\Bigg\}^{\frac{1}{2}}. \qquad (2.3.21)\end{aligned}$$

□

Remark 2.21. *If we use uniform time intervals $k_n = k$, and uniform polynomial degrees $q_n = q \ge 2, s_n = s \ge 2$ for $n = 1, \ldots, N$, then the error bound becomes*

$$|||u - u_{\mathrm{DG}}||| \le C(u)\left(h^{2r} + \frac{k^{2q-1}}{q^{2s-6}}\right)^{\frac{1}{2}},$$

where $C(u)$ is a positive constant depending on u and $1 \le r \le \min(p, m-1)$.

Remark 2.22. *If we only consider the L^2 error estimate at the end point, we can deduce from the proof of Theorem 2.20 that*

$$\begin{aligned}\|\dot{u}(t_N^-) - \dot{u}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)} &\le \|\dot{\theta}(t_N^-)\|_{L^2(\Omega)} + \|\dot{\rho}_1(t_N^-)\|_{L^2(\Omega)} + \|\dot{\rho}_2(t_N^-)\|_{L^2(\Omega)}\\
&\le C\left(h^{2(r+1)} + \frac{k^{2q}}{q^{2s-2}}\right)^{\frac{1}{2}}\|u\|_{H^s(0,T;H^{r+1}(\Omega))} + Ch^{r+1}\|\dot{u}(t_N^-)\|_{H^{r+1}(\Omega)}\end{aligned}$$

$$\leq C(u) \left(h^{2(r+1)} + \frac{k^{2q}}{q^{2s-2}} \right)^{\frac{1}{2}}.$$

Analogously,

$$\|\dot{u}(t_N^-) - \dot{u}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)} + \|u(t_N^-) - u_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)} \leq C(u) \left(h^{2(r+1)} + \frac{k^{2q}}{q^{2s-2}} \right)^{\frac{1}{2}}.$$

2.4. Numerical experiments

In this section, we show some numerical results to verify the convergence properties of the DG-in-time scheme introduced in Section 2.3.

2.4.1. Numerical results for a scalar linear wave equation We consider the one-dimensional wave problem

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

Here, we set $T = 1$, $\gamma = 1$ and let u_0, u_1 and f be chosen such that the exact solution is $u(x,t) = \sin(\sqrt{2}\pi t)\sin(\pi x)$. That is, $u_0(x) \equiv 0$, $u_1(x) = \sqrt{2}\pi \sin(\pi x)$, and

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

We first discretise the problem in the spatial direction using continuous piecewise polynomials of degree $p \geq 1$, and compute the corresponding mass and stiffness matrices using FEniCS. Let $\mathcal{V}_h$ be the finite element function space with h being the spatial discretisation parameter specified previously. The numerical approximation of the one-dimensional wave-type equation is the following: find $u_h \in \mathcal{V}_h$ such that

$$\int_\Omega \ddot{u}_h \cdot v_h \,\mathrm{d}x + \int_\Omega 2\gamma\dot{u}_h \cdot v_h \,\mathrm{d}x + \int_\Omega \gamma^2 u_h \cdot v_h \,\mathrm{d}x + \int_\Omega \partial_x u_h \cdot \partial_x v_h \,\mathrm{d}x = \int_\Omega f \cdot v_h \,\mathrm{d}x,$$

for all $v_h \in \mathcal{V}_h$. After discretisation in space based on the above weak formulation, the problem results in the following second-order differential system for the nodal displacement $\mathbf{U}(t)$:

$$\begin{cases} \tilde{M}\ddot{\mathbf{U}}(t) + 2\gamma\tilde{M}\dot{\mathbf{U}}(t) + \gamma^2\tilde{M}\mathbf{U}(t) + \tilde{K}\mathbf{U}(t) = \mathbf{F}(t), \quad t \in (0,T], \\ \dot{\mathbf{U}}(0) = \mathbf{U}_1, \quad \mathbf{U}(0) = \mathbf{U}_0, \end{cases} \tag{2.4.1}$$

where $\ddot{\mathbf{U}}(t)$ (respectively $\dot{\mathbf{U}}(t)$) represents the vector of nodal acceleration (respectively velocity) and $\mathbf{F}(t)$ is the vector of externally applied loads. Here, $\tilde{M}$ and $\tilde{K}$ are the mass and stiffness matrices (with a Dirichlet boundary condition applied) whose entries are, respectively,

$$\tilde{M}_{ij} := \int_0^1 \psi_i(x) \cdot \psi_j(x)\,\mathrm{d}x, \quad \tilde{K}_{ij} := \int_0^1 \partial_x\psi_i(x) \cdot \partial_x\psi_j(x)\,\mathrm{d}x,$$

where $\{\psi_i\}_{i=1}^{\hat{d}}$ are the basis functions in the spatial direction. Multiplying the above algebraic formulation by $\tilde{M}^{-\frac{1}{2}}$ and setting $\mathbf{Z}(t) = \tilde{M}^{\frac{1}{2}}\mathbf{U}(t)$, we obtain

$$\ddot{\mathbf{Z}}(t) + L\dot{\mathbf{Z}}(t) + K\mathbf{Z}(t) = \mathbf{G}(t), \quad t \in (0,T], \tag{2.4.2}$$

$$\dot{\mathbf{Z}}(0) = \tilde{M}^{\frac{1}{2}}\mathbf{U}_1, \quad \mathbf{Z}(0) = \tilde{M}^{\frac{1}{2}}\mathbf{U}_0, \tag{2.4.3}$$

where $\mathbf{U}_0 = [0,\ldots,0]^{\mathrm{T}} \in \mathbb{R}^{\hat{d}}$ and $\mathbf{U}_1$ is the $\hat{d}$-vector corresponding to u_1 at the grid points. Here $L = 2\gamma\mathrm{Id}$, $K = \gamma^2\mathrm{Id} + \tilde{M}^{-\frac{1}{2}}\tilde{K}\tilde{M}^{-\frac{1}{2}}$ and $\mathbf{G}(t) = \tilde{M}^{-\frac{1}{2}}\mathbf{F}(t)$. Note that $\tilde{M}^{-\frac{1}{2}}$ can be easily computed using `scipy.linalg.fractional_matrix_power` in Python. We subdivided $[0,T)$ with $T = 1$ into N sub-intervals 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, $q_1 = \cdots = q_N = q \geq 2$. 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 follows: find $\mathbf{Z} \in \mathcal{V}_{kh}^{q_n}$ such that

$$\left(\ddot{\mathbf{Z}}(t), \dot{\mathbf{v}}\right)_{L^2(I_n)} + \left(L\dot{\mathbf{Z}}(t), \dot{\mathbf{v}}\right)_{L^2(I_n)} + (K\mathbf{Z}(t), \dot{\mathbf{v}})_{L^2(I_n)} + \dot{\mathbf{Z}}(t_{n-1}^+) \cdot \dot{\mathbf{v}}(t_{n-1}^+)$$

$$+ K\mathbf{Z}(t_{n-1}^{+}) \cdot \mathbf{v}(t_{n-1}^{+}) = (\mathbf{G}(t), \dot{\mathbf{v}})_{L^2(I_n)} + \dot{\mathbf{Z}}(t_{n-1}^{-}) \cdot \dot{\mathbf{v}}(t_{n-1}^{+}) + K\mathbf{Z}(t_{n-1}^{-}) \cdot \mathbf{v}(t_{n-1}^{+}), \tag{2.4.4}$$

for all $\mathbf{v} \in \mathcal{V}_{kh}^{q_n}$, where on the right-hand side the values $\dot{\mathbf{Z}}(t_{n-1}^{-})$ and $\mathbf{Z}(t_{n-1}^{-})$ computed for I_{n-1} are used as initial conditions for the current time interval. For I_1, we set $\dot{\mathbf{Z}}(t_0^{-}) = \dot{\mathbf{Z}}(0)$ and $\mathbf{Z}(t_0^{-}) = \mathbf{Z}(0)$. Before we start the algebraic formulation of this problem, we need to show that equation (2.4.4) is equivalent to the DG formulation of our original problem. After semi-discretisation in spatial variables, we are left with equation (2.4.1). On each time interval, our DG formulation is

$$\begin{aligned}
&\left(\tilde{M}\ddot{\mathbf{U}}(t), \dot{\boldsymbol{\varphi}}\right)_{L^2(I_n)} + 2\gamma \left(\tilde{M}\dot{\mathbf{U}}(t), \dot{\boldsymbol{\varphi}}\right)_{L^2(I_n)} + \gamma^2 \left(\tilde{M}\mathbf{U}(t), \dot{\boldsymbol{\varphi}}\right)_{L^2(I_n)} + \left(\tilde{K}\mathbf{U}(t), \dot{\boldsymbol{\varphi}}\right)_{L^2(I_n)} \\
&\quad + \tilde{M}\dot{\mathbf{U}}(t_{n-1}^{+}) \cdot \dot{\boldsymbol{\varphi}}(t_{n-1}^{+}) + \gamma^2 \tilde{M}\mathbf{U}(t_{n-1}^{+}) \cdot \boldsymbol{\varphi}(t_{n-1}^{+}) + \tilde{K}\mathbf{U}(t_{n-1}^{+}) \cdot \boldsymbol{\varphi}(t_{n-1}^{+}) \\
&= (\mathbf{F}(t), \dot{\boldsymbol{\varphi}})_{L^2(I_n)} + \tilde{M}\dot{\mathbf{U}}(t_{n-1}^{-}) \cdot \dot{\boldsymbol{\varphi}}(t_{n-1}^{+}) + \gamma^2 \tilde{M}\mathbf{U}(t_{n-1}^{-}) \cdot \boldsymbol{\varphi}(t_{n-1}^{+}) + \tilde{K}\mathbf{U}(t_{n-1}^{-}) \cdot \boldsymbol{\varphi}(t_{n-1}^{+}),
\end{aligned}$$

for all $\boldsymbol{\varphi} \in [\mathcal{V}_h]^{\hat{d}}$. Since both $\tilde{M}$ and $\tilde{K}$ are positive definite and symmetric matrices, we can replace $\tilde{M}^{\frac{1}{2}}\mathbf{U}$ by $\mathbf{Z}$ to have

$$\begin{aligned}
&\left(\ddot{\mathbf{Z}}(t), \tilde{M}^{\frac{1}{2}}\dot{\boldsymbol{\varphi}}\right)_{L^2(I_n)} + 2\gamma \left(\dot{\mathbf{Z}}(t), \tilde{M}^{\frac{1}{2}}\dot{\boldsymbol{\varphi}}\right)_{L^2(I_n)} + \left(\gamma^2\mathbf{Z}(t) + \tilde{M}^{-\frac{1}{2}}\tilde{K}\tilde{M}^{-\frac{1}{2}}\mathbf{Z}(t), \tilde{M}^{\frac{1}{2}}\dot{\boldsymbol{\varphi}}\right)_{L^2(I_n)} \\
&\quad + \dot{\mathbf{Z}}(t_{n-1}^{+}) \cdot \tilde{M}^{\frac{1}{2}}\dot{\boldsymbol{\varphi}}(t_{n-1}^{+}) + \gamma^2\mathbf{Z}(t_{n-1}^{+}) \cdot \tilde{M}^{\frac{1}{2}}\boldsymbol{\varphi}(t_{n-1}^{+}) + \tilde{M}^{-\frac{1}{2}}\tilde{K}\tilde{M}^{-\frac{1}{2}}\mathbf{Z}(t_{n-1}^{+}) \cdot \tilde{M}^{\frac{1}{2}}\boldsymbol{\varphi}(t_{n-1}^{+}) \\
&= \left(\tilde{M}^{-\frac{1}{2}}\mathbf{F}(t), \tilde{M}^{\frac{1}{2}}\dot{\boldsymbol{\varphi}}\right)_{L^2(I_n)} + \dot{\mathbf{Z}}(t_{n-1}^{-}) \cdot \tilde{M}^{\frac{1}{2}}\dot{\boldsymbol{\varphi}}(t_{n-1}^{+}) + \gamma^2\mathbf{Z}(t_{n-1}^{-}) \cdot \tilde{M}^{\frac{1}{2}}\boldsymbol{\varphi}(t_{n-1}^{+}) \\
&\quad + \tilde{M}^{-\frac{1}{2}}\tilde{K}\tilde{M}^{-\frac{1}{2}}\mathbf{Z}(t_{n-1}^{-}) \cdot \tilde{M}^{\frac{1}{2}}\boldsymbol{\varphi}(t_{n-1}^{+}),
\end{aligned} \tag{2.4.5}$$

for all $\boldsymbol{\varphi} \in [\mathcal{V}_h]^{\hat{d}}$. By substituting $\mathbf{G}(t) = \tilde{M}^{-\frac{1}{2}}\mathbf{F}(t)$, $K = \gamma^2\mathrm{Id} + \tilde{M}^{-\frac{1}{2}}\tilde{K}\tilde{M}^{-\frac{1}{2}}$ and $\mathbf{v} = \tilde{M}^{-\frac{1}{2}}\boldsymbol{\varphi}$, we obtain equation (2.4.4). Here K is also a symmetric and positive definite matrix. From now on, we will consider the algebraic formulation of equation (2.4.4) only. 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 $\mathcal{V}_{kh}^{q_n}$. We also introduce the vectorial basis $\{\Phi_m^j(t)\}_{m=1,\ldots,\hat{d}}^{j=1,\ldots,q_n+1}$, 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

$$\mathbf{Z}(t) = \sum_{m=1}^{\hat{d}} \sum_{j=1}^{q_n+1} \alpha_m^j \Phi_m^j(t), \tag{2.4.6}$$

where $\alpha_m^j \in \mathbb{R}$ for $m = 1, \ldots, \hat{d}$, $j = 1, \ldots, q_n + 1$. By choosing $\mathbf{v}(t) = \Phi_m^j(t)$ for each $m = 1, \ldots, \hat{d}$, $j = 1, \ldots, q_n + 1$, we obtain the following algebraic system

$$\mathbf{A}\mathbf{z} = \mathbf{b}, \tag{2.4.7}$$

where $\mathbf{z} \in \mathbb{R}^D = \mathbb{R}^{(q_n+1)\hat{d}}$ is the solution vector (values for α_m^j). Here $\mathbf{b} \in \mathbb{R}^D$ corresponds to the right-hand side, which is given componentwise as

$$\mathbf{b}_m^j = \Big(\mathbf{G}(t), \dot{\Phi}_m^j\Big)_{L^2(I_n)} + \dot{\mathbf{Z}}(t_{n-1}^-) \cdot \dot{\Phi}_m^j(t_{n-1}^+) + K\mathbf{Z}(t_{n-1}^-) \cdot \Phi_m^j(t_{n-1}^+), \tag{2.4.8}$$

for $m = 1, \ldots, \hat{d}$, $j = 1, \ldots, q_n+1$. $\mathbf{A}$ is the local stiffness matrix of size $(q_n+1)\hat{d} \times (q_n+1)\hat{d}$ with its structure being discussed below. For $l, j = 1, \ldots, q_n + 1$,

$$\begin{aligned}
M_{lj}^1 &= \Big(\ddot{\phi}^j, \dot{\phi}^l\Big)_{L^2(I_n)}, \quad M_{lj}^2 = \Big(\dot{\phi}^j, \dot{\phi}^l\Big)_{L^2(I_n)}, \quad M_{lj}^3 = \Big(\phi^j, \dot{\phi}^l\Big)_{L^2(I_n)}, \\
M_{lj}^4 &= \dot{\phi}^j(t_{n-1}^+) \cdot \dot{\phi}^l(t_{n-1}^+), \quad M_{lj}^5 = \phi^j(t_{n-1}^+) \cdot \phi^l(t_{n-1}^+).
\end{aligned}$$

Setting

$$M = M^1 + M^4, \quad B_{ij} = L_{ij}M^2 + K_{ij}(M^3 + M^5),$$

with $M, B_{ij} \in \mathbb{R}^{(q_n+1)\times(q_n+1)}$ for any $i, j = 1, \ldots, \hat{d}$, we can write the matrix $\mathbf{A}$ as

$$\mathbf{A} = \begin{bmatrix} M & 0 & 0 & \cdots & 0 \\ 0 & M & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & M \end{bmatrix} + \begin{bmatrix} B_{1,1} & B_{1,2} & \cdots & B_{1,\hat{d}} \\ B_{2,1} & B_{2,2} & \cdots & B_{2,\hat{d}} \\ \vdots & \ddots & \ddots & \vdots \\ B_{\hat{d},1} & B_{\hat{d},2} & \cdots & B_{\hat{d},\hat{d}} \end{bmatrix}.$$

For each time interval $I_n = (t_{n-1}, t_n]$, we use the following shifted and scaled Legendre polynomials $\{\phi_i\}$ as the basis polynomials.

$$\phi^1(t) = 1, \quad \phi^2(t) = \frac{2(t - t_{n-1}^+)}{k_n} - 1, \quad \phi^3(t) = \frac{6(t - t_{n-1}^+)^2}{k_n^2} - \frac{6(t - t_{n-1}^+)}{k_n} + 1,$$
$$\phi^4(t) = \frac{20(t - t_{n-1}^+)^3}{k_n^3} - \frac{30(t - t_{n-1}^+)^2}{k_n^2} + \frac{12(t - t_{n-1}^+)}{k_n} - 1,$$
$$\phi^5(t) = \frac{70(t - t_{n-1}^+)^4}{k_n^4} - \frac{140(t - t_{n-1}^+)^3}{k_n^3} + \frac{90(t - t_{n-1}^+)^2}{k_n^2} - \frac{20(t - t_{n-1}^+)}{k_n} + 1,$$
$$\phi^6(t) = \frac{252(t - t_{n-1}^+)^5}{k_n^5} - \frac{630(t - t_{n-1}^+)^4}{k_n^4} + \frac{560(t - t_{n-1}^+)^3}{k_n^3} - \frac{210(t - t_{n-1}^+)^2}{k_n^2} + \frac{30(t - t_{n-1}^+)}{k_n} - 1.$$

This implies that

$$\dot{\phi}^1(t) = 0, \quad \dot{\phi}^2(t) = \frac{2}{k_n}, \quad \dot{\phi}^3(t) = \frac{12(t - t_{n-1}^+)}{k_n^2} - \frac{6}{k_n},$$
$$\dot{\phi}^4(t) = \frac{60(t - t_{n-1}^+)^2}{k_n^3} - \frac{60(t - t_{n-1}^+)}{k_n^2} + \frac{12}{k_n},$$
$$\dot{\phi}^5(t) = \frac{280(t - t_{n-1}^+)^3}{k_n^4} - \frac{420(t - t_{n-1}^+)^2}{k_n^3} + \frac{180(t - t_{n-1}^+)}{k_n^2} - \frac{20}{k_n},$$
$$\dot{\phi}^6(t) = \frac{1260(t - t_{n-1}^+)^4}{k_n^5} - \frac{2520(t - t_{n-1}^+)^3}{k_n^4} + \frac{1680(t - t_{n-1}^+)^2}{k_n^3} - \frac{420(t - t_{n-1}^+)}{k_n^2} + \frac{30}{k_n},$$

and

$$\ddot{\phi}^1(t) = 0, \quad \ddot{\phi}^2(t) = 0, \quad \ddot{\phi}^3(t) = \frac{12}{k_n^2}, \quad \ddot{\phi}^4(t) = \frac{120(t - t_{n-1}^+)}{k_n^3} - \frac{60}{k_n^2},$$
$$\ddot{\phi}^5(t) = \frac{840(t - t_{n-1}^+)^2}{k_n^4} - \frac{840(t - t_{n-1}^+)}{k_n^3} + \frac{180}{k_n^2},$$

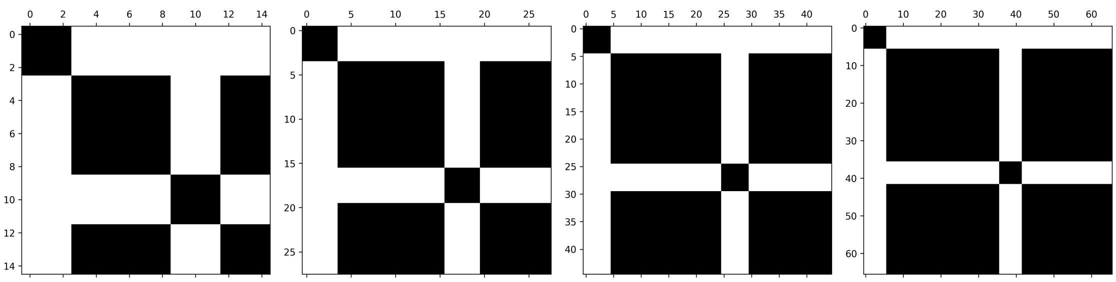

Figure 2.2: Sparsity pattern of the matrix $\mathbf{A}$ for $p = q_n = 2, 3, 4, 5$ from left to right, where p is the polynomial degree for the spatial discretisation.

$$\ddot{\phi}^6(t) = \frac{5040(t-t_{n-1}^+)^3}{k_n^5} - \frac{7560(t-t_{n-1}^+)^2}{k_n^4} + \frac{3360(t-t_{n-1}^+)}{k_n^3} - \frac{420}{k_n^2}.$$

We compute $M_1, \ldots, M_5$ according to the above-mentioned formulae for $q_n = 2,3,4,5$ respectively. The corresponding matrix $\mathbf{A}$ has a symmetric density structure and its sparsity depends on the polynomial degree of the fully discrete finite element space $\mathcal{V}_{kh}^{q_n}$ for each time step. We plot the sparsity of $\mathbf{A}$ in Figure 2.2 for $q_n = 2,3,4,5$ with CG–p (where $p = q_n$) elements in space. It shows that the matrix gets less sparse as the polynomial degree increases. The linear system $\mathbf{Az} = \mathbf{b}$ is solved using `numpy.linalg.solve` in Python.

Table 2.1: 1D linear wave equation with $p = q$: computed errors $|||\mathbf{Z} - \mathbf{Z}_{\mathrm{DG}}|||$, $||\dot{u}(T) - \dot{u}_{\mathrm{DG}}(t_N^-)||_{L^2(\Omega)}$ and corresponding convergence rates with respect to polynomial degrees $q = 2,3,4,5$.

q	k	energy-norm error	rate	L^2-error	rate
2	5.000e − 1	1.6104e − 0	—	2.7764e − 2	—
	2.500e − 1	6.3133e − 1	1.3509	3.1785e − 3	3.1268
	1.250e − 1	2.3095e − 1	1.4508	2.9815e − 4	3.4142
	6.250e − 2	8.2525e − 2	1.4847	3.8145e − 5	2.9665
3	5.000e − 1	3.3107e − 1	—	9.2796e − 3	—
	2.500e − 1	6.6906e − 2	2.3069	3.4250e − 4	4.7599
	1.250e − 1	1.2170e − 2	2.4588	1.4774e − 5	4.5350
	6.250e − 2	2.1682e − 3	2.4888	7.6664e − 7	4.2684
4	5.000e − 1	6.3773e − 2	—	2.4916e − 4	—
	2.500e − 1	5.7582e − 3	3.4693	4.3113e − 6	5.8528
	1.250e − 1	5.1473e − 4	3.4837	1.1657e − 7	5.2089
	6.250e − 2	4.5654e − 5	3.4950	3.8021e − 9	4.9383
5	5.000e − 1	6.8152e − 3	—	1.3549e − 5	—
	2.500e − 1	3.3249e − 4	4.3574	2.8591e − 7	5.5665
	1.250e − 1	1.4968e − 5	4.4734	4.9007e − 9	5.8664
	6.250e − 2	6.6462e − 7	4.4932	7.7459e − 11	5.9834

Here we use CG–p elements (continuous piecewise polynomials of degree p) in space with $h = k$, $T = 1$ and $\gamma = 1$, and compute the errors $|||\mathbf{Z} - \mathbf{Z}_{\mathrm{DG}}|||$ and $\|\dot{u}(T) - \dot{u}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)}$ versus k for $k = 2^{-l}$, $l = 1,2,3,4$, with respect to polynomial degrees $q = 2,3,4,5$. We first choose $p = q$, i.e., the polynomial degree in the spatial direction is the same as the polynomial degree in the time direction. Both the semi-discrete errors in the energy norm and the fully discrete errors at end nodal points in the L^2-norm are reported in Table 2.1 and plotted in a log-log scale in Figure 2.3 and Figure 2.4. As expected, the convergence rate in the energy norm is of order $O(k^{q-\frac{1}{2}})$, which is consistent with Remark 2.18. For the fully discrete errors in the L^2-norm, the numerical experiments show a better convergence rate of $O(k^{q+1})$ rather than $O(k^q)$ (cf. Remark 2.22). This hints that our error estimates derived at the nodal points are not optimal. In fact, an error estimate of $O(h^{p+1} + k^{2q-1})$ at the nodal points was proved by Eriksson *et al.* [34] for the linear parabolic equation. It is thus possible to derive similar error estimates for the linear wave equation. To verify our conjecture, we test with $p = 2q - 2$ and $p = 2q - 1$ respectively (cf. Table 2.2 and Table 2.3). Both tests show a convergence rate of $O(k^{2q-1})$, which is consistent with our conjecture. Note that for the case $p = 2q - 1$, with $h = k$, the error estimate is still of $O(k^{2q-1})$ because $O(k^{2q-1})$ dominates $O(h^{2q})$.

Remark 2.23. *Note that $\gamma > 0$ (i.e., L is a positive definite differential operator) ensures the stability of our numerical scheme. If we set $\gamma = 0$ and consider the following wave equation*

$$\ddot{u}(x,t) - \partial_{xx}u(x,t) = f(x,t) \qquad \text{in } (0,1) \times (0,T],$$

with

$$f(x,t) = -\pi^2 \sin(\sqrt{2}\pi t)\sin(\pi x),$$

similar convergence rates are observed. This suggests that our numerical scheme still works even if the coercivity of the bilinear form does not induce a norm.

Table 2.2: 1D linear wave equation with $p = 2q - 2$: computed errors $|||\mathbf{Z} - \mathbf{Z}_{\mathrm{DG}}|||$, $||\dot{u}(T) - \dot{u}_{\mathrm{DG}}(t_N^-)||_{L^2(\Omega)}$ and corresponding convergence rates with respect to polynomial degrees $q = 2, 3, 4$.

q	k	energy-norm error	rate	L^2-error	rate
2	5.000e − 1	1.6104e − 0	—	2.7764e − 2	—
	2.500e − 1	6.3133e − 1	1.3509	3.1785e − 3	3.1268
	1.250e − 1	2.3095e − 1	1.4508	2.9815e − 4	3.4142
	6.250e − 2	8.2525e − 2	1.4847	3.8145e − 5	2.9665
3	5.000e − 1	3.3067e − 1	—	8.4203e − 3	—
	2.500e − 1	6.6901e − 2	2.3053	2.5352e − 4	5.0537
	1.250e − 1	1.2169e − 2	2.4588	7.4658e − 6	5.0857
	6.250e − 2	2.1682e − 3	2.4886	2.2401e − 7	5.0587
4	5.000e − 1	6.3765e − 2	—	1.5878e − 4	—
	2.500e − 1	5.7580e − 3	3.4691	1.2464e − 6	6.9931
	1.250e − 1	5.1472e − 4	3.4837	9.5992e − 9	7.0206
	6.250e − 2	4.5654e − 5	3.4950	7.2384e − 11	7.0511

Table 2.3: 1D linear wave equation with $p = 2q - 1$: computed errors $|||\mathbf{Z} - \mathbf{Z}_{\mathrm{DG}}|||$, $||\dot{u}(T) - \dot{u}_{\mathrm{DG}}(t_N^-)||_{L^2(\Omega)}$ and corresponding convergence rates with respect to polynomial degrees $q = 2, 3, 4$.

q	k	energy-norm error	rate	L^2-error	rate
2	5.000e − 1	1.6125e − 0	—	9.4398e − 2	—
	2.500e − 1	6.3137e − 1	1.3527	1.3508e − 2	2.8049
	1.250e − 1	2.3095e − 1	1.4509	1.7750e − 3	2.9279
	6.250e − 2	8.2525e − 2	1.4847	2.2554e − 4	2.9763
3	5.000e − 1	3.3075e − 1	—	2.1981e − 3	—
	2.500e − 1	6.6902e − 2	2.3056	9.6398e − 5	4.5111
	1.250e − 1	1.2169e − 2	2.4588	3.2432e − 6	4.8935
	6.250e − 2	2.1682e − 3	2.4886	1.0376e − 7	4.9661
4	5.000e − 1	6.3765e − 2	—	7.9749e − 5	—
	2.500e − 1	5.7580e − 3	3.4691	6.0876e − 7	7.0334
	1.250e − 1	5.1472e − 4	3.4837	4.8693e − 9	6.9660
	6.250e − 2	4.5654e − 5	3.4950	3.9113e − 11	6.9599

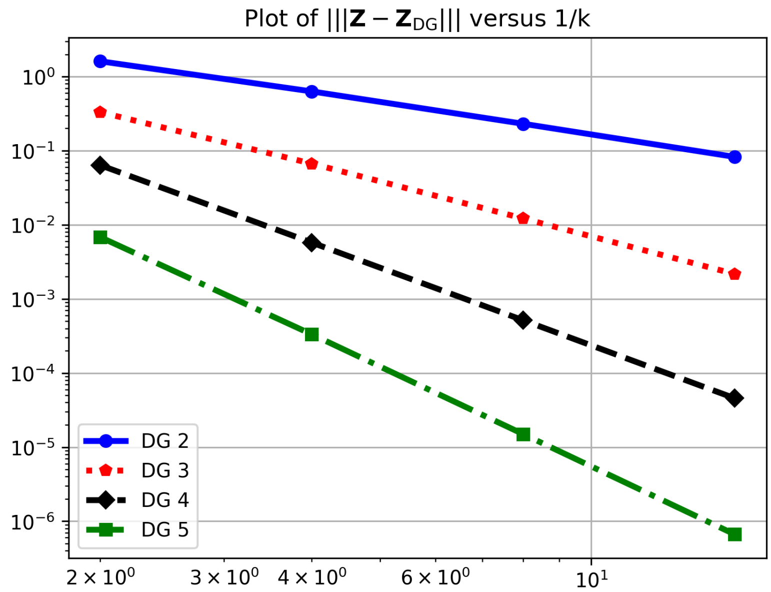

Figure 2.3: 1D linear wave equation with $p = q$: computed errors $|||\mathbf{Z} - \mathbf{Z}_{\mathrm{DG}}|||$ plotted against $1/k$ for $q = 2, 3, 4, 5$ in a log-log scale.

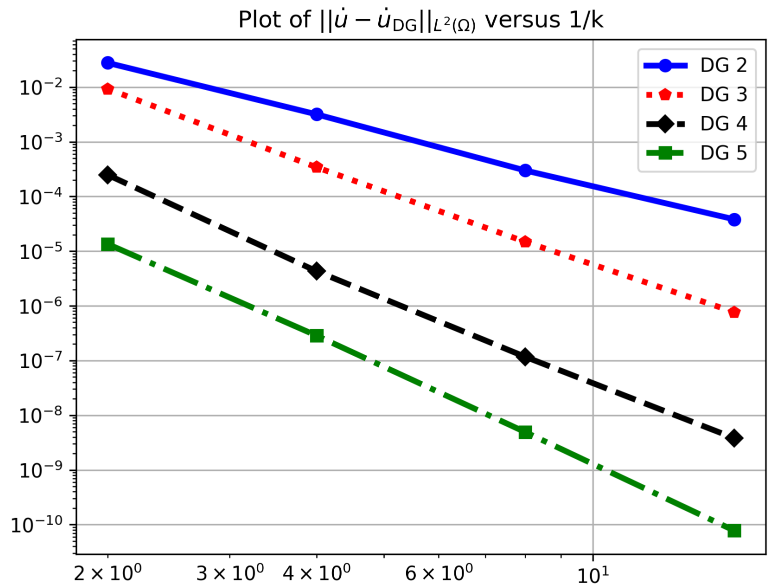

Figure 2.4: 1D linear wave equation with $p = q$: computed errors $\|\dot{u}(T) - \dot{u}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)}$ plotted against $1/k$ for $q = 2, 3, 4, 5$ in a log-log scale.

Table 2.4: 1D linear wave equation with the generalised–α and the Newmark–β methods: computed errors $||\dot{u}(T)-\dot{u}_{\mathrm{DG}}(t_N^-)||_{L^2(\Omega)}$ and corresponding convergence rates with respect to polynomial degrees $p = 2, 3, 4$.

p	$k = h$	Newmark–β error	rate	generalised–α error	rate
2	5.000e − 1	1.2325e − 0	—	1.2283e − 0	—
	2.500e − 1	5.1116e − 1	1.2697	5.4672e − 1	1.1678
	1.250e − 1	1.4317e − 1	1.8360	1.5308e − 1	1.8365
	6.250e − 2	3.6778e − 2	1.9608	3.9282e − 2	1.9623
3	5.000e − 1	1.2455e − 0	—	1.2411e − 0	—
	2.500e − 1	5.1247e − 1	1.2812	5.4793e − 1	1.1796
	1.250e − 1	1.4327e − 1	1.8387	1.5316e − 1	1.8390
	6.250e − 2	3.6784e − 2	1.9616	3.9288e − 2	1.9629
4	5.000e − 1	1.2438e − 0	—	1.2397e − 0	—
	2.500e − 1	5.1238e − 1	1.2795	5.4785e − 1	1.1781
	1.250e − 1	1.4326e − 1	1.8386	1.5316e − 1	1.8387
	6.250e − 2	3.6784e − 2	1.9615	3.9288e − 2	1.9629

Table 2.5: 1D linear wave equation with $p = 2$: comparison of computational time between the DG method and the generalised–α method

q	k	single time step	total time	L^2-error
DG method	1.250e − 1	5.82e − 2s	7.06e − 1s	2.9815e − 4
Generalised α	1.250e − 1	2.26e − 2s	3.67e − 1s	1.5308e − 1

We also compare our proposed DG method with the generalised–α method [35] (with $\alpha_m = 0.2$ and $\alpha_f = 0.4$) and the Newmark–β method (i.e., the generalised–α method with $\alpha_m = \alpha_f = 0$), which are second-order accurate in the temporal domain (cf. Table 2.4 and Table 2.5). As expected, our DG method outperforms these two finite difference schemes in terms of convergence rate and accuracy. Though each time step is more expensive in our approach because a linear system needs to be solved for each time step, this can be balanced using a larger time step.

2.4.2. Numerical results for a two-dimensional elastodynamics problem Now we consider a two-dimensional linearised elastodynamics problem. For $T > 0$, find $\mathbf{u}\colon \Omega \times [0, T] \to \mathbb{R}^2$ such that

$$\begin{cases} \rho\ddot{\mathbf{u}} + 2\rho\gamma\dot{\mathbf{u}} + \rho\gamma^2\mathbf{u} - \nabla\cdot\boldsymbol{\sigma} = \mathbf{f} & \text{in } \Omega \times (0, T], \\ \mathbf{u} = 0 & \text{on } \partial\Omega \times (0, T], \\ \mathbf{u}(x, 0) = \mathbf{u}_0(x), \quad \dot{\mathbf{u}}(x, 0) = \mathbf{u}_1(x) & \text{in } \Omega. \end{cases}$$

Here $\mathbf{f} \in L^2(0, T, L^2(\Omega))$ is the source term, and $\rho \in L^\infty(\Omega)$ is such that $\rho = \rho(\mathbf{x}) > 0$ for almost any $\mathbf{x} \in \Omega$. The stress tensor $\boldsymbol{\sigma}$ is defined through Hooke's law, that is, $\boldsymbol{\sigma} = 2\mu\boldsymbol{\varepsilon} + \lambda\mathrm{tr}(\boldsymbol{\varepsilon})\mathrm{Id}$, where Id is the identity matrix, tr is the trace operator, and

$$\boldsymbol{\varepsilon}(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^T).$$

Note that we assume $\partial\Omega \in C^2$ in the problem set up (cf. equation 2.0.1) to apply the duality argument in the convergence analysis (cf. Section 2.3). In fact, it is sufficient to assume that Ω is a convex polygon for $d = 2$. We consider $\Omega = (0, 1) \times (0, 1)$, $T = 1$, and set $\rho = 1, \gamma = 1, \lambda = 1$ and $\mu = 1$. Here we choose $\mathbf{u}_0$, $\mathbf{u}_1$ and $\mathbf{f}$ such that the exact solution is

$$\mathbf{u}(x, t) = \sin(\sqrt{2}\pi t)\begin{bmatrix} -\sin^2(\pi x)\sin(2\pi y) \\ \sin(2\pi x)\sin^2(\pi y) \end{bmatrix}.$$

We discretise the linearised elastodynamics equations in the same manner as in Section 2.4.1. Again, we choose $p = q$, $h = k$ and compute the errors $|||\mathbf{Z} - \mathbf{Z}_{\mathrm{DG}}|||$ and $\|\mathbf{u}(T) - \mathbf{u}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)} + \|\dot{\mathbf{u}}(T) - \dot{\mathbf{u}}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)}$ versus k for $k = 0.500, 0.250, 0.125$ and 0.100 with respect to polynomial degrees $q = 2, 3, 4$. Note that here we use $k = h = 0.100$ instead of 0.0625 for the smallest time step; this is due to the computational limits of FEniCs for the high order approximation of non-scalar problems. We choose the last step to be 0.100 to ensure that we still have sufficient data to compute the convergence rates. Both the semi–discrete errors in the energy norm and the fully discrete errors at nodal points in the L^2-norm are computed in Table 2.6 and plotted in a log-log scale in Figure 2.5 and Figure 2.6.

Table 2.6: 2D linearised elastodynamics problem with $p = q$: computed errors $|||\mathbf{Z} - \mathbf{Z}_{\mathrm{DG}}|||$, $||\mathbf{u}(T) - \mathbf{u}_{\mathrm{DG}}(t_N^-)||_{L^2(\Omega)} + ||\dot{\mathbf{u}}(T) - \dot{\mathbf{u}}_{\mathrm{DG}}(t_N^-)||_{L^2(\Omega)}$ and corresponding convergence rates with respect to polynomial degrees $q = 2, 3, 4$.

q	k	energy-norm error	rate	L^2-error	rate
2	5.000e − 1	1.2668e − 0	—	7.2172e − 1	—
	2.500e − 1	5.2863e − 1	1.2609	9.5802e − 2	2.9133
	1.250e − 1	1.9754e − 1	1.4201	1.2390e − 2	2.9509
	1.000e − 1	1.4262e − 2	1.4599	6.8663e − 3	2.6452
3	5.000e − 1	3.1328e − 1	—	1.3788e − 1	—
	2.500e − 1	6.0999e − 2	2.3606	1.2789e − 2	3.4304
	1.250e − 1	1.0548e − 2	2.5318	6.1569e − 4	4.3765
	1.000e − 1	6.0522e − 3	2.4895	2.4334e − 4	4.1600
4	5.000e − 1	1.5241e − 1	—	8.4535e − 2	—
	2.500e − 1	6.1539e − 3	4.6303	1.7324e − 3	5.6087
	1.250e − 1	4.5846e − 4	3.7466	5.4731e − 5	4.9843
	1.000e − 1	2.0732e − 4	3.5565	1.7987e − 5	4.9868

Table 2.7: 2D linearised elastodynamics problem with $p = 2q - 2$: computed errors $|||\mathbf{Z} - \mathbf{Z}_{\mathrm{DG}}|||$, $||\mathbf{u}(T) - \mathbf{u}_{\mathrm{DG}}(t_N^-)||_{L^2(\Omega)} + ||\dot{\mathbf{u}}(T) - \dot{\mathbf{u}}_{\mathrm{DG}}(t_N^-)||_{L^2(\Omega)}$ and corresponding convergence rates with respect to polynomial degrees $q = 2, 3, 4$.

q	k	energy-norm error	rate	L^2-error	rate
2	5.000e − 1	1.2668e − 0	—	7.2172e − 1	—
	2.500e − 1	5.2863e − 1	1.2609	9.5802e − 2	2.9133
	1.250e − 1	1.9754e − 1	1.4201	1.2390e − 2	2.9509
	1.000e − 1	1.4262e − 2	1.4599	6.8663e − 3	2.6452
3	5.000e − 1	2.9135e − 1	—	7.3446e − 2	—
	2.500e − 1	5.6740e − 2	2.3603	3.7267e − 3	4.3007
	1.250e − 1	1.0486e − 2	2.4359	1.1836e − 4	4.9766
	1.000e − 1	6.0356e − 3	2.4754	3.8835e − 5	4.9941
4	5.000e − 1	5.2822e − 2	—	6.7767e − 3	—
	2.500e − 1	4.9297e − 3	3.4216	4.4456e − 5	7.2521
	1.250e − 1	4.4454e − 4	3.4711	3.7959e − 7	6.8718

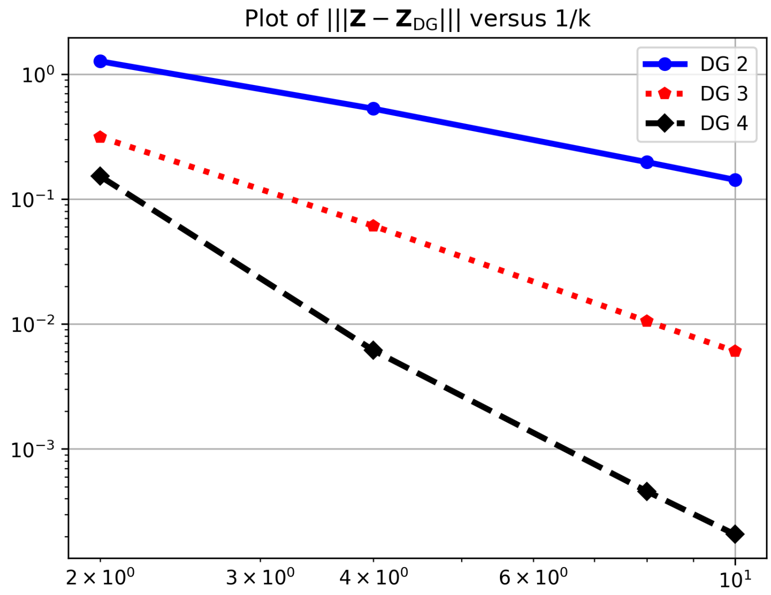

Figure 2.5: 2D linearised elastodynamics problem with $p = q$: computed errors $|||\mathbf{Z} - \mathbf{Z}_{\mathrm{DG}}|||$ plotted against $1/k$ for $q = 2, 3, 4$ in a log-log scale.

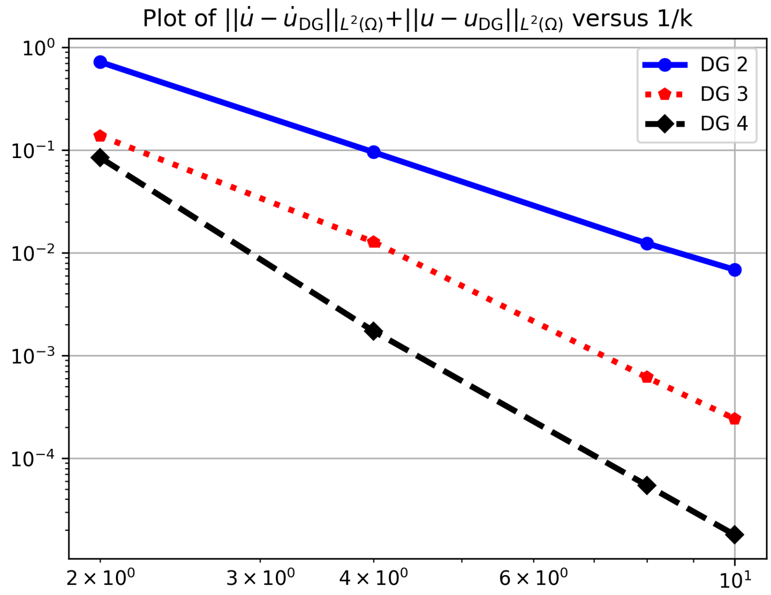

Figure 2.6: 2D linearised elastodynamics problem with $p = q$: computed errors $\|\mathbf{u}(T) - \mathbf{u}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)} + \|\dot{\mathbf{u}}(T) - \dot{\mathbf{u}}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)}$ plotted against $1/k$ for $q = 2, 3, 4$ in a log-log scale.

Table 2.8: 2D linearised elastodynamics problem with the generalised–α and the Newmark–β methods: computed errors $\|\mathbf{u}(T) - \mathbf{u}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)} + \|\dot{\mathbf{u}}(T) - \dot{\mathbf{u}}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)}$ and corresponding convergence rates with respect to polynomial degrees $p = 2, 3, 4$.

p	$k = h$	Newmark–β error	rate	generalised–α error	rate
2	5.000e − 1	1.7657e − 0	—	3.3111e − 0	—
	2.500e − 1	6.6997e − 1	1.3981	1.3432e − 0	1.3016
	1.250e − 1	1.5397e − 1	2.1214	2.9609e − 1	2.1816
	6.250e − 2	3.5884e − 2	2.1012	6.4175e − 2	2.2060
3	5.000e − 1	2.3486e − 0	—	4.2408e − 0	—
	2.500e − 1	8.0663e − 1	1.5418	1.5214e − 0	1.4789
	1.250e − 1	1.6678e − 1	2.2740	3.1034e − 1	2.2934
	6.250e − 2	3.6695e − 2	2.1843	6.5030e − 2	2.2547
4	5.000e − 1	2.3966e − 0	—	4.3113e − 0	—
	2.500e − 1	8.1573e − 1	1.5548	1.5310e − 0	1.4936
	1.250e − 1	1.6703e − 1	2.2880	3.1059e − 1	2.3014
	6.250e − 2	3.6703e − 2	2.1861	6.5038e − 2	2.2557

Table 2.9: 2D linearised elastodynamics equation with $p = 2$: comparison of computational time between the DG method and the generalised–α method.

q	k	single time step	total time	L^2-error
DG method	1.250e − 1	9.81e + 1s	1.04e + 3s	1.2390e − 2
Generalised α	1.250e − 1	3.86e − 1s	3.44e − 0s	2.9609e − 1

A convergence rate of $O(k^{q-\frac{1}{2}})$ for the energy norm is observed, in accordance with the theoretical result (cf. Remark 2.18). Again, the numerical experiments show a better nodal convergence rate in the L^2 norm of $O(k^{q+1})$ rather than $O(k^q)$. If we repeat the numerical tests with $p = 2q - 2$ and $h = k$ (cf. Table 2.7), we obtain a convergence rate of $O(k^{2q-1})$ as in the one-dimensional example. Note that for DG–4 elements, we did not compute the errors for the last time step $k = 0.100$ due to computational limit. In particular, each mass matrix M takes around 1.38GB storage space, and this means that the linear system $\mathbf{Az} = \mathbf{b}$ we are solving takes at least $1.38 \times (4 + 1)^2 = 34.5$GB physical memory. Thus, it is not within the capabilities of my computational resources.

Again, we compare the proposed DG method with the generalised–α [35] and the

Newmark–β [58] time integration schemes. For the two-dimensional example, our DG approach is much more expensive than the generalised-α scheme. Thus, it is necessary to introduce some preconditioning or parallelisation strategies to improve the efficiency of computations.

Chapter 3

Nonlinear elastodynamics equations

This chapter 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 to second-order hyperbolic equations. In particular, Kato [49] established the existence of solutions to the initial-boundary-value problem for quasilinear hyperbolic equations using semigroup theory. Building on Kato's work [49], Hughes, Kato and Marsden [43] analysed the existence, uniqueness and well-posedness of 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 [24] used energy methods to establish the 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 [18]. Concerning numerical approximations of such equations, Makridakis [56] proved optimal L^2 error estimates for both a semi-discrete and a class of fully discrete schemes for the elastodynamics problem. In [61], Ortner and Süli developed the convergence analysis of semidiscrete discontinuous Galerkin finite element approximations of second-order quasilinear hyperbolic systems. In this chapter, we will focus on the equations of nonlinear elastodynamics, though the ideas and techniques can be easily generalised to other second-order nonlinear hyperbolic equations, for instance, Einstein's equations, provided the assumptions on the nonlinearity assumed herein are satisfied.

We begin by formulating the time-dependent problem resulting from nonlinear elasticity. Consider a homogeneous elastic body with reference configuration Ω, a bounded domain in $\mathbb{R}^d$ for $d = 1, 2, 3$, with sufficiently smooth boundary $\partial\Omega$. Particles in this reference configuration are denoted by their coordinates x. A motion $\boldsymbol{\chi}(\cdot, t)\colon \Omega \to \mathbb{R}^n$ over the time interval $[0, T]$ for $0 < T < \infty$ is a mapping of the body to the current configuration, which satisfies the equations of balance of linear momentum

$$\ddot{\boldsymbol{\chi}}(x, t) = \mathrm{Div}\hat{S}(x, t) + \mathbf{f}(x, t) \text{ in } \Omega \times (0, T]. \tag{3.0.1}$$

Here $\boldsymbol{\chi}(x, t)$ is the position of the particle with reference position x at time t, $\hat{S}$ represents the Piola-Kirchhoff stress and $\mathbf{f}$ is the body force. Here the dot over $\boldsymbol{\chi}$ denotes differentiation with respect to time and Div is the divergence operator. The Piola-Kirchhoff stress $\hat{S}$ is determined from the deformation gradient $\nabla\boldsymbol{\chi}$ through the following constitutive relation

$$\hat{S}(x, t) = S(\nabla\boldsymbol{\chi}(x, t)), \tag{3.0.2}$$

where S is a given smooth $d \times d$ matrix-valued function defined on $\mathbb{R}^{d\times d}$, which characterises the material.

A typical problem of interest is to determine a motion of the elastic body when its boundary is held fixed, its initial configuration $\boldsymbol{\chi}(\cdot, 0)$ and initial velocity $\dot{\boldsymbol{\chi}}(\cdot, 0)$ are known, and the body force $\mathbf{f}$ is given. We consider the displacement field $\mathbf{u}\colon \Omega \times [0, T] \to$

$\mathbb{R}^n$ defined by

$$\mathbf{u}(x,t) := \boldsymbol{\chi}(x,t) - x. \tag{3.0.3}$$

For simplicity, we assume that the reference configuration is an actual configuration of the body, that is, we write $S(\nabla\mathbf{u})$ instead of $S(\nabla\mathbf{u}+\mathrm{Id})$. Then this displacement field $\mathbf{u}$ satisfies the following initial-boundary-value problem:

$$\ddot{u}_i(x,t) - \sum_{\alpha=1}^{d} \partial_\alpha S_{i\alpha}(\nabla\mathbf{u}(x,t)) = f_i(x,t) \text{ in } \Omega\times(0,T], \tag{3.0.4}$$

for each $i = 1,\dots,d$, where $\mathbf{u} = [u_1,\dots,u_d]^{\mathrm{T}}$ represents the displacement field and $\mathbf{f} = [f_1,\dots,f_d]^{\mathrm{T}}$ is the given body force which is sufficiently smooth, and

$$\mathbf{u}(x,t) = \mathbf{0} \text{ on } \partial\Omega\times(0,T], \tag{3.0.5}$$

$$\mathbf{u}(x,0) = \mathbf{u}_0(x) \in [H_0^1(\Omega)]^d \cap [H^m(\Omega)]^d,\ \ \dot{\mathbf{u}}(x,0) = \mathbf{u}_1(x) \in [H^{m-1}(\Omega)]^d, \tag{3.0.6}$$

are prescribed boundary and initial conditions, and m is an integer which will be specified later. Here and in (3.0.4) the dots over $\mathbf{u}$ denote differentiation with respect to time t, and ∂_α is the partial derivative with respect to x_α. For a complete discussion of the relevant mechanical background, we refer the reader to [5, 40].

For hyperelastic materials, S is the gradient of a scalar-valued 'stored energy function', that is,

$$S_{i\alpha}(\eta) := \frac{\partial}{\partial\eta_{i\alpha}} W(\eta),\ \text{ for } \eta \in \mathbb{R}^{d\times d}.$$

Let

$$A_{i\alpha j\beta}(\eta) := \frac{\partial}{\partial\eta_{j\beta}} S_{i\alpha}(\eta),\ \text{ for } \eta \in \mathbb{R}^{d\times d}.$$

Clearly, we have

$$A_{i\alpha j\beta} = A_{j\beta i\alpha},\ \text{ for } 1 \le i,\alpha,j,\beta \le d. \tag{S1a}$$

We assume that the *elasticities* $A_{i\alpha j\beta}$ satisfy the *strict Legendre–Hadamard condition*

$$\sum_{i,\alpha,\beta,j=1}^{d} A_{i\alpha j\beta}(\eta)\zeta_\alpha\zeta_\beta\xi_i\xi_j \ge M_0|\zeta|^2|\xi|^2 \quad \text{ for all } \eta\in\mathcal{O} \text{ and } \zeta,\xi\in\mathbb{R}^d, \tag{S1b}$$

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 $\mathbb{R}^d$.

The initial-boundary-value problem (3.0.4)–(3.0.6) does not have a global smooth solution as a result of breaking waves and shocks no matter how smooth $\mathbf{u}_0$, $\mathbf{u}_1$ and $\mathbf{f}$ are. It was proved by Dafermos and Hrusa [24] that there exists a unique local solution to the problem (3.0.4)–(3.0.6) provided that (S1a,b) are satisfied. We summarise this existence result in the following theorem.

Theorem 3.1. *Let Ω 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 $\mathbf{u}_0 \in [H^m(\Omega)]^d$ and $\mathbf{u}_1 \in [H^{m-1}(\Omega)]^d$ for some integer $m \geq [\frac{d}{2}] + 3$. Assume further that the initial values of the time derivatives of $\mathbf{u}$ up to order $m-1$ vanish on $\partial\Omega$ and that $\nabla\mathbf{u}_0[\overline{\Omega}] \subset \mathcal{O}$. Then, there exists a finite time $T > 0$ for which (3.0.4)–(3.0.6) has a unique solution $\mathbf{u}$ such that*

$$\mathbf{u} \in \bigcap_{s=0}^{m} C^{m-s}([0,T];[H^s(\Omega)]^d). \tag{3.0.7}$$

By the Sobolev embedding theorem, (3.0.7) implies that

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

where $\beta = m - [\frac{d}{2}] - 1$. Note that our assumption on m implies that $\beta \geq 2$. We shall assume throughout this chapter that the above assumptions are satisfied for m sufficiently large so that a unique solution of (3.0.4) exists. In fact, by Theorem 3.1, this unique solution satisfies

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

3.1. Definition and assumptions

In order to find a numerical approximation to the solution of the hyperbolic system (3.0.4)–(3.0.6), we discretise 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 $\mathbf{u} = \mathrm{e}^{\gamma t}\mathbf{v}$, with $\gamma > 0$, resulting in the equivalent equation:

$$\ddot{v}_i(x,t) + 2\gamma\dot{v}_i(x,t) + \gamma^2 v_i(x,t) - \mathrm{e}^{-\gamma t}\sum_{\alpha=1}^{d}\partial_\alpha S_{i\alpha}(\mathrm{e}^{\gamma t}\nabla\mathbf{v}(x,t)) = \tilde{f}_i(x,t) \text{ in } \Omega\times(0,T], \tag{3.1.1}$$

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

$$\mathbf{v}(x,t) = \mathbf{0} \text{ on } \partial\Omega\times(0,T], \tag{3.1.2}$$

$$\mathbf{v}(x,0) = \mathbf{v}_0(x) \in [H_0^1(\Omega)]^d \cap [H^m(\Omega)]^d,\ \ \dot{\mathbf{v}}(x,0) = \mathbf{v}_1(x) \in [H^{m-1}(\Omega)]^d, \tag{3.1.3}$$

where $\mathbf{v}_0(x) = \mathbf{u}_0(x)$ and $\mathbf{v}_1(x) = \mathbf{u}_1(x) - \gamma\mathbf{u}_0(x)$. Here γ can be any positive number and it does not have any effect on the convergence of the numerical scheme. From now on, we focus on the system of equations (3.1.1)–(3.1.3) only. By Theorem 3.1, we have

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

This shows that the initial conditions stated as (3.1.3) in the above definition are meaningful.

We define the following time-dependent semilinear form

$$a(\mathbf{v}(t),\boldsymbol{\varphi}) := \sum_{i,\alpha=1}^{d}\mathrm{e}^{-\gamma t}(S_{i\alpha}(e^{\gamma t}\nabla\mathbf{v}(t)),\partial_\alpha\boldsymbol{\varphi}_i)_{L^2}, \qquad \text{for } \boldsymbol{\varphi} \in [H_0^1(\Omega)]^d.$$

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

$$\mathcal{M}_\delta := \{\eta \in \mathbb{R}^{d\times d} \colon \inf_{\sigma\in\mathcal{M}}|\eta - \sigma| \le \delta\}, \tag{3.1.4}$$

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. [60] for the proof). 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

$$K_\delta := \sup_{\eta \in \mathcal{M}_\delta} \left(\sum_{i,\alpha,j,\beta=1}^{d} |A_{i\alpha j\beta}(\eta)|^2 \right)^{\frac{1}{2}}, \tag{3.1.5}$$

and the local Lipschitz constant of the fourth-order elasticity tensor $A = \nabla S$ by

$$L_\delta := \sup_{\eta_1,\eta_2 \in \mathcal{M}_\delta,\, \eta_1 \neq \eta_2} |\eta_1 - \eta_2|^{-1} \left(\sum_{i,\alpha,j,\beta=1}^{d} |A_{i\alpha j\beta}(\eta_1) - A_{i\alpha j\beta}(\eta_2)|^2 \right)^{\frac{1}{2}}. \tag{3.1.6}$$

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_δ and L_δ are finite. We also define

$$\mathcal{Z}_\delta := \{\Phi \in L^\infty(\overline{\Omega})^{d\times d} \colon \Phi(x) \in \mathcal{M}_\delta, x \in \overline{\Omega}\}. \tag{3.1.7}$$

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

$$\tilde{a}(\boldsymbol{\varphi}; \boldsymbol{\phi}, \boldsymbol{\psi}) := \sum_{i,\alpha,j,\beta=1}^{d} \left(A_{i\alpha j\beta}(\nabla\boldsymbol{\varphi})\partial_\beta \boldsymbol{\phi}_j, \partial_\alpha \boldsymbol{\psi}_i\right)_{L^2}, \quad \text{for } \boldsymbol{\varphi}, \boldsymbol{\phi}, \boldsymbol{\psi} \in [H_0^1(\Omega)]^d. \tag{3.1.8}$$

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

$$\tilde{a}(\boldsymbol{\varphi}; \boldsymbol{\phi}, \boldsymbol{\psi}) = \tilde{a}(\boldsymbol{\varphi}; \boldsymbol{\psi}, \boldsymbol{\phi}), \quad \text{for } \boldsymbol{\varphi}, \boldsymbol{\phi}, \boldsymbol{\psi} \in [H_0^1(\Omega)]^d, \nabla\boldsymbol{\varphi} \in \mathcal{Z}_\delta. \tag{S2a}$$

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

$$\tilde{a}(\boldsymbol{\varphi}; \boldsymbol{\phi}, \boldsymbol{\phi}) \geq M_1 \|\nabla\boldsymbol{\phi}\|^2_{L^2(\Omega)}, \quad \text{for } \boldsymbol{\varphi}, \boldsymbol{\phi} \in [H_0^1(\Omega)]^d, \nabla\boldsymbol{\varphi} \in \mathcal{Z}_\delta. \tag{S2b}$$

In general, (S1b) does not imply (S2b) for $d > 1$. We refer the reader to [23, 76, 80] for counterexamples. In fact, (S1b) only implies the following Gårding's inequality:

$$\tilde{a}(\boldsymbol{\varphi}; \boldsymbol{\phi}, \boldsymbol{\phi}) \geq \frac{1}{2} M_0 \|\nabla\boldsymbol{\phi}\|^2_{L^2} - \mu\|\boldsymbol{\phi}\|^2_{L^2} \quad \text{for } \mu \geq 0, \boldsymbol{\varphi}, \boldsymbol{\phi} \in [H_0^1(\Omega)]^d, \nabla\boldsymbol{\varphi} \in \mathcal{Z}_\delta,$$

cf. Theorem 6.5.1 in [57] and Lemma 5 in [60]. The techniques of this chapter can be extended so that our results are still valid under this weak condition.

3.2. Numerical scheme

3.2.1. Semi-discrete approximation We shall discretise the problem (3.1.1)–(3.1.3) in space using a continuous Galerkin method. For the spatial discretisation parameter $h \in (0,1)$, we define $\mathcal{V}_h$ to be a given family of finite-dimensional subspaces of $[H_0^1(\Omega)]^d \cap [H^m(\Omega)]^d$ with polynomial degree $p \geq 1$. We shall assume that the triangulation $\{\mathcal{T}_h\}_{h>0}$ of Ω 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 [15] that, for $\mathbf{v} \in [H^m(\Omega)]^d \cap [H_0^1(\Omega)]^d$,

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

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 $\mathbf{v}_h \in \mathcal{V}_h$,

$$\|\nabla \mathbf{v}_h\|_{L^2(\Omega)} \leq C_0 h^{-1} \|\mathbf{v}_h\|_{L^2(\Omega)} \quad \text{and} \quad \|\nabla \mathbf{v}_h\|_{L^\infty(\Omega)} \leq C_0 h^{-1} \|\mathbf{v}_h\|_{L^\infty(\Omega)}. \tag{ii,a}$$

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

$$\|\nabla \mathbf{v}_h\|_{L^\infty(\Omega)} \leq C_1 h^{-\frac{d}{2}} \|\nabla \mathbf{v}_h\|_{L^2(\Omega)} \quad \text{and} \quad \|\mathbf{v}_h\|_{L^\infty(\Omega)} \leq C_1 h^{-\frac{d}{2}} \|\mathbf{v}_h\|_{L^2(\Omega)}. \tag{ii,b}$$

With the above assumptions, we are ready to construct the continuous-in-time finite element approximation $\mathbf{v}_h$ of $\mathbf{v}$. The semi-discrete approximation $\mathbf{v}_h \colon [0,T] \to \mathcal{V}_h$ of the solution of (3.1.1)–(3.1.3) satisfies the following initial-value problem in $\mathcal{V}_h$:

$$(\ddot{\mathbf{v}}_h(t), \boldsymbol{\varphi})_{L^2} + a(\mathbf{v}_h(t), \boldsymbol{\varphi}) + 2\gamma\,(\dot{\mathbf{v}}_h(t), \boldsymbol{\varphi})_{L^2} + \gamma^2\,(\mathbf{v}_h(t), \boldsymbol{\varphi})_{L^2} = (\tilde{\mathbf{f}}(t), \boldsymbol{\varphi})_{L^2} \tag{3.2.1}$$

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

$$\mathbf{v}_h(0) = \mathbf{v}_{0,h} \in \mathcal{V}_h, \ \ \dot{\mathbf{v}}_h(0) = \mathbf{v}_{1,h} \in \mathcal{V}_h, \tag{3.2.2}$$

where $\mathbf{v}_{0,h}$ and $\mathbf{v}_{1,h}$ are specially chosen initial values. It was proved by Makridakis [56] that the semi-discrete problem (3.2.1), (3.2.2) with $\gamma = 0$ (the semi-discrete form based on a continuous finite element approximation of the original problem (3.0.4)–(3.0.6)) has a locally unique solution and that the optimal-order L^2 error estimate

$$\max_{0\le t\le T} \|\mathbf{v}(t) - \mathbf{v}_h(t)\|_{L^2(\Omega)} \le C(\mathbf{v})h^{p+1} \tag{3.2.3}$$

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.

3.2.2. Discontinuous-in-time fully discrete scheme In this section we shall construct a fully discrete approximation of the solution of (3.1.1)–(3.1.3) by applying a discontinuous Galerkin method in time. For this purpose, we partition the time interval $I = (0, T]$ into N subintervals in the same manner as the linear case in Section 2.2.1 and define the jump operator

$$[\mathbf{v}_h]_n := \mathbf{v}_h(t_n^+) - \mathbf{v}_h(t_n^-) \quad \text{ for } n = 0, 1, \dots, N-1,$$

where

$$\mathbf{v}_h(t_n^\pm) := \lim_{\varepsilon\to 0^\pm} \mathbf{v}_h(t_n + \varepsilon) \quad \text{ for } n = 0, 1, \dots, N-1.$$

Again, we assume that $\mathbf{v}_h(0^-) = \mathbf{v}_{0,h}$ and $\dot{\mathbf{v}}_h(0^-) = \mathbf{v}_{1,h}$. Moreover, we define $\mathbf{v}_{h,n}^+ := \mathbf{v}_h(t_n^+)$ and $\mathbf{v}_{h,n}^- := \mathbf{v}_h(t_n^-)$. Taking $\boldsymbol{\varphi} = \dot{\mathbf{v}}_h(t)$ in (3.2.1) and integrating on I_n, we have

$$\begin{aligned}\int_{t_{n-1}}^{t_n} (\ddot{\mathbf{v}}_h(t), \dot{\mathbf{v}}_h(t))_{L^2}\,\mathrm{d}t + \int_{t_{n-1}}^{t_n} a(\mathbf{v}_h(t), \dot{\mathbf{v}}_h(t))\mathrm{d}t + 2\gamma \int_{t_{n-1}}^{t_n} \|\dot{\mathbf{v}}_h(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \\ + \gamma^2 \int_{t_{n-1}}^{t_n} (\mathbf{v}_h(t), \dot{\mathbf{v}}_h(t))_{L^2}\,\mathrm{d}t = \int_{t_{n-1}}^{t_n} \Big(\tilde{\mathbf{f}}(t), \dot{\mathbf{v}}_h(t)\Big)_{L^2}\,\mathrm{d}t, \quad \text{ for } n = 1, 2, \dots, N.\end{aligned} \tag{3.2.4}$$

That is,

$$\int_{t_{n-1}}^{t_n} (\ddot{\mathbf{v}}_h(t), \dot{\mathbf{v}}_h(t))_{L^2}\,\mathrm{d}t + 2\gamma \int_{t_{n-1}}^{t_n} \|\dot{\mathbf{v}}_h(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t + \gamma^2 \int_{t_{n-1}}^{t_n} (\mathbf{v}_h(t), \dot{\mathbf{v}}_h(t))_{L^2}\,\mathrm{d}t$$

$$+\int_{t_{n-1}}^{t_n}\sum_{i,\alpha=1}^{d}\mathrm{e}^{-\gamma t}(S_{i\alpha}(\nabla\mathbf{v}_h(t)\mathrm{e}^{\gamma t}),\partial_\alpha\dot{\mathbf{v}}_{h,i}(t))_{L^2}\,\mathrm{d}t=\int_{t_{n-1}}^{t_n}\Big(\tilde{\mathbf{f}}(t),\dot{\mathbf{v}}_h(t)\Big)_{L^2}\,\mathrm{d}t, \tag{3.2.5}$$

for $n=1,2,\ldots,N$. To determine the appropriate strong consistent terms corresponding to $\sum_{i,\alpha=1}^{d}\mathrm{e}^{-\gamma t}(S_{i\alpha}(\nabla\mathbf{v}_h(t)\mathrm{e}^{\gamma t}),\partial_\alpha\dot{\mathbf{v}}_{h,i}(t))_{L^2}$ in the numerical scheme, we linearise this semilinear term first. By Taylor's theorem with an integral remainder, we have

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

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

$$\begin{aligned}
&\int_{t_{n-1}}^{t_n}\sum_{i,\alpha=1}^{d}\mathrm{e}^{-\gamma t}\left(S_{i\alpha}(\nabla\mathbf{v}_h(t)\mathrm{e}^{\gamma t}),\partial_\alpha\dot{\mathbf{v}}_{h,i}(t)\right)_{L^2}\,\mathrm{d}t\\
&=\int_{t_{n-1}}^{t_n}\sum_{i,\alpha,j,\beta=1}^{d}\left(\partial_\beta\mathbf{v}_{h,j}(t)\int_0^1 A_{i\alpha j\beta}(\tau\nabla\mathbf{v}_h(t)\mathrm{e}^{\gamma t})\mathrm{d}\tau,\partial_\alpha\dot{\mathbf{v}}_{h,i}(t)\right)_{L^2}\,\mathrm{d}t\\
&=\int_{t_{n-1}}^{t_n}\int_0^1\sum_{i,\alpha,j,\beta=1}^{d}\left(A_{i\alpha j\beta}(\tau\nabla\mathbf{v}_h(t)\mathrm{e}^{\gamma t})\partial_\beta\mathbf{v}_{h,j}(t),\partial_\alpha\dot{\mathbf{v}}_{h,i}(t)\right)_{L^2}\,\mathrm{d}\tau\,\mathrm{d}t\\
&:=\int_{t_{n-1}}^{t_n}\left\{\int_0^1\tilde{a}(\tau\mathbf{v}_h(t)\mathrm{e}^{\gamma t};\mathbf{v}_h(t),\dot{\mathbf{v}}_h(t))\mathrm{d}\tau\right\}\mathrm{d}t.
\end{aligned}\tag{3.2.6}$$

Note that

$$\begin{aligned}
&\int_0^1\tilde{a}(\tau\mathbf{v}_h(t)\mathrm{e}^{\gamma t};\mathbf{v}_h(t),\dot{\mathbf{v}}_h(t))\,\mathrm{d}\tau=\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_0^1\tilde{a}(\tau\mathbf{v}_h(t)\mathrm{e}^{\gamma t};\mathbf{v}_h(t),\mathbf{v}_h(t))\,\mathrm{d}\tau\\
&\quad-\frac{1}{2}\int_0^1\sum_{i,\alpha,j,\beta=1}^{d}\left(\partial_t A_{i\alpha j\beta}(\tau\nabla\mathbf{v}_h(t)\mathrm{e}^{\gamma t})\partial_\beta\mathbf{v}_{h,j}(t),\partial_\alpha\mathbf{v}_{h,i}(t)\right)_{L^2}\,\mathrm{d}\tau.
\end{aligned}\tag{3.2.7}$$

Using Theorem 1.7 and Theorem 1.8 in (3.2.5), we have

$$\begin{aligned}
&\frac{1}{2}\left(\int_0^1\tilde{a}(\tau\mathbf{v}_{h,n}^{-}\mathrm{e}^{\gamma t_n};\mathbf{v}_{h,n}^{-},\mathbf{v}_{h,n}^{-})\,\mathrm{d}\tau-\int_0^1\tilde{a}(\tau\mathbf{v}_{h,n-1}^{+}\mathrm{e}^{\gamma t_{n-1}};\mathbf{v}_{h,n-1}^{+},\mathbf{v}_{h,n-1}^{+})\,\mathrm{d}\tau\right)\\
&\quad+\frac{1}{2}\left(\|\dot{\mathbf{v}}_{h,n}^{-}\|_{L^2(\Omega)}^2-\|\dot{\mathbf{v}}_{h,n-1}^{+}\|_{L^2(\Omega)}^2\right)+\frac{\gamma^2}{2}\left(\|\mathbf{v}_{h,n}^{-}\|_{L^2(\Omega)}^2-\|\mathbf{v}_{h,n-1}^{+}\|_{L^2(\Omega)}^2\right)\\
&\quad+2\gamma\int_{t_{n-1}}^{t_n}\|\dot{\mathbf{v}}_h(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t\\
&=\frac{1}{2}\int_{t_{n-1}}^{t_n}\int_0^1\sum_{i,\alpha,j,\beta=1}^{d}\left(\partial_t A_{i\alpha j\beta}(\tau\nabla\mathbf{v}_h(t)\mathrm{e}^{\gamma t})\partial_\beta\mathbf{v}_{h,j}(t),\partial_\alpha\mathbf{v}_{h,i}(t)\right)_{L^2}\,\mathrm{d}\tau\,\mathrm{d}t
\end{aligned}$$

$$+\int_{t_{n-1}}^{t_n}\Big(\tilde{\mathbf{f}}(t),\dot{\mathbf{v}}_h(t)\Big)_{L^2}\,\mathrm{d}t.$$

Summing up over $n=1,2,\ldots,N$, we have

$$\begin{aligned}
&\frac{1}{2}\bigg(\int_0^1 \tilde{a}(\tau\mathbf{v}_{h,N}^{-}\mathrm{e}^{\gamma t_N};\mathbf{v}_{h,N}^{-},\mathbf{v}_{h,N}^{-})\,\mathrm{d}\tau-\int_0^1 \tilde{a}(\tau\mathbf{v}_{h,0}^{+}\mathrm{e};\mathbf{v}_{h,0}^{+},\mathbf{v}_{h.0}^{+})\,\mathrm{d}\tau\bigg)\\
&\quad+\frac{1}{2}\bigg(\|\dot{\mathbf{v}}_h(t_N^-)\|_{L^2(\Omega)}^2-\|\dot{\mathbf{v}}_h(t_0^+)\|_{L^2(\Omega)}^2\bigg)+\frac{\gamma^2}{2}\bigg(\|\mathbf{v}_h(t_N^-)\|_{L^2(\Omega)}^2-\|\mathbf{v}_h(t_0^+)\|_{L^2(\Omega)}^2\bigg)\\
&\quad-\frac{1}{2}\bigg\{\|\dot{\mathbf{v}}_h(t_{N-1}^+)\|_{L^2(\Omega)}^2-\|\dot{\mathbf{v}}_h(t_{N-1}^-)\|_{L^2(\Omega)}^2+\cdots+\|\dot{\mathbf{v}}_h(t_1^+)\|_{L^2(\Omega)}^2-\|\dot{\mathbf{v}}_h(t_1^-)\|_{L^2(\Omega)}^2\bigg\}\\
&\quad-\frac{\gamma^2}{2}\bigg\{\|\mathbf{v}_h(t_{N-1}^+)\|_{L^2(\Omega)}^2-\|\mathbf{v}_h(t_{N-1}^-)\|_{L^2(\Omega)}^2+\cdots+\|\mathbf{v}_h(t_1^+)\|_{L^2(\Omega)}^2-\|\mathbf{v}_h(t_1^-)\|_{L^2(\Omega)}^2\bigg\}\\
&\quad-\frac{1}{2}\bigg\{\int_0^1 \tilde{a}(\tau\mathbf{v}_{h,N-1}^{+}\mathrm{e}^{\gamma t_{N-1}};\mathbf{v}_{h,N-1}^{+},\mathbf{v}_{h,N-1}^{+})\,\mathrm{d}\tau-\int_0^1 \tilde{a}(\tau\mathbf{v}_{h,N-1}^{-}\mathrm{e}^{\gamma t_{N-1}};\mathbf{v}_{h,N-1}^{-},\mathbf{v}_{h,N-1}^{-})\,\mathrm{d}\tau\\
&\quad+\cdots+\int_0^1 \tilde{a}(\tau\mathbf{v}_{h,1}^{+}\mathrm{e}^{\gamma t_1};\mathbf{v}_{h,1}^{+},\mathbf{v}_{h,1}^{+})\,\mathrm{d}\tau-\int_0^1 \tilde{a}(\tau\mathbf{v}_{h,1}^{-}\mathrm{e}^{\gamma t_1};\mathbf{v}_{h,1}^{-},\mathbf{v}_{h,1}^{-})\,\mathrm{d}\tau\bigg\}\\
&\quad+2\gamma\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\|\dot{\mathbf{v}}_h(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t\\
&=\frac{1}{2}\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\int_0^1\sum_{i,\alpha,j,\beta=1}^{d}\big(\partial_t A_{i\alpha j\beta}(\tau\nabla\mathbf{v}_h(t)\mathrm{e}^{\gamma t})\partial_\beta\mathbf{v}_{h,j}(t),\partial_\alpha\mathbf{v}_{h,i}(t)\big)_{L^2}\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(\tilde{\mathbf{f}}(t),\dot{\mathbf{v}}_h(t))_{L^2}\,\mathrm{d}t.
\end{aligned}$$

The terms in the large brackets motivate us to add extra terms to achieve a well-posed variational form. By Carathéodory's existence theorem for systems of ODEs, we deduce from (3.2.1) that both $\mathbf{v}_h$ and $\dot{\mathbf{v}}_h$ are absolutely continuous in time. This implies that $[\mathbf{v}_h]_n=0$ and $[\dot{\mathbf{v}}_h]_n=0$ for all $n=0,1,\ldots,N-1$. Slightly different from the linear case in Chapter 2, we add

$$\frac{1}{2}\bigg\{\int_0^1 \tilde{a}(\tau\mathbf{v}_{h,n}^{+}\mathrm{e}^{\gamma t_n};\mathbf{v}_{h,n}^{+},\mathbf{v}_{h,n}^{+})\,\mathrm{d}\tau-\int_0^1 \tilde{a}(\tau\mathbf{v}_{h,n}^{-}\mathrm{e}^{\gamma t_n};\mathbf{v}_{h,n}^{-},\mathbf{v}_{h,n}^{-})\,\mathrm{d}\tau\bigg\}$$

to compensate the terms

$$-\frac{1}{2}\bigg\{\int_0^1 \tilde{a}(\tau\mathbf{v}_{h,n}^{+}\mathrm{e}^{\gamma t_n};\mathbf{v}_{h,n}^{+},\mathbf{v}_{h,n}^{+})\,\mathrm{d}\tau-\int_0^1 \tilde{a}(\tau\mathbf{v}_{h,n}^{-}\mathrm{e}^{\gamma t_n};\mathbf{v}_{h,n}^{-},\mathbf{v}_{h,n}^{-})\,\mathrm{d}\tau\bigg\}$$

arising from integration by parts of the linearised version of

$$\int_{t_{n-1}}^{t_n} \sum_{i,\alpha=1}^{d} \mathrm{e}^{-\gamma t} \left(S_{i\alpha}(\nabla \mathbf{v}_h(t)\mathrm{e}^{\gamma t}), \partial_\alpha \dot{\mathbf{v}}_{h,i}(t)\right)_{L^2} \, \mathrm{d}t$$

for each $n = 0, 1, \ldots, N-1$. Adding

$$\sum_{n=0}^{N-1} ([\dot{\mathbf{v}}_h]_n, \dot{\mathbf{v}}_{h,n}^+)_{L^2} = \left(\dot{\mathbf{v}}_{h,N-1}^+ - \dot{\mathbf{v}}_{h,N-1}^-, \dot{\mathbf{v}}_{h,N-1}^+\right)_{L^2} + \cdots + \left(\dot{\mathbf{v}}_{h,0}^+ - \dot{\mathbf{v}}_{h,0}^-, \dot{\mathbf{v}}_{h,0}^+\right)_{L^2},$$

$$\gamma^2 \sum_{n=0}^{N-1} ([\mathbf{v}_h]_n, \mathbf{v}_{h,n}^+)_{L^2} = \gamma^2 \left(\mathbf{v}_{h,N-1}^+ - \mathbf{v}_{h,N-1}^-, \mathbf{v}_{h,N-1}^+\right)_{L^2} + \cdots + \gamma^2 \left(\mathbf{v}_{h,0}^+ - \mathbf{v}_{h,0}^-, \mathbf{v}_{h,0}^+\right)_{L^2},$$

and

$$\sum_{n=0}^{N-1} \frac{1}{2} \left\{ \int_0^1 \tilde{a}(\tau \mathbf{v}_{h,n}^+ \mathrm{e}^{\gamma t_n}; \mathbf{v}_{h,n}^+, \mathbf{v}_{h,n}^+) \, \mathrm{d}\tau - \int_0^1 \tilde{a}(\tau \mathbf{v}_{h,n}^- \mathrm{e}^{\gamma t_n}; \mathbf{v}_{h,n}^-, \mathbf{v}_{h,n}^-) \, \mathrm{d}\tau \right\}$$

to the left-hand side, we have

$$\begin{aligned}
&\frac{1}{2}\|\dot{\mathbf{v}}_h(t_0^+)\|_{L^2(\Omega)}^2 + \frac{1}{2}\sum_{n=1}^{N-1} \|[\dot{\mathbf{v}}_h]_n\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\dot{\mathbf{v}}_h(t_N^-)\|_{L^2(\Omega)}^2 + \frac{\gamma^2}{2}\|\mathbf{v}_h(t_0^+)\|_{L^2(\Omega)}^2 \\
&\quad + \frac{\gamma^2}{2}\sum_{n=1}^{N-1} \|[\mathbf{v}_h]_n\|_{L^2(\Omega)}^2 + \frac{\gamma^2}{2}\|\mathbf{v}_h(t_N^-)\|_{L^2(\Omega)}^2 + \frac{1}{2}\int_0^1 \tilde{a}(\tau \mathbf{v}_{h,N}^- \mathrm{e}^{\gamma t_N}; \mathbf{v}_{h,N}^-, \mathbf{v}_{h,N}^-) \, \mathrm{d}\tau \\
&\quad + 2\gamma \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|\dot{\mathbf{v}}_h(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t \\
&= \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\tilde{\mathbf{f}}(t), \dot{\mathbf{v}}_h(t))_{L^2} \, \mathrm{d}t + \left(\dot{\mathbf{v}}_h(t_0^-), \dot{\mathbf{v}}_h(t_0^+)\right)_{L^2} + \gamma^2 \left(\mathbf{v}_h(t_0^-), \mathbf{v}_h(t_0^+)\right)_{L^2} \\
&\quad + \frac{1}{2}\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_0^1 \sum_{i,\alpha,j,\beta=1}^{d} \left(\partial_t A_{i\alpha j\beta}(\tau \nabla \mathbf{v}_h(t)\mathrm{e}^{\gamma t}) \partial_\beta \mathbf{v}_{h,j}(t), \partial_\alpha \mathbf{v}_{h,i}(t)\right)_{L^2} \, \mathrm{d}\tau \, \mathrm{d}t \\
&\quad + \frac{1}{2}\int_0^1 \tilde{a}(\tau \mathbf{v}_{h,0}^-; \mathbf{v}_{h,0}^-, \mathbf{v}_{h,0}^-) \, \mathrm{d}\tau.
\end{aligned}$$

We now focus on the generic time slab I_n and assume that the solution on I_{n-1} is known. Testing the equation (3.1.1) against $\dot{\boldsymbol{\varphi}}$ for $\boldsymbol{\varphi} \in H^1(I_n; [H_0^1(\Omega)]^d)$ and integrating on I_n, we have

$$\int_{t_{n-1}}^{t_n} (\ddot{\mathbf{v}}(t), \dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t + \int_{t_{n-1}}^{t_n} \sum_{i,\alpha=1}^{d} \mathrm{e}^{-\gamma t}\left(S_{i\alpha}(\nabla \mathbf{v}(t)\mathrm{e}^{\gamma t}), \partial_\alpha \dot{\boldsymbol{\varphi}}_i(t)\right)_{L^2}\,\mathrm{d}t$$
$$+ 2\gamma \int_{t_{n-1}}^{t_n} (\dot{\mathbf{v}}(t), \dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t + \gamma^2 \int_{t_{n-1}}^{t_n} (\mathbf{v}(t), \dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t = \int_{t_{n-1}}^{t_n} \left(\tilde{\mathbf{f}}(t), \dot{\boldsymbol{\varphi}}(t)\right)_{L^2}\,\mathrm{d}t. \tag{3.2.8}$$

Rewriting (3.2.8) by adding suitable (strongly consistent) terms, we have

$$\int_{t_{n-1}}^{t_n} (\ddot{\mathbf{v}}(t), \dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t + \left([\dot{\mathbf{v}}(t)]_{n-1}, \dot{\boldsymbol{\varphi}}(t_{n-1}^+)\right)_{L^2} + \int_{t_{n-1}}^{t_n} \sum_{i,\alpha=1}^{d} e^{-\gamma t}\left(S_{i\alpha}(\nabla \mathbf{v}(t)e^{\gamma t}), \partial_\alpha \dot{\boldsymbol{\varphi}}_i(t)\right)_{L^2}\,\mathrm{d}t$$
$$+ 2\gamma \int_{t_{n-1}}^{t_n} (\dot{\mathbf{v}}(t), \dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t + \gamma^2 \int_{t_{n-1}}^{t_n} (\mathbf{v}(t), \dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t + \gamma^2\left([\mathbf{v}(t)]_{n-1}, \boldsymbol{\varphi}(t_{n-1}^+)\right)_{L^2}$$
$$+ \frac{1}{2}\left\{\int_0^1 \tilde{a}(\tau \mathbf{v}_{h,n-1}^+ \mathrm{e}^{\gamma t_{n-1}}; \mathbf{v}_{h,n-1}^+, \mathbf{v}_{h,n-1}^+)\,\mathrm{d}\tau - \int_0^1 \tilde{a}(\tau \mathbf{v}_{h,n-1}^- \mathrm{e}^{\gamma t_{n-1}}; \mathbf{v}_{h,n-1}^-, \mathbf{v}_{h,n-1}^-)\,\mathrm{d}\tau\right\}$$
$$= \int_{t_{n-1}}^{t_n} \left(\tilde{\mathbf{f}}(t), \dot{\boldsymbol{\varphi}}(t)\right)_{L^2}\,\mathrm{d}t. \tag{3.2.9}$$

Summing over all time intervals in (3.2.9) leads us to define the following semilinear form $\mathcal{A}\colon \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ with

$$\mathcal{H} := H^2(0,T; [H^m(\Omega)]^d \cap [H_0^1(\Omega)]^d),$$

by

$$\begin{aligned}
\mathcal{A}(\mathbf{v}, \boldsymbol{\varphi}) :=& \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\ddot{\mathbf{v}}(t), \dot{\boldsymbol{\varphi}}(t))\,\mathrm{d}t + \sum_{n=1}^{N-1} \left([\dot{\mathbf{v}}(t)]_n, \dot{\boldsymbol{\varphi}}(t_n^+)\right)_{L^2} + \left(\dot{\mathbf{v}}(t_0^+), \dot{\boldsymbol{\varphi}}(t_0^+)\right)_{L^2} \\
&+ \gamma^2 \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\mathbf{v}(t), \dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t + \gamma^2 \sum_{n=1}^{N-1} \left([\mathbf{v}(t)]_n, \boldsymbol{\varphi}(t_n^+)\right)_{L^2} + \gamma^2\left(\mathbf{v}(t_0^+), \boldsymbol{\varphi}(t_0^+)\right)_{L^2} \\
&+ \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \sum_{i,\alpha=1}^{d} \mathrm{e}^{-\gamma t}\left(S_{i\alpha}(\nabla \mathbf{v}(t)\mathrm{e}^{\gamma t}), \partial_\alpha \dot{\boldsymbol{\varphi}}_i(t)\right)_{L^2}\,\mathrm{d}t + 2\gamma \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\dot{\mathbf{v}}(t), \dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t \\
&+ \sum_{n=1}^{N-1} \frac{1}{2}\left\{\int_0^1 \tilde{a}(\tau \mathbf{v}_n^+ \mathrm{e}^{\gamma t_n}; \mathbf{v}_n^+, \boldsymbol{\varphi}(t_n^+))\,\mathrm{d}\tau - \int_0^1 \tilde{a}(\tau \mathbf{v}_n^- \mathrm{e}^{\gamma t_n}; \mathbf{v}_n^-, \boldsymbol{\varphi}(t_n^-))\,\mathrm{d}\tau\right\} \\
&+ \frac{1}{2}\int_0^1 \tilde{a}(\tau \mathbf{v}(t_0^+); \mathbf{v}(t_0^+), \boldsymbol{\varphi}(t_0^+))\,\mathrm{d}\tau
\end{aligned}$$

for $\boldsymbol{\varphi} \in \mathcal{H}$. Note that the consistent terms added in the fourth and fifth lines of the definition of $\mathcal{A}$ correspond to the linearised version of $\sum_{i,\alpha=1}^{d} \mathrm{e}^{-\gamma t}(S_{i\alpha}(\nabla \mathbf{v}_h(t)\mathrm{e}^{\gamma t}), \partial_\alpha \dot{\mathbf{v}}_{h,i}(t))_{L^2}$. Let, further, F be the linear functional defined by

$$
\begin{aligned}
F(\boldsymbol{\varphi}) :=& \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left(\tilde{\mathbf{f}}(t), \dot{\boldsymbol{\varphi}}(t)\right)_{L^2} \, \mathrm{d}t + \left(\mathbf{v}_1, \dot{\boldsymbol{\varphi}}(t_0^+)\right)_{L^2} + \gamma^2 \left(\mathbf{v}_0, \boldsymbol{\varphi}(t_0^+)\right)_{L^2} \\
&+ \frac{1}{2} \int_0^1 \tilde{a}(\tau \mathbf{v}_0; \mathbf{v}_0, \boldsymbol{\varphi}(t_0^-)) \, \mathrm{d}\tau.
\end{aligned}
$$

Again, the semilinear form $\mathcal{A}(\cdot,\cdot)$ works for piecewise in time version of $\mathcal{H} \times \mathcal{H}$ as well. Analogously to Section 2.3, the fully discrete finite element space is defined as

$$
\mathcal{V}_{kh}^{\mathbf{q}} := \{\mathbf{v} \colon [0,T] \to \mathcal{V}_h; \mathbf{v}|_{I_n} \in \mathcal{V}_{kh}^{q_n} \text{ for } n = 1, 2, \ldots, N\},
$$

for $\mathbf{q} := [q_1, q_2, \ldots, q_N]^T \in \mathbb{N}^N$, where

$$
\mathcal{V}_{kh}^{q_n} := \{\mathbf{v} \colon [0,T] \to \mathcal{V}_h; \mathbf{v}|_{I_n} = \sum_{j=0}^{q_n} \mathbf{v}_j t^j, \mathbf{v}_j \in \mathcal{V}_h\},
$$

with $q_n \geq 2$ for each $1 \leq n \leq N$. Then, the discontinuous-in-time fully discrete approximation of the problem reads as follows: find $\mathbf{v}_{\mathrm{DG}} \in \mathcal{V}_{kh}^{\mathbf{q}}$ such that

$$
\mathcal{A}(\mathbf{v}_{\mathrm{DG}}, \boldsymbol{\varphi}) = \tilde{F}(\boldsymbol{\varphi}) \qquad \text{for all } \boldsymbol{\varphi} \in \mathcal{V}_{kh}^{\mathbf{q}}, \tag{3.2.10}
$$

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

$$
\begin{aligned}
\tilde{F}(\boldsymbol{\varphi}) :=& \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left(\tilde{\mathbf{f}}(t), \dot{\boldsymbol{\varphi}}(t)\right)_{L^2} \, \mathrm{d}t + \left(\mathbf{v}_{1,h}, \dot{\boldsymbol{\varphi}}(t_0^+)\right)_{L^2} + \gamma^2 \left(\mathbf{v}_{0,h}, \boldsymbol{\varphi}(t_0^+)\right)_{L^2} \\
&+ \frac{1}{2} \int_0^1 \tilde{a}(\tau \mathbf{v}_{0,h}; \mathbf{v}_{0,h}, \boldsymbol{\varphi}(t_0^-)) \, \mathrm{d}\tau.
\end{aligned}
$$

3.3. Convergence analysis

By using the ideas introduced in [56] based on Banach's fixed point theorem, we will show the existence and uniqueness of $\mathbf{v}_{\mathrm{DG}}$. We shall also prove *a priori* error estimates as summarised in the following theorem.

Theorem 3.2. *Let* $\mathbf{v} \in W^{s,\infty}([0,T]; [H^m(\Omega)]^d \cap [H_0^1(\Omega)]^d)$ *be the solution of (3.1.1)–*

(3.1.3). Assume that $\frac{d}{2}+1 < r \le \min(p, m-1)$, $s \ge q_i + 1$, $k_i^{q_i - \frac{1}{2}} = o(h^{\frac{d}{2}+1})$, *and there exist positive constants* μ_i, ν_i *such that* $\mu_i k_i \le h^2 \le \nu_i k_i$ *for each* $i = 1, 2, \dots N$. *Suppose that we choose the initial data* $\mathbf{v}_{0,h}$, $\mathbf{v}_{1,h} \in \mathcal{V}_h$ *to be*

$$\mathbf{v}_{0,h} = \mathbf{W}(0), \quad \mathbf{v}_{1,h} = \dot{\mathbf{W}}(0), \tag{3.3.1}$$

where $\mathbf{W}(t) \in \mathcal{V}_h$ *is the nonlinear elliptic projection of* $\mathbf{v}(t)$ *such that*

$$a(\mathbf{W}(t), \boldsymbol{\varphi}) = a(\mathbf{v}(t), \boldsymbol{\varphi}) \quad \textit{for all } \boldsymbol{\varphi} \in \mathcal{V}_h. \tag{3.3.2}$$

Then there exists a unique solution to (3.2.10) such that

$$\|\mathbf{v}_{\mathrm{DG}}(t_j^-) - \mathbf{v}(t_j^-)\|_{L^2(\Omega)} + \|\dot{\mathbf{v}}_{\mathrm{DG}}(t_j^-) - \dot{\mathbf{v}}(t_j^-)\|_{L^2(\Omega)} \le C(\mathbf{v}) \left(h^{2r+2} + \sum_{n=1}^{j} \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} \right)^{\frac{1}{2}} \tag{3.3.3}$$

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

Remark 3.3. *If we use uniform time intervals* $k_n = k = h^2$, *and uniform polynomial degrees* $q_n = q \ge 2$, *for* $n = 1, \dots, N$, *then the error bound at the end nodal point becomes*

$$\|\mathbf{v}_{\mathrm{DG}}(t_N^-) - \mathbf{v}(t_N^-)\|_{L^2(\Omega)} + \|\dot{\mathbf{v}}_{\mathrm{DG}}(t_N^-) - \dot{\mathbf{v}}(t_N^-)\|_{L^2(\Omega)} \le C(\mathbf{v}) \left(k^{\frac{r+1}{2}} + k^q \right).$$

Remark 3.4. *The assumptions that* $k_i^{q_i - \frac{1}{2}} = o(h^{1+\frac{d}{2}})$ *and* $\mu_i k_i \le h^2 \le \nu_i k_i$ *for each* $i = 1, \dots N$ *require that* $q_i > 1 + \frac{d}{4}$ *for each* $i = 1, \dots, N$. *That is, we need the polynomial degree in time satisfies* $q_i \ge 2$ *for* $d = 1, 2, 3$ *on each time interval* I_i, *with* $i = 1, \dots, N$.

Remark 3.5. *By the Sobolev embedding theorem,* $\mathbf{v} \in W^{s,\infty}(0, T; [H^m(\Omega)]^d)$ *for* $m > \frac{d}{2}+2$ *implies that* $\mathbf{v} \in W^{s,\infty}(0, T; C^{2,\alpha}(\overline{\Omega})^d)$ *for some* $\alpha \in (0, 1)$. *Note that the assumption* $m > \frac{d}{2} + 2$ *is consistent with the assumption* $m \ge [\frac{d}{2}] + 3$ *in Theorem 3.1. That is, we need* $m \ge 3$ *for* $d = 1$ *and* $m \ge 4$ *for* $d = 2, 3$.

It will be assumed throughout the convergence analysis that

$$\mathbf{v} \in W^{s,\infty}(0, T; [H^m(\Omega)]^d \cap [H_0^1(\Omega)]^d).$$

3.3.1. Definition of the fixed point map It is known, see [26] and [62], that (3.3.2) has, for h sufficiently small, a locally unique solution $\mathbf{W}(t) \in \mathcal{V}_h$ for $0 \le t \le T$. Furthermore, $\mathbf{W}$ satisfies the following properties, which are established in Section 3.4.

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

$$\|\nabla \mathbf{v}^{(j)}(t) - \nabla \mathbf{W}^{(j)}(t)\|_{L^2(\Omega)} \le C_r(\mathbf{v}) h^r, \quad 0 \le t \le T, \tag{iii,a}$$

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

$$\|\mathbf{v}^{(j)}(t) - \mathbf{W}^{(j)}(t)\|_{L^2(\Omega)} \le \tilde{C}_r(\mathbf{v}) h^{r+1}, \quad 0 \le t \le T, \tag{iii,b}$$

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

$$\|\nabla \mathbf{W}(t)\|_{L^\infty(\Omega)} \le c_0 \text{ and } \|\nabla \dot{\mathbf{W}}(t)\|_{L^\infty(\Omega)} \le c_1, \quad 0 \le t \le T. \tag{iii,c}$$

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

$$(\Pi_k \mathbf{W} - \mathbf{W})(x, t_{n-1}^+) = 0; \tag{3.3.4}$$

$$(\Pi_k \mathbf{W} - \mathbf{W})(x, t_n^-) = 0; \tag{3.3.5}$$

$$\partial_t (\Pi_k \mathbf{W} - \mathbf{W})(x, t_n^-) = 0; \tag{3.3.6}$$

$$\int_{t_{n-1}}^{t_n} (\partial_t (\Pi_k \mathbf{W} - \mathbf{W}), \boldsymbol{\chi})_{L^2} \,\mathrm{d}t = 0 \quad \text{ for } \boldsymbol{\chi} \in \mathcal{V}_{kh}^{q_n - 2}. \tag{3.3.7}$$

It was proved in Section 2.2 that for each $\mathbf{W} \in H^s(I_n; [L^2(\Omega)]^d)$, there exists a positive constant C such that

$$\int_{t_{n-1}}^{t_n} \|\partial_{tt}(\mathbf{W}(\cdot, t) - \Pi_k \mathbf{W}(\cdot, t))\|_{L^2(\Omega)}^2 \,\mathrm{d}t \le C \frac{k_n^{2(\mu-2)}}{q_n^{2(s-3)}} \|\mathbf{W}\|_{H^s(I_n, [L^2(\Omega)]^d)}^2, \tag{3.3.8}$$

$$\int_{t_{n-1}}^{t_n} \|\partial_t(\mathbf{W}(\cdot, t) - \Pi_k \mathbf{W}(\cdot, t))\|_{L^2(\Omega)}^2 \,\mathrm{d}t \le C \frac{k_n^{2(\mu-1)}}{q_n^{2(s-1)}} \|\mathbf{W}\|_{H^s(I_n, [L^2(\Omega)]^d)}^2, \tag{3.3.9}$$

$$\int_{t_{n-1}}^{t_n} \|\mathbf{W}(\cdot, t) - \Pi_k \mathbf{W}(\cdot, t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \le C \frac{k_n^{2\mu}}{q_n^{2(s-1)}} \|\mathbf{W}\|_{H^s(I_n, [L^2(\Omega)]^d)}^2, \tag{3.3.10}$$

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 $[L^2(\Omega)]^d$ to $[H_0^1(\Omega)]^d$ in (3.3.8)–(3.3.10), 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 $\boldsymbol{\varphi} \in \mathcal{V}_{kh}^{\mathbf{q}}$,

$$\|\partial_t \boldsymbol{\varphi}(x,\cdot)\|_{L^2(I_n)} \leq C_2 k_n^{-1} \|\boldsymbol{\varphi}(x,\cdot)\|_{L^2(I_n)}, \tag{3.3.11}$$

$$\|\partial_t \boldsymbol{\varphi}(x,\cdot)\|_{L^\infty(I_n)} \leq C_2 k_n^{-1} \|\boldsymbol{\varphi}(x,\cdot)\|_{L^\infty(I_n)}, \tag{3.3.12}$$

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

$$\begin{aligned}
\mathbf{v}_{\mathrm{DG}}(t) - \mathbf{v}(t) &= (\mathbf{v}_{\mathrm{DG}}(t) - \Pi_k \mathbf{W}(t)) + (\Pi_k \mathbf{W}(t) - \mathbf{W}(t)) + (\mathbf{W}(t) - \mathbf{v}(t)) \\
&:= \boldsymbol{\theta}(t) + \boldsymbol{\rho}_1(t) + \boldsymbol{\rho}_2(t)
\end{aligned}$$

for $t \in I_n, n = 1, 2, \ldots, N$. First note that

$$\begin{aligned}
\int_{t_{n-1}}^{t_n} \|\boldsymbol{\rho}_1^{(j)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t &= \int_{t_{n-1}}^{t_n} \|\Pi_k \mathbf{W}^{(j)}(t) - \mathbf{W}^{(j)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\leq C \frac{k_n^{2(\mu-j)}}{q_n^{2(s-1)}} \int_{t_{n-1}}^{t_n} \sum_{\alpha=0}^{s} \|\mathbf{W}^{(\alpha)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\leq C_1(\mathbf{v}) \frac{k_n^{2(\mu-j)+1}}{q_n^{2(s-1)}} \qquad \text{for } j = 0, 1,
\end{aligned} \tag{3.3.13}$$

where we have applied (3.3.9) and (3.3.10). Here $\mu = \min(s, q_n+1)$ and $\Pi_k \mathbf{W}^{(j)} = \frac{\mathrm{d}^j \Pi_k \mathbf{W}}{\mathrm{d}t^j}$, $\mathbf{W}^{(j)} = \frac{\mathrm{d}^j \mathbf{W}}{\mathrm{d}t^j}$, with $j = 0, 1$. If we assume that the solution

$$\mathbf{v} \in W^{s,\infty}(0, T; [H^m(\Omega)]^d \cap [H_0^1(\Omega)]^d)$$

of (3.1.1)–(3.1.3) is sufficiently smooth (i.e. $s > q_n + 1$), then we can write

$$\int_{t_{n-1}}^{t_n} \|\boldsymbol{\rho}_1^{(j)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \leq C \frac{k_n^{2(q_n+1-j)}}{q_n^{2(s-1)}} \int_{t_{n-1}}^{t_n} \sum_{\alpha=0}^{s} \|\mathbf{W}^{(\alpha)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \leq C_1(\mathbf{v}) \frac{k_n^{2(q_n+1-j)+1}}{q_n^{2(s-1)}} \tag{3.3.14}$$

for $j = 0, 1$. By the property (iii,b) of the elliptic projection, we know that

$$\int_{t_{n-1}}^{t_n} \|\boldsymbol{\rho}_2^{(j)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t = \int_{t_{n-1}}^{t_n} \|\mathbf{W}^{(j)}(t) - \mathbf{v}^{(j)}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \leq C_2(\mathbf{v}) k_n h^{2r+2} \qquad (3.3.15)$$

for $j = 0, 1, 2$. Here $C_i(\mathbf{v})$ for $i = 1, 2$ are constants depending on the exact solution $\mathbf{v}$. Recall that the fully discrete scheme is

$$\begin{aligned}
&\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \langle \ddot{\mathbf{v}}_{\mathrm{DG}}(t), \dot{\boldsymbol{\varphi}}(t) \rangle \,\mathrm{d}t + \sum_{n=1}^{N-1} \left([\dot{\mathbf{v}}_{\mathrm{DG}}(t)]_n, \dot{\boldsymbol{\varphi}}_n^+ \right)_{L^2} + \left(\dot{\mathbf{v}}_{\mathrm{DG},0}^+, \dot{\boldsymbol{\varphi}}_0^+ \right)_{L^2} \\
&\quad + \gamma^2 \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left(\mathbf{v}_{\mathrm{DG}}(t), \dot{\boldsymbol{\varphi}}(t) \right)_{L^2} \,\mathrm{d}t + \gamma^2 \sum_{n=1}^{N-1} \left([\mathbf{v}_{\mathrm{DG}}(t)]_n, \boldsymbol{\varphi}_n^+ \right)_{L^2} + \gamma^2 \left(\mathbf{v}_{\mathrm{DG},0}^+, \boldsymbol{\varphi}_0^+ \right)_{L^2} \\
&\quad + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \sum_{i,\alpha=1}^{d} \mathrm{e}^{-\gamma t} \left(S_{i\alpha}(\nabla \mathbf{v}_{\mathrm{DG}}(t) \mathrm{e}^{\gamma t}), \partial_\alpha \dot{\boldsymbol{\varphi}}_i(t) \right)_{L^2} \,\mathrm{d}t + 2\gamma \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left(\dot{\mathbf{v}}_{\mathrm{DG}}(t), \dot{\boldsymbol{\varphi}}(t) \right)_{L^2} \,\mathrm{d}t \\
&\quad + \sum_{n=0}^{N-1} \frac{1}{2} \left\{ \int_0^1 \tilde{a}(\tau \mathbf{v}_{\mathrm{DG},n}^+ \mathrm{e}^{\gamma t_n}; \mathbf{v}_{\mathrm{DG},n}^+, \boldsymbol{\varphi}_n^+) \,\mathrm{d}\tau - \int_0^1 \tilde{a}(\tau \mathbf{v}_{\mathrm{DG},n}^- \mathrm{e}^{\gamma t_n}; \mathbf{v}_{\mathrm{DG},n}^-, \boldsymbol{\varphi}_n^-) \,\mathrm{d}\tau \right\} \\
&= \left(\mathbf{v}_{1,h}, \dot{\boldsymbol{\varphi}}_0^+ \right)_{L^2} + \gamma^2 \left(\mathbf{v}_{0,h}, \boldsymbol{\varphi}_0^+ \right)_{L^2} + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left(\tilde{\mathbf{f}}(t), \dot{\boldsymbol{\varphi}}(t) \right)_{L^2} \,\mathrm{d}t, \qquad \text{for } \boldsymbol{\varphi} \in \mathcal{V}_{kh}^{\mathbf{q}}.
\end{aligned} \qquad (3.3.16)$$

The variational form of the original problem is written as

$$\begin{aligned}
&\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \langle \ddot{\mathbf{v}}(t), \dot{\boldsymbol{\varphi}}(t) \rangle \,\mathrm{d}t + \sum_{n=1}^{N-1} \left([\dot{\mathbf{v}}(t)]_n, \dot{\boldsymbol{\varphi}}_n^+ \right)_{L^2} + \left(\dot{\mathbf{v}}_0^+, \dot{\boldsymbol{\varphi}}_0^+ \right)_{L^2} \\
&\quad + \gamma^2 \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left(\mathbf{v}(t), \dot{\boldsymbol{\varphi}}(t) \right)_{L^2} \,\mathrm{d}t + \gamma^2 \sum_{n=1}^{N-1} \left([\mathbf{v}(t)]_n, \boldsymbol{\varphi}_n^+ \right)_{L^2} + \gamma^2 \left(\mathbf{v}_0^+, \boldsymbol{\varphi}_0^+ \right)_{L^2} \\
&\quad + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \sum_{i,\alpha=1}^{d} \mathrm{e}^{-\gamma t} \left(S_{i\alpha}(\nabla \mathbf{v}(t) \mathrm{e}^{\gamma t}), \partial_\alpha \dot{\boldsymbol{\varphi}}_i(t) \right)_{L^2} \,\mathrm{d}t + 2\gamma \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left(\dot{\mathbf{v}}(t), \dot{\boldsymbol{\varphi}}(t) \right)_{L^2} \,\mathrm{d}t \\
&\quad + \sum_{n=0}^{N-1} \frac{1}{2} \left\{ \int_0^1 \tilde{a}(\tau \mathbf{v}_n^+ \mathrm{e}^{\gamma t_n}; \mathbf{v}_n^+, \boldsymbol{\varphi}_n^+) \,\mathrm{d}\tau - \int_0^1 \tilde{a}(\tau \mathbf{v}_n^- \mathrm{e}^{\gamma t_n}; \mathbf{v}_n^-, \boldsymbol{\varphi}_n^-) \,\mathrm{d}\tau \right\} \\
&= \left(\mathbf{v}_1, \dot{\boldsymbol{\varphi}}_0^+ \right)_{L^2} + \gamma^2 \left(\mathbf{v}_0, \boldsymbol{\varphi}_0^+ \right)_{L^2} + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left(\tilde{\mathbf{f}}(t), \dot{\boldsymbol{\varphi}}(t) \right)_{L^2} \,\mathrm{d}t, \qquad \text{for } \boldsymbol{\varphi} \in \mathcal{V}_{kh}^{\mathbf{q}}.
\end{aligned} \qquad (3.3.17)$$

By considering the nonlinear elliptic projection of $\mathbf{v}(t)$ (cf. equality (3.3.2)), we can replace

$$\sum_{i,\alpha=1}^{d} \mathrm{e}^{-\gamma t} \left(S_{i\alpha}(\nabla \mathbf{v}(t)\mathrm{e}^{\gamma t}), \partial_\alpha \dot{\boldsymbol{\varphi}}_i(t)\right)_{L^2} \mathrm{d}t \text{ by } \sum_{i,\alpha=1}^{d} \mathrm{e}^{-\gamma t} \left(S_{i\alpha}(\nabla \mathbf{W}(t)\mathrm{e}^{\gamma t}), \partial_\alpha \dot{\boldsymbol{\varphi}}_i(t)\right)_{L^2} \mathrm{d}t$$

in (3.3.17). Using the continuity of $\mathbf{v}(t)$ and $\mathbf{W}(t)$ in time, we can also replace

$$\sum_{n=0}^{N-1} \left\{ \int_0^1 \tilde{a}(\tau \mathbf{v}_n^+ \mathrm{e}^{\gamma t_n}; \mathbf{v}_n^+, \boldsymbol{\varphi}_n^+)\,\mathrm{d}\tau - \int_0^1 \tilde{a}(\tau \mathbf{v}_n^- \mathrm{e}^{\gamma t_n}; \mathbf{v}_n^-, \boldsymbol{\varphi}_n^-)\,\mathrm{d}\tau \right\}$$

by

$$\sum_{n=0}^{N-1} \left\{ \int_0^1 \tilde{a}(\tau \mathbf{W}_n^+ \mathrm{e}^{\gamma t_n}; \mathbf{W}_n^+, \boldsymbol{\varphi}_n^+)\,\mathrm{d}\tau - \int_0^1 \tilde{a}(\tau \mathbf{W}_n^- \mathrm{e}^{\gamma t_n}; \mathbf{W}_n^-, \boldsymbol{\varphi}_n^-)\,\mathrm{d}\tau \right\}$$

in (3.3.17). Subtracting the resulting equality from (3.3.16), we have

$$\begin{aligned}
&\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left\langle \ddot{\boldsymbol{\theta}}(t) + \ddot{\boldsymbol{\rho}}_1(t) + \ddot{\boldsymbol{\rho}}_2(t), \dot{\boldsymbol{\varphi}}(t) \right\rangle \mathrm{d}t + \sum_{n=1}^{N-1} \left([\dot{\boldsymbol{\theta}}(t) + \dot{\boldsymbol{\rho}}_1(t) + \dot{\boldsymbol{\rho}}_2(t)]_n, \dot{\boldsymbol{\varphi}}_n^+ \right)_{L^2} \\
&\quad + \left(\dot{\mathbf{v}}_{\mathrm{DG},0}^+ - \dot{\mathbf{v}}_0^+, \dot{\boldsymbol{\varphi}}_0^+ \right)_{L^2} + 2\gamma \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left(\dot{\boldsymbol{\theta}}(t) + \dot{\boldsymbol{\rho}}_1(t) + \dot{\boldsymbol{\rho}}_2(t), \dot{\boldsymbol{\varphi}}(t) \right)_{L^2} \mathrm{d}t \\
&\quad + \gamma^2 \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left(\boldsymbol{\theta}(t) + \boldsymbol{\rho}_1(t) + \boldsymbol{\rho}_2(t), \dot{\boldsymbol{\varphi}}(t) \right)_{L^2} \mathrm{d}t \\
&\quad + \gamma^2 \sum_{n=1}^{N-1} \left([\boldsymbol{\theta}(t) + \boldsymbol{\rho}_1(t) + \boldsymbol{\rho}_2(t)]_n, \boldsymbol{\varphi}_n^+ \right)_{L^2} + \gamma^2 \left(\mathbf{v}_{\mathrm{DG},0}^+ - \mathbf{v}_0^+, \boldsymbol{\varphi}_0^+ \right)_{L^2} \\
&\quad + \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \sum_{i,\alpha=1}^{d} \mathrm{e}^{-\gamma t} \left(S_{i\alpha}(\nabla \mathbf{v}_{\mathrm{DG}}(t)\mathrm{e}^{\gamma t}) - S_{i\alpha}(\nabla \mathbf{W}(t)\mathrm{e}^{\gamma t}), \partial_\alpha \dot{\boldsymbol{\varphi}}_i \right)_{L^2} \mathrm{d}t \\
&\quad + \sum_{n=0}^{N-1} \frac{1}{2} \left\{ \int_0^1 \tilde{a}(\tau \mathbf{v}_{\mathrm{DG},n}^+ \mathrm{e}^{\gamma t_n}; \mathbf{v}_{\mathrm{DG},n}^+, \boldsymbol{\varphi}_n^+)\,\mathrm{d}\tau - \int_0^1 \tilde{a}(\tau \mathbf{W}_n^+ \mathrm{e}^{\gamma t_n}; \mathbf{W}_n^+, \boldsymbol{\varphi}_n^+)\,\mathrm{d}\tau \right\} \\
&\quad + \sum_{n=0}^{N-1} \frac{1}{2} \left\{ \int_0^1 \tilde{a}(\tau \mathbf{W}_n^- \mathrm{e}^{\gamma t_n}; \mathbf{W}_n^-, \boldsymbol{\varphi}_n^-)\,\mathrm{d}\tau - \int_0^1 \tilde{a}(\tau \mathbf{v}_{\mathrm{DG},n}^- \mathrm{e}^{\gamma t_n}; \mathbf{v}_{\mathrm{DG},n}^-, \boldsymbol{\varphi}_n^-)\,\mathrm{d}\tau \right\} \\
&= \left(\mathbf{v}_{1,h} - \mathbf{v}_1, \dot{\boldsymbol{\varphi}}_0^+ \right)_{L^2} + \gamma^2 \left(\mathbf{v}_{0,h} - \mathbf{v}_0, \boldsymbol{\varphi}_0^+ \right)_{L^2}
\end{aligned} \tag{3.3.18}$$

for $\boldsymbol{\varphi} \in \mathcal{V}_{kh}^{\mathbf{q}}$. Now we consider the integral on $I_n = (t_{n-1}, t_n]$ only,

$$\int_{t_{n-1}}^{t_n} \left\langle \ddot{\boldsymbol{\theta}}(t), \dot{\boldsymbol{\varphi}}(t) \right\rangle \mathrm{d}t + \left([\dot{\boldsymbol{\theta}}(t)]_{n-1}, \dot{\boldsymbol{\varphi}}_{n-1}^+ \right)_{L^2} + 2\gamma \int_{t_{n-1}}^{t_n} \left(\dot{\boldsymbol{\theta}}(t), \dot{\boldsymbol{\varphi}}(t) \right)_{L^2} \mathrm{d}t$$

$$
\begin{aligned}
&+\gamma^2\int_{t_{n-1}}^{t_n}(\boldsymbol{\theta}(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}+\gamma^2\left([\boldsymbol{\theta}(t)]_{n-1},\boldsymbol{\varphi}_{n-1}^{+}\right)_{L^2}\\
&+\int_{t_{n-1}}^{t_n}\sum_{i,\alpha=1}^{d}\mathrm{e}^{-\gamma t}\left(S_{i\alpha}(\nabla\mathbf{v}_{\mathrm{DG}}(t)\mathrm{e}^{\gamma t})-S_{i\alpha}(\nabla\mathbf{W}(t)\mathrm{e}^{\gamma t}),\partial_\alpha\dot{\boldsymbol{\varphi}}_i(t)\right)_{L^2}\,\mathrm{d}t\\
&+\frac{1}{2}\left\{\int_0^1\tilde{a}(\tau\mathbf{v}_{\mathrm{DG},n-1}^{+}\mathrm{e}^{\gamma t_{n-1}};\mathbf{v}_{\mathrm{DG},n-1}^{+},\boldsymbol{\varphi}_{n-1}^{+})\,\mathrm{d}\tau-\int_0^1\tilde{a}(\tau\mathbf{W}_{n-1}^{+}\mathrm{e}^{\gamma t_{n-1}};\mathbf{W}_{n-1}^{+},\boldsymbol{\varphi}_{n-1}^{+})\,\mathrm{d}\tau\right\}\\
&+\frac{1}{2}\left\{\int_0^1\tilde{a}(\tau\mathbf{W}_{n-1}^{-}\mathrm{e}^{\gamma t_{n-1}};\mathbf{W}_{n-1}^{-},\boldsymbol{\varphi}_{n-1}^{-})\,\mathrm{d}\tau-\int_0^1\tilde{a}(\tau\mathbf{v}_{\mathrm{DG},n-1}^{-}\mathrm{e}^{\gamma t_{n-1}};\mathbf{v}_{\mathrm{DG},n-1}^{-},\boldsymbol{\varphi}_{n-1}^{-})\,\mathrm{d}\tau\right\}\\
=&-\int_{t_{n-1}}^{t_n}\langle\ddot{\boldsymbol{\rho}}_1(t),\dot{\boldsymbol{\varphi}}(t)\rangle\,\mathrm{d}t-\gamma^2\int_{t_{n-1}}^{t_n}(\boldsymbol{\rho}_1(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t-2\gamma\int_{t_{n-1}}^{t_n}(\dot{\boldsymbol{\rho}}_1(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t\\
&-\left([\dot{\boldsymbol{\rho}}_1]_{n-1},\dot{\boldsymbol{\varphi}}_{n-1}^{+}\right)_{L^2}-\gamma^2\left([\boldsymbol{\rho}_1]_{n-1},\boldsymbol{\varphi}_{n-1}^{+}\right)_{L^2}\\
&-\int_{t_{n-1}}^{t_n}\langle\ddot{\boldsymbol{\rho}}_2(t),\dot{\boldsymbol{\varphi}}(t)\rangle\,\mathrm{d}t-\gamma^2\int_{t_{n-1}}^{t_n}(\boldsymbol{\rho}_2(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t-2\gamma\int_{t_{n-1}}^{t_n}(\dot{\boldsymbol{\rho}}_2(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t\\
=&\int_{t_{n-1}}^{t_n}\langle\dot{\boldsymbol{\rho}}_1(t),\ddot{\boldsymbol{\varphi}}(t)\rangle\,\mathrm{d}t-\gamma^2\int_{t_{n-1}}^{t_n}(\boldsymbol{\rho}_1(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t-2\gamma\int_{t_{n-1}}^{t_n}(\dot{\boldsymbol{\rho}}_1(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t\\
&-\left(\dot{\boldsymbol{\rho}}_1(t_n^{-}),\dot{\boldsymbol{\varphi}}_n^{-}\right)_{L^2}+\left(\dot{\boldsymbol{\rho}}_1(t_{n-1}^{-}),\dot{\boldsymbol{\varphi}}_{n-1}^{+}\right)_{L^2}-\gamma^2\left([\boldsymbol{\rho}_1]_{n-1},\boldsymbol{\varphi}_{n-1}^{+}\right)_{L^2}\\
&-\int_{t_{n-1}}^{t_n}\langle\ddot{\boldsymbol{\rho}}_2(t),\dot{\boldsymbol{\varphi}}(t)\rangle\,\mathrm{d}t-\gamma^2\int_{t_{n-1}}^{t_n}(\boldsymbol{\rho}_2(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t-2\gamma\int_{t_{n-1}}^{t_n}(\dot{\boldsymbol{\rho}}_2(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t\\
=&-\gamma^2\int_{t_{n-1}}^{t_n}(\boldsymbol{\rho}_1(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t-2\gamma\int_{t_{n-1}}^{t_n}(\dot{\boldsymbol{\rho}}_1(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t-\int_{t_{n-1}}^{t_n}\langle\ddot{\boldsymbol{\rho}}_2(t),\dot{\boldsymbol{\varphi}}(t)\rangle\,\mathrm{d}t\\
&-\gamma^2\int_{t_{n-1}}^{t_n}(\boldsymbol{\rho}_2(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t-2\gamma\int_{t_{n-1}}^{t_n}(\dot{\boldsymbol{\rho}}_2(t),\dot{\boldsymbol{\varphi}}(t))_{L^2}\,\mathrm{d}t,
\end{aligned}
\tag{3.3.19}
$$

where we have used the fact that $\boldsymbol{\rho}_2(t)$ and $\dot{\boldsymbol{\rho}}_2(t)$ are continuous in time and properties (3.3.4)–(3.3.7). By Taylor's theorem with an integral remainder, we have

$$
\begin{aligned}
&S_{i\alpha}(\nabla\mathbf{v}_{\mathrm{DG}}(t)\mathrm{e}^{\gamma t})=S_{i\alpha}(\nabla\mathbf{W}(t)\mathrm{e}^{\gamma t})\\
&+\sum_{j,\beta=1}^{d}\mathrm{e}^{\gamma t}\partial_\beta(\mathbf{v}_{\mathrm{DG}}(t)-\mathbf{W}(t))_j\int_0^1\frac{\partial}{\partial\eta_{j\beta}}S_{i\alpha}(\nabla\mathbf{W}(t)\mathrm{e}^{\gamma t}+\tau(\nabla\mathbf{v}_{\mathrm{DG}}(t)-\nabla\mathbf{W}(t))\mathrm{e}^{\gamma t})\,\mathrm{d}\tau.
\end{aligned}
$$

If $\nabla\mathbf{v}_{\mathrm{DG}}(t)\mathrm{e}^{\gamma t}\in\mathcal{Z}_\delta,\nabla\mathbf{W}(t)\mathrm{e}^{\gamma t}\in\mathcal{Z}_\delta$ for each $t\in[0,T]$, we have $\nabla\mathbf{W}(t)\mathrm{e}^{\gamma t}+\tau(\nabla\mathbf{v}_{\mathrm{DG}}(t)-\nabla\mathbf{W}(t))\mathrm{e}^{\gamma t}\in\mathcal{Z}_\delta$ for $0\le\tau\le1$ by the convexity of $\mathcal{Z}_\delta$. This implies that the term in the integral remainder is well-defined. Thus, we can write

$$\begin{aligned}
&\sum_{i,\alpha=1}^{d} \mathrm{e}^{-\gamma t} \left(S_{i\alpha}(\nabla \mathbf{v}_{\mathrm{DG}}(t)\mathrm{e}^{\gamma t}) - S_{i\alpha}(\nabla \mathbf{W}(t)\mathrm{e}^{\gamma t}), \partial_\alpha \dot{\boldsymbol{\varphi}}_i(t)\right)_{L^2} \,\mathrm{d}t \\
&= \int_0^1 \sum_{i,\alpha,j,\beta=1}^{d} \left(A_{i\alpha j\beta}(\nabla \mathbf{W}(t)\mathrm{e}^{\gamma t} + \tau(\nabla \mathbf{v}_{\mathrm{DG}}(t) - \nabla \mathbf{W}(t))\mathrm{e}^{\gamma t})\partial_\beta(\mathbf{v}_{\mathrm{DG}}(t) - \mathbf{W}(t))_j, \partial_\alpha \dot{\boldsymbol{\varphi}}_i(t)\right)_{L^2} \,\mathrm{d}\tau \\
&= \int_0^1 \tilde{a}(\mathbf{W}(t)\mathrm{e}^{\gamma t} + \tau(\mathbf{v}_{\mathrm{DG}}(t) - \mathbf{W}(t))\mathrm{e}^{\gamma t}; \mathbf{v}_{\mathrm{DG}}(t) - \mathbf{W}(t), \dot{\boldsymbol{\varphi}}(t)) \,\mathrm{d}\tau \\
&:= \mathsf{A}(\mathbf{v}_{\mathrm{DG}}(t)\mathrm{e}^{\gamma t}; \mathbf{v}_{\mathrm{DG}}(t) - \mathbf{W}(t), \dot{\boldsymbol{\varphi}}(t)),
\end{aligned}$$

for $\boldsymbol{\varphi} \in \mathcal{V}_{kh}^{\mathbf{q}}$. Analogously,

$$\begin{aligned}
&\int_0^1 \tilde{a}(\tau \mathbf{v}^{\pm}_{\mathrm{DG},n-1}\mathrm{e}^{\gamma t_{n-1}}; \mathbf{v}^{\pm}_{\mathrm{DG},n-1}, \boldsymbol{\varphi}^{\pm}_{n-1}) \,\mathrm{d}\tau - \int_0^1 \tilde{a}(\tau \mathbf{W}^{\pm}_{n-1}\mathrm{e}^{\gamma t_{n-1}}; \mathbf{W}^{\pm}_{n-1}, \boldsymbol{\varphi}^{\pm}_{n-1}) \,\mathrm{d}\tau \\
&= \sum_{i,\alpha=1}^{d} \mathrm{e}^{-\gamma t_{n-1}} \left(S_{i\alpha}(\nabla \mathbf{v}^{\pm}_{\mathrm{DG},n-1}\mathrm{e}^{\gamma t_{n-1}}) - S_{i\alpha}(\nabla \mathbf{W}^{\pm}_{n-1}\mathrm{e}^{\gamma t_{n-1}}), \partial_\alpha \boldsymbol{\varphi}_i(t^{\pm}_{n-1})\right)_{L^2} \\
&:= \mathsf{A}(\mathbf{v}^{\pm}_{\mathrm{DG},n-1}\mathrm{e}^{\gamma t_{n-1}}; \mathbf{v}^{\pm}_{\mathrm{DG},n-1} - \mathbf{W}^{\pm}_{n-1}, \boldsymbol{\varphi}^{\pm}_{n-1}).
\end{aligned}$$

Since $\mathbf{W}(t_n^{\pm}) - \Pi_k \mathbf{W}(t_n^{\pm}) = 0$ for each $n = 1, 2, \ldots, N$, our equation (3.3.19) becomes

$$\begin{aligned}
&\int_{t_{n-1}}^{t_n} \left\langle \ddot{\boldsymbol{\theta}}(t), \dot{\boldsymbol{\varphi}}(t) \right\rangle \,\mathrm{d}t + \left([\dot{\boldsymbol{\theta}}(t)]_{n-1}, \dot{\boldsymbol{\varphi}}^{+}_{n-1}\right)_{L^2} + 2\gamma \int_{t_{n-1}}^{t_n} \left(\dot{\boldsymbol{\theta}}(t), \dot{\boldsymbol{\varphi}}(t)\right)_{L^2} \,\mathrm{d}t \\
&\quad + \gamma^2 \int_{t_{n-1}}^{t_n} (\boldsymbol{\theta}(t), \dot{\boldsymbol{\varphi}}(t))_{L^2} \,\mathrm{d}t + \gamma^2 \left([\boldsymbol{\theta}(t)]_{n-1}, \boldsymbol{\varphi}^{+}_{n-1}\right)_{L^2} \\
&\quad + \int_{t_{n-1}}^{t_n} \mathsf{A}(\mathbf{v}_{\mathrm{DG}}(t)\mathrm{e}^{\gamma t}; \mathbf{v}_{\mathrm{DG}}(t) - \Pi_k \mathbf{W}(t), \dot{\boldsymbol{\varphi}}(t)) \,\mathrm{d}t + \frac{1}{2}\mathsf{A}(\mathbf{v}^{+}_{\mathrm{DG},n-1}\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}^{+}_{n-1}, \boldsymbol{\varphi}^{+}_{n-1}) \\
&= -\gamma^2 \int_{t_{n-1}}^{t_n} (\boldsymbol{\rho}_1(t), \dot{\boldsymbol{\varphi}}(t))_{L^2} \,\mathrm{d}t - 2\gamma \int_{t_{n-1}}^{t_n} (\dot{\boldsymbol{\rho}}_1(t), \dot{\boldsymbol{\varphi}}(t))_{L^2} \,\mathrm{d}t - \int_{t_{n-1}}^{t_n} \langle \ddot{\boldsymbol{\rho}}_2(t), \dot{\boldsymbol{\varphi}}(t) \rangle \,\mathrm{d}t \\
&\quad - \gamma^2 \int_{t_{n-1}}^{t_n} (\boldsymbol{\rho}_2(t), \dot{\boldsymbol{\varphi}}(t))_{L^2} \,\mathrm{d}t - 2\gamma \int_{t_{n-1}}^{t_n} (\dot{\boldsymbol{\rho}}_2(t), \dot{\boldsymbol{\varphi}}(t))_{L^2} \,\mathrm{d}t \\
&\quad + \int_{t_{n-1}}^{t_n} \mathsf{A}(\mathbf{v}_{\mathrm{DG}}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\boldsymbol{\varphi}}(t)) \,\mathrm{d}t + \frac{1}{2}\mathsf{A}(\mathbf{v}^{-}_{\mathrm{DG},n-1}\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}^{-}_{n-1}, \boldsymbol{\varphi}^{-}_{n-1})
\end{aligned}$$

for $\boldsymbol{\varphi} \in \mathcal{V}_{kh}^{\mathbf{q}}$. Consider the following subset of $\mathcal{V}_{kh}^{\mathrm{q}}$ defined by

$$\mathcal{F} := \Bigg\{ \boldsymbol{\psi} \in \mathcal{V}_{kh}^{\mathbf{q}} \mid \text{ for each } j = 1, 2, \ldots, N, \ \|\boldsymbol{\psi}(t_j^-) - \Pi_k \mathbf{W}(t_j^-)\|_{H^1(\Omega)}^2$$
$$+ \|\partial_t(\boldsymbol{\psi}(t_j^-) - \Pi_k \mathbf{W}(t_j^-))\|_{L^2(\Omega)}^2 + \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\partial_t(\boldsymbol{\psi}(t) - \Pi_k \mathbf{W}(t))\|_{L^2(\Omega)}^2 \, \mathrm{d}t$$
$$\leq C_*(\mathbf{v}) \left(\sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} \right), \mathrm{e}^{\gamma t} \nabla \boldsymbol{\psi}(t) \in \mathcal{Z}_\delta \Bigg\}, \tag{3.3.20}$$

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

$$\|\boldsymbol{\varphi}\|_{\mathcal{F}} = \max_{t \in I_n, 1 \leq n \leq N} (\|\boldsymbol{\varphi}(t)\|_{H^1(\Omega)} + \|\dot{\boldsymbol{\varphi}}(t)\|_{L^2(\Omega)})$$

for $\boldsymbol{\varphi} \in \mathcal{V}_{kh}^{\mathbf{q}}$. With this notation, we are ready to define a fixed point mapping $\mathcal{N}$ on $\mathcal{F}$ as follows: if $\boldsymbol{\phi} \in \mathcal{F}$, the image $\mathbf{v}_\phi := \mathcal{N}(\boldsymbol{\phi})$ is given by the relation

$$\mathbf{v}_\phi(0) = \mathbf{v}_{0,h}, \ \ \dot{\mathbf{v}}_\phi(0) = \mathbf{v}_{1,h}, \tag{3.3.21}$$

$$\begin{aligned}
&\int_{t_{n-1}}^{t_n} \left\langle \ddot{\boldsymbol{\theta}}_\phi(t), \dot{\boldsymbol{\varphi}}(t) \right\rangle \mathrm{d}t + \left([\dot{\boldsymbol{\theta}}_\phi(t)]_{n-1}, \dot{\boldsymbol{\varphi}}(t_{n-1}^+) \right)_{L^2} + 2\gamma \int_{t_{n-1}}^{t_n} \left(\dot{\boldsymbol{\theta}}_\phi(t), \dot{\boldsymbol{\varphi}}(t) \right)_{L^2} \mathrm{d}t \\
&\quad + \gamma^2 \int_{t_{n-1}}^{t_n} (\boldsymbol{\theta}_\phi(t), \dot{\boldsymbol{\varphi}}(t))_{L^2} \, \mathrm{d}t + \gamma^2 \left([\boldsymbol{\theta}_\phi(t)]_{n-1}, \boldsymbol{\varphi}(t_{n-1}^+) \right)_{L^2} \\
&\quad + \int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_\phi(t), \dot{\boldsymbol{\varphi}}(t)) \, \mathrm{d}t + \frac{1}{2} \mathsf{A}(\boldsymbol{\phi}(t_{n-1}^+)\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_\phi(t_{n-1}^+), \boldsymbol{\varphi}(t_{n-1}^+)) \\
&= -\gamma^2 \int_{t_{n-1}}^{t_n} (\boldsymbol{\rho}_1(t), \dot{\boldsymbol{\varphi}}(t))_{L^2} \, \mathrm{d}t - 2\gamma \int_{t_{n-1}}^{t_n} (\dot{\boldsymbol{\rho}}_1(t), \dot{\boldsymbol{\varphi}}(t))_{L^2} \, \mathrm{d}t - \int_{t_{n-1}}^{t_n} \langle \ddot{\boldsymbol{\rho}}_2(t), \dot{\boldsymbol{\varphi}}(t) \rangle \, \mathrm{d}t \\
&\quad - \gamma^2 \int_{t_{n-1}}^{t_n} (\boldsymbol{\rho}_2(t), \dot{\boldsymbol{\varphi}}(t))_{L^2} \, \mathrm{d}t - 2\gamma \int_{t_{n-1}}^{t_n} (\dot{\boldsymbol{\rho}}_2(t), \dot{\boldsymbol{\varphi}}(t))_{L^2} \, \mathrm{d}t \\
&\quad + \int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\boldsymbol{\varphi}}(t)) \, \mathrm{d}t + \frac{1}{2} \mathsf{A}(\boldsymbol{\phi}(t_{n-1}^-)\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_\phi(t_{n-1}^-), \boldsymbol{\varphi}(t_{n-1}^-)),
\end{aligned} \tag{3.3.22}$$

where $\boldsymbol{\theta}_\phi = \mathbf{v}_\phi - \Pi_k \mathbf{W}$.

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 (3.3.22) has a unique fixed point in $\mathcal{F}$. If $\mathbf{v}_{\mathrm{DG}} \in \mathcal{F}$ is this fixed point, then $\mathbf{v}_{\mathrm{DG}}$ is a solution to (3.2.10).

3.3.2. Auxiliary results If we take $\boldsymbol{\varphi} = \boldsymbol{\theta}_{\boldsymbol{\phi}}$ in (3.3.22), then the nonlinear term inside the integral becomes

$$\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_\phi(t), \dot{\boldsymbol{\varphi}}(t)) = \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_\phi(t), \dot{\boldsymbol{\theta}}_\phi(t)).$$

Following the proof in [56], it is crucial to replace the expression $\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_\phi(t), \dot{\boldsymbol{\theta}}_\phi(t))$ by

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_\phi(t), \boldsymbol{\theta}_\phi(t)) - \frac{1}{2}\sum_{i,\alpha,j,\beta=1}^{d}\int_0^1 \left(\partial_t A^\tau_{i\alpha j\beta}\partial_\beta \boldsymbol{\theta}_{\phi,j}(t), \partial_\alpha \boldsymbol{\theta}_{\phi,i}(t)\right)_{L^2} \mathrm{d}\tau,$$

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

$$\mathsf{A}_t(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\varphi}(t), \boldsymbol{\psi}(t)) := \frac{1}{2}\sum_{i,\alpha,j,\beta=1}^{d}\int_0^1 \left(\partial_t A^\tau_{i\alpha j\beta}\partial_\beta \boldsymbol{\varphi}_j(t), \partial_\alpha \boldsymbol{\psi}_i(t)\right)_{L^2} \mathrm{d}\tau$$

for $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathcal{V}^{\mathbf{q}}_{hk}$, $t \in I_n, n = 1, 2, \ldots, N$.

Lemma 3.6. *Under the assumptions stated in Theorem 3.2, there exists a constant* $C_\tau > 0$ *such that, for* $t \in I_n$, $n = 1, 2, \ldots, N$,

$$|\mathsf{A}_t(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\varphi}(t), \boldsymbol{\psi}(t))| \le C_\tau \|\nabla\boldsymbol{\varphi}(t)\|_{L^2(\Omega)}\|\nabla\boldsymbol{\psi}(t)\|_{L^2(\Omega)}. \tag{3.3.23}$$

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

$$\begin{aligned}
&\partial_t[A_{i\alpha j\beta}(\nabla\mathbf{W}(t)\mathrm{e}^{\gamma t} + \tau(\nabla\boldsymbol{\phi}(t) - \nabla\mathbf{W}(t))\mathrm{e}^{\gamma t})] \\
&= \sum_{k,m=1}^{d} \frac{\partial A^\tau_{i\alpha j\beta}}{\partial \eta_{km}}\partial_m \left(\partial_t(\mathbf{W}_k(t)\mathrm{e}^{\gamma t}) + \tau\partial_t((\boldsymbol{\phi}_k(t) - \mathbf{W}_k(t))\mathrm{e}^{\gamma t})\right).
\end{aligned}$$

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

$$\Big| \sum_{i,\alpha,j,\beta=1}^{d} \frac{1}{2} \int_0^1 \left(\partial_t [A_{i\alpha j\beta}(\nabla \mathbf{W}(t) \mathrm{e}^{\gamma t} + \tau \nabla (\boldsymbol{\phi}(t) - \mathbf{W}(t)) \mathrm{e}^{\gamma t})] \partial_\beta \boldsymbol{\varphi}_j(t), \partial_\alpha \boldsymbol{\psi}_i(t)\right)_{L^2} \mathrm{d}\tau \Big|$$
$$\leq c \left(\|\nabla \dot{\mathbf{W}}(t)\|_{L^\infty(\Omega)} + \|\nabla \mathbf{W}(t)\|_{L^\infty(\Omega)} \right) \|\nabla \boldsymbol{\varphi}(t)\|_{L^2(\Omega)} \|\nabla \boldsymbol{\psi}(t)\|_{L^2(\Omega)}$$
$$+ c \left(\|\nabla (\dot{\boldsymbol{\phi}}(t) - \dot{\mathbf{W}}(t))\|_{L^\infty(\Omega)} + \|\nabla (\boldsymbol{\phi}(t) - \mathbf{W}(t))\|_{L^\infty(\Omega)} \right) \|\nabla \boldsymbol{\varphi}(t)\|_{L^2(\Omega)} \|\nabla \boldsymbol{\psi}(t)\|_{L^2(\Omega)}$$
$$\leq \tilde{C} h^{-\frac{d}{2}} \left(\|\nabla (\dot{\boldsymbol{\phi}}(t) - \dot{\mathbf{W}}(t))\|_{L^2(\Omega)} + \|\nabla (\boldsymbol{\phi}(t) - \mathbf{W}(t))\|_{L^2(\Omega)} \right) \|\nabla \boldsymbol{\varphi}(t)\|_{L^2(\Omega)} \|\nabla \boldsymbol{\psi}(t)\|_{L^2(\Omega)}$$
$$+ \tilde{C} \|\nabla \boldsymbol{\varphi}(t)\|_{L^2(\Omega)} \|\nabla \boldsymbol{\psi}(t)\|_{L^2(\Omega)},$$

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

$$\|\nabla (\boldsymbol{\phi}(t) - \mathbf{W}(t))\|_{L^2(\Omega)} \leq \|\nabla (\boldsymbol{\phi}(t) - \Pi_k \mathbf{W}(t))\|_{L^2(\Omega)} + \|\nabla (\Pi_k \mathbf{W}(t) - \mathbf{W}(t))\|_{L^2(\Omega)}.$$

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

$$\|\nabla (\boldsymbol{\phi}(t) - \Pi_k \mathbf{W}(t))\|_{L^2(\Omega)}$$
$$\leq \|\nabla (\boldsymbol{\phi}(t_n^-) - \Pi_k \mathbf{W}(t_n^-))\|_{L^2(\Omega)} + \int_t^{t_n} \|\partial_s (\nabla \boldsymbol{\phi}(s) - \nabla \Pi_k \mathbf{W}(s))\|_{L^2(\Omega)} \, \mathrm{d}s$$
$$\leq \|\nabla (\boldsymbol{\phi}(t_n^-) - \Pi_k \mathbf{W}(t_n^-))\|_{L^2(\Omega)} + \int_{t_{n-1}}^{t_n} \|\partial_t (\nabla \boldsymbol{\phi}(t) - \nabla \Pi_k \mathbf{W}(t))\|_{L^2(\Omega)} \, \mathrm{d}t$$
$$\leq \|\nabla (\boldsymbol{\phi}(t_n^-) - \Pi_k \mathbf{W}(t_n^-))\|_{L^2(\Omega)} + C_0 h^{-1} \sqrt{k_n} \left(\int_{t_{n-1}}^{t_n} \|\partial_t (\boldsymbol{\phi}(t) - \Pi_k \mathbf{W}(t))\|_{L^2(\Omega)}^2 \, \mathrm{d}t \right)^{\frac{1}{2}}$$
$$\leq C(\mathbf{v}) \left(\sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}},$$

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

$$\|\nabla (\Pi_k \mathbf{W}(t) - \mathbf{W}(t))\|_{L^2(\Omega)}$$

$$
\begin{aligned}
&\leq \|\nabla(\Pi_k \mathbf{W}(t_n^-) - \mathbf{W}(t_n^-))\|_{L^2(\Omega)} + \int_t^{t_n} \|\partial_s(\nabla\Pi_k\mathbf{W}(s) - \nabla\mathbf{W}(s))\|_{L^2(\Omega)}\,\mathrm{d}s \\
&\leq \int_{t_{n-1}}^{t_n} \|\partial_t(\nabla\Pi_k\mathbf{W}(t) - \nabla\mathbf{W}(t))\|_{L^2(\Omega)}\,\mathrm{d}t \qquad (\text{ since } \Pi_k\mathbf{W}(t_n^-) = \mathbf{W}(t_n^-)\ ) \\
&\leq \sqrt{k_n}\bigg(\int_{t_{n-1}}^{t_n} \|\partial_t(\nabla\Pi_k\mathbf{W}(t) - \nabla\mathbf{W}(t))\|^2_{L^2(\Omega)}\,\mathrm{d}t\bigg)^{\frac{1}{2}} \\
&\leq C\frac{k_n^{q_n+\frac{1}{2}}}{q_n^{s-1}}\|\mathbf{W}\|_{H^s(I_n;[H_0^1(\Omega)]^d)}, \qquad \text{for } \mathbf{W} \in H^s([0,T];[H_0^1(\Omega)]^d),
\end{aligned}
$$

where we have used inequality (3.3.9) with the $[L^2(\Omega)]^d$ norm in space replaced by the $[H^1(\Omega)]^d$ semi-norm. Thus,

$$
\max_{t\in I_n, 1\leq n\leq N} \|\nabla(\boldsymbol{\phi}(t) - \mathbf{W}(t))\|_{L^2(\Omega)} \leq C(\mathbf{v})\left(\sum_{i=1}^n k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}}\right)^{\frac{1}{2}} + C\frac{k_n^{q_n+\frac{1}{2}}}{q_n^{s-1}}\|\mathbf{W}\|_{H^s(I_n;[H_0^1(\Omega)]^d)}. \tag{3.3.24}
$$

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

$$
\|\nabla(\dot{\boldsymbol{\phi}}(t) - \dot{\mathbf{W}}(t))\|_{L^2(\Omega)} \leq \|\partial_t(\nabla\boldsymbol{\phi}(t) - \nabla\Pi_k\mathbf{W}(t))\|_{L^2(\Omega)} + \|\partial_t(\nabla\Pi_k\mathbf{W}(t) - \nabla\mathbf{W}(t))\|_{L^2(\Omega)}.
$$

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

$$
\begin{aligned}
&\|\partial_t(\nabla\boldsymbol{\phi}(t) - \nabla\Pi_k\mathbf{W}(t))\|_{L^2(\Omega)} \\
&\leq C_0 h^{-1}\|\partial_t(\boldsymbol{\phi}(t) - \Pi_k\mathbf{W}(t))\|_{L^2(\Omega)} \\
&\leq C_0 h^{-1}\|\partial_t(\boldsymbol{\phi}(t_n^-) - \Pi_k\mathbf{W}(t_n^-))\|_{L^2(\Omega)} + C_0 h^{-1}\int_t^{t_n} \|\partial_{ss}(\boldsymbol{\phi}(s) - \Pi_k\mathbf{W}(s))\|_{L^2(\Omega)}\,\mathrm{d}s \\
&\leq C_0 h^{-1}\|\partial_t(\boldsymbol{\phi}(t_n^-) - \Pi_k\mathbf{W}(t_n^-))\|_{L^2(\Omega)} + C_0 h^{-1}\int_{t_{n-1}}^{t_n} \|\partial_{tt}(\boldsymbol{\phi}(t) - \Pi_k\mathbf{W}(t))\|_{L^2(\Omega)}\,\mathrm{d}t.
\end{aligned}
$$

Since $\boldsymbol{\phi} \in \mathcal{F}$, we have

$$
\|\partial_t(\boldsymbol{\phi}(t_n^-) - \Pi_k\mathbf{W}(t_n^-))\|_{L^2(\Omega)} \leq C_*(\mathbf{v})\left(\sum_{i=1}^n k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}}\right)^{\frac{1}{2}}.
$$

Using the inverse inequality in time (cf. inequality (3.3.11)), we obtain

$$
\int_{t_{n-1}}^{t_n} \|\partial_{tt}(\boldsymbol{\phi}(t) - \Pi_k\mathbf{W}(t))\|_{L^2(\Omega)}\,\mathrm{d}t \leq \sqrt{k_n}\left(\int_{t_{n-1}}^{t_n} \|\partial_{tt}(\boldsymbol{\phi}(t) - \Pi_k\mathbf{W}(t))\|^2_{L^2(\Omega)}\,\mathrm{d}t\right)^{\frac{1}{2}}
$$

$$
\begin{aligned}
&\le C_2 \frac{1}{\sqrt{k_n}} \left(\int_{t_{n-1}}^{t_n} \|\partial_t(\boldsymbol{\phi}(t) - \Pi_k \mathbf{W}(t))\|_{L^2(\Omega)}^2 \,\mathrm{d}t \right)^{\frac{1}{2}} \\
&\le C(\mathbf{v}) \frac{1}{\sqrt{k_n}} \left(\sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}}.
\end{aligned}
$$

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

$$
\begin{aligned}
&\|\partial_t(\nabla \Pi_k \mathbf{W}(t) - \nabla \mathbf{W}(t))\|_{L^2(\Omega)} \\
&\le \|\partial_t(\nabla \Pi_k \mathbf{W}(t_n^-) - \nabla \mathbf{W}(t_n^-))\|_{L^2(\Omega)} + \int_t^{t_n} \|\partial_{ss}(\nabla \Pi_k \mathbf{W}(s) - \nabla \mathbf{W}(s))\|_{L^2(\Omega)} \,\mathrm{d}s \\
&\le \int_{t_{n-1}}^{t_n} \|\partial_{tt}(\nabla \Pi_k \mathbf{W}(t) - \nabla \mathbf{W}(t))\|_{L^2(\Omega)} \,\mathrm{d}t \\
&\le \sqrt{k_n} \left(\int_{t_{n-1}}^{t_n} \|\partial_{tt}(\nabla \Pi_k \mathbf{W}(t) - \nabla \mathbf{W}(t))\|_{L^2(\Omega)}^2 \,\mathrm{d}t \right)^{\frac{1}{2}} \\
&\le C \frac{k_n^{q_n - \frac{1}{2}}}{q_n^{s-3}} \|\mathbf{W}\|_{H^s(I_n;[H_0^1(\Omega)]^d)}, \qquad \text{for } \mathbf{W} \in H^s(0,T;[H_0^1(\Omega)]^d).
\end{aligned}
$$

Thus

$$
\begin{aligned}
\max_{t \in I_n, 1 \le n \le N} \|\nabla \dot{\boldsymbol{\phi}}(t) - \nabla \dot{\mathbf{W}}(t)\|_{L^2(\Omega)} \le\ & C(\mathbf{v})(h^{-1} + h^{-1} k_n^{-\frac{1}{2}}) \left(\sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} \\
&+ C \frac{k_n^{q_n - \frac{1}{2}}}{q_n^{s-3}} \|\mathbf{W}\|_{H^s(0,T;[H_0^1(\Omega)]^d)}. \qquad (3.3.25)
\end{aligned}
$$

Combining (3.3.24) and (3.3.25), we have

$$
\begin{aligned}
&\tilde{C} h^{-\frac{d}{2}} \max_{t \in I_n, 1 \le n \le N} \Big(\|\nabla \boldsymbol{\phi}(t) - \nabla \mathbf{W}(t)\|_{L^2(\Omega)} + \|\nabla \dot{\boldsymbol{\phi}}(t) - \nabla \dot{\mathbf{W}}(t)\|_{L^2(\Omega)} \Big) \\
&\le \tilde{C} h^{-\frac{d}{2}} C(\mathbf{v})(1 + h^{-1} + h^{-1} k_n^{-\frac{1}{2}}) \left(\sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} \\
&\quad + \tilde{C} h^{-\frac{d}{2}} C \left(\frac{k_n^{q_n + \frac{1}{2}}}{q_n^{s-1}} + \frac{k_n^{q_n - \frac{1}{2}}}{q_n^{s-3}} \right) \|\mathbf{W}\|_{H^s(0,T;[H_0^1(\Omega)]^d)}. \qquad (3.3.26)
\end{aligned}
$$

Since $r > \frac{d}{2} + 1$, $k_i^{q_i - \frac{1}{2}} = o(h^{1+\frac{d}{2}})$ and $\mu k_i \le h^2 \le \nu_i k_i$ for each $i = 1, 2, \ldots, N$, we can choose $h_0 > 0$ such that for $h \le h_0$, the right-hand side of (3.3.26) is bounded by 1. Thus, (3.3.23) follows by taking $C_\tau = \tilde{C} + 1$. The constant C_τ defined in this way does not depend on $C_*(\mathbf{v})$. $\square$

3.3.3. 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 $\boldsymbol{\phi}, \boldsymbol{\varphi} \in \mathcal{F}$,

$$d(\boldsymbol{\phi}, \boldsymbol{\varphi}) := \max_{t \in I_n, 1 \le n \le N} \Big(\|\boldsymbol{\phi}(t) - \boldsymbol{\varphi}(t)\|_{H^1(\Omega)} + \|\dot{\boldsymbol{\phi}}(t) - \dot{\boldsymbol{\varphi}}(t)\|_{L^2(\Omega)} \Big).$$

Existence of a fixed point of $\mathcal{N}$ in $\mathcal{F}$

For (a), we first observe that $\mathcal{N}$ is well-defined. Indeed, if $\boldsymbol{\phi} \in \mathcal{F}$, since $\nabla \mathbf{W} \mathrm{e}^{\gamma t} \in \mathcal{Z}_\delta$, $\nabla \mathbf{W} \mathrm{e}^{\gamma t} + \tau(\nabla \boldsymbol{\phi} - \nabla \mathbf{W}) \mathrm{e}^{\gamma t} \in \mathcal{Z}_\delta$ for $0 \le \tau \le 1$, and the bilinear form $\mathsf{A}(\boldsymbol{\phi}(t)\mathbf{e}^{\gamma t}; \cdot, \cdot)$ is symmetric and positive definite. Taking $\boldsymbol{\varphi} = \boldsymbol{\theta}_\phi$ in (3.3.22) and replacing $\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_\phi(t), \dot{\boldsymbol{\theta}}_\phi(t))$ by

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_\phi(t), \boldsymbol{\theta}_\phi(t)) - \frac{1}{2}\sum_{i,\alpha,j,\beta=1}^{d} \int_0^1 \left(\partial_t A^{\tau}_{i\alpha j\beta} \partial_\beta \boldsymbol{\theta}_{\phi,j}(t), \partial_\alpha \boldsymbol{\theta}_{\phi,i}(t)\right)_{L^2} \mathrm{d}\tau,$$

we obtain

$$\begin{aligned}
&\|\dot{\boldsymbol{\theta}}_\phi(t_n^-)\|^2_{L^2(\Omega)} + \|\dot{\boldsymbol{\theta}}_\phi(t_{n-1}^+)\|^2_{L^2(\Omega)} + \gamma^2\|\boldsymbol{\theta}_\phi(t_n^-)\|^2_{L^2(\Omega)} + \gamma^2\|\boldsymbol{\theta}_\phi(t_{n-1}^+)\|^2_{L^2(\Omega)} \\
&\quad + 4\gamma \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \mathsf{A}(\boldsymbol{\phi}(t_n^-)\mathrm{e}^{\gamma t_n}; \boldsymbol{\theta}_\phi(t_n^-), \boldsymbol{\theta}_\phi(t_n^-)) \\
&= -2\gamma^2 \int_{t_{n-1}}^{t_n} \Big(\boldsymbol{\rho}_1(t), \dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\Big)_{L^2} \mathrm{d}t - 4\gamma \int_{t_{n-1}}^{t_n} \Big(\dot{\boldsymbol{\rho}}_1(t), \dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\Big)_{L^2} \mathrm{d}t \\
&\quad - 2\int_{t_{n-1}}^{t_n} \Big\langle \ddot{\boldsymbol{\rho}}_2(t), \dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\Big\rangle \,\mathrm{d}t - 2\gamma^2 \int_{t_{n-1}}^{t_n} \Big(\boldsymbol{\rho}_2(t), \dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\Big)_{L^2} \mathrm{d}t \\
&\quad - 4\gamma \int_{t_{n-1}}^{t_n} \Big(\dot{\boldsymbol{\rho}}_2(t), \dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\Big)_{L^2} \mathrm{d}t + 2\Big(\dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t_{n-1}^-), \dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t_{n-1}^+)\Big)_{L^2} + 2\gamma^2 \left(\boldsymbol{\theta}_{\boldsymbol{\phi}}(t_{n-1}^-), \boldsymbol{\theta}_{\boldsymbol{\phi}}(t_{n-1}^+)\right)_{L^2} \\
&\quad + 2\int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t))\,\mathrm{d}t + \mathsf{A}(\boldsymbol{\phi}(t_{n-1}^-)e^{\gamma t_{n-1}}; \boldsymbol{\theta}_\phi(t_{n-1}^-), \boldsymbol{\theta}_\phi(t_{n-1}^-))
\end{aligned} \tag{3.3.27}$$

$$+\sum_{i,\alpha,j,\beta=1}^{d}\int_{t_{n-1}}^{t_n}\int_0^1 \left(\partial_t A^{\tau}_{i\alpha j\beta}\partial_\beta \boldsymbol{\theta}_{\phi,j}(t), \partial_\alpha \boldsymbol{\theta}_{\phi,i}(t)\right)_{L^2}\,\mathrm{d}\tau\,\mathrm{d}t.$$

By using (3.3.14), (3.3.15) and Young's inequality, we have

$$\begin{aligned}
&\Bigg| -2\gamma^2\int_{t_{n-1}}^{t_n}\Big(\boldsymbol{\rho}_1(t),\dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\Big)_{L^2}\,\mathrm{d}t - 4\gamma\int_{t_{n-1}}^{t_n}\Big(\dot{\boldsymbol{\rho}}_1(t),\dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\Big)_{L^2}\,\mathrm{d}t \\
&\quad -2\int_{t_{n-1}}^{t_n}\Big\langle\ddot{\boldsymbol{\rho}}_2(t),\dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\Big\rangle\,\mathrm{d}t - 2\gamma^2\int_{t_{n-1}}^{t_n}\Big(\boldsymbol{\rho}_2(t),\dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\Big)_{L^2}\,\mathrm{d}t - 4\gamma\int_{t_{n-1}}^{t_n}\Big(\dot{\boldsymbol{\rho}}_2(t),\dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\Big)_{L^2}\,\mathrm{d}t\Bigg| \\
&\le 3\gamma\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + C_1(\gamma)\int_{t_{n-1}}^{t_n}\Big(\|\dot{\boldsymbol{\rho}}_1(t)\|^2_{L^2(\Omega)}+\|\boldsymbol{\rho}_1(t)\|^2_{L^2(\Omega)}\Big)\,\mathrm{d}t \\
&\quad + C_2(\gamma)\int_{t_{n-1}}^{t_n}\Big(\|\ddot{\boldsymbol{\rho}}_2(t)\|^2_{L^2(\Omega)}+\|\dot{\boldsymbol{\rho}}_2(t)\|^2_{L^2(\Omega)}+\|\boldsymbol{\rho}_2(t)\|^2_{L^2(\Omega)}\Big)\,\mathrm{d}t \\
&\le 3\gamma\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + c_1(\gamma,\mathbf{v})\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} + c_2(\gamma,\mathbf{v})k_n h^{2r+2} \\
&\le 3\gamma\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + C(\gamma,\mathbf{v})\left(k_n h^{2r+2}+\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right),
\end{aligned} \tag{3.3.28}$$

where $C_i(\gamma)$ for $i = 1, 2$ are constants depending on γ only, while $C(\gamma, \mathbf{v})$ and $c_i(\gamma, \mathbf{v})$ for $i = 1, 2$ are constants depending on both γ and the exact solution $\mathbf{v}$. By the Cauchy–Schwarz inequality, we obtain

$$2\left(\dot{\boldsymbol{\theta}}_\phi(t^-_{n-1}),\dot{\boldsymbol{\theta}}_\phi(t^+_{n-1})\right)_{L^2} \le \|\dot{\boldsymbol{\theta}}_\phi(t^+_{n-1})\|^2_{L^2(\Omega)} + \|\dot{\boldsymbol{\theta}}_\phi(t^-_{n-1})\|^2_{L^2(\Omega)}, \tag{3.3.29}$$

$$2\gamma^2\left(\boldsymbol{\theta}_\phi(t^-_{n-1}),\boldsymbol{\theta}_\phi(t^+_{n-1})\right)_{L^2} \le \gamma^2\|\boldsymbol{\theta}_\phi(t^+_{n-1})\|^2_{L^2(\Omega)} + \gamma^2\|\boldsymbol{\theta}_\phi(t^-_{n-1})\|^2_{L^2(\Omega)}. \tag{3.3.30}$$

Note that

$$\begin{aligned}
\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t};\mathbf{W}(t)-\Pi_k\mathbf{W}(t),\dot{\boldsymbol{\theta}}_\phi(t)) &= \frac{\mathrm{d}}{\mathrm{d}t}\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t};\mathbf{W}(t)-\Pi_k\mathbf{W}(t),\boldsymbol{\theta}_\phi(t)) \\
&\quad - \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t};\partial_t(\mathbf{W}(t)-\Pi_k\mathbf{W}(t)),\boldsymbol{\theta}_\phi(t)) \\
&\quad - \sum_{i,\alpha,j,\beta=1}^{d}\int_0^1\Big(\partial_t A^{\tau}_{i\alpha j\beta}\partial_\beta(\mathbf{W}-\Pi_k\mathbf{W})_j,\partial_\alpha\boldsymbol{\theta}_{\phi,i}(t)\Big)_{L^2}\,\mathrm{d}\tau.
\end{aligned}$$

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

$$\int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k\mathbf{W}(t), \dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t)) = -\int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \partial_t\left(\mathbf{W}(t) - \Pi_k\mathbf{W}(t)\right), \boldsymbol{\theta}_{\boldsymbol{\phi}}(t))\,\mathrm{d}t$$
$$-\int_{t_{n-1}}^{t_n} \sum_{i,\alpha,j,\beta=1}^{d} \int_0^1 \left(\partial_t A^{\tau}_{i\alpha j\beta}\partial_\beta(\mathbf{W}(t) - \Pi_k\mathbf{W}(t))_j, \partial_\alpha\boldsymbol{\theta}_{\phi,i}(t)\right)_{L^2} \mathrm{d}\tau\,\mathrm{d}t. \tag{3.3.31}$$

Then

$$\begin{aligned}
&\left|\int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k\mathbf{W}(t), \dot{\boldsymbol{\theta}}_{\boldsymbol{\phi}}(t))\,\mathrm{d}t\right| \\
&\le \left|-\int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \partial_t\left(\mathbf{W}(t) - \Pi_k\mathbf{W}(t)\right), \boldsymbol{\theta}_{\boldsymbol{\phi}}(t))\,\mathrm{d}t\right| \\
&\quad + \left|\int_{t_{n-1}}^{t_n} \sum_{i,\alpha,j,\beta=1}^{d} \int_0^1 \left(\partial_t A^{\tau}_{i\alpha j\beta}\partial_\beta(\mathbf{W}(t) - \Pi_k\mathbf{W}(t))_j, \partial_\alpha\boldsymbol{\theta}_{\phi,i}(t)\right)_{L^2} \mathrm{d}\tau\,\mathrm{d}t\right| \\
&\le \int_{t_{n-1}}^{t_n} K_\delta \|\nabla\partial_t(\mathbf{W}(t) - \Pi_k\mathbf{W}(t))\|_{L^2(\Omega)} \|\nabla\boldsymbol{\theta}_{\boldsymbol{\phi}}(t)\|_{L^2(\Omega)}\,\mathrm{d}t \\
&\quad + C_\tau \int_{t_{n-1}}^{t_n} \|\nabla(\mathbf{W}(t) - \Pi_k\mathbf{W}(t))\|_{L^2(\Omega)} \|\nabla\boldsymbol{\theta}_{\boldsymbol{\phi}}(t)\|_{L^2(\Omega)}\,\mathrm{d}t \\
&\le \left(\frac{K_\delta}{2} + \frac{C_\tau}{2}\right) \int_{t_{n-1}}^{t_n} \|\nabla\boldsymbol{\theta}_{\phi}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{K_\delta}{2}\int_{t_{n-1}}^{t_n} \|\nabla(\partial_t(\mathbf{W}(t) - \Pi_k\mathbf{W}(t)))\|^2_{L^2(\Omega)}\,\mathrm{d}t \\
&\quad + \frac{C_\tau}{2}\int_{t_{n-1}}^{t_n} \|\nabla(\mathbf{W}(t) - \Pi_k\mathbf{W}(t))\|^2_{L^2(\Omega)}\,\mathrm{d}t \\
&\le \left(\frac{K_\delta}{2} + \frac{C_\tau}{2}\right) \int_{t_{n-1}}^{t_n} \|\nabla\boldsymbol{\theta}_{\boldsymbol{\phi}}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + C(\mathbf{v})\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} \qquad \text{for } \mathbf{W} \in W^{s,\infty}(0,T;[H_0^1(\Omega)]^d).
\end{aligned} \tag{3.3.32}$$

To bound the terms involving $\partial_t A_{i\alpha j\beta}$, we apply Lemma 3.6 to get

$$\left|\int_{t_{n-1}}^{t_n} \sum_{i,\alpha,j,\beta=1}^{d} \int_0^1 \left(\partial_t A^{\tau}_{i\alpha j\beta}\partial_\beta\boldsymbol{\theta}_{\phi,j}(t), \partial_\alpha\boldsymbol{\theta}_{\phi,i}(t)\right)_{L^2} \mathrm{d}\tau\mathrm{d}t\right| \le C_\tau \int_{t_{n-1}}^{t_n} \|\nabla\boldsymbol{\theta}_{\phi}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t. \tag{3.3.33}$$

Combining (3.3.27)–(3.3.33), we obtain

$$\|\dot{\boldsymbol{\theta}}_{\phi}(t_n^-)\|^2_{L^2(\Omega)} + \gamma^2\|\boldsymbol{\theta}_{\phi}(t_n^-)\|^2_{L^2(\Omega)} + \gamma\int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_{\phi}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$

$$
\begin{aligned}
&+ \mathsf{A}(\boldsymbol{\phi}(t_n^-)\mathrm{e}^{\gamma t_n}; \boldsymbol{\theta}_\phi(t_n^-), \boldsymbol{\theta}_\phi(t_n^-)) - \mathsf{A}(\boldsymbol{\phi}(t_{n-1}^-); \boldsymbol{\theta}_\phi(t_{n-1}^-), \boldsymbol{\theta}_\phi(t_{n-1}^-)) \\
&\leq \tilde{C} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + \|\dot{\boldsymbol{\theta}}_\phi(t_{n-1}^-)\|_{L^2(\Omega)}^2 + \gamma^2 \|\boldsymbol{\theta}_\phi(t_{n-1}^-)\|_{L^2(\Omega)}^2 \\
&+ C(\gamma, \mathbf{v}) \left(k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} \right) + C(\mathbf{v}) \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} + C(\mathbf{v}) \frac{k_n^{2(q_n+1)+1}}{q_n^{2(s-1)}}. \qquad (3.3.34)
\end{aligned}
$$

Summing up over $n = 1, \ldots, j$, we obtain

$$
\begin{aligned}
&\|\dot{\boldsymbol{\theta}}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \gamma^2 \|\boldsymbol{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \gamma \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + \mathsf{A}(\boldsymbol{\phi}(t_j^-)\mathrm{e}^{\gamma t_j}; \boldsymbol{\theta}_\phi(t_j^-), \boldsymbol{\theta}_\phi(t_j^-)) \\
&\leq \tilde{C} \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + C(\mathbf{v}) \sum_{n=1}^{j} \left(k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} \right). \qquad (3.3.35)
\end{aligned}
$$

Using the coercivity of $\mathsf{A}(\boldsymbol{\phi}(t_j^-)\mathrm{e}^{\gamma t_n}; \boldsymbol{\theta}_\phi(t_j^-), \boldsymbol{\theta}_\phi(t_j^-))$, we have

$$
\begin{aligned}
&\|\dot{\boldsymbol{\theta}}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \gamma^2 \|\boldsymbol{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \gamma \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + M_1 \|\nabla \boldsymbol{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 \\
&\leq \tilde{C} \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + C(\mathbf{v}) \sum_{n=1}^{j} \left(k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} \right). \qquad (3.3.36)
\end{aligned}
$$

By the fundamental theorem of calculus, the triangle inequality and the inverse inequality, we have for each $t \in I_n$, with $n = 1, \ldots, N$,

$$
\begin{aligned}
\|\nabla \boldsymbol{\theta}_\phi(t)\|_{L^2(\Omega)}^2 &\leq \left(\|\nabla \boldsymbol{\theta}_\phi(t_n^-)\|_{L^2(\Omega)} + \int_{t_{n-1}}^{t_n} \|\nabla \dot{\boldsymbol{\theta}}_\phi(t)\|_{L^2(\Omega)} \,\mathrm{d}t \right)^2 \\
&\leq 2 \|\nabla \boldsymbol{\theta}_\phi(t_n^-)\|_{L^2(\Omega)}^2 + 2C_0^2 h^{-2} k_n \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t.
\end{aligned}
$$

This implies that

$$
\begin{aligned}
&\|\dot{\boldsymbol{\theta}}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \gamma^2 \|\boldsymbol{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \gamma \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + M_1 \|\nabla \boldsymbol{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 \\
&\leq 2\tilde{C} \sum_{n=1}^{j-1} k_n \|\nabla \boldsymbol{\theta}_\phi(t_n^-)\|_{L^2(\Omega)}^2 + 2\tilde{C} k_j \|\nabla \boldsymbol{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + 2\tilde{C} C_0^2 \sum_{n=1}^{j} h^{-2} k_n^2 \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&+ C(\mathbf{v}) \sum_{n=1}^{j} \left(k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} \right)
\end{aligned}
$$

$$
\begin{aligned}
&\leq 2\tilde{C}\sum_{n=1}^{j-1} k_n\|\nabla\boldsymbol{\theta}_\phi(t_n^-)\|_{L^2(\Omega)}^2 + 2\tilde{C}k_j\|\nabla\boldsymbol{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \hat{C}\sum_{n=1}^{j} k_n\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\theta}}_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \\
&\quad + C(\mathbf{v})\sum_{n=1}^{j}\left(k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right),
\end{aligned}
\tag{3.3.37}
$$

where the last inequality follows from the assumption that $\mu_i k_i \leq h^2$ for each $i = 1, \ldots, N$, with $\hat{C} = 2\tilde{C}C_0^2 \max_{1\leq i\leq j}\frac{1}{\mu_i}$. The term $2\tilde{C}k_j\|\nabla\boldsymbol{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2$ and the sum of integrals on the right-hand side of (3.3.37) can be absorbed into the third and fourth terms of the left-hand side of (3.3.37) if we choose each time step k_n to be sufficiently small. That is,

$$
\begin{aligned}
&\|\dot{\boldsymbol{\theta}}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \gamma^2\|\boldsymbol{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \sum_{n=1}^{j}(\gamma - \hat{C}k_n)\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\theta}}_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \\
&\quad + (M_1 - 2\tilde{C}k_n)\|\nabla\boldsymbol{\theta}_\phi(t_n^-)\|_{L^2(\Omega)}^2 \\
&\leq 2\tilde{C}\sum_{n=1}^{j-1} k_n\|\nabla\boldsymbol{\theta}_\phi(t_n^-)\|_{L^2(\Omega)}^2 + C(\mathbf{v})\sum_{n=1}^{j}\left(k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right).
\end{aligned}
\tag{3.3.38}
$$

By choosing $k_n \leq \min\{\frac{\gamma}{2\hat{C}}, \frac{M_1}{4\tilde{C}}\}$ for each $n = 1, \ldots, N$ and applying the discrete Grönwall lemma, we have

$$
\begin{aligned}
&\|\dot{\boldsymbol{\theta}}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \|\boldsymbol{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\theta}}_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t + \|\nabla\boldsymbol{\theta}_\phi(t_j^-))\|_{L^2(\Omega)}^2 \\
&\leq C_j(\mathbf{v})\exp\left(C\sum_{n=1}^{j}k_n\right)\sum_{n=1}^{j}\left(k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right) \leq\ C_{\max}(\mathbf{v})\sum_{n=1}^{j}\left(k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right),
\end{aligned}
\tag{3.3.39}
$$

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

$$
\|\nabla\mathbf{v}_\phi(t) - \nabla\mathbf{v}(t)\|_{L^\infty(\Omega)} \leq \|\nabla\mathbf{v}_\phi(t) - \nabla\mathbf{W}(t)\|_{L^\infty(\Omega)} + \|\nabla\mathbf{v}(t) - \nabla\mathbf{W}(t)\|_{L^\infty(\Omega)}.
$$

By the inverse estimate (ii,b), the error bound (iii,a), and the approximation properties of $\mathcal{P}_h$ in the $[W^{1,\infty}(\Omega)]^d$ and $[H^1(\Omega)]^d$ semi-norms, we can find an $h_1 > 0$ such that, for

$h < h_1$,

$$
\begin{aligned}
\|\nabla \mathbf{W}(t) - \nabla \mathbf{v}(t)\|_{L^\infty(\Omega)} &\le \|\nabla \mathbf{W}(t) - \nabla \mathcal{P}_h \mathbf{v}(t)\|_{L^\infty(\Omega)} + \|\nabla \mathcal{P}_h \mathbf{v}(t) - \mathbf{v}(t)\|_{L^\infty(\Omega)} \\
&\le C_1 h^{-\frac{d}{2}} \|\nabla(\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t))\|_{L^2(\Omega)} + C(\mathbf{v}) h^{r-\frac{d}{2}} \quad \text{(by (ii,b))} \\
&\le C_1 h^{-\frac{d}{2}} \|\nabla(\mathbf{W}(t) - \mathbf{v}(t))\|_{L^2(\Omega)} + C_1 h^{-\frac{d}{2}} \|\nabla(\mathcal{P}_h \mathbf{v}(t) - \mathbf{v}(t))\|_{L^2(\Omega)} \\
&\quad + C(\mathbf{v}) h^{r-\frac{d}{2}} \\
&\le \tilde{C}(\mathbf{v}) h^{r-\frac{d}{2}} \le \frac{\delta}{2} \mathrm{e}^{-\gamma T}.
\end{aligned}
$$

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

$$
\begin{aligned}
&\|\nabla \mathbf{v}_\phi(t) - \nabla \mathbf{W}(t)\|_{L^\infty(\Omega)} \\
&\le C_1 h^{-\frac{d}{2}} \Big(\|\nabla \boldsymbol{\theta}_\phi(t)\|_{L^2(\Omega)} + \|\nabla \Pi_k \mathbf{W}(t) - \nabla \mathbf{W}(t)\|_{L^2(\Omega)} \Big) \\
&\le C_1 h^{-\frac{d}{2}} \Big(\|\nabla \boldsymbol{\theta}_\phi(t_n^-)\|_{L^2(\Omega)} + \int_{t_{n-1}}^{t_n} \|\partial_t \nabla \boldsymbol{\theta}_\phi(t)\|_{L^2(\Omega)} \,\mathrm{d}t \\
&\quad + \|\nabla \Pi_k \mathbf{W}(t_n^-) - \nabla \mathbf{W}(t_n^-)\|_{L^2(\Omega)} + \int_{t_{n-1}}^{t_n} \|\partial_t (\nabla \Pi_k \mathbf{W} - \nabla \mathbf{W})(t)\|_{L^2(\Omega)} \,\mathrm{d}t \Big) \\
&\le C_1 h^{-\frac{d}{2}} \Big(\|\nabla \boldsymbol{\theta}_\phi(t_n^-)\|_{L^2(\Omega)} + \int_{t_{n-1}}^{t_n} \|\partial_t \nabla \boldsymbol{\theta}_\phi(t)\|_{L^2(\Omega)} \,\mathrm{d}t \Big) \\
&\quad + C_1 h^{-\frac{d}{2}} \sqrt{k_n} \left(\int_{t_{n-1}}^{t_n} \|\partial_t (\nabla \Pi_k \mathbf{W} - \nabla \mathbf{W})(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \right)^{\frac{1}{2}} \\
&\le C_1 h^{-\frac{d}{2}} \left(\|\nabla \boldsymbol{\theta}_\phi(t_n^-)\|_{L^2(\Omega)} + C_0 h^{-1} \sqrt{k_n} \int_{t_{n-1}}^{t_n} \|\partial_t \boldsymbol{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \right)^{\frac{1}{2}} + C(\mathbf{W}) h^{-\frac{d}{2}} \frac{k_n^{q_n+1}}{q_n^{s-1}} \\
&\le C(\mathbf{v}) h^{-\frac{d}{2}} \left(\sum_{i=1}^{n} k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} < \frac{\delta}{2} \mathrm{e}^{-\gamma T}.
\end{aligned}
$$

By choosing $h < h_* = \min\{h_0, h_1, h_2\}$, we obtain $\nabla \mathbf{v}_\phi \mathrm{e}^{\gamma t} \in \mathcal{Z}_\delta$.

Verification that $\mathcal{N}$ is a contraction mapping

To show the contraction property, we consider $\mathbf{R} = \boldsymbol{\phi} - \boldsymbol{\phi}'$ and $\boldsymbol{\Theta} = \mathbf{v}_\phi - \mathbf{v}_{\phi'}$ where $\boldsymbol{\phi}, \boldsymbol{\phi}' \in \mathcal{F}$. Replacing $\boldsymbol{\phi}$ in (3.3.22) by $\boldsymbol{\phi}'$ and subtracting the new equation from (3.3.22), we have

$$
\begin{aligned}
&\int_{t_{n-1}}^{t_n} \left\langle \ddot{\mathbf{\Theta}}(t), \dot{\boldsymbol{\varphi}}(t) \right\rangle \, \mathrm{d}t + \left([\dot{\mathbf{\Theta}}(t)]_{n-1}, \dot{\boldsymbol{\varphi}}(t_{n-1}^{+}) \right)_{L^2} + 2\gamma \int_{t_{n-1}}^{t_n} \left(\dot{\mathbf{\Theta}}(t), \dot{\boldsymbol{\varphi}}(t) \right)_{L^2} \, \mathrm{d}t \\
&\quad + \gamma^2 \int_{t_{n-1}}^{t_n} (\mathbf{\Theta}(t), \dot{\boldsymbol{\varphi}}(t))_{L^2} \, \mathrm{d}t + \gamma^2 \left([\mathbf{\Theta}(t)]_{n-1}, \boldsymbol{\varphi}(t_{n-1}^{+}) \right)_{L^2} \\
&\quad + \int_{t_{n-1}}^{t_n} \left(\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi}(t), \dot{\boldsymbol{\varphi}}(t)) - \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\boldsymbol{\varphi}}(t)) \right) \mathrm{d}t \\
&\quad + \frac{1}{2} \left\{ \mathsf{A}(\boldsymbol{\phi}(t_{n-1}^{+})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi}(t_{n-1}^{+}), \boldsymbol{\varphi}(t_{n-1}^{+})) - \mathsf{A}(\boldsymbol{\phi'}(t_{n-1}^{+})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^{+}), \boldsymbol{\varphi}(t_{n-1}^{+})) \right\} \\
&= \int_{t_{n-1}}^{t_n} \left(\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\boldsymbol{\varphi}}(t)) - \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\boldsymbol{\varphi}}(t)) \right) \mathrm{d}t \\
&\quad + \frac{1}{2} \left\{ \mathsf{A}(\boldsymbol{\phi}(t_{n-1}^{-})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi}(t_{n-1}^{-}), \boldsymbol{\varphi}(t_{n-1}^{-})) - \mathsf{A}(\boldsymbol{\phi'}(t_{n-1}^{-})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^{-}), \boldsymbol{\varphi}(t_{n-1}^{-})) \right\}.
\end{aligned}
$$

Taking $\boldsymbol{\varphi} = \mathbf{\Theta}$ and replacing

$$
\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi}(t), \dot{\mathbf{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\mathbf{\Theta}}(t))
$$

by

$$
\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{\Theta}(t), \dot{\mathbf{\Theta}}(t)) + \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\mathbf{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\mathbf{\Theta}}(t))
$$

in the resulting equation, we have

$$
\begin{aligned}
&\frac{1}{2}\|\dot{\mathbf{\Theta}}(t_n^{-})\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\dot{\mathbf{\Theta}}(t_{n-1}^{+})\|_{L^2(\Omega)}^2 + \frac{\gamma^2}{2}\|\mathbf{\Theta}(t_n^{-})\|_{L^2(\Omega)}^2 + \frac{\gamma^2}{2}\|\mathbf{\Theta}(t_{n-1}^{+})\|_{L^2(\Omega)}^2 \\
&\quad + 2\gamma \int_{t_{n-1}}^{t_n} \|\dot{\mathbf{\Theta}}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t + \int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{\Theta}(t), \dot{\mathbf{\Theta}}(t)) + \frac{1}{2}\mathsf{A}(\boldsymbol{\phi}_{n-1}^{+}\mathrm{e}^{\gamma t_{n-1}}; \mathbf{\Theta}_{n-1}^{+}, \mathbf{\Theta}_{n-1}^{+}) \\
&= \int_{t_{n-1}}^{t_n} \left(\mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\mathbf{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\mathbf{\Theta}}(t)) \right) \mathrm{d}t \\
&\quad + \frac{1}{2}\mathsf{A}(\boldsymbol{\phi}_{n-1}^{-}\mathrm{e}^{\gamma t_{n-1}}; \mathbf{\Theta}_{n-1}^{-}, \mathbf{\Theta}_{n-1}^{-}) + \left(\dot{\mathbf{\Theta}}(t_{n-1}^{-}), \dot{\mathbf{\Theta}}(t_{n-1}^{+}) \right)_{L^2} + \gamma^2 \left(\mathbf{\Theta}(t_{n-1}^{-}), \mathbf{\Theta}(t_{n-1}^{+}) \right)_{L^2} \\
&\quad + \int_{t_{n-1}}^{t_n} \left(\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\mathbf{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\mathbf{\Theta}}(t)) \right) \mathrm{d}t \\
&\quad + \frac{1}{2} \left\{ \mathsf{A}(\boldsymbol{\phi'}(t_{n-1}^{+})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^{+}), \mathbf{\Theta}(t_{n-1}^{+})) - \mathsf{A}(\boldsymbol{\phi}(t_{n-1}^{+})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^{+}), \mathbf{\Theta}(t_{n-1}^{+})) \right\} \\
&\quad + \frac{1}{2} \left\{ \mathsf{A}(\boldsymbol{\phi}(t_{n-1}^{-})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^{-}), \mathbf{\Theta}(t_{n-1}^{-})) - \mathsf{A}(\boldsymbol{\phi'}(t_{n-1}^{-})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^{-}), \mathbf{\Theta}(t_{n-1}^{-})) \right\}.
\end{aligned}
$$

By writing

$$\int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\Theta}(t), \dot{\boldsymbol{\Theta}}(t))\,\mathrm{d}t = \frac{1}{2}\int_{t_{n-1}}^{t_n} \frac{\mathrm{d}}{\mathrm{d}t}\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\Theta}(t), \boldsymbol{\Theta}(t))\,\mathrm{d}t$$
$$- \frac{1}{2}\sum_{i,\alpha,j,\beta=1}^{d}\int_{t_{n-1}}^{t_n}\int_0^1 \Big(\partial_t A^{\tau}_{i\alpha j\beta}\partial_\beta \boldsymbol{\Theta}_j(t), \partial_\alpha \boldsymbol{\Theta}_i(t)\Big)_{L^2}\,\mathrm{d}\tau\,\mathrm{d}t,$$

where $A^{\tau}_{i\alpha j\beta} = A_{i\alpha j\beta}(\nabla \mathbf{W}(t)\mathrm{e}^{\gamma t} + \tau(\nabla\boldsymbol{\phi}(t) - \nabla\mathbf{W}(t))\mathrm{e}^{\gamma t})$, we have

$$\begin{aligned}
&\|\dot{\boldsymbol{\Theta}}(t_n^-)\|^2_{L^2(\Omega)} + \|\dot{\boldsymbol{\Theta}}(t_{n-1}^+)\|^2_{L^2(\Omega)} + \gamma^2\|\boldsymbol{\Theta}(t_n^-)\|^2_{L^2(\Omega)} + \gamma^2\|\boldsymbol{\Theta}(t_{n-1}^+)\|^2_{L^2(\Omega)}\\
&\quad + 4\gamma\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\Theta}}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \mathsf{A}(\boldsymbol{\phi}_n^-\mathrm{e}^{\gamma t_n}; \boldsymbol{\Theta}_n^-, \boldsymbol{\Theta}_n^-)\\
&= 2\Big(\dot{\boldsymbol{\Theta}}(t_{n-1}^-), \dot{\boldsymbol{\Theta}}(t_{n-1}^+)\Big)_{L^2} + 2\gamma^2\left(\boldsymbol{\Theta}(t_{n-1}^-), \boldsymbol{\Theta}(t_{n-1}^+)\right)_{L^2} + \mathsf{A}(\boldsymbol{\phi}_{n-1}^-\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\Theta}_{n-1}^-, \boldsymbol{\Theta}_{n-1}^-)\\
&\quad + \sum_{i,\alpha,j,\beta=1}^{d}\int_{t_{n-1}}^{t_n}\int_0^1\Big(\partial_t A^{\tau}_{i\alpha j\beta}\partial_\beta\boldsymbol{\Theta}_j(t), \partial_\alpha\boldsymbol{\Theta}_i(t)\Big)_{L^2}\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad + 2\int_{t_{n-1}}^{t_n}\Big(\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k\mathbf{W}(t), \dot{\boldsymbol{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k\mathbf{W}(t), \dot{\boldsymbol{\Theta}}(t))\Big)\,\mathrm{d}t\\
&\quad + 2\int_{t_{n-1}}^{t_n}\Big(\mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\boldsymbol{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\boldsymbol{\Theta}}(t))\Big)\,\mathrm{d}t\\
&\quad + \Big(\mathsf{A}(\boldsymbol{\phi'}(t_{n-1}^+)\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^+), \boldsymbol{\Theta}(t_{n-1}^+)) - \mathsf{A}(\boldsymbol{\phi}(t_{n-1}^+)\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^+), \boldsymbol{\Theta}(t_{n-1}^+))\Big)\\
&\quad + \Big(\mathsf{A}(\boldsymbol{\phi}(t_{n-1}^-)\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^-), \boldsymbol{\Theta}(t_{n-1}^-)) - \mathsf{A}(\boldsymbol{\phi'}(t_{n-1}^-)\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^-), \boldsymbol{\Theta}(t_{n-1}^-))\Big).
\end{aligned}$$

Note that

$$\begin{aligned}
&\int_{t_{n-1}}^{t_n}\mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\boldsymbol{\Theta}}(t))\,\mathrm{d}t + \mathsf{A}(\boldsymbol{\phi'}(t_{n-1}^+)\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^+), \boldsymbol{\Theta}(t_{n-1}^+))\\
&\quad - \mathsf{A}(\boldsymbol{\phi'}(t_{n-1}^-)\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^-), \boldsymbol{\Theta}(t_{n-1}^-))\\
&= -\int_{t_{n-1}}^{t_n}\mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \dot{\boldsymbol{\theta}}_{\phi'}(t), \boldsymbol{\Theta}(t))\,\mathrm{d}t\\
&\quad - \sum_{i,\alpha,j,\beta=1}^{d}\int_{t_{n-1}}^{t_n}\int_0^1\Big(\partial_t\tilde{A}^{\tau}_{i\alpha j\beta}\partial_\beta\boldsymbol{\theta}_{\phi',j}(t), \partial_\alpha\boldsymbol{\Theta}_i(t)\Big)_{L^2}\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad + \mathsf{A}(\boldsymbol{\phi'}(t_n^-)\mathrm{e}^{\gamma t_n}; \boldsymbol{\theta}_{\phi'}(t_n^-), \boldsymbol{\Theta}(t_n^-)) - \mathsf{A}(\boldsymbol{\phi'}(t_{n-1}^-)\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t_{n-1}^-), \boldsymbol{\Theta}(t_{n-1}^-)),
\end{aligned}$$

where $\tilde{A}^{\tau}_{i\alpha j\beta} := A_{i\alpha j\beta}(\nabla \mathbf{W}(t)\mathrm{e}^{\gamma t} + \tau(\nabla \boldsymbol{\phi'}(t) - \nabla \mathbf{W}(t))\mathrm{e}^{\gamma t})$. Analogously, we have

$$
\begin{aligned}
&-\int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\boldsymbol{\Theta}}(t))\,\mathrm{d}t - \mathsf{A}(\boldsymbol{\phi}(t^+_{n-1})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t^+_{n-1}), \boldsymbol{\Theta}(t^+_{n-1}))\\
&\quad + \mathsf{A}(\boldsymbol{\phi}(t^-_{n-1})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t^-_{n-1}), \boldsymbol{\Theta}(t^-_{n-1}))\\
&= \int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \dot{\boldsymbol{\theta}}_{\boldsymbol{\phi'}}(t), \boldsymbol{\Theta}(t))\,\mathrm{d}t\\
&\quad + \sum_{i,\alpha,j,\beta=1}^{d} \int_{t_{n-1}}^{t_n}\int_0^1 \Big(\partial_t A^{\tau}_{i\alpha j\beta}\partial_\beta \boldsymbol{\theta}_{\phi',j}(t), \partial_\alpha \boldsymbol{\Theta}_i(t)\Big)_{L^2}\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad - \mathsf{A}(\boldsymbol{\phi}(t^-_n)\mathrm{e}^{\gamma t_n}; \boldsymbol{\theta}_{\phi'}(t^-_n), \boldsymbol{\Theta}(t^-_n)) + \mathsf{A}(\boldsymbol{\phi}(t^-_{n-1})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t^-_{n-1}), \boldsymbol{\Theta}(t^-_{n-1})).
\end{aligned}
$$

This implies that

$$
\begin{aligned}
&\|\dot{\boldsymbol{\Theta}}(t^-_n)\|^2_{L^2(\Omega)} + \|\dot{\boldsymbol{\Theta}}(t^+_{n-1})\|^2_{L^2(\Omega)} + \gamma^2\|\boldsymbol{\Theta}(t^-_n)\|^2_{L^2(\Omega)} + \gamma^2\|\boldsymbol{\Theta}(t^+_{n-1})\|^2_{L^2(\Omega)}\\
&\quad + 4\gamma\int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\Theta}}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \mathsf{A}(\boldsymbol{\phi}^-_n\mathrm{e}^{\gamma t_n}; \boldsymbol{\Theta}^-_n, \boldsymbol{\Theta}^-_n)\\
&= 2\Big(\dot{\boldsymbol{\Theta}}(t^-_{n-1}), \dot{\boldsymbol{\Theta}}(t^+_{n-1})\Big)_{L^2} + 2\gamma^2\left(\boldsymbol{\Theta}(t^-_{n-1}), \boldsymbol{\Theta}(t^+_{n-1})\right)_{L^2} + \mathsf{A}(\boldsymbol{\phi}^-_{n-1}\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\Theta}^-_{n-1}, \boldsymbol{\Theta}^-_{n-1})\\
&\quad + \sum_{i,\alpha,j,\beta=1}^{d} \int_{t_{n-1}}^{t_n}\int_0^1 \Big(\partial_t A^{\tau}_{i\alpha j\beta}\partial_\beta \boldsymbol{\Theta}_j(t), \partial_\alpha \boldsymbol{\Theta}_i(t)\Big)_{L^2}\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad + \int_{t_{n-1}}^{t_n} \Big(\mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\boldsymbol{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\boldsymbol{\Theta}}(t))\Big)\,\mathrm{d}t\\
&\quad + \int_{t_{n-1}}^{t_n} \Big(\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \dot{\boldsymbol{\theta}}_{\phi'}(t), \boldsymbol{\Theta}(t)) - \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \dot{\boldsymbol{\theta}}_{\phi'}(t), \boldsymbol{\Theta}(t))\Big)\,\mathrm{d}t\\
&\quad + \int_{t_{n-1}}^{t_n}\int_0^1 \sum_{i,\alpha,j,\beta=1}^{d} \Big((\partial_t A^{\tau}_{i,\alpha,j,\beta} - \partial_t \tilde{A}^{\tau}_{i\alpha j\beta})\partial_\beta \boldsymbol{\theta}_{\phi',j}(t), \partial_\alpha \boldsymbol{\Theta}_i(t)\Big)_{L^2}\,\mathrm{d}\tau\,\mathrm{d}t\\
&\quad + 2\int_{t_{n-1}}^{t_n} \Big(\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k\mathbf{W}(t), \dot{\boldsymbol{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k\mathbf{W}(t), \dot{\boldsymbol{\Theta}}(t))\Big)\,\mathrm{d}t\\
&\quad + \mathsf{A}(\boldsymbol{\phi'}(t^-_n)\mathrm{e}^{\gamma t_n}; \boldsymbol{\theta}_{\phi'}(t^-_n), \boldsymbol{\Theta}(t^-_n)) - \mathsf{A}(\boldsymbol{\phi'}(t^-_{n-1})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t^-_{n-1}), \boldsymbol{\Theta}(t^-_{n-1}))\\
&\quad - \mathsf{A}(\boldsymbol{\phi}(t^-_n)\mathrm{e}^{\gamma t_n}; \boldsymbol{\theta}_{\phi'}(t^-_n), \boldsymbol{\Theta}(t^-_n)) + \mathsf{A}(\boldsymbol{\phi}(t^-_{n-1})\mathrm{e}^{\gamma t_{n-1}}; \boldsymbol{\theta}_{\phi'}(t^-_{n-1}), \boldsymbol{\Theta}(t^-_{n-1})). \qquad (3.3.40)
\end{aligned}
$$

By the Cauchy–Schwarz inequality, we have

$$
2\Big(\dot{\boldsymbol{\Theta}}(t^-_{n-1}), \dot{\boldsymbol{\Theta}}(t^+_{n-1})\Big)_{L^2} \le \|\dot{\boldsymbol{\Theta}}(t^-_{n-1})\|^2_{L^2(\Omega)} + \|\dot{\boldsymbol{\Theta}}(t^+_{n-1})\|^2_{L^2(\Omega)}, \qquad (3.3.41)
$$

$$2\gamma^2 \left(\mathbf{\Theta}(t_{n-1}^-), \mathbf{\Theta}(t_{n-1}^+)\right)_{L^2} \leq \gamma^2 \|\mathbf{\Theta}(t_{n-1}^-)\|_{L^2(\Omega)}^2 + \gamma^2 \|\mathbf{\Theta}(t_{n-1}^+)\|_{L^2(\Omega)}^2. \tag{3.3.42}$$

By Lemma 3.6, for $\boldsymbol{\phi} \in \mathcal{F}$, we have

$$\left| \int_{t_{n-1}}^{t_n} \int_0^1 \sum_{i,\alpha,j,\beta=1}^{d} \left(\partial_t A_{i\alpha j\beta}^{\tau} \partial_\beta \mathbf{\Theta}_j(t), \partial_\alpha \mathbf{\Theta}_i(t) \right)_{L^2} \mathrm{d}\tau \, \mathrm{d}t \right| \leq C_\tau \int_{t_{n-1}}^{t_n} \|\nabla \mathbf{\Theta}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t. \tag{3.3.43}$$

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

$$\begin{aligned}
&\left| \int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\mathbf{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\mathbf{\Theta}}(t)) \, \mathrm{d}t \right| \\
&\leq \int_{t_{n-1}}^{t_n} \left| \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\mathbf{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\mathbf{\Theta}}(t)) \right| \mathrm{d}t \\
&\leq L_\delta \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\phi}(t) - \nabla \boldsymbol{\phi'}(t)\|_{L^2(\Omega)} \|\nabla \boldsymbol{\theta}_{\phi'}(t)\|_{L^\infty(\Omega)} \|\nabla \dot{\mathbf{\Theta}}(t)\|_{L^2(\Omega)} \, \mathrm{d}t \\
&\leq L_\delta C_0 C_1 \int_{t_{n-1}}^{t_n} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)} h^{-1-\frac{d}{2}} \|\nabla \boldsymbol{\theta}_{\boldsymbol{\phi'}}(t)\|_{L^2(\Omega)} \|\dot{\mathbf{\Theta}}(t)\|_{L^2(\Omega)} \, \mathrm{d}t \\
&\leq 2\gamma \int_{t_{n-1}}^{t_n} \|\dot{\mathbf{\Theta}}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t + C(\gamma) \max_{t \in I_n} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)}^2 h^{-2-d} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\theta}_{\boldsymbol{\phi'}}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t,
\end{aligned} \tag{3.3.44}$$

where we have used the inverse inequalities (ii,a) and (ii,b), Young's inequality. By the fundamental theorem of calculus, we have

$$\begin{aligned}
\int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\theta}_{\phi'}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t &\leq 2k_n \|\nabla \boldsymbol{\theta}_{\phi'}(t_n^-)\|_{L^2(\Omega)}^2 + 2k_n \left(\int_{t_{n-1}}^{t_n} \|\nabla \dot{\boldsymbol{\theta}}_{\phi'}(t)\|_{L^2(\Omega)} \, \mathrm{d}t \right)^2 \\
&\leq 2k_n \|\nabla \boldsymbol{\theta}_{\phi'}(t_n^-)\|_{L^2(\Omega)}^2 + 2C_0^2 k_n^2 h^{-2} \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_{\phi'}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t.
\end{aligned}$$

This implies that

$$\left| \int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\mathbf{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \boldsymbol{\theta}_{\phi'}(t), \dot{\mathbf{\Theta}}(t)) \, \mathrm{d}t \right| \leq 2\gamma \int_{t_{n-1}}^{t_n} \|\dot{\mathbf{\Theta}}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t$$

$$+ C \max_{t\in I_n} \|\nabla \mathbf{R}(t)\|^2_{L^2(\Omega)} h^{-2-d} k_n \left(\|\nabla \boldsymbol{\theta}_{\phi'}(t_n^-)\|^2_{L^2(\Omega)} + \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_{\phi'}(t)\|^2_{L^2(\Omega)} \,\mathrm{d}t \right), \tag{3.3.45}$$

where we have used the assumption that $\mu_i k_i \leq h^2$ for each $i = 1, \ldots, N$. Here C is a generic positive constant. Analogously, we obtain

$$\begin{aligned}
&\left| \int_{t_{n-1}}^{t_n} \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \dot{\boldsymbol{\theta}}_{\phi'}(t), \boldsymbol{\Theta}(t)) - \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \dot{\boldsymbol{\theta}}_{\phi'}(t), \boldsymbol{\Theta}(t)) \,\mathrm{d}t \right| \\
&\leq \int_{t_{n-1}}^{t_n} \left| \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \dot{\boldsymbol{\theta}}_{\phi'}(t), \boldsymbol{\Theta}(t)) - \mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \dot{\boldsymbol{\theta}}_{\phi'}(t), \boldsymbol{\Theta}(t)) \right| \mathrm{d}t \\
&\leq L_\delta \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\phi}(t) - \nabla \boldsymbol{\phi'}(t)\|_{L^2(\Omega)} \|\nabla \dot{\boldsymbol{\theta}}_{\phi'}(t)\|_{L^\infty(\Omega)} \|\nabla \boldsymbol{\Theta}(t)\|_{L^2(\Omega)} \,\mathrm{d}t \\
&\leq L_\delta C_1 \int_{t_{n-1}}^{t_n} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)} h^{-\frac{d}{2}} \|\nabla \dot{\boldsymbol{\theta}}_{\phi'}(t)\|_{L^2(\Omega)} \|\nabla \boldsymbol{\Theta}(t)\|_{L^2(\Omega)} \,\mathrm{d}t \\
&\leq \tilde{C} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\Theta}(t)\|^2_{L^2(\Omega)} \,\mathrm{d}t + C \max_{t\in I_n} \|\nabla \mathbf{R}(t)\|^2_{L^2(\Omega)} h^{-d-2} \left(\int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_{\phi'}(t)\|^2_{L^2(\Omega)} \,\mathrm{d}t \right).
\end{aligned} \tag{3.3.46}$$

Applying the Lipschitz continuity of $A_{i\alpha j\beta}$ again, we have

$$\begin{aligned}
&2\left| \int_{t_{n-1}}^{t_n} \big(\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\boldsymbol{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\boldsymbol{\Theta}}(t)) \big) \,\mathrm{d}t \right| \\
&\leq 2 \int_{t_{n-1}}^{t_n} \left| \big(\mathsf{A}(\boldsymbol{\phi}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\boldsymbol{\Theta}}(t)) - \mathsf{A}(\boldsymbol{\phi'}(t)\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \Pi_k \mathbf{W}(t), \dot{\boldsymbol{\Theta}}(t)) \big) \right| \mathrm{d}t \\
&\leq 2 L_\delta \int_{t_{n-1}}^{t_n} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)} \|\nabla \mathbf{W}(t) - \nabla(\Pi_k \mathbf{W})(t)\|_{L^\infty(\Omega)} \|\nabla \dot{\boldsymbol{\Theta}}(t)\|_{L^2(\Omega)} \,\mathrm{d}t \\
&\leq 2 L_\delta C_0 C_1 h^{-\frac{d}{2}-1} \int_{t_{n-1}}^{t_n} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)} \|\nabla \mathbf{W}(t) - \nabla(\Pi_k \mathbf{W})(t)\|_{L^2(\Omega)} \|\dot{\boldsymbol{\Theta}}(t)\|_{L^2(\Omega)} \,\mathrm{d}t \\
&\leq C(\gamma) h^{-d-2} \max_{t\in I_n} \|\nabla \mathbf{R}(t)\|^2_{L^2(\Omega)} \int_{t_{n-1}}^{t_n} \|\nabla \mathbf{W}(t) - \nabla(\Pi_k \mathbf{W})(t)\|^2_{L^2(\Omega)} \,\mathrm{d}t + \gamma \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\Theta}}(t)\|^2_{L^2(\Omega)} \,\mathrm{d}t \\
&\leq \gamma \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\Theta}}(t)\|^2_{L^2(\Omega)} \,\mathrm{d}t + C(\gamma, \mathbf{W}) h^{-d-2} \frac{k_n^{2(q_n+1)+1}}{q_n^{2(s-1)}} \max_{t\in I_n} \|\nabla \mathbf{R}(t)\|^2_{L^2(\Omega)},
\end{aligned} \tag{3.3.47}$$

for $\mathbf{W} \in W^{s,\infty}(0,T;[H_0^1(\Omega)]^d)$, where $C(\gamma,\mathbf{W})$ is a positive constant depending on both γ and the nonlinear projection $\mathbf{W}$. Recall that

$$\partial_t A^{\tau}_{i\alpha j\beta} := \partial_t A_{i\alpha j\beta}(\nabla\mathbf{W}(t)\mathrm{e}^{\gamma t} + \tau(\nabla\boldsymbol{\phi}(t) - \nabla\mathbf{W}(t))\mathrm{e}^{\gamma t}),$$

$$\partial_t \tilde{A}^{\tau}_{i\alpha j\beta} := \partial_t A_{i\alpha j\beta}(\nabla\mathbf{W}(t)\mathrm{e}^{\gamma t} + \tau(\nabla\boldsymbol{\phi}'(t) - \nabla\mathbf{W}(t))\mathrm{e}^{\gamma t}).$$

By Taylor's theorem, we have

$$\begin{aligned}
&\partial_t A^{\tau}_{i\alpha j\beta} - \partial_t \tilde{A}^{\tau}_{i\alpha j\beta} \\
&= \int_0^1 \sum_{k,\gamma,l,\delta=1}^{d} \frac{\partial^2}{\partial\eta_{k\gamma}\partial\eta_{l\delta}} A_{i\alpha j\beta}(\nabla\mathbf{W}(t)\mathrm{e}^{\gamma t} + \tau(\nabla\boldsymbol{\phi}'(t) - \nabla\mathbf{W}(t))\mathrm{e}^{\gamma t} + \tilde{\tau}\tau(\nabla\boldsymbol{\phi} - \nabla\boldsymbol{\phi}')\mathrm{e}^{\gamma t}) \\
&\quad \times \partial_t\partial_\delta(\mathbf{W}_l(t)\mathrm{e}^{\gamma t} + \tau(\boldsymbol{\phi}'_l(t) - \mathbf{W}_l(t))\mathrm{e}^{\gamma t} + \tilde{\tau}\tau(\boldsymbol{\phi}_l - \boldsymbol{\phi}'_l)\mathrm{e}^{\gamma t})\tau\partial_\gamma(\boldsymbol{\phi}_k(t) - \boldsymbol{\phi}'_k(t))\,\mathrm{d}\tilde{\tau}.
\end{aligned}$$

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

$$\begin{aligned}
&|\partial_t A^{\tau}_{i\alpha j\beta} - \partial_t \tilde{A}^{\tau}_{i\alpha j\beta}| \le \hat{C}\|\nabla\mathbf{R}(t)\|_{L^\infty(\Omega)}\big(\|\nabla\mathbf{W}(t)\|_{L^\infty(\Omega)} + \|\nabla\dot{\mathbf{W}}(t)\|_{L^\infty(\Omega)}\big) \\
&+ \hat{C}\|\nabla\mathbf{R}(t)\|_{L^\infty(\Omega)}\big(\|\nabla\boldsymbol{\phi}(t) - \nabla\boldsymbol{\phi}'(t)\|_{L^\infty(\Omega)} + \|\nabla\dot{\boldsymbol{\phi}}(t) - \nabla\dot{\boldsymbol{\phi}}'(t)\|_{L^\infty(\Omega)}\big) \\
&+ \hat{C}\|\nabla\mathbf{R}(t)\|_{L^\infty(\Omega)}\big(\|\nabla\boldsymbol{\phi}'(t) - \nabla\mathbf{W}(t)\|_{L^\infty(\Omega)} + \|\nabla\dot{\boldsymbol{\phi}}'(t) - \nabla\dot{\mathbf{W}}(t)\|_{L^\infty(\Omega)}\big).
\end{aligned}$$

Similarly to the proof of Lemma 3.6, we can show that

$$\|\nabla\boldsymbol{\phi}'(t) - \nabla\mathbf{W}(t)\|_{L^\infty(\Omega)} + \|\nabla\dot{\boldsymbol{\phi}}'(t) - \nabla\dot{\mathbf{W}}(t)\|_{L^\infty(\Omega)} \le C_\tau. \tag{3.3.48}$$

Property (iii, c) of the nonlinear projection $\mathbf{W}$ implies that

$$\|\nabla\mathbf{W}(t)\|_{L^\infty(\Omega)} + \|\nabla\dot{\mathbf{W}}(t)\|_{L^\infty(\Omega)} \le c_0 + c_1. \tag{3.3.49}$$

In the view of triangle inequality, we have

$$\begin{aligned}
&\|\nabla\boldsymbol{\phi}(t) - \nabla\boldsymbol{\phi}'(t)\|_{L^\infty(\Omega)} + \|\nabla\dot{\boldsymbol{\phi}}(t) - \nabla\dot{\boldsymbol{\phi}}'(t)\|_{L^\infty(\Omega)} \\
&\le \|\nabla\boldsymbol{\phi}(t) - \nabla\mathbf{W}(t)\|_{L^\infty(\Omega)} + \|\nabla\boldsymbol{\phi}'(t) - \nabla\mathbf{W}(t)\|_{L^\infty(\Omega)} + \|\nabla\dot{\boldsymbol{\phi}}(t) - \nabla\dot{\mathbf{W}}(t)\|_{L^\infty(\Omega)} \\
&\quad + \|\nabla\dot{\boldsymbol{\phi}}'(t) - \nabla\dot{\mathbf{W}}(t)\|_{L^\infty(\Omega)} \le 2C_\tau.
\end{aligned} \tag{3.3.50}$$

Combining (3.3.48)–(3.3.50) and applying the inverse inequality (ii,b) and Young's inequality, we obtain

$$
\begin{aligned}
&\left|\int_{t_{n-1}}^{t_n}\int_0^1 \sum_{i,\alpha,j,\beta=1}^{d}\Big((\partial_t A^\tau_{i\alpha j\beta}-\partial_t \tilde{A}^\tau_{i\alpha j\beta})\partial_\beta \boldsymbol{\theta}_{\phi',j}(t),\partial_\alpha \boldsymbol{\Theta}_i(t)\Big)_{L^2}\,\mathrm{d}\tau\,\mathrm{d}t\right|\\
&\le \tilde{C}_{\mathrm{Lip}}\int_{t_{n-1}}^{t_n}\|\nabla\mathbf{R}(t)\|_{L^\infty(\Omega)}\|\nabla\boldsymbol{\theta}_{\phi'}(t)\|_{L^2(\Omega)}\|\nabla\boldsymbol{\Theta}(t)\|_{L^2(\Omega)}\,\mathrm{d}t\\
&\le \tilde{C}_{\mathrm{Lip}}C_1h^{-\frac{d}{2}}\int_{t_{n-1}}^{t_n}\|\nabla\mathbf{R}(t)\|_{L^2(\Omega)}\|\nabla\boldsymbol{\theta}_{\phi'}(t)\|_{L^2(\Omega)}\|\nabla\boldsymbol{\Theta}(t)\|_{L^2(\Omega)}\,\mathrm{d}t\\
&\le \tilde{C}\int_{t_{n-1}}^{t_n}\|\nabla\boldsymbol{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t+Ch^{-d}\int_{t_{n-1}}^{t_n}\|\nabla\mathbf{R}(t)\|^2_{L^2(\Omega)}\|\nabla\boldsymbol{\theta}_{\phi'}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\
&\le \tilde{C}\int_{t_{n-1}}^{t_n}\|\nabla\boldsymbol{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t+2C\max_{t\in I_n}\|\nabla\mathbf{R}(t)\|^2_{L^2(\Omega)}h^{-d}k_n\|\nabla\boldsymbol{\theta}_{\phi'}(t_n^-)\|^2_{L^2(\Omega)}\\
&\quad+2C\max_{t\in I_n}\|\nabla\mathbf{R}(t)\|^2_{L^2(\Omega)}h^{-d}k_n\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\theta}}_{\phi'}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t, \qquad (3.3.51)
\end{aligned}
$$

where $\tilde{C}_{\mathrm{Lip}}=\hat{C}(3C_\tau+c_0+c_1)$ and C is a generic positive constant. Combining the estimates (3.3.41)–(3.3.47) and (3.3.51), we obtain

$$
\begin{aligned}
&\|\dot{\boldsymbol{\Theta}}(t_n^-)\|^2_{L^2(\Omega)}+\gamma^2\|\boldsymbol{\Theta}(t_n^-)\|^2_{L^2(\Omega)}+\mathsf{A}(\boldsymbol{\phi}_n^-\mathrm{e}^{\gamma t_n};\boldsymbol{\Theta}_n^-,\boldsymbol{\Theta}_n^-)+\gamma\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\Theta}}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\
&\le \|\dot{\boldsymbol{\Theta}}(t_{n-1}^-)\|^2_{L^2(\Omega)}+\gamma^2\|\boldsymbol{\Theta}(t_{n-1}^-)\|^2_{L^2(\Omega)}+\mathsf{A}(\boldsymbol{\phi}_{n-1}^-\mathrm{e}^{\gamma t_{n-1}};\boldsymbol{\Theta}_{n-1}^-,\boldsymbol{\Theta}_{n-1}^-)+C\int_{t_{n-1}}^{t_n}\|\nabla\boldsymbol{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\
&\quad+C(h^{-d-2}+h^{-d})k_n\left(\|\nabla\boldsymbol{\theta}_{\phi'}(t_n^-)\|^2_{L^2(\Omega)}+\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\theta}}_{\phi'}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\right)\max_{t\in I_n}\|\nabla\mathbf{R}(t)\|^2_{L^2(\Omega)}\\
&\quad+C\max_{t\in I_n}\|\nabla\mathbf{R}(t)\|^2_{L^2(\Omega)}h^{-d-2}\left(\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\theta}}_{\phi'}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\right)+C(\gamma,\mathbf{W})h^{-d-2}\frac{k_n^{2q_n+3}}{q_n^{2(s-1)}}\max_{t\in I_n}\|\nabla\mathbf{R}(t)\|^2_{L^2(\Omega)}\\
&\quad+\mathsf{A}(\boldsymbol{\phi'}(t_n^-)\mathrm{e}^{\gamma t_n};\boldsymbol{\theta}_{\phi'}(t_n^-),\boldsymbol{\Theta}(t_n^-))-\mathsf{A}(\boldsymbol{\phi'}(t_{n-1}^-)\mathrm{e}^{\gamma t_{n-1}};\boldsymbol{\theta}_{\phi'}(t_{n-1}^-),\boldsymbol{\Theta}(t_{n-1}^-))\\
&\quad-\mathsf{A}(\boldsymbol{\phi}(t_n^-)\mathrm{e}^{\gamma t_n};\boldsymbol{\theta}_{\phi'}(t_n^-),\boldsymbol{\Theta}(t_n^-))+\mathsf{A}(\boldsymbol{\phi}(t_{n-1}^-)\mathrm{e}^{\gamma t_{n-1}};\boldsymbol{\theta}_{\phi'}(t_{n-1}^-),\boldsymbol{\Theta}(t_{n-1}^-)). \qquad (3.3.52)
\end{aligned}
$$

Summing up over $n=1,\dots,j$, we obtain

$$
\begin{aligned}
&\|\dot{\boldsymbol{\Theta}}(t_j^-)\|^2_{L^2(\Omega)}+\gamma^2\|\boldsymbol{\Theta}(t_j^-)\|^2_{L^2(\Omega)}+\mathsf{A}(\boldsymbol{\phi}(t_j^-)\mathrm{e}^{\gamma t_j};\boldsymbol{\Theta}(t_j^-),\boldsymbol{\Theta}(t_j^-))+\gamma\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\Theta}}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\
&\le C\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\boldsymbol{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t+C(\gamma,\mathbf{W})h^{-d-2}\sum_{n=1}^{j}\frac{k_n^{2q_n+3}}{q_n^{2(s-1)}}\max_{t\in I_n}\|\nabla\mathbf{R}(t)\|^2_{L^2(\Omega)}
\end{aligned}
$$

$$
\begin{aligned}
&+ C\left(h^{-d-2}+h^{-d}\right)\left(\sum_{n=1}^{j} k_n \|\nabla \boldsymbol{\theta}_{\phi'}(t_{n-1}^-)\|_{L^2(\Omega)}^2\right) \max_{t\in I_n, 1\le n\le j} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)}^2 \\
&+ C\left(h^{-d-2}+h^{-d}\right)\left(\sum_{n=1}^{j} k_n \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_{\phi'}\|_{L^2(\Omega)}^2 \,\mathrm{d}t\right) \max_{t\in I_n, 1\le n\le j} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)}^2 \\
&+ C h^{-d-2}\left(\sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\theta}}_{\phi'}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t\right) \max_{t\in I_n, 1\le n\le j} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)}^2 \\
&+ \mathsf{A}(\boldsymbol{\phi'}(t_j^-)\mathrm{e}^{\gamma t_j}; \boldsymbol{\theta}_{\phi'}(t_j^-), \boldsymbol{\Theta}(t_j^-)) - \mathsf{A}(\boldsymbol{\phi}(t_j^-)\mathrm{e}^{\gamma t_j}; \boldsymbol{\theta}_{\phi'}(t_j^-), \boldsymbol{\Theta}(t_j^-)) \\
\le\ & C \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\Theta}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + C(\gamma, \mathbf{W}) h^{-d-2} \sum_{n=1}^{j} \frac{k_n^{2q_n+3}}{q_n^{2(s-1)}} \max_{t\in I_n, 1\le n\le j} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)}^2 \\
&+ C(\mathbf{v})\left(h^{-d-2}+h^{-d}\right)\left(\sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right) \max_{t\in I_n, 1\le n\le j} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)}^2 \\
&+ C(\mathbf{v})\left(h^{-d-2}+h^{-d}\right) \max_{1\le n\le j} k_n \left(\sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right) \max_{t\in I_n, 1\le n\le j} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)}^2 \\
&+ \mathsf{A}(\boldsymbol{\phi'}(t_j^-)\mathrm{e}^{\gamma t_j}; \boldsymbol{\theta}_{\phi'}(t_j^-), \boldsymbol{\Theta}(t_j^-)) - \mathsf{A}(\boldsymbol{\phi}(t_j^-)\mathrm{e}^{\gamma t_j}; \boldsymbol{\theta}_{\phi'}(t_j^-), \boldsymbol{\Theta}(t_j^-)) \\
\le\ & C \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\Theta}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + C(\mathbf{v}) h^{-d-2} \left(\sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right) \max_{t\in I_n, 1\le n\le j} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)}^2 \\
&+ \mathsf{A}(\boldsymbol{\phi'}(t_j^-)\mathrm{e}^{\gamma t_j}; \boldsymbol{\theta}_{\phi'}(t_j^-), \boldsymbol{\Theta}(t_j^-)) - \mathsf{A}(\boldsymbol{\phi}(t_j^-)\mathrm{e}^{\gamma t_j}; \boldsymbol{\theta}_{\phi'}(t_j^-)\boldsymbol{\Theta}(t_j^-)), \qquad (3.3.53)
\end{aligned}
$$

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

$$
\begin{aligned}
&|A(\boldsymbol{\phi'}(t_j^-)\mathrm{e}^{\gamma t_j}; \boldsymbol{\theta}_{\phi'}(t_j^-), \boldsymbol{\Theta}(t_j^-)) - A(\boldsymbol{\phi}(t_j^-)\mathrm{e}^{\gamma t_j}; \boldsymbol{\theta}_{\phi'}(t_j^-), \boldsymbol{\Theta}(t_j^-))| \\
&\le L_\delta \|\nabla \mathbf{R}(t_j^-)\|_{L^2(\Omega)} \|\nabla \boldsymbol{\theta}_{\phi'}(t_j^-)\|_{L^\infty(\Omega)} \|\nabla \boldsymbol{\Theta}(t_j^-)\|_{L^2(\Omega)} \\
&\le C(\mathbf{v}) h^{-\frac{d}{2}} \left(\sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)^{\frac{1}{2}} \max_{t\in I_n, 1\le n\le j} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)} \|\nabla \boldsymbol{\Theta}(t_j^-)\|_{L^2(\Omega)}.
\end{aligned}
\qquad (3.3.54)
$$

Combining the estimates (3.3.53) and (3.3.54) and applying the assumption (S2b) to $\mathsf{A}(\boldsymbol{\phi}(t_j^-)\mathrm{e}^{\gamma t_j}; \boldsymbol{\Theta}(t_j^-), \boldsymbol{\Theta}(t_j^-))$ on the left-hand-side of the resulting inequality yield

$$\begin{aligned}
&\|\dot{\boldsymbol{\Theta}}(t_j^-)\|_{L^2(\Omega)}^2 + \gamma^2\|\boldsymbol{\Theta}(t_j^-)\|_{L^2(\Omega)}^2 + M_1\|\nabla\boldsymbol{\Theta}(t_j^-)\|_{L^2(\Omega)}^2 + \gamma\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\Theta}}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \\
&\leq C\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\boldsymbol{\Theta}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t + C(\mathbf{v})h^{-d-2}\left(\sum_{n=1}^{j}k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n,\,1\leq n\leq j}\|\nabla\mathbf{R}(t)\|_{L^2(\Omega)}^2 \\
&\quad + C(\mathbf{v})h^{-\frac{d}{2}}\left(\sum_{n=1}^{j}k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}}\right)^{\frac{1}{2}}\max_{t\in I_n,\,1\leq n\leq j}\|\nabla\mathbf{R}(t)\|_{L^2(\Omega)}\|\nabla\boldsymbol{\Theta}(t_j^-)\|_{L^2(\Omega)}.
\end{aligned} \tag{3.3.55}$$

By applying Young's inequality on the right-hand side of (3.3.55), we have

$$\begin{aligned}
&\|\dot{\boldsymbol{\Theta}}(t_j^-)\|_{L^2(\Omega)}^2 + \gamma^2\|\boldsymbol{\Theta}(t_j^-)\|_{L^2(\Omega)}^2 + M_1\|\nabla\boldsymbol{\Theta}(t_j^-)\|_{L(\Omega)}^2 + \gamma\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\Theta}}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \\
&\leq C\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\boldsymbol{\Theta}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t + C(\mathbf{v})h^{-d-2}\left(\sum_{n=1}^{j}k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n,\,1\leq n\leq j}\|\nabla\mathbf{R}(t)\|_{L^2(\Omega)}^2 \\
&\quad + C(M_1,\mathbf{v})h^{-d}\left(\sum_{n=1}^{j}k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}}\right)\max_{t\in I_n,\,1\leq n\leq j}\|\nabla\mathbf{R}(t)\|_{L^2(\Omega)}^2 + \frac{M_1}{2}\|\nabla\boldsymbol{\Theta}(t_j^-)\|_{L^2(\Omega)},
\end{aligned} \tag{3.3.56}$$

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

$$\begin{aligned}
&\|\dot{\boldsymbol{\Theta}}(t_j^-)\|_{L^2(\Omega)}^2 + \gamma^2\|\boldsymbol{\Theta}(t_j^-)\|_{L^2(\Omega)}^2 + \frac{M_1}{2}\|\nabla\boldsymbol{\Theta}(t_j^-)\|_{L^2(\Omega)}^2 + \gamma\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\Theta}}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \\
&\leq C\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\boldsymbol{\Theta}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t + C(\mathbf{v})h^{-d-2}\left(\sum_{n=1}^{j}k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n,\,1\leq n\leq j}\|\nabla\mathbf{R}(t)\|_{L^2(\Omega)}^2 \\
&\quad + C(M_1,\mathbf{v})h^{-d}\left(\sum_{n=1}^{j}k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}}\right)\max_{t\in I_n,\,1\leq n\leq j}\|\nabla\mathbf{R}(t)\|_{L^2(\Omega)}^2 \\
&\leq C\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\boldsymbol{\Theta}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t + C(\mathbf{v})h^{-d-2}\left(\sum_{n=1}^{j}k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n,\,1\leq n\leq j}\|\nabla\mathbf{R}(t)\|_{L^2(\Omega)}^2.
\end{aligned} \tag{3.3.57}$$

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

$$\|\dot{\boldsymbol{\Theta}}(t_j^-)\|^2_{L^2(\Omega)} + \|\boldsymbol{\Theta}(t_j^-)\|^2_{L^2(\Omega)} + \|\nabla\boldsymbol{\Theta}(t_j^-)\|^2_{L^2(\Omega)} + \sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \|\dot{\boldsymbol{\Theta}}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$
$$\leq \tilde{C}(\mathbf{v})h^{-d-2}\left(\sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n,\,1\leq n\leq j}\|\nabla\mathbf{R}(t)\|^2_{L^2(\Omega)}, \tag{3.3.58}$$

where $\tilde{C}(\mathbf{v})$ is a positive constant depending on $\mathbf{v}$ 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 $n = 1, \dots N$,

$$\|\nabla\boldsymbol{\Theta}(t)\|_{L^2(\Omega)} \leq \|\nabla\boldsymbol{\Theta}(t_n^-)\|_{L^2(\Omega)} + \int_{t_{n-1}}^{t_n}\|\nabla\dot{\boldsymbol{\Theta}}(t)\|_{L^2(\Omega)}\,\mathrm{d}t$$
$$\leq \|\nabla\boldsymbol{\Theta}(t_n^-)\|_{L^2(\Omega)} + C_0h^{-1}k_n^{\frac{1}{2}}\left(\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\Theta}}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\right)^{\frac{1}{2}}.$$

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

$$\|\nabla\boldsymbol{\Theta}(t)\|_{L^2(\Omega)} \leq (1 + C_0h^{-1}k_n^{\frac{1}{2}})\tilde{C}(\mathbf{v})^{\frac{1}{2}}h^{-\frac{d}{2}-1}\left(\sum_{i=1}^{N} k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}}\right)^{\frac{1}{2}}\max_{t\in I_n,\,1\leq n\leq N}\|\nabla\mathbf{R}(t)\|_{L^2(\Omega)}$$
$$\leq \tilde{C}_1(\mathbf{v})h^{-\frac{d}{2}-1}\left(\sum_{i=1}^{N} k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}}\right)^{\frac{1}{2}}\max_{t\in I_n,\,1\leq n\leq N}\|\nabla\mathbf{R}(t)\|_{L^2(\Omega)}, \tag{3.3.59}$$

where the last inequality follows from the assumption that $\mu_i k_i \leq h^2$ for each $1 \leq i \leq N$. Analogously, we have

$$\|\boldsymbol{\Theta}(t)\|_{L^2(\Omega)} \leq \|\boldsymbol{\Theta}(t_n)\|_{L^2(\Omega)} + k_n^{\frac{1}{2}}\left(\int_{t_{n-1}}^{t_n}\|\dot{\boldsymbol{\Theta}}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\right)^{\frac{1}{2}}$$
$$\leq \tilde{C}_2(\mathbf{v})h^{-\frac{d}{2}-1}\left(\sum_{i=1}^{N} k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}}\right)^{\frac{1}{2}}\max_{t\in I_n,\,1\leq n\leq N}\|\nabla\mathbf{R}(t)\|_{L^2(\Omega)}, \tag{3.3.60}$$

and

$$\|\dot{\mathbf{\Theta}}(t)\|_{L^2(\Omega)} \leq \|\dot{\mathbf{\Theta}}(t_n)\|_{L^2(\Omega)} + C_2 k_n^{-\frac{1}{2}} \left(\int_{t_{n-1}}^{t_n} \|\dot{\mathbf{\Theta}}(t)\|^2_{L^2(\Omega)} \, \mathrm{d}t \right)^{\frac{1}{2}}$$

$$\leq \tilde{C}_3(\mathbf{v}) h^{-\frac{d}{2}-1} k_n^{-\frac{1}{2}} \left(\sum_{i=1}^{N} k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} \max_{t \in I_n,\, 1 \leq n \leq N} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)}. \tag{3.3.61}$$

Summing up (3.3.59)–(3.3.61) and taking maximum on the left-hand for $t \in I_n$, with $1 \leq n \leq N$, we have

$$\max_{t \in I_n, 1 \leq n \leq N} \left(\|\mathbf{\Theta}(t)\|_{H^1(\Omega)} + \|\dot{\mathbf{\Theta}}(t)\|_{L^2(\Omega)} \right)$$

$$\leq \widehat{C}(\mathbf{v}) h^{-\frac{d}{2}-1} \max_{1 \leq n \leq N} k_n^{-\frac{1}{2}} \left(\sum_{i=1}^{N} k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} \max_{t \in I_n, 1 \leq n \leq N} \|\nabla \mathbf{R}(t)\|_{L^2(\Omega)}. \tag{3.3.62}$$

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

$$\widehat{C}(\mathbf{v}) h^{-\frac{d}{2}-1} \max_{1 \leq n \leq N} k_n^{-\frac{1}{2}} \left(\sum_{i=1}^{N} k_i h^{2r+2} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} < 1, \tag{3.3.63}$$

we obtain

$$\max_{t \in I_n, 1 \leq n \leq N} \left(\|\mathbf{\Theta}(t)\|_{H^1(\Omega)} + \|\dot{\mathbf{\Theta}}(t)\|_{L^2(\Omega)} \right) < \max_{t \in I_n, 1 \leq n \leq N} \left(\|\mathbf{R}(t)\|_{H^1(\Omega)} + \|\dot{\mathbf{R}}(t)\|_{L^2(\Omega)} \right). \tag{3.3.64}$$

Indeed, the inequality (3.3.63) follows from our assumptions that $r > \frac{d}{2} + 1$, $k_i^{q_i - \frac{1}{2}} = o(h^{1+\frac{d}{2}})$ for each $i = 1, 2, \ldots, N$.

Therefore, by Banach's fixed point theorem, $\mathbf{v}_{\mathrm{DG}} = \mathbf{v}_\phi$ is the unique solution to (3.2.10). By the triangle inequality, properties of the modified projection operator, and property (iii,b) of the nonlinear projection $\mathbf{W}$, we have

$$\|\mathbf{v}_{\mathrm{DG}}(t_j^-) - \mathbf{v}(t_j^-)\|_{L^2(\Omega)} + \|\dot{\mathbf{v}}_{\mathrm{DG}}(t_j^-) - \dot{\mathbf{v}}(t_j^-)\|_{L^2(\Omega)}$$

$$\leq \|\boldsymbol{\theta}(t_j^-)\|_{L^2(\Omega)} + \|\dot{\boldsymbol{\theta}}(t_j^-)\|_{L^2(\Omega)} + \|\mathbf{W}(t_j^-) - \mathbf{v}(t_j^-)\|_{L^2(\Omega)} + \|\dot{\mathbf{W}}(t_j^-) - \dot{\mathbf{v}}(t_j^-)\|_{L^2(\Omega)}$$

$$\le C_*(\mathbf{v})\bigg(\sum_{n=1}^{j} k_n h^{2r+2} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\bigg)^{\frac{1}{2}} + 2\tilde{C}_r(\mathbf{v})h^{r+1} \qquad \text{(by (iii,b))}$$
$$\le C(\mathbf{v})\bigg(h^{2r+2} + \sum_{n=1}^{j} \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\bigg)^{\frac{1}{2}}.$$

3.4. Approximation properties of the elliptic projection

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

3.4.1. L^2 bound on $\nabla(\mathbf{v} - \mathbf{W})$ and L^∞ bound on $\nabla\mathbf{W}$ Recall that for each $t \in [0, T]$, $a(\mathbf{W}(t), \boldsymbol{\varphi}) = a(\mathbf{v}(t), \boldsymbol{\varphi})$ for all $\boldsymbol{\varphi} \in \mathcal{V}_h$. Let $\mathcal{P}_h \colon L^2 \to \mathcal{V}_h$ denote the standard L^2-projection operator in the spatial direction. Then we have

$$a(\mathbf{W}(t), \boldsymbol{\varphi}) - a(\mathcal{P}_h\mathbf{v}(t), \boldsymbol{\varphi}) = a(\mathbf{v}(t), \boldsymbol{\varphi}) - a(\mathcal{P}_h\mathbf{v}(t), \boldsymbol{\varphi}) \quad \text{for all } \boldsymbol{\varphi} \in \mathcal{V}_h. \tag{3.4.1}$$

That is,

$$\int_0^1 \tilde{a}(\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t), \boldsymbol{\varphi})\,\mathrm{d}\tau$$
$$= \int_0^1 \tilde{a}(\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}; \mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t), \boldsymbol{\varphi})\,\mathrm{d}\tau.$$

For $\mathbf{v} \in [H^m(\Omega)]^d \cap [H_0^1(\Omega)]^d$, we define the following subset of $[H_0^1(\Omega)]^d$,

$$\mathcal{F} = \{\boldsymbol{\phi} \in \mathcal{V}_h \colon \|\nabla(\boldsymbol{\phi} - \mathcal{P}_h\mathbf{v})\|_{L^2(\Omega)} \le C_* h^r \|\mathbf{v}\|_{H^{r+1}(\Omega)} \text{ for } \frac{d}{2} < r \le \min(p, m-1)\}$$

where C_* is a constant independent of h. The set $\mathcal{F}$ is non-empty since for each fixed $t \in [0, T]$, $\mathcal{P}_h\mathbf{v}(t) \in \mathcal{F}$. Furthermore, $\mathcal{F}$ is a closed and convex subset of $[H_0^1(\Omega)]^d$. We define the fixed point mapping $\mathcal{N}$ on $\mathcal{F}$ as follows. Given $\boldsymbol{\phi} \in \mathcal{F}$, we denote by $\mathbf{W}_\phi \in \mathcal{V}_h$, the solution to the following variational problem: find $\mathbf{W}_\phi \in \mathcal{V}_h$ such that

$$\int_0^1 \tilde{a}(\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\boldsymbol{\phi}(t) - \mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}; \mathbf{W}_\phi(t) - \mathcal{P}_h\mathbf{v}(t), \boldsymbol{\varphi})\,\mathrm{d}\tau$$

$$= \int_0^1 \tilde{a}(\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}; \mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t), \boldsymbol{\varphi})\,\mathrm{d}\tau \qquad \text{for all } \boldsymbol{\varphi} \in \mathcal{V}_h.$$

Since $\mathcal{V}_h$ is a finite-dimensional linear space, the existence and uniqueness of $\mathbf{W}_\phi(t) \in \mathcal{V}_h$ for each $t \in [0,T]$ follow if we can show that $\int_0^1 \tilde{a}(\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\boldsymbol{\phi}(t) - \mathcal{P}_h\mathbf{v}(t)); \cdot, \cdot)$ 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 $\mathbf{W}(t) = \mathbf{W}_\phi(t)$, we have

$$\|\nabla(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^2(\Omega)} \leq C_* h^r \|\mathbf{v}(t)\|_{H^{r+1}(\Omega)}, \quad \frac{d}{2} < r \leq \min(p, m-1). \tag{3.4.2}$$

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

$$\|\nabla(\mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^2(\Omega)} \lesssim h^r \|\mathbf{v}(t)\|_{H^{r+1}(\Omega)}, \quad \frac{d}{2} < r \leq \min(p, m-1). \tag{3.4.3}$$

It follows from the triangle inequality that

$$\|\nabla(\mathbf{W}(t) - \mathbf{v}(t))\|_{L^2(\Omega)} \lesssim h^r \|\mathbf{v}(t)\|_{H^{r+1}(\Omega)}, \quad \frac{d}{2} < r \leq \min(p, m-1). \tag{3.4.4}$$

By the approximation properties of $\mathcal{P}_h$ in the $|\cdot|_{W^{1,\infty}}$ semi-norm, we have

$$\|\nabla(\mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^\infty(\Omega)} \lesssim h^{r-\frac{d}{2}} \|\mathbf{v}(t)\|_{H^{r+1}(\Omega)}, \quad \frac{d}{2} < r \leq \min(p, m-1). \tag{3.4.5}$$

Combining (3.4.4) and (3.4.5), we obtain

$$\begin{aligned}
\|\nabla\mathbf{W}(t)\|_{L^\infty(\Omega)} &\leq \|\nabla\mathbf{v}(t)\|_{L^\infty(\Omega)} + \|\nabla(\mathbf{W}(t) - \mathbf{v}(t))\|_{L^\infty(\Omega)} \\
&\leq \|\nabla\mathbf{v}(t)\|_{L^\infty(\Omega)} + \|\nabla(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^\infty(\Omega)} + \|\nabla(\mathcal{P}_h\mathbf{v}(t) - \mathbf{v}(t))\|_{L^\infty(\Omega)} \\
&\leq \|\nabla\mathbf{v}(t)\|_{L^\infty(\Omega)} + C_1 h^{-\frac{d}{2}} \|\nabla(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^2(\Omega)} + C(\mathbf{v}) h^{r-\frac{d}{2}} \\
&\leq c_0,
\end{aligned}$$

for some constant c_0. The last inequality follows from the boundedness of $\nabla\mathbf{v}$ and the fact that $r > \frac{d}{2}$, while the second last line follows from (ii,b) and (3.4.5).

3.4.2. L^2 bound on $\nabla(\dot{\mathbf{W}} - \dot{\mathbf{v}})$ and L^∞ bound on $\nabla\dot{\mathbf{W}}$ For the estimate of the L^2–bound on $\nabla(\dot{\mathbf{W}} - \dot{\mathbf{v}})$, we follow the proof from Section 6 in [60]. We need to show that

$t \mapsto \mathbf{W}(t)$ is differentiable with respect to t. For $\mathbf{U} \in \mathcal{V}_h$ and $t \in [0,T]$, we notice that the mapping $\boldsymbol{\varphi} \mapsto a(\mathbf{U},\boldsymbol{\varphi}) - a(\mathbf{v}(t),\boldsymbol{\varphi})$ is a bounded linear functional on $\mathcal{V}_h$; hence by the *Riesz representation theorem*, there exists a unique $\mathcal{A}(t,\mathbf{U}) \in \mathcal{V}_h$ such that

$$(\mathcal{A}(t,\mathbf{U}),\boldsymbol{\varphi}) = a(\mathbf{U},\boldsymbol{\varphi}) - a(\mathbf{v}(t),\boldsymbol{\varphi}).$$

It follows from the linearisation process that the derivative of the nonlinear mapping $(t,\mathbf{U}) \mapsto \mathcal{A}(t,\mathbf{U})$ with respect to $\mathbf{U}$, evaluated at $\mathbf{U} = \mathbf{W}(t)$, exists and is invertible for any $t \in [0,T]$. We also have $\mathcal{A}(t,\mathbf{W}(t)) = 0$. Since $\mathbf{v}(t)$ is differentiable with respect to t, it follows that $\mathcal{A}(t,\mathbf{U})$ is differentiable in a neighbourhood of $(t_0,\mathbf{W}(t_0))$ for any $t_0 \in (0,T)$. We then deduce from the *implicit function theorem* that $t \mapsto \mathbf{W}(t)$ is differentiable in $(0,T)$. Next, we derive an error bound on $\|\nabla(\dot{\mathbf{W}}(t) - \dot{\mathbf{v}}(t))\|_{L^2(\Omega)}$. By the definition of $\mathbf{W}(t)$, we have

$$\begin{aligned}
&\int_0^1 \tilde{a}(\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}; \mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t), \boldsymbol{\varphi})\,\mathrm{d}\tau \\
&\quad = \int_0^1 \tilde{a}(\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}; \mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t), \boldsymbol{\varphi})\,\mathrm{d}\tau,
\end{aligned}$$

for all $\boldsymbol{\varphi} \in \mathcal{V}_h$. After differentiation with respect to t, we have

$$\begin{aligned}
&\int_0^1 \tilde{a}(\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}; \dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t), \boldsymbol{\varphi})\,\mathrm{d}\tau \\
&\quad + \int_0^1 \int_\Omega \sum_{i,\alpha,j,\beta,k,m=1}^{d} \bigg(\bigg\{ \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{km}} (\nabla\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\nabla\mathbf{W}(t) - \nabla\mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}) \\
&\quad \times \partial_m \left(\partial_t[\mathcal{P}_h\mathbf{v}_k(t)\mathrm{e}^{\gamma t}] + \tau\partial_t[(\mathbf{W}_k(t) - \mathcal{P}_h\mathbf{v}_k(t))\mathrm{e}^{\gamma t}]\right) \bigg\} \partial_j(\mathbf{W} - \mathcal{P}_h\mathbf{v})_\beta, \partial_j\boldsymbol{\varphi}_\alpha \bigg)_{L^2} \mathrm{d}x\,\mathrm{d}\tau \\
&= \int_0^1 \tilde{a}(\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}; \dot{\mathbf{v}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t), \boldsymbol{\varphi})\,\mathrm{d}\tau \\
&\quad + \int_0^1 \int_\Omega \sum_{i,\alpha,j,\beta,k,m=1}^{d} \bigg(\bigg\{ \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{km}} (\nabla\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\nabla\mathbf{v}(t) - \nabla\mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}) \\
&\quad \times \partial_m \left(\partial_t[\mathcal{P}_h\mathbf{v}_k(t)\mathrm{e}^{\gamma t}] + \tau\partial_t[(\mathbf{v}_k(t) - \mathcal{P}_h\mathbf{v}_k(t))\mathrm{e}^{\gamma t}]\right) \bigg\} \partial_j(\mathbf{v} - \mathcal{P}_h\mathbf{v})_\beta, \partial_j\boldsymbol{\varphi}_\alpha \bigg)_{L^2} \mathrm{d}x\,\mathrm{d}\tau,
\end{aligned}$$

for all $\boldsymbol{\varphi} \in \mathcal{V}_h$. Rearranging gives

$$
\begin{aligned}
&\int_0^1 \tilde{a}(\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}; \dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t), \boldsymbol{\varphi})\,\mathrm{d}\tau \\
&= \int_0^1 \tilde{a}(\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}; \dot{\mathbf{v}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t), \boldsymbol{\varphi})\,\mathrm{d}\tau \\
&\quad + \int_0^1 \int_\Omega \sum_{i,\alpha,j,\beta,k,m=1}^{d} \Bigg(\Bigg\{ \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{km}}(\nabla\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\nabla\mathbf{v}(t) - \nabla\mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}) \\
&\quad \times \partial_m\left(\partial_t[\mathcal{P}_h\mathbf{v}_k(t)\mathrm{e}^{\gamma t}] + \tau\partial_t[(\mathbf{v}_k(t) - \mathcal{P}_h\mathbf{v}_k(t))\mathrm{e}^{\gamma t}]\right)\Bigg\}\partial_j(\mathbf{v} - \mathcal{P}_h\mathbf{v})_\beta, \partial_j\boldsymbol{\varphi}_\alpha\Bigg)_{L^2}\,\mathrm{d}x\,\mathrm{d}\tau \\
&\quad - \int_0^1 \int_\Omega \sum_{i,\alpha,j,\beta,k,m=1}^{d} \Bigg(\Bigg\{ \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{km}}(\nabla\mathcal{P}_h\mathbf{v}(t)\mathrm{e}^{\gamma t} + \tau(\nabla\mathbf{W}(t) - \nabla\mathcal{P}_h\mathbf{v}(t))\mathrm{e}^{\gamma t}) \\
&\quad \times \partial_m\left(\partial_t[\mathcal{P}_h\mathbf{v}_k(t)\mathrm{e}^{\gamma t}] + \tau\partial_t[(\mathbf{W}_k(t) - \mathcal{P}_h\mathbf{v}_k(t))\mathrm{e}^{\gamma t}]\right)\Bigg\}\partial_j(\mathbf{W} - \mathcal{P}_h\mathbf{v})_\beta, \partial_j\boldsymbol{\varphi}_\alpha\Bigg)_{L^2}\,\mathrm{d}x\,\mathrm{d}\tau \\
&:= T_1 + T_2 + T_3.
\end{aligned}
$$

Taking $\boldsymbol{\varphi}(t) = \dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t)$, we have

$$
\begin{aligned}
T_1 &\lesssim h^r\|\dot{\mathbf{v}}(t)\|_{H^{r+1}(\Omega)}\|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^2(\Omega)}, \\
T_2 &\lesssim \Bigg(\|\nabla\mathcal{P}_h\dot{\mathbf{v}}(t)\|_{L^\infty(\Omega)} + \|\nabla(\dot{\mathbf{v}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^\infty(\Omega)} + \|\nabla\mathcal{P}_h\mathbf{v}(t)\|_{L^\infty(\Omega)} \\
&\qquad + \|\nabla(\mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^\infty(\Omega)}\Bigg) \times \|\nabla\mathbf{v}(t) - \nabla\mathcal{P}_h\mathbf{v}(t)\|_{L^2(\Omega)}\|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \\
&\lesssim h^r\|\mathbf{v}(t)\|_{H^{r+1}(\Omega)}\|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^2(\Omega)}, \\
T_3 &\lesssim \Bigg(\|\nabla\mathcal{P}_h\dot{\mathbf{v}}(t)\|_{L^\infty(\Omega)} + \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^\infty(\Omega)} + \|\nabla\mathcal{P}_h\mathbf{v}(t)\|_{L^\infty(\Omega)} \\
&\qquad + \|\nabla(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^\infty(\Omega)}\Bigg) \times \|\nabla(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^2(\Omega)}\|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \\
&\lesssim h^r\|\mathbf{v}(t)\|_{H^{r+1}(\Omega)}\Bigg(\|\nabla\mathcal{P}_h\dot{\mathbf{v}}(t)\|_{L^\infty(\Omega)} + \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^\infty(\Omega)} + \|\nabla\mathcal{P}_h\mathbf{v}(t)\|_{L^\infty(\Omega)} \\
&\qquad + \|\nabla(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^\infty(\Omega)}\Bigg) \times \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \\
&\lesssim h^r\|\mathbf{v}(t)\|_{H^{r+1}(\Omega)}\Bigg(\|\nabla\dot{\mathbf{v}}(t)\|_{L^\infty(\Omega)} + \|\nabla(\dot{\mathbf{v}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^\infty(\Omega)} + \|\nabla\mathbf{v}(t)\|_{L^\infty(\Omega)} \\
&\qquad + \|\nabla(\mathbf{v}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^\infty(\Omega)}\Bigg) \times \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^2(\Omega)}
\end{aligned}
$$

$$
\begin{aligned}
&+ h^{r-\frac{d}{2}} \|\mathbf{v}(t)\|_{H^{r+1}(\Omega)} \|\nabla(\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t))\|_{L^2(\Omega)} \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \\
&+ h^{r-\frac{d}{2}} \|\mathbf{v}(t)\|_{H^{r+1}(\Omega)} \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|^2_{L^2(\Omega)} \\
\lesssim{}& h^r \|\mathbf{v}(t)\|_{H^{r+1}(\Omega)} \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \\
&+ h^{r-\frac{d}{2}} \|\mathbf{v}(t)\|_{H^{r+1}(\Omega)} \|\nabla(\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t))\|_{L^2(\Omega)} \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \\
&+ h^{r-\frac{d}{2}} \|\mathbf{v}(t)\|_{H^{r+1}(\Omega)} \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|^2_{L^2(\Omega)}.
\end{aligned}
$$

Combining the estimates for T_1, T_2 and T_3, we have

$$
\begin{aligned}
&\int_0^1 \tilde{a}(\mathcal{P}_h \mathbf{v}(t) \mathrm{e}^{\gamma t} + \tau(\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t)) \mathrm{e}^{\gamma t}; \dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t), \dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \, \mathrm{d}\tau \\
&\lesssim h^r \left(\|\mathbf{v}(t)\|_{H^{r+1}(\Omega)} + \|\dot{\mathbf{v}}\|_{H^{r+1}(\Omega)} \right) \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \qquad (3.4.6) \\
&\quad + h^{r-\frac{d}{2}} \|\mathbf{v}\|_{H^{r+1}(\Omega)} \|\nabla(\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t))\|_{L^2(\Omega)} \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \\
&\quad + h^{r-\frac{d}{2}} \|\mathbf{v}\|_{H^{r+1}(\Omega)} \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|^2_{L^2(\Omega)}.
\end{aligned}
$$

Applying the strong ellipticity condition (S2b) on the left-hand side of (3.4.6), we have

$$
\begin{aligned}
&M_1 \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|^2_{L^2(\Omega)} \\
&\leq \int_0^1 \tilde{a}(\mathcal{P}_h \mathbf{v}(t) \mathrm{e}^{\gamma t} + \tau(\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t)) \mathrm{e}^{\gamma t}; \dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t), \dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t)) \, \mathrm{d}\tau \\
&\lesssim h^r \left(\|\mathbf{v}(t)\|_{H^{r+1}(\Omega)} + \|\dot{\mathbf{v}}\|_{H^{r+1}(\Omega)} \right) \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \\
&\quad + h^{r-\frac{d}{2}} \|\mathbf{v}\|_{H^{r+1}(\Omega)} \|\nabla(\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t))\|_{L^2(\Omega)} \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \\
&\quad + h^{r-\frac{d}{2}} \|\mathbf{v}\|_{H^{r+1}(\Omega)} \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|^2_{L^2(\Omega)}.
\end{aligned}
$$

Dividing by $\|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|_{L^2(\Omega)}$ on both sides yields

$$
\begin{aligned}
&M_1 \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \\
&\lesssim h^r \left(\|\mathbf{v}(t)\|_{H^{r+1}(\Omega)} + \|\dot{\mathbf{v}}\|_{H^{r+1}(\Omega)} \right) + h^{r-\frac{d}{2}} \|\mathbf{v}\|_{H^{r+1}(\Omega)} \|\nabla(\mathbf{W}(t) - \mathcal{P}_h \mathbf{v}(t))\|_{L^2(\Omega)} \\
&\quad + h^{r-\frac{d}{2}} \|\mathbf{v}\|_{H^{r+1}(\Omega)} \|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h \dot{\mathbf{v}}(t))\|_{L^2(\Omega)}.
\end{aligned}
$$

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

$$\begin{aligned}
&\|\nabla(\dot{\mathbf{W}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \\
&\lesssim h^r \left(\|\mathbf{v}(t)\|_{H^{r+1}(\Omega)} + \|\dot{\mathbf{v}}(t)\|_{H^{r+1}(\Omega)}\right) + h^{r-\frac{d}{2}}\|\mathbf{v}\|_{H^{r+1}(\Omega)}\|\nabla(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^2(\Omega)} \\
&\lesssim h^r \left(\|\mathbf{v}(t)\|_{H^{r+1}(\Omega)} + \|\dot{\mathbf{v}}(t)\|_{H^{r+1}(\Omega)}\right), \qquad (3.4.7)
\end{aligned}$$

where we have applied $\|\nabla(\mathbf{W}(t) - \mathcal{P}_h\mathbf{v}(t))\|_{L^2(\Omega)} \leq C_* h^r\|\mathbf{v}(t)\|_{H^{r+1}(\Omega)}$ for the last inequality. Again, by the approximation property of $\mathcal{P}_h$, we have, for each $t \in [0, T]$,

$$\|\nabla(\dot{\mathbf{v}}(t) - \mathcal{P}_h\dot{\mathbf{v}}(t))\|_{L^2(\Omega)} \lesssim h^r\|\dot{\mathbf{v}}(t)\|_{H^{r+1}(\Omega)}, \quad \frac{d}{2} < r \leq \min(p, m-1). \qquad (3.4.8)$$

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

$$\|\nabla\dot{\mathbf{W}}(t) - \nabla\dot{\mathbf{v}}(t)\|_{L^2(\Omega)} \lesssim h^r \left(\|\mathbf{v}(t)\|_{H^{r+1}(\Omega)} + \|\dot{\mathbf{v}}(t)\|_{H^{r+1}(\Omega)}\right), \quad \frac{d}{2} < r \leq \min(p, m-1). \qquad (3.4.9)$$

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

$$\|\nabla\dot{\mathbf{W}}(t)\|_{L^\infty(\Omega)} \leq c_1. \qquad (3.4.10)$$

3.4.3. L^2 bounds on $(\mathbf{v} - \mathbf{W})$, $(\dot{\mathbf{v}} - \dot{\mathbf{W}})$ and $(\ddot{\mathbf{v}} - \ddot{\mathbf{W}})$ It was proved by Dobrowolski and Rannacher in [26] that for each $t \in [0, T]$,

$$\|\mathbf{v}(t) - \mathbf{W}(t)\|_{L^2(\Omega)} \leq C_r(\mathbf{v})h^{r+1}, \quad \frac{d}{2} < r \leq \min(p, m-1). \qquad (3.4.11)$$

We shall focus on proving the L^2 error bound on the time derivative using a duality argument in this section.

Consider the following boundary value problem: for a given $\mathbf{g} \in [L^2(\Omega)]^d$, solve $\boldsymbol{\psi} \in [H_0^1(\Omega)]^d$ such that

$$\tilde{a}(\mathbf{v}; \boldsymbol{\psi}, \boldsymbol{\phi}) = (\mathbf{g}, \boldsymbol{\phi})_{L^2} \quad \text{for all } \boldsymbol{\phi} \in [H_0^1(\Omega)]^d, \qquad (3.4.12)$$

where $\mathbf{v}$ is the solution of (3.1.1)–(3.1.3) and

$$\tilde{a}(\mathbf{v};\boldsymbol{\psi},\boldsymbol{\phi}) := \sum_{i,\alpha,j,\beta=1}^{d} (A_{i\alpha j\beta}(\nabla\mathbf{v})\partial_\beta\boldsymbol{\psi}, \partial_\alpha\boldsymbol{\phi}_i)_{L^2}. \tag{3.4.13}$$

Since $A_{i\alpha j\beta}(\nabla\mathbf{v}) \in W^{1,\infty}(\Omega)$ provided that $A_{i\alpha j\beta}$ is sufficiently smooth and $\nabla\mathbf{v} \in [C^{2,\alpha}(\overline{\Omega})]^{d\times d}$ (cf. Remark 3.5), the adjoint problem (3.4.12) has a unique solution, which satisfies the following elliptic regularity conditions, cf. Theorem 1.1 and Theorem 2.6 of Chapter 8 in [19],

$$\|\boldsymbol{\psi}\|_{H^2(\Omega)} \leq \hat{c}\left(\|\boldsymbol{\psi}\|_{L^2(\Omega)} + \|\mathbf{g}\|_{L^2(\Omega)}\right) \tag{3.4.14}$$

for some positive constant $\hat{c}$. Taking $\boldsymbol{\phi} = \boldsymbol{\psi} \in [H_0^1(\Omega)]^d$ in (3.4.13) and applying the coercivity condition (S2b), we have

$$M_1\|\nabla\boldsymbol{\psi}\|_{L^2(\Omega)}^2 \leq \|\mathbf{g}\|_{L^2(\Omega)}\|\boldsymbol{\psi}\|_{L^2(\Omega)}. \tag{3.4.15}$$

Applying Poincaré's inequality in (3.4.15), we deduce that

$$\|\boldsymbol{\psi}\|_{L^2(\Omega)} \leq M_1^{-1}C_{\text{poin}}\|\mathbf{g}\|_{L^2(\Omega)}.$$

Thus

$$\|\boldsymbol{\psi}\|_{H^2(\Omega)} \leq c\|\mathbf{g}\|_{L^2(\Omega)}, \tag{3.4.16}$$

for some positive constant c. The corresponding discrete problem is formulated as follows: find $\boldsymbol{\psi}_h \in \mathcal{V}_h$ such that

$$\tilde{a}(\mathbf{v};\boldsymbol{\psi}_h,\boldsymbol{\phi}) = (\mathbf{g},\boldsymbol{\phi})_{L^2}, \qquad \text{for all } \boldsymbol{\phi} \in \mathcal{V}_h. \tag{3.4.17}$$

It is known (cf. [26]) that we have

$$\|\boldsymbol{\psi} - \boldsymbol{\psi}_h\|_{L^2(\Omega)} + h\|\boldsymbol{\psi} - \boldsymbol{\psi}_h\|_{H^1(\Omega)} \leq Ch^{r+1}\|\boldsymbol{\psi}\|_{H^{r+1}(\Omega)}, \tag{3.4.18}$$

for some constant C. Let $\mathbf{g} = \dot{\mathbf{v}} - \dot{\mathbf{W}}$; then (3.4.12) becomes

$$\tilde{a}(\mathbf{v},\boldsymbol{\psi},\boldsymbol{\phi}) = (\dot{\mathbf{v}} - \dot{\mathbf{W}},\boldsymbol{\phi})_{L^2}. \tag{3.4.19}$$

Plugging $\boldsymbol{\phi} = \dot{\mathbf{v}} - \dot{\mathbf{W}}$ into (3.4.19), we obtain

$$\|\dot{\mathbf{v}} - \dot{\mathbf{W}}\|^2_{L^2(\Omega)} = \tilde{a}(\mathbf{v}; \boldsymbol{\psi}, \dot{\mathbf{v}} - \dot{\mathbf{W}}). \tag{3.4.20}$$

Using (S2a) and the definition of the elliptic projection, we have,

$$\begin{aligned}
\tilde{a}(\mathbf{v}; \boldsymbol{\psi}, \dot{\mathbf{v}} - \dot{\mathbf{W}}) &= \tilde{a}(\mathbf{v}; \boldsymbol{\psi} - \boldsymbol{\psi}_h, \dot{\mathbf{v}} - \dot{\mathbf{W}}) + \tilde{a}(\mathbf{v}; \boldsymbol{\psi}_h, \dot{\mathbf{v}} - \dot{\mathbf{W}}) \\
&= \tilde{a}(\mathbf{v}; \boldsymbol{\psi} - \boldsymbol{\psi}_h, \dot{\mathbf{v}} - \dot{\mathbf{W}}) + \tilde{a}(\mathbf{v}; \dot{\mathbf{v}} - \dot{\mathbf{W}}, \boldsymbol{\psi}_h) \\
&= \tilde{a}(\mathbf{v}; \boldsymbol{\psi} - \boldsymbol{\psi}_h, \dot{\mathbf{v}} - \dot{\mathbf{W}}) + \tilde{a}(\mathbf{W}; \dot{\mathbf{W}}, \boldsymbol{\psi}_h) - \tilde{a}(\mathbf{v}; \dot{\mathbf{W}}, \boldsymbol{\psi}_h).
\end{aligned} \tag{3.4.21}$$

By (iii,a), (3.4.16) and (3.4.18), we have

$$\begin{aligned}
|\tilde{a}(\mathbf{v}; \boldsymbol{\psi} - \boldsymbol{\psi}_h, \dot{\mathbf{v}} - \dot{\mathbf{W}})| &\le K_\delta \|\nabla(\dot{\mathbf{v}} - \dot{\mathbf{W}})\|_{L^2(\Omega)} \|\nabla(\boldsymbol{\psi} - \boldsymbol{\psi}_{\boldsymbol{h}})\|_{L^2(\Omega)} \\
&\le K_\delta C_r(\mathbf{v}) h^r \cdot Ch \|\boldsymbol{\psi}\|_{H^2(\Omega)} \quad \text{(by (iii,a) and (3.4.18))} \\
&\le K_\delta c C_r(\mathbf{v}) C h^{r+1} \|\dot{\mathbf{v}} - \dot{\mathbf{W}}\|_{L^2(\Omega)} \quad \text{(by (3.4.16))}.
\end{aligned}$$

For the remaining terms in (3.4.21), we observe that

$$\begin{aligned}
|\tilde{a}(\mathbf{W}; \dot{\mathbf{W}}, \boldsymbol{\psi}_h) - \tilde{a}(\mathbf{v}; \dot{\mathbf{W}}, \boldsymbol{\psi}_h)| &\le |\tilde{a}(\mathbf{W}; \dot{\mathbf{W}}, \boldsymbol{\psi}_h - \boldsymbol{\psi}) - \tilde{a}(\mathbf{v}; \dot{\mathbf{W}}, \boldsymbol{\psi}_h - \boldsymbol{\psi})| \\
&\quad + |\tilde{a}(\mathbf{W}; \dot{\mathbf{W}} - \dot{\mathbf{v}}, \boldsymbol{\psi}) - \tilde{a}(\mathbf{v}; \dot{\mathbf{W}} - \dot{\mathbf{v}}, \boldsymbol{\psi})| \\
&\quad + |\tilde{a}(\mathbf{W}; \dot{\mathbf{v}}, \boldsymbol{\psi}) - \tilde{a}(\mathbf{v}; \dot{\mathbf{v}}, \boldsymbol{\psi})| \\
&:= T_4 + T_5 + T_6.
\end{aligned}$$

By the Lipschitz continuity of $A_{i\alpha j\beta}$, we have

$$\begin{aligned}
T_4 &\le L_\delta \|\nabla(\mathbf{W} - \mathbf{v})\|_{L^2(\Omega)} \|\nabla \dot{\mathbf{W}}\|_{L^\infty(\Omega)} \|\nabla(\boldsymbol{\psi}_h - \boldsymbol{\psi})\|_{L^2(\Omega)} \\
&\le C_r(\mathbf{v}) h^r c_1 \|\nabla(\boldsymbol{\psi}_h - \boldsymbol{\psi})\|_{L^2(\Omega)} \quad \text{(by (iii,a) and (iii,c))} \\
&\le C_r(\mathbf{v}) c_1 h^r \cdot Ch \|\boldsymbol{\psi}\|_{H^2(\Omega)} \quad \text{(by (3.4.18))} \\
&\le cC C_r(\mathbf{v}) c_1 h^{r+1} \|\dot{\mathbf{W}} - \dot{\mathbf{v}}\|_{L^2(\Omega)} \quad \text{(by (3.4.16))}.
\end{aligned}$$

Similarly, we have

$$T_5 \leq L_\delta \|\nabla(\mathbf{W}-\mathbf{v})\|_{L^\infty(\Omega)} \|\nabla(\dot{\mathbf{W}}-\dot{\mathbf{v}})\|_{L^2(\Omega)} \|\nabla \boldsymbol{\psi}\|_{L^2(\Omega)}.$$

Following the analysis in [63] and Chapter 8 of [17], it can be shown that

$$\|\mathbf{v}-\mathbf{W}\|_{W^{1,\infty}(\Omega)} \leq c(\mathbf{v}) h^r, \tag{3.4.22}$$

where $c(\mathbf{v})$ is a positive constant depending on the exact solution $\mathbf{v}$. Therefore, we can bound T_5 by

$$T_5 \leq L_\delta c(\mathbf{v}) h^r C_r(\mathbf{v}) h^r \|\boldsymbol{\psi}\|_{H^2(\Omega)} \leq L_\delta c(\mathbf{v}) C_r(\mathbf{v}) c h^{r+1} \|\dot{\mathbf{W}}-\dot{\mathbf{v}}\|_{L^2(\Omega)}, \tag{3.4.23}$$

for any $r \geq 1$ provided that h is sufficiently small. We bound T_6 by

$$\begin{aligned}
T_6 &= |\tilde{a}(\mathbf{W}; \dot{\mathbf{v}}, \boldsymbol{\psi}) - \tilde{a}(\mathbf{v}; \dot{\mathbf{v}}, \boldsymbol{\psi})| \\
&\leq \left| \sum_{i,\alpha,j,\beta,k,\gamma=1}^{d} \left(\partial_\gamma(\mathbf{W}_k - \mathbf{v}_k) \partial_\beta \dot{\mathbf{v}}_j \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{k\gamma}}(\nabla \mathbf{v}), \partial_\alpha \boldsymbol{\psi}_i \right)_{L^2} \right| \\
&\quad + \left| \sum_{i,\alpha,j,\beta,k,\gamma,l,\delta=1}^{d} \left(\partial_\delta(\mathbf{W}_l - \mathbf{v}_l) \partial_\gamma(\mathbf{W}_k - \mathbf{v}_k) \partial_\beta \dot{\mathbf{v}}_j \int_0^1 \frac{\partial^2 A_{i\alpha j\beta}}{\partial \eta_{l\delta} \partial \eta_{k\gamma}}(\nabla \mathbf{v} + \tau \nabla(\mathbf{W}-\mathbf{v}))\, \mathrm{d}\tau, \partial_\alpha \boldsymbol{\psi}_i \right)_{L^2} \right| \\
&:= b(\mathbf{W}, \mathbf{v}; \dot{\mathbf{v}}, \boldsymbol{\psi}) + d(\mathbf{W}, \mathbf{v}; \dot{\mathbf{v}}, \boldsymbol{\psi}).
\end{aligned}$$

To ensure that $\nabla \mathbf{W} \in \mathcal{Z}_\delta$, we take h sufficiently small. By the convexity of $\mathcal{Z}_\delta$ and $\mathcal{M}_\delta$, we know that $\nabla \mathbf{v}(x) + \tau \nabla(\mathbf{W}(x) - \mathbf{v}(x)) \in \mathcal{M}_\delta$ for $x \in \bar{\Omega}$. Since $A_{i\alpha j\beta}$ is sufficiently smooth (in particular twice continuously differentiable on $\mathcal{M}_\delta$), we have

$$\begin{aligned}
d(\mathbf{W}, \mathbf{v}; \dot{\mathbf{v}}, \boldsymbol{\psi}) &\leq C_A \|\nabla(\mathbf{W}-\mathbf{v})\|_{L^\infty(\Omega)}^2 \|\nabla \dot{\mathbf{v}}\|_{L^2(\Omega)} \|\nabla \boldsymbol{\psi}\|_{L^2(\Omega)} \\
&\leq C_A c(\mathbf{v})^2 h^{2r} \|\nabla \dot{\mathbf{v}}\|_{L^2(\Omega)} \|\nabla \boldsymbol{\psi}\|_{L^2(\Omega)} \quad \text{(by (3.4.22))} \\
&\leq C_A c(\mathbf{v})^2 h^{2r} \|\nabla \dot{\mathbf{v}}\|_{L^2(\Omega)} \|\boldsymbol{\psi}\|_{H^2(\Omega)} \\
&\leq C_A c(\mathbf{v})^2 c h^{2r} \|\nabla \dot{\mathbf{v}}\|_{L^2(\Omega)} \|\dot{\mathbf{W}}-\dot{\mathbf{v}}\|_{L^2(\Omega)} \quad \text{(by (3.4.16))} \\
&\leq C h^{r+1} \|\dot{\mathbf{W}}-\dot{\mathbf{v}}\|_{L^2(\Omega)}, \quad \text{for } r \geq 1.
\end{aligned}$$

For the estimation of $b(\mathbf{W}, \mathbf{v}; \dot{\mathbf{v}}, \boldsymbol{\psi})$, we apply integration by parts and the fact that $\mathcal{V}_h \subset [H_0^1(\Omega)]^d$ to obtain

$$
\begin{aligned}
b(\mathbf{W}, \mathbf{v}; \dot{\mathbf{v}}, \boldsymbol{\psi}) &= \left| \sum_{i,\alpha,j,\beta,k,\gamma=1}^{d} \left(\partial_\gamma \big[\partial_\beta \dot{\mathbf{v}}_j \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{k\gamma}}(\nabla \mathbf{v}), \partial_\alpha \boldsymbol{\psi}_i \big] (\mathbf{W}_k - \mathbf{v}_k) \right)_{L^2} \right| \\
&\leq C \|\mathbf{W} - \mathbf{v}\|_{L^2(\Omega)} \|\dot{\mathbf{v}}\|_{W^{2,\infty}(\Omega)} \|\mathbf{v}\|_{W^{2,\infty}(\Omega)} \|\boldsymbol{\psi}\|_{H^2(\Omega)} \\
&\leq \tilde{C} h^{r+1} \|\dot{\mathbf{W}} - \dot{\mathbf{v}}\|_{L^2(\Omega)} \quad \text{(by (3.4.16) and (3.4.11))}.
\end{aligned}
$$

Combining the above estimates for T_4, T_5 and T_6, we have

$$
\|\dot{\mathbf{W}} - \dot{\mathbf{v}}\|_{L^2(\Omega)} \leq \tilde{C}_r(\mathbf{v}) h^{r+1}, \tag{3.4.24}
$$

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

By a similar argument, we can easily show that $\dot{\mathbf{W}}(t)$ is differentiable with respect to t and derive a similar L^2 error estimate for $\ddot{\mathbf{W}} - \ddot{\mathbf{v}}$. The proof of this estimate can be found in [54] and [55]. We omit the details here.

3.5. Numerical experiments

In this section, we show some numerical experiments on a simple version of the nonlinear elastodynamics equation to verify the convergence result.

3.5.1. Numerical results for a nonlinear elastodynamics problem We consider the one-dimensional nonlinear equation

$$
\ddot{u}(x,t) + 2\gamma \dot{u}(x,t) + \gamma^2 u(x,t) - \partial_x[S(\partial_x u(x,t))] = f(x,t) \quad \text{in } (0,1) \times (0,T], \tag{3.5.1}
$$

$$
u(0,t) = u(1,t) = 0 \quad \text{for all } t \in (0,T], \tag{3.5.2}
$$

$$
u(x,0) = u_0(x), \quad \dot{u}(x,0) = u_1(x). \tag{3.5.3}
$$

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

$$u(x,t) = \sin(\sqrt{2}\pi t)\sin(\pi x).$$

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

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

Analogously to Section 2.4.1, we first discretise 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 discretisation parameter. The numerical approximation of the nonlinear wave-type equation following a Picard-type linearisation in the nonlinear term is the following: find $u_h \in \mathcal{V}_h$ such that

$$\int_\Omega \ddot{u}_h \cdot v_h \,\mathrm{d}x + \int_\Omega 2\gamma \dot{u}_h \cdot v_h \,\mathrm{d}x + \int_\Omega \gamma^2 u_h \cdot v_h \,\mathrm{d}x + \frac{1}{3}\int_\Omega (\partial_x u_h^*)^2 \partial_x u_h \cdot \partial_x v_h \,\mathrm{d}x = \int_\Omega f \cdot v_h \,\mathrm{d}x,$$

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 $\mathbf{U}(t)$:

$$\begin{cases} \tilde{M}\ddot{\mathbf{U}}(t) + 2\gamma\tilde{M}\dot{\mathbf{U}}(t) + \gamma^2\tilde{M}\mathbf{U}(t) + \frac{1}{3}\tilde{K}(t)\mathbf{U}(t) = \mathbf{F}(t), \quad t \in (0,T], \\ \dot{\mathbf{U}}(0) = \mathbf{U}_1, \quad \mathbf{U}(0) = \mathbf{U}_0. \end{cases}$$

Again, $\ddot{\mathbf{U}}(t)$ (respectively $\dot{\mathbf{U}}(t)$) represents the vector of nodal acceleration (respectively velocity) and $\mathbf{F}(t)$ is the vector of externally applied loads. $\tilde{M}$ is the mass matrix defined as before. However, in contrast to the discretisation for linear equations, the stiffness matrix $\tilde{K}(t)$ is now time-dependent:

$$\tilde{K}_{ij}(t) := \int_0^1 (\partial_x u_h^*(t))^2 \partial_x \psi_i(x) \cdot \partial_x \psi_j(x) \,\mathrm{d}x,$$

which is computed using u_h^* at each time step. Similarly to Section 2.4, we have

$$\ddot{\mathbf{Z}}(t) + L\dot{\mathbf{Z}}(t) + K_0\mathbf{Z}(t) + K(t)\mathbf{Z}(t) = \mathbf{G}(t), \quad t \in (0, T], \tag{3.5.4}$$

$$\dot{\mathbf{Z}}(0) = \tilde{M}^{\frac{1}{2}}\mathbf{U}_1, \quad \mathbf{Z}(0) = \tilde{M}^{\frac{1}{2}}\mathbf{U}_0. \tag{3.5.5}$$

Here

$$L = 2\gamma \mathrm{Id}, \quad K_0 = \gamma^2 \mathrm{Id}, \quad K(t) = \frac{1}{3}\tilde{M}^{-\frac{1}{2}}\tilde{K}(t)\tilde{M}^{-\frac{1}{2}}, \quad \mathbf{G}(t) = \tilde{M}^{-\frac{1}{2}}\mathbf{F}(t).$$

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 and assume that $q_1 = \cdots = q_N = q \geq 2$ on each time interval. If we consider the time integration on a generic time interval I_n for each $n = 1, \ldots N$, our DG in time formulation reads as: find $\mathbf{Z} \in \mathcal{V}_{kh}^{q_n}$ such that

$$\begin{aligned}
&\left(\ddot{\mathbf{Z}}(t), \dot{\mathbf{v}}\right)_{L^2(I_n)} + \left(L\dot{\mathbf{Z}}(t), \dot{\mathbf{v}}\right)_{L^2(I_n)} + (K_0\mathbf{Z}(t), \dot{\mathbf{v}})_{L^2(I_n)} + (K(t)\mathbf{Z}(t), \dot{\mathbf{v}})_{L^2(I_n)} \\
&\quad + \dot{\mathbf{Z}}(t_{n-1}^+) \cdot \dot{\mathbf{v}}(t_{n-1}^+) + K_0\mathbf{Z}(t_{n-1}^+) \cdot \mathbf{v}(t_{n-1}^+) + K(t_{n-1}^+)\mathbf{Z}(t_{n-1}^+) \cdot \mathbf{v}(t_{n-1}^+) \\
&= (\mathbf{G}(t), \dot{\mathbf{v}})_{L^2(I_n)} + \dot{\mathbf{Z}}(t_{n-1}^-) \cdot \dot{\mathbf{v}}(t_{n-1}^+) + K_0\mathbf{Z}(t_{n-1}^-) \cdot \mathbf{v}(t_{n-1}^+) + K(t_{n-1}^-)\mathbf{Z}(t_{n-1}^-) \cdot \mathbf{v}(t_{n-1}^+),
\end{aligned}$$

for all $\mathbf{v} \in \mathcal{V}_{kh}^{q_n}$. By choosing appropriate test functions as in the previous chapter, we obtain an algebraic system $\mathbf{Az} = \mathbf{b}$, where $\mathbf{z} \in \mathbb{R}^D = \mathbb{R}^{(q_n+1)\hat{d}}$ is the solution vector; $\mathbf{b} \in \mathbb{R}^D$ corresponds to the right-hand side, and $\mathbf{A}$ is the local stiffness matrix with its structure defined in Section 2.4. In this case, we set

$$M = M^1 + M^4, \quad B_{ij} = L_{ij}M^2 + K_{0ij}(M^3 + M^5) + (\tilde{M}^3 + \tilde{M}^5),$$

with $M, B_{ij} \in \mathbb{R}^{(q_n+1)\times(q_n+1)}$ for any $i, j = 1, \ldots, \hat{d}$. The new time-dependent matrices $\tilde{M}^3$ and $\tilde{M}^5$ are defined as

$$\tilde{M}_{lj}^3 = \left(K(t)\phi^j, \dot{\phi}^l\right)_{L^2(I_n)}, \quad \tilde{M}_{lj}^5 = K(t_{n-1}^+)\phi^j(t_{n-1}^+) \cdot \phi^l(t_{n-1}^+),$$

where ϕ^l, ϕ^j for $l, j = 1, \ldots, q_n + 1$ are the basis polynomials as defined in Chapter 2. The computation of $\tilde{M}^3(t)$ and $\tilde{M}^5(t)$ depends on the time-dependent matrix $K(t)$. In order to compute this matrix $K(t)$ (correspondingly $\tilde{K}(t)$), we apply a *Picard iteration* at each time interval. We set the maximal number of Picard iterations to be 30 at each time step and the tolerance to be $1\mathrm{e}-10$. The details of the algorithm are summarised in Algorithm 1.

We use CG–p elements where $p = q$ in space with $k = h^2$, $T = 1$ and $\gamma = 1$, and compute the errors $\|u(T) - u_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)} + \|\dot{u}(T) - \dot{u}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)}$ versus k for $k = h^2$ with respect to polynomial degrees $2, 3, 4$ in Table 3.1. Note that here we use $h = 2.50 \times 10^{-1}$, 2.00×10^{-1}, 1.25×10^{-1} and 6.25×10^{-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 3.1 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 3.3, 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 3.1.

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

Initialisation: $\partial_x u_h^* = \partial_x u_0$ and

$$[\tilde{K}^0]_{ij} = \int_0^1 (\partial_x u_0)^2 \partial_x \psi_i(x) \partial_x \psi_j(x) \, \mathrm{d}x.$$

Iteration: On each interval $I_n = (t_{n-1}, t_n]$ for $n = 1, 2, \ldots, N$, we solve

$$\tilde{M}\ddot{\mathbf{U}}(t) + 2\gamma \tilde{M}\dot{\mathbf{U}}(t) + \gamma^2 \tilde{M}\mathbf{U}(t) + \frac{1}{3}\tilde{K}^n(t)\mathbf{U}(t) = \mathbf{F}(t)$$

iteratively (using Picard iterations) by applying the discontinuous-in-time integration. Here

$$[\tilde{K}_0^n]_{ij} = \int_0^1 [\partial_x u_{\mathrm{DG}}^{n-1}(t)]^2 \partial_x \psi_i(x) \partial_x \psi_j(x) \, \mathrm{d}x,$$

where $u_{\mathrm{DG}}^{n-1}(t)$ is the solution we obtained from the previous time interval I_{n-1}.

$$[\tilde{K}_k^n]_{ij} = \int_0^1 [\partial_x u_{\mathrm{DG}}^{n-1,k-1}(t)]^2 \partial_x \psi_i(x) \partial_x \psi_j(x) \, \mathrm{d}x,$$

for $k = 1, 2, \ldots,$ where $u_{\mathrm{DG}}^{n-1,k-1}(t)$ is computed from the previous Picard iteration by using the stiffness matrix $[\tilde{K}_{k-1}^n]_{ij}$.

Update:

$$[\tilde{K}_0^{n+1}]_{ij} = \int_0^1 [\partial_x u_{\mathrm{DG}}^n(t)]^2 \partial_x \psi_i(x) \partial_x \psi_j(x) \, \mathrm{d}x,$$

where $\partial_x u_{\mathrm{DG}}^n(t)$ is computed using $[\tilde{K}_{k_{\mathrm{end}}}^n]$. Here k_{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}.

Table 3.1: 1D nonlinear elastodynamics equation with $p = q$: computed error $||u(T) - u_{\mathrm{DG}}(t_N^-)||_{L^2(\Omega)} + ||\dot{u}(T) - \dot{u}_{\mathrm{DG}}(t_N^-)||_{L^2(\Omega)}$ and corresponding convergence rates with respect to polynomial degrees $q = 2, 3, 4$.

q	h	$k = h^2$	L^2-error	rate
2	2.50000e − 1	6.25000e − 2	1.2123e − 2	—
	2.00000e − 1	4.00000e − 2	4.9774e − 3	1.9948
	1.25000e − 1	1.56250e − 2	1.1643e − 3	1.5455
	6.25000e − 2	3.90625e − 3	1.2454e − 4	1.6124
3	2.50000e − 1	6.25000e − 2	4.1533e − 4	—
	2.00000e − 1	4.00000e − 2	1.8590e − 4	1.8012
	1.25000e − 1	1.56250e − 2	2.5283e − 5	2.1224
	6.25000e − 2	3.90625e − 3	1.5609e − 6	2.0089
4	2.50000e − 1	6.25000e − 2	2.5498e − 5	—
	2.00000e − 1	4.00000e − 2	8.3999e − 6	2.4881
	1.25000e − 1	1.56250e − 2	9.7790e − 7	2.2878
	6.25000e − 2	3.90625e − 3	3.1115e − 8	2.4870

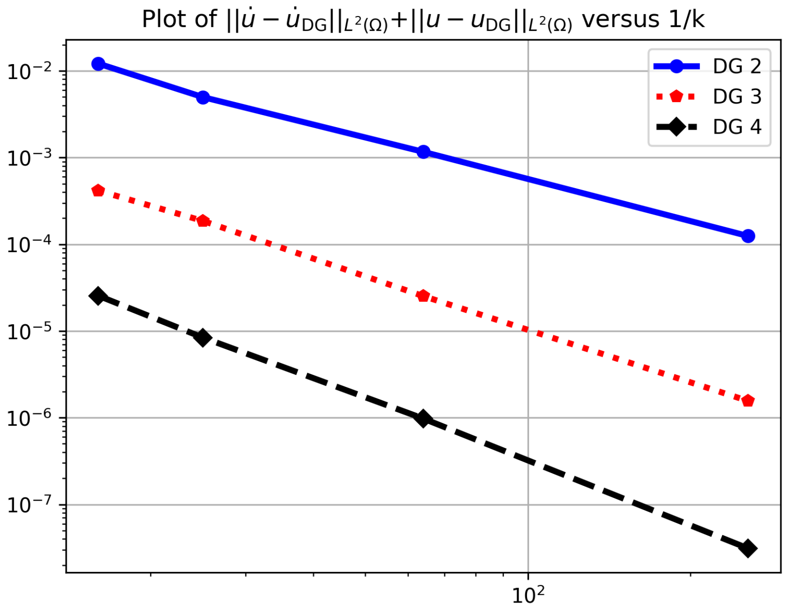

Figure 3.1: 1D nonlinear elastodynamics equation with $p = q$: computed error $\|u(T) - u_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)} + \|\dot{u}(T) - \dot{u}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)}$ plotted against $1/k$ for $q = 2, 3, 4$ in a log-log scale.

Chapter 4

Nonlinear damped wave equations

Let Ω be a bounded domain in $\mathbb{R}^d$ for $d = 1, 2, 3$, with smooth boundary $\partial\Omega$, and let $0 < T < \infty$. We consider the damped wave equation given by

$$\ddot{u}(x,t) + \partial_t a(u(x,t)) - \Delta\dot{u}(x,t) - \Delta b(u(x,t)) = 0 \text{ in } \Omega \times (0,T], \tag{4.0.1}$$

subject to the initial conditions

$$u(\cdot,0) = u_0(\cdot) \in H_0^1(\Omega), \;\; \dot{u}(\cdot,0) = u_1(\cdot) \in L^2(\Omega) \text{ in } \Omega, \tag{4.0.2}$$

and the homogeneous Dirichlet boundary conditions

$$u(x,t) = 0 \text{ on } \partial\Omega \times [0,T], \tag{4.0.3}$$

where $a \in C^3(\mathbb{R})$, $b(u) = u + c(u) \in C^4(\mathbb{R})$ with $c'(u) \geq M_0 > 0$ and $a'(u) \geq M_1 > 0$. We assume that $|a'(\cdot)| \leq C_a$ and $|c'(\cdot)| \leq C_c$ for some positive constants C_a and C_c. The motivation for studying the nonlinear wave equation of (4.0.1) comes from the nonlinear wave-dielectric interaction problem, derived from Maxwell's equations by assuming a linearly polarised wave propagating on an infinite cylindrical domain Ω (see Bloom [16]) of the form

$$\Omega = \{(x,y,z) \mid -\infty < x < \infty, \quad f(y,z) = c > 0\}.$$

The resulting one-dimensional nonlinear wave equation is of the following form.

$$\frac{\partial^2 D}{\partial t^2} + \Sigma'(D)\frac{\partial D}{\partial t} = \frac{1}{\mu_0}\frac{\partial^2 b(D)}{\partial x^2}, \tag{4.0.4}$$

where $D(x,t)$ is the scalar electric displacement, and μ_0 is a constant representing the permeability of free space. The nonlinear function b is defined as $b(D) = \lambda_0 D + \lambda_2 D^3$ with $\lambda_i > 0$ for $i = 0, 2$. Here $\Sigma(D) = \tilde{\sigma}(b(D))b(D)$ where $\tilde{\sigma}(\xi) = \boldsymbol{\sigma}(0,\xi,0)$ with $\boldsymbol{\sigma}$ being the nonlinear function appearing in Ohm's law. That is, $\mathbf{J} = \boldsymbol{\sigma}(\mathbf{E})\mathbf{E}$, where $\mathbf{J}$ is the current and $\mathbf{E}$ is the electric field. The detailed derivation from Maxwell's equations of this one-dimensional nonlinear equation can be found in Section 2.6 of [16]. The existence and uniqueness of a solution to the nonlinear wave equation (4.0.4) follows from [18] and Chapter 3 of Bloom's book [16]. The $\Delta\dot{u}$ term is added in (4.0.1) for the sake of ensuring stability of our numerical scheme. By performing an abstract Galerkin semi-discretisation, we will prove in the next section that the same local existence result holds for (4.0.1)–(4.0.3). After that, we will approximate (4.0.1)–(4.0.3) by a continuous Galerkin method in space first, then apply a DG discretisation in time.

4.1. Existence of a unique weak solution

In this section, we aim to show the existence of a unique weak solution to (4.0.1)–(4.0.3) on any space-time domain $\Omega \times [0,T]$.

Definition 4.1. *Suppose that $u_0 \in H_0^1(\Omega)$ and $u_1 \in L^2(\Omega)$, and that a and b satisfy the assumptions stated in the previous section. We say that a function $u \in L^2(0,T;H_0^1(\Omega))$, with $\dot{u} \in L^2(0,T;H_0^1(\Omega))$, $\ddot{u} \in L^2(0,T;H^{-1}(\Omega))$, is a weak solution of the hyperbolic initial-boundary-value problem (4.0.1)–(4.0.3) if*

(a) $\langle \ddot{u}(t), \varphi\rangle + (\partial_t a(u(t)), \varphi)_{L^2} + (\nabla\dot{u}(t), \nabla\varphi)_{L^2} + (\nabla b(u(t)), \nabla\varphi)_{L^2} = 0$ *for all* $\varphi \in H_0^1(\Omega)$ *and a.e. time* $0 \le t \le T$, *and*

(b) $u(x,0) = u_0(x)$, $\dot{u}(x,0) = u_1(x)$ *for* $x \in \overline{\Omega}$.

Here $\langle \cdot, \cdot \rangle$ denotes the dual pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$.

Remark 4.2. *It follows directly from Sobolev embedding in one-dimensional space that $u \in H^1(0,T;H_0^1(\Omega)) \subset C([0,T];H_0^1(\Omega))$. By Theorem 3 in Sect.5.9.2 of Evans [36], we know that $\dot{u} \in L^2(0,T;H_0^1(\Omega))$ with $\ddot{u} \in L^2(0,T;H^{-1}(\Omega))$ implies $\dot{u} \in C([0,T];L^2(\Omega))$. This shows that the initial conditions stated as (b) in the above definition are meaningful.*

4.1.1. Existence of weak solutions We show the existence of weak solutions using an abstract *Galerkin's method* by selecting smooth functions $w_k = w_k(x)$ $(k = 1, \ldots)$ such that

$$\{w_k\}_{k=1}^{\infty} \text{ is an orthogonal basis of } H_0^1(\Omega) \tag{4.1.1}$$

and

$$\{w_k\}_{k=1}^{\infty} \text{ is an orthonormal basis of } L^2(\Omega). \tag{4.1.2}$$

For a positive integer m, we write

$$u_m(t) := \sum_{k=1}^{m} d_m^k(t) w_k, \tag{4.1.3}$$

where we intend to select the coefficients $d_m^k(t)(0 \le t \le T, k = 1, \ldots, m)$ to satisfy

$$d_m^k(0) = (u_0, w_k)_{L^2} \quad (k = 1, \ldots, m), \tag{4.1.4}$$

$$\dot{d}_m^k(0) = (u_1, w_k)_{L^2} \quad (k = 1, \ldots, m), \tag{4.1.5}$$

and

$$(\ddot{u}_m(t), w_k)_{L^2} + (\partial_t a(u_m(t)), w_k)_{L^2} + (\nabla \dot{u}_m(t), \nabla w_k)_{L^2} + (\nabla b(u_m(t)), \nabla w_k)_{L^2} = 0, \tag{4.1.6}$$

for $0 \le t \le T, k = 1, \ldots, m$.

Theorem 4.3 (Existence of approximate solutions)**.** *For each integer $m = 1, 2, \ldots,$ there exists at least one function u_m of the form (4.1.3) satisfying (4.1.4)–(4.1.6).*

Proof. We assume that u_m is given by (4.1.3), then using the fact that $\{w_k\}_{k=1}^{\infty}$ is an orthonormal basis of $L^2(\Omega)$, we have

$$(\ddot{u}_m(t), w_k)_{L^2} = \ddot{d}^k_m(t), \tag{4.1.7}$$

$$(\nabla \dot{u}_m(t), \nabla w_k)_{L^2} = \dot{d}^k_m(t) \left(\nabla w_k, \nabla w_k\right)_{L^2}, \tag{4.1.8}$$

$$\begin{aligned}(\partial_t a(u_m(t)), w_k)_{L^2} &= \left(a'\left(\sum_{k=1}^m d^k_m(t) w_k\right) \sum_{l=1}^m \dot{d}^l_m(t) w_l, w_k \right)_{L^2} \\ &= \sum_{l=1}^m \dot{d}^l_m(t) \left(a'\left(\sum_{k=1}^m d^k_m(t) w_k\right) w_l, w_k \right)_{L^2}, \end{aligned}\tag{4.1.9}$$

$$(\nabla b(u_m), \nabla w_k)_{L^2} = \sum_{l=1}^m d^l_m(t) \left(b'\left(\sum_{k=1}^m d^k_m(t) w_k\right) \nabla w_l, \nabla w_k \right)_{L^2}. \tag{4.1.10}$$

Then we can rewrite (4.1.6) as

$$\ddot{d}^k_m(t) + \|\nabla w_k\|^2_{L^2(\Omega)} \dot{d}^k_m(t) = f^k\left(\mathbf{d}_m(t), \dot{\mathbf{d}}_m(t)\right) \tag{4.1.11}$$

subject to the initial conditions (4.1.4) and (4.1.5). Here f^k is a nonlinear function in terms of $\mathbf{d}_m$ and $\dot{\mathbf{d}}_m$ defined using the relations (4.1.9) and (4.1.10). By *Peano's theorem*, there exists at least one C^2 function $\mathbf{d}_m(t) = (d^1_m(t), \dots, d^m_m(t))$ satisfying the initial conditions (4.1.4), (4.1.5) and solving (4.1.11) for $0 \le t \le T_0$ where $T_0 \in (0, T)$.

Multiplying (4.1.6) by $\dot{d}^k_m(t)$ and summing over $k = 1, \dots, m$ give

$$(\ddot{u}_m, \dot{u}_m)_{L^2} + (\partial_t a(u_m), \dot{u}_m)_{L^2} + (\nabla \dot{u}_m, \nabla \dot{u}_m)_{L^2} + \left(b'(u_m)\nabla u_m, \nabla \dot{u}_m\right)_{L^2} = 0, \tag{4.1.12}$$

for a.e. $0 \le t \le T_0$. Recall that $b(u_m) = u_m + c(u_m)$, so we can write

$$\left(b'(u_m)\nabla u_m, \nabla \dot{u}_m\right)_{L^2} = (\nabla u_m, \nabla \dot{u}_m)_{L^2} + \left(c'(u_m)\nabla u_m, \nabla \dot{u}_m\right)_{L^2}.$$

We note that

$$\begin{aligned}(\ddot{u}_m, \dot{u}_m)_{L^2} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\|\dot{u}_m\|^2_{L^2(\Omega)}\right), \\ (\nabla u_m, \nabla \dot{u}_m)_{L^2} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\|\nabla u_m\|^2_{L^2(\Omega)}\right).\end{aligned}$$

Integrating (4.1.12) from 0 to t, we have

$$\begin{aligned}
&\frac{1}{2}\|\dot{u}_m(t)\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla u_m(t)\|^2_{L^2(\Omega)} + \int_0^t \left(a'(u_m)\dot{u}_m(s), \dot{u}_m(s)\right)_{L^2} \mathrm{d}s + \int_0^t \|\nabla \dot{u}_m(s)\|^2_{L^2(\Omega)} \mathrm{d}s \\
&= \frac{1}{2}\|\dot{u}_m(0)\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla u_m(0)\|^2_{L^2(\Omega)} - \int_0^t \left(c'(u_m(s))\nabla u_m(s), \nabla \dot{u}_m(s)\right)_{L^2} \mathrm{d}s.
\end{aligned}$$

Applying the Cauchy–Schwarz inequality and using the fact that $a'(u_m) \geq M_1$ and $|c'(u_m)| \leq C_c$, we have

$$\begin{aligned}
&\|\dot{u}_m(t)\|^2_{L^2(\Omega)} + \|\nabla u_m(t)\|^2_{L^2(\Omega)} + 2M_1 \int_0^t \|\dot{u}_m(s)\|^2_{L^2(\Omega)} \mathrm{d}s + 2\int_0^t \|\nabla \dot{u}_m(s)\|^2_{L^2(\Omega)} \mathrm{d}s \\
&\leq \|\dot{u}_m(0)\|^2_{L^2(\Omega)} + \|\nabla u_m(0)\|^2_{L^2(\Omega)} + \int_0^t \|\nabla \dot{u}_m(s)\|^2_{L^2(\Omega)} \mathrm{d}s + C_c^2 \int_0^t \|\nabla u_m(s)\|^2_{L^2(\Omega)} \mathrm{d}s.
\end{aligned} \tag{4.1.13}$$

That is,

$$\begin{aligned}
&\|\dot{u}_m(t)\|^2_{L^2(\Omega)} + \|\nabla u_m(t)\|^2_{L^2(\Omega)} + 2M_1 \int_0^t \|\dot{u}_m(s)\|^2_{L^2(\Omega)} \mathrm{d}s + \int_0^t \|\nabla \dot{u}_m\|^2_{L^2(\Omega)} \mathrm{d}s \\
&\leq \|\dot{u}_m(0)\|^2_{L^2(\Omega)} + \|\nabla u_m(0)\|^2_{L^2(\Omega)} + C_c^2 \int_0^t \|\nabla u_m(s)\|^2_{L^2(\Omega)} \mathrm{d}s \\
&\leq \|u_1\|^2_{L^2(\Omega)} + \|u_0\|^2_{H^1_0(\Omega)} + C_c^2 \int_0^t \|\nabla u_m(s)\|^2_{L^2(\Omega)} \mathrm{d}s.
\end{aligned} \tag{4.1.14}$$

Thus, Grönwall's inequality (cf. Lemma 1.5) implies that

$$\|\dot{u}_m(t)\|^2_{L^2(\Omega)} + \|\nabla u_m(t)\|^2_{L^2(\Omega)} + \int_0^t \|\nabla \dot{u}_m\|^2_{L^2(\Omega)} \mathrm{d}s \leq \exp(C_c^2 T_0)\left(\|u_1\|^2_{L^2(\Omega)} + \|u_0\|^2_{H^1_0(\Omega)}\right). \tag{4.1.15}$$

Since $0 \leq t \leq T_0$ is arbitrary, we see from the above estimate that

$$\begin{aligned}
&\max_{0\leq t\leq T_0}\left(\|\dot{u}_m(t)\|^2_{L^2(\Omega)} + \|u_m(t)\|^2_{H^1_0(\Omega)}\right) + \int_0^{T_0} \|\nabla \dot{u}_m\|^2_{L^2(\Omega)} \mathrm{d}s \\
&\leq \exp(C_c^2 T_0)\left(\|u_1\|^2_{L^2(\Omega)} + \|u_0\|^2_{H^1_0(\Omega)}\right).
\end{aligned} \tag{4.1.16}$$

Fix $v \in H^1_0(\Omega)$, $\|v\|_{H^1_0(\Omega)} \leq 1$ and write $v = v^1 + v^2$ where $v^1 \in \operatorname{span}\{w_k\}_{k=1}^m$ (i.e. $v^1 = \sum_{k=1}^m c^k w_k$ for some constants $\{c_k\}_{k=1}^m$) and $\left(v^2, w_k\right)_{L^2} = 0$ for $k = 1, \ldots, m$. Note that there exists a positive constant C_v such that

$$\|v^1\|_{H^1(\Omega)} = \|v^1\|_{L^2(\Omega)} + \|\nabla v^1\|_{L^2(\Omega)} \leq \|v\|_{L^2(\Omega)} + \left(\sum_{k=1}^{m} c_k^2 \|\nabla w_k\|_{L^2(\Omega)}^2\right)^{\frac{1}{2}} \leq C_v,$$

where we have used the orthogonality relation $\left(v^2, w_k\right)_{L^2} = 0$ for $k = 1, \ldots, m$ in the first inequality. Then we have

$$\begin{aligned}
\langle \ddot{u}_m, v\rangle &= (\ddot{u}_m, v)_{L^2} = \left(\ddot{u}_m, v^1\right)_{L^2} \\
&= -\left(a'(u_m)\dot{u}_m, v^1\right)_{L^2} - \left(\nabla \dot{u}_m, \nabla v^1\right)_{L^2} - \left(\nabla u_m, \nabla v^1\right)_{L^2} - \left(c'(u_m)\nabla u_m, \nabla v^1\right)_{L^2} \\
&\leq C_a \|\dot{u}_m\|_{L^2(\Omega)} \|v^1\|_{L^2(\Omega)} + \|\nabla \dot{u}_m\|_{L^2(\Omega)} \|\nabla v^1\|_{L^2(\Omega)} + \|\nabla u_m\|_{L^2(\Omega)} \|\nabla v^1\|_{L^2(\Omega)} \\
&\quad + C_c \|\nabla u_m\|_{L^2(\Omega)} \|\nabla v^1\|_{L^2(\Omega)} \\
&\leq C\left(\|\dot{u}_m\|_{L^2(\Omega)} + \|\nabla \dot{u}_m\|_{L^2(\Omega)} + \|\nabla u_m\|_{L^2(\Omega)}\right), \qquad (4.1.17)
\end{aligned}$$

where we have used the fact that $\|v^1\|_{H^1(\Omega)} \leq C_v$. Here C is a positive constant which may vary from line to line. Consequently

$$\begin{aligned}
\int_0^{T_0} \|\ddot{u}_m\|_{H^{-1}(\Omega)}^2 \,\mathrm{d}t &\leq C \int_0^{T_0} \left(\|\dot{u}_m\|_{L^2(\Omega)}^2 + \|\nabla \dot{u}_m\|_{L^2(\Omega)}^2 + \|\nabla u_m\|_{L^2(\Omega)}^2\right) \mathrm{d}t \\
&\leq C\left(\|u_1\|_{L^2(\Omega)}^2 + \|u_0\|_{H_0^1(\Omega)}^2\right). \qquad (4.1.18)
\end{aligned}$$

Combining (4.1.18) and (4.1.15), we deduce the following energy inequality:

$$\begin{aligned}
&\max_{0\leq t\leq T_0} \left(\|u_m(t)\|_{H_0^1(\Omega)}^2 + \|\dot{u}_m(t)\|_{L^2(\Omega)}^2\right) + \|\ddot{u}_m\|_{L^2(0,T_0;H^{-1}(\Omega))}^2 + \|\dot{u}_m\|_{L^2(0,T_0;H_0^1(\Omega))}^2 \\
&\leq C\left(\|u_1\|_{L^2(\Omega)}^2 + \|u_0\|_{H_0^1(\Omega)}^2\right), \quad \text{for } m = 1, 2, \ldots. \qquad (4.1.19)
\end{aligned}$$

This implies that $u_m(T_0, \cdot) \in H_0^1(\Omega)$ and $\dot{u}_m(T_0, \cdot) \in L^2(\Omega)$. Thus, we can use $t = T_0$ as initial data to extend u_m beyond T_0. By applying the same argument on the intervals $[T_0, 2T_0], [2T_0, 3T_0]$ etc., we can eventually deduce the existence of u_m on the whole time interval $[0, T]$, with the following energy estimate.

$$\max_{0\leq t\leq T} \left(\|u_m(t)\|_{H_0^1(\Omega)}^2 + \|\dot{u}_m(t)\|_{L^2(\Omega)}^2\right) + \|\ddot{u}_m\|_{L^2(0,T;H^{-1}(\Omega))}^2 + \|\dot{u}_m\|_{L^2(0,T;H_0^1(\Omega))}^2$$

$$\leq C\left(\|u_1\|^2_{L^2(\Omega)} + \|u_0\|^2_{H^1_0(\Omega)}\right), \qquad \text{for } m = 1, 2, \ldots. \tag{4.1.20}$$

□

Now we are ready to pass to the limits in our Galerkin approximations. In order to pass to the limits in the nonlinear terms, we need strong convergence results, which arise from the following *Aubin–Lions lemma* [10, 52, 72].

Lemma 4.4 (Aubin–Lions Lemma)**.** *Let X_0, X and X_1 be three Banach spaces with $X_0 \subseteq X \subseteq X_1$. Suppose that X_0 is compactly embedded into X and X is continuously embedded into X_1. For $1 \leq p, q \leq +\infty$, let*

$$W = \{u \in L^p(0, T; X_0) \mid \dot{u} \in L^q(0, T; X_1)\}.$$

(a) *If $p < +\infty$, then the embedding of W into $L^p(0, T; X)$ is compact.*

(b) *If $p = +\infty$ and $q > 1$, then the embedding of W into $C([0, T]; X)$ is compact.*

With this lemma, we can prove the following existence result.

Theorem 4.5. *There exists at least one weak solution to (4.0.1)–(4.0.3).*

Proof. By the energy estimates (4.1.20), we see that the $\{u_m\}_{m=1}^{\infty}$ is bounded in $L^2(0, T; H^1_0(\Omega))$, $\{\dot{u}_m\}_{m=1}^{\infty}$ is bounded in $L^2(0, T; H^1_0(\Omega))$ and $\{\ddot{u}_m\}_{m=1}^{\infty}$ is bounded in $L^2(0, T; H^{-1}(\Omega))$. As a consequence of the *Banach–Alaoglu theorem*, there exist a subsequence $\{u_{ml}\}_{l=1}^{\infty} \subset \{u_m\}_{m=1}^{\infty}$ and $u \in L^2(0, T; H^1_0(\Omega))$, with $\dot{u} \in L^2(0, T; H^1_0(\Omega))$, $\ddot{u} \in L^2(0, T; H^{-1}(\Omega))$ such that

$$\begin{cases} u_{ml} \rightharpoonup u & \text{weakly in } L^2(0, T; H^1_0(\Omega)), \\ \dot{u}_{ml} \rightharpoonup \dot{u} & \text{weakly in } L^2(0, T; H^1_0(\Omega)), \\ \ddot{u}_{ml} \rightharpoonup \ddot{u} & \text{weakly in } L^2(0, T; H^{-1}(\Omega)). \end{cases}$$

Consider

$$W_1 := \{u \in L^\infty(0, T; H^1_0(\Omega)) \mid \dot{u} \in L^\infty(0, T; L^2(\Omega))\}.$$

Applying the *Aubin–Lions lemma* with $X_0 = H_0^1(\Omega)$, $X = X_1 = L^2(\Omega)$ and $p = q = \infty$, we know that W_1 is compactly embedded into $C([0,T];L^2(\Omega))$. According to the energy estimates (4.1.19), we have

$$u_{ml} \to u \text{ in } C([0,T];L^2(\Omega)).$$

Analogously, if we consider

$$W_2 := \{v \in L^\infty(0,T;L^2(\Omega)) \mid \dot{v} \in L^2(0,T;H^{-1}(\Omega))\},$$

W_2 is compactly embedded into $C([0,T];H^{-1}(\Omega))$ by taking $X_0 = L^2(\Omega), X = X_1 = H^{-1}(\Omega)$, $p = \infty$ and $q = 2$ in the *Aubin–Lions lemma.* Again, using the energy estimates but with $v = \dot{u}$, we have

$$\dot{u}_{ml} \to \dot{u} \text{ in } C([0,T];H^{-1}(\Omega)).$$

Since $\{w_k\}_{k=1}^\infty$ is an orthogonal basis in $H_0^1(\Omega)$, (4.1.6) implies that

$$\int_0^T \Big[(\ddot{u}_m(t),\varphi(t))_{L^2} + (\partial_t a(u_m(t)),\varphi(t))_{L^2} + (\nabla \dot{u}_m(t),\nabla\varphi(t))_{L^2} + (\nabla b(u_m(t)),\nabla\varphi(t))_{L^2} \Big] \,\mathrm{d}t = 0, \tag{4.1.21}$$

for all $\varphi \in L^2(0,T;H_0^1(\Omega))$. We know that $H_0^1(0,T;C_0^\infty(\Omega))$ is a dense subset of $L^2(0,T;H_0^1(\Omega))$, thus

$$\int_0^T \Big[(\ddot{u}_m(t),\varphi(t))_{L^2} + (\partial_t a(u_m(t)),\varphi(t))_{L^2} + (\nabla \dot{u}_m(t),\nabla\varphi(t))_{L^2} + (\nabla b(u_m(t)),\nabla\varphi(t))_{L^2} \Big] \,\mathrm{d}t = 0, \tag{4.1.22}$$

for all $\varphi \in H_0^1(0,T;C_0^\infty(\Omega))$. Here $C_0^\infty(\Omega)$ represents the set of smooth functions with compact support in Ω. Integration by parts for the nonlinear terms in (4.1.22) gives

$$\int_0^T \Big[(\ddot{u}_m(t),\varphi(t))_{L^2} - (a(u_m(t)),\dot{\varphi}(t))_{L^2} + (\nabla \dot{u}_m(t),\nabla\varphi(t))_{L^2} - (b(u_m(t)),\Delta\varphi(t))_{L^2} \Big] \,\mathrm{d}t = 0, \tag{4.1.23}$$

for all $\varphi \in H_0^1(0,T;C_0^\infty(\Omega))$. Note that there are no boundary terms resulting from integration by parts because $\varphi \in H_0^1(0,T;C_0^\infty(\Omega))$. Recall that

$$\begin{cases} u_{ml} \rightharpoonup u & \text{weakly in } L^2(0,T;H_0^1(\Omega)), \\ \dot{u}_{ml} \rightharpoonup \dot{u} & \text{weakly in } L^2(0,T;H_0^1(\Omega)), \\ \ddot{u}_{ml} \rightharpoonup \ddot{u} & \text{weakly in } L^2(0,T;H^{-1}(\Omega)). \end{cases}$$

By setting $m = m_l$, we have

$$\int_0^T (\ddot{u}_m(t), \varphi(t))_{L^2} \, \mathrm{d}t \to \int_0^T (\ddot{u}(t), \varphi(t))_{L^2} \, \mathrm{d}t \tag{4.1.24}$$

and

$$\int_0^T (\nabla \dot{u}_m(t), \nabla \varphi(t))_{L^2} \, \mathrm{d}t \to \int_0^T (\nabla \dot{u}(t), \nabla \varphi(t))_{L^2} \, \mathrm{d}t \tag{4.1.25}$$

as $m \to \infty$, for all $\varphi \in H_0^1(0,T;C_0^\infty(\Omega))$. For the two nonlinear terms, we notice that

$$\begin{aligned} &\left| \int_0^T (a(u_m(t)), \dot{\varphi}(t))_{L^2} \, \mathrm{d}t - \int_0^T (a(u(t)), \dot{\varphi}(t))_{L^2} \, \mathrm{d}t \right| \\ &= \left| \int_0^T (a(u_m(t)) - a(u(t)), \dot{\varphi}(t))_{L^2} \, \mathrm{d}t \right| \\ &\leq \int_0^T \|a(u_m(t)) - a(u(t))\|_{L^2(\Omega)} \|\dot{\varphi}(t)\|_{L^2(\Omega)} \, \mathrm{d}t \\ &\leq \int_0^T C_a \|u_m(t) - u(t)\|_{L^2(\Omega)} \|\dot{\varphi}(t)\|_{L^2(\Omega)} \, \mathrm{d}t \\ &\leq C_a \|\dot{\varphi}\|_{L^2(0,T;L^2(\Omega))} \|u_m - u\|_{L^2(0,T;L^2(\Omega))} \to 0 \quad \text{as } m \to \infty, \end{aligned} \tag{4.1.26}$$

since $\|\dot{\varphi}\|_{L^2(0,T;L^2(\Omega))} < \infty$ and $u_m \to u$ in $C([0,T];L^2(\Omega))$. For the second last line of the inequality, we have applied the global Lipschitz continuity of a with C_a being the Lipschitz constant. Analogously, with $\tilde{C}_{\mathrm{Lip}} := 1 + C_c$,

$$\begin{aligned} &\left| \int_0^T (b(u_m(t)), \Delta\varphi(t))_{L^2} \, \mathrm{d}t - \int_0^T (b(u(t)), \Delta\varphi(t))_{L^2} \, \mathrm{d}t \right| \\ &= \left| \int_0^T (b(u_m(t)) - b(u(t)), \Delta\varphi(t))_{L^2} \, \mathrm{d}t \right| \\ &\leq \int_0^T \|b(u_m(t)) - b(u(t))\|_{L^2(\Omega)} \|\Delta\varphi(t)\|_{L^2(\Omega)} \, \mathrm{d}t \\ &\leq \int_0^T \tilde{C}_{\mathrm{Lip}} \|u_m(t) - u(t)\|_{L^2(\Omega)} \|\Delta\varphi(t)\|_{L^2(\Omega)} \, \mathrm{d}t \end{aligned}$$

$$\leq \tilde{C}_{\mathrm{Lip}} \|\Delta\varphi\|_{L^2(0,T;L^2(\Omega))} \|u_m - u\|_{L^2(0,T;L^2(\Omega))} \to 0 \quad \text{as } m \to \infty, \tag{4.1.27}$$

since $\|\Delta\varphi\|_{L^2(0,T;L^2(\Omega))} < \infty$ and $u_m \to u$ in $C([0,T];L^2(\Omega))$. Combining (4.1.24)–(4.1.27) and applying integration by parts in the reverse direction, we obtain

$$\int_0^T \Big[\langle \ddot{u}(t), \varphi(t) \rangle + (\partial_t a(u(t)), \varphi(t))_{L^2} + (\nabla \dot{u}(t), \nabla\varphi(t))_{L^2} + (\nabla b(u(t)), \nabla\varphi(t))_{L^2} \Big] \,\mathrm{d}t = 0, \tag{4.1.28}$$

for all $\varphi \in H_0^1(0,T;C_0^\infty(\Omega))$. The equality (4.1.28) then holds for all functions $\varphi \in L^2(0,T;H_0^1(\Omega))$ since $H_0^1(0,T;C_0^\infty(\Omega))$ is dense in $L^2(0,T;H_0^1(\Omega))$. We consider functions $\varphi(x,t)$ of the factorised form $\varphi(x,t) = \phi(t)\psi(x)$ in (4.1.28); then it follows that

$$\langle \ddot{u}(t), \psi \rangle + (\partial_t a(u(t)), \psi)_{L^2} + (\nabla \dot{u}(t), \nabla\psi)_{L^2} + (\nabla b(u(t)), \nabla\psi)_{L^2} = 0 \tag{4.1.29}$$

for all $\psi \in H_0^1(\Omega)$ and a.e. $t \in [0,T]$. Furthermore, $u \in C([0,T];L^2(\Omega))$ and $\dot{u} \in C([0,T];H^{-1}(\Omega))$.

Next, we need to verify (4.0.2). Now we choose any test function $\varphi \in C^2([0,T];H_0^1(\Omega))$, with $\varphi(T) = \dot{\varphi}(T) = 0$. Then, integrating by parts twice with respect to t in (4.1.28), we have

$$\begin{aligned}&\int_0^T \Big[(u(t), \ddot{\varphi}(t))_{L^2} + (\partial_t a(u(t)), \varphi(t))_{L^2} + (\nabla \dot{u}(t), \nabla\varphi(t))_{L^2} + (\nabla b(u(t)), \nabla\varphi(t))_{L^2} \Big] \,\mathrm{d}t \\ &= -(u(0), \dot{\varphi}(0))_{L^2} + (\dot{u}(0), \varphi(0))_{L^2}. \end{aligned} \tag{4.1.30}$$

Similarly, applying integration by parts to the first term in (4.1.22) twice yields

$$\begin{aligned}&\int_0^T \Big[(u_m(t), \ddot{\varphi}(t))_{L^2} + (\partial_t a(u_m(t)), \varphi(t))_{L^2} + (\nabla \dot{u}_m(t), \nabla\varphi(t))_{L^2} + (\nabla b(u_m(t)), \nabla\varphi(t))_{L^2} \Big] \,\mathrm{d}t \\ &= -(u_m(0), \dot{\varphi}(0))_{L^2} + (\dot{u}_m(0), \varphi(0))_{L^2}. \end{aligned} \tag{4.1.31}$$

By setting $m = m_l$ in (4.1.31) and using (4.1.4), (4.1.5) and the convergence of integrals asserted in (4.1.26) and (4.1.27), we have

$$\int_0^T \Big[(u(t), \ddot{\varphi}(t))_{L^2} + (\partial_t a(u(t)), \varphi(t))_{L^2} + (\nabla \dot{u}(t), \nabla\varphi(t))_{L^2} + (\nabla b(u(t)), \nabla\varphi(t))_{L^2} \Big] \,\mathrm{d}t$$

$$= -\left(u_0, \dot{\varphi}(0)\right)_{L^2} + \left(u_1, \varphi(0)\right)_{L^2}. \tag{4.1.32}$$

Comparing (4.1.30) and (4.1.32), we conclude that $u(0) = u_0$ and $\dot{u}(0) = u_1$ since $\varphi(0)$ and $\dot{\varphi}(0)$ are arbitrary. Therefore, u is a weak solution of (4.0.1)–(4.0.3). □

4.1.2. Uniqueness of the weak solution

Theorem 4.6. *A weak solution of (4.0.1)–(4.0.3) is unique in the class of functions* $u \in C([0,T]; W^{1,\infty}(\Omega)) \cap C^1([0,T]; L^\infty(\Omega))$.

Proof. First note that for $u \in C([0,T]; W^{1,\infty}(\Omega)) \cap C^1([0,T]; L^\infty(\Omega))$, $\max_{t\in[0,T]} \|u(t)\|_{L^\infty(\Omega)}$, $\max_{t\in[0,T]} \|\nabla u(t)\|_{L^\infty(\Omega)}$ and $\max_{t\in[0,T]} \|\dot{u}(t)\|_{L^\infty(\Omega)}$ are all bounded.

Assume that both u and $\tilde{u}$ are the solutions of (4.0.1)–(4.0.3), then

$$\langle \ddot{u}, \varphi \rangle + \left(\partial_t a(u), \varphi\right)_{L^2} + \left(\nabla \dot{u}, \nabla \varphi\right)_{L^2} + \left(\nabla u, \nabla \varphi\right)_{L^2} + \left(\nabla c(u), \nabla \varphi\right)_{L^2} = 0, \tag{4.1.33}$$

and

$$\left\langle \ddot{\tilde{u}}, \varphi \right\rangle + \left(\partial_t a(\tilde{u}), \varphi\right)_{L^2} + \left(\nabla \dot{\tilde{u}}, \nabla \varphi\right)_{L^2} + \left(\nabla \tilde{u}, \nabla \varphi\right)_{L^2} + \left(\nabla c(\tilde{u}), \nabla \varphi\right)_{L^2} = 0, \tag{4.1.34}$$

for all $\varphi \in H_0^1(\Omega)$. We also have $u(0) = \tilde{u}(0) = u_0$ and $\dot{u}(0) = \dot{\tilde{u}}(0)$. Subtracting (4.1.34) from (4.1.33) gives

$$\begin{aligned}
&\left\langle \ddot{u} - \ddot{\tilde{u}}, \varphi \right\rangle + \left(\nabla \dot{u} - \nabla \dot{\tilde{u}}, \nabla \varphi\right)_{L^2} + \left(\nabla u - \nabla \tilde{u}, \nabla \varphi\right)_{L^2} \\
&= \left(\partial_t a(\tilde{u}) - \partial_t a(u), \varphi\right)_{L^2} + \left(\nabla c(\tilde{u}) - \nabla c(u), \nabla \varphi\right)_{L^2} \\
&= \left(a'(\tilde{u})\dot{\tilde{u}} - a'(\tilde{u})\dot{u}, \varphi\right)_{L^2} + \left([a'(\tilde{u}) - a'(u)]\dot{u}, \varphi\right)_{L^2} \\
&\quad + \left(c'(\tilde{u})\nabla \tilde{u} - c'(\tilde{u})\nabla u, \nabla \varphi\right)_{L^2} + \left([c'(\tilde{u}) - c'(u)]\nabla u, \nabla \varphi\right)_{L^2}.
\end{aligned} \tag{4.1.35}$$

Note that

$$|a'(u) - a'(\tilde{u})| \le \max_{\tau\in[0,1]} |a''(\tau u + (1-\tau)\tilde{u})||u - \tilde{u}| \le C_{\mathrm{Lip1}}|u - \tilde{u}|, \tag{4.1.36}$$

with $C_{\mathrm{Lip1}} := \max_\xi |a''(\xi)|$ where $|\xi| \le \max_{t\in[0,T]} \left(\|u(t)\|_{L^\infty(\Omega)} + \|\tilde{u}(t)\|_{L^\infty(\Omega)}\right)$. Analogously,

$$|c'(u) - c'(\tilde{u})| \leq \max_{\tau\in[0,1]} |c''(\tau u + (1-\tau)\tilde{u})||u - \tilde{u}| \leq C_{\text{Lip2}}|u - \tilde{u}|, \tag{4.1.37}$$

with $C_{\text{Lip2}} := \max_\xi |c''(\xi)|$ where $|\xi| \leq \max_{t\in[0,T]} \left(\|u(t)\|_{L^\infty(\Omega)} + \|\tilde{u}(t)\|_{L^\infty(\Omega)}\right)$. Since $a \in C^3(\mathbb{R})$ and $c \in C^4(\mathbb{R})$, both C_{Lip1} and C_{Lip2} are bounded positive constants.

If we write $\Theta = u - \tilde{u}$, take $\varphi = \dot{\Theta}$, and integrate the resulting equality from 0 to t, we have

$$\begin{aligned}
&\int_0^t \left[\left\langle \ddot{\Theta}(s), \dot{\Theta}(s) \right\rangle + \|\nabla\dot{\Theta}(s)\|^2_{L^2(\Omega)} + \left(\nabla\Theta(s), \nabla\dot{\Theta}(s)\right)_{L^2} \right] \mathrm{d}s + \int_0^t \left(a'(\tilde{u})\dot{\Theta}(s), \dot{\Theta}(s)\right)_{L^2} \mathrm{d}s \\
&= -\int_0^t \left(c'(\tilde{u}(s))\nabla\Theta(s), \nabla\dot{\Theta}(s)\right)_{L^2} \mathrm{d}s + \int_0^t \left([c'(\tilde{u}(s)) - c'(u(s))]\nabla u(s), \nabla\dot{\Theta}(s)\right)_{L^2} \mathrm{d}s \\
&\quad + \int_0^t \left([a'(\tilde{u}(s)) - a'(u(s))]\dot{u}, \dot{\Theta}(s)\right)_{L^2} \mathrm{d}s.
\end{aligned} \tag{4.1.38}$$

Using the assumption $a(\cdot) \geq M_1$, the Lipschitz continuity of $a'(\cdot)$ and $c'(\cdot)$ and Young's inequality, we have

$$\begin{aligned}
&\frac{1}{2}\|\dot{\Theta}(t)\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla\Theta(t)\|^2_{L^2(\Omega)} + \int_0^t \|\nabla\dot{\Theta}(s)\|^2_{L^2(\Omega)} \mathrm{d}s + M_1 \int_0^t \|\dot{\Theta}(t)\|^2_{L^2} \mathrm{d}s \\
&\leq \frac{1}{2}\|\dot{\Theta}(0)\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla\Theta(0)\|^2_{L^2(\Omega)} - \int_0^t \left(c'(\tilde{u}(s))\nabla\Theta(s), \nabla\dot{\Theta}(s)\right)_{L^2} \mathrm{d}s \\
&\quad + \int_0^t \left([c'(\tilde{u}(s)) - c'(u(s))]\nabla u(s), \nabla\dot{\Theta}(s)\right)_{L^2} \mathrm{d}s + \int_0^t \left([a'(\tilde{u}(s)) - a'(u(s))]\dot{u}, \dot{\Theta}(s)\right)_{L^2} \mathrm{d}s \\
&\leq C_c \int_0^t \|\nabla\Theta(s)\|_{L^2(\Omega)} \|\nabla\dot{\Theta}(s)\|_{L^2(\Omega)} \mathrm{d}s \\
&\quad + C_{\text{Lip2}} \max_{t\in[0,T]} \|\nabla u(t)\|_{L^\infty(\Omega)} \int_0^t \|\Theta(s)\|_{L^2(\Omega)} \|\nabla\dot{\Theta}(s)\|_{L^2(\Omega)} \mathrm{d}s \\
&\quad + C_{\text{Lip1}} \max_{t\in[0,T]} \|\dot{u}(t)\|_{L^\infty(\Omega)} \int_0^t \|\Theta(s)\|_{L^2(\Omega)} \|\dot{\Theta}(s)\|_{L^2(\Omega)} \mathrm{d}s \\
&\leq M_1 \int_0^t \|\dot{\Theta}(s)\|^2_{L^2(\Omega)} \mathrm{d}s + \frac{C^2_{\text{Lip1}} \max_{t\in[0,T]} \|\dot{u}(t)\|^2_{L^\infty(\Omega)}}{4M_1} \int_0^t \|\Theta(s)\|^2_{L^2(\Omega)} \mathrm{d}s \\
&\quad + \frac{1}{2} \int_0^t \|\nabla\dot{\Theta}(s)\|^2_{L^2(\Omega)} \mathrm{d}s + \frac{C^2_{\text{Lip2}} \max_{t\in[0,T]} \|\nabla u(t)\|^2_{L^\infty(\Omega)}}{2} \int_0^t \|\Theta(s)\|^2_{L^2(\Omega)} \mathrm{d}s \\
&\quad + \frac{1}{2} \int_0^t \|\nabla\dot{\Theta}(s)\|^2_{L^2(\Omega)} \mathrm{d}s + \frac{C_c^2}{2} \int_0^t \|\nabla\Theta(s)\|^2_{L^2(\Omega)} \mathrm{d}s.
\end{aligned} \tag{4.1.39}$$

This implies that

$$\|\dot{\Theta}(t)\|_{L^2(\Omega)}^2 + \|\nabla\Theta(t)\|_{L^2(\Omega)}^2 \leq C \int_0^t \left(\|\nabla\Theta(s)\|_{L^2(\Omega)}^2 + \|\Theta(s)\|_{L^2(\Omega)}^2 \right) \mathrm{d}s, \tag{4.1.40}$$

where C is a positive constant depending on C_{Lip1}, C_{Lip2}, C_c, M_1, $\max_{t\in[0,T]} \|\dot{u}(t)\|_{L^\infty(\Omega)}$ and $\max_{t\in[0,T]} \|\nabla u(t)\|_{L^\infty(\Omega)}$. Applying the Poincaré inequality, we obtain

$$\|\dot{\Theta}(t)\|_{L^2(\Omega)}^2 + \|\Theta(t)\|_{H^1(\Omega)}^2 \leq C_p \int_0^t \|\Theta(s)\|_{H^1(\Omega)}^2 \,\mathrm{d}s, \tag{4.1.41}$$

where C_p depends on both the previous constant C and the Poincaré constant C_{poin}. Finally, by Grönwall's inequality, we achieve

$$\|\dot{\Theta}(t)\|_{L^2(\Omega)}^2 + \|\Theta(t)\|_{H^1(\Omega)}^2 = 0. \tag{4.1.42}$$

This implies that $\Theta(t) = 0$ for arbitrary $0 \leq t \leq T$. Thus $\Theta(t) \equiv 0$ for all $0 \leq t \leq T$, which in turn shows that $u = \tilde{u}$ a.e. $0 \leq t \leq T$. $\square$

4.2. Numerical scheme

4.2.1. Semi-discrete approximation Analogously to Section 3.2, we approximate (4.0.1)–(4.0.3) in space using a continuous Galerkin method first. Here, we define $\mathcal{V}_h$ to be a given family of finite-dimensional subspaces of $W^{1,\infty}(\Omega) \cap H_0^1(\Omega)$ with polynomial degree $p \geq 1$ and assume that the triangulation $\mathcal{T}_h$ of Ω into d-dimensional simplices, which are possibly curved along the boundary $\partial\Omega$, is shape regular and quasi-uniform. If we replace $\mathbf{v}_h$ by $v_h \in \mathcal{V}_h$ in (ii,a) and (ii,b) of Section 3.2, similar inverse inequalities hold. The semi-discrete approximation $u_h \colon [0,T] \to \mathcal{V}_h$ of the solution of (4.0.1)–(4.0.3) satisfies the following initial-value problem in $\mathcal{V}_h$:

$$\langle \ddot{u}_h(t), \varphi\rangle + (\partial_t a(u_h(t)), \varphi)_{L^2} + (\nabla \dot{u}_h(t), \nabla\varphi)_{L^2} + (\nabla b(u_h(t)), \nabla\varphi)_{L^2} = 0 \tag{4.2.1}$$

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

$$u_h(0) = u_{0,h} \in \mathcal{V}_h, \;\; \dot{u}_h(0) = u_{1,h} \in \mathcal{V}_h, \tag{4.2.2}$$

where $u_{0,h}$ and $u_{1,h}$ are specially chosen initial values. Note that the dots over u_h mean $\dot{u}_h = \partial_t u_h$ and $\ddot{u}_h = \partial_{tt} u_h$. We use both notations for differentiation with respect to time t throughout this chapter.

The existence of a local-in-time solution to the system of ODEs (4.2.1) follows from *Caratheodory's theorem*. This local-in-time solution can be extended to the whole of $[0, T]$ based on the energy estimates of the form (4.1.19) by replacing u_m with u_h. The uniqueness of the solution follows from Theorem 4.6 if we replace u by u_h in the proof. Alternatively, we can apply Süli and Wilkins' [74] argument based on *Banach's fixed point theorem* to show the existence and uniqueness of the semi-discrete problem (4.2.1), (4.2.2). We omit the details here.

4.2.2. Discontinuous-in-time fully discrete scheme Similarly to previous chapters, we partition the time interval $I = (0, T]$ into N time slabs $I_n = (t_{n-1}, t_n]$ having length $k_n = t_n - t_{n-1}$ for $n = 1, 2, \ldots, N$, with $t_0 = 0$ and $t_N = T$. We focus on the generic time slab I_n and assume that the solution on I_{n-1} is known. Following the discontinuous-in-time formulation introduced in previous chapters, we first test the equation (4.0.1) against $\dot{\varphi}$ for $\varphi \in H^1(I_n; H_0^1(\Omega))$ and integrate on I_n to obtain the following weak formulation:

$$\begin{aligned}
&\int_{t_{n-1}}^{t_n} \langle \ddot{u}(t), \dot{\varphi}(t) \rangle \, \mathrm{d}t + \int_{t_{n-1}}^{t_n} (\nabla \dot{u}(t), \nabla \dot{\varphi}(t))_{L^2} \, \mathrm{d}t + \int_{t_{n-1}}^{t_n} (\partial_t (a(u(t)), \dot{\varphi}(t))_{L^2} \, \mathrm{d}t \\
&\quad + \int_{t_{n-1}}^{t_n} (\nabla u(t), \nabla \dot{\varphi}(t))_{L^2} \, \mathrm{d}t + \int_{t_{n-1}}^{t_n} (\nabla c(u(t)), \nabla \dot{\varphi}(t))_{L^2} \, \mathrm{d}t = 0.
\end{aligned} \tag{4.2.3}$$

Now we rewrite (4.2.3) by adding suitable (strongly consistent) terms:

$$\begin{aligned}
&\int_{t_{n-1}}^{t_n} \langle \ddot{u}(t), \dot{\varphi}(t) \rangle \, \mathrm{d}t + \big([\dot{u}]_{n-1}, \dot{\varphi}(t_{n-1}^+) \big)_{L^2} + \int_{t_{n-1}}^{t_n} (\nabla \dot{u}(t), \nabla \dot{\varphi}(t))_{L^2} \, \mathrm{d}t \\
&\quad + \int_{t_{n-1}}^{t_n} (\partial_t (a(u(t)), \dot{\varphi}(t))_{L^2} \, \mathrm{d}t + \int_{t_{n-1}}^{t_n} (\nabla u(t), \nabla \dot{\varphi}(t))_{L^2} \, \mathrm{d}t + \big([\nabla u]_{n-1}, \nabla \varphi(t_{n-1}^+) \big)_{L^2} \\
&\quad + \int_{t_{n-1}}^{t_n} (\nabla c(u(t)), \nabla \dot{\varphi}(t))_{L^2} \, \mathrm{d}t = 0.
\end{aligned} \tag{4.2.4}$$

Summing over all time intervals in (4.2.4) leads us to define the following semilinear form $\mathcal{A}\colon \mathcal{H}\times\mathcal{H}\to\mathbb{R}$, with

$$\mathcal{H} := H^2(0,T;H^{-1}(\Omega))\cap H^1(0,T;H_0^1(\Omega))\cap L^2(0,T;H_0^1(\Omega)),$$

by

$$\begin{aligned}
\mathcal{A}(u,\varphi) := &\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\langle \ddot{u}(t),\dot{\varphi}(t)\rangle\,\mathrm{d}t + \sum_{n=1}^{N-1}\big([\dot{u}]_n,\dot{\varphi}(t_n^+)\big)_{L^2} + \big(\dot{u}(t_0^+),\dot{\varphi}(t_0^+)\big)_{L^2}\\
&+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(\nabla\dot{u}(t),\nabla\dot{\varphi}(t))_{L^2}\,\mathrm{d}t + \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(\partial_t(a(u(t)),\dot{\varphi}(t))_{L^2}\,\mathrm{d}t\\
&+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(\nabla u(t),\nabla\dot{\varphi}(t))\,\mathrm{d}t + \sum_{n=1}^{N-1}\big([\nabla u]_n,\nabla\varphi(t_n^+)\big)_{L^2} + \big(\nabla u(t_0^+),\nabla\varphi(t_0^+)\big)_{L^2}\\
&+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(\nabla c(u),\nabla\dot{\varphi}(t))_{L^2}\,\mathrm{d}t
\end{aligned}$$

for $\varphi\in\mathcal{H}$. Again, the semilinear form $\mathcal{A}(\cdot,\cdot)$ works for piecewise in time version of $\mathcal{H}\times\mathcal{H}$ as well. Let, further, F be the linear functional defined by

$$F(\varphi) := \big(u_1,\dot{\varphi}(t_0^+)\big)_{L^2} + \big(\nabla u_0,\nabla\varphi(t_0^+)\big)_{L^2}.$$

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

$$\mathcal{A}(u_{\mathrm{DG}},\varphi) = \tilde{F}(\varphi)\quad\text{ for all }\varphi\in\mathcal{V}_{kh}^{\mathbf{q}}, \tag{4.2.5}$$

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

$$\tilde{F}(\varphi) := \big(u_{1,h},\dot{\varphi}(t_0^+)\big)_{L^2} + \big(\nabla u_{0,h},\nabla\varphi(t_0^+)\big)_{L^2}.$$

Here the finite-dimensional space $\mathcal{V}_{kh}^{\mathbf{q}}$ is defined the same as before.

4.3. Convergence analysis

Similarly to the nonlinear elastodynamics problem (cf. Section 3.3), we shall also prove *a priori* error estimates for the nonlinear damped wave equation.

Theorem 4.7. *Let $u \in W^{s,\infty}(0,T;H^m(\Omega)\cap H_0^1(\Omega))$ be the solution of (4.0.1)–(4.0.3). Assume that $\frac{d}{2}+1 < r \le \min(p, m-1)$, $s \ge q_i+1$, $k_i^{q_i-\frac{1}{2}} = o(h^{\frac{d}{2}})$, and there exist positive constants μ_i, ν_i such that $\mu_i k_i \le h^2 \le \nu_i k_i$ for each $i = 1, 2, \ldots N$. Suppose that we choose the initial data $u_{0,h}$, $u_{1,h} \in \mathcal{V}_h$ to be*

$$u_{0,h} = W(0), \quad u_{1,h} = \dot{W}(0), \tag{4.3.1}$$

where $W(t) \in \mathcal{V}_h$ is the nonlinear elliptic projection of $u(t)$ such that

$$(\nabla c(W(t)), \nabla\varphi)_{L^2} = (\nabla c(u(t)), \nabla\varphi)_{L^2} \qquad \text{for all } \varphi \in \mathcal{V}_h. \tag{4.3.2}$$

Then we have for the solutions of (4.2.5) that

$$\|u_{\mathrm{DG}}(t_j^-) - u(t_j^-)\|_{H^1(\Omega)} + \|\dot{u}_{\mathrm{DG}}(t_j^-) - \dot{u}(t_j^-)\|_{L^2(\Omega)} \le C(u)\left(h^{2r} + \sum_{n=1}^{j} \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)^{\frac{1}{2}} \tag{4.3.3}$$

for each $j = 1, \ldots, N$, where $C(u)$ is a positive constant depending on u.

Remark 4.8. *If we use uniform time intervals $k_n = k = h^2$ (this is to ensure that the assumption $\mu_i k_i \le h^2 \le \nu_i k_i$ for each $i = 1, 2, \ldots N$ is satisfied), and uniform polynomial degrees $q_n = q \ge 2$ for $n = 1, \ldots, N$, then the error bound at the end nodal points becomes*

$$\|u_{\mathrm{DG}}(t_j^-) - u(t_j^-)\|_{H^1(\Omega)} + \|\dot{u}_{\mathrm{DG}}(t_j^-) - \dot{u}(t_j^-)\|_{L^2(\Omega)} \le C(u)\left(k^{\frac{r}{2}} + k^q\right). \tag{4.3.4}$$

That is, if we use DG–2 *elements in the temporal domain and consider $r = 2, 3$ and 4 (the polynomial degree p in the spatial discretisation is at least $2, 3$ or 4 correspondingly), the expected convergence rates are of order* 1.0, 1.5 *and* 2.0 *respectively. For $q = 3$, the expected convergence rates are of order* 1.5, 2.0, 2.5 *and* 3.0 *for $r = 3, 4, 5$ and 6 respectively. Similarly, for $q = 4$, the expected convergence rates are of order* 2.0, 2.5 *and* 3.0 *for $r = 4, 5$ and 6 respectively.*

Remark 4.9. *The assumptions that $k_i^{q_i-\frac{1}{2}} = o(h^{\frac{d}{2}})$ and $\mu_i k_i \le h^2 \le \nu_i k_i$ for each $i = 1, \ldots, N$ require that $q_i > \frac{d}{4} + \frac{1}{2}$ for each $i = 1, \ldots, N$. That is, we need that the polynomial degree in time satisfies $q_i \ge 1$ for $d = 1$ and $q_i \ge 2$ for $d = 2, 3$ on each time* *interval* I_i, *with* $i = 1, \ldots, N$.

It will be assumed throughout the convergence analysis that

$$u \in W^{s,\infty}([0,T]; H^m(\Omega) \cap H_0^1(\Omega))$$

for $m > \frac{d}{2} + 2$ and $s \geq q_i + 1$ for each $i = 1, \ldots, N$.

4.3.1. Definition of the fixed point map We will prove in Section 4.4 that the nonlinear projection W satisfies the following inequalities:

$$\|\nabla u^j(t) - \nabla W^j(t)\|_{L^2(\Omega)} \leq C_r(u) h^r, \quad 0 \leq t \leq T, \text{ for } j = 0, 1, \tag{4.3.5}$$

$$\|u(t) - W(t)\|_{L^2(\Omega)} \leq \tilde{C}_r(u) h^{r+1}, \quad 0 \leq t \leq T, \tag{4.3.6}$$

$$\|u^{(j)}(t) - W^{(j)}(t)\|_{L^2(\Omega)} \leq \tilde{C}_r(u) h^r, \quad 0 \leq t \leq T, \text{ for } j = 1, 2, \tag{4.3.7}$$

$$\|W(t)\|_{W^{1,\infty}(\Omega)} \leq c_0 \text{ and } \|\dot{W}(t)\|_{W^{1,\infty}(\Omega)} \leq c_0, \quad 0 \leq t \leq T, \tag{4.3.8}$$

where $\frac{d}{2} + 1 < r \leq \min(p, m-1)$, $C_r(u), \tilde{C}_r(u)$ are postive constants depending on u, and c_0 is a universal constant independent of h. Analogously to the convergence proof in Section 3.3, for each $W \in H^s(I_n; L^2(\Omega))$, there exist positive constants C_{proj1}, C_{proj2} and C_{proj3} such that,

$$\int_{t_{n-1}}^{t_n} \|\partial_{tt}(W(\cdot,t) - \Pi_k W(\cdot,t))\|_{L^2(\Omega)}^2 \, \mathrm{d}t \leq C_{\text{proj1}} \frac{k_n^{2(\mu-2)}}{q_n^{2(s-3)}} \|W\|_{H^s(I_n, L^2(\Omega))}^2, \tag{4.3.9}$$

$$\int_{t_{n-1}}^{t_n} \|\partial_t(W(\cdot,t) - \Pi_k W(\cdot,t))\|_{L^2(\Omega)}^2 \, \mathrm{d}t \leq C_{\text{proj2}} \frac{k_n^{2(\mu-1)}}{q_n^{2(s-1)}} \|W\|_{H^s(I_n, L^2(\Omega))}^2, \tag{4.3.10}$$

$$\int_{t_{n-1}}^{t_n} \|W(\cdot,t) - \Pi_k W(\cdot,t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t \leq C_{\text{proj3}} \frac{k_n^{2\mu}}{q_n^{2(s-1)}} \|W\|_{H^s(I_n, L^2(\Omega))}^2, \tag{4.3.11}$$

where $\mu = \min(q_n + 1, s)$ and q_n is the polynomial degree with respect to the variable t. For simplicity of notation, we define C_{inv} to be the maximum of the constants appeared in inverse inequalities, and $C_{\text{proj}} = \max \left\{ C_{\text{proj1}}, C_{\text{proj2}}, C_{\text{proj3}} \right\}$.

Again, we need to decompose the error as

$$u_{\mathrm{DG}}(t)-u(t)=(u_{\mathrm{DG}}(t)-\Pi_k W(t))+(\Pi_k W(t)-W(t))+(W(t)-u(t))$$
$$:=\theta(t)+\rho_1(t)+\rho_2(t) \text{ for } t\in I_n, n=1,2,\ldots,N.$$

Similarly to the argument in Section 3.3, we can bound the error terms invovling ρ_1 and ρ_2 by the following:

$$\int_{t_{n-1}}^{t_n}\|\rho_1^{(j)}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t\le C_1(u)\frac{k_n^{2(q_n+1-j)+1}}{q_n^{2(s-1)}}\quad \text{for } j=0,1, \tag{4.3.12}$$

$$\int_{t_{n-1}}^{t_n}\|\nabla\rho_2^{(j)}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t\le C_r(u)k_nh^{2r}\quad \text{for } j=0,1, \tag{4.3.13}$$

$$\int_{t_{n-1}}^{t_n}\|\rho_2^{(j)}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t\le \tilde{C}_r(u)k_nh^{2r}\quad \text{for } j=1,2, \tag{4.3.14}$$

$$\int_{t_{n-1}}^{t_n}\|\rho_2(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t\le \tilde{C}_r(u)k_nh^{2r+2}. \tag{4.3.15}$$

We define $C_2(u)=\max\{C_r(u),\tilde{C}_r(u)\}$. Both $C_1(u)$ and $C_2(u)$ are constants depending on the exact solution u.

Analogously to the convergence analysis for nonlinear elastodynamics problem, we replace

$$\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(\nabla c(u(t)),\nabla\dot{\varphi}(t))_{L^2}\,\mathrm{d}t \quad \text{by} \quad \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(\nabla c(W(t)),\nabla\dot{\varphi}(t))_{L^2}\,\mathrm{d}t$$

in the variational formulation and subtract the resulting equality from the fully discrete scheme to obtain the following equation:

$$\begin{aligned}
&\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\Big\langle\ddot{\theta}(t)+\ddot{\rho}_1(t)+\ddot{\rho}_2(t),\dot{\varphi}(t)\Big\rangle\,\mathrm{d}t+\sum_{n=1}^{N-1}\Big([\dot{\theta}+\dot{\rho}_1+\dot{\rho}_2]_n,\dot{\varphi}(t_n^+)\Big)_{L^2}\\
&\quad+\big(\dot{u}_{\mathrm{DG}}(t_0^+)-\dot{u}(t_0^+),\dot{\varphi}(t_0^+)\big)_{L^2}+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(\nabla\theta(t)+\nabla\rho_1(t)+\nabla\rho_2(t),\nabla\dot{\varphi}(t))_{L^2}\,\mathrm{d}t\\
&\quad+\sum_{n=1}^{N-1}\big([\nabla\theta+\nabla\rho_1+\nabla\rho_2]_n,\nabla\varphi(t_n^+)\big)_{L^2}+\big(\nabla u_{\mathrm{DG}}(t_0^+)-\nabla u(t_0^+),\nabla\varphi(t_0^+)\big)_{L^2}\\
&\quad+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(\partial_t a(u_{\mathrm{DG}}(t))-\partial_t a(u(t)),\dot{\varphi}(t))_{L^2}\,\mathrm{d}t+\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}(\nabla\dot{u}_{\mathrm{DG}}(t)-\nabla\dot{u}(t),\nabla\dot{\varphi}(t))_{L^2}\,\mathrm{d}t
\end{aligned}$$

$$= \big(u_{1,h} - u_1, \dot\varphi(t_0^+)\big)_{L^2} + \big(\nabla u_{0,h} - \nabla u_0, \nabla\varphi(t_0^+)\big)_{L^2}$$
$$- \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\nabla c(u_{\mathrm{DG}}(t)) - \nabla c(W(t)), \nabla\dot\varphi(t))_{L^2}\, \mathrm{d}t, \tag{4.3.16}$$

for $\varphi \in \mathcal{V}_{kh}^{\mathbf{q}}$. To deal with the nonlinearity, we apply Taylor's theorem to have

$$\begin{aligned}
\partial_t a(u_{\mathrm{DG}}(t)) - \partial_t a(u(t)) &= \partial_t \left[a(u_{\mathrm{DG}}(t)) - a(u(t))\right] \\
&= \partial_t \left[(u_{\mathrm{DG}}(t) - u(t)) \int_0^1 a'(u(t) + \tau(u_{\mathrm{DG}}(t) - u(t)))\, \mathrm{d}\tau\right] \\
&= \partial_t \left[(u_{\mathrm{DG}}(t) - u(t)) a'(u_{\mathrm{DG}}(t), u(t))\right] \\
&= \partial_t \left[\theta(t) + \rho_1(t) + \rho_2(t)\right] a'(u_{\mathrm{DG}}(t), u(t)) \\
&\quad + \left[\theta(t) + \rho_1(t) + \rho_2(t)\right] \partial_t a'(u_{\mathrm{DG}}(t), u(t)),
\end{aligned}$$

where

$$a'(u_{\mathrm{DG}}(t), u(t)) := \int_0^1 a'(u(t) + \tau(u_{\mathrm{DG}}(t) - u(t)))\, \mathrm{d}\tau,$$

and

$$\begin{aligned}
&\nabla c(u_{\mathrm{DG}}(t)) - \nabla c(W(t)) \\
&= \nabla \left[(u_{\mathrm{DG}}(t) - W(t)) \int_0^1 c'(W(t) + \tau(u_{\mathrm{DG}}(t) - W(t)))\, \mathrm{d}\tau\right] \\
&= \nabla \left[(u_{\mathrm{DG}}(t) - W(t)) c'(u_{\mathrm{DG}}(t), W(t))\right] \\
&= \nabla \left[\theta(t) + \rho_1(t)\right] c'(u_{\mathrm{DG}}(t), W(t)) + \left[\theta(t) + \rho_1(t)\right] \nabla c'(u_{\mathrm{DG}}(t), W(t)),
\end{aligned}$$

where

$$c'(u_{\mathrm{DG}}(t), W(t)) := \int_0^1 c'(W(t) + \tau(u_{\mathrm{DG}}(t) - W(t)))\, \mathrm{d}\tau.$$

In order to define the fixed point mapping, we first need to consider a fixed point subset $\mathcal{F}$ of $\mathcal{V}_{kh}^{\mathrm{q}}$ as in Section 3.3. We define $\mathcal{F}$ by

$$\begin{aligned}
\mathcal{F} := \Big\{ \psi \in \mathcal{V}_{kh}^{\mathbf{q}} \colon &\|\psi(t_j^-) - \Pi_k W(t_j^-)\|_{H^1(\Omega)}^2 + \|\partial_t(\psi(t_j^-) - \Pi_k W(t_j^-))\|_{L^2(\Omega)}^2 \\
&+ \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\partial_t(\psi(t) - \Pi_k W(t))\|_{L^2(\Omega)}^2\, \mathrm{d}t + \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\partial_t(\nabla\psi(t) - \nabla\Pi_k W(t))\|_{L^2(\Omega)}^2\, \mathrm{d}t
\end{aligned}$$

$$\leq C_*(u)\left(\sum_{n=1}^{j} k_n h^{2r} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right) \text{ for all } j = 1, 2 \ldots, N\Big\}, \tag{4.3.17}$$

where $C_*(u)$ is a positive constant depending on the solution u, 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 $\mathcal{V}_{kh}^{\mathrm{q}}$ in the topology induced by the norm $\|\cdot\|_{\mathcal{F}}$, which is defined by $\|\varphi\|_{\mathcal{F}} = \max_{t\in I_n, 1\leq n\leq N}(\|\varphi(t)\|_{H^1} + \|\dot{\varphi}(t)\|_{L^2})$ for $\varphi \in \mathcal{V}_{kh}^{\mathbf{q}}$. Next, we consider equation (4.3.16) on the integral of $I_n = (t_{n-1}, t_n]$ only, rearrange terms to the right-hand side, and apply integration by parts appropriately. The resulting equation motivates us to define the following fixed point mapping $\mathcal{N}$ on $\mathcal{F}$: if $\phi \in \mathcal{F}$, the image $u_\phi := \mathcal{N}(\phi)$ is given by the relation

$$u_\phi(0) = u_{0,h},\;\; \dot{u}_\phi(0) = u_{1,h}, \tag{4.3.18}$$

$$\begin{aligned}
&\int_{t_{n-1}}^{t_n} \left\langle \ddot{\theta}_\phi(t), \dot{\varphi}(t) \right\rangle \mathrm{d}t + \left([\dot{\theta}_\phi]_{n-1}, \dot{\varphi}_{n-1}^+\right)_{L^2} + \int_{t_{n-1}}^{t_n} (\nabla\theta_\phi(t), \nabla\dot{\varphi}(t))_{L^2}\,\mathrm{d}t \\
&\quad + \left([\nabla\theta_\phi]_{n-1}, \nabla\varphi_{n-1}^+\right)_{L^2} + \int_{t_{n-1}}^{t_n} \left(a'(\phi(t), u(t))\dot{\theta}_\phi(t), \dot{\varphi}(t)\right)_{L^2} \mathrm{d}t + \int_{t_{n-1}}^{t_n} \left(\nabla\dot{\theta}_\phi(t), \nabla\dot{\varphi}(t)\right)_{L^2} \mathrm{d}t \\
&= -\int_{t_{n-1}}^{t_n} \left(a'(\phi(t), u(t))\dot{\rho}_1(t), \dot{\varphi}(t)\right)_{L^2} \mathrm{d}t - \int_{t_{n-1}}^{t_n} \left(a'(\phi(t), u(t))\dot{\rho}_2(t), \dot{\varphi}(t)\right)_{L^2} \mathrm{d}t \\
&\quad - \int_{t_{n-1}}^{t_n} \left(\partial_t a'(\phi(t), u(t))\theta_\phi(t), \dot{\varphi}(t)\right)_{L^2} \mathrm{d}t - \int_{t_{n-1}}^{t_n} \left(\partial_t a'(\phi(t), u(t))\rho_1(t), \dot{\varphi}(t)\right)_{L^2} \mathrm{d}t \\
&\quad - \int_{t_{n-1}}^{t_n} \left(\partial_t a'(\phi(t), u(t))\rho_2(t), \dot{\varphi}(t)\right)_{L^2} \mathrm{d}t - \int_{t_{n-1}}^{t_n} \langle \ddot{\rho}_2(t), \dot{\varphi}(t)\rangle\,\mathrm{d}t - \int_{t_{n-1}}^{t_n} (\nabla\rho_1(t), \nabla\dot{\varphi}(t))_{L^2}\,\mathrm{d}t \\
&\quad - \int_{t_{n-1}}^{t_n} (\nabla\rho_2(t), \nabla\dot{\varphi}(t))_{L^2}\,\mathrm{d}t - \int_{t_{n-1}}^{t_n} (\nabla\dot{\rho}_1(t), \nabla\dot{\varphi}(t))_{L^2}\,\mathrm{d}t - \int_{t_{n-1}}^{t_n} (\nabla\dot{\rho}_2(t), \nabla\dot{\varphi}(t))_{L^2}\,\mathrm{d}t \\
&\quad - \int_{t_{n-1}}^{t_n} \left(c'(\phi(t), W(t))\nabla\theta_\phi(t), \nabla\dot{\varphi}(t)\right)_{L^2} \mathrm{d}t - \int_{t_{n-1}}^{t_n} \left(c'(\phi(t), W(t))\nabla\rho_1(t), \nabla\dot{\varphi}(t)\right)_{L^2} \mathrm{d}t \\
&\quad - \int_{t_{n-1}}^{t_n} \left(\nabla c'(\phi(t), W(t))\theta_\phi(t), \nabla\dot{\varphi}(t)\right)_{L^2} \mathrm{d}t - \int_{t_{n-1}}^{t_n} \left(\nabla c'(\phi(t), W(t))\rho_1(t), \nabla\dot{\varphi}(t)\right)_{L^2} \mathrm{d}t,
\end{aligned} \tag{4.3.19}$$

where $\theta_\phi = u_\phi - \Pi_k W$.

4.3.2. Convergence proof Similarly to Section 3.3, we will show 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 $\phi, \varphi \in \mathcal{F}$,

$$d(\phi,\varphi) := \max_{t\in I_n, 1\le n\le N} \left(\|\phi(t)-\varphi(t)\|_{H^1(\Omega)} + \|\dot\phi(t)-\dot\varphi(t)\|_{L^2(\Omega)} \right).$$

Existence of a fixed point of $\mathcal{N}$ in $\mathcal{F}$

Taking $\varphi = \theta_\phi$ in (4.3.19), we obtain

$$\begin{aligned}
&\|\dot\theta_\phi(t_n^-)\|^2_{L^2(\Omega)} + \|\dot\theta_\phi(t_{n-1}^+)\|^2_{L^2(\Omega)} + \|\nabla\theta_\phi(t_n^-)\|^2_{L^2(\Omega)} + \|\nabla\theta_\phi(t_{n-1}^+)\|^2_{L^2(\Omega)} \\
&\quad + 2\int_{t_{n-1}}^{t_n} \|\nabla\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + 2\int_{t_{n-1}}^{t_n} \left(a'(\phi(t),u(t))\dot\theta_\phi(t),\dot\theta_\phi(t)\right)_{L^2}\,\mathrm{d}t \\
&= 2\left(\dot\theta_\phi(t_{n-1}^-),\dot\theta_\phi(t_{n-1}^+)\right)_{L^2} + 2\left(\nabla\theta_\phi(t_{n-1}^-),\nabla\theta_\phi(t_{n-1}^+)\right)_{L^2} \\
&\quad - 2\int_{t_{n-1}}^{t_n} \left(a'(\phi(t),u(t))\dot\rho_1(t),\dot\theta_\phi(t)\right)_{L^2}\,\mathrm{d}t - 2\int_{t_{n-1}}^{t_n} \left(a'(\phi(t),u(t))\dot\rho_2(t),\dot\theta_\phi(t)\right)_{L^2}\,\mathrm{d}t \\
&\quad - 2\int_{t_{n-1}}^{t_n} \left(\partial_t a'(\phi(t),u(t))\theta_\phi(t),\dot\theta_\phi(t)\right)_{L^2}\,\mathrm{d}t - 2\int_{t_{n-1}}^{t_n} \left(\partial_t a'(\phi(t),u(t))\rho_1(t),\dot\theta_\phi(t)\right)_{L^2}\,\mathrm{d}t \\
&\quad - 2\int_{t_{n-1}}^{t_n} \left(\partial_t a'(\phi(t),u(t))\rho_2(t),\dot\theta_\phi(t)\right)_{L^2}\,\mathrm{d}t - 2\int_{t_{n-1}}^{t_n} \left\langle\ddot\rho_2(t),\dot\theta_\phi(t)\right\rangle\,\mathrm{d}t \\
&\quad - 2\int_{t_{n-1}}^{t_n} \left(\nabla\rho_1(t),\nabla\dot\theta_\phi(t)\right)_{L^2}\,\mathrm{d}t - 2\int_{t_{n-1}}^{t_n} \left(\nabla\rho_2(t),\nabla\dot\theta_\phi(t)\right)_{L^2}\,\mathrm{d}t \\
&\quad - 2\int_{t_{n-1}}^{t_n} \left(\nabla\dot\rho_1(t),\nabla\dot\theta_\phi(t)\right)_{L^2}\,\mathrm{d}t - 2\int_{t_{n-1}}^{t_n} \left(\nabla\dot\rho_2(t),\nabla\dot\theta_\phi(t)\right)_{L^2}\,\mathrm{d}t \\
&\quad - 2\int_{t_{n-1}}^{t_n} \left(c'(\phi(t),W(t))\nabla\theta_\phi(t),\nabla\dot\theta_\phi\right)_{L^2}\,\mathrm{d}t - 2\int_{t_{n-1}}^{t_n} \left(c'(\phi(t),W(t))\nabla\rho_1(t),\nabla\dot\theta_\phi\right)_{L^2}\,\mathrm{d}t \\
&\quad - 2\int_{t_{n-1}}^{t_n} \left(\nabla c'(\phi(t),W(t))\theta_\phi(t),\nabla\dot\theta_\phi\right)_{L^2}\,\mathrm{d}t - 2\int_{t_{n-1}}^{t_n} \left(\nabla c'(\phi(t),W(t))\rho_1(t),\nabla\dot\theta_\phi\right)_{L^2}\,\mathrm{d}t.
\end{aligned} \tag{4.3.20}$$

By the Cauchy–Schwarz inequality, we obtain

$$2\left(\dot\theta_\phi(t_{n-1}^-),\dot\theta_\phi(t_{n-1}^+)\right)_{L^2} \le \|\dot\theta_\phi(t_{n-1}^+)\|^2_{L^2(\Omega)} + \|\dot\theta_\phi(t_{n-1})\|^2_{L^2(\Omega)}, \tag{4.3.21}$$

$$2\left(\nabla\theta_\phi(t_{n-1}^-),\nabla\theta_\phi(t_{n-1}^+)\right)_{L^2}\le\|\nabla\theta_\phi(t_{n-1}^+)\|^2_{L^2(\Omega)}+\|\nabla\theta_\phi(t_{n-1})\|^2_{L^2(\Omega)}.\qquad(4.3.22)$$

By applying (4.3.12), (4.3.14) and Young's inequality, we have

$$\begin{aligned}\left|-2\int_{t_{n-1}}^{t_n}\Big(a'(\phi(t),u(t))\dot\rho_1(t),\dot\theta_\phi(t)\Big)_{L^2}\,\mathrm{d}t\right|&\le\delta\int_{t_{n-1}}^{t_n}\|\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t+\frac{c_1^2}{\delta}\int_{t_{n-1}}^{t_n}\|\dot\rho_1(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\&\le\delta\int_{t_{n-1}}^{t_n}\|\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t+\frac{c_1^2C_1(u)}{\delta}\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}},\end{aligned}\qquad(4.3.23)$$

$$\begin{aligned}\left|-2\int_{t_{n-1}}^{t_n}\Big(a'(\phi(t),u(t))\dot\rho_2(t),\dot\theta_\phi(t)\Big)_{L^2}\,\mathrm{d}t\right|&\le\delta\int_{t_{n-1}}^{t_n}\|\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t+\frac{c_1^2}{\delta}\int_{t_{n-1}}^{t_n}\|\dot\rho_2(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\&\le\delta\int_{t_{n-1}}^{t_n}\|\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t+\frac{c_1^2C_2(u)}{\delta}k_nh^{2r},\end{aligned}\qquad(4.3.24)$$

for some $\delta>0$ that will be specified later. Here

$$c_1:=\max_{\tau\in[0,1]}\max_{t\in I_n,1\le n\le N}\|a'(\tau\phi(t)+(1-\tau)u(t))\|_{L^\infty(\Omega)}.$$

Analogously,

$$\begin{aligned}\left|-2\int_{t_{n-1}}^{t_n}\Big(\partial_t a'(\phi(t),u(t))\rho_1(t),\dot\theta_\phi(t)\Big)_{L^2}\,\mathrm{d}t\right|&\le\delta\int_{t_{n-1}}^{t_n}\|\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t+\frac{c_2^2}{\delta}\int_{t_{n-1}}^{t_n}\|\rho_1(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\&\le\delta\int_{t_{n-1}}^{t_n}\|\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t+\frac{c_2^2C_1(u)}{\delta}\frac{k_n^{2q_n+3}}{q_n^{2(s-1)}},\end{aligned}\qquad(4.3.25)$$

$$\begin{aligned}\left|-2\int_{t_{n-1}}^{t_n}\Big(\partial_t a'(\phi(t),u(t))\rho_2(t),\dot\theta_\phi(t)\Big)_{L^2}\,\mathrm{d}t\right|&\le\delta\int_{t_{n-1}}^{t_n}\|\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t+\frac{c_2^2}{\delta}\int_{t_{n-1}}^{t_n}\|\rho_2(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\&\le\delta\int_{t_{n-1}}^{t_n}\|\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t+\frac{c_2^2C_2(u)}{\delta}k_nh^{2r+2},\end{aligned}\qquad(4.3.26)$$

and

$$\left| -2\int_{t_{n-1}}^{t_n} \Big(\partial_t a'(\phi(t),u(t))\theta_\phi(t),\dot\theta_\phi(t)\Big)_{L^2}\,\mathrm{d}t\right| \le \delta\int_{t_{n-1}}^{t_n}\|\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{c_2^2}{\delta}\int_{t_{n-1}}^{t_n}\|\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t, \tag{4.3.27}$$

where

$$c_2 := \max_{\tau\in[0,1]}\max_{t\in I_n,1\le n\le N}\Big(\|a''(\tau\phi(t)+(1-\tau)u(t))\|_{L^\infty(\Omega)}\|\tau\dot\phi(t)+(1-\tau)\dot u(t)\|_{L^\infty(\Omega)}\Big).$$

Applying Young's inequality again, we have,

$$\begin{aligned}\left| -2\int_{t_{n-1}}^{t_n}\Big\langle \ddot\rho_2(t),\dot\theta_\phi(t)\Big\rangle\,\mathrm{d}t\right| &\le \delta\int_{t_{n-1}}^{t_n}\|\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{1}{\delta}\int_{t_{n-1}}^{t_n}\|\ddot\rho_2(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\ &\le \delta\int_{t_{n-1}}^{t_n}\|\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{C_2(u)}{\delta}k_nh^{2r},\end{aligned} \tag{4.3.28}$$

$$\begin{aligned}\left| -2\int_{t_{n-1}}^{t_n}\Big(\nabla\rho_1(t),\nabla\dot\theta_\phi(t)\Big)\,\mathrm{d}t\right| &\le \delta'\int_{t_{n-1}}^{t_n}\|\nabla\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{1}{\delta'}\int_{t_{n-1}}^{t_n}\|\nabla\rho_1(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\ &\le \delta'\int_{t_{n-1}}^{t_n}\|\nabla\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{C_1(u)}{\delta'}\frac{k_n^{2q_n+3}}{q_n^{2(s-1)}},\end{aligned} \tag{4.3.29}$$

$$\begin{aligned}\left| -2\int_{t_{n-1}}^{t_n}\Big(\nabla\rho_2(t),\nabla\dot\theta_\phi(t)\Big)\,\mathrm{d}t\right| &\le \delta'\int_{t_{n-1}}^{t_n}\|\nabla\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{1}{\delta'}\int_{t_{n-1}}^{t_n}\|\nabla\rho_2(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\ &\le \delta'\int_{t_{n-1}}^{t_n}\|\nabla\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{C_2(u)}{\delta'}k_nh^{2r},\end{aligned} \tag{4.3.30}$$

$$\begin{aligned}\left| -2\int_{t_{n-1}}^{t_n}\Big(\nabla\dot\rho_1(t),\nabla\dot\theta_\phi(t)\Big)\,\mathrm{d}t\right| &\le \delta'\int_{t_{n-1}}^{t_n}\|\nabla\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{1}{\delta'}\int_{t_{n-1}}^{t_n}\|\nabla\dot\rho_1(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\ &\le \delta'\int_{t_{n-1}}^{t_n}\|\nabla\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{C_1(u)}{\delta'}\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}},\end{aligned} \tag{4.3.31}$$

and

$$\begin{aligned}\left| -2\int_{t_{n-1}}^{t_n}\Big(\nabla\dot\rho_2(t),\nabla\dot\theta_\phi(t)\Big)\,\mathrm{d}t\right| &\le \delta'\int_{t_{n-1}}^{t_n}\|\nabla\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{1}{\delta'}\int_{t_{n-1}}^{t_n}\|\nabla\dot\rho_2(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t\\ &\le \delta'\int_{t_{n-1}}^{t_n}\|\nabla\dot\theta_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \frac{C_2(u)}{\delta'}k_nh^{2r},\end{aligned} \tag{4.3.32}$$

for some $\delta' > 0$ that will be specified later. Similarly, we have

$$\begin{aligned}
&\left| - 2 \int_{t_{n-1}}^{t_n} \left(c'(\phi(t), W(t)) \nabla \theta_\phi(t), \nabla \dot{\theta}_\phi(t) \right)_{L^2} \, \mathrm{d}t \right| \\
&\leq \delta' \int_{t_{n-1}}^{t_n} \|\nabla \dot{\theta}_\phi(t)\|^2_{L^2(\Omega)} \, \mathrm{d}t + \frac{c_3^2}{\delta'} \int_{t_{n-1}}^{t_n} \|\nabla \theta_\phi(t)\|^2_{L^2(\Omega)} \, \mathrm{d}t,
\end{aligned} \tag{4.3.33}$$

$$\begin{aligned}
&\left| - 2 \int_{t_{n-1}}^{t_n} \left(c'(\phi(t), W(t)) \nabla \rho_1(t), \nabla \dot{\theta}_\phi(t) \right)_{L^2} \, \mathrm{d}t \right| \\
&\leq \delta' \int_{t_{n-1}}^{t_n} \|\nabla \dot{\theta}_\phi(t)\|^2_{L^2(\Omega)} \, \mathrm{d}t + \frac{c_3^2}{\delta'} \int_{t_{n-1}}^{t_n} \|\nabla \rho_1(t)\|^2_{L^2(\Omega)} \, \mathrm{d}t \\
&\leq \delta' \int_{t_{n-1}}^{t_n} \|\nabla \dot{\theta}_\phi(t)\|^2_{L^2(\Omega)} \, \mathrm{d}t + \frac{c_3^2 C_1(u)}{\delta'} \frac{k_n^{2q_n+3}}{q_n^{2(s-1)}},
\end{aligned} \tag{4.3.34}$$

and

$$\begin{aligned}
&\left| - 2 \int_{t_{n-1}}^{t_n} \left(\nabla c'(\phi(t), W(t)) \theta_\phi(t), \nabla \dot{\theta}_\phi(t) \right)_{L^2} \, \mathrm{d}t \right| \\
&\leq \delta' \int_{t_{n-1}}^{t_n} \|\nabla \dot{\theta}_\phi(t)\|^2_{L^2(\Omega)} \, \mathrm{d}t + \frac{c_4^2}{\delta'} \int_{t_{n-1}}^{t_n} \|\theta_\phi(t)\|^2_{L^2(\Omega)} \, \mathrm{d}t,
\end{aligned} \tag{4.3.35}$$

$$\begin{aligned}
&\left| - 2 \int_{t_{n-1}}^{t_n} \left(\nabla c'(\phi(t), W(t)) \rho_1(t), \nabla \dot{\theta}_\phi(t) \right)_{L^2} \, \mathrm{d}t \right| \\
&\leq \delta' \int_{t_{n-1}}^{t_n} \|\nabla \dot{\theta}_\phi(t)\|^2_{L^2(\Omega)} \, \mathrm{d}t + \frac{c_4^2}{\delta'} \int_{t_{n-1}}^{t_n} \|\rho_1(t)\|^2_{L^2(\Omega)} \, \mathrm{d}t \\
&\leq \delta' \int_{t_{n-1}}^{t_n} \|\nabla \dot{\theta}_\phi(t)\|^2_{L^2(\Omega)} \, \mathrm{d}t + \frac{c_4^2 C_1(u)}{\delta'} \frac{k_n^{2q_n+3}}{q_n^{2(s-1)}},
\end{aligned} \tag{4.3.36}$$

where

$$c_3 := \max_{\tau \in [0,1]} \max_{t \in I_n, 1 \leq n \leq N} \|c'(\tau \phi(t) + (1 - \tau) W(t))\|_{L^\infty(\Omega)}$$

and

$$c_4 := \max_{\tau \in [0,1]} \max_{t \in I_n, 1 \leq n \leq N} \left(\|c''(\tau \phi(t) + (1 - \tau) W(t))\|_{L^\infty(\Omega)} \|\tau \nabla \phi(t) + (1 - \tau) \nabla W(t)\|_{L^\infty(\Omega)} \right).$$

Before we proceed further, we estimate the values of the constants c_1–c_4; this contributes to the restriction on h, r and k_n for each $n = 1, \ldots, N$. First note that $c_1 \leq C_a$ and

$c_3 \leq C_c$. For each $t \in I_n, n = 1, \ldots, N$,

$$\begin{aligned}
&\|\tau\phi(t) + (1-\tau)u(t)\|_{L^\infty(\Omega)} \leq \|\phi(t) - u(t)\|_{L^\infty(\Omega)} + \|u(t)\|_{L^\infty(\Omega)} \\
&\leq \|\phi(t) - \Pi_k W(t)\|_{L^\infty(\Omega)} + \|\Pi_k W(t) - W(t)\|_{L^\infty(\Omega)} + \|W(t) - u(t)\|_{L^\infty(\Omega)} + \|u(t)\|_{L^\infty(\Omega)} \\
&\leq C_{\text{inv}} h^{-\frac{d}{2}} \Big(\|\phi(t) - \Pi_k W(t)\|_{L^2(\Omega)} + \|\Pi_k W(t) - W(t)\|_{L^2(\Omega)} \Big) \\
&\quad + \|W(t) - u(t)\|_{L^\infty(\Omega)} + \|u(t)\|_{L^\infty(\Omega)},
\end{aligned}$$

where we have applied the inverse inequality (i,a) to to the first two terms of the above inequality. Now we bound the first L^2-norm term by

$$\begin{aligned}
\|\phi(t) - \Pi_k W(t)\|_{L^2(\Omega)} &\leq \|\phi(t_n^-) - \Pi_k W(t_n^-)\|_{L^2(\Omega)} + \int_t^{t_n} \|\partial_s(\phi(s) - \Pi_k W(s))\|_{L^2(\Omega)} \,\mathrm{d}t \\
&\leq \|\phi(t_n^-) - \Pi_k W(t_n^-)\|_{L^2(\Omega)} + \sqrt{k_n} \left(\int_{t_{n-1}}^{t_n} \|\partial_t(\phi(t) - \Pi_k W(t))\|^2_{L^2(\Omega)} \,\mathrm{d}t \right)^{\frac{1}{2}} \\
&\leq C_*(u) \left(\sum_{i=1}^n k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}},
\end{aligned}$$

since $\sqrt{k_n} < 1$. Here $C_*(u)$ is the constant which appeared in the definition of the fixed point subset $\mathcal{F}$. For the second L^2-norm term, we have

$$\begin{aligned}
\|\Pi_k W(t) - W(t)\|_{L^2(\Omega)} &\leq \|\Pi_k W(t_n^-) - W(t_n^-)\|_{L^2(\Omega)} + \int_t^{t_n} \|\partial_s(\Pi_k W(s) - W(s))\|_{L^2(\Omega)} \,\mathrm{d}s \\
&\leq \int_t^{t_n} \|\partial_s(\Pi_k W(s) - W(s))\|_{L^2(\Omega)} \,\mathrm{d}s \quad (\text{since } \Pi_k W(t_n^-) = W(t_n^-)) \\
&\leq \sqrt{k_n} \left(\int_{t_{n-1}}^{t_n} \|\partial_t(\Pi_k W(t) - W(t))\|^2_{L^2(\Omega)} \,\mathrm{d}t \right)^{\frac{1}{2}} \\
&\leq \sqrt{k_n} \left(C_{\text{proj}} \frac{k_n^{2q_n}}{q_n^{2(s-1)}} \|W\|^2_{H^s(I_n;L^2(\Omega))} \right)^{\frac{1}{2}} \\
&\leq \sqrt{C_1(u)} \frac{k_n^{q_n+\frac{1}{2}}}{q_n^{s-1}}.
\end{aligned}$$

Note that

$$\|W(t) - u(t)\|_{L^\infty(\Omega)} \leq \|W(t)\|_{L^\infty(\Omega)} + \|u(t)\|_{L^\infty(\Omega)} \leq c_0 + \|u(t)\|_{L^\infty(\Omega)}.$$

Thus, for each $t \in I_n, n = 1, \dots, N$,

$$
\begin{aligned}
&\|\tau\phi(t) + (1-\tau)u(t)\|_{L^\infty(\Omega)} \\
&\leq 2\|u(t)\|_{L^\infty(\Omega)} + c_0 + C_{\text{inv}} h^{-\frac{d}{2}} \left[C_*(u) \left(\sum_{i=1}^{n} k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} + \sqrt{C_1(u)} \frac{k_n^{q_n+\frac{1}{2}}}{q_n^{s-1}} \right] \\
&\leq 2\|u(t)\|_{L^\infty(\Omega)} + \tilde{c}_1,
\end{aligned}
$$

for some positive constant $\tilde{c}_1$ since we have assumed that $r > \frac{d}{2} + 1$ (thus $r > \frac{d}{2}$) and $k_i^{q_i - \frac{1}{2}} = o(h^{\frac{d}{2}})$ (which implies that $k_i^{q_i} = o(h^{\frac{d}{2}})$ for each $i = 1, \dots, N$). Analogously, we have, for each $t \in I_n, 1 \leq n \leq N$,

$$
\begin{aligned}
&\|\tau\dot{\phi}(t) + (1-\tau)\dot{u}(t)\|_{L^\infty(\Omega)} \\
&\leq \|\partial_t\phi(t) - \partial_t\Pi_k W(t)\|_{L^\infty(\Omega)} + \|\partial_t\Pi_k W(t) - \partial_t W(t)\|_{L^\infty(\Omega)} + \|\partial_t u(t) - \partial_t W(t)\|_{L^\infty(\Omega)} \\
&\quad + \|\partial_t u(t)\|_{L^\infty(\Omega)} \\
&\leq C_{\text{inv}} h^{-\frac{d}{2}} \Big(\|\partial_t\phi(t) - \partial_t\Pi_k W(t)\|_{L^2(\Omega)} + \|\partial_t\Pi_k W(t) - \partial_t W(t)\|_{L^2(\Omega)} \Big) \\
&\quad + \|\partial_t u(t) - \partial_t W(t)\|_{L^\infty(\Omega)} + \|\partial_t u(t)\|_{L^\infty(\Omega)}.
\end{aligned}
$$

Note that

$$
\|\partial_t\phi(t) - \partial_t\Pi_k W(t)\|_{L^2(\Omega)} \leq \|\partial_t(\phi(t_n^-) - \Pi_k W(t_n^-))\|_{L^2(\Omega)} + \int_{t_{n-1}}^{t_n} \|\partial_{tt}(\phi(t) - \Pi_k W(t))\|_{L^2(\Omega)}\,\mathrm{d}t.
$$

Since $\phi \in \mathcal{F}$, we have

$$
\|\partial_t(\phi(t_n^-) - \Pi_k W(t_n^-))\|_{L^2(\Omega)} \leq C_*(u) \left(\sum_{i=1}^{n} k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right).
$$

Using the inverse inequality, we obtain

$$
\begin{aligned}
&\int_{t_{n-1}}^{t_n} \|\partial_{tt}(\phi(t) - \Pi_k W(t))\|_{L_2(\Omega)}\,\mathrm{d}t \leq \sqrt{k_n} \left(\int_{t_{n-1}}^{t_n} \|\partial_{tt}(\phi(t) - \Pi_k W(t))\|_{L^2(\Omega)}^2\,\mathrm{d}t \right)^{\frac{1}{2}} \\
&\leq C_{\text{inv}} \frac{1}{\sqrt{k_n}} \left(\int_{t_{n-1}}^{t_n} \|\partial_t(\phi(t) - \Pi_k W(t))\|_{L^2(\Omega)}^2\,\mathrm{d}t \right)^{\frac{1}{2}} \leq C_{\text{inv}} C_*(u) \frac{1}{\sqrt{k_n}} \left(\sum_{i=1}^{n} k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}}.
\end{aligned}
$$

On the other hand, we have

$$
\begin{aligned}
&\|\partial_t \Pi_k W(t) - \partial_t W(t)\|_{L^2(\Omega)} \\
&\leq \|\partial_t \Pi_k W(t_n^-) - \partial_t W(t_n^-)\|_{L^2(\Omega)} + \int_t^{t_n} \|\partial_{ss}(\Pi_k W(s) - W(s))\|_{L^2(\Omega)}\,\mathrm{d}s \\
&\leq \int_t^{t_n} \|\partial_{ss}(\Pi_k W(s) - W(s))\|_{L^2(\Omega)}\,\mathrm{d}s \text{ (since } \partial_t \Pi_k W(t_n^-) = \partial_t W(t_n^-)) \\
&\leq \sqrt{k_n}\left(\int_{t_{n-1}}^{t_n} \|\partial_{tt}(\Pi_k W(t) - W(t))\|^2_{L^2(\Omega)}\,\mathrm{d}t\right)^{\frac{1}{2}} \\
&\leq \sqrt{k_n}\left(C_{\mathrm{proj}}\frac{k_n^{2(q_n-1)}}{q_n^{2(s-3)}}\|W\|^2_{H^s(I_n;L^2(\Omega))}\right)^{\frac{1}{2}} \\
&\leq \sqrt{C_1(u)}\frac{k_n^{q_n-\frac{1}{2}}}{q_n^{s-3}}.
\end{aligned}
$$

Note also that

$$
\|\partial_t W(t) - \partial_t u(t)\|_{L^\infty(\Omega)} \leq \|\partial_t W(t)\|_{L^\infty(\Omega)} + \|\partial_t u(t)\|_{L^\infty(\Omega)} \leq\ c_0 + \|\partial_t u(t)\|_{L^\infty(\Omega)}.
$$

Thus

$$
\begin{aligned}
&\max_{t\in I_n, 1\leq n\leq N} \|\tau\dot{\phi}(t) + (1-\tau)\dot{u}(t)\|_{L^\infty(\Omega)} \\
&\leq C_{\mathrm{inv}} h^{-\frac{d}{2}}\left[C_*(u)(1+\frac{C_{\mathrm{inv}}}{\sqrt{k_n}})\left(\sum_{i=1}^n k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}}\right)^{\frac{1}{2}} + \sqrt{C_1(u)}\frac{k_n^{q_n-\frac{1}{2}}}{q_n^{s-3}}\right] + c_0 + 2\|\partial_t u(t)\|_{L^\infty(\Omega)} \\
&\leq 2\max_{t\in I_n, 1\leq n\leq N} \|\partial_t u(t)\|_{L^\infty(\Omega)} + \tilde{c}_2,
\end{aligned}
$$

for some positive bounded constant $\tilde{c}_2$ since $r > \frac{d}{2}+1$, $k_i^{q_i-\frac{1}{2}} = o(h^{\frac{d}{2}})$, $h^2 \leq \nu_i k_i$ for each $i = 1, \ldots, N$. With this estimate, we can bound c_2 by

$$
\begin{aligned}
c_2 &:= \max_{\tau\in[0,1]} \max_{t\in I_n, 1\leq n\leq N} \|a''(\tau\phi(t) + (1-\tau)u(t))\|_{L^\infty(\Omega)} \|\tau\dot{\phi}(t) + (1-\tau)\dot{u}(t)\|_{L^\infty(\Omega)} \\
&\leq \max_{\xi} |a''(\xi)| \left(2\max_{t\in I_n, 1\leq n\leq N} \|\partial_t u(t)\|_{L^\infty(\Omega)} + \tilde{c}_2\right),
\end{aligned}
$$

where $|\xi| \leq \tilde{c}_1 + 2\max_{t\in I_n, 1\leq n\leq N} \|u(t)\|_{L^\infty(\Omega)}$. Similarly, we have, for each $t \in I_n, 1 \leq n \leq N$,

$$\begin{aligned}
&\|\tau\phi(t) + (1-\tau)W(t)\|_{L^\infty(\Omega)} \\
&\leq \|W(t)\|_{L^\infty(\Omega)} + \|\phi(t) - W(t)\|_{L^\infty(\Omega)} \\
&\leq \|W(t)\|_{L^\infty(\Omega)} + \|\phi(t) - \Pi_k W(t)\|_{L^\infty(\Omega)} + \|\Pi_k W(t) - W(t)\|_{L^\infty(\Omega)} \\
&\leq c_0 + C_{\mathrm{inv}} h^{-\frac{d}{2}} \Big(\|\phi(t) - \Pi_k W(t)\|_{L^2(\Omega)} + \|\Pi_k W(t) - W(t)\|_{L^2(\Omega)} \Big),
\end{aligned}$$

where we have used (ii,c) and the inverse inequality (i,a). Thus

$$\begin{aligned}
&\|\tau\phi(t) + (1-\tau)W(t)\|_{L^\infty(\Omega)} \\
&\leq c_0 + C_{\mathrm{inv}} h^{-\frac{d}{2}} \left(C_*(u) \left(\sum_{i=1}^{n} k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} + \sqrt{C_1(u)} \frac{k_n^{q_n+\frac{1}{2}}}{q_n^{s-1}} \right) \\
&\leq \tilde{c}_3,
\end{aligned}$$

for some bounded positive constant $\tilde{c}_3$. For the $W^{1,\infty}$-term which appeared in c_4, we have

$$\|\tau\nabla\phi(t) + (1-\tau)\nabla W(t)\|_{L^\infty(\Omega)} \leq \|\nabla W(t)\|_{L^\infty(\Omega)} + \|\nabla\phi(t) - \nabla W(t)\|_{L^\infty(\Omega)}.$$

Applying the inverse inequality gives

$$\|\nabla\phi(t) - \nabla W(t)\|_{L^\infty(\Omega)} \leq C_{\mathrm{inv}} h^{-\frac{d}{2}} \|\nabla\phi(t) - \nabla W(t)\|_{L^2(\Omega)}.$$

Using the triangle inequality, we can write

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

Note that for each $t \in I_n$, $n = 1, \ldots, N$,

$$\begin{aligned}
&\|\nabla\phi(t) - \nabla\Pi_k W(t)\|_{L^2(\Omega)} \\
&\leq \|\nabla\phi(t_n^-) - \nabla\Pi_k W(t_n^-)\|_{L^2(\Omega)} + \int_t^{t_n} \|\partial_s(\nabla\phi(s) - \nabla\Pi_k W(s))\|_{L^2(\Omega)} \,\mathrm{d}t \\
&\leq \|\nabla\phi(t_n^-) - \nabla\Pi_k W(t_n^-)\|_{L^2(\Omega)} + C_{\mathrm{inv}} h^{-1}\sqrt{k_n} \left(\int_{t_{n-1}}^{t_n} \|\partial_t(\phi(t) - \Pi_k W(t))\|_{L^2(\Omega)}^2 \,\mathrm{d}t \right)^{\frac{1}{2}}
\end{aligned}$$

$$\leq \tilde{C}_*(u) \left(\sum_{i=1}^{n} k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}},$$

where we have used the inverse inequality, Hölder's inequality, the fact that $\phi \in \mathcal{F}$ and the assumption that $\mu_i k_i \leq h^2$ for each $i = 1, \ldots, N$. On the other hand, we have

$$\begin{aligned}
&\|\nabla(\Pi_k W(t) - W(t))\|_{L^2(\Omega)} \\
&\leq \|\nabla(\Pi_k W(t_n^-) - W(t_n^-))\|_{L^2(\Omega)} + \int_t^{t_n} \|\partial_s(\nabla \Pi_k W(s) - \nabla W(s))\|_{L^2(\Omega)} \, \mathrm{d}s \\
&\leq \int_t^{t_n} \|\partial_s(\nabla \Pi_k W(s) - \nabla W(s))\|_{L^2(\Omega)} \, \mathrm{d}s \quad (\text{ since } \Pi_k W(t_n^-) = W(t_n^-)) \\
&\leq \sqrt{k_n} \left(\int_{t_{n-1}}^{t_n} \|\partial_t(\nabla \Pi_k W(t) - \nabla W(t))\|^2_{L^2(\Omega)} \, \mathrm{d}t \right)^{\frac{1}{2}} \\
&\leq \sqrt{k_n} \left(C_{\mathrm{proj}} \frac{k_n^{2q_n}}{q_n^{2(s-1)}} \|W\|^2_{H^s(I_n; H_0^1(\Omega))} \right)^{\frac{1}{2}} \\
&\leq \sqrt{C_1(u)} \frac{k_n^{q_n + \frac{1}{2}}}{q_n^{s-1}}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\max_{t \in I_n, 1 \leq n \leq N} \|\nabla(\phi(t) - W(t))\|_{L^2(\Omega)} &\leq C_*(u) \left(\sum_{i=1}^{n} k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} + \sqrt{C_1(u)} \frac{k_n^{q_n + \frac{1}{2}}}{q_n^{s-1}} \\
&\leq \tilde{C}_*(u) \left(\sum_{i=1}^{n} k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}}.
\end{aligned}$$

Then we can bound the $W^{1,\infty}$-norm term by

$$\begin{aligned}
&\max_{t \in I_n, 1 \leq n \leq N} \|\tau \nabla \phi(t) + (1-\tau) \nabla W(t)\|_{L^\infty(\Omega)} \\
&\leq \max_{t \in I_n, 1 \leq n \leq N} \left(\|\nabla W(t)\|_{L^\infty(\Omega)} + \|\nabla \phi(t) - \nabla W(t)\|_{L^\infty(\Omega)} \right) \\
&\leq c_0 + C_{\mathrm{inv}} h^{-\frac{d}{2}} \tilde{C}_*(u) \left(\sum_{i=1}^{n} k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} \leq \tilde{c}_4,
\end{aligned}$$

where $\tilde{c}_4$ is a bounded positive constant. Now we can bound c_4 by

$$c_4 := \max_{\tau \in [0,1]} \max_{t \in I_n, 1 \leq n \leq N} \left(\|c''(\tau \phi(t) + (1-\tau) W(t))\|_{L^\infty(\Omega)} \|\tau \nabla \phi(t) + (1-\tau) \nabla W(t)\|_{L^\infty(\Omega)} \right)$$

$$\leq |c''(\xi)|\tilde{c}_4,$$

where $|\xi| \leq \tilde{c}_3$. Now combing the above estimates (4.3.21)–(4.3.36), we have

$$\begin{aligned}
&\|\dot{\theta}_\phi(t_n^-)\|_{L^2(\Omega)}^2 + \|\nabla\theta_\phi(t_n^-)\|_{L^2(\Omega)}^2 + 2\int_{t_{n-1}}^{t_n} \|\nabla\dot{\theta}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + 2\int_{t_{n-1}}^{t_n} \Big(a'(\phi(t), u(t)), |\dot{\theta}_\phi(t)|^2\Big)_{L^2} \,\mathrm{d}t \\
&\leq \|\dot{\theta}_\phi(t_{n-1}^-)\|_{L^2(\Omega)}^2 + \|\nabla\theta_\phi(t_{n-1}^-)\|_{L^2(\Omega)}^2 + 6\delta\int_{t_{n-1}}^{t_n} \|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + 8\delta'\int_{t_{n-1}}^{t_n} \|\nabla\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\quad + \left(\frac{c_2^2}{\delta} + \frac{c_4^2}{\delta'}\right)\int_{t_{n-1}}^{t_n} \|\theta_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + \frac{c_3^2}{\delta'}\int_{t_{n-1}}^{t_n} \|\nabla\theta_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + C_{1,*}\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} + C_{2,*}k_n h^{2r},
\end{aligned} \tag{4.3.37}$$

where

$$\begin{aligned}
C_{1,*} &:= \frac{c_1^2 C_1(u)}{\delta} + \frac{c_2^2 C_1(u)}{\delta} + \frac{2C_1(u)}{\delta'} + \frac{c_3^2 C_1(u)}{\delta'} + \frac{c_4^2 C_1(u)}{\delta'}, \\
C_{2,*} &:= \frac{c_1^2 C_2(u)}{\delta} + \frac{c_2^2 C_2(u)}{\delta} + \frac{C_2(u)}{\delta} + \frac{2C_2(u)}{\delta'}.
\end{aligned}$$

Summing up over $n = 1, 2, \dots j$, we obtain

$$\begin{aligned}
&\|\dot{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \|\nabla\theta_\phi(t_j^-)\|_{L^2(\Omega)}^2 + 2\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \|\nabla\dot{\theta}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\quad + 2\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \Big(a'(\phi(t), u(t)), |\dot{\theta}_\phi(t)|^2\Big)_{L^2} \,\mathrm{d}t \\
&\leq C_{1,*}\sum_{n=1}^{j}\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} + C_{2,*}\sum_{n=1}^{j} k_n h^{2r} + 6\delta\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\quad + \left(\frac{c_2^2}{\delta} + \frac{c_4^2}{\delta'}\right)\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \|\theta_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\quad + 8\delta'\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \|\nabla\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + \frac{c_3^2}{\delta'}\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \|\nabla\theta_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t.
\end{aligned} \tag{4.3.38}$$

Using the assumption that $a'(\cdot) \geq M_1 > 0$, we obtain

$$\begin{aligned}
&\|\dot{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \|\nabla\theta_\phi(t_j^-)\|_{L^2(\Omega)}^2 + 2\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \|\nabla\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + 2M_1\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\leq C_{1,*}\sum_{n=1}^{j}\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} + C_{2,*}\sum_{n=1}^{j} k_n h^{2r} + 6\delta\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t
\end{aligned}$$

$$+\left(\frac{c_2^2}{\delta}+\frac{c_4^2}{\delta'}\right)\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\theta_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t$$
$$+8\delta'\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t+\frac{c_3^2}{\delta'}\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\theta_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t. \tag{4.3.39}$$

Choosing $\delta=\frac{M_1}{6}$, $\delta'=\frac{1}{8}$ and applying the Poincaré inequality, we have

$$\|\dot{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2+\frac{1}{C_{\text{poin}}}\|\theta_\phi(t_j^-)\|_{H^1(\Omega)}^2+\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t+M_1\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t$$
$$\leq\left(\frac{6c_2^2}{M_1}+8c_4^2\right)\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\theta_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t+8c_3^2\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\theta_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t$$
$$+C_{1,*}\sum_{n=1}^{j}\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}+C_{2,*}\sum_{n=1}^{j}k_nh^{2r}. \tag{4.3.40}$$

By the fundamental theorem of calculus and the triangle inequality, we have

$$\begin{aligned}\|\theta_\phi(t)\|_{L^2(\Omega)}^2&\leq\left(\|\theta_\phi(t_n^-)\|_{L^2(\Omega)}+\int_{t_{n-1}}^{t_n}\|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}\,\mathrm{d}t\right)^2\\&\leq 2\|\theta_\phi(t_n^-)\|_{L^2(\Omega)}^2+2\left(\int_{t_{n-1}}^{t_n}\|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}\,\mathrm{d}t\right)^2\\&\leq 2\|\theta_\phi(t_n^-)\|_{L^2(\Omega)}^2+2k_n\int_{t_{n-1}}^{t_n}\|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t.\end{aligned}$$

This implies that

$$\left(\frac{6c_2^2}{M_1}+8c_4^2\right)\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\theta_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t\leq 2\left(\frac{6c_2^2}{M_1}+8c_4^2\right)\sum_{n=1}^{j-1}k_n\|\theta_\phi(t_n^-)\|_{L^2(\Omega)}^2$$
$$+2\left(\frac{6c_2^2}{M_1}+8c_4^2\right)k_j\|\theta_\phi(t_j^-)\|_{L^2(\Omega)}+2\left(\frac{6c_2^2}{M_1}+8c_4^2\right)\sum_{n=1}^{j}k_n^2\int_{t_{n-1}}^{t_n}\|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t, \tag{4.3.41}$$

where the last two terms on the right-hand of the inequality can be absorbed into terms on the left-hand side. Analogously, we have

$$8c_3^2\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\theta_\phi(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t\leq 16c_3^2\sum_{n=1}^{j-1}k_n\|\nabla\theta_\phi(t_n^-)\|_{L^2(\Omega)}^2+16c_3^2k_j\|\nabla\theta_\phi(t_j^-)\|_{L^2(\Omega)}^2$$

$$+ 16c_3^2 \sum_{n=1}^{j} k_n^2 \int_{t_{n-1}}^{t_n} \|\nabla \dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t. \tag{4.3.42}$$

Combining (4.3.41) and (4.3.42), we can further simplify (4.3.40) to

$$\begin{aligned}
&\|\dot{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \left(\frac{1}{C_{\text{poin}}} - k_j \max\left\{\frac{12c_2^2}{M_1} + 16c_4^2, 16c_3^2\right\}\right) \|\theta_\phi(t_j^-)\|_{H^1(\Omega)}^2 \\
&\quad + \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \left(1 - 16c_3^2 k_n^2\right) \|\nabla \dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\quad + \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \left(M_1 - k_n^2 \left(\frac{12c_2^2}{M_1} + 16c_4^2\right)\right) \|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\leq C_{1,*} \sum_{n=1}^{j} \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} + C_{2,*} \sum_{n=1}^{j} k_n h^{2r} + \left(\frac{12c_2^2}{M_1} + 16c_4^2\right) \sum_{n=1}^{j-1} k_n \|\theta_\phi(t_n^-)\|_{L^2(\Omega)}^2 \\
&\quad + 16c_3^2 \sum_{n=1}^{j-1} k_n \|\nabla \theta_\phi(t_n^-)\|_{L^2(\Omega)}^2.
\end{aligned} \tag{4.3.43}$$

By choosing k_n for each $n = 1, 2, \ldots, N$ sufficiently small such that $k_n^2 \left(\frac{12c_2^2}{M_1} + 16c_4^2\right) \leq \frac{M_1}{2}$, $16c_3^2 k_n^2 \leq \frac{1}{2}$ and $k_j \max\left\{\frac{12c_2^2}{M_1} + 16c_4^2, 16c_3^2\right\} \leq \frac{1}{2C_{\text{poin}}}$, we have

$$\begin{aligned}
&\|\dot{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \|\theta_\phi(t_j^-)\|_{H^1(\Omega)}^2 + \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\nabla \dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\leq C_* \left(\sum_{n=1}^{j} \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} + k_n h^{2r}\right) + \tilde{C} \sum_{n=1}^{j-1} k_n \|\theta_\phi(t_n^-)\|_{H^1(\Omega)}^2.
\end{aligned}$$

By applying the discrete Grönwall lemma, we have

$$\begin{aligned}
&\|\dot{\theta}_\phi(t_j^-)\|_{L^2(\Omega)}^2 + \|\theta_\phi(t_j^-)\|_{H^1(\Omega)}^2 + \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\nabla \dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + \sum_{n=1}^{j} \int_{t_{n-1}}^{t_n} \|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\leq C_* \exp\left(\tilde{C} \sum_{n=1}^{j} k_n\right) \left(\sum_{n=1}^{j} \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} + k_n h^{2r}\right) \leq C_{\max}(u) \left(\sum_{n=1}^{j} \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}} + k_n h^{2r}\right),
\end{aligned} \tag{4.3.44}$$

where $C_{\max}(u) = C_* \exp(\tilde{C}T)$.

Now tracing back constants through the previous estimates, we notice that $C_{\max}(u)$ does not depend on $C_*(u)$, so we can define $C_*(u) := C_{\max}(u)$. That is, we have shown that $\mathcal{N}(\mathcal{F}) \subset \mathcal{F}$.

Verification that $\mathcal{N}$ is a contraction mapping

In this subsection, we establish the contraction property of $\mathcal{N}$ as stated in (b). Before we start the proof, we note that, since $a \in C^3(\mathbb{R})$ and $c \in C^4(\mathbb{R})$, a', a'', c' and c'' are locally Lipschitz continuous. We claim that there exist positive constants $C_{L1} - C_{L4}$ such that

$$|a'(\phi(t), u(t)) - a'(\phi'(t), u(t))| \leq C_{L1}|\phi(t) - \phi'(t)|, \tag{4.3.45}$$

$$|a''(\phi(t), u(t)) - a''(\phi'(t), u(t))| \leq C_{L2}|\phi(t) - \phi'(t)|, \tag{4.3.46}$$

$$|c'(\phi(t), W(t)) - c'(\phi'(t), W(t))| \leq C_{L3}|\phi(t) - \phi'(t)|, \tag{4.3.47}$$

$$|c''(\phi(t), W(t)) - c''(\phi'(t), W(t))| \leq C_{L4}|\phi(t) - \phi'(t)|, \tag{4.3.48}$$

for all $t \in I_n$ where $n = 1, \ldots, N$. We prove the first inequality (4.3.45); the proofs for the other three inequalities follow in exactly the same manner. By definition of $a'(\cdot, u(t))$, we have

$$\begin{aligned}
&|a'(\phi(t), u(t)) - a'(\phi'(t), u(t))| \\
&= \left| \int_0^1 a'(u(t) + \tau(\phi(t) - u(t))) - a'(u(t) + \tau(\phi'(t) - u(t)))\, \mathrm{d}\tau \right| \\
&= \left| \int_0^1 \int_0^1 \tau(\phi(t) - \phi'(t)) a''(\theta[u(t) + \tau(\phi(t) - u(t))] + (1-\theta)[u(t) + \tau(\phi'(t) - u(t))])\, \mathrm{d}\theta\, \mathrm{d}\tau \right| \\
&\leq \frac{1}{2}|\phi(t) - \phi'(t)| \max_{\theta,\tau \in [0,1]} \max_{t \in [0,T]} \|a''((1-\tau)u(t) + \theta\tau\phi(t) + (1-\theta)\tau\phi'(t))\|_{L^\infty(\Omega)}.
\end{aligned}$$

Working on the argument of a'', we have

$$\begin{aligned}
&\|(1-\tau)u(t) + \theta\tau\phi(t) + (1-\theta)\tau\phi'(t)\|_{L^\infty(\Omega)} \\
&\leq \|u(t)\|_{L^\infty(\Omega)} + \theta\tau\|\phi(t) - u(t)\|_{L^\infty(\Omega)} + (1-\theta)\tau\|\phi'(t) - u(t)\|_{L^\infty(\Omega)} \\
&\leq \|u(t)\|_{L^\infty(\Omega)} + \theta\tau\|\phi(t) - \Pi_k W(t)\|_{L^\infty(\Omega)} + \theta\tau\|\Pi_k W(t) - u(t)\|_{L^\infty(\Omega)} \\
&\quad + (1-\theta)\tau\|\phi'(t) - \Pi_k W(t)\|_{L^\infty(\Omega)} + (1-\theta)\tau\|\Pi_k W(t) - u(t)\|_{L^\infty(\Omega)} \\
&\leq \|u(t)\|_{L^\infty(\Omega)} + \theta\|\phi(t) - \Pi_k W(t)\|_{L^\infty(\Omega)} + (1-\theta)\|\phi'(t) - \Pi_k W(t)\|_{L^\infty(\Omega)}
\end{aligned}$$

$$
\begin{aligned}
&\quad + \|\Pi_k W(t) - W(t)\|_{L^\infty(\Omega)} + \|W(t) - u(t)\|_{L^\infty(\Omega)} \\
&\leq C_{\text{inv}} h^{-\frac{d}{2}} \left[C_*(u) \left(\sum_{i=1}^{n} k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} + \sqrt{C_1(u)} \frac{k_n^{q_n+\frac{1}{2}}}{q_n^{s-1}} \right] + c_0 + 2\|u(t)\|_{L^\infty(\Omega)} \\
&\leq 2\|u(t)\|_{L^\infty(\Omega)} + \tilde{c}_1
\end{aligned}
$$

as before. This implies that

$$
|a'(\phi(t), u(t)) - a'(\phi(t), u(t))| \leq \frac{1}{2} \max_{\xi} |a''(\xi)| |\phi(t) - \phi'(t)|, \tag{4.3.49}
$$

where $|\xi| \leq 2 \max_{t \in I_n, 1 \leq n \leq N} \|u(t)\|_{L^\infty(\Omega)} + \tilde{c}_1$. Hence (4.3.45) is proved with $C_{L1} = \frac{1}{2} \max_\xi |a''(\xi)|$. The proof of (4.3.46) follows in exactly the same manner, so we omit the proof here. The proof of (4.3.47) and (4.3.48) also follow immediately since

$$
\begin{aligned}
&\|(1-\tau)W(t) + \theta\tau\phi(t) + (1-\theta)\tau\phi'(t)\|_{L^\infty(\Omega)} \\
&\leq \|W(t)\|_{L^\infty(\Omega)} + \theta\tau\|\phi(t) - W(t)\|_{L^\infty(\Omega)} + (1-\theta)\tau\|\phi'(t) - W(t)\|_{L^\infty(\Omega)} \\
&\leq \|W(t)\|_{L^\infty(\Omega)} + \theta\tau\|\phi(t) - \Pi_k W(t)\|_{L^\infty(\Omega)} + (1-\theta)\tau\|\phi'(t) - \Pi_k W(t)\|_{L^\infty(\Omega)} \\
&\quad + \tau\|W(t) - \Pi_k W(t)\|_{L^\infty(\Omega)} \\
&\leq c_0 + C_{\text{inv}} h^{-\frac{d}{2}} \Big(\theta\tau\|\phi(t) - \Pi_k W(t)\|_{L^2(\Omega)} + (1-\theta)\tau\|\phi'(t) - \Pi_k W(t)\|_{L^2(\Omega)} \Big) \\
&\quad + C_{\text{inv}} h^{-\frac{d}{2}} \tau\|W(t) - \Pi_k W(t)\|_{L^2(\Omega)} \\
&\leq c_0 + C_{\text{inv}} h^{-\frac{d}{2}} \left(C_*(u) \left(\sum_{i=1}^{n} k_i h^{2r} + \frac{k_i^{2q_i+1}}{q_i^{2(s-1)}} \right)^{\frac{1}{2}} + \sqrt{C_1(u)} \frac{k_n^{q_n+\frac{1}{2}}}{q_n^{s-1}} \right) \\
&\leq \tilde{c}_3,
\end{aligned}
$$

where $\tilde{c}_3$ is a bounded positive constant.

To show the contraction property, we consider $R = \phi - \phi'$ and $\Theta = u_\phi - u_{\phi'}$, where $\phi, \phi' \in \mathcal{F}$. Then

$$
\int_{t_{n-1}}^{t_n} \left\langle \ddot{\Theta}(t), \dot{\varphi}(t) \right\rangle \, \mathrm{d}t + \Big([\dot{\Theta}]_{n-1}, \dot{\varphi}(t_{n-1}^+) \Big)_{L^2} + \int_{t_{n-1}}^{t_n} (\nabla\Theta(t), \nabla\dot{\varphi}(t))_{L^2} \, \mathrm{d}t + \big([\nabla\Theta]_{n-1}, \nabla\varphi(t_{n-1}^+) \big)_{L^2}
$$

$$
\begin{aligned}
&+\int_{t_{n-1}}^{t_n}\Big(\nabla\dot{\Theta}(t),\nabla\dot{\varphi}(t)\Big)_{L^2}\,\mathrm{d}t+\int_{t_{n-1}}^{t_n}\Big(\Big[a'(\phi(t),u(t))\dot{\theta}_\phi(t)-a'(\phi'(t),u(t))\dot{\theta}_{\phi'}(t)\Big],\dot{\varphi}(t)\Big)_{L^2}\,\mathrm{d}t\\
=&-\int_{t_{n-1}}^{t_n}\big(\big[a'(\phi(t),u(t))-a'(\phi'(t),u(t))\big]\left(\dot{\rho}_1(t)+\dot{\rho}_2(t)\right),\dot{\varphi}(t)\big)_{L^2}\,\mathrm{d}t\\
&-\int_{t_{n-1}}^{t_n}\big(\big[\partial_t a'(\phi(t),u(t))-\partial_t a'(\phi'(t),u(t))\big]\left(\rho_1(t)+\rho_2(t)\right),\dot{\varphi}(t)\big)_{L^2}\,\mathrm{d}t\\
&-\int_{t_{n-1}}^{t_n}\big(\big[\partial_t a'(\phi(t),u(t))\theta_\phi(t)-\partial_t a'(\phi'(t),u(t))\theta_{\phi'}(t)\big],\dot{\varphi}(t)\big)_{L^2}\,\mathrm{d}t\\
&-\int_{t_{n-1}}^{t_n}\big(\big[c'(\phi(t),W(t))-c'(\phi'(t),W(t))\big]\nabla\rho_1(t),\nabla\dot{\varphi}(t)\big)_{L^2}\,\mathrm{d}t\\
&-\int_{t_{n-1}}^{t_n}\big(\big[\nabla c'(\phi(t),W(t))-\nabla c'(\phi'(t),W(t))\big]\rho_1(t),\nabla\dot{\varphi}(t)\big)_{L^2}\,\mathrm{d}t\\
&-\int_{t_{n-1}}^{t_n}\big(\big[c'(\phi(t),W(t))\nabla\theta_\phi(t)-c'(\phi'(t),W(t))\nabla\theta_{\phi'}(t)\big],\nabla\dot{\varphi}(t)\big)_{L^2}\,\mathrm{d}t\\
&-\int_{t_{n-1}}^{t_n}\big(\big[\nabla c'(\phi(t),W(t))\theta_\phi(t)-\nabla c'(\phi'(t),W(t))\theta_{\phi'}(t)\big],\nabla\dot{\varphi}(t)\big)_{L^2}\,\mathrm{d}t.
\end{aligned}
\tag{4.3.50}
$$

Taking $\varphi=\Theta$, we obtain

$$
\begin{aligned}
&\frac{1}{2}\|\dot{\Theta}(t_n^-)\|^2_{L^2(\Omega)}+\frac{1}{2}\|\dot{\Theta}(t_{n-1}^+)\|^2_{L^2(\Omega)}+\frac{1}{2}\|\nabla\Theta(t_n^-)\|^2_{L^2(\Omega)}+\frac{1}{2}\|\nabla\Theta(t_{n-1}^+)\|^2_{L^2(\Omega)}\\
&\quad+\int_{t_{n-1}}^{t_n}\|\nabla\dot{\Theta}(t)\|^2_{L^2(\Omega)}\mathrm{d}t+\int_{t_{n-1}}^{t_n}\Big(\Big[a'(\phi(t),u(t))\dot{\theta}_\phi(t)-a'(\phi'(t),u(t))\dot{\theta}_{\phi'}(t)\Big],\dot{\Theta}(t)\Big)_{L^2}\mathrm{d}t\\
&=-\int_{t_{n-1}}^{t_n}\Big(\big[a'(\phi(t),u(t))-a'(\phi'(t),u(t))\big]\dot{\rho}_1(t),\dot{\Theta}(t)\Big)_{L^2}\mathrm{d}t\\
&\quad-\int_{t_{n-1}}^{t_n}\Big(\big[a'(\phi(t),u(t))-a'(\phi'(t),u(t))\big]\dot{\rho}_2(t),\dot{\Theta}(t)\Big)_{L^2}\mathrm{d}t\\
&\quad-\int_{t_{n-1}}^{t_n}\Big(\big[\partial_t a'(\phi(t),u(t))-\partial_t a'(\phi'(t),u(t))\big]\rho_1(t),\dot{\Theta}(t)\Big)_{L^2}\mathrm{d}t\\
&\quad-\int_{t_{n-1}}^{t_n}\Big(\big[\partial_t a'(\phi(t),u(t))-\partial_t a'(\phi'(t),u(t))\big]\rho_2(t),\dot{\Theta}(t)\Big)_{L^2}\mathrm{d}t\\
&\quad-\int_{t_{n-1}}^{t_n}\Big(\big[\partial_t a'(\phi(t),u(t))\theta_\phi(t)-\partial_t a'(\phi'(t),u(t))\theta_{\phi'}(t)\big],\dot{\Theta}(t)\Big)_{L^2}\mathrm{d}t\\
&\quad-\int_{t_{n-1}}^{t_n}\Big(\big[c'(\phi(t),W(t))-c'(\phi'(t),W(t))\big]\nabla\rho_1(t),\nabla\dot{\Theta}(t)\Big)_{L^2}\mathrm{d}t\\
&\quad-\int_{t_{n-1}}^{t_n}\Big(\big[\nabla c'(\phi(t),W(t))-\nabla c'(\phi'(t),W(t))\big]\rho_1(t),\nabla\dot{\Theta}(t)\Big)_{L^2}\mathrm{d}t
\end{aligned}
$$

$$
\begin{aligned}
&- \int_{t_{n-1}}^{t_n} \Big(\left[c'(\phi(t), W(t)) \nabla \theta_\phi(t) - c'(\phi'(t), W(t)) \nabla \theta_{\phi'}(t) \right], \nabla \dot{\Theta}(t) \Big)_{L^2} \,\mathrm{d}t \\
&- \int_{t_{n-1}}^{t_n} \Big(\left[\nabla c'(\phi(t), W(t)) \theta_\phi(t) - \nabla c'(\phi'(t), W(t)) \theta_{\phi'}(t) \right], \nabla \dot{\Theta}(t) \Big)_{L^2} \,\mathrm{d}t \\
&+ \Big(\dot{\Theta}(t_{n-1}^-), \dot{\Theta}(t_{n-1}^+) \Big)_{L^2} + \left(\nabla \Theta(t_{n-1}^-), \nabla \Theta(t_{n-1}^+) \right)_{L^2}. \qquad (4.3.51)
\end{aligned}
$$

Note that

$$
\begin{aligned}
&\Big(a'(\phi(t), u(t)) \dot{\theta}_\phi(t), \dot{\Theta}(t) \Big)_{L^2} - \Big(a'(\phi'(t), u(t)) \dot{\theta}_{\phi'}(t), \dot{\Theta}(t) \Big)_{L^2} \\
&= \Big(a'(\phi(t), u(t)) \dot{\Theta}(t), \dot{\Theta}(t) \Big)_{L^2} + \Big(\left[a'(\phi(t), u(t)) - a'(\phi'(t), u(t)) \right] \dot{\theta}_{\phi'}(t), \dot{\Theta}(t) \Big)_{L^2}.
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&\Big(\partial_t a'(\phi(t), u(t)) \theta_\phi(t), \dot{\Theta}(t) \Big)_{L^2} - \Big(\partial_t a'(\phi'(t), u(t)) \theta_{\phi'}(t), \dot{\Theta}(t) \Big)_{L^2} \\
&= \Big(\partial_t a'(\phi(t), u(t)) \Theta(t), \dot{\Theta}(t) \Big)_{L^2} + \Big(\left[\partial_t a'(\phi(t), u(t)) - \partial_t a'(\phi'(t), u(t)) \right] \theta_{\phi'}(t), \dot{\Theta}(t) \Big)_{L^2},
\end{aligned}
$$

$$
\begin{aligned}
&\Big(c'(\phi(t), W(t)) \nabla \theta_\phi(t), \dot{\Theta}(t) \Big)_{L^2} - \Big(c'(\phi'(t), W(t)) \nabla \theta_{\phi'}(t), \dot{\Theta}(t) \Big)_{L^2} \\
&= \Big(c'(\phi(t), W(t)) \nabla \Theta(t), \dot{\Theta}(t) \Big)_{L^2} + \Big(\left[c'(\phi(t), W(t)) - c'(\phi'(t), W(t)) \right] \nabla \theta_{\phi'}(t), \dot{\Theta}(t) \Big)_{L^2},
\end{aligned}
$$

and

$$
\begin{aligned}
&\Big(\nabla c'(\phi(t), W(t)) \theta_\phi(t), \dot{\Theta}(t) \Big)_{L^2} - \Big(\nabla c'(\phi'(t), W(t)) \theta_{\phi'}(t), \dot{\Theta}(t) \Big)_{L^2} \\
&= \Big(\nabla c'(\phi(t), W(t)) \Theta(t), \dot{\Theta}(t) \Big)_{L^2} + \Big(\left[\nabla c'(\phi(t), W(t)) - \nabla c'(\phi'(t), W(t)) \right] \theta_{\phi'}(t), \dot{\Theta}(t) \Big)_{L^2}.
\end{aligned}
$$

By the Cauchy–Schwarz inequality, we have

$$
2 \Big(\dot{\Theta}(t_{n-1}^-), \dot{\Theta}(t_{n-1}^+) \Big)_{L^2} \leq \|\dot{\Theta}(t_{n-1}^-)\|_{L^2(\Omega)}^2 + \|\dot{\Theta}(t_{n-1}^+)\|_{L^2(\Omega)}^2, \qquad (4.3.52)
$$

$$
2 \left(\nabla \Theta(t_{n-1}^-), \nabla \Theta(t_{n-1}^+) \right)_{L^2} \leq \|\nabla \Theta(t_{n-1}^-)\|_{L^2(\Omega)}^2 + \|\nabla \Theta(t_{n-1}^+)\|_{L^2(\Omega)}^2. \qquad (4.3.53)
$$

Combining (4.3.52), (4.3.53) and the above equalities, we deduce from (4.3.51) that

$$
\|\dot{\Theta}(t_n^-)\|_{L^2(\Omega)}^2 + \|\nabla \Theta(t_n^-)\|_{L^2(\Omega)}^2 + 2 \int_{t_{n-1}}^{t_n} \|\nabla \dot{\Theta}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + 2 \int_{t_{n-1}}^{t_n} \Big(a'(\phi(t), u(t)) \dot{\Theta}(t), \dot{\Theta}(t) \Big)_{L^2} \,\mathrm{d}t
$$

$$\leq \|\dot{\Theta}(t_{n-1}^-)\|^2_{L^2(\Omega)} + \|\nabla\Theta(t_{n-1}^-)\|^2_{L^2(\Omega)} + |\mathrm{I}_a| + |\mathrm{II}_a| + |\mathrm{III}_a| + |\mathrm{IV}_a| + |\mathrm{V}_a| + |\mathrm{VI}_a| + |\mathrm{VII}_a|$$
$$+ |\mathrm{I}_c| + |\mathrm{II}_c| + |\mathrm{III}_c| + |\mathrm{IV}_c| + |\mathrm{V}_c| + |\mathrm{VI}_c|, \tag{4.3.54}$$

where $\mathrm{I}_a, \ldots, \mathrm{VI}_c$ are terms defined and estimated below.

$$\begin{aligned}|\mathrm{I}_a| &:= \left| -2\int_{t_{n-1}}^{t_n} \Big(\big[a'(\phi(t),u(t)) - a'(\phi'(t),u(t))\big]\, \dot{\theta}_\phi(t), \dot{\Theta}(t)\Big)_{L^2}\, \mathrm{d}t\right| \\ &\leq 2C_{L1}\int_{t_{n-1}}^{t_n} \|\phi(t)-\phi'(t)\|_{L^\infty(\Omega)}\|\dot{\theta}_\phi(t)\|_{L^2(\Omega)}\|\dot{\Theta}(t)\|_{L^2(\Omega)}\,\mathrm{d}t \\ &\leq \frac{C_{L1}^2}{\tilde{\delta}}\max_{t\in I_n}\|R(t)\|^2_{L^\infty(\Omega)}\int_{t_{n-1}}^{t_n}\|\dot{\theta}_\phi(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \tilde{\delta}\int_{t_{n-1}}^{t_n}\|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t, \end{aligned} \tag{4.3.55}$$

for some $\tilde{\delta} > 0$ that will be specified later.

$$\begin{aligned}|\mathrm{II}_a| &:= \left| -2\int_{t_{n-1}}^{t_n} \Big(\big[a'(\phi(t),u(t)) - a'(\phi'(t),u(t))\big]\, \dot{\rho}_1(t), \dot{\Theta}(t)\Big)_{L^2}\, \mathrm{d}t\right| \\ &\leq 2C_{L1}\int_{t_{n-1}}^{t_n} \|\phi(t)-\phi'(t)\|_{L^\infty(\Omega)}\|\dot{\rho}_1(t)\|_{L^2(\Omega)}\|\dot{\Theta}(t)\|_{L^2(\Omega)}\,\mathrm{d}t \\ &\leq \frac{C_{L1}^2}{\tilde{\delta}}\max_{t\in I_n}\|R(t)\|^2_{L^\infty(\Omega)}\int_{t_{n-1}}^{t_n}\|\dot{\rho}_1(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \tilde{\delta}\int_{t_{n-1}}^{t_n}\|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t. \end{aligned} \tag{4.3.56}$$

$$\begin{aligned}|\mathrm{III}_a| &:= \left| -2\int_{t_{n-1}}^{t_n} \Big(\big[a'(\phi(t),u(t)) - a'(\phi'(t),u(t))\big]\, \dot{\rho}_2(t), \dot{\Theta}(t)\Big)_{L^2}\, \mathrm{d}t\right| \\ &\leq 2C_{L1}\int_{t_{n-1}}^{t_n} \|\phi(t)-\phi'(t)\|_{L^\infty(\Omega)}\|\dot{\rho}_2(t)\|_{L^2(\Omega)}\|\dot{\Theta}(t)\|_{L^2(\Omega)}\,\mathrm{d}t \\ &\leq \frac{C_{L1}^2}{\tilde{\delta}}\max_{t\in I_n}\|R(t)\|^2_{L^\infty(\Omega)}\int_{t_{n-1}}^{t_n}\|\dot{\rho}_2(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \tilde{\delta}\int_{t_{n-1}}^{t_n}\|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t. \end{aligned} \tag{4.3.57}$$

$$\begin{aligned}|\mathrm{IV}_a| &:= \left| -2\int_{t_{n-1}}^{t_n} \Big(\big[\partial_t a'(\phi(t),u(t)) - \partial_t a'(\phi'(t),u(t))\big]\, \rho_1(t), \dot{\Theta}(t)\Big)_{L^2}\, \mathrm{d}t\right| \\ &= \Bigg| -2\int_{t_{n-1}}^{t_n} \Bigg(\int_0^1 \big[a''(u(t)+\tau(\phi(t)-u(t))) - a''(u(t)+\tau(\phi'(t)-u(t)))\big]\,(\dot{u}(t)+\tau(\dot{\phi}(t)-\dot{u}(t))) \\ &\quad + a''(u(t)+\tau(\phi'(t)-u(t)))\tau(\dot{\phi}(t)-\dot{\phi}'(t))\mathrm{d}\tau\rho_1(t), \dot{\Theta}(t)\Bigg)_{L^2}\mathrm{d}t\Bigg| \\ &\leq \frac{C_{L2}^2c_5^2}{\tilde{\delta}}\int_{t_{n-1}}^{t_n}\|\phi(t)-\phi'(t)\|^2_{L^\infty(\Omega)}\|\rho_1(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \tilde{\delta}\int_{t_{n-1}}^{t_n}\|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t \\ &\quad + \frac{c_6^2}{\tilde{\delta}}\int_{t_{n-1}}^{t_n}\|\dot{\phi}(t)-\dot{\phi}'(t)\|^2_{L^\infty(\Omega)}\|\rho_1(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \tilde{\delta}\int_{t_{n-1}}^{t_n}\|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t \end{aligned}$$

$$\leq \left(\frac{C_{L2}^2 c_5^2}{\tilde{\delta}} \max_{t\in I_n} \|R(t)\|^2_{L^\infty(\Omega)} + \frac{c_6^2}{\tilde{\delta}} \max_{t\in I_n} \|\dot{R}(t)\|^2_{L^\infty(\Omega)} \right) \int_{t_{n-1}}^{t_n} \|\rho_1(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$
$$+ 2\tilde{\delta} \int_{t_{n-1}}^{t_n} \|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t, \tag{4.3.58}$$

where

$$c_5 := \max_{\tau\in[0,1]} \max_{t\in I_n, 1\leq n\leq N} \|\dot{u}(t) + \tau(\dot{\phi}(t) - \dot{u}(t))\|_{L^\infty(\Omega)}$$

and

$$c_6 := \max_{\tau\in[0,1]} \max_{t\in I_n, 1\leq n\leq N} \|a''(u(t) + \tau(\phi'(t) - u(t)))\|_{L^\infty(\Omega)}.$$

Analogously, we have

$$|\mathrm{V}_a| := \left| -2 \int_{t_{n-1}}^{t_n} \Big(\left[\partial_t a'(\phi(t), u(t)) - \partial_t a'(\phi'(t), u(t))\right] \rho_2(t), \dot{\Theta}(t) \Big)_{L^2}\,\mathrm{d}t \right|$$
$$\leq \frac{C_{L2}^2 c_5^2}{\tilde{\delta}} \int_{t_{n-1}}^{t_n} \|\phi(t) - \phi'(t)\|^2_{L^\infty(\Omega)} \|\rho_2(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \tilde{\delta} \int_{t_{n-1}}^{t_n} \|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$
$$+ \frac{c_6^2}{\tilde{\delta}} \int_{t_{n-1}}^{t_n} \|\dot{\phi}(t) - \dot{\phi}'(t)\|^2_{L^\infty(\Omega)} \|\rho_2(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \tilde{\delta} \int_{t_{n-1}}^{t_n} \|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$
$$\leq \left(\frac{C_{L2}^2 c_5^2}{\tilde{\delta}} \max_{t\in I_n} \|R(t)\|^2_{L^\infty(\Omega)} + \frac{c_6^2}{\tilde{\delta}} \max_{t\in I_n} \|\dot{R}(t)\|^2_{L^\infty(\Omega)} \right) \int_{t_{n-1}}^{t_n} \|\rho_2(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$
$$+ 2\tilde{\delta} \int_{t_{n-1}}^{t_n} \|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t, \tag{4.3.59}$$

and

$$|\mathrm{VI}_a| := \left| -2 \int_{t_{n-1}}^{t_n} \Big(\left[\partial_t a'(\phi(t), u(t)) - \partial_t a'(\phi'(t), u(t))\right] \theta_{\phi'}(t), \dot{\Theta}(t) \Big)_{L^2}\,\mathrm{d}t \right|$$
$$\leq \frac{C_{L2}^2 c_5^2}{\tilde{\delta}} \int_{t_{n-1}}^{t_n} \|\phi(t) - \phi'(t)\|^2_{L^\infty(\Omega)} \|\theta_{\phi'}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \tilde{\delta} \int_{t_{n-1}}^{t_n} \|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$
$$+ \frac{c_6^2}{\tilde{\delta}} \int_{t_{n-1}}^{t_n} \|\dot{\phi}(t) - \dot{\phi}'(t)\|^2_{L^\infty(\Omega)} \|\theta_{\phi'}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \tilde{\delta} \int_{t_{n-1}}^{t_n} \|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$
$$\leq \left(\frac{C_{L2}^2 c_5^2}{\tilde{\delta}} \max_{t\in I_n} \|R(t)\|^2_{L^\infty(\Omega)} + \frac{c_6^2}{\tilde{\delta}} \max_{t\in I_n} \|\dot{R}(t)\|^2_{L^\infty(\Omega)} \right) \int_{t_{n-1}}^{t_n} \|\theta_{\phi'}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$
$$+ 2\tilde{\delta} \int_{t_{n-1}}^{t_n} \|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t. \tag{4.3.60}$$

$$|\mathrm{VII}_a| := \left| -2 \int_{t_{n-1}}^{t_n} \Big(\partial_t a'(\phi(t), u(t)) \Theta(t), \dot{\Theta}(t) \Big)_{L^2}\,\mathrm{d}t \right|$$

$$\leq \frac{c_7^2}{\tilde{\delta}} \int_{t_{n-1}}^{t_n} \|\Theta(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t + \tilde{\delta} \int_{t_{n-1}}^{t_n} \|\dot{\Theta}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t, \tag{4.3.61}$$

where

$$c_7 := \max_{\tau \in [0,1]} \max_{t \in I_n, 1 \leq n \leq N} \Big(\|a''(u(t) + \tau(\phi(t) - u(t)))\|_{L^\infty(\Omega)} \|\dot{u}(t) + \tau(\dot{\phi}(t) - \dot{u}(t))\|_{L^\infty(\Omega)} \Big).$$

As for the terms involving $c'(\phi(t), W(t))$ and $\nabla c'(\phi(t), W(t))$, we have

$$\begin{aligned} |\mathrm{I}_c| &:= \left| -2 \int_{t_{n-1}}^{t_n} \Big(\big[c'(\phi(t), W(t)) - c'(\phi'(t), W(t))\big] \nabla \rho_1(t), \nabla \dot{\Theta}(t) \Big)_{L^2} \, \mathrm{d}t \right| \\ &\leq \frac{C_{L3}^2}{\hat{\delta}} \max_{t \in I_n} \|R(t)\|_{L^\infty(\Omega)}^2 \int_{t_{n-1}}^{t_n} \|\nabla \rho_1(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t + \hat{\delta} \int_{t_{n-1}}^{t_n} \|\nabla \dot{\Theta}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t, \end{aligned} \tag{4.3.62}$$

for some constant $\hat{\delta} > 0$ which will be specified later.

$$\begin{aligned} |\mathrm{II}_c| &:= \left| -2 \int_{t_{n-1}}^{t_n} \Big(\big[c'(\phi(t), W(t)) - c'(\phi'(t), W(t))\big] \nabla \theta_{\phi'}(t), \nabla \dot{\Theta}(t) \Big)_{L^2} \, \mathrm{d}t \right| \\ &\leq \frac{C_{L3}^2}{\hat{\delta}} \max_{t \in I_n} \|R(t)\|_{L^\infty(\Omega)}^2 \int_{t_{n-1}}^{t_n} \|\nabla \theta_{\phi'}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t + \hat{\delta} \int_{t_{n-1}}^{t_n} \|\nabla \dot{\Theta}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t, \end{aligned} \tag{4.3.63}$$

$$\begin{aligned} |\mathrm{III}_c| &:= \left| -2 \int_{t_{n-1}}^{t_n} \Big(c'(\phi(t), W(t)) \nabla \Theta(t), \nabla \dot{\Theta}(t) \Big)_{L^2} \, \mathrm{d}t \right| \\ &\leq \frac{c_8^2}{\hat{\delta}} \int_{t_{n-1}}^{t_n} \|\nabla \Theta(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t + \hat{\delta} \int_{t_{n-1}}^{t_n} \|\nabla \dot{\Theta}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t, \end{aligned} \tag{4.3.64}$$

where

$$c_8 := \max_{\tau \in [0,1]} \max_{t \in I_n, 1 \leq n \leq N} \|c'(W(t) + \tau(\phi(t) - W(t)))\|_{L^\infty(\Omega)}.$$

$$\begin{aligned} |\mathrm{IV}_c| &:= \left| -2 \int_{t_{n-1}}^{t_n} \Big(\big[\nabla c'(\phi(t), W(t)) - \nabla c'(\phi'(t), W(t))\big] \rho_1(t), \nabla \dot{\Theta}(t) \Big)_{L^2} \, \mathrm{d}t \right| \\ &= \Bigg| -2 \int_{t_{n-1}}^{t_n} \Bigg(\int_0^1 \big[c''(W + \tau(\phi - W)) - c''(W + \tau(\phi' - W))\big] (\nabla W + \tau(\nabla \phi - \nabla W)) \\ &\quad + c''(W + \tau(\phi' - W)) \tau (\nabla \phi - \nabla \phi') \mathrm{d}\tau \rho_1(t), \nabla \dot{\Theta}(t) \Bigg)_{L^2} \mathrm{d}t \Bigg| \\ &\leq \frac{C_{L4}^2 c_9^2}{\hat{\delta}} \int_{t_{n-1}}^{t_n} \|\phi(t) - \phi'(t)\|_{L^\infty(\Omega)}^2 \|\rho_1(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t + \hat{\delta} \int_{t_{n-1}}^{t_n} \|\nabla \dot{\Theta}(t)\|_{L^2(\Omega)}^2 \, \mathrm{d}t \end{aligned}$$

$$
\begin{aligned}
&+ \frac{c_{10}^2}{\hat{\delta}} \int_{t_{n-1}}^{t_n} \|\nabla\phi(t) - \nabla\phi'(t)\|_{L^\infty(\Omega)}^2 \|\rho_1(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + \hat{\delta} \int_{t_{n-1}}^{t_n} \|\nabla\dot{\Theta}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\leq \left(\frac{C_{L4}^2 c_9^2}{\hat{\delta}} \max_{t\in I_n} \|R(t)\|_{L^\infty(\Omega)}^2 + \frac{c_{10}^2}{\hat{\delta}} \max_{t\in I_n} \|\nabla R(t)\|_{L^\infty(\Omega)}^2 \right) \int_{t_{n-1}}^{t_n} \|\rho_1(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&+ 2\hat{\delta} \int_{t_{n-1}}^{t_n} \|\nabla\dot{\Theta}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t, \qquad (4.3.65)
\end{aligned}
$$

where

$$
c_9 := \max_{\tau\in[0,1]} \max_{t\in I_n, 1\leq n\leq N} \|\nabla W(t) + \tau(\nabla\phi(t) - \nabla W(t))\|_{L^\infty(\Omega)}
$$

and

$$
c_{10} := \max_{\tau\in[0,1]} \max_{t\in I_n, 1\leq n\leq N} \|c''(W(t) + \tau(\phi'(t) - W(t)))\|_{L^\infty(\Omega)}.
$$

$$
\begin{aligned}
|\mathrm{V}_c| &:= \left| -2 \int_{t_{n-1}}^{t_n} \left(\nabla c'(\phi(t), W(t))\Theta(t), \nabla\dot{\Theta}(t) \right)_{L^2} \mathrm{d}t \right| \\
&\leq \frac{c_{11}^2}{\hat{\delta}} \int_{t_{n-1}}^{t_n} \|\Theta(t)\|_{L^2(\Omega)}^2 \mathrm{d}t + \hat{\delta} \int_{t_{n-1}}^{t_n} \|\nabla\dot{\Theta}(t)\|_{L^2(\Omega)}^2 \mathrm{d}t, \qquad (4.3.66)
\end{aligned}
$$

where

$$
c_{11} := \max_{\tau\in[0,1]} \max_{t\in I_n, 1\leq n\leq N} \left(\|c''(W(t) + \tau(\phi(t) - W(t)))\|_{L^\infty(\Omega)} \|\nabla W(t) + \tau(\nabla\phi(t) - \nabla W(t))\|_{L^\infty(\Omega)} \right).
$$

$$
\begin{aligned}
|\mathrm{VI}_c| &:= \left| -2 \int_{t_{n-1}}^{t_n} \left(\left[\nabla c'(\phi(t), W(t)) - \nabla c'(\phi'(t), W(t))\right] \theta_{\phi'}(t), \nabla\dot{\Theta}(t) \right)_{L^2} \mathrm{d}t \right| \\
&= \Bigg| -2 \int_{t_{n-1}}^{t_n} \Bigg(\int_0^1 \left[c''(W + \tau(\phi - W)) - c''(W + \tau(\phi' - W)) \right] (\nabla W + \tau(\nabla\phi - \nabla W)) \\
&\quad + c''(W + \tau(\phi' - W))\tau(\nabla\phi - \nabla\phi')\mathrm{d}\tau\, \theta_{\phi'}(t), \nabla\dot{\Theta}(t) \Bigg)_{L^2} \mathrm{d}t \Bigg| \\
&\leq \frac{C_{L4}^2 c_9^2}{\hat{\delta}} \int_{t_{n-1}}^{t_n} \|\phi(t) - \phi'(t)\|_{L^\infty(\Omega)}^2 \|\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + \hat{\delta} \int_{t_{n-1}}^{t_n} \|\nabla\dot{\Theta}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\quad + \frac{c_{10}^2}{\hat{\delta}} \int_{t_{n-1}}^{t_n} \|\nabla\phi(t) - \nabla\phi'(t)\|_{L^\infty(\Omega)}^2 \|\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t + \hat{\delta} \int_{t_{n-1}}^{t_n} \|\nabla\dot{\Theta}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\leq \left(\frac{C_{L4}^2 c_9^2}{\hat{\delta}} \max_{t\in I_n} \|R(t)\|_{L^\infty(\Omega)}^2 + \frac{c_{10}^2}{\hat{\delta}} \max_{t\in I_n} \|\nabla R(t)\|_{L^\infty(\Omega)}^2 \right) \int_{t_{n-1}}^{t_n} \|\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t \\
&\quad + 2\hat{\delta} \int_{t_{n-1}}^{t_n} \|\nabla\dot{\Theta}(t)\|_{L^2(\Omega)}^2 \,\mathrm{d}t. \qquad (4.3.67)
\end{aligned}
$$

Again, we need to verify that all these constant c_5–c_{11} appearing in the above estimates are bounded. First note that

$$c_5 := \max_{\tau\in[0,1]} \max_{t\in I_n, 1\le n\le N} \|\dot{u}(t) + \tau(\dot{\phi}(t) - \dot{u}(t))\|_{L^\infty(\Omega)} \le 2 \max_{t\in I_n, 1\le n\le N} \|\partial_t u(t)\|_{L^\infty(\Omega)} + \tilde{c}_2.$$

$$c_6 := \max_{\tau\in[0,1]} \max_{t\in I_n, 1\le n\le N} \|a''(u(t) + \tau(\phi'(t) - u(t))\|_{L^\infty(\Omega)} \le \max_{\xi} |a''(\xi)|,$$

and

$$\begin{aligned} c_7 &:= \max_{\tau\in[0,1]} \max_{t\in I_n, 1\le n\le N} \left(\|a''(u(t) + \tau(\phi(t) - u(t))\|_{L^\infty(\Omega)} \|\dot{u}(t) + \tau(\dot{\phi}(t) - \dot{u}(t))\|_{L^\infty(\Omega)} \right) \\ &\le \max_{\xi} |a''(\xi)| \left(2 \max_{t\in I_n, 1\le n\le N} \|\partial_t u(t)\|_{L^\infty(\Omega)} + \tilde{c}_2 \right), \end{aligned}$$

where $|\xi| \le \tilde{c}_1 + 2\max_{t\in I_n, 1\le n\le N} \|u(t)\|_{L^\infty(\Omega)}$. Clearly,

$$c_8 := \max_{\tau\in[0,1]} \max_{t\in I_n, 1\le n\le N} \|c'(W(t) + \tau(\phi(t) - W(t)))\|_{L^\infty(\Omega)} \le C_c < \infty.$$

Note that for each $t \in I_n, 1 \le n \le N$,

$$\|W(t) + \tau(\phi(t) - W(t))\|_{L^\infty(\Omega)} = \|(1-\tau)W(t) + \tau\phi(t)\|_{L^\infty(\Omega)} \le \tilde{c}_3,$$

$$\|\nabla W(t) + \tau(\nabla\phi(t) - \nabla W(t))\|_{L^\infty(\Omega)} = \|\tau\nabla\phi(t) + (1-\tau)\nabla W(t)\|_{L^\infty(\Omega)} \le \tilde{c}_4.$$

This implies that

$$c_9 := \max_{\tau\in[0,1]} \max_{t\in I_n, 1\le n\le N} \|\nabla W(t) + \tau(\nabla\phi(t) - \nabla W(t))\|_{L^\infty(\Omega)} \le \tilde{c}_4,$$

$$c_{10} := \max_{\tau\in[0,1]} \max_{t\in I_n, 1\le n\le N} \|c''(W(t) + \tau(\phi(t) - W(t))\|_{L^\infty(\Omega)} \le |c''(\xi)|,$$

and

$$\begin{aligned} c_{11} &:= \max_{\tau\in[0,1]} \max_{t\in I_n, 1\le n\le N} \left(\|c''(W(t) + \tau(\phi(t) - W(t))\|_{L^\infty(\Omega)} \|\nabla W(t) + \tau(\nabla\phi(t) - \nabla W(t))\|_{L^\infty(\Omega)} \right) \\ &\le |c''(\xi)|\tilde{c}_4, \end{aligned}$$

where $|\xi| \le \tilde{c}_3$. Here, c_2, $\tilde{c}_4$, $\tilde{c}_3$, $\tilde{c}_2$ and $\tilde{c}_1$ are defined as in the proof of (a).

Combining all these estimates, we have

$$
\begin{aligned}
&\|\dot{\Theta}(t_n^-)\|_{L^2}^2 + \|\nabla\Theta(t_n^-)\|_{L^2(\Omega)}^2 + (2-8\hat{\delta})\int_{t_{n-1}}^{t_n} \|\nabla\dot{\Theta}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t + (2M_1 - 10\tilde{\delta})\int_{t_{n-1}}^{t_n} \|\dot{\Theta}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \\
&\le \|\dot{\Theta}(t_{n-1}^-)\|_{L^2(\Omega)}^2 + \|\nabla\Theta(t_{n-1}^-)\|_{L^2(\Omega)}^2 + \left(\frac{c_{11}^2}{\hat{\delta}} + \frac{c_7^2}{\tilde{\delta}}\right)\int_{t_{n-1}}^{t_n} \|\Theta(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t + \frac{c_8^2}{\tilde{\delta}}\int_{t_{n-1}}^{t_n} \|\nabla\Theta(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \\
&\quad + \frac{C_{L1}^2}{\tilde{\delta}} \max_{t\in I_n} \|R(t)\|_{L^\infty(\Omega)}^2 \left\{ \int_{t_{n-1}}^{t_n} \|\dot{\theta}_{\phi'}(t)\|_{L^2(\Omega)}^2 + \|\dot{\rho}_1(t)\|_{L^2(\Omega)}^2 + \|\dot{\rho}_2(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\} \\
&\quad + \frac{C_{L2}^2 c_5^2}{\tilde{\delta}} \max_{t\in I_n} \|R(t)\|_{L^\infty(\Omega)}^2 \left\{ \int_{t_{n-1}}^{t_n} \|\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 + \|\rho_1(t)\|_{L^2(\Omega)}^2 + \|\rho_2(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\} \\
&\quad + \frac{c_6^2}{\tilde{\delta}} \max_{t\in I_n} \|\dot{R}(t)\|_{L^\infty(\Omega)}^2 \left\{ \int_{t_{n-1}}^{t_n} \|\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 + \|\rho_1(t)\|_{L^2(\Omega)}^2 + \|\rho_2(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\} \\
&\quad + \frac{C_{L3}^2}{\hat{\delta}} \max_{t\in I_n} \|R(t)\|_{L^\infty(\Omega)}^2 \left\{ \int_{t_{n-1}}^{t_n} \|\nabla\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 + \|\nabla\rho_1(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\} \\
&\quad + \frac{C_{L4}^2 c_9^2}{\hat{\delta}} \max_{t\in I_n} \|R(t)\|_{L^\infty(\Omega)}^2 \left\{ \int_{t_{n-1}}^{t_n} \|\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 + \|\rho_1(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\} \\
&\quad + \frac{c_{10}^2}{\hat{\delta}} \max_{t\in I_n} \|\nabla R(t)\|_{L^\infty(\Omega)}^2 \left\{ \int_{t_{n-1}}^{t_n} \|\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 + \|\rho_1(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\} \\
&\le \|\dot{\Theta}(t_{n-1}^-)\|_{L^2(\Omega)}^2 + \|\nabla\Theta(t_{n-1}^-)\|_{L^2(\Omega)}^2 + \left(\frac{c_{11}^2}{\hat{\delta}} + \frac{c_7^2}{\tilde{\delta}}\right)\int_{t_{n-1}}^{t_n} \|\Theta(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t + \frac{c_8^2}{\tilde{\delta}}\int_{t_{n-1}}^{t_n} \|\nabla\Theta(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \\
&\quad + C_{\mathrm{inv}}^2 \frac{C_{L1}^2}{\delta} \max_{t\in I_n} \|R(t)\|_{L^2(\Omega)}^2 h^{-d} \left\{ \int_{t_{n-1}}^{t_n} \|\dot{\theta}_{\phi'}(t)\|_{L^2}^2 + \|\dot{\rho}_1(t)\|_{L^2(\Omega)}^2 + \|\dot{\rho}_2(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\} \\
&\quad + C_{\mathrm{inv}}^2 \frac{C_{L2}^2 c_5^2}{\tilde{\delta}} \max_{t\in I_n} \|R(t)\|_{L^2(\Omega)}^2 h^{-d} \left\{ \int_{t_{n-1}}^{t_n} \|\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 + \|\rho_1(t)\|_{L^2(\Omega)}^2 + \|\rho_2(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\} \\
&\quad + C_{\mathrm{inv}}^2 \frac{c_6^2}{\tilde{\delta}} \max_{t\in I_n} \|\dot{R}(t)\|_{L^2(\Omega)}^2 h^{-d} \left\{ \int_{t_{n-1}}^{t_n} \|\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 + \|\rho_1(t)\|_{L^2(\Omega)}^2 + \|\rho_2(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\} \\
&\quad + C_{\mathrm{inv}}^2 \frac{C_{L3}^2}{\hat{\delta}} \max_{t\in I_n} \|R(t)\|_{L^2(\Omega)}^2 h^{-d} \left\{ \int_{t_{n-1}}^{t_n} \|\nabla\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 + \|\nabla\rho_1(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\} \\
&\quad + C_{\mathrm{inv}}^2 \frac{C_{L4}^2 c_9^2}{\hat{\delta}} \max_{t\in I_n} \|R(t)\|_{L^2(\Omega)}^2 h^{-d} \left\{ \int_{t_{n-1}}^{t_n} \|\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 + \|\rho_1(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\} \\
&\quad + C_{\mathrm{inv}}^2 \frac{c_{10}^2}{\hat{\delta}} \max_{t\in I_n} \|\nabla R(t)\|_{L^2(\Omega)}^2 h^{-d} \left\{ \int_{t_{n-1}}^{t_n} \|\theta_{\phi'}(t)\|_{L^2(\Omega)}^2 + \|\rho_1(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t \right\}. \qquad (4.3.68)
\end{aligned}
$$

Taking $\hat{\delta} = \frac{1}{8}$, $\tilde{\delta} = \frac{M_1}{10}$ and summing up over $n = 1, \ldots, j$, we obtain

$$
\|\dot{\Theta}(t_j^-)\|_{L^2(\Omega)}^2 + \|\nabla\Theta(t_j^-)\|_{L^2(\Omega)}^2 + \sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \|\nabla\dot{\Theta}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t + \sum_{n=1}^{j}\int_{t_{n-1}}^{t_n} \|\dot{\Theta}(t)\|_{L^2(\Omega)}^2\,\mathrm{d}t
$$

$$\leq \tilde{C}_{1,*}(u)h^{-d}\left(\sum_{n=1}^{j} k_n h^{2r} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n, 1\leq n\leq j}\left(\|\dot{R}(t)\|^2_{L^2(\Omega)} + \|R(t)\|^2_{H^1(\Omega)}\right)$$
$$+ \tilde{C}_{2,*}\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\Theta(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \tilde{C}_{3,*}\sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\Theta(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t, \tag{4.3.69}$$

where $\tilde{C}_{1,*}$ is a positive constant depending on the solution u while $\tilde{C}_{2,*}$ and $\tilde{C}_{3,*}$ are two generic constants independent of u and the discretisation parameters. 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 $j = 1, \dots, N$,

$$\|\dot{\Theta}(t_j^-)\|^2_{L^2(\Omega)} + \|\nabla\Theta(t_j^-)\|^2_{L^2(\Omega)} + \sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\nabla\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t + \sum_{n=1}^{j}\int_{t_{n-1}}^{t_n}\|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$
$$\leq \tilde{C}(u)h^{-d}\left(\sum_{n=1}^{j} k_n h^{2r} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n, 1\leq n\leq j}\left(\|\dot{R}(t)\|^2_{L^2(\Omega)} + \|R(t)\|^2_{H^1(\Omega)}\right). \tag{4.3.70}$$

By the fundamental theorem of calculus, we have for $t \in I_j$,

$$\|\nabla\Theta(t)\|^2_{L^2(\Omega)} \leq 2\|\nabla\Theta(t_j^-)\|^2_{L^2(\Omega)} + 2\left(\int_{t_{j-1}}^{t_j}\|\nabla\dot{\Theta}(t)\|_{L^2(\Omega)}\,\mathrm{d}t\right)^2$$
$$\leq 2\tilde{C}(u)h^{-d}\left(\sum_{n=1}^{j} k_n h^{2r} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n, 1\leq n\leq j}\left(\|\dot{R}(t)\|^2_{L^2(\Omega)} + \|R(t)\|^2_{H^1(\Omega)}\right),$$

$$\|\Theta(t)\|^2_{L^2(\Omega)} \leq 2\|\Theta(t_j^-)\|^2_{L^2(\Omega)} + 2\left(\int_{t_{j-1}}^{t_j}\|\dot{\Theta}(t)\|_{L^2(\Omega)}\,\mathrm{d}t\right)^2$$
$$\leq 2\tilde{C}(u)h^{-d}\left(\sum_{n=1}^{j} k_n h^{2r} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n, 1\leq n\leq j}\left(\|\dot{R}(t)\|^2_{L^2(\Omega)} + \|R(t)\|^2_{H^1(\Omega)}\right),$$

$$\|\dot{\Theta}(t)\|^2_{L^2(\Omega)} \leq 2\|\dot{\Theta}(t_j^-)\|^2_{L^2(\Omega)} + 2\left(\int_{t_{j-1}}^{t_j}\|\ddot{\Theta}(t)\|_{L^2(\Omega)}\,\mathrm{d}t\right)^2$$
$$\leq 2\|\dot{\Theta}(t_j^-)\|^2_{L^2(\Omega)} + 2C_{\mathrm{inv}}^2 k_j^{-1}\int_{t_{j-1}}^{t_j}\|\dot{\Theta}(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t$$
$$\leq \hat{C}(u)h^{-d-2}\left(\sum_{n=1}^{j} k_n h^{2r} + \frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n, 1\leq n\leq j}\left(\|\dot{R}(t)\|^2_{L^2(\Omega)} + \|R(t)\|^2_{H^1(\Omega)}\right),$$

where we have used the assumption that $\mu_j k_j \le h^2 \le \nu_j k_j$. Thus,

$$\max_{t\in I_j}\left(\|\dot\Theta(t)\|^2_{L^2(\Omega)}+\|\Theta(t)\|^2_{H^1(\Omega)}\right)$$
$$\le C^*(u)h^{-d-2}\left(\sum_{n=1}^N k_n h^{2r}+\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n,1\le n\le N}\left(\|\dot R(t)\|^2_{L^2(\Omega)}+\|R(t)\|^2_{H^1(\Omega)}\right). \tag{4.3.71}$$

Taking maximum over $j=1,\dots,N$, we have

$$\max_{t\in I_j,1\le j\le N}\left(\|\dot\Theta(t)\|^2_{L^2(\Omega)}+\|\Theta(t)\|^2_{H^1(\Omega)}\right)$$
$$\le C^*(u)h^{-d-2}\left(\sum_{n=1}^N k_n h^{2r}+\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)\max_{t\in I_n,1\le n\le N}\left(\|\dot R(t)\|^2_{L^2(\Omega)}+\|R(t)\|^2_{H^1(\Omega)}\right). \tag{4.3.72}$$

We can choose h and k_n sufficiently small and q_n and r sufficiently large for each $n=1,\dots,N$ such that $C^*(u)h^{-d-2}\left(\sum_{n=1}^N k_n h^{2r}+\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)<1$, which is possible under our assumption that $r>\frac{d}{2}+1$, $k_n^{q_n-\frac{1}{2}}=o(h^{\frac{d}{2}})$ and $\mu_n k_n\le h^2\le \nu_n k_n$ for each $n=1,\dots,N$. This completes the proof of the contraction property.

By Banach's fixed point theorem, there exists a unique $u_{\mathrm{DG}}\in\mathcal{F}$ such that $\mathcal{N}(u_{\mathrm{DG}})=u_{\mathrm{DG}}$. By the triangle inequality, the properties of the modified projection operators and the properties of W, we have

$$\begin{aligned}
&\|u_{\mathrm{DG}}(t_j)-u(t_j^-)\|_{H^1(\Omega)}+\|\dot u_{\mathrm{DG}}(t_j)-\dot u(t_j^-)\|_{L^2(\Omega)}\\
&\le \|\theta(t_j^-)\|_{H^1(\Omega)}+\|\dot\theta(t_j^-)\|_{L^2(\Omega)}+\|W(t_j^-)-u(t_j^-)\|_{H^1(\Omega)}+\|\dot W(t_j^-)-\dot u(t_j^-)\|_{L^2(\Omega)}\\
&\le C_*(u)\left(\sum_{n=1}^j k_n h^{2r}+\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)^{\frac{1}{2}}+3C_2(u)h^r\le C(u)\left(h^{2r}+\sum_{n=1}^j\frac{k_n^{2q_n+1}}{q_n^{2(s-1)}}\right)^{\frac{1}{2}}.
\end{aligned}$$

4.4. Approximation properties of the nonlinear projection

Here we derive the properties (4.3.5)–(4.3.8) of the nonlinear projection W. We write $a\lesssim b$ if there exists a universal constant $C>0$ independent of the spatial discretisation parameter h such that $a \le Cb$.

4.4.1. Bounds on $u - W$ and W Recall that for each $t \in [0,T]$, the nonlinear projection $W(t) \in \mathcal{V}_h$ satisfies that

$$(\nabla c(W(t)), \nabla\varphi)_{L^2} = (\nabla c(u(t)), \nabla\varphi)_{L^2} \qquad \text{for all } \varphi \in \mathcal{V}_h. \tag{4.4.1}$$

We can rewrite (4.4.1) as

$$\big(c'(W(t))\nabla W(t), \nabla\varphi\big)_{L^2} = \big(c'(u(t))\nabla u(t), \nabla\varphi\big)_{L^2} \qquad \text{for all } \varphi \in \mathcal{V}_h, \tag{4.4.2}$$

for $c \in C^4(\mathbb{R})$ with $c'(\cdot) \ge M_0 > 0$. This arises from the following elliptic problem:

$$-\nabla \cdot (c'(u)\nabla u) = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega, \tag{4.4.3}$$

where f is sufficiently smooth. For each $t \in [0,T]$, $W(t)$ is the Galerkin approximation of the solution $u(t)$ of the above nonlinear Dirichlet problem. For dimensions $d \le 3$ and finite elements of degree $p \ge 2$, Douglas and Dupont [27] proved that

$$\|u - W\|_{L^2(\Omega)} + h\|\nabla u - \nabla W\|_{L^2(\Omega)} \lesssim h^{p+1}\|u\|_{H^{p+1}(\Omega)}. \tag{4.4.4}$$

Thus, we conclude that for each $t \in [0,T]$, $\frac{d}{2} + 1 < r \le \min(p, m-1)$,

$$\|u(t) - W(t)\|_{L^2(\Omega)} \lesssim h^{r+1}\|u(t)\|_{H^{r+1}(\Omega)}, \tag{4.4.5}$$

$$\|\nabla u(t) - \nabla W(t)\|_{L^2(\Omega)} \lesssim h^{r}\|u(t)\|_{H^{r+1}(\Omega)}. \tag{4.4.6}$$

Let $\mathcal{P}_h \colon H_0^1(\Omega) \to \mathcal{V}_h$ be the standard finite element interpolation operator; then we have

$$\begin{aligned}
\|u - \mathcal{P}_h u\|_{H^1(\Omega)} &\lesssim h^r |u|_{H^{r+1}(\Omega)}, \quad \frac{d}{2} + 1 < r \le \min(p, m-1),\\
\|u - \mathcal{P}_h u\|_{W^{1,\infty}(\Omega)} &\lesssim h^{r-\frac{d}{2}} |u|_{H^{r+1}(\Omega)}, \quad \frac{d}{2} + 1 < r \le \min(p, m-1).
\end{aligned}$$

See Brenner and Scott [17] for the proof of these results. Now we can bound the $W^{1,\infty}$-norm of $W(t)$ for each $t \in [0,T]$ by

$$\|W(t)\|_{W^{1,\infty}(\Omega)} \le \|W(t) - \mathcal{P}_h u(t)\|_{W^{1,\infty}(\Omega)} + \|\mathcal{P}_h u(t) - u(t)\|_{W^{1,\infty}(\Omega)} + \|u(t)\|_{W^{1,\infty}(\Omega)}$$

$$\leq C_{\rm inv} h^{-\frac{d}{2}} \|W(t) - \mathcal{P}_h u(t)\|_{H^1(\Omega)} + Ch^{r-\frac{d}{2}} + \|u(t)\|_{W^{1,\infty}(\Omega)}$$
$$\leq C_{\rm inv} h^{-\frac{d}{2}} \left(\|W(t) - u(t)\|_{H^1(\Omega)} + \|u(t) - \mathcal{P}_h u(t)\|_{H^1(\Omega)}\right) + Ch^{r-\frac{d}{2}} + \|u(t)\|_{W^{1,\infty}(\Omega)}$$
$$\leq Ch^{r-\frac{d}{2}} + \|u(t)\|_{W^{1,\infty}(\Omega)}$$
$$\leq c_0,$$

for some positive constant c_0. Note that here C is a generic positive constant, which may vary from line to line.

4.4.2. Bounds on $\dot{u} - \dot{W}$ and $\dot{W}$ We first need to show that $t \mapsto W(t)$ is differentiable with respect to t. Here we follow the proof from Section 6 of Ortner and Süli's work [60]. For $U \in \mathcal{V}_h$ and $t \in [0, T]$, we notice that the mapping $\varphi \mapsto (\nabla c(U), \nabla\varphi)_{L^2} - (\nabla c(u(t)), \nabla\varphi)_{L^2}$ is a bounded linear functional on $\mathcal{V}_h$; hence by the *Riesz representation theorem*, there exists a unique $\mathcal{A}(t, U) \in \mathcal{V}_h$ such that

$$(\mathcal{A}(t, U), \varphi) = (\nabla c(U), \nabla\varphi)_{L^2} - (\nabla c(u(t)), \nabla\varphi)_{L^2} . \tag{4.4.7}$$

Here $(\cdot, \cdot)$ is the inner product associated with the Hilbert space $\mathcal{V}_h \subset H^1_0(\Omega)$. That is,

$$(\nabla\mathcal{A}(t, U), \nabla\varphi)_{L^2} = (\nabla c(U), \nabla\varphi)_{L^2} - (\nabla c(u(t)), \nabla\varphi)_{L^2} . \tag{4.4.8}$$

It follows from the linearisation process that the derivative of $(t, U) \mapsto \mathcal{A}(t, U)$ with respect to U, evaluated at $U = W(t)$, exists and is invertible. We also have $\mathcal{A}(t, W(t)) = 0$. Since $t \mapsto u(t)$ 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 $t_0 \in (0, T)$. It can be therefore deduced from the *implicit function theorem* that $t \mapsto W(t)$ is differentiable in $(0, T)$.

Before we proceed to the proof of the L^2-estimate of $\dot{u} - \dot{W}$, we first note that since $c \in C^4(\mathbb{R})$, c', c'' and c''' are locally Lipschitz continuous. That is, there exists a positive constant $C_{\rm Lip} < \infty$ such that

$$|c'(u(t)) - c'(W(t))| \leq C_{\rm Lip} |u(t) - W(t)|, \tag{4.4.9}$$

$$|c''(u(t)) - c''(W(t))| \leq C_{\text{Lip}}|u(t) - W(t)|, \tag{4.4.10}$$

$$|c'''(u(t)) - c'''(W(t))| \leq C_{\text{Lip}}|u(t) - W(t)|, \tag{4.4.11}$$

where $C_{\text{Lip}} = \max_{\xi} \left(|c''(\xi)|, |c'''(\xi)|, |c''''(\xi)|\right)$, with $|\xi| \leq \max_{t\in[0,T]} \left(\|u\|_{L^\infty(\Omega)} + \|W\|_{L^\infty(\Omega)}\right)$.

Now, by differentiating (4.4.2) with respect to t, we have

$$\begin{aligned}
&\Big(c'(W(t))\nabla\dot{W}(t), \nabla\varphi\Big)_{L^2} + \Big(c''(W(t))\dot{W}(t)\nabla W(t), \nabla\varphi\Big)_{L^2} \\
&= \Big(c'(u(t))\nabla\dot{u}(t), \nabla\varphi\Big)_{L^2} + \Big(c''(u(t))\dot{u}(t)\nabla u(t), \nabla\varphi\Big)_{L^2} \quad \text{for all } \varphi \in \mathcal{V}_h.
\end{aligned} \tag{4.4.12}$$

Using the assumption that $c'(\cdot) \geq M_0 > 0$, we have

$$\begin{aligned}
M_0\|\nabla\dot{W}(t) - \nabla\dot{u}(t)\|^2_{L^2(\Omega)} &\leq \Big(c'(W(t))(\nabla\dot{W}(t) - \nabla\dot{u}(t)), \nabla\dot{W}(t) - \nabla\dot{u}(t)\Big)_{L^2} \\
&= \Big(c'(W(t))(\nabla\dot{W}(t) - \nabla\dot{u}(t)), \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} \\
&\quad + \Big(c'(W(t))(\nabla\dot{W}(t) - \nabla\dot{u}(t)), \nabla\mathcal{P}_h\dot{u}(t) - \nabla\dot{u}(t)\Big)_{L^2} \\
&:= T_1 + T_2.
\end{aligned}$$

Since $\nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t) \in \mathcal{V}_h$, we can apply (4.4.12) to have

$$\begin{aligned}
T_1 &= \Big(c'(W(t))(\nabla\dot{W}(t) - \nabla\dot{u}(t)), \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} \\
&= \Big(c'(u(t))\nabla\dot{u}(t), \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} - \Big(c'(W(t))\nabla\dot{u}(t), \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} \\
&\quad + \Big(c''(u(t))\dot{u}(t)\nabla u(t), \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} - \Big(c''(W(t))\dot{W}(t)\nabla W(t), \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} \\
&= \Big([c'(u(t)) - c'(W(t)]\nabla\dot{u}(t), \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} \\
&\quad + \Big(c''(u(t))\dot{u}(t)[\nabla u(t) - \nabla W(t)], \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} \\
&\quad + \Big(c''(u(t))\dot{u}(t)\nabla W(t), \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} - \Big(c''(W(t))\dot{W}(t)\nabla W(t), \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} \\
&= \Big([c'(u(t)) - c'(W(t)]\nabla\dot{u}(t), \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} \\
&\quad + \Big(c''(u(t))\dot{u}(t)[\nabla u(t) - \nabla W(t)], \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2} \\
&\quad + \Big([c''(u(t)) - c''(W(t))]\dot{u}(t)\nabla W(t), \nabla\dot{W}(t) - \nabla\mathcal{P}_h\dot{u}(t)\Big)_{L^2}
\end{aligned}$$

$$
\begin{aligned}
&+ \Big(c''(W(t))[\dot u(t) - \dot W(t)]\nabla W(t), \nabla \dot W(t) - \nabla \mathcal{P}_h \dot u(t)\Big)_{L^2} \\
&\le C_{\mathrm{Lip}} \|u(t) - W(t)\|_{L^2(\Omega)} \|\nabla \dot u(t)\|_{L^\infty(\Omega)} \|\nabla \dot W(t) - \nabla \mathcal{P}_h \dot u(t)\|_{L^2(\Omega)} \\
&+ C_{\mathrm{Lip}} \|\dot u(t)\|_{L^\infty(\Omega)} \|\nabla u(t) - \nabla W(t)\|_{L^2(\Omega)} \|\nabla \dot W(t) - \nabla \mathcal{P}_h \dot u(t)\|_{L^2(\Omega)} \\
&+ C_{\mathrm{Lip}} \|u(t) - W(t)\|_{L^2(\Omega)} \|\dot u(t)\|_{L^\infty(\Omega)} \|\nabla W(t)\|_{L^\infty(\Omega)} \|\nabla \dot W(t) - \nabla \mathcal{P}_h \dot u(t)\|_{L^2(\Omega)} \\
&+ C_{\mathrm{Lip}} \|\dot u(t) - \dot W(t)\|_{L^2(\Omega)} \|\nabla W(t)\|_{L^\infty(\Omega)} \|\nabla \dot W(t) - \nabla \mathcal{P}_h \dot u(t)\|_{L^2(\Omega)} \\
&\le C \left(h^r + \|\dot W(t) - \dot u(t)\|_{L^2(\Omega)}\right) \|\nabla \dot W(t) - \nabla \mathcal{P}_h \dot u(t)\|_{L^2(\Omega)} \\
&\le C \left(h^r + \|\dot W(t) - \dot u(t)\|_{L^2(\Omega)}\right) \|\nabla \dot W(t) - \nabla \dot u(t)\|_{L^2(\Omega)} \\
&+ C \left(h^r + \|\dot W(t) - \dot u(t)\|_{L^2(\Omega)}\right) \|\nabla \mathcal{P}_h \dot u(t) - \nabla \dot u(t)\|_{L^2(\Omega)} \\
&\le C \left(h^r + \|\dot W(t) - \dot u(t)\|_{L^2(\Omega)}\right) \|\nabla \dot W(t) - \nabla \dot u(t)\|_{L^2(\Omega)} + C \left(h^{2r} + h^r \|\dot W(t) - \dot u(t)\|_{L^2(\Omega)}\right).
\end{aligned}
$$

Note that

$$
\begin{aligned}
T_2 &= \Big(c'(W(t))(\nabla \dot W(t) - \nabla \dot u(t)), \nabla \mathcal{P}_h \dot u(t) - \nabla \dot u(t)\Big)_{L^2} \\
&\le C_c \|\nabla \dot W(t) - \nabla \dot u(t)\|_{L^2(\Omega)} \|\nabla \mathcal{P}_h \dot u(t) - \nabla \dot u(t)\|_{L^2(\Omega)} \\
&\le C h^r \|\nabla \dot W(t) - \nabla \dot u(t)\|_{L^2(\Omega)}.
\end{aligned}
$$

Combining the estimates for T_1 and T_2 we have

$$
\begin{aligned}
M_0 \|\nabla \dot W(t) - \nabla \dot u(t)\|^2_{L^2(\Omega)} &\le C \left(h^r + \|\dot W(t) - \dot u(t)\|_{L^2(\Omega)}\right) \|\nabla \dot W(t) - \nabla \dot u(t)\|_{L^2(\Omega)} \\
&\quad + C \left(h^{2r} + h^r \|\dot W(t) - \dot u(t)\|_{L^2(\Omega)}\right) \\
&\le \frac{M_0}{2} \|\nabla \dot W(t) - \nabla \dot u(t)\|^2_{L^2(\Omega)} + C h^{2r} + C \|\dot W(t) - \dot u(t)\|^2_{L^2(\Omega)}.
\end{aligned} \tag{4.4.13}
$$

This implies that

$$
\|\nabla \dot W(t) - \nabla \dot u(t)\|_{L^2(\Omega)} \le C h^r + C \|\dot W(t) - \dot u(t)\|_{L^2(\Omega)}, \text{ for } 1 + \frac{d}{2} < r \le \min(p, m-1). \tag{4.4.14}
$$

Now we focus on deriving the L^2-bound on $\dot u(t) - \dot W(t)$ using a duality argument. Following the ideas introduced in [27], we consider the formal derivative,

$$L\varphi = -\nabla \cdot (c'(u)\nabla\varphi + \varphi c''(u)\nabla u) \tag{4.4.15}$$

of the elliptic operator in (4.4.3). The formal adjoint of L is defined as

$$L^*\psi := -\nabla \cdot (c'(u)\nabla\psi) + c''(u)\nabla u \cdot \nabla\psi. \tag{4.4.16}$$

It follows from the *weak maximum principle* (cf. Theorem 8.1 and Corollary 8.2 from [38]) that the following homogeneous Dirichlet boundary value problem

$$L^*\psi = 0 \text{ in } \Omega, \quad \psi = 0 \text{ on } \partial\Omega \tag{4.4.17}$$

has only the trival solution $\psi \equiv 0$ in Ω. By the *Fredholm alternative* (see Theorem 4, Section 6.2.3 from [36] for more details), for any $\xi \in L^2(\Omega)$, there exists a unique $\psi \in H^2(\Omega)$ such that

$$L^*\psi = \xi \text{ in } \Omega, \quad \psi = 0 \text{ on } \partial\Omega. \tag{4.4.18}$$

By the well-known elliptic regularity (cf. Theorem 8.8 from [38]), there exists a constant C independent of ξ such that

$$\|\psi\|_{H^2(\Omega)} \le C\|\xi\|_{L^2(\Omega)}. \tag{4.4.19}$$

Taking $\xi = \dot{u} - \dot{W}$ in (4.4.18), we have

$$\begin{aligned}
\|\dot{u}(t) - \dot{W}(t)\|^2_{L^2(\Omega)} &= \left(\dot{u}(t) - \dot{W}(t), L^*\psi\right)_{L^2} \\
&= \left(L(\dot{u}(t) - \dot{W}(t)), \psi\right)_{L^2} \\
&= \left(c'(u)[\nabla\dot{u}(t) - \nabla\dot{W}(t)], \nabla\psi\right)_{L^2} + \left([\dot{u}(t) - \dot{W}(t)]c''(u)\nabla u, \nabla\psi\right)_{L^2} \\
&= \left(c'(u)\nabla\dot{u}(t), \nabla\psi\right)_{L^2} - \left(c'(W)\nabla\dot{W}(t), \nabla\psi\right)_{L^2} \\
&\quad + \left([c'(W) - c'(u)]\nabla\dot{W}(t), \nabla\psi\right)_{L^2} + \left([\dot{u}(t) - \dot{W}(t)]c''(u)\nabla u, \nabla\psi\right)_{L^2}.
\end{aligned}$$

Now let $\psi_h \in \mathcal{V}_h$ be the Galerkin approximation of ψ; then it is known that we have

$$\|\psi - \psi_h\|_{L^2(\Omega)} + h\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \le Ch^2\|\psi\|_{H^2(\Omega)}. \tag{4.4.20}$$

From (4.4.12), we know that

$$\begin{aligned}
&\big(c'(u)\nabla\dot{u}(t), \nabla\psi_h\big)_{L^2} - \Big(c'(W)\nabla\dot{W}(t), \nabla\psi_h\Big)_{L^2} \\
&= \Big(c''(W)\dot{W}(t)\nabla W(t), \nabla\psi_h\Big)_{L^2} - \big(c''(u)\dot{u}(t)\nabla u(t), \nabla\psi_h\big)_{L^2}.
\end{aligned} \tag{4.4.21}$$

Then, we have

$$\begin{aligned}
&\|\dot{u}(t) - \dot{W}(t)\|^2_{L^2(\Omega)} \\
&= \big(c'(u)\nabla\dot{u}(t), \nabla\psi - \nabla\psi_h\big)_{L^2} - \Big(c'(W)\nabla\dot{W}(t), \nabla\psi - \nabla\psi_h\Big)_{L^2} \\
&\quad + \Big(c''(W)\dot{W}(t)\nabla W(t), \nabla\psi_h\Big)_{L^2} - \big(c''(u)\dot{u}(t)\nabla u(t), \nabla\psi_h\big)_{L^2} \\
&\quad + \Big([c'(W) - c'(u)]\nabla\dot{W}(t), \nabla\psi\Big)_{L^2} + \Big([\dot{u}(t) - \dot{W}(t)]c''(u)\nabla u, \nabla\psi\Big)_{L^2} \\
&= \Big(c'(W)[\nabla\dot{u}(t) - \nabla\dot{W}(t)], \nabla\psi - \nabla\psi_h\Big)_{L^2} + \big([c'(u) - c'(W)]\nabla\dot{u}(t), \nabla\psi - \nabla\psi_h\big)_{L^2} \\
&\quad + \Big([c'(W) - c'(u)]\nabla\dot{W}(t), \nabla\psi\Big)_{L^2} + \big(c''(u)\dot{u}(t)[\nabla u(t) - \nabla W(t)], \nabla\psi - \nabla\psi_h\big)_{L^2} \\
&\quad + \Big([c''(u)\dot{u}(t) - c''(W)\dot{W}(t)]\nabla W(t), \nabla\psi - \nabla\psi_h\Big)_{L^2} \\
&\quad + \Big(c''(W)\dot{W}(t)\nabla W(t) - c''(u)\dot{W}(t)\nabla u(t), \nabla\psi\Big)_{L^2} \\
&:= T_3 + T_4 + T_5 + T_6 + T_7 + T_8.
\end{aligned}$$

Note that

$$\begin{aligned}
T_3 &= \Big(c'(W)[\nabla\dot{u}(t) - \nabla\dot{W}(t)], \nabla\psi - \nabla\psi_h\Big)_{L^2} \\
&\le C_c\|\nabla\dot{u}(t) - \nabla\dot{W}(t)\|_{L^2(\Omega)}\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \\
&\le C\|\nabla\dot{u}(t) - \nabla\dot{W}(t)\|_{L^2(\Omega)} h\|\psi\|_{H^2(\Omega)} \\
&\le Ch\|\nabla\dot{u}(t) - \nabla\dot{W}(t)\|_{L^2(\Omega)}\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)}.
\end{aligned}$$

$$T_4 = \big([c'(u) - c'(W)]\nabla\dot{u}(t), \nabla\psi - \nabla\psi_h\big)_{L^2}$$

$$\le C_{\text{Lip}}\|u(t)-W(t)\|_{L^2(\Omega)}\|\nabla\dot{u}(t)\|_{L^\infty(\Omega)}\|\nabla\psi-\nabla\psi_h\|_{L^2(\Omega)}$$
$$\le C\|u(t)-W(t)\|_{L^2(\Omega)}h\|\psi\|_{H^2(\Omega)}$$
$$\le Ch^{r+2}\|\dot{u}(t)-\dot{W}(t)\|_{L^2(\Omega)}.$$

$$\begin{aligned}T_5 &= \Big([c'(W)-c'(u)]\nabla\dot{W}(t),\nabla\psi\Big)_{L^2}\\ &\le C_{\text{Lip}}\|W(t)-u(t)\|_{L^\infty(\Omega)}\|\nabla\dot{W}(t)-\nabla\dot{u}\|_{L^2(\Omega)}\|\nabla\psi\|_{L^2(\Omega)}\\ &\quad + C_{\text{Lip}}\|W(t)-u(t)\|_{L^2(\Omega)}\|\nabla\dot{u}\|_{L^\infty(\Omega)}\|\nabla\psi\|_{L^2(\Omega)}.\end{aligned}$$

It was proved by Nitsche [59] that

$$\|W(t)-u(t)\|_{L^\infty(\Omega)}\le Ch^s, \tag{4.4.22}$$

with $s=\min(n+\sigma,p+1)$, where p is the polynomial degree of the finite element space, n is an integer and $0\le\sigma\le 1$ such that $u\in C^{n,\sigma}(\overline{\Omega})$ (i.e., u has continuous derivatives up to the order n which are Hölder-continuous with an exponent σ). Note that in our case, $u\in H^m(\Omega)$ with $m>\frac{d}{2}+2$. By Sobolev embedding theorem, we have $u\in C^{m-1-[\frac{d}{2}],\sigma}(\overline{\Omega})$, with

$$\sigma=\begin{cases}[\frac{d}{2}]+1-\frac{d}{2} & \text{if } \frac{d}{2}\notin\mathbb{Z},\\ \text{any element in } (0,1) & \text{if } \frac{d}{2}\in\mathbb{Z}.\end{cases}$$

This implies that $s\ge m-2>1$ for $d=2,3$. Thus

$$T_5\le Ch\|\nabla\dot{W}(t)-\nabla\dot{u}(t)\|_{L^2(\Omega)}\|\dot{u}(t)-\dot{W}(t)\|_{L^2(\Omega)}+Ch^{r+1}\|\dot{u}(t)-\dot{W}(t)\|_{L^2(\Omega)}.$$

$$\begin{aligned}T_6 &= \big(c''(u)\dot{u}(t)[\nabla u(t)-\nabla W(t)],\nabla\psi-\nabla\psi_h\big)_{L^2}\\ &\le C_{\text{Lip}}\|\dot{u}(t)\|_{L^\infty(\Omega)}\|\nabla u(t)-\nabla W(t)\|_{L^2(\Omega)}\|\nabla\psi-\nabla\psi_h\|_{L^2(\Omega)}\\ &\le C\|\nabla u(t)-\nabla W(t)\|_{L^2(\Omega)}h\|\psi\|_{H^2(\Omega)}\\ &\le Ch^{r+1}\|\dot{u}(t)-\dot{W}(t)\|_{L^2(\Omega)}.\end{aligned}$$

$$
\begin{aligned}
T_7 &= \Big([c''(u)\dot{u}(t) - c''(W)\dot{W}(t)]\nabla W(t), \nabla\psi - \nabla\psi_h\Big)_{L^2} \\
&= \big([c''(u)\dot{u}(t) - c''(W)\dot{u}(t)]\nabla W(t), \nabla\psi - \nabla\psi_h\big)_{L^2} + \Big(c''(W)[\dot{u}(t) - \dot{W}(t)]\nabla W(t), \nabla\psi - \nabla\psi_h\Big)_{L^2} \\
&\leq C_{\mathrm{Lip}}\|u(t) - W(t)\|_{L^2(\Omega)}\|\dot{u}(t)\|_{L^\infty(\Omega)}\|\nabla W(t)\|_{L^\infty(\Omega)}\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \\
&\quad + C_{\mathrm{Lip}}\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)}\|\nabla W\|_{L^\infty(\Omega)}\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \\
&\leq C\|u(t) - W(t)\|_{L^2(\Omega)}h\|\psi\|_{H^2(\Omega)} + C\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)}h\|\psi\|_{H^2(\Omega)} \\
&\leq Ch^{r+2}\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)} + Ch\|\dot{u}(t) - \dot{W}(t)\|^2_{L^2(\Omega)}.
\end{aligned}
$$

$$
\begin{aligned}
T_8 &= \Big(c''(W)\dot{W}(t)\nabla W(t) - c''(u)\dot{W}(t)\nabla u(t), \nabla\psi\Big)_{L^2} \\
&= \Big([c''(W) - c''(u)][\dot{W}(t) - \dot{u}(t)]\nabla W(t), \nabla\psi\Big)_{L^2} + \big([c''(W) - c''(u)]\dot{u}(t)\nabla W(t), \nabla\psi\big)_{L^2} \\
&\quad + \Big(c''(u)[\dot{W}(t) - \dot{u}(t)][\nabla W(t) - \nabla u(t)], \nabla\psi\Big)_{L^2} + \big(c''(u)\dot{u}(t)[\nabla W(t) - \nabla u(t)], \nabla\psi\big)_{L^2} \\
&\leq C_{\mathrm{Lip}}\|W(t) - u(t)\|_{L^\infty(\Omega)}\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)}\|\nabla W\|_{L^\infty(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\quad + C_{\mathrm{Lip}}\|W(t) - u(t)\|_{L^2(\Omega)}\|\dot{u}(t)\|_{L^\infty(\Omega)}\|\nabla W\|_{L^\infty(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\quad + \Big(c''(u)[\dot{W}(t) - \dot{u}(t)][\nabla W(t) - \nabla u(t)], \nabla\psi\Big)_{L^2} \\
&\quad + C_{\mathrm{Lip}}\|\dot{u}(t)\|_{L^\infty(\Omega)}\|\nabla W(t) - \nabla u(t)\|_{L^2(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\leq \Big(c''(u)[\dot{W}(t) - \dot{u}(t)][\nabla W(t) - \nabla u(t)], \nabla\psi\Big)_{L^2} \\
&\quad + Ch\|\dot{u}(t) - \dot{W}(t)\|^2_{L^2(\Omega)} + Ch^{r+1}\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)} + Ch^r\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)}.
\end{aligned}
$$

Note that for $d = 2, 3$, by Sobolev embedding theorem and the Gagliardo-Nirenberg interpolation inequality

$$
\|\varphi\|_{L^3(\Omega)} \leq C\|\varphi\|^{\frac{1}{2}}_{L^2(\Omega)}\|\nabla\varphi\|^{\frac{1}{2}}_{L^2(\Omega)}, \tag{4.4.23}
$$

we have

$$
\begin{aligned}
&\Big(c''(u)[\dot{W}(t) - \dot{u}(t)][\nabla W(t) - \nabla u(t)], \nabla\psi\Big)_{L^2} \\
&\leq C_{\mathrm{Lip}}\|\dot{u}(t) - \dot{W}(t)\|_{L^3(\Omega)}\|\nabla W(t) - \nabla u(t)\|_{L^2(\Omega)}\|\nabla\psi\|_{L^6(\Omega)}
\end{aligned}
$$

$$
\begin{aligned}
&\le Ch^r \|\dot u(t) - \dot W(t)\|_{L^2(\Omega)}^{\frac12} \|\nabla \dot u(t) - \nabla \dot W(t)\|_{L^2(\Omega)}^{\frac12} \|\psi\|_{H^2(\Omega)} \\
&\le Ch^r \|\dot u(t) - \dot W(t)\|_{L^2(\Omega)}^{\frac32} \|\nabla \dot u(t) - \nabla \dot W(t)\|_{L^2(\Omega)}^{\frac12}.
\end{aligned}
$$

For $d = 1$, $H^2(\Omega)$ is compactly embedded into $W^{1,\infty}(\Omega)$, thus

$$
\begin{aligned}
&\Big(c''(u)[\dot W(t) - \dot u(t)][\nabla W(t) - \nabla u(t)], \nabla \psi\Big)_{L^2} \\
&\le C_{\mathrm{Lip}} \|\dot u(t) - \dot W(t)\|_{L^2(\Omega)} \|\nabla W(t) - \nabla u(t)\|_{L^2(\Omega)} \|\nabla \psi\|_{L^\infty(\Omega)} \\
&\le Ch^r \|\dot u(t) - \dot W(t)\|_{L^2(\Omega)}.
\end{aligned}
$$

Combining these two cases, we have that

$$
\begin{aligned}
T_8 \le{}& Ch\|\dot u(t) - \dot W(t)\|_{L^2(\Omega)}^2 + Ch^{r+1}\|\dot u(t) - \dot W(t)\|_{L^2(\Omega)} + Ch^r\|\dot u(t) - \dot W(t)\|_{L^2(\Omega)} \\
&+ Ch^r \|\dot u(t) - \dot W(t)\|_{L^2(\Omega)}^{\frac32} \|\nabla \dot u(t) - \nabla \dot W(t)\|_{L^2(\Omega)}^{\frac12}.
\end{aligned}
$$

Combining the estimates for T_3–T_8, we have

$$
\begin{aligned}
\|\dot u(t) - \dot W(t)\|_{L^2(\Omega)} &\le Ch^r + Ch\|\dot u(t) - \dot W(t)\|_{L^2(\Omega)} + Ch\|\nabla \dot u(t) - \nabla \dot W(t)\|_{L^2(\Omega)} \\
&\quad + Ch^r \|\dot u(t) - \dot W(t)\|_{L^2(\Omega)}^{\frac12} \|\nabla \dot u(t) - \nabla \dot W(t)\|_{L^2(\Omega)}^{\frac12} \\
&\le Ch^r + Ch\|\dot u(t) - \dot W(t)\|_{L^2(\Omega)} + Ch\|\nabla \dot u(t) - \nabla \dot W(t)\|_{L^2(\Omega)} \\
&\quad + Ch^r \Big(\|\dot u(t) - \dot W(t)\|_{L^2(\Omega)} + \|\nabla \dot u(t) - \nabla \dot W(t)\|_{L^2(\Omega)}\Big).
\end{aligned}
$$

This implies that

$$
\|\dot u(t) - \dot W(t)\|_{L^2(\Omega)} \le Ch^r + Ch\|\nabla \dot u(t) - \nabla \dot W(t)\|_{L^2(\Omega)}. \tag{4.4.24}
$$

Substituting (4.4.24) into (4.4.14), we have

$$
\|\nabla \dot u(t) - \nabla \dot W(t)\|_{L^2(\Omega)} \le Ch^r. \tag{4.4.25}
$$

If we substitute (4.4.25) back to (4.4.24), we have

$$
\|\dot u(t) - \dot W(t)\|_{L^2(\Omega)} \le Ch^r. \tag{4.4.26}
$$

The L^2 error estimate obtained here is not of optimal order, but it is sufficiently sharp for our purposes. Analogously, we have for each $t \in [0,T]$ that

$$\begin{aligned}
\|\dot{W}(t)\|_{W^{1,\infty}(\Omega)} &\le \|\dot{W}(t) - \mathcal{P}_h\dot{u}(t)\|_{W^{1,\infty}(\Omega)} + \|\mathcal{P}_h\dot{u}(t) - \dot{u}(t)\|_{W^{1,\infty}(\Omega)} + \|\dot{u}(t)\|_{W^{1,\infty}(\Omega)} \\
&\le C_{\mathrm{inv}} h^{-\frac{d}{2}} \|\dot{W}(t) - \mathcal{P}_h\dot{u}(t)\|_{H^1(\Omega)} + Ch^{r-\frac{d}{2}} + \|\dot{u}(t)\|_{W^{1,\infty}(\Omega)} \\
&\le C_{\mathrm{inv}} h^{-\frac{d}{2}} \left(\|\dot{W}(t) - \dot{u}(t)\|_{H^1(\Omega)} + \|\dot{u}(t) - \mathcal{P}_h\dot{u}(t)\|_{H^1(\Omega)}\right) + Ch^{r-\frac{d}{2}} + \|\dot{u}(t)\|_{W^{1,\infty}(\Omega)} \\
&\le Ch^{r-\frac{d}{2}} + \|\dot{u}(t)\|_{W^{1,\infty}(\Omega)} \\
&\le c_0.
\end{aligned}$$

4.4.3. Bounds on $\ddot{u} - \ddot{W}$ By an analogous argument as in the previous section, we find that $\dot{W}(t)$ is differentiable with respect to t and a similar L^2-error estimate holds. By differentiating (4.4.12), we have

$$\begin{aligned}
&\left(c'(W(t))\nabla\ddot{W}(t), \nabla\varphi\right)_{L^2} + 2\left(c''(W(t))\dot{W}(t)\nabla\dot{W}(t), \nabla\varphi\right)_{L^2} + \left(c''(W(t))\ddot{W}(t)\nabla W(t), \nabla\varphi\right)_{L^2} \\
&\quad + \left(c'''(W(t))\dot{W}(t)^2\nabla W(t), \nabla\varphi\right)_{L^2} \\
&= \left(c'(u(t))\nabla\ddot{u}(t), \nabla\varphi\right)_{L^2} + 2\left(c''(u(t))\dot{u}(t)\nabla\dot{u}(t), \nabla\varphi\right)_{L^2} + \left(c''(u(t))\ddot{u}(t)\nabla u(t), \nabla\varphi\right)_{L^2} \\
&\quad + \left(c'''(u(t))\dot{u}(t)^2\nabla u(t), \nabla\varphi\right)_{L^2} \quad \text{for all } \varphi \in \mathcal{V}_h. \qquad (4.4.27)
\end{aligned}$$

Now we derive an L^2-bound on $\ddot{u} - \ddot{W}$ as in the previous section. In particular, we take $\xi = \ddot{u} - \ddot{W}$ in the adjoint problem (4.4.18); then we have

$$\begin{aligned}
&\|\ddot{u}(t) - \ddot{W}(t)\|^2_{L^2(\Omega)} = \left(\ddot{u}(t) - \ddot{W}(t), L^*\psi\right)_{L^2} \\
&= \left(L(\ddot{u}(t) - \ddot{W}(t)), \psi\right)_{L^2} \\
&= \left(c'(u)[\nabla\ddot{u}(t) - \nabla\ddot{W}(t)], \nabla\psi\right)_{L^2} + \left([\ddot{u}(t) - \ddot{W}(t)]c''(u)\nabla u(t), \nabla\psi\right)_{L^2} \\
&= \left(c'(u)\nabla\ddot{u}(t), \nabla\psi\right)_{L^2} - \left(c'(W)\nabla\ddot{W}(t), \nabla\psi\right)_{L^2} + \left([c'(W) - c'(u)]\nabla\ddot{W}(t), \nabla\psi\right)_{L^2} \\
&\quad + \left([\ddot{u}(t) - \ddot{W}(t)]c''(u)\nabla u(t), \nabla\psi\right)_{L^2}. \qquad (4.4.28)
\end{aligned}$$

Again, we let $\psi_h \in \mathcal{V}_h$ be the Galerkin approximation of ψ; then it satisfies that

$$\|\psi - \psi_h\|_{L^2(\Omega)} + h\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \le Ch^2\|\psi\|_{H^2(\Omega)}. \tag{4.4.29}$$

From (4.4.27), we know that, for $\psi_h \in \mathcal{V}_h$,

$$\begin{aligned}
&\left(c'(W(t))\nabla\ddot{W}(t), \nabla\psi_h\right)_{L^2} - \left(c'(u(t))\nabla\ddot{u}(t), \nabla\psi_h\right)_{L^2} \\
&= 2\left(c''(u(t))\dot{u}(t)\nabla\dot{u}(t), \nabla\psi_h\right)_{L^2} - 2\left(c''(W(t))\dot{W}(t)\nabla\dot{W}(t), \nabla\psi_h\right)_{L^2} \\
&\quad + \left(c''(u(t))\ddot{u}(t)\nabla u(t), \nabla\psi_h\right)_{L^2} - \left(c''(W(t))\ddot{W}(t)\nabla W(t), \nabla\psi_h\right)_{L^2} \\
&\quad + \left(c'''(u(t))\dot{u}(t)^2\nabla u(t), \nabla\psi_h\right)_{L^2} - \left(c'''(W(t))\dot{W}(t)^2\nabla W(t), \nabla\psi_h\right)_{L^2}.
\end{aligned} \tag{4.4.30}$$

Substituting (4.4.30) into (4.4.28) gives

$$\begin{aligned}
&\|\ddot{u}(t) - \ddot{W}(t)\|^2_{L^2(\Omega)} \\
&= \left(c'(u)\nabla\ddot{u}(t), \nabla\psi - \nabla\psi_h\right)_{L^2} - \left(c'(W)\nabla\ddot{W}(t), \nabla\psi - \nabla\psi_h\right)_{L^2} + 2\left(c''(W)\dot{W}(t)\nabla\dot{W}(t), \nabla\psi_h\right)_{L^2} \\
&\quad - 2\left(c''(u)\dot{u}(t)\nabla\dot{u}(t), \nabla\psi_h\right)_{L^2} + \left(c''(W)\ddot{W}(t)\nabla W(t), \nabla\psi_h\right)_{L^2} - \left(c''(u)\ddot{u}(t)\nabla u(t), \nabla\psi_h\right)_{L^2} \\
&\quad + \left(c'''(W)\dot{W}(t)^2\nabla W(t), \nabla\psi_h\right)_{L^2} - \left(c'''(u)\dot{u}(t)^2\nabla u(t), \nabla\psi_h\right)_{L^2} + \left([c'(W) - c'(u)]\nabla\ddot{W}(t), \nabla\psi\right)_{L^2} \\
&\quad + \left([\ddot{u}(t) - \ddot{W}(t)]c''(u)\nabla u(t), \nabla\psi\right)_{L^2} \\
&= \left([c'(u) - c'(W)]\nabla\ddot{u}(t), \nabla\psi - \nabla\psi_h\right)_{L^2} + \left(c'(W)[\nabla\ddot{u}(t) - \nabla\ddot{W}(t)], \nabla\psi - \nabla\psi_h\right)_{L^2} \\
&\quad + \left([c'(W) - c'(u)]\nabla\ddot{W}(t), \nabla\psi\right)_{L^2} + \left(c''(u)\ddot{u}(t)[\nabla u(t) - \nabla W(t)], \nabla\psi - \nabla\psi_h\right)_{L^2} \\
&\quad + \left([c''(u)\ddot{u} - c''(W)\ddot{W}(t)]\nabla W(t), \nabla\psi - \nabla\psi_h\right)_{L^2} + \left([c''(W)\nabla W(t) - c''(u)\nabla u(t)]\ddot{W}(t), \nabla\psi\right)_{L^2} \\
&\quad + 2\left(c''(u)\dot{u}(t)\nabla\dot{u}(t) - c''(W)\dot{W}(t)\nabla\dot{W}(t), \nabla\psi - \nabla\psi_h\right)_{L^2} \\
&\quad + 2\left(c''(W)\dot{W}(t)\nabla\dot{W}(t) - c''(u)\dot{u}(t)\nabla\dot{u}(t), \nabla\psi\right)_{L^2} \\
&\quad + \left(c'''(u)\dot{u}(t)^2\nabla u(t) - c'''(W)\dot{W}(t)^2\nabla W(t), \nabla\psi - \nabla\psi_h\right)_{L^2} \\
&\quad + \left(c'''(W)\dot{W}(t)^2\nabla W(t) - c'''(u)\dot{u}(t)^2\nabla u(t), \nabla\psi\right)_{L^2} \\
&:= E_1 + E_2 + E_3 + E_4 + E_5 + E_6 + E_7 + E_8 + E_9 + E_{10}.
\end{aligned}$$

Note that

$$
\begin{aligned}
E_1 &= \big([c'(u)-c'(W)]\nabla\ddot{u}(t), \nabla\psi-\nabla\psi_h\big)_{L^2} \\
&\le C_{\mathrm{Lip}}\|u(t)-W(t)\|_{L^2(\Omega)}\|\nabla\ddot{u}(t)\|_{L^\infty(\Omega)}\|\nabla\psi-\nabla\psi_h\|_{L^2(\Omega)} \\
&\le C\|u(t)-W(t)\|_{L^2(\Omega)}\|\nabla\ddot{u}(t)\|_{L^\infty(\Omega)}h\|\psi\|_{H^2(\Omega)} \\
&\le Ch^{r+2}\|\ddot{u}(t)-\ddot{W}(t)\|_{L^2(\Omega)},
\end{aligned}
$$

$$
\begin{aligned}
E_2 &= \Big(c'(W)[\nabla\ddot{u}(t)-\nabla\ddot{W}(t)], \nabla\psi-\nabla\psi_h\Big)_{L^2} \\
&\le C_c\|\nabla\ddot{u}(t)-\nabla\ddot{W}(t)\|_{L^2(\Omega)}\|\nabla\psi-\nabla\psi_h\|_{L^2(\Omega)} \\
&\le C\|\nabla\ddot{u}(t)-\nabla\ddot{W}(t)\|_{L^2(\Omega)}h\|\psi\|_{H^2(\Omega)} \\
&\le Ch\|\nabla\ddot{u}(t)-\nabla\ddot{W}(t)\|_{L^2(\Omega)}\|\ddot{u}-\ddot{W}\|_{L^2(\Omega)},
\end{aligned}
$$

$$
\begin{aligned}
E_3 &= \Big([c'(W)-c'(u)]\nabla\ddot{W}(t), \nabla\psi\Big)_{L^2} \\
&= \Big([c'(W)-c'(u)][\nabla\ddot{W}(t)-\nabla\ddot{u}(t)], \nabla\psi\Big)_{L^2} + \big([c'(W)-c'(u)]\nabla\ddot{u}(t), \nabla\psi\big)_{L^2} \\
&\le C_{\mathrm{Lip}}\|W(t)-u(t)\|_{L^\infty(\Omega)}\|\nabla\ddot{W}(t)-\nabla\ddot{u}(t)\|_{L^2(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\quad + C_{\mathrm{Lip}}\|W(t)-u(t)\|_{L^2(\Omega)}\|\nabla\ddot{u}\|_{L^\infty(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\le Ch\|\nabla\ddot{W}(t)-\nabla\ddot{u}(t)\|_{L^2}\|\ddot{W}(t)-\ddot{u}(t)\|_{L^2(\Omega)} + Ch^{r+1}\|\ddot{W}(t)-\ddot{u}(t)\|_{L^2(\Omega)},
\end{aligned}
$$

$$
\begin{aligned}
E_4 &= \big(c''(u)\ddot{u}[\nabla u(t)-\nabla W(t)], \nabla\psi-\nabla\psi_h\big) \\
&\le C_{\mathrm{Lip}}\|\ddot{u}\|_{L^\infty(\Omega)}\|\nabla u(t)-\nabla W(t)\|_{L^2(\Omega)}\|\nabla\psi-\nabla\psi_h\|_{L^2(\Omega)} \\
&\le C\|\nabla u(t)-\nabla W(t)\|_{L^2(\Omega)}h\|\psi\|_{H^2(\Omega)} \\
&\le Ch^{r+1}\|\ddot{u}(t)-\ddot{W}(t)\|_{L^2(\Omega)},
\end{aligned}
$$

$$
\begin{aligned}
E_5 &= \Big([c''(u)\ddot{u}(t)-c''(W)\ddot{W}(t)]\nabla W(t), \nabla\psi-\nabla\psi_h\Big)_{L^2} \\
&= \big([c''(u)-c''(W)]\ddot{u}(t)\nabla W(t), \nabla\psi-\nabla\psi_h\big)_{L^2} + \Big(c''(W)[\ddot{u}(t)-\ddot{W}(t)]\nabla W(t), \nabla\psi-\nabla\psi_h\Big)_{L^2} \\
&\le C_{\mathrm{Lip}}\|u(t)-W(t)\|_{L^2(\Omega)}\|\ddot{u}(t)\|_{L^\infty(\Omega)}\|\nabla W(t)\|_{L^\infty(\Omega)}\|\nabla\psi-\nabla\psi_h\|_{L^2(\Omega)}
\end{aligned}
$$

$$
\begin{aligned}
&+ C_{\mathrm{Lip}}\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)}\|\nabla W(t)\|_{L^\infty(\Omega)}\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \\
&\le C\|u(t) - W(t)\|_{L^2(\Omega)} h\|\psi\|_{H^2(\Omega)} + C\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)} h\|\psi\|_{H^2(\Omega)} \\
&\le Ch^{r+1}\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)} + Ch\|\ddot{u}(t) - \ddot{W}(t)\|^2_{L^2(\Omega)},
\end{aligned}
$$

$$
\begin{aligned}
E_6 &= \Big([c''(W)\nabla W(t) - c''(u)\nabla u(t)]\ddot{W}(t), \nabla\psi\Big)_{L^2} \\
&= \Big([c''(W) - c''(u)]\nabla W(t)\ddot{W}(t), \nabla\psi\Big)_{L^2} + \Big(c''(u)\ddot{W}(t)[\nabla W(t) - \nabla u(t)], \nabla\psi\Big)_{L^2} \\
&= \Big([c''(W) - c''(u)]\nabla W(t)[\ddot{W}(t) - \ddot{u}(t)], \nabla\psi\Big)_{L^2} + \big([c''(W) - c''(u)]\nabla W(t), \ddot{u}(t), \nabla\psi\big)_{L^2} \\
&\quad + \Big(c''(u)[\ddot{W}(t) - \ddot{u}(t)][\nabla W(t) - \nabla u(t)], \nabla\psi\Big)_{L^2} + \big(c''(u)\ddot{u}(t)[\nabla W(t) - \nabla u(t)], \nabla\psi\big)_{L^2} \\
&\le C_{\mathrm{Lip}}\|W(t) - u(t)\|_{L^\infty(\Omega)}\|\nabla W(t)\|_{L^\infty(\Omega)}\|\ddot{W}(t) - \ddot{u}(t)\|_{L^2(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\quad + C_{\mathrm{Lip}}\|W(t) - u(t)\|_{L^2(\Omega)}\|\nabla W(t)\|_{L^\infty(\Omega)}\|\ddot{u}(t)\|_{L^\infty(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\quad + C_{\mathrm{Lip}}\|\ddot{u}(t)\|_{L^\infty(\Omega)}\|\nabla W(t) - \nabla u(t)\|_{L^2(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\quad + \Big(c''(u)[\ddot{u}(t) - \ddot{W}(t)][\nabla W(t) - \nabla u(t)], \nabla\psi\Big)_{L^2}.
\end{aligned}
$$

Note that for $d = 2, 3$, by Sobolev embedding theorem and the Gagliardo-Nirenberg interpolation inequality

$$
\|\varphi\|_{L^3(\Omega)} \le C\|\varphi\|^{\frac{1}{2}}_{L^2(\Omega)}\|\nabla\varphi\|^{\frac{1}{2}}_{L^2(\Omega)}, \tag{4.4.31}
$$

we have

$$
\begin{aligned}
&\Big(c''(u)[\ddot{W}(t) - \ddot{u}(t)][\nabla W(t) - \nabla u(t)], \nabla\psi\Big)_{L^2} \\
&\le C_{\mathrm{Lip}}\|\ddot{u}(t) - \ddot{W}(t)\|_{L^3(\Omega)}\|\nabla W(t) - \nabla u(t)\|_{L^2(\Omega)}\|\nabla\psi\|_{L^6(\Omega)} \\
&\le Ch^r\|\ddot{u}(t) - \ddot{W}(t)\|^{\frac{1}{2}}_{L^2(\Omega)}\|\nabla\ddot{u}(t) - \nabla\ddot{W}(t)\|^{\frac{1}{2}}_{L^2(\Omega)}\|\psi\|_{H^2(\Omega)} \\
&\le Ch^r\|\ddot{u}(t) - \ddot{W}(t)\|^{\frac{3}{2}}_{L^2(\Omega)}\|\nabla\ddot{u}(t) - \nabla\ddot{W}(t)\|^{\frac{1}{2}}_{L^2(\Omega)}.
\end{aligned}
$$

For $d = 1$, $H^2(\Omega)$ is compactly embedded into $W^{1,\infty}(\Omega)$, thus

$$
\Big(c''(u)[\ddot{W}(t) - \ddot{u}(t)][\nabla W(t) - \nabla u(t)], \nabla\psi\Big)_{L^2}
$$

$$
\begin{aligned}
&\le C_{\mathrm{Lip}} \|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)} \|\nabla W(t) - \nabla u(t)\|_{L^2(\Omega)} \|\nabla \psi\|_{L^\infty(\Omega)} \\
&\le Ch^r \|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)}.
\end{aligned}
$$

Combining these two cases, we have that

$$
\begin{aligned}
E_6 &\le Ch\|\ddot{u}(t) - \ddot{W}(t)\|^2_{L^2(\Omega)} + Ch^{r+1}\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)} + Ch^r\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)} \\
&\quad + Ch^r\|\ddot{u}(t) - \ddot{W}(t)\|^{\frac{3}{2}}_{L^2(\Omega)} \|\nabla\ddot{u}(t) - \nabla\ddot{W}(t)\|^{\frac{1}{2}}_{L^2(\Omega)}.
\end{aligned}
$$

$$
\begin{aligned}
E_7 &= 2\left(c''(u)\dot{u}(t)\nabla u(t) - c''(W)\dot{W}(t)\nabla W(t), \nabla\psi - \nabla\psi_h\right)_{L^2} \\
&= 2\left([c''(u) - c''(W)]\dot{u}(t)\nabla u(t), \nabla\psi - \nabla\psi_h\right)_{L^2} \\
&\quad + 2\left(c''(W)\dot{W}(t)[\nabla u(t) - \nabla W(t)], \nabla\psi - \nabla\psi_h\right)_{L^2} \\
&\quad + 2\left(c''(W)[\dot{u}(t) - \dot{W}(t)]\nabla u(t), \nabla\psi - \nabla\psi_h\right)_{L^2} \\
&\le 2C_{\mathrm{Lip}}\|u(t) - W(t)\|_{L^2(\Omega)}\|\dot{u}(t)\|_{L^\infty(\Omega)}\|\nabla u(t)\|_{L^\infty(\Omega)}\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \\
&\quad + 2C_{\mathrm{Lip}}\|\dot{W}(t)\|_{L^\infty(\Omega)}\|\nabla u(t) - \nabla W(t)\|_{L^2(\Omega)}\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \\
&\quad + 2C_{\mathrm{Lip}}\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)}\|\nabla u(t)\|_{L^\infty(\Omega)}\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \\
&\le C\|u(t) - W(t)\|_{L^2(\Omega)} h\|\psi\|_{H^2(\Omega)} + C\|\nabla u(t) - \nabla W(t)\|_{L^2(\Omega)} h\|\psi\|_{H^2(\Omega)} \\
&\quad + C\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)} h\|\psi\|_{H^2(\Omega)} \\
&\le Ch^{r+1}\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)}.
\end{aligned}
$$

$$
\begin{aligned}
E_8 &= 2\left(c''(W)\dot{W}(t)\nabla W(t) - c''(u)\dot{u}(t)\nabla u(t), \nabla\psi\right)_{L^2} \\
&\le 2C_{\mathrm{Lip}}\|u(t) - W(t)\|_{L^2(\Omega)}\|\dot{u}(t)\|_{L^\infty(\Omega)}\|\nabla u(t)\|_{L^\infty(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\quad + 2C_{\mathrm{Lip}}\|\dot{W}(t)\|_{L^\infty(\Omega)}\|\nabla u(t) - \nabla W(t)\|_{L^2(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\quad + 2C_{\mathrm{Lip}}\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)}\|\nabla W(t)\|_{L^\infty(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\le C\|u(t) - W(t)\|_{L^2(\Omega)}\|\psi\|_{H^2(\Omega)} + C\|\nabla u(t) - \nabla W(t)\|_{L^2(\Omega)}\|\psi\|_{H^2(\Omega)} \\
&\quad + C\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)}\|\psi\|_{H^2(\Omega)} \\
&\le Ch^r\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)}.
\end{aligned}
$$

$$\begin{aligned}
E_9 &= \Big(c'''(u)\dot{u}(t)^2\nabla u(t) - c'''(W)\dot{W}(t)^2\nabla W(t), \nabla\psi - \nabla\psi_h\Big)_{L^2} \\
&= \big([c'''(u) - c''(W)]\dot{u}(t)^2\nabla u(t), \nabla\psi - \nabla\psi_h\big)_{L^2} + \big(c'''(W)\dot{u}(t)^2[\nabla u(t) - \nabla W(t)], \nabla\psi - \nabla\psi_h\big)_{L^2} \\
&\quad + \Big(c'''(W)\nabla W(t)[\dot{u}(t)^2 - \dot{W}(t)^2], \nabla\psi - \nabla\psi_h\Big)_{L^2} \\
&\le C_{\mathrm{Lip}}\|u(t) - W(t)\|_{L^2(\Omega)}\|\dot{u}(t)\|^2_{L^\infty(\Omega)}\|\nabla u(t)\|_{L^\infty(\Omega)}\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \\
&\quad + C_{\mathrm{Lip}}\|\dot{u}(t)\|^2_{L^\infty(\Omega)}\|\nabla u(t) - \nabla W(t)\|_{L^2(\Omega)}\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \\
&\quad + C_{\mathrm{Lip}}\|\nabla W(t)\|_{L^\infty(\Omega)}\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)}\left(\|\dot{u}(t)\|_{L^\infty(\Omega)} + \|\dot{W}(t)\|_{L^\infty(\Omega)}\right)\|\nabla\psi - \nabla\psi_h\|_{L^2(\Omega)} \\
&\le Ch^{r+2}\|\psi\|_{H^2(\Omega)} + Ch^{r+1}\|\psi\|_{H^2(\Omega)} \\
&\le Ch^{r+1}\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)}.
\end{aligned}$$

$$\begin{aligned}
E_{10} &= \Big(c'''(W)\dot{W}(t)^2\nabla W(t) - c'''(u)\dot{u}(t)^2\nabla u(t), \nabla\psi\Big)_{L^2} \\
&= \big([c'''(W) - c'''(u)]\dot{u}(t)^2\nabla u(t), \nabla\psi\big)_{L^2} + \big(c'''(W)\dot{u}(t)^2[\nabla W(t) - \nabla u(t)], \nabla\psi\big)_{L^2} \\
&\quad + \Big(c'''(W)\nabla W(t)[\dot{W}(t)^2 - \dot{u}(t)^2], \nabla\psi\Big)_{L^2(\Omega)} \\
&\le C_{\mathrm{Lip}}\|u(t) - W(t)\|_{L^2(\Omega)}\|\dot{u}(t)\|^2_{L^\infty(\Omega)}\|\nabla u(t)\|_{L^\infty(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\quad + C_{\mathrm{Lip}}\|\dot{u}(t)\|^2_{L^\infty(\Omega)}\|\nabla u(t) - \nabla W(t)\|_{L^2(\Omega)}\|\nabla\psi\|_{L^2(\Omega)} \\
&\quad + C_{\mathrm{Lip}}\|\nabla W(t)\|_{L^\infty(\Omega)}\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)}\left(\|\dot{u}(t)\|_{L^\infty(\Omega)} + \|\dot{W}(t)\|_{L^\infty(\Omega)}\right)\|\nabla\psi\|_{L^2(\Omega)} \\
&\le Ch^{r+1}\|\psi\|_{H^2(\Omega)} + Ch^r\|\psi\|_{H^2(\Omega)} \le Ch^r\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)}.
\end{aligned}$$

Combining the estimates for E_1–E_{10}, we have

$$\begin{aligned}
\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)} &\le Ch^r + Ch\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)} + Ch\|\nabla\ddot{u}(t) - \nabla\ddot{W}(t)\|_{L^2(\Omega)} \\
&\quad + Ch^r\|\ddot{u}(t) - \ddot{W}(t)\|^{\frac{1}{2}}_{L^2(\Omega)}\|\nabla\ddot{u}(t) - \nabla\ddot{W}(t)\|^{\frac{1}{2}}_{L^2(\Omega)} \\
&\le Ch^r + Ch\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)} + Ch\|\nabla\ddot{u}(t) - \nabla\ddot{W}(t)\|_{L^2(\Omega)} \\
&\quad + Ch^r\left(\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)} + \|\nabla\ddot{u}(t) - \nabla\ddot{W}(t)\|_{L^2(\Omega)}\right).
\end{aligned}$$

This implies that

$$\|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)} \le Ch^r + Ch\|\nabla\ddot{u}(t) - \nabla\ddot{W}(t)\|_{L^2(\Omega)}. \tag{4.4.32}$$

Again, applying the assumption that $c'(\cdot) \geq M_0 > 0$, we have

$$
\begin{aligned}
M_0\|\nabla\ddot{W}(t)-\nabla\ddot{u}(t)\|^2_{L^2(\Omega)} &\leq \Big(c'(W(t))(\nabla\ddot{W}(t)-\nabla\ddot{u}(t)),\nabla\ddot{W}(t)-\nabla\ddot{u}(t)\Big)_{L^2} \\
&= \Big(c'(W(t))(\nabla\ddot{W}(t)-\nabla\ddot{u}(t)),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&\quad + \Big(c'(W(t))(\nabla\ddot{W}(t)-\nabla\ddot{u}(t)),\nabla\mathcal{P}_h\ddot{u}(t)-\nabla\ddot{u}(t)\Big)_{L^2} \\
&:= \mathcal{E}_1+\mathcal{E}_2.
\end{aligned}
$$

Note that

$$
\begin{aligned}
\mathcal{E}_1 &= \Big(c'(W)\nabla\ddot{W}(t),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} - \Big(c'(W)\nabla\ddot{u}(t),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&= \Big([c'(u)-c'(W)]\nabla\ddot{u}(t),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&\quad + 2\Big(c''(u)\nabla\dot{u}(t)\dot{u}(t)-c''(W)(t)\nabla\dot{W}\dot{W}(t),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&\quad + \Big(c''(u)\nabla u(t)\ddot{u}(t)-c''(W)\nabla W(t)\ddot{W}(t),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&\quad + \Big(c'''(u)\nabla u(t)\dot{u}(t)^2-c'''(W)\nabla W(t)\dot{W}(t)^2,\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&:= \mathcal{E}_3+\mathcal{E}_4+\mathcal{E}_5+\mathcal{E}_6.
\end{aligned}
$$

We can bound $\mathcal{E}_3$–$\mathcal{E}_6$ by the following estimates:

$$
\begin{aligned}
\mathcal{E}_3 &= \Big([c'(u)-c'(W)]\nabla\ddot{u}(t),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&\leq C_{\mathrm{Lip}}\|u(t)-W(t)\|_{L^2(\Omega)}\|\nabla\ddot{u}(t)\|_{L^\infty(\Omega)}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)} \\
&\leq Ch^{r+1}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)},
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{E}_4 &= 2\Big(c''(u)\nabla\dot{u}(t)\dot{u}(t)-c''(W)\nabla\dot{W}(t)\dot{W}(t),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2(\Omega)} \\
&= 2\Big([c''(u)-c''(W)]\nabla\dot{u}(t)\dot{u}(t),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2(\Omega)} \\
&\quad + 2\Big(c''(W)\nabla\dot{u}(t)[\dot{u}(t)-\dot{W}(t)],\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2(\Omega)} \\
&\quad + 2\Big(c''(W)[\nabla\dot{u}(t)-\nabla\dot{W}(t)]\dot{W}(t),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2(\Omega)} \\
&\leq 2C_{\mathrm{Lip}}\|u(t)-W(t)\|_{L^2(\Omega)}\|\nabla\dot{u}(t)\|_{L^\infty(\Omega)}\|\dot{u}(t)\|_{L^\infty(\Omega)}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)}
\end{aligned}
$$

$$
\begin{aligned}
&+ 2C_{\text{Lip}}\|\nabla\dot{u}(t)\|_{L^\infty(\Omega)}\|\dot{u}(t)-\dot{W}(t)\|_{L^2(\Omega)}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)} \\
&+ 2C_{\text{Lip}}\|\dot{W}(t)\|_{L^\infty(\Omega)}\|\nabla\dot{u}(t)-\nabla\dot{W}(t)\|_{L^2(\Omega)}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)} \\
&\le Ch^r\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)},
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{E}_5 &= \Big(c''(u)\nabla u(t)\ddot{u}(t)-c''(W)\nabla W(t)\ddot{W}(t),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&= \Big([c''(u)-c''(W)]\nabla u(t)\ddot{u}(t),\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&\quad+ \Big(c''(W)\ddot{u}(t)[\nabla u(t)-\nabla W(t)],\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&\quad+ \Big(c''(W)\nabla W(t)[\ddot{u}(t)-\ddot{W}(t)],\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&\le C_{\text{Lip}}\|u(t)-W(t)\|_{L^2(\Omega)}\|\ddot{u}(t)\|_{L^\infty(\Omega)}\|\nabla u(t)\|_{L^\infty(\Omega)}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)} \\
&\quad+ C_{\text{Lip}}\|\ddot{u}(t)\|_{L^\infty(\Omega)}\|\nabla u(t)-\nabla W(t)\|_{L^2(\Omega)}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)} \\
&\quad+ C_{\text{Lip}}\|\nabla W(t)\|_{L^\infty(\Omega)}\|\ddot{u}(t)-\ddot{W}(t)\|_{L^2(\Omega)}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)} \\
&\le Ch^{r+1}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)} + Ch^r\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)} \\
&\quad+ C\|\ddot{W}(t)-\ddot{u}(t)\|_{L^2(\Omega)}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)},
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{E}_6 &= \Big(c'''(u)\nabla u(t)\dot{u}(t)^2-c'''(W)\nabla W(t)\dot{W}(t)^2,\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&= \Big([c'''(u)-c'''(W)]\nabla u(t)\dot{u}(t)^2,\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&\quad+ \Big(c'''(W)\dot{u}(t)^2[\nabla u(t)-\nabla W(t)],\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&\quad+ \Big(c'''(W)\nabla W(t)[\dot{u}(t)^2-\dot{W}(t)^2],\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\Big)_{L^2} \\
&\le C_{\text{Lip}}\|u(t)-W(t)\|_{L^2(\Omega)}\|\dot{u}(t)\|_{L^\infty(\Omega)}^2\|\nabla u(t)\|_{L^\infty(\Omega)}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)} \\
&\quad+ C_{\text{Lip}}\|\dot{u}(t)\|_{L^\infty(\Omega)}^2\|\nabla u(t)-\nabla W(t)\|_{L^2(\Omega)}\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)} \\
&\quad+ C_{\text{Lip}}\|\nabla W(t)\|_{L^\infty(\Omega)}\|\dot{u}(t)-\dot{W}(t)\|_{L^2(\Omega)}\Big(\|\dot{u}(t)\|_{L^\infty(\Omega)}+\|\dot{W}\|_{L^\infty(\Omega)}\Big)\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)} \\
&\le Ch^r\|\nabla\ddot{W}(t)-\nabla\mathcal{P}_h\ddot{u}(t)\|_{L^2(\Omega)}.
\end{aligned}
$$

For the $\mathcal{E}_2$ term, we have

$$
\begin{aligned}
\mathcal{E}_2 &= \Big(c'(W)(\nabla \ddot{W}(t) - \nabla \ddot{u}(t)), \nabla \mathcal{P}_h \ddot{u}(t) - \nabla \ddot{u}(t)\Big)_{L^2} \\
&\leq C_c \|\nabla \ddot{W}(t) - \nabla \ddot{u}(t)\|_{L^2(\Omega)} \|\nabla \mathcal{P}_h \ddot{u}(t) - \nabla \ddot{u}(t)\|_{L^2(\Omega)} \\
&\leq C h^r \|\nabla \ddot{W}(t) - \nabla \ddot{u}(t)\|_{L^2(\Omega)}.
\end{aligned}
$$

Combining the estimates $\mathcal{E}_1$–$\mathcal{E}_6$, we have

$$
\begin{aligned}
& M_0 \|\nabla \ddot{W}(t) - \nabla \ddot{u}(t)\|^2_{L^2(\Omega)} \\
&\leq C h^r \|\nabla \ddot{W}(t) - \nabla \ddot{u}(t)\|_{L^2(\Omega)} + C\left(h^r + \|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)}\right) \|\nabla \ddot{W}(t) - \nabla \mathcal{P}_h \ddot{u}(t)\|_{L^2(\Omega)} \\
&\leq C h^r \|\nabla \ddot{W}(t) - \nabla \ddot{u}(t)\|_{L^2(\Omega)} + C\left(h^r + \|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)}\right) \|\nabla \ddot{W}(t) - \nabla \ddot{u}(t)\|_{L^2(\Omega)} \\
&\quad + C\left(h^r + \|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)}\right) \|\nabla \ddot{u}(t) - \nabla \mathcal{P}_h \ddot{u}(t)\|_{L^2(\Omega)} \\
&\leq C h^r \|\nabla \ddot{W}(t) - \nabla \ddot{u}(t)\|_{L^2(\Omega)} + C h^{2r} + C \|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)} \left(\|\nabla \ddot{W}(t) - \nabla \ddot{u}(t)\|_{L^2} + h^r\right).
\end{aligned}
$$

This implies that

$$
\|\nabla \ddot{W}(t) - \nabla \ddot{u}(t)\|_{L^2(\Omega)} \leq \; C h^r + C \|\ddot{u}(t) - \ddot{W}(t)\|_{L^2(\Omega)}. \tag{4.4.33}
$$

Substituting (4.4.33) into (4.4.32) gives

$$
\|\ddot{W}(t) - \ddot{u}(t)\|_{L^2(\Omega)} \leq C h^r. \tag{4.4.34}
$$

This completes our proof.

4.5. Numerical experiments

4.5.1. Numerical results for a nonlinear damped wave equation We consider the following one-dimensional nonlinear wave equation

$$\ddot{u}(x,t) + a'(u)\dot{u}(x,t) - \Delta\dot{u}(x,t) - \Delta b(u(x,t)) = f(x,t) \quad \text{in } (0,1)\times(0,T], \tag{4.5.1}$$

$$u(0,t) = u(1,t) = 0 \quad \text{for all } t \in (0,T], \tag{4.5.2}$$

$$u(x,0) = u_0(x), \quad \dot{u}(x,0) = u_1(x). \tag{4.5.3}$$

We take $a(u) = u + u^3$, $b(u) = u + u^3$, and the time interval $I = (0,T]$ with $T = 1$. Let u_0, u_1 and f be chosen such that the exact solution is $u(x,t) = \mathrm{e}^t \sin(\pi x)$. That is, $u_0(x) = u_1(x) = \sin(\pi x)$, and $f(x,t) = (2\pi^2+2)\mathrm{e}^t\sin(\pi x) + (2\pi^2+3)\mathrm{e}^{3t}\sin^3(\pi x) - 6\pi^2\mathrm{e}^{3t}\cos^2(\pi x)\sin(\pi x)$.

Analogously to Section 3.5, the numerical approximation of the nonlinear damped wave equation following a Picard-type linearisation in the nonlinear term is the following: find $u_h \in \mathcal{V}_h$ such that

$$\int_\Omega \ddot{u}_h \cdot v_h \mathrm{d}x + \int_\Omega (3{u_h^*}^2+1)\dot{u}_h\cdot v_h\mathrm{d}x + \int_\Omega (3{u_h^*}^2+1)\partial_x u_h\cdot\partial_x v_h\mathrm{d}x + \int_\Omega \partial_x\dot{u}_h\cdot\partial_x v_h\mathrm{d}x = \int_\Omega f\cdot v_h\mathrm{d}x,$$

for all $v_h \in \mathcal{V}_h$, where $\mathcal{V}_h$ is the finite element function space with polynomial degree $p \geq 1$. We assume that 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. This time, the nonlinear problem results in the following ODE system for $\mathbf{U}(t)$:

$$\begin{cases} \tilde{M}\ddot{\mathbf{U}}(t) + [\tilde{K} + \tilde{M} + \tilde{M}^*(t)]\dot{\mathbf{U}}(t) + [\tilde{K} + \tilde{K}^*(t)]\mathbf{U}(t) = \mathbf{F}(t), \quad t \in (0,T], \\ \dot{\mathbf{U}}(0) = \mathbf{U}_1, \quad \mathbf{U}(0) = \mathbf{U}_0. \end{cases}$$

The mass and stiffness matrices are defined as before. The time-dependent matrices $\tilde{K}^*(t)$ and $\tilde{M}^*(t)$ are defined by

$$\tilde{M}^*_{ij}(t) := \int_0^1 3u_h^*(t)^2\psi_i(x)\cdot\psi_j(x)\,\mathrm{d}x, \quad \tilde{K}^*_{ij}(t) := \int_0^1 3u_h^*(t)^2\partial_x\psi_i(x)\cdot\partial_x\psi_j(x)\,\mathrm{d}x.$$

Again, by setting $\mathbf{Z}(t) = \tilde{M}^{\frac{1}{2}}\mathbf{U}(t)$, we obtain

$$\ddot{\mathbf{Z}}(t) + (K_0 + \mathrm{Id})\dot{\mathbf{Z}}(t) + M(t)\dot{\mathbf{Z}}(t) + K_0\mathbf{Z}(t) + K(t)\mathbf{Z}(t) = \mathbf{G}(t), \quad t \in (0, T], \tag{4.5.4}$$

$$\dot{\mathbf{Z}}(0) = \tilde{M}^{\frac{1}{2}}\mathbf{U}_1, \quad \mathbf{Z}(0) = \tilde{M}^{\frac{1}{2}}\mathbf{U}_0, \tag{4.5.5}$$

where

$$K_0 = \tilde{M}^{-\frac{1}{2}}\tilde{K}\tilde{M}^{-\frac{1}{2}}, \quad K(t) = \tilde{M}^{-\frac{1}{2}}\tilde{K}^*(t)\tilde{M}^{-\frac{1}{2}},$$
$$M(t) = \tilde{M}^{-\frac{1}{2}}\tilde{M}^*(t)\tilde{M}^{-\frac{1}{2}}, \quad \mathbf{G}(t) = \tilde{M}^{-\frac{1}{2}}\mathbf{F}(t).$$

Analogously to Section 3.5, our DG in time formulation reads as: find $\mathbf{Z} \in \mathcal{V}_{kh}^{q_n}$ such that

$$\begin{aligned}
&\left(\ddot{\mathbf{Z}}(t), \dot{\mathbf{v}}\right)_{L^2(I_n)} + \left((K_0 + \mathrm{Id})\dot{\mathbf{Z}}(t), \dot{\mathbf{v}}\right)_{L^2(I_n)} + \left(M(t)\dot{\mathbf{Z}}(t), \dot{\mathbf{v}}\right)_{L^2(I_n)} + (K_0\mathbf{Z}(t), \dot{\mathbf{v}})_{L^2(I_n)} \\
&\quad + (K(t)\mathbf{Z}(t), \dot{\mathbf{v}})_{L^2(I_n)} + \dot{\mathbf{Z}}(t_{n-1}^+) \cdot \dot{\mathbf{v}}(t_{n-1}^+) + K_0\mathbf{Z}(t_{n-1}^+) \cdot \mathbf{v}(t_{n-1}^+) \\
&= (\mathbf{G}(t), \dot{\mathbf{v}})_{L^2(I_n)} + \dot{\mathbf{Z}}(t_{n-1}^-) \cdot \dot{\mathbf{v}}(t_{n-1}^+) + K_0\mathbf{Z}(t_{n-1}^-) \cdot \mathbf{v}(t_{n-1}^+),
\end{aligned} \tag{4.5.6}$$

for all $\mathbf{v} \in \mathcal{V}_{kh}^{q_n}$. Similarly to previous chapters, we obtain the algebraic system $\mathbf{Az} = \mathbf{b}$, where $\mathbf{z} \in \mathbb{R}^D = \mathbb{R}^{(q_n+1)\hat{d}}$ is the solution vector, $\mathbf{b} \in \mathbb{R}^D$ corresponds to the right-hand side, and $\mathbf{A}$ is the local stiffness matrix with its structure defined in Section 2.4. Here, we set

$$M = M^1 + M^4, \quad B_{ij} = L_{ij}M^2 + \tilde{M}^2 + K_{0ij}(M^3 + M^5) + \tilde{M}^3,$$

where $L = K_0 + \mathrm{Id}$, $M, B_{ij} \in \mathbb{R}^{(q_n+1)\times(q_n+1)}$ for any $i, j = 1, \ldots, \hat{d}$. Here, M^1–M^5 are defined the same as in Section 2.4, $\tilde{M}^3$ and $\tilde{M}^5$ are defined in Section 3.5, and the new term $\tilde{M}^2$ is defined by $\tilde{M}^2_{lj} = \left(M(t)\dot{\phi}^j, \dot{\phi}^l\right)_{L^2(I_n)}$, for $l, j = 1, \ldots, q_n + 1$. In order to compute the time-dependent matrices $\tilde{M}_2(t)$ and $\tilde{M}_3(t)$, we apply a *Picard iteration* at each time interval. This is summarised in Algorithm 2.

Algorithm 2 Iterative Algorithm (Multiple Picard iterations at each time interval)

Initialisation: $u_h^* = u_0$ and

$$[\tilde{K}^0]_{ij} = 3\int_0^1 (u_0)^2 \partial_x \psi_i(x) \partial_x \psi_j(x)\,\mathrm{d}x,$$
$$[\tilde{M}^0]_{ij} = 3\int_0^1 (u_0)^2 \psi_i(x) \psi_j(x)\,\mathrm{d}x.$$

Iteration: On each interval $I_n = (t_{n-1}, t_n]$ for $n = 1, 2, \ldots, N$, we solve

$$\tilde{M}\ddot{\mathbf{U}}(t) + [\tilde{K} + \tilde{M} + \tilde{M}^n(t)]\dot{\mathbf{U}}(t) + [\tilde{K} + \tilde{K}^n(t)]\mathbf{U}(t) = \mathbf{F}(t),$$

iteratively (using Picard iterations) by applying the discontinuous-in-time integration. Here

$$[\tilde{K}_0^n]_{ij} = 3\int_0^1 [u_{\mathrm{DG}}^{n-1}(t)]^2 \partial_x \psi_i(x) \partial_x \psi_j(x)\,\mathrm{d}x,$$
$$[\tilde{M}_0^n]_{ij} = 3\int_0^1 [u_{\mathrm{DG}}^{n-1}(t)]^2 \psi_i(x) \psi_j(x)\,\mathrm{d}x,$$

where $u_{\mathrm{DG}}^{n-1}(t)$ is the solution we obtained from the previous time interval I_{n-1}.

$$[\tilde{K}_k^n]_{ij} = 3\int_0^1 [u_{\mathrm{DG}}^{n-1,k-1}(t)]^2 \partial_x \psi_i(x) \partial_x \psi_j(x)\,\mathrm{d}x,$$
$$[\tilde{M}_k^n]_{ij} = 3\int_0^1 [u_{\mathrm{DG}}^{n-1,k-1}(t)]^2 \psi_i(x) \psi_j(x)\,\mathrm{d}x,$$

for $k = 1, 2, \ldots$, where $u_{\mathrm{DG}}^{n-1,k-1}(t)$ is computed from the previous Picard iteration by using $[\tilde{K}_{k-1}^n]_{ij}$ and $[\tilde{M}_{k-1}^n]_{ij}$.

Update:

$$[\tilde{K}_0^{n+1}]_{ij} = 3\int_0^1 [u_{\mathrm{DG}}^{n}(t)]^2 \partial_x \psi_i(x) \partial_x \psi_j(x)\,\mathrm{d}x,$$
$$[\tilde{M}_0^{n+1}]_{ij} = 3\int_0^1 [u_{\mathrm{DG}}^{n}(t)]^2 \psi_i(x) \psi_j(x)\,\mathrm{d}x,$$

where $u_{\mathrm{DG}}^n(t)$ is computed using $\tilde{K}_{k_{\mathrm{end}}}^n$ and $\tilde{M}_{k_{\mathrm{end}}}^n$. Here k_{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}.

We use CG–p elements for $p \geq 2$ in space with $k = h^2$, $T = 1$ and compute the errors $\|u(T) - u_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)} + \|\dot{u}(T) - \dot{u}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)}$ versus k for $k = h^2$ with respect to polynomial degrees p for $q = 2, 3$ and 4 respectively. Note that here we use $h = 2.50\times10^{-1}$, 2.00×10^{-1}, 1.25×10^{-1} and 1.00×10^{-1} 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.

For $q = 2$ (i.e., DG–2 elements in the temporal domain), we consider $p = 2, 3$ and 4. The computed errors are presented in Table 4.1 and plotted in Figure 4.1 with a log-log scale. By Remark 4.8, we expect convergence rates of order 1.0, 1.5 and 2.0 for $p = 2, 3$ and 4 respectively, which are consistent with the numerical results shown in Table 4.1.

Similarly, for $q = 3$, the computed errors for $p = 3, 4, 5$ and 6 are shown in Figure 4.2, and they agree with the expected convergence rates of 1.5, 2.0, 2.5 and 3.0 respectively.

For DG–4 elements in the temporal domain, we would expect convergence rates of $2.0, 2.5, 3.0$ for $p = 4, 5,$ and 6 respectively since the error is of order $O(k^{\frac{p}{2}} + k^4)$ with $k^{\frac{p}{2}}$ being the dominating term for $p \leq 8$. The numerical results agree with our theoretical findings (cf. Figure 4.3). In general, the accuracy of this discontinuous-in-time integration scheme outperforms many other existing finite difference schemes. If we compare Table 4.1, Table 4.2 and Table 4.3, we observe that the error decreases as we increase the polynomial degree q or decrease the time step k, thus permitting up to exponential rates of convergence.

Remark 4.10. *Note that the $\Delta\dot{u}$ term is added in (4.0.1) for the sake of ensuring stability of our numerical scheme. If we remove this term and consider the following equation*

$$\ddot{u}(x,t) + a'(u)\dot{u}(x,t) - \Delta b(u(x,t)) = f(x,t) \qquad \text{in } (0,1) \times (0,T],$$

with $f = (\pi^2 + 2)\mathrm{e}^t \sin(\pi x) + (2\pi^2 + 3)\mathrm{e}^{3t} \sin^3(\pi x) - 6\pi^2\mathrm{e}^{3t} \cos^2(\pi x)\sin(\pi x)$ in this case, similar convergence rates are observed. This suggests that it is possible to directly apply our proposed discontinuous time-stepping method to the original damped wave equation considered in Süli and Wilkins' work [74].

Table 4.1: 1D nonlinear damped wave equation with $q = 2$: computed error $||u(T) - u_{\mathrm{DG}}(t_N^-)||_{H^1(\Omega)} + ||\dot{u}(T) - \dot{u}_{\mathrm{DG}}(t_N^-)||_{L^2(\Omega)}$ and corresponding convergence rates with respect to polynomial degrees $p = 2, 3, 4$.

p	h	k	L^2-error	rate
2	5.0000e − 1	2.5000e − 2	6.0903e − 1	—
	2.5000e − 1	6.2500e − 2	1.4774e − 1	1.0217
	2.0000e − 1	4.0000e − 2	8.9097e − 2	1.3332
	1.2500e − 1	1.5625e − 2	3.5502e − 2	0.9789
	1.0000e − 1	1.0000e − 2	2.2374e − 2	1.0345
3	5.0000e − 1	2.5000e − 2	8.5172e − 2	—
	2.5000e − 1	6.2500e − 2	9.5072e − 3	1.5816
	2.0000e − 1	4.0000e − 2	4.6403e − 3	1.6072
	1.2500e − 1	1.5625e − 2	1.1684e − 3	1.4672
	1.0000e − 1	1.0000e − 2	5.9048e − 4	1.5292
4	5.0000e − 1	2.5000e − 2	1.2530e − 2	—
	2.5000e − 1	6.2500e − 2	5.4227e − 4	2.2651
	2.0000e − 1	4.0000e − 2	2.0021e − 4	2.2326
	1.2500e − 1	1.5625e − 2	2.9801e − 5	2.0264
	1.0000e − 1	1.0000e − 2	1.1899e − 5	2.0572

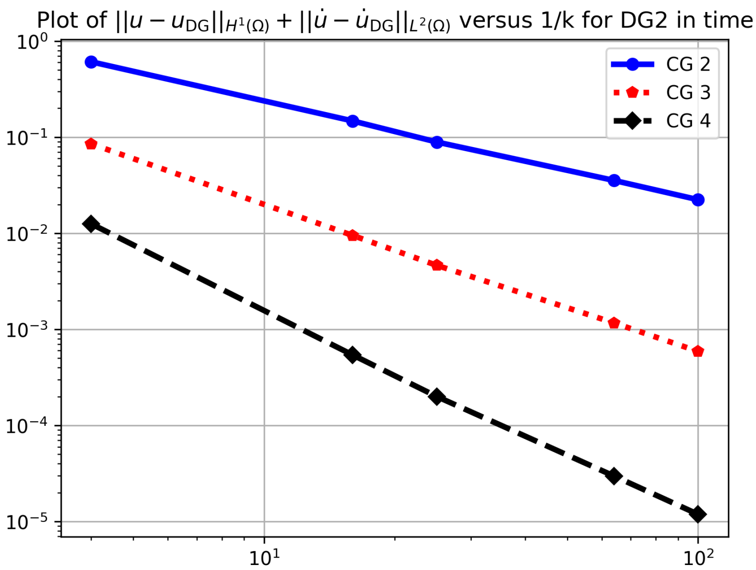

Figure 4.1: 1D nonlinear damped wave equation with $q = 2$: computed error $\|u(T) - u_{\mathrm{DG}}(t_N^-)\|_{H^1(\Omega)} + \|\dot{u}(T) - \dot{u}_{\mathrm{DG}}(t_N^-)\|_{L^2(\Omega)}$ plotted against $1/k$ for $p = 2, 3, 4$ in a log-log scale.

Table 4.2: 1D nonlinear damped wave equation with $q = 3$: computed error $||u(T) - u_{\text{DG}}(t_N^-)||_{H^1(\Omega)} + ||\dot{u}(T) - \dot{u}_{\text{DG}}(t_N^-)||_{L^2(\Omega)}$ and corresponding convergence rates with respect to polynomial degrees $p = 3, 4, 5, 6$.

p	h	k	L^2-error	rate
3	5.0000e − 1	2.5000e − 2	8.3459e − 2	—
	2.5000e − 1	6.2500e − 2	9.4976e − 2	1.5677
	2.0000e − 1	4.0000e − 2	4.6388e − 3	1.6057
	1.2500e − 1	1.5625e − 2	1.1684e − 3	1.4668
	1.0000e − 1	1.0000e − 2	5.9047e − 4	1.5292
4	5.0000e − 1	2.5000e − 2	8.2612e − 3	—
	2.5000e − 1	6.2500e − 2	4.8451e − 4	2.0459
	2.0000e − 1	4.0000e − 2	1.8660e − 4	2.1380
	1.2500e − 1	1.5625e − 2	2.9049e − 5	1.9787
	1.0000e − 1	1.0000e − 2	1.1718e − 5	2.0342
5	5.0000e − 1	2.5000e − 2	6.6916e − 4	—
	2.5000e − 1	6.2500e − 2	1.8487e − 5	2.5888
	2.0000e − 1	4.0000e − 2	5.7761e − 6	2.6067
	1.2500e − 1	1.5625e − 2	5.6859e − 7	2.4663
	1.0000e − 1	1.0000e − 2	1.8397e − 7	2.5284
6	5.0000e − 1	2.5000e − 2	6.7489e − 5	—
	2.5000e − 1	6.2500e − 2	6.8508e − 7	3.3111
	2.0000e − 1	4.0000e − 2	1.6209e − 7	3.2297
	1.2500e − 1	1.5625e − 2	9.5069e − 9	3.0171
	1.0000e − 1	1.0000e − 2	2.4616e − 9	3.0277

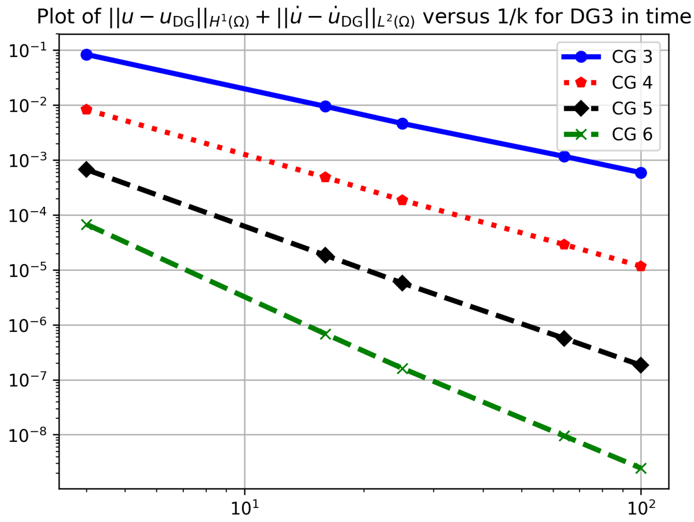

Figure 4.2: 1D nonlinear damped wave equation with $q = 3$: computed error $\|u(T) - u_{\text{DG}}(t_N^-)\|_{H^1(\Omega)} + \|\dot{u}(T) - \dot{u}_{\text{DG}}(t_N^-)\|_{L^2(\Omega)}$ plotted against $1/k$ for $p = 3, 4, 5, 6$ in a log-log scale.

Table 4.3: 1D nonlinear damped wave equation with $q = 4$: computed error $||u(T) - u_{\rm DG}(t_N^-)||_{H^1(\Omega)} + ||\dot{u}(T) - \dot{u}_{\rm DG}(t_N^-)||_{L^2(\Omega)}$ and corresponding convergence rates with respect to polynomial degrees $p = 4, 5, 6$.

p	h	k	L^2-error	rate
4	5.0000e − 1	2.5000e − 2	8.2580e − 3	—
	2.5000e − 1	6.2500e − 2	4.8451e − 4	2.0456
	2.0000e − 1	4.0000e − 2	1.8660e − 4	2.1380
	1.2500e − 1	1.5625e − 2	2.9049e − 5	1.9787
	1.0000e − 1	1.0000e − 2	1.1718e − 5	2.0342
5	5.0000e − 1	2.5000e − 2	6.5806e − 4	—
	2.5000e − 1	6.2500e − 2	1.8482e − 5	2.5770
	2.0000e − 1	4.0000e − 2	5.7757e − 6	2.6063
	1.2500e − 1	1.5625e − 2	5.6859e − 7	2.4662
	1.0000e − 1	1.0000e − 2	1.8397e − 7	2.5284
6	5.0000e − 1	2.5000e − 2	4.4105e − 5	—
	2.5000e − 1	6.2500e − 2	6.3450e − 7	3.0596
	2.0000e − 1	4.0000e − 2	1.5579e − 7	3.1467
	1.2500e − 1	1.5625e − 2	9.4363e − 9	2.9829
	1.0000e − 1	1.0000e − 2	2.4728e − 9	3.0008

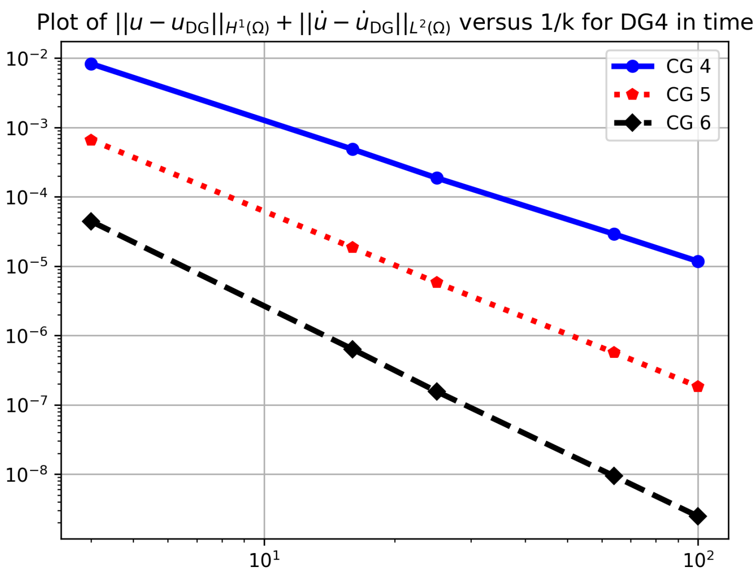

Figure 4.3: 1D nonlinear damped wave equation with $q = 4$: computed error $\|u(T) - u_{\rm DG}(t_N^-)\|_{H^1(\Omega)} + \|\dot{u}(T) - \dot{u}_{\rm DG}(t_N^-)\|_{L^2(\Omega)}$ plotted against $1/k$ for $p = 4, 5, 6$ in a log-log scale.

Chapter 5

Conclusion

In this thesis, we have proposed an arbitrarily high-order accurate time integration scheme for a broad range of second-order hyperbolic PDEs, supported with a complete mathematical analysis in terms of stability and convergence, as well as various numerical examples to verify our theoretical findings. This proposed numerical scheme is based on the *hp*-version of the DGFEM. It overcomes many challenges in the mathematical analysis of high-order numerical methods for second-order nonlinear hyperbolic PDEs and overperforms many existing finite difference time-stepping schemes in terms of flexibility and accuracy. Our numerical experiments on hyperbolic equations of second order demonstrate the high-order convergence rates in the temporal domain and the adaptability of this numerical algorithm to different nonlinear problems.

We have shown that this discontinuous-in-time scheme is an implicit, unconditionally stable and convergent method for the linear problem. In contrast, we require a weak restriction on the temporal step in terms of the spatial mesh size (i.e. $k_n = O(h^2)$ for each $n = 1, \ldots, N$) for the nonlinear case. It might be possible to relax these conditions if one could improve the fixed point argument that we used in the derivation of the error estimates.

This discontinuous-in-time discretisation approach to nonlinear elastodynamics systems presented in Chapter 3 also shows its potential to solve more challenging Einstein equations. For instance, Hughes, Kato and Marsden [43] mentioned that the following Einstein equations for a Lorentz metric $g_{\alpha,\beta}$ on $\mathbb{R}^4$, $0 \leq \alpha, \beta \leq 3$,

$$-\frac{1}{2} g^{\alpha\beta} \frac{\partial^2 g_{u\mu}}{\partial x^\alpha \partial x^\beta} + H_{\mu\nu}\left(g_{\alpha\beta}, \frac{\partial g_{\alpha\beta}}{\partial x^\lambda} \right) = 0 \tag{5.0.1}$$

can be transformed into a quasilinear second-order hyperbolic systems of the form

$$a_{00} \frac{\partial^2 \varphi}{\partial t^2} = \sum_{i,j=1}^{d} a_{ij} \frac{\partial^2 \varphi}{\partial x_i \partial x_j} + \sum_{i=1}^{d} (a_{0i} + a_{i0}) \frac{\partial^2 \varphi}{\partial t \partial x_i} + b, \tag{5.0.2}$$

where the unknown $\varphi = (\varphi_1, \ldots, \varphi_N)$ is an N–vector valued function of $t \in [0, T]$ and of $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, where $\{a_{ij} \mid i, j = 1, \ldots, d\}$ is a collection of $(N \times N)$-matrix valued functions of the suppressed arguments $t, x, \varphi, \frac{\partial \varphi}{\partial t}, \nabla \varphi$, and b is an N–vector valued functions of the same arguments. Despite the fact that the extra coefficient terms a_{ij} may lead to more tedious calculations, a slightly modified version of our proposed scheme would be naturally suited for this more general class of quasilinear hyperbolic systems.

Another possible direction for future research would be exploring the preconditioning and parallelisation strategies to improve the efficiency of computations. Werder, Gerdes, and their collaborators [77] introduced a decoupling procedure for a discontinuous time-stepping method for parabolic problems. Their numerical examples showed that the decoupling process is helpful in terms of the reduction of both the computational time and memory requirements. They also briefly discussed several parallelisation strategies, including local static condensation, parallel computing on shared memory machines and distributed memory machines. All of the possible parallelisation strategies considered for parabolic problems open up the possibility of designing similar robust and efficient solvers for nonlinear hyperbolic equations of second order, which would be of importance for the numerical approximation of second-order linear and nonlinear hyperbolic problems that arise in real-world applications.

Bibliography

[1] R. A. Adams, *Sobolev spaces*, Academic Press, New York, 1975.

[2] S. Adjerid and H. Temimi, *A discontinuous Galerkin method for higher-order ordinary differential equations*, Comput. Methods Appl. Mech. Engrg., 197 (2007), pp. 202–218.

[3] ——, *A discontinuous Galerkin method for the wave equation*, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 837–849.

[4] M. Ainsworth, P. Monk, and W. Muniz, *Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation*, J. Sci. Comput., 27 (2006), pp. 5–40.

[5] S. S. Antman, *Nonlinear problems of elasticity*, vol. 107 of Applied Mathematical Sciences, Springer, New York, second ed., 2005.

[6] P. F. Antonietti, I. Mazzieri, N. Dal Santo, and A. Quarteroni, *A high-order discontinuous Galerkin approximation to ordinary differential equations with applications to elastodynamics*, IMA J. Numer. Anal., 38 (2018), pp. 1709–1734.

[7] P. F. Antonietti, I. Mazzieri, and F. Migliorini, *A space-time discontinuous Galerkin method for the elastic wave equation*, J. Comput. Phys., 419 (2020), pp. 109685, 26.

[8] P. F. Antonietti, I. Mazzieri, A. Quarteroni, and F. Rapetti, *Nonconforming high order approximations of the elastodynamics equation*, Comput. Methods Appl. Mech. Engrg., 209/212 (2012), pp. 212–238.

[9] D. N. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.

[10] J.-P. Aubin, *Un théorème de compacité*, C. R. Acad. Sci. Paris, 256 (1963), pp. 5042–5044.

[11] I. Babuška and M. Suri, *The h-p version of the finite element method with quasi-uniform meshes*, RAIRO Modél. Math. Anal. Numér., 21 (1987), pp. 199–238.

[12] I. Babuška and M. Zlámal, *Nonconforming elements in the finite element method with penalty*, SIAM J. Numer. Anal., 10 (1973), pp. 863–875.

[13] 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), pp. 27–43.

[14] S. Banach, *Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales*, Fund. Math, 3 (1922), pp. 133–181.

[15] C. Bernardi, *Optimal finite-element interpolation on curved domains*, SIAM J. Numer. Anal., 26 (1989), pp. 1212–1240.

[16] F. Bloom, *Mathematical problems of classical nonlinear electromagnetic theory*, vol. 63 of Pitman Monographs and Surveys in Pure and Applied Mathematics, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1993.

[17] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, vol. 15 of Texts in Applied Mathematics, Springer, New York, third ed., 2008.

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

[19] Y.-Z. Chen and L.-C. Wu, *Second order elliptic equations and elliptic systems*, vol. 174 of Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 1998. Translated from the 1991 Chinese original by Bei Hu.

[20] B. Cockburn and C.-W. Shu, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework*, Math. Comp., 52 (1989), pp. 411–435.

[21] ——, *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM J. Numer. Anal., 35 (1998), pp. 2440–2463.

[22] R. Courant, K. Friedrichs, and H. Lewy, *Über die partiellen Differenzengleichungen der mathematischen Physik*, Math. Ann., 100 (1928), pp. 32–74.

[23] B. Dacorogna and P. Marcellini, *A counterexample in the vectorial calculus of variations*, in Material instabilities in continuum mechanics (Edinburgh, 1985–1986), Oxford Sci. Publ., Oxford Univ. Press, New York, 1988, pp. 77–83.

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

[25] M. Delfour, W. Hager, and F. Trochu, *Discontinuous Galerkin methods for ordinary differential equations*, Math. Comp., 36 (1981), pp. 455–473.

[26] M. Dobrowolski and R. Rannacher, *Finite element methods for nonlinear elliptic systems of second order*, Math. Nachr., 94 (1980), pp. 155–172.

[27] J. Douglas, Jr. and T. Dupont, *A Galerkin method for a nonlinear Dirichlet problem*, Math. Comp., 29 (1975), pp. 689–696.

[28] K. Eriksson and C. Johnson, *Error estimates and automatic time step control for nonlinear parabolic problems. I*, SIAM J. Numer. Anal., 24 (1987), pp. 12–23.

[29] ——, *Adaptive finite element methods for parabolic problems. I. A linear model problem*, SIAM J. Numer. Anal., 28 (1991), pp. 43–77.

[30] ——, *Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal., 32 (1995), pp. 706–740.

[31] ——, *Adaptive finite element methods for parabolic problems. IV. Nonlinear problems*, SIAM J. Numer. Anal., 32 (1995), pp. 1729–1749.

[32] ——, *Adaptive finite element methods for parabolic problems. V. Long-time integration*, SIAM J. Numer. Anal., 32 (1995), pp. 1750–1763.

[33] K. Eriksson, C. Johnson, and S. Larsson, *Adaptive finite element methods for parabolic problems. VI. Analytic semigroups*, SIAM J. Numer. Anal., 35 (1998), pp. 1315–1325.

[34] K. Eriksson, C. Johnson, and V. Thomée, *Time discretization of parabolic problems by the discontinuous Galerkin method*, RAIRO Modél. Math. Anal. Numér., 19 (1985), pp. 611–643.

[35] S. Erlicher, L. Bonaventura, and O. S. Bursi, *The analysis of the generalized-α method for non-linear dynamic problems*, Comput. Mech., 28 (2002), pp. 83–104.

[36] L. C. Evans, *Partial differential equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2010.

[37] R. S. Falk and G. R. Richter, *Local error estimates for a finite element method for hyperbolic and convection-diffusion equations*, SIAM J. Numer. Anal., 29 (1992), pp. 730–754.

[38] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[39] M. J. Grote, A. Schneebeli, and D. Schötzau, *Discontinuous Galerkin finite element method for the wave equation*, SIAM J. Numer. Anal., 44 (2006), pp. 2408–2431.

[40] M. E. Gurtin, *An introduction to continuum mechanics*, vol. 158 of Mathematics in Science and Engineering, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1981.

[41] J. M. Holte, *Discrete Gronwall lemma and applications*, in MAA-NCS meeting at the University of North Dakota, vol. 24, 2009, pp. 1–7.

[42] T. J. R. Hughes and G. M. Hulbert, *Space-time finite element methods for elastodynamics: formulations and error estimates*, Comput. Methods Appl. Mech. Engrg., 66 (1988), pp. 339–363.

[43] 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. Rational Mech. Anal., 63 (1976), pp. 273–294 (1977).

[44] G. M. Hulbert, *Space-time finite element methods for second-order hyperbolic equations*, PhD thesis, Standford University, 1989.

[45] G. M. Hulbert and T. J. R. Hughes, *Space-time finite element methods for second-order hyperbolic equations*, Comput. Methods Appl. Mech. Engrg., 84 (1990), pp. 327–348.

[46] P. Jamet, *Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain*, SIAM J. Numer. Anal., 15 (1978), pp. 912–928.

[47] C. Johnson, *Discontinuous Galerkin finite element methods for second order hyperbolic problems*, Comput. Methods Appl. Mech. Engrg., 107 (1993), pp. 117–129.

[48] C. Johnson and J. Pitkäranta, *An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation*, Math. Comp., 46 (1986), pp. 1–26.

[49] T. Kato, *Quasi-linear equations of evolution, with applications to partial differential equations*, in Spectral theory and differential equations, Springer, 1975, pp. 25–70.

[50] W. Kutta, *Beitrag zur näherungsweisen integration totaler differentialgleichungen*, Z. Math. Phys., 46 (1901), pp. 435–453.

[51] P. Lasaint and P.-A. Raviart, *On a finite element method for solving the neutron transport equation*, in Mathematical aspects of finite elements in partial differential equations (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1974), 1974, pp. 89–123. Publication No. 33.

[52] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Paris; Gauthier-Villars, Paris, 1969.

[53] J.-L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications. Vol. I*, Die Grundlehren der mathematischen Wissenschaften, Band 181, Springer-Verlag, New York-Heidelberg, 1972. Translated from the French by P. Kenneth.

[54] C. G. Makridakis, *Galerkin/finite element methods for the equations of elastodynamics*, PhD thesis, Univ. of Crete (Greek), 1989.

[55] ——, *Finite element approximations of nonlinear elastic waves*, tech. rep., Dept. of Mathematics, Univ. of Crete, 1992.

[56] C. G. Makridakis, *Finite element approximations of nonlinear elastic waves*, Math. Comp., 61 (1993), pp. 569–594.

[57] C. B. Morrey, Jr., *Multiple integrals in the calculus of variations*, Die Grundlehren der mathematischen Wissenschaften, Band 130, Springer-Verlag New York, Inc., New York, 1966.

[58] N. M. Newmark, *A method of computation for structural dynamics*, Journal of the engineering mechanics division, 85 (1959), pp. 67–94.

[59] J. A. Nitsche, *On l^∞–convergence of finite element approximations to the solution of a nonlinear boundary value problem*, tech. rep., In:Roc. Conf. Numerical Analysis, Dublin, 1976.

[60] C. Ortner and E. Süli, *Discontinuous Galerkin finite element approximation of nonlinear second-order elliptic and hyperbolic systems*, tech. rep., Oxford University Computing Laboratory, London, 2006.

[61] C. Ortner and E. Süli, *Discontinuous Galerkin finite element approximation of nonlinear second-order elliptic and hyperbolic systems*, SIAM J. Numer. Anal., 45 (2007), pp. 1370–1397.

[62] R. Rannacher, *On finite element approximation of general boundary value problems in nonlinear elasticity*, Calcolo, 17 (1980), pp. 175–193 (1981).

[63] R. Rannacher and R. Scott, *Some optimal error estimates for piecewise linear finite element approximations*, Math. Comp., 38 (1982), pp. 437–445.

[64] W. H. Reed and T. R. Hill, *Triangular mesh methods for the neutron transport equation*, tech. rep., Los Alamos Scientific Lab., N. Mex.(USA), 1973.

[65] G. R. Richter, *The discontinuous Galerkin method with diffusion*, Math. Comp., 58 (1992), pp. 631–643.

[66] B. Rivière, *Discontinuous Galerkin methods for solving elliptic and parabolic equations*, vol. 35 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[67] C. Runge, *Üeber die numerische Auflösung von Differentialgleichungen*, Math. Ann., 46 (1895), pp. 167–178.

[68] D. Schötzau and C. Schwab, *An hp a priori error analysis of the DG time-stepping method for initial value problems*, Calcolo, 37 (2000), pp. 207–232.

[69] D. Schötzau and C. Schwab, *Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method*, SIAM J. Numer. Anal., 38 (2000), pp. 837–875.

[70] C. Schwab, *p- and hp-finite element methods*, Numerical Mathematics and Scientific Computation, The Clarendon Press, Oxford University Press, New York, 1998. Theory and applications in solid and fluid mechanics.

[71] G. Seregin, *Parabolic pdes*, Lecture Notes from University of Oxford, (2019).

[72] J. Simon, *Compact sets in the space $l^p(o, t; b)$*, Annali di Matematica pura ed applicata, 146 (1986), pp. 65–96.

[73] E. Süli, P. Houston, and C. Schwab, *hp-finite element methods for hyperbolic problems*, in The mathematics of finite elements and applications, X, MAFELAP 1999 (Uxbridge), Elsevier, Oxford, 2000, pp. 143–162.

[74] E. Süli and C. Wilkins, *A priori analysis for the semi-discrete approximation to the nonlinear damped wave equation*, Tech. Rep. 99/09, Oxford University Computing Laboratory, 2000.

[75] V. Thomée, *Galerkin finite element methods for parabolic problems*, vol. 25 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 2006.

[76] V. Šverák, *Rank-one convexity does not imply quasiconvexity*, Proc. Roy. Soc. Edinburgh Sect. A, 120 (1992), pp. 185–189.

[77] T. Werder, K. Gerdes, D. Schötzau, and C. Schwab, *hp-discontinuous Galerkin time stepping for parabolic problems*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 6685–6708.

[78] M. F. Wheeler, *An elliptic collocation-finite element method with interior penalties*, SIAM J. Numer. Anal., 15 (1978), pp. 152–161.

[79] J. Wloka, *Partial differential equations*, Cambridge University Press, Cambridge, 1987. Translated from the German by C. B. Thomas and M. J. Thomas.

[80] K. Zhang, *On the coercivity of elliptic systems in two-dimensional spaces*, Bull. Austral. Math. Soc., 54 (1996), pp. 423–430.