

Discontinuous Galerkin Finite Element Approximation of Nonlinear Second-Order Elliptic and Hyperbolic Systems

Christoph Ortner & Endre Süli

We develop the convergence analysis of discontinuous Galerkin finite element approximations to second-order quasilinear elliptic and hyperbolic systems of partial differential equations of the form, respectively, $-\sum_{\alpha=1}^d \partial_{x_\alpha} S_{i\alpha}(\nabla u(x)) = f_i(x)$, $i = 1, \dots, d$, and $\partial_t^2 u_i - \sum_{\alpha=1}^d \partial_{x_\alpha} S_{i\alpha}(\nabla u(t, x)) = f_i(t, x)$, $i = 1, \dots, d$, with $\partial_{x_\alpha} = \partial/\partial x_\alpha$, in a bounded spatial domain in \mathbb{R}^d , subject to mixed Dirichlet–Neumann boundary conditions, and assuming that $S = (S_{i\alpha})$ is uniformly monotone on $\mathbb{R}^{d \times d}$. The associated energy functional is then uniformly convex. An optimal order bound is derived on the discretization error in each case without requiring the global Lipschitz continuity of the tensor S . We then further relax our hypotheses: using a broken Gårding inequality we extend our optimal error bounds to the case of quasilinear hyperbolic systems where, instead of assuming that S is uniformly monotone, we only require that the fourth-order tensor $A = \nabla S$ satisfies a Legendre–Hadamard condition. The associated energy functional is then only rank-1 convex. Evolution problems of this kind arise as mathematical models in nonlinear elastic wave propagation.

Key words and phrases: Nonlinear elliptic and hyperbolic systems of partial differential equations, discontinuous Galerkin methods, Legendre–Hadamard condition, broken Gårding inequality

Oxford University Computing Laboratory
Numerical Analysis Group
Wolfson Building
Parks Road
Oxford, England OX1 3QD

May, 2006

*The authors acknowledge the financial support received from the European research project HPRN-CT-2002-00284: *New Materials, Adaptive Systems and their Nonlinearities. Modelling, Control and Numerical Simulation*, and the kind hospitality of Carlo Lovadina and Matteo Negri (University of Pavia).

1 Introduction

Second-order nonlinear elliptic and hyperbolic systems of partial differential equations arise in numerous applications, and a substantial body of research has been devoted to their analytical and computational study. This paper is concerned with the construction and convergence analysis of a class of numerical algorithms — discontinuous Galerkin finite element methods — for the approximate solution of quasilinear elliptic and hyperbolic systems. Nonlinear elasticity is a particularly fertile source of equations of this type. In order to motivate the discussion that will follow, we begin by formulating a static problem from nonlinear elasticity which results in a mixed Dirichlet–Neumann boundary-value problem for a system of second-order quasilinear elliptic partial differential equations. We shall then state the associated dynamic problem, which is a mixed initial-boundary-value problem for a second-order quasilinear hyperbolic system.

Suppose that Ω is a bounded open set in \mathbb{R}^d , $d \in \{2, 3\}$, with Lipschitz continuous boundary $\partial\Omega$. We shall seek a displacement field $u : \overline{\Omega} \rightarrow \mathbb{R}^d$ such that u is a stationary point of the energy functional

$$J : v \mapsto J(v) := \int_{\Omega} W(\nabla v(x)) \, dx - \int_{\Omega} f(x) \cdot v(x) \, dx - \int_{\Gamma_N} g_N(s) \cdot v(s) \, ds,$$

defined over the set of all (sufficiently smooth) d -component vector functions v on $\overline{\Omega}$ satisfying the boundary condition $v = g_D$ on Γ_D , where $\Gamma_D \subset \Gamma = \partial\Omega$ has positive $(d-1)$ -dimensional surface measure $\mathcal{H}^{d-1}(\Gamma_D)$, $\Gamma_N = \Gamma \setminus \Gamma_D$, $W \in C^4(\mathbb{R}^{d \times d}; \mathbb{R})$ is the stored energy function, $f \in L^2(\Omega)^d$ is a given body force, and $g_N \in L^2(\Gamma_N)^d$. Let us define the Piola–Kirchhoff stress tensor S as the gradient of W , that is,

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

and let

$$A_{i\alpha j\beta}(\eta) := \frac{\partial}{\partial \eta_{j\beta}} S_{i\alpha}(\eta) = \frac{\partial^2}{\partial \eta_{i\alpha} \partial \eta_{j\beta}} W(\eta), \quad \eta \in \mathbb{R}^{d \times d}.$$

Clearly, $A_{i\alpha j\beta}(\eta) = A_{j\beta i\alpha}(\eta)$ for all $\eta \in \mathbb{R}^{d \times d}$ and $i, \alpha, j, \beta = 1, \dots, d$.

Formal calculations show that sufficiently smooth stationary points $u = u(x)$ of the functional J satisfy the following Euler–Lagrange equation

$$-\sum_{\alpha=1}^d \partial_{x_\alpha} S_{i\alpha}(\nabla u(x)) = f_i(x), \quad i = 1, \dots, d, \quad x \in \Omega, \quad (1.1)$$

subject to the boundary conditions

$$u = g_D \quad \text{on } \Gamma_D \quad \text{and} \quad S(\nabla u)\nu = g_N \quad \text{on } \Gamma_N, \quad (1.2)$$

on the Dirichlet and Neumann parts Γ_D and Γ_N of the boundary Γ , respectively. Here ν is the unit outward normal vector to Γ and $\partial_{x_\alpha} = \partial/\partial x_\alpha$.

The weak formulation of the boundary-value problem (1.1), (1.2) is posed as follows: find $u \in H_{D,g_D}^1(\Omega)^d = \{v \in H^1(\Omega)^d : v|_{\Gamma_D} = g_D\}$ such that

$$\int_{\Omega} S(\nabla u) : \nabla v \, dx = \int_{\Omega} f \cdot v \, dx + \int_{\Gamma_N} g_N \cdot v \, ds \quad \forall v \in H_{D,0}^1(\Omega)^d.$$

We shall assume that this problem has a solution $u \in H^{m+1}(\Omega)^d \cap H_{D,g_D}^1(\Omega)^d$ with $m > d/2$. By the Sobolev embedding theorem, u is then, in fact, contained in $C^{1,\hat{\alpha}}(\overline{\Omega})^d$ for some $\hat{\alpha} \in (0, 1)$.

Let \mathcal{M} be a convex open set in $\mathbb{R}^{d \times d}$ such that $\nabla u(\overline{\Omega}) \subset \mathcal{M}$; $\delta > 0$ we define

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

where $|\cdot|$ denotes the Frobenius norm on $\mathbb{R}^{d \times d}$ defined, for $\eta \in \mathbb{R}^{d \times d}$, by $|\eta| = (\eta : \eta)^{1/2}$. Consider the local Lipschitz constant K_{δ} of S in the δ -neighbourhood \mathcal{M}_{δ} of $\nabla u(\overline{\Omega})$,

$$K_{\delta} := \max_{\eta \in \mathcal{M}_{\delta}} \left(\sum_{i,\alpha,j,\beta=1}^d |A_{i\alpha j\beta}(\eta)|^2 \right)^{1/2}. \quad (1.3)$$

We note here that we do not require S to be globally Lipschitz continuous. We also define the set

$$\mathcal{Z}_{\delta} := \{\Phi \in C_{pw}(\overline{\Omega})^{d \times d} : \Phi(x) \in \mathcal{M}_{\delta}, \quad x \in \overline{\Omega}\},$$

where $C_{pw}(\overline{\Omega})$ denotes the set of bounded piecewise continuous functions defined on $\overline{\Omega}$. The set \mathcal{Z}_{δ} will be required in the error analysis of the finite element method: it will contain the *piecewise* gradients of discontinuous Galerkin finite element approximations to u .

Lemma 1 *The set \mathcal{M}_{δ} is a convex subset of $\mathbb{R}^{d \times d}$, and \mathcal{Z}_{δ} is a convex subset of $C_{pw}(\overline{\Omega})^{d \times d}$.*

Proof Trivially, $\mathcal{M} \subset \mathcal{M}_{\delta}$. In order to show the convexity of \mathcal{M}_{δ} , let $\eta, \zeta \in \mathcal{M}_{\delta}$ and consider $\eta_0, \zeta_0 \in \overline{\mathcal{M}}$ such that $\inf_{\sigma \in \mathcal{M}} |\eta - \sigma| = |\eta - \eta_0|$ and $\inf_{\sigma \in \mathcal{M}} |\zeta - \sigma| = |\zeta - \zeta_0|$. As $\overline{\mathcal{M}}$ is closed, such η_0, ζ_0 always exist. Since the closure $\overline{\mathcal{M}}$ of the convex set \mathcal{M} is convex, $\tau\eta_0 + (1 - \tau)\zeta_0 \in \overline{\mathcal{M}}$. Hence, for any $\tau \in [0, 1]$,

$$\begin{aligned} \inf_{\sigma \in \mathcal{M}} |(\tau\eta + (1 - \tau)\zeta) - \sigma| &\leq |\tau\eta + (1 - \tau)\zeta - (\tau\eta_0 + (1 - \tau)\zeta_0)| \\ &= |\tau(\eta - \eta_0) + (1 - \tau)(\zeta - \zeta_0)| \\ &\leq \tau|\eta - \eta_0| + (1 - \tau)|\zeta - \zeta_0| \\ &= \tau \inf_{\sigma \in \mathcal{M}} |\eta - \sigma| + (1 - \tau) \inf_{\sigma \in \mathcal{M}} |\zeta - \sigma| \\ &\leq \tau\delta + (1 - \tau)\delta = \delta, \end{aligned}$$

and thereby $\tau\eta + (1 - \tau)\zeta \in \mathcal{M}_{\delta}$, meaning that \mathcal{M}_{δ} is convex. Finally, the convexity of \mathcal{Z}_{δ} follows from that of \mathcal{M}_{δ} . ■

We shall also consider the dynamic counterpart of the boundary-value problem (1.1), (1.2), — the initial-boundary-value problem for the second-order nonlinear evolution equation

$$\partial_t^2 u_i - \sum_{\alpha=1}^d \partial_{x_\alpha} S_{i\alpha}(\nabla u) = f_i(t, x), \quad i = 1, \dots, d, \quad x \in \Omega, \quad t \in (0, T], \quad (1.4)$$

subject to the initial conditions $u(0, x) = u_0(x)$, $\partial_t(0, x) = u_1(x)$, $x \in \Omega$, and the same boundary conditions as in the static problem above. Here $\partial_t^2 u = \frac{\partial^2 u}{\partial t^2}$; we shall also write \ddot{u} instead of $\partial_t^2 u$ and \dot{u} instead of $\partial_t u = \frac{\partial u}{\partial t}$. For a detailed discussion concerning the physical background to these equations in the field of nonlinear elasticity we refer to [8], for example.

Throughout Sections 2–6 of the paper, we shall assume *uniform ellipticity* in the sense that there exist constants $\delta > 0$ and $M_1 > 0$ such that

$$\sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta}(\chi) \eta_{i\alpha} \eta_{j\beta} \geq M_1 |\eta|^2 \quad \forall \eta \in \mathbb{R}^{d \times d} \quad \forall \chi \in \mathcal{M}_\delta. \quad (1.5)$$

The energy functional $v \mapsto J(v)$ is then uniformly convex over the set $\{v \in W^{1,\infty}(\Omega)^d : \nabla v(x) \in \mathcal{M}_\delta \text{ for a.e. } x \in \Omega\}$. Our main objective here is to show how one can pursue the analysis of discontinuous Galerkin finite element approximations to quasilinear elliptic and hyperbolic problems of the form (1.1) and (1.4) without assuming that the mapping $\eta \mapsto S(\eta)$ is globally Lipschitz continuous and monotone on all of $\mathbb{R}^{d \times d}$.

In Section 7, we shall extend these ideas further by weakening the condition (1.5) to one that is physically more realistic for nonlinear elasticity. We note that in the special case when η is a rank-one matrix, $\eta = \zeta \xi^\top$, with $\zeta, \xi \in \mathbb{R}^d$, (1.5) implies the *strong Legendre–Hadamard condition* or *strong ellipticity condition*: there exist constants $\delta > 0$ and $M_1 > 0$ such that

$$\sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta}(\chi) \zeta_i \zeta_j \xi_\alpha \xi_\beta \geq M_1 |\zeta|^2 |\xi|^2 \quad \forall \zeta, \xi \in \mathbb{R}^d, \quad \forall \chi \in \mathcal{M}_\delta. \quad (1.6)$$

The energy functional $v \mapsto J(v)$ is then only locally rank-1 convex. Nevertheless, strongly elliptic systems of linear partial differential operators of the form

$$u \mapsto - \sum_{\alpha,j,\beta=1}^d \frac{\partial}{\partial x_\alpha} \left(A_{i\alpha j\beta}(x) \frac{\partial u_j}{\partial x_\beta} \right), \quad i = 1, \dots, d,$$

with $A_{i\alpha j\beta} \in C(\overline{\Omega})$, $i, \alpha, j, \beta = 1, \dots, d$, satisfy a Gårding inequality on $H_0^1(\Omega)^d$ (see Theorem 6.5.1 on p.253 in [12]). Gårding's inequality plays a crucial role in the convergence analysis of H^1 -conforming finite element approximations to strongly elliptic systems of linear partial differential equations; however, the finite element space of a discontinuous Galerkin method is not a subspace of $H^1(\Omega)^d$, and therefore we are unable to use

Gårding's inequality in its classical form. In Section 7 we shall show that, in fact, the strong Legendre–Hadamard condition (1.6) implies a *broken Gårding inequality* over the approximation space $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ of the discontinuous Galerkin finite element method consisting of d -component discontinuous piecewise polynomial functions of degree p defined over a partition \mathcal{T}_h of Ω , where \mathcal{T}_h consists of disjoint open Lipschitz domains κ such that $\text{meas}(\kappa) \leq h$, $\overline{\Omega} = \cup \overline{\kappa}$ (cf. Section 2). We shall then use this broken Gårding inequality to extend the results of Section 5 to second-order quasilinear hyperbolic systems assuming the Legendre–Hadamard condition, and still requiring only local (as opposed to global) Lipschitz continuity of the mapping $\eta \mapsto S(\eta)$ on $\mathbb{R}^{d \times d}$. To the best of our knowledge, the analysis of discontinuous Galerkin finite element approximations to second-order quasilinear systems of partial differential equations has not been previously considered in the literature under such weak structural assumptions.

In recent years there has been considerable interest in discontinuous Galerkin finite element methods for the numerical solution of a wide range of partial differential equations which arise from continuum mechanics. We shall not attempt to give a detailed review of this area of research: the reader is referred to [6] for a comprehensive historical survey of the field and [1] and [10] for convergence analyses of the method for second-order linear elliptic problems and partial differential equations with nonnegative characteristic form. Discontinuous Galerkin finite element methods were introduced in the early 1970s for the numerical solution of first-order hyperbolic problems. Simultaneously, but quite independently, they were proposed as nonstandard schemes for the approximation of second-order elliptic equations. The recent upsurge of interest in this class of techniques has been stimulated by the computational convenience of discontinuous Galerkin methods due to their high degree of locality and the presence of associated local conservation properties, as well as the need to accommodate high-order hp - and spectral element discretizations on irregular finite element meshes. The present work has been stimulated by our ongoing research on discontinuous Galerkin methods in the field of fracture mechanics.

The paper is structured as follows. The next section is devoted to the construction of the discontinuous Galerkin method for the nonlinear elliptic boundary-value problem (1.1), (1.2). In Section 3 we develop the linearization of the semilinear form appearing in the definition of the finite element method. In Section 4 we perform the convergence analysis of the discontinuous Galerkin finite element approximation of the uniformly elliptic boundary-value problem (1.1), (1.2). We note, in particular, that our analysis does not assume the global Lipschitz continuity of the functions $S_{i\alpha}$, $i, \alpha = 1, \dots, d$, with respect to ∇u . Building on the work of Makridakis [11] for classical conforming methods, in Section 5 we develop the convergence analysis of semidiscrete discontinuous Galerkin finite element approximations of mixed Dirichlet–Neumann initial-boundary-value problems for systems of second-order quasilinear hyperbolic equations of the form (1.4). In Section 6 we establish a set of auxiliary approximation results for a nonlinear elliptic projector onto the finite element space of discontinuous piecewise polynomials which we require in the analysis of the discontinuous Galerkin finite element method for the hyperbolic problem (1.4) presented in Section 5. In Section 7 we use the Legendre–Hadamard condition to establish a broken Gårding inequality over $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ which

we then apply to extend our results to more general second-order quasilinear hyperbolic systems, such as those that arise in the modelling of nonlinear elastic waves.

2 Finite element spaces

Let \mathcal{T}_h be a subdivision of Ω into disjoint open *element domains* (or, simply, *elements*) κ such that $\overline{\Omega} = \cup_{\kappa \in \mathcal{T}_h} \overline{\kappa}$. Here, $h = \max_{\kappa \in \mathcal{T}_h} h_\kappa$ where $h_\kappa = \text{diam}(\kappa)$. We suppose, without loss of generality, that $h \in (0, 1]$. Each $\kappa \in \mathcal{T}_h$ is assumed to be the image, under a smooth bijective mapping F_κ , of a fixed master element $\hat{\kappa}$; i.e., $\kappa = F_\kappa(\hat{\kappa})$ for all $\kappa \in \mathcal{T}_h$, where $\hat{\kappa}$ is either the open unit simplex or the open unit hypercube in \mathbb{R}^d .

For a nonnegative integer k , we denote by $\mathcal{P}_k(\hat{\kappa})$ the set of polynomials of total degree k on $\hat{\kappa}$. When $\hat{\kappa}$ is the unit hypercube, we also consider $\mathcal{Q}_k(\hat{\kappa})$, the set of all tensor-product polynomials on $\hat{\kappa}$ of degree k in each coordinate direction. To each $\kappa \in \mathcal{T}_h$ we assign a nonnegative integer s_κ (local Sobolev index), collect the s_κ and F_κ in the vectors $\mathbf{s} = \{s_\kappa : \kappa \in \mathcal{T}_h\}$ and $\mathbf{F} = \{F_\kappa : \kappa \in \mathcal{T}_h\}$, respectively, and consider, for $p \geq 1$, the finite element space

$$S^p(\Omega, \mathcal{T}_h, \mathbf{F}) = \{v \in L^2(\Omega)^d : v|_\kappa \circ F_\kappa \in \mathcal{R}_p(\hat{\kappa}) \quad \forall \kappa \in \mathcal{T}_h\},$$

where \mathcal{R} is either \mathcal{P} or \mathcal{Q} .

We assign to the subdivision \mathcal{T}_h the broken (Hilbertian) Sobolev space of composite index \mathbf{s} ,

$$H^{\mathbf{s}}(\Omega, \mathcal{T}_h)^d = \{v \in L^2(\Omega)^d : v|_\kappa \in H^{s_\kappa}(\kappa)^d \quad \forall \kappa \in \mathcal{T}_h\},$$

equipped with the broken Sobolev norm and corresponding seminorm, respectively,

$$\|v\|_{\mathbf{s}, \mathcal{T}_h} = \left(\sum_{\kappa \in \mathcal{T}_h} \|v\|_{H^{s_\kappa}(\kappa)}^2 \right)^{\frac{1}{2}}, \quad |v|_{\mathbf{s}, \mathcal{T}_h} = \left(\sum_{\kappa \in \mathcal{T}_h} |v|_{H^{s_\kappa}(\kappa)}^2 \right)^{\frac{1}{2}}.$$

When $s_\kappa = s$ for all $\kappa \in \mathcal{T}_h$, we shall write $H^s(\Omega, \mathcal{T}_h)^d$, $\|v\|_{s, \mathcal{T}_h}$ and $|v|_{s, \mathcal{T}_h}$. Similarly, one can define the function space $C^k(\overline{\Omega}, \mathcal{T}_h)^d$ of d -component bounded piecewise C^k functions and the broken Sobolev space $W^{s,p}(\Omega, \mathcal{T}_h)^d$ of d -component piecewise $W^{s,p}$ functions, $s > 0$, $1 \leq p < \infty$, over \mathcal{T}_h .

Let us consider the set \mathcal{E} of all $(d-1)$ -dimensional open faces — or, simply, *faces*, — of all elements $\kappa \in \mathcal{T}_h$. Since hanging nodes are permitted (cf. Fig. 1), \mathcal{T}_h may be irregular, and therefore \mathcal{E} will be understood to contain the smallest common $(d-1)$ -dimensional open faces of neighbouring elements. Further, we denote by \mathcal{E}_{int} the set of all e in \mathcal{E} that are contained in Ω , we let $\Gamma_{\text{int}} = \{x \in \Omega : x \in e \text{ for some } e \in \mathcal{E}_{\text{int}}\}$ and we introduce the set \mathcal{E}_D of $(d-1)$ -dimensional boundary faces contained in the subset Γ_D of Γ . Implicit in these definitions is the assumption that \mathcal{T}_h respects the decomposition of Γ in the sense that each $e \in \mathcal{E}$ that lies on Γ belongs to the interior of exactly one of Γ_D or Γ_N . Given $e \in \mathcal{E}$, we define $h_e := \text{diam}(e)$.

In the convergence analysis of the discontinuous Galerkin finite element approximations to the partial differential equations considered here, we shall adopt the following

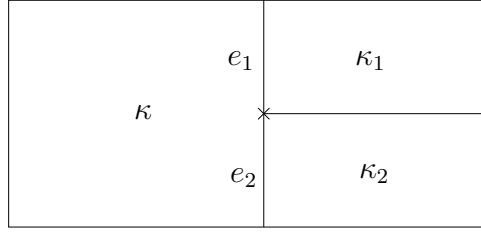


Figure 1: Hanging node \times and faces $e_1, e_2 \in \mathcal{E}_{\text{int}}$.

three hypotheses, the first of which controls the number of hanging nodes which any one element may have, the second strengthens our original hypothesis that the family $\{\mathcal{T}_h\}_{h>0}$ is shape-regular, while the third is a technical condition on the lowest polynomial degree which our analysis admits.

(H1) The family of subdivisions $\{\mathcal{T}_h\}_{h>0}$ is *contact-regular*, i.e., there exist positive constants c_d and c_e independent of h such that, for each $\kappa \in \mathcal{T}_h$,

$$\#\{\kappa' \in \mathcal{T}_h : \kappa' \neq \kappa, \mathcal{H}^{d-1}(\overline{\kappa'} \cap \overline{\kappa}) > 0\} \leq c_d, \quad \text{and} \quad c_e h_\kappa \leq h_e \quad \text{for every face } e \text{ of } \kappa.$$

(H2) The family of subdivisions $\{\mathcal{T}_h\}_{h>0}$ is *quasiuniform*; i.e., there exist positive constants c_0 and c_1 , independent of h , such that for each $\kappa \in \mathcal{T}_h$ there exist open balls $B(x_0, c_0 h)$ and $B(x_1, c_1 h)$ such that $B(x_0, c_0 h) \subset \kappa \subset B(x_1, c_1 h)$.

(H3) In the case of the elliptic problem (1.1) the polynomial degree $p > d/2$, and in the case of the hyperbolic problem (1.4) $p > (d/2) + 1$ (viz. $p \geq 2$ for $d = 2, 3$, and $p \geq 3$ for $d = 2, 3$, respectively).

Remark 1 The last two assumptions become redundant if S is assumed to be globally Lipschitz-continuous in the sense that $|S(\eta) - S(\zeta)| \leq M_2 |\eta - \zeta|$ for all $\eta, \zeta \in \mathbb{R}^d$; see, for example, Houston, Robson & Süli [9]. We note in particular that **H2** implies the existence of a fixed constant $C_4 \geq 1$, independent of h , such that $h/h_\kappa \leq C_4$ for all $\kappa \in \mathcal{T}_h$. **H3** is required in order to deduce, by the use of inverse inequalities from bounds in a broken H^1 norm, that the gradient of the numerical solution lies \mathcal{M}_δ . \blacktriangleleft

Suppose that e is a $(d-1)$ -dimensional open face of an element $\kappa \in \mathcal{T}_h$ and recall the notation introduced above: $h_\kappa = \text{diam}(\kappa)$ and $h_e = \text{diam}(e)$. The following *inverse inequalities* hold: there exists a positive constant C_3 , independent of the discretization parameter h , such that

$$\|\nabla w\|_{L^\infty(\kappa)} \leq \frac{C_3}{h_\kappa^{d/2}} \|\nabla w\|_{L^2(\kappa)}, \quad \|w\|_{L^2(e)}^2 \leq \frac{C_3}{h_e} \|w\|_{L^2(\kappa)}^2, \quad \|\nabla w\|_{L^2(e)}^2 \leq \frac{C_3}{h_e} \|\nabla w\|_{L^2(\kappa)}^2 \quad (2.1)$$

for all $w \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$. In the case of the first inverse inequality C_3 depends only on the shape-regularity parameters of \mathcal{T}_h , while in the case of the other two inequalities it also

depends on the contact-regularity parameter c_e . In fact, h_e in the last two inequalities can be replaced by h_κ , at the expense of possibly altering the value of the constant C_3 .

In the discussion that follows, we shall frequently need to consider the element-wise weak derivative (called broken derivative) and element-wise weak gradient (called broken gradient) of a function that belongs to a broken Sobolev space. In order to simplify the presentation, our notation will not distinguish these from weak derivatives and weak gradients; the implied meaning of the notation will always be clear from the context. Thus, we adopt the following definition.

Definition 1 For $v \in W^{1,1}(\Omega, \mathcal{T}_h)^d$, we denote by $\nabla v \in L^1(\Omega)^{d \times d}$ the broken gradient of v , defined by $(\nabla v)|_\kappa = \nabla(v|_\kappa)$ for each $\kappa \in \mathcal{T}_h$, where the ∇ on the right-hand side signifies a weak gradient applied to $v|_\kappa \in W^{1,1}(\kappa)$. The broken partial derivative $\partial_j v_i = \partial v_i / \partial x_j$ of $v \in W^{1,1}(\Omega, \mathcal{T}_h)^d$ is the (i, j) component of its broken gradient ∇v .

For each $e \in \mathcal{E}_{\text{int}}$ there exist indices i and j such that $i > j$ and κ_i and κ_j share the face e ; we define the (element-numbering-dependent) jump of $v \in H^1(\Omega, \mathcal{T}_h)^d$ across e and the mean value of v on e by

$$[[v]]_e = v|_{\partial\kappa_i \cap e} - v|_{\partial\kappa_j \cap e} \quad \text{and} \quad \langle v \rangle_e = \frac{1}{2} (v|_{\partial\kappa_i \cap e} + v|_{\partial\kappa_j \cap e}),$$

respectively. For the sake of simplicity, the subscript e will be suppressed and we shall simply write $[[v]]$ and $\langle v \rangle$; the implied choice of e will be clear from the context. In addition, we associate with the face e the unit normal vector ν which points from κ_i to κ_j , $i > j$.

We introduce the semilinear form

$$\begin{aligned} B(w, v) = & \int_{\Omega} S(\nabla w) : \nabla v \, dx - \int_{\Gamma_D} S(\nabla w) \nu \cdot v \, ds - \int_{\Gamma_{\text{int}}} \langle S(\nabla w) \nu \rangle \cdot [[v]] \, ds \\ & + \int_{\Gamma_D} \sigma w \cdot v \, ds + \int_{\Gamma_{\text{int}}} \sigma [[w]] \cdot [[v]] \, ds, \quad w \in C^1(\overline{\Omega}, \mathcal{T}_h)^d, \quad v \in W^{1,1}(\Omega, \mathcal{T}_h)^d, \end{aligned} \quad (2.2)$$

and the linear functional

$$\ell(v) = \int_{\Omega} f \cdot v \, dx + \int_{\Gamma_D} \sigma g_D \cdot v \, ds + \int_{\Gamma_N} g_N \cdot v \, ds, \quad v \in W^{1,1}(\Omega, \mathcal{T}_h)^d. \quad (2.3)$$

Here, $h^{-1}|_e = h_e^{-1}$ for all $e \subset \Gamma_D \cup \Gamma_{\text{int}}$. Let $\kappa \in \mathcal{T}_h$ and let e be a $(d-1)$ -dimensional face of $\partial\kappa$. The *discontinuity penalization parameter* σ , featuring in $B(\cdot, \cdot)$ and $\ell(\cdot)$ above, is defined by

$$\sigma|_e = \sigma_e = \frac{\alpha}{h_e} \quad \text{for } e \subset \Gamma_D \cup \Gamma_{\text{int}}. \quad (2.4)$$

Here α is a positive constant whose size will be fixed later on.

The discontinuous Galerkin finite element approximation of problem (1.1), (1.2) is posed as follows: find $u_{\text{DG}} \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ such that

$$B(u_{\text{DG}}, v) = \ell(v) \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \quad (2.5)$$

Remark 2 In (2.2), the role of the fourth and fifth integral is, respectively, to weakly and approximately enforce the Dirichlet boundary condition $u = g_D$ on Γ_D and the continuity condition $\llbracket u \rrbracket = 0$ on Γ_{int} satisfied by the analytical solution u . If the problem were linear, our discretization would correspond to the incomplete interior penalty method of Sun [14]. ◀

3 Linearization

Before embarking on the analysis of the discontinuous Galerkin finite element method (2.5), we shall make some preliminary remarks.

We begin by noting that for any $\eta, \zeta \in \mathcal{M}_\delta$ we have

$$\begin{aligned} S_{i\alpha}(\eta) - S_{i\alpha}(\zeta) &= \sum_{j,\beta=1}^d (\eta_{j\beta} - \zeta_{j\beta}) \int_0^1 \frac{\partial S_{i\alpha}}{\partial \eta_{j\beta}} (\tau\eta + (1-\tau)\zeta) d\tau \\ &= \sum_{j,\beta=1}^d (\eta_{j\beta} - \zeta_{j\beta}) \int_0^1 A_{i\alpha j\beta} (\tau\eta + (1-\tau)\zeta) d\tau, \end{aligned} \quad (3.1)$$

where $\tau\eta + (1-\tau)\zeta \in \mathcal{M}_\delta$ by Lemma 1.

Therefore, for any $w_i \in C^1(\overline{\Omega}, \mathcal{T}_h)^d$ such that $\nabla w_i \in \mathcal{Z}_\delta$, $i = 1, 2$, we have that

$$\begin{aligned} S_{i\alpha}(\nabla w_1(x)) - S_{i\alpha}(\nabla w_2(x)) &= \sum_{j,\beta=1}^d \frac{\partial z_j}{\partial x_\beta} \int_0^1 A_{i\alpha j\beta} (\tau \nabla w_1(x) + (1-\tau) \nabla w_2(x)) d\tau, \\ &\quad x \in \overline{\kappa}, \kappa \in \mathcal{T}_h, \end{aligned}$$

where we have used the abbreviation $z = w_1 - w_2$. By summing over all admissible indices $i, \alpha = 1, \dots, d$, we deduce that

$$\begin{aligned} \sum_{i,\alpha=1}^d (S_{i\alpha}(\nabla w_1(x)) - S_{i\alpha}(\nabla w_2(x))) \frac{\partial v_i}{\partial x_\alpha} \\ = \sum_{i,\alpha,j,\beta=1}^d \left(\int_0^1 A_{i\alpha j\beta} (\nabla w_2(x) + \tau \nabla (w_1(x) - w_2(x))) d\tau \right) \frac{\partial v_i}{\partial x_\alpha} \frac{\partial z_j}{\partial x_\beta}. \end{aligned}$$

Hence, for any $w_1, w_2 \in C^1(\overline{\Omega}, \mathcal{T}_h)^d$ such that $\nabla w_i \in \mathcal{Z}_\delta$, $i = 1, 2$, and all $v \in W^{1,1}(\Omega, \mathcal{T}_h)^d$,

$$\begin{aligned}
B(w_1, v) - B(w_2, v) &= \int_{\Omega} \sum_{i,\alpha,j,\beta=1}^d \left(\int_0^1 A_{i\alpha j\beta}(\tau \nabla w_1 + (1-\tau) \nabla w_2) d\tau \right) \frac{\partial v_i}{\partial x_\alpha} \frac{\partial z_j}{\partial x_\beta} dx \\
&\quad - \int_{\Gamma_D} \sum_{i,\alpha,j,\beta=1}^d \left(\int_0^1 A_{i\alpha j\beta}(\tau \nabla w_1 + (1-\tau) \nabla w_2) d\tau \right) v_i \nu_\alpha \frac{\partial z_j}{\partial x_\beta} ds \\
&\quad - \int_{\Gamma_{\text{int}}} \sum_{i,\alpha,j,\beta=1}^d \left\langle \left(\int_0^1 A_{i\alpha j\beta}(\tau \nabla w_1 + (1-\tau) \nabla w_2) d\tau \right) \nu_\alpha \frac{\partial z_j}{\partial x_\beta} \right\rangle \llbracket v_i \rrbracket ds \\
&\quad + \int_{\Gamma_D} \sigma z \cdot v ds + \int_{\Gamma_{\text{int}}} \sigma \llbracket z \rrbracket \cdot \llbracket v \rrbracket ds, \tag{3.2}
\end{aligned}$$

where $z = w_1 - w_2$. Equivalently,

$$B(w_1, v) - B(w_2, v) = \int_0^1 \tilde{b}(w_2 + \tau(w_1 - w_2); w_1 - w_2, v) d\tau \tag{3.3}$$

where, for $w \in C^1(\overline{\Omega}, \mathcal{T}_h)^d$ with $\nabla w \in \mathcal{Z}_\delta$, $\tilde{b}(w; \cdot, \cdot)$ is a bilinear form defined by

$$\begin{aligned}
\tilde{b}(w; z, v) &= \int_{\Omega} \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta}(\nabla w) \frac{\partial v_i}{\partial x_\alpha} \frac{\partial z_j}{\partial x_\beta} dx - \int_{\Gamma_D} \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta}(\nabla w) v_i \nu_\alpha \frac{\partial z_j}{\partial x_\beta} ds \\
&\quad - \int_{\Gamma_{\text{int}}} \sum_{i,\alpha,j,\beta=1}^d \left\langle A_{i\alpha j\beta}(\nabla w) \nu_\alpha \frac{\partial z_j}{\partial x_\beta} \right\rangle \llbracket v_i \rrbracket ds + \int_{\Gamma_D} \sigma z \cdot v ds + \int_{\Gamma_{\text{int}}} \sigma \llbracket z \rrbracket \cdot \llbracket v \rrbracket ds. \tag{3.4}
\end{aligned}$$

In the next section, we shall use \tilde{b} to perform a convergence analysis of the method (2.5).

4 Convergence analysis

The convergence analysis will be based on Banach's fixed point theorem. We begin by constructing a nonlinear mapping whose unique fixed point is the numerical solution u_{DG} . For this purpose, let us suppose for the moment that $\nabla u_{\text{DG}} \in \mathcal{Z}_\delta$. Further, let $\Pi_h u$ denote the finite element interpolant, from $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$, of the analytical solution u , defined by $(\Pi_h u)|_\kappa := \Pi_p^\kappa(u|_\kappa \circ F_\kappa) \in \mathcal{R}_p(\hat{\kappa})$, where $\Pi_p^\kappa(u|_\kappa \circ F_\kappa)$ is the classical finite element interpolant of $u|_\kappa \circ F_\kappa$ from $\mathcal{R}_p(\hat{\kappa})$. Since for $h \in (0, 1]$ sufficiently small $\nabla \Pi_h u \in \mathcal{Z}_\delta$, we can then take $w_1 = u_{\text{DG}}$ and $w_2 = \Pi_h u$ in the identities (3.3) and (3.4) above. Hence,

$$B(u_{\text{DG}}, v) - B(\Pi_h u, v) = \int_0^1 \tilde{b}(\Pi_h u + \tau(u_{\text{DG}} - \Pi_h u); u_{\text{DG}} - \Pi_h u, v) d\tau.$$

Let us write

$$u - u_{\text{DG}} = (u - \Pi_h u) - (u_{\text{DG}} - \Pi_h u) \equiv \eta - \xi.$$

Note that since $u \in C^1(\overline{\Omega})^d \cap H^2(\Omega, \mathcal{T}_h)^d$, we have that $B(u, v) = \ell(v)$ for all v in $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$; in particular, $B(u, \xi) = \ell(\xi)$. We begin by estimating ξ . Clearly,

$$B(u_{\text{DG}}, v) - B(\Pi_h u, v) = \ell(v) - B(\Pi_h u, v) = B(u, v) - B(\Pi_h u, v) \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).$$

Therefore, for all $v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$,

$$\begin{aligned} & \int_0^1 \tilde{b}(\Pi_h u + \tau(u_{\text{DG}} - \Pi_h u); u_{\text{DG}} - \Pi_h u, v) \, d\tau \\ &= \int_0^1 \tilde{b}(\Pi_h u + \tau(u - \Pi_h u); u - \Pi_h u, v) \, d\tau. \end{aligned} \quad (4.1)$$

Let us consider $H^1(\Omega, \mathcal{T}_h)^d (\supset S^p(\Omega, \mathcal{T}_h, \mathbf{F}))$ equipped with the norm $\|\cdot\|_{1,h}$ defined by

$$\|v\|_{1,h} = \left(\int_{\Omega} |\nabla v|^2 \, dx + \int_{\Gamma_D} \sigma |v|^2 \, ds + \int_{\Gamma_{\text{int}}} \sigma \llbracket v \rrbracket^2 \, ds \right)^{1/2},$$

induced by the inner product $(\cdot, \cdot)_{1,h}$, where

$$(w, v)_{1,h} = \int_{\Omega} \nabla w : \nabla v \, dx + \int_{\Gamma_D} \sigma w \cdot v \, ds + \int_{\Gamma_{\text{int}}} \sigma \llbracket w \rrbracket \cdot \llbracket v \rrbracket \, ds.$$

4.1 Construction of the fixed-point map

Let us recall our hypotheses that $u \in H^{m+1}(\Omega)^d$ with $m > d/2$ and that the polynomial degree $p > d/2$. Let $d/2 < r \leq \min(m, p)$, and define the following subset of the broken Sobolev space $H^1(\Omega, \mathcal{T}_h)^d$:

$$\mathcal{J} = \{v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}) : \|v - \Pi_h u\|_{1,h} \leq C_* h^r \|u\|_{H^{r+1}(\Omega)}\},$$

where C_* is a fixed positive constant whose value will be made explicit below (cf. (4.18)). We note that since $\Pi_h u \in \mathcal{J}$, the set \mathcal{J} is nonempty. Further, \mathcal{J} is a closed, convex subset of $H^1(\Omega, \mathcal{T}_h)^d$ in the topology induced by the norm $\|\cdot\|_{1,h}$. Finally, we note that for each $v \in \mathcal{J}$, using (2.1), we have

$$\begin{aligned} \|\nabla v - \nabla u\|_{L^\infty(\Omega)} &\leq \|\nabla v - \nabla \Pi_h u\|_{L^\infty(\Omega)} + \|\nabla \Pi_h u - \nabla u\|_{L^\infty(\Omega)} \\ &\leq C_* C_3 h^{r-d/2} \|u\|_{H^{r+1}(\Omega)} + \|\nabla \Pi_h u - \nabla u\|_{L^\infty(\Omega)}. \end{aligned}$$

Hence, given $\delta > 0$, there exists $h_0 \in (0, 1]$ such that, for all $h \in (0, h_0]$,

$$\varphi \in \mathcal{J} \Rightarrow \nabla \varphi \in \mathcal{Z}_\delta. \quad (4.2)$$

Motivated by the form of (4.1), we define the fixed point mapping \mathcal{N} on \mathcal{J} as follows. Given $\varphi \in \mathcal{J}$, we denote by $u_\varphi = \mathcal{N}(\varphi) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ the solution to the following linear variational problem: find $u_\varphi \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ such that

$$\begin{aligned} \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); u_\varphi - \Pi_h u, v) \, d\tau &= \int_0^1 \tilde{b}(\Pi_h u + \tau(u - \Pi_h u); u - \Pi_h u, v) \, d\tau, \\ &\quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \end{aligned} \quad (4.3)$$

Equivalently, find $u_\varphi \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ such that

$$\begin{aligned} \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); u_\varphi, v) \, d\tau &= \int_0^1 \tilde{b}(\Pi_h u + \tau(u - \Pi_h u); u - \Pi_h u, v) \, d\tau \\ &+ \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); \Pi_h u, v) \, d\tau, \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \end{aligned}$$

Since $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ is a finite-dimensional linear space, the existence and uniqueness of a solution $u_\varphi \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ to problem (4.3) will follow once we have shown that the bilinear form

$$(w, v) \mapsto \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); w, v) \, d\tau$$

is coercive on $S^p(\Omega, \mathcal{T}_h, \mathbf{F}) \times S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ in the norm $\|\cdot\|_{1,h}$. Next, we shall show that this can indeed be ensured by taking the factor $\alpha > 0$ featuring in the definition of the discontinuity penalization parameter σ defined by (2.4), which enters into the semilinear form (2.2) of the discontinuous Galerkin method (2.5), sufficiently large.

4.2 Coercivity of the bilinear form \tilde{b}

For $\varphi \in \mathcal{J}$ fixed and $v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ we consider

$$\begin{aligned} \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); v, v) \, d\tau &= \\ &\int_0^1 \left[\int_\Omega \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta} (\nabla \Pi_h u + \tau \nabla (\varphi - \Pi_h u)) \frac{\partial v_i}{\partial x_\alpha} \frac{\partial v_j}{\partial x_\beta} \, dx \right. \\ &\quad - \int_{\Gamma_D} \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta} (\nabla \Pi_h u + \tau \nabla (\varphi - \Pi_h u)) v_i \nu_\alpha \frac{\partial v_j}{\partial x_\beta} \, ds \\ &\quad - \left. \int_{\Gamma_{\text{int}}} \sum_{i,\alpha,j,\beta=1}^d \left\langle A_{i\alpha j\beta} (\nabla \Pi_h u + \tau \nabla (\varphi - \Pi_h u)) \nu_\alpha \frac{\partial v_j}{\partial x_\beta} \right\rangle \llbracket v_i \rrbracket \, ds \right] \, d\tau \\ &\quad + \int_{\Gamma_D} \sigma |v|^2 \, ds + \int_{\Gamma_{\text{int}}} \sigma \llbracket v \rrbracket^2 \, ds \\ &\equiv T_1 + T_2 + T_3 + T_4 + T_5. \end{aligned}$$

By (1.5) we have that

$$T_1 \geq M_1 \|\nabla v\|_{L^2(\Omega)}^2.$$

Next, we bound T_2 . Since \mathcal{J} is a convex set, $\nabla \Pi_h u(x) + \tau \nabla (\varphi(x) - \Pi_h u(x)) \in \mathcal{M}_\delta$ for all $x \in \bar{\kappa}$ and all $\kappa \in \mathcal{T}_h$, — provided that $h \in (0, h_0]$, where h_0 is as in (4.2).

Since $\theta \in \mathcal{M}_\delta \mapsto A_{i\alpha j\beta}(\theta) \in \mathbb{R}$ is a continuous function and \mathcal{M}_δ is a compact subset of $\mathbb{R}^{d \times d}$, we have that

$$\max_{\kappa \in \mathcal{T}_h} \max_{x \in \bar{\kappa}} \max_{\theta \in \text{conv}(\nabla \varphi(x), \nabla \Pi_h u(x))} \left(\sum_{i,\alpha,j,\beta=1}^d |A_{i\alpha j\beta}(\theta)|^2 \right)^{1/2} \leq K_\delta \quad \forall h \in (0, h_0],$$

where K_δ is the positive constant defined in (1.3); clearly, K_δ is independent of h and φ . Hence,

$$\begin{aligned} |\mathbf{T}_2| &\leq K_\delta \int_{\Gamma_D} \left(\sum_{i,\alpha,j,\beta=1}^d |v_i|^2 |\nu_\alpha|^2 \left| \frac{\partial v_j}{\partial x_\beta} \right|^2 \right)^{1/2} ds \\ &\leq K_\delta \left(\int_{\Gamma_D} \sigma^{-1} |\nabla v|^2 ds \right)^{1/2} \left(\int_{\Gamma_D} \sigma |v|^2 ds \right)^{1/2}. \end{aligned}$$

Hence, using the second of the inverse inequalities (2.1) and recalling the definition of the penalty parameter σ_e on $e \subset \Gamma_D$, we have that

$$\mathbf{T}_2 \leq K_\delta (C_3 \alpha^{-1} 2d)^{1/2} \left(\int_{\Omega} |\nabla v|^2 dx \right)^{1/2} \left(\int_{\Gamma_D} \sigma |v|^2 ds \right)^{1/2},$$

where $2d$ stands for the maximum number of faces any one element may have on Γ_D .

Analogously,

$$\mathbf{T}_3 \leq K_\delta \int_{\Gamma_{\text{int}}} \langle |\nabla v| \rangle \llbracket v \rrbracket ds \leq K_\delta \left(\int_{\Gamma_{\text{int}}} \sigma^{-1} \langle |\nabla v| \rangle^2 ds \right)^{1/2} \left(\int_{\Gamma_{\text{int}}} \sigma \llbracket v \rrbracket^2 ds \right)^{1/2}.$$

Let us note that

$$\int_{\Gamma_{\text{int}}} \sigma^{-1} \langle |\nabla v| \rangle^2 ds = \sum_{e \in \mathcal{E}_{\text{int}}} \sigma_e^{-1} \int_e \langle |\nabla v| \rangle^2 ds,$$

and, for $e \in \mathcal{E}_{\text{int}}$, let κ and κ' be the two elements that share e . Then,

$$\begin{aligned} \int_e \langle |\nabla v| \rangle^2 ds &\leq \frac{1}{2} \int_e |\nabla v|_\kappa|^2 ds + \frac{1}{2} \int_e |\nabla v|_{\kappa'}|^2 ds \\ &\leq \frac{C_3}{2h_e} \int_\kappa |\nabla v|^2 dx + \frac{C_3}{2h_e} \int_{\kappa'} |\nabla v|^2 dx \\ &\leq \frac{C_3}{h_e} \max \left\{ \int_\kappa |\nabla v|^2 dx, \int_{\kappa'} |\nabla v|^2 dx \right\}. \end{aligned}$$

On recalling from the definition of σ that $\sigma_e = \alpha/h_e$ for $e \in \mathcal{E}_{\text{int}}$, we have that

$$\sum_{e \in \mathcal{E}_{\text{int}}} \sigma_e^{-1} \int_e \langle |\nabla v| \rangle^2 ds \leq C_3 \alpha^{-1} \sum_{e \in \mathcal{E}_{\text{int}}} \max_{\{\kappa : e \subset \partial \kappa\}} \int_\kappa |\nabla v|^2 dx.$$

Thanks to our assumption of contact-regularity, it follows that no element κ can have more than c_d faces, where c_d is a finite number independent of h . We have

$$\sum_{e \in \mathcal{E}_{\text{int}}} \sigma_e^{-1} \int_e \langle |\nabla v| \rangle^2 ds \leq C_3 \alpha^{-1} c_d \sum_{\kappa \in \mathcal{T}_h} \int_\kappa |\nabla v|^2 dx,$$

and therefore,

$$\mathbf{T}_3 \leq K_\delta (C_3 \alpha^{-1} c_d)^{1/2} \left(\int_{\Omega} |\nabla v|^2 dx \right)^{1/2} \left(\int_{\Gamma_{\text{int}}} \sigma \llbracket v \rrbracket^2 ds \right)^{1/2}. \quad (4.4)$$

Using the lower bound on T_1 and the upper bounds on T_2 and T_3 , we thus deduce that

$$\begin{aligned} \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); v, v) d\tau &\geq M_1 \int_{\Omega} |\nabla v|^2 dx + \int_{\Gamma_D} \sigma |v|^2 ds + \int_{\Gamma_{\text{int}}} \sigma \llbracket v \rrbracket^2 ds \\ &\quad - K_{\delta} (C_3 \alpha^{-1} 2d)^{1/2} \left(\int_{\Omega} |\nabla v|^2 dx \right)^{1/2} \left(\int_{\Gamma_D} \sigma |v|^2 ds \right)^{1/2} \\ &\quad - K_{\delta} (C_3 \alpha^{-1} c_d)^{1/2} \left(\int_{\Omega} |\nabla v|^2 dx \right)^{1/2} \left(\int_{\Gamma_{\text{int}}} \sigma \llbracket v \rrbracket^2 ds \right)^{1/2}. \end{aligned}$$

Applying Cauchy's inequality $ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$ to the last two terms on the right-hand side and defining $C_d = c_d + 2d$, we have

$$\begin{aligned} \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); v, v) d\tau &\geq M_1 \left(1 - \frac{K_{\delta}^2 C_3 C_d}{2M_1 \alpha} \right) \int_{\kappa} |\nabla v|^2 dx \\ &\quad + \frac{1}{2} \int_{\Gamma_D} \sigma |v|^2 ds + \frac{1}{2} \int_{\Gamma_{\text{int}}} \sigma \llbracket v \rrbracket^2 ds. \end{aligned}$$

Thus, on selecting α such that $\alpha \geq K_{\delta}^2 M_1^{-1} C_3 C_d$, we deduce that, for all $h \in (0, h_0]$,

$$\int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); v, v) d\tau \geq \frac{1}{2} \min(1, M_1) \|v\|_{1,h}^2 \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \quad (4.5)$$

which is the required coercivity of the bilinear form

$$(v, w) \mapsto \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); v, w) d\tau$$

on the finite-dimensional space $S^p(\Omega, \mathcal{T}_h, \mathbf{F}) \times S^p(\Omega, \mathcal{T}_h, \mathbf{F})$. Hence we deduce that for any $\varphi \in \mathcal{J}$ there exists $u_{\varphi} \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ which solves (4.3), so the fixed point mapping \mathcal{N} is correctly defined. Next, we refine this statement by showing that, in fact, $u_{\varphi} \in \mathcal{J}$ (which is a subset of $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$).

4.3 Proof that \mathcal{N} maps \mathcal{J} into itself

We begin by noting that since $u_{\varphi} - \Pi_h u$ is an element of $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$, the inequality (4.5) implies that

$$\begin{aligned} \frac{1}{2} \min(1, M_1) \|u_{\varphi} - \Pi_h u\|_{1,h}^2 &\leq \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); u_{\varphi} - \Pi_h u, u_{\varphi} - \Pi_h u) d\tau \\ &= B(u, u_{\varphi} - \Pi_h u) - B(\Pi_h u, u_{\varphi} - \Pi_h u). \end{aligned} \quad (4.6)$$

Next, we establish an upper bound on the right-hand side of (4.6). Using the fact that $|\nabla v \cdot \nu| \leq |\nabla v|$, we have

$$\begin{aligned}
|B(u, u_\varphi - \Pi_h u) - B(\Pi_h u, u_\varphi - \Pi_h u)| &\leq \int_{\Omega} |S(\nabla u) - S(\nabla \Pi_h u)| |\nabla u_\varphi - \nabla \Pi_h u| dx \\
&+ \int_{\Gamma_D} |S(\nabla u) - S(\nabla \Pi_h u)| |u_\varphi - \Pi_h u| ds + \int_{\Gamma_{\text{int}}} \langle |S(\nabla u) - S(\nabla \Pi_h u)| \rangle \llbracket u_\varphi - \Pi_h u \rrbracket ds \\
&+ \int_{\Gamma_D} \sigma |u - \Pi_h u| |u_\varphi - \Pi_h u| ds + \int_{\Gamma_{\text{int}}} \sigma \llbracket u - \Pi_h u \rrbracket \llbracket u_\varphi - \Pi_h u \rrbracket ds \\
&\equiv T_1 + T_2 + T_3 + T_4 + T_5.
\end{aligned} \tag{4.7}$$

We note that for any $\eta, \zeta \in \mathcal{M}_\delta$ we have from (3.1) the bound

$$\sum_{i, \alpha=1}^d |S_{i\alpha}(\eta) - S_{i\alpha}(\zeta)|^2 \leq K_\delta^2 |\eta - \zeta|^2.$$

This implies, on taking square-roots, that

$$|S(\eta) - S(\zeta)| \leq K_\delta |\eta - \zeta|, \quad \eta, \zeta \in \mathcal{M}_\delta. \tag{4.8}$$

On selecting $\eta = \nabla u$ and $\zeta = \nabla \Pi_h u$, we deduce that, for all $x \in \bar{\kappa}$ and all $\kappa \in \mathcal{T}_h$,

$$|S(\nabla u(x)) - S(\nabla \Pi_h u(x))| \leq K_\delta |\nabla u(x) - \nabla \Pi_h u(x)|.$$

Therefore, for $h \in (0, h_0]$,

$$\left(\int_{\Omega} |S(\nabla u(x)) - S(\nabla \Pi_h u(x))|^2 dx \right)^{1/2} \leq K_\delta \left(\int_{\Omega} |\nabla u(x) - \nabla \Pi_h u(x)|^2 dx \right)^{1/2}. \tag{4.9}$$

Applying (4.9) and the Cauchy-Schwarz inequality, it follows that

$$T_1 \leq K_\delta \left(\int_{\Omega} |\nabla u(x) - \nabla \Pi_h u(x)|^2 dx \right)^{1/2} \left(\int_{\Omega} |\nabla u_\varphi - \nabla \Pi_h u|^2 dx \right)^{1/2}. \tag{4.10}$$

For T_4 and T_5 , we have

$$T_4 \leq \left(\int_{\Gamma_D} \sigma |u - \Pi_h u|^2 ds \right)^{1/2} \left(\int_{\Gamma_D} \sigma |u_\varphi - \Pi_h u|^2 ds \right)^{1/2}, \tag{4.11}$$

$$T_5 \leq \left(\int_{\Gamma_{\text{int}}} \sigma \llbracket u - \Pi_h u \rrbracket^2 ds \right)^{1/2} \left(\int_{\Gamma_{\text{int}}} \sigma \llbracket u_\varphi - \Pi_h u \rrbracket^2 ds \right)^{1/2}. \tag{4.12}$$

For T_2 , we use (4.8) again to estimate

$$\begin{aligned}
T_2 &\leq \left(\int_{\Gamma_D} \sigma^{-1} |S(\nabla u) - S(\nabla \Pi_h u)|^2 ds \right)^{1/2} \left(\int_{\Gamma_D} \sigma |u_\varphi - \Pi_h u|^2 ds \right)^{1/2} \\
&\leq K_\delta \left(\int_{\Gamma_D} \sigma^{-1} |\nabla u - \nabla \Pi_h u|^2 ds \right)^{1/2} \left(\int_{\Gamma_D} \sigma |u_\varphi - \Pi_h u|^2 ds \right)^{1/2},
\end{aligned}$$

for all $h \in (0, h_0]$.

Hence, recalling the definition of the penalty parameter σ_e on $e \in \Gamma_D$, we have that

$$T_2 \leq K_\delta \alpha^{-1/2} \left(\sum_{e \in \mathcal{E}_D} h_e \int_e |\nabla u - \nabla \Pi_h u|^2 ds \right)^{1/2} \left(\int_{\Gamma_D} \sigma |u_\varphi - \Pi_h u|^2 ds \right)^{1/2}. \quad (4.13)$$

We proceed analogously for T_3 . We begin by observing that

$$\begin{aligned} T_3 &\leq K_\delta \int_{\Gamma_{\text{int}}} \langle |\nabla u - \nabla \Pi_h u| \rangle \llbracket u_\varphi - \Pi_h u \rrbracket ds \\ &\leq K_\delta \left(\int_{\Gamma_{\text{int}}} \sigma^{-1} \langle |\nabla u - \nabla \Pi_h u| \rangle^2 ds \right)^{1/2} \left(\int_{\Gamma_{\text{int}}} \sigma \llbracket u_\varphi - \Pi_h u \rrbracket^2 ds \right)^{1/2}. \end{aligned}$$

Let us write $\eta = u - \Pi_h u$ and note that

$$\int_{\Gamma_{\text{int}}} \sigma^{-1} \langle |\nabla \eta| \rangle^2 ds = \sum_{e \in \mathcal{E}_{\text{int}}} \sigma_e^{-1} \int_e \langle |\nabla \eta| \rangle^2 ds.$$

For $e \in \mathcal{E}_{\text{int}}$, let κ and κ' be the two elements that share e . Then,

$$\int_e \langle |\nabla \eta| \rangle^2 ds \leq \frac{1}{2} \int_e |\nabla \eta|_\kappa|^2 ds + \frac{1}{2} \int_e |\nabla \eta|_{\kappa'}|^2 ds \leq \max \left\{ \int_e |\nabla \eta|_\kappa|^2 ds, \int_e |\nabla \eta|_{\kappa'}|^2 ds \right\}.$$

On recalling from the definition of σ that $\sigma_e = \alpha/h_e$ for $e \in \mathcal{E}_{\text{int}}$, we have that

$$\sum_{e \in \mathcal{E}_{\text{int}}} \sigma_e^{-1} \int_e \langle |\nabla \eta| \rangle^2 ds \leq \alpha^{-1} \sum_{e \in \mathcal{E}_{\text{int}}} h_e \max_{\{\kappa : e \subset \partial \kappa\}} \int_e |\nabla \eta|_\kappa|^2 dx,$$

and hence

$$\begin{aligned} T_3 &\leq K_\delta \alpha^{-1/2} \left(\sum_{e \in \mathcal{E}_{\text{int}}} h_e \max_{\{\kappa : e \subset \partial \kappa\}} \int_e |(\nabla u - \Pi_h u)|_\kappa|^2 ds \right)^{1/2} \\ &\quad \times \left(\int_{\Gamma_{\text{int}}} \sigma \llbracket u_\varphi - \Pi_h u \rrbracket^2 ds \right)^{1/2}. \end{aligned} \quad (4.14)$$

Substituting the bounds on T_1, \dots, T_5 into (4.7), we deduce that, for any $h \in (0, h_0]$,

$$\begin{aligned}
|B(u, u_\varphi - \Pi_h u) - B(\Pi_h u, u_\varphi - \Pi_h u)| &\leq \\
&K_\delta \left(\int_\Omega |\nabla u(x) - \nabla \Pi_h u(x)|^2 dx \right)^{1/2} \left(\int_\Omega |\nabla u_\varphi - \nabla \Pi_h u|^2 dx \right)^{1/2} \\
&+ K_\delta \alpha^{-1/2} \left(\sum_{e \in \mathcal{E}_D} h_e \int_e |\nabla u - \nabla \Pi_h u|^2 ds \right)^{1/2} \left(\int_{\Gamma_D} \sigma |u_\varphi - \Pi_h u|^2 ds \right)^{1/2} \\
&+ K_\delta \alpha^{-1/2} \left(\sum_{e \in \mathcal{E}_{\text{int}}} h_e \max_{\{\kappa : e \subset \partial \kappa\}} \int_e |(\nabla u - \Pi_h u)|_\kappa|^2 ds \right)^{1/2} \left(\int_{\Gamma_{\text{int}}} \sigma \llbracket u_\varphi - \Pi_h u \rrbracket^2 ds \right)^{1/2} \\
&+ \alpha^{1/2} \left(\sum_{e \in \mathcal{E}_D} h_e^{-1} \int_e |u - \Pi_h u|^2 ds \right)^{1/2} \left(\int_{\Gamma_D} \sigma |u_\varphi - \Pi_h u|^2 ds \right)^{1/2} \\
&+ \alpha^{1/2} \left(\sum_{e \in \mathcal{E}_{\text{int}}} h_e^{-1} \int_e \llbracket u - \Pi_h u \rrbracket^2 ds \right)^{1/2} \left(\int_{\Gamma_{\text{int}}} \sigma \llbracket u_\varphi - \Pi_h u \rrbracket^2 ds \right)^{1/2}. \quad (4.15)
\end{aligned}$$

Hence,

$$\begin{aligned}
|B(u, u_\varphi - \Pi_h u) - B(\Pi_h u, u_\varphi - \Pi_h u)| &\leq \max(K_\delta, (K_\delta^2/\alpha + \alpha)^{1/2}) \|u_\varphi - \Pi_h u\|_{1,h} \\
&\times \left(\int_\Omega |\nabla u(x) - \nabla \Pi_h u(x)|^2 dx + \sum_{e \in \mathcal{E}_D} h_e \int_e |\nabla u - \nabla \Pi_h u|^2 ds \right. \\
&+ \sum_{e \in \mathcal{E}_D} h_e^{-1} \int_e |u - \Pi_h u|^2 ds + \sum_{e \in \mathcal{E}_{\text{int}}} h_e \max_{\{\kappa : e \subset \partial \kappa\}} \int_e |(\nabla u - \Pi_h u)|_\kappa|^2 ds \\
&\left. + \sum_{e \in \mathcal{E}_{\text{int}}} h_e^{-1} \int_e \llbracket u - \Pi_h u \rrbracket^2 ds \right)^{1/2}. \quad (4.16)
\end{aligned}$$

Now, using the approximation properties of the projector $\Pi_h u$, we deduce that

$$\begin{aligned}
|B(u, u_\varphi - \Pi_h u) - B(\Pi_h u, u_\varphi - \Pi_h u)| \\
\leq C_5 \max(K_\delta, (K_\delta^2/\alpha + \alpha)^{1/2}) h^r \|u\|_{H^{r+1}(\Omega)} \|u_\varphi - \Pi_h u\|_{1,h}.
\end{aligned}$$

Therefore, by (4.6),

$$\|u_\varphi - \Pi_h u\|_{1,h} \leq C_* h^r \|u\|_{H^{r+1}(\Omega)}, \quad d/2 < r \leq \min(m, p), \quad (4.17)$$

for all $h \in (0, h_0]$, where

$$C_* = \frac{2C_5}{\min\{1, M_1\}} \max \left\{ K_\delta, \left(\frac{K_\delta^2}{\alpha} + \alpha \right)^{\frac{1}{2}} \right\}. \quad (4.18)$$

Note that, while h_0 depends on C_* , the constant C_* does not depend on h_0 .

4.4 Proof of contractivity

It remains to show that \mathcal{N} is a contraction in the norm $\|\cdot\|_{1,h}$. To do so, let us suppose that φ and ψ belong to \mathcal{J} and let $u_\varphi = \mathcal{N}(\varphi)$ and $u_\psi = \mathcal{N}(\psi)$. Then,

$$\begin{aligned} \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); u_\varphi - u_\psi, v) \, d\tau &= \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); \Pi_h u - u_\psi, v) \, d\tau \\ &\quad - \int_0^1 \tilde{b}(\Pi_h u + \tau(\psi - \Pi_h u); \Pi_h u - u_\psi, v) \, d\tau. \end{aligned}$$

Upon choosing $v = u_\varphi - u_\psi$ and setting $w = \Pi_h u - u_\psi$, we deduce that

$$\begin{aligned} \frac{1}{2} \min\{1, M_1\} \|u_\varphi - u_\psi\|_{1,h}^2 &\leq \int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); \Pi_h u - u_\psi, u_\varphi - u_\psi) \, d\tau \\ &\quad - \int_0^1 \tilde{b}(\Pi_h u + \tau(\psi - \Pi_h u); \Pi_h u - u_\psi, u_\varphi - u_\psi) \, d\tau \\ &= \int_0^1 \left[\int_\Omega \sum_{i,\alpha,j,\beta=1}^d [A_{i\alpha j\beta}(\nabla \Pi_h u + \tau \nabla(\varphi - \Pi_h u)) \right. \\ &\quad \left. - A_{i\alpha j\beta}(\nabla \Pi_h u + \tau \nabla(\psi - \Pi_h u))] \frac{\partial v_i}{\partial x_\alpha} \frac{\partial w_j}{\partial x_\beta} \, dx \right. \\ &\quad - \int_{\Gamma_D} \sum_{i,\alpha,j,\beta=1}^d [A_{i\alpha j\beta}(\nabla \Pi_h u + \tau \nabla(\varphi - \Pi_h u)) \\ &\quad \left. - A_{i\alpha j\beta}(\nabla \Pi_h u + \tau \nabla(\psi - \Pi_h u))] v_i \nu_\alpha \frac{\partial w_j}{\partial x_\beta} \, ds \right. \\ &\quad \left. - \int_{\Gamma_{\text{int}}} \sum_{i,\alpha,j,\beta=1}^d \left\langle [A_{i\alpha j\beta}(\nabla \Pi_h u + \tau \nabla(\varphi - \Pi_h u)) \right. \right. \\ &\quad \left. \left. - A_{i\alpha j\beta}(\nabla \Pi_h u + \tau \nabla(\psi - \Pi_h u))] \nu_\alpha \frac{\partial w_j}{\partial x_\beta} \right\rangle \llbracket v_i \rrbracket \, ds \right] \, d\tau \\ &\equiv T_1 + T_2 + T_3. \end{aligned} \tag{4.19}$$

Since both $\nabla \Pi_h u + \tau \nabla(\varphi - \Pi_h u)$ and $\nabla \Pi_h u + \tau \nabla(\psi - \Pi_h u)$ belong to \mathcal{Z}_δ , and \mathcal{Z}_δ is a convex set, any convex combination of these two elements belongs to \mathcal{Z}_δ . As $A_{i\alpha j\beta}$ is a C^1 function on the compact set \mathcal{Z}_δ , it is, in particular, a Lipschitz continuous function on the closed convex hull $\text{conv}(\nabla \Pi_h u(x), \nabla \varphi(x), \nabla \psi(x))$ of $\nabla \Pi_h u(x)$, $\nabla \varphi(x)$, and $\nabla \psi(x)$ for any fixed $x \in \bar{\kappa}$, $\kappa \in \mathcal{T}_h$. Therefore,

$$L_{i\alpha j\beta} = \max_{\varphi, \psi \in \mathcal{J}} \max_{\kappa \in \mathcal{T}_h} \max_{x \in \bar{\kappa}} \max_{\eta \in \text{conv}(\nabla \Pi_h u(x), \nabla \varphi(x), \nabla \psi(x))} \left(\sum_{k,\gamma=1}^d \left| \frac{\partial}{\partial \eta_{k\gamma}} A_{i\alpha j\beta}(\eta) \right|^2 \right)^{1/2} < \infty.$$

Hence, for any $\tau \in [0, 1]$,

$$|A_{i\alpha j\beta}(\nabla \Pi_h u + \tau \nabla(\varphi - \Pi_h u)) - A_{i\alpha j\beta}(\nabla \Pi_h u + \tau \nabla(\psi - \Pi_h u))| \leq L_{i\alpha j\beta} |\nabla \varphi - \nabla \psi|.$$

Writing $|L| = \left(\sum_{i,\alpha,j,\beta=1}^d |L_{i\alpha j\beta}|^2 \right)^{1/2}$, we deduce that

$$\begin{aligned}
T_1 &\leq \int_{\Omega} \sum_{i,\alpha,j,\beta=1}^d L_{i\alpha j\beta} |\nabla \varphi - \nabla \psi| \left| \frac{\partial v_i}{\partial x_{\alpha}} \right| \left| \frac{\partial w_j}{\partial x_{\beta}} \right| dx \\
&\leq |L| \int_{\Omega} |\nabla \varphi - \nabla \psi| |\nabla v| |\nabla w| dx \\
&\leq |L| \|\nabla w\|_{L^{\infty}(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} \\
&= |L| \|\nabla u_{\psi} - \nabla \Pi_h u\|_{L^{\infty}(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\nabla u_{\varphi} - \nabla u_{\psi}\|_{L^2(\Omega)}.
\end{aligned}$$

For T_2 , we have, completely analogously, and recalling the definitions of v and w ,

$$\begin{aligned}
T_2 &\leq \int_{\Gamma_D} \sum_{i,\alpha,j,\beta=1}^d L_{i\alpha j\beta} |\nabla \varphi - \nabla \psi| |v_i| |\nu_{\alpha}| \left| \frac{\partial w_j}{\partial x_{\beta}} \right| ds \\
&\leq |L| \int_{\Gamma_D} |\nabla \varphi - \nabla \psi| |v| |\nabla w| ds \\
&\leq |L| \|\nabla u_{\psi} - \nabla \Pi_h u\|_{L^{\infty}(\Gamma_D)} \|\sigma^{-1/2}(\nabla \varphi - \nabla \psi)\|_{L^2(\Gamma_D)} \|\sigma^{1/2}(u_{\varphi} - u_{\psi})\|_{L^2(\Gamma_D)} \\
&\leq |L| \|\nabla u_{\psi} - \nabla \Pi_h u\|_{L^{\infty}(\Omega)} \|\sigma^{-1/2}(\nabla \varphi - \nabla \psi)\|_{L^2(\Gamma_D)} \|\sigma^{1/2}(u_{\varphi} - u_{\psi})\|_{L^2(\Gamma_D)} \\
&\leq (C_3 \alpha^{-1} 2d)^{1/2} |L| \|\nabla u_{\psi} - \nabla \Pi_h u\|_{L^{\infty}(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}(u_{\varphi} - u_{\psi})\|_{L^2(\Gamma_D)}.
\end{aligned}$$

For T_3 , an identical argument gives that

$$T_3 \leq (C_3 \alpha^{-1} c_d)^{1/2} |L| \|\nabla u_{\psi} - \nabla \Pi_h u\|_{L^{\infty}(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}(u_{\varphi} - u_{\psi})\|_{L^2(\Gamma_{\text{int}})},$$

where c_d is as before. Inserting the bounds on T_1, T_2, T_3 into (4.19) we get

$$\begin{aligned}
\frac{1}{2} \min\{1, M_1\} \|u_{\varphi} - u_{\psi}\|_{1,h}^2 &\leq \\
&|L| \|\nabla u_{\psi} - \nabla \Pi_h u\|_{L^{\infty}(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\nabla u_{\varphi} - \nabla u_{\psi}\|_{L^2(\Omega)} \\
&+ (C_3 \alpha^{-1} 2d)^{1/2} |L| \|\nabla u_{\psi} - \nabla \Pi_h u\|_{L^{\infty}(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}(u_{\varphi} - u_{\psi})\|_{L^2(\Gamma_D)} \\
&+ (C_3 \alpha^{-1} c_d)^{1/2} |L| \|\nabla u_{\psi} - \nabla \Pi_h u\|_{L^{\infty}(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}(u_{\varphi} - u_{\psi})\|_{L^2(\Gamma_{\text{int}})}.
\end{aligned}$$

This yields

$$\begin{aligned}
\frac{1}{2} \min\{1, M_1\} \|u_{\varphi} - u_{\psi}\|_{1,h}^2 &\leq |L| (1 + C_3 \alpha^{-1} (2d + c_d))^{1/2} \\
&\times \|\nabla u_{\psi} - \nabla \Pi_h u\|_{L^{\infty}(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|u_{\varphi} - u_{\psi}\|_{1,h},
\end{aligned}$$

whereby,

$$\|u_{\varphi} - u_{\psi}\|_{1,h} \leq \frac{2 |L| (1 + C_3 \alpha^{-1} (2d + c_d))^{1/2}}{\min\{1, M_1\}} \|\nabla u_{\psi} - \nabla \Pi_h u\|_{L^{\infty}(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)}.$$

Bounding the last factor on the right-hand side further using the definition of the norm $\|\cdot\|_{1,h}$, we have that

$$\|u_{\varphi} - u_{\psi}\|_{1,h} \leq L(h) \|\varphi - \psi\|_{1,h},$$

where

$$L(h) = \frac{2|L|}{\min\{1, M_1\}} (1 + C_3 \alpha^{-1} (2d + c_d))^{1/2} \|\nabla u_\psi - \nabla \Pi_h u\|_{L^\infty(\Omega)}.$$

Now, using the fact that $u_\psi \in \mathcal{J}$ we can bound the last term on the right-hand side in an identical manner as in the proof of (4.2) to get

$$L(h) \leq \frac{2|L|}{\min\{1, M_1\}} (1 + C_3 \alpha^{-1} (2d + c_d))^{1/2} C_* C_3 (C_4)^{d/2} h^{r-d/2} \|u\|_{H^{r+1}(\Omega)},$$

$$d/2 < r \leq \min(m, p).$$

Since $r > d/2$, there exists a positive constant $h_1 \in (0, 1]$ such that $L = \sup_{h \in (0, h_1]} L(h) < 1$. Thus, for $h \in (0, \min(h_0, h_1)]$, the mapping \mathcal{N} is a contraction in the norm $\|\cdot\|_{1,h}$ of the closed set \mathcal{J} . By Banach's fixed point theorem, \mathcal{N} has a unique fixed point u_{DG} in \mathcal{J} ; in particular, by the definition of the set \mathcal{J} , the finite element approximation u_{DG} of u satisfies the bound

$$\|u_{\text{DG}} - \Pi_h u\|_{1,h} \leq C_* h^r \|u\|_{H^{r+1}(\Omega)}, \quad d/2 < r \leq \min(m, p), \quad (4.20)$$

where C_* is defined by (4.18); furthermore $\nabla u_{\text{DG}} \in \mathcal{Z}_\delta$, for all $h \in (0, \min(h_0, h_1)]$.

Let us write $a \lesssim b$ to express the fact that, for real numbers a and b , there exists a positive constant C , depending on the analytical solution u but *independent* of the discretization parameter h , such that $a \leq Cb$ for all h in a closed subinterval of $[0, 1]$ containing 0. We shall write $a \approx b$ if, and only if, $a \lesssim b$ and $b \lesssim a$. Since

$$\|u - \Pi_h u\|_{1,h} \lesssim h^r \|u\|_{H^{r+1}(\Omega)}, \quad d/2 < r \leq \min(m, p), \quad (4.21)$$

we deduce from (4.20) and (4.21) via the triangle inequality that, for all $h \in (0, \min(h_0, h_1)]$,

$$\|u - u_{\text{DG}}\|_{1,h} \lesssim h^r \|u\|_{H^{r+1}(\Omega)}, \quad d/2 < r \leq \min(m, p),$$

which is the required optimal bound on the error in the discontinuous Galerkin finite element method.

5 The hyperbolic problem

Now consider the hyperbolic problem

$$\partial_t^2 u_i - \sum_{\alpha=1}^d \partial_{x_\alpha} (S_{i\alpha}(\nabla u)) = f_i(t, x), \quad i = 1, \dots, d, \quad t \in (0, T], \quad x \in \Omega,$$

subject to the initial condition $u(0, x) = u_0(x)$, $\partial_t u(0, x) = u_1(x)$, $x \in \Omega$, where $u_0 \in H^{m+1}(\Omega)^d$ and $u_1 \in H^m(\Omega)^d$ and analogous boundary conditions as in the case of the static problem considered earlier; that is,

$$u(t, x) = g_D(t, x) \text{ on } (0, T] \times \Gamma_D \quad \text{and} \quad S(\nabla u(t, x))\nu = g_N(t, x) \text{ on } (0, T] \times \Gamma_N. \quad (5.1)$$

We refer to the papers of Dafermos and Hrusa [7] and Chen and von Wahl [5] for theoretical results concerning the existence of a unique local solution to (5.1) in the special case of homogeneous Dirichlet boundary condition on Γ .

We shall suppose throughout that

$$u \in C^2([0, T]; H^{m+1}(\Omega)^d), \quad m > (d/2) + 1.$$

As in the elliptic case, let \mathcal{M} be a convex open set such that $\nabla u([0, T] \times \overline{\Omega}) \subset \mathcal{M}$, and we define δ , \mathcal{M}_δ , \mathcal{Z}_δ and K_δ similarly as before. For simplicity, when there is no danger of confusion, we shall suppress the x -dependence in our notation and write $u(t)$, $v(t)$, etc., instead of $u(t, x)$, $v(t, x)$, etc.; we shall, on occasion, suppress both the x - and the t -dependence, and write u , v , and so on.

Let us consider, for $t \in [0, T]$ and $p > (d/2) + 1$, the (semidiscrete) discontinuous Galerkin finite element approximation $u_{\text{DG}}(t, \cdot) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ to $u(t, \cdot)$, such that

$$(\ddot{u}_{\text{DG}}, v) + B(u_{\text{DG}}, v) + \int_{\Gamma_{\text{D}}} \sigma \dot{u}_{\text{DG}} \cdot v \, ds + \int_{\Gamma_{\text{int}}} \sigma [\![\dot{u}_{\text{DG}}]\!] \cdot [\![v]\!] \, ds = \ell(v) + \int_{\Gamma_{\text{D}}} \sigma \dot{g}_{\text{DG}} \cdot v \, ds \quad (5.2)$$

for all $v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ and all $t \in (0, T]$, and

$$u_{\text{DG}}(0, x) = u_{\text{DG}}^0(x), \quad \dot{u}_{\text{DG}}(0, x) = u_{\text{DG}}^1(x), \quad x \in \Omega,$$

with u_{DG}^0 and u_{DG}^1 in $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$.

We highlight the presence of the last two terms on the left-hand side and the final term on the right-hand side of (5.2) which did not feature in the definition of our discontinuous Galerkin approximation of the elliptic problem considered in the earlier sections. The inclusion of these terms does not affect the consistency of the method. On the other hand, they play a crucial role in ensuring the validity of energy estimates in sufficiently strong norms.

We denote by $W(t) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ the nonlinear projection of $u(t)$ defined by

$$B(W(t), v) = B(u(t), v) \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \quad 0 \leq t \leq T,$$

and we select u_{DG}^0 and u_{DG}^1 in $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ such that

$$\|u_{\text{DG}}^0 - W(0)\|_{1,h} + \|u_{\text{DG}}^1 - \dot{W}(0)\|_{L^2(\Omega)} \lesssim h^r, \quad (d/2) + 1 < r \leq \min(m, p).$$

The existence, uniqueness, approximation properties and differentiability with respect to t of $W(t)$ are established in Section 6. For the sake of simplicity of presentation, we choose $u_{\text{DG}}^0 = W(0)$ and $u_{\text{DG}}^1 = \dot{W}(0)$ here. By using an argument based on Banach's fixed point theorem, similar to the one presented in the previous section, and stimulated by the ideas in [11], we will show the existence and uniqueness of u_{DG} . We shall also show that u_{DG} converges to the analytical solution u with optimal order as the spatial discretization parameter h converges to 0.

5.1 Convergence analysis

We decompose

$$u - u_{\text{DG}} = (u - W) - (u_{\text{DG}} - W) \equiv \eta - \xi.$$

Then, with our choice of the numerical initial conditions u_{DG}^0 and u_{DG}^1 , we have $\xi(0) = 0$ and $\dot{\xi}(0) = 0$. Hence,

$$\begin{aligned} (\ddot{\xi}, v) + B(u_{\text{DG}}, v) - B(W, v) + \int_{\Gamma_{\text{int}}} \sigma[\dot{\xi}] \cdot [v] \, ds + \int_{\Gamma_{\text{D}}} \sigma \dot{\xi} \cdot v \, ds \\ = (\ddot{\eta}, v) + \int_{\Gamma_{\text{int}}} \sigma[\dot{\eta}] \cdot [v] \, ds + \int_{\Gamma_{\text{D}}} \sigma \dot{\eta} \cdot v \, ds \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \end{aligned}$$

Let us assume for the moment that $\nabla u_{\text{DG}}(t, \cdot) \in \mathcal{Z}_\delta$ for all $t \in [0, T]$. In terms of our earlier notation we have

$$\begin{aligned} (\ddot{\xi}, v) + \int_0^1 \tilde{b}(W(t) + \tau(u_{\text{DG}}(t) - W(t)); u_{\text{DG}} - W, v) \, d\tau + \int_{\Gamma_{\text{int}}} \sigma[\dot{\xi}] \cdot [v] \, ds + \int_{\Gamma_{\text{D}}} \sigma \dot{\xi} \cdot v \, ds \\ = (\ddot{\eta}, v) + \int_{\Gamma_{\text{int}}} \sigma[\dot{\eta}] \cdot [v] \, ds + \int_{\Gamma_{\text{D}}} \sigma \dot{\eta} \cdot v \, ds \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \end{aligned}$$

We consider the subset \mathcal{J} of

$$Y := C([0, T]; H^1(\Omega, \mathcal{T}_h)^d) \cap C^1([0, T]; L^2(\Omega)^d)$$

defined by

$$\begin{aligned} \mathcal{J} &= \{ \psi : [0, T] \mapsto S^p(\Omega, \mathcal{T}_h, \mathbf{F}) : \\ &\quad \|\psi - W\|_Y := \max_{t \in [0, T]} \left(\|\psi(t) - W(t)\|_{1,h} + \|\dot{\psi}(t) - \dot{W}(t)\|_{L^2(\Omega)} \right) \leq C_*(u) h^r \} \end{aligned}$$

where $C_*(u)$ is a positive constant and $(d/2) + 1 < r \leq \min(m, p)$. As in the elliptic case, there exists $h_0 > 0$ such that, for all $h \in (0, h_0]$,

$$\psi \in \mathcal{J} \Rightarrow \nabla \psi(t) \in \mathcal{Z}_\delta \text{ for all } t \in [0, T]. \quad (5.3)$$

In addition, \mathcal{J} is a closed, convex subset of Y . Since

$$\begin{aligned} (\ddot{u}_{\text{DG}} - \ddot{W}, v) + \int_0^1 \tilde{b}(W(t) + \tau(u_{\text{DG}}(t) - W(t)); u_{\text{DG}} - W, v) \, d\tau \\ + \int_{\Gamma_{\text{int}}} \sigma[\dot{u}_{\text{DG}} - \dot{W}] \cdot [v] \, ds + \int_{\Gamma_{\text{D}}} \sigma(\dot{u}_{\text{DG}} - \dot{W}) \cdot v \, ds \\ = (\ddot{\eta}, v) + \int_{\Gamma_{\text{int}}} \sigma[\dot{\eta}] \cdot [v] \, ds + \int_{\Gamma_{\text{D}}} \sigma \dot{\eta} \cdot v \, ds \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \end{aligned}$$

with $u_{\text{DG}}(0) = u_{\text{DG}}^0$, $\dot{u}_{\text{DG}}(0) = u_{\text{DG}}^1$, we are led to the following definition of the fixed point map \mathcal{N} on \mathcal{J} : if $\varphi \in \mathcal{J}$, the image $u_\varphi = \mathcal{N}(\varphi)$ is defined as the solution to the

following linear problem:

$$\begin{aligned}
& (\ddot{u}_\varphi - \ddot{W}, v) + \int_0^1 \tilde{b}(W(t) + \tau(\varphi(t) - W(t)); u_\varphi - W, v) \, d\tau \\
& + \int_{\Gamma_{\text{int}}} \sigma[\dot{u}_\varphi - \dot{W}] \cdot [v] \, ds + \int_{\Gamma_D} \sigma(\dot{u}_\varphi - \dot{W}) \cdot v \, ds \\
& = (\ddot{\eta}, v) + \int_{\Gamma_{\text{int}}} \sigma[\dot{\eta}] \cdot [v] \, ds + \int_{\Gamma_D} \sigma \dot{\eta} \cdot v \, ds \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}),
\end{aligned}$$

with $u_\varphi(0) = u_{\text{DG}}^0$, $\dot{u}_\varphi(0) = u_{\text{DG}}^1$.

For the sake of notational simplicity, we define

$$\xi_\varphi = u_\varphi - W,$$

and rewrite the last identity as follows:

$$\begin{aligned}
& (\ddot{\xi}_\varphi, v) + \int_0^1 \tilde{b}(W(t) + \tau(\varphi(t) - W(t)); \xi_\varphi, v) \, d\tau + \int_{\Gamma_{\text{int}}} \sigma[\dot{\xi}_\varphi] \cdot [v] \, ds + \int_{\Gamma_D} \sigma \dot{\xi}_\varphi \cdot v \, ds \\
& = (\ddot{\eta}, v) + \int_{\Gamma_{\text{int}}} \sigma[\dot{\eta}] \cdot [v] \, ds + \int_{\Gamma_D} \sigma \dot{\eta} \cdot v \, ds \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).
\end{aligned}$$

Next, we take $v = \dot{\xi}_\varphi$ above to deduce that

$$\begin{aligned}
& (\ddot{\xi}_\varphi, \dot{\xi}_\varphi) + \int_0^1 \tilde{b}(W(t) + \tau(\varphi(t) - W(t)); \xi_\varphi, \dot{\xi}_\varphi) \, d\tau + \int_{\Gamma_{\text{int}}} \sigma[\dot{\xi}_\varphi] \cdot [\dot{\xi}_\varphi] \, ds + \int_{\Gamma_D} \sigma |\dot{\xi}_\varphi|^2 \, ds \\
& = (\ddot{\eta}, \dot{\xi}_\varphi) + \int_{\Gamma_{\text{int}}} \sigma[\dot{\eta}] \cdot [\dot{\xi}_\varphi] \, ds + \int_{\Gamma_D} \sigma \dot{\eta} \cdot \dot{\xi}_\varphi \, ds.
\end{aligned}$$

This identity can be rewritten as follows:

$$\begin{aligned}
& \frac{1}{2} \frac{d}{dt} \left[\|\dot{\xi}_\varphi\|_{L^2(\Omega)}^2 + \int_\Omega \int_0^1 \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta} (\nabla W(t) + \tau \nabla(\varphi(t) - W(t))) \frac{\partial \xi_{\varphi,i}}{\partial x_\alpha} \frac{\partial \xi_{\varphi,j}}{\partial x_\beta} \, d\tau \, dx \right. \\
& \quad \left. + \int_{\Gamma_D} \sigma |\xi_\varphi|^2 \, ds + \int_{\Gamma_{\text{int}}} \sigma [\xi_\varphi]^2 \, ds \right] \\
& \quad + \int_{\Gamma_{\text{int}}} \sigma [\dot{\xi}_\varphi]^2 \, ds + \int_{\Gamma_D} \sigma |\dot{\xi}_\varphi|^2 \, ds \\
& = (\ddot{\eta}, \dot{\xi}_\varphi) + \int_{\Gamma_{\text{int}}} \sigma[\dot{\eta}] \cdot [\dot{\xi}_\varphi] \, ds + \int_{\Gamma_D} \sigma \dot{\eta} \cdot \dot{\xi}_\varphi \, ds \tag{5.4} \\
& \quad + \frac{1}{2} \int_\Omega \int_0^1 \sum_{i,\alpha,j,\beta=1}^d \left\{ \frac{d}{dt} [A_{i\alpha j\beta} (\nabla W(t) + \tau \nabla(\varphi(t) - W(t)))] \right\} \frac{\partial \xi_{\varphi,i}}{\partial x_\alpha} \frac{\partial \xi_{\varphi,j}}{\partial x_\beta} \, d\tau \, dx \\
& \quad + \frac{1}{2} \int_{\Gamma_D} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta} (\nabla W(t) + \tau \nabla(\varphi(t) - W(t))) \dot{\xi}_{\varphi,i} \nu_\alpha \frac{\partial \xi_{\varphi,j}}{\partial x_\beta} \, d\tau \, ds \\
& \quad + \frac{1}{2} \int_{\Gamma_{\text{int}}} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d \left\langle A_{i\alpha j\beta} (\nabla W(t) + \tau \nabla(\varphi(t) - W(t))) \nu_\alpha \frac{\partial \xi_{\varphi,j}}{\partial x_\beta} \right\rangle [\dot{\xi}_{\varphi,i}] \, d\tau \, ds.
\end{aligned}$$

On noting that

$$\xi_\varphi(0) = 0 \quad \text{and} \quad \dot{\xi}_\varphi(0) = 0,$$

integrating the above identity in t and multiplying by 2, we easily deduce (see inequality (4.5)) that for $\alpha \geq K_\delta^2 M_1^{-1} C_3 C_d$ and $h \in (0, h_0]$,

$$\begin{aligned} & \|\dot{\xi}_\varphi(t)\|_{L^2(\Omega)}^2 + \frac{1}{2} \min(1, M_1) \|\xi_\varphi(t)\|_{1,h}^2 + \int_{\Gamma_D} \sigma |\xi_\varphi(t)|^2 ds + \int_{\Gamma_{\text{int}}} \sigma \|\xi_\varphi(t)\|^2 ds \\ & \quad + 2 \int_0^t \int_{\Gamma_{\text{int}}} \sigma \|\dot{\xi}_\varphi\|^2 ds d\tau + 2 \int_0^t \int_{\Gamma_D} \sigma |\dot{\xi}_\varphi|^2 ds d\tau \\ & \leq 2 \int_0^t (\ddot{\eta}(s), \dot{\xi}_\varphi(s)) ds + 2 \int_0^t \int_{\Gamma_{\text{int}}} \sigma \|\dot{\eta}\| \cdot \|\dot{\xi}_\varphi\| ds d\tau + 2 \int_0^t \int_{\Gamma_D} \sigma \dot{\eta} \cdot \dot{\xi}_\varphi ds d\tau \\ & \quad + \int_0^t \int_\Omega \int_0^1 \sum_{i,\alpha,j,\beta=1}^d \left\{ \frac{d}{ds} [A_{i\alpha j\beta}(\nabla W(s) + \tau \nabla(\varphi(s) - W(s)))] \right\} \frac{\partial \xi_{\varphi,i}}{\partial x_\alpha} \frac{\partial \xi_{\varphi,j}}{\partial x_\beta} d\tau dx ds \\ & \quad + \int_0^t \int_{\Gamma_D} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta}(\nabla W(s) + \tau \nabla(\varphi(s) - W(s))) \dot{\xi}_{\varphi,i} \nu_\alpha \frac{\partial \xi_{\varphi,j}}{\partial x_\beta} d\tau d\sigma ds \quad (5.5) \\ & \quad + \int_0^t \int_{\Gamma_{\text{int}}} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d \left\langle A_{i\alpha j\beta}(\nabla W(s) + \tau \nabla(\varphi(s) - W(s))) \nu_\alpha \frac{\partial \xi_{\varphi,j}}{\partial x_\beta} dx \right\rangle \|\dot{\xi}_{\varphi,i}\| d\tau d\sigma ds. \end{aligned}$$

Next, we estimate the last three terms on the right-hand side. First, note that for each $s \in [0, T]$ and each $\kappa \in \mathcal{T}_h$, we have

$$\begin{aligned} & \frac{d}{ds} [A_{i\alpha j\beta}(\nabla W(s) + \tau \nabla(\varphi(s) - W(s)))] \\ & = \sum_{k,\gamma=1}^d \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{k\gamma}} ((\nabla W(s) + \tau \nabla(\varphi(s) - W(s))) \partial_{x_\gamma} (\dot{W}_k + \tau(\dot{\varphi}_k - \dot{W}_k))). \end{aligned}$$

Recalling that the values of the function $\nabla W(t) + \tau \nabla(\varphi(t) - W(t))$, $t \in [0, T]$, $\tau \in [0, 1]$, belong to the compact convex subset \mathcal{M}_δ of $\mathbb{R}^{d \times d}$, and that the $A_{i\alpha j\beta}$ are C^1 functions, it follows that

$$\begin{aligned} & \left| \int_0^t \int_\Omega \int_0^1 \sum_{i,\alpha,j,\beta=1}^d \left\{ \frac{d}{ds} [A_{i\alpha j\beta}(\nabla W(s) + \tau \nabla(\varphi(s) - W(s)))] \right\} \frac{\partial \xi_{\varphi,i}}{\partial x_\alpha} \frac{\partial \xi_{\varphi,j}}{\partial x_\beta} d\tau dx ds \right| \\ & \lesssim \int_0^t \left(\|\nabla \dot{W}(s)\|_{L^\infty(\Omega)} + \|\nabla \dot{\varphi}(s) - \nabla \dot{W}(s)\|_{L^\infty(\Omega)} \right) \|\nabla \xi_\varphi(s)\|_{L^2(\Omega)}^2 ds \\ & \lesssim \int_0^t \left(1 + h^{-1-d/2} \|\dot{\varphi}(s) - \dot{W}(s)\|_{L^2(\Omega)} \right) \|\nabla \xi_\varphi(s)\|_{L^2(\Omega)}^2 ds \\ & \lesssim \int_0^t \left(1 + h^{r-1-d/2} \right) \|\nabla \xi_\varphi(s)\|_{L^2(\Omega)}^2 ds \\ & \lesssim \int_0^t \|\xi_\varphi(s)\|_{1,h}^2 ds, \end{aligned}$$

since, by hypothesis, $r > (d/2) + 1$. Here we made use of the fact that $\|\nabla \dot{W}(s)\|_{L^\infty(\Omega)} \lesssim 1$ and $\|\dot{\varphi}(s) - \dot{W}(s)\|_{L^2(\Omega)} \lesssim h^r$, $(d/2) + 1 < r \leq \min(m, p)$, which follow from our assumption that $\varphi \in \mathcal{J}$. Similarly,

$$\begin{aligned} & \left| \int_0^t \int_{\Gamma_D} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta} (\nabla W(s) + \tau \nabla(\varphi(s) - W(s))) \dot{\xi}_{\varphi,i} \nu_\alpha \frac{\partial \xi_{\varphi,j}}{\partial x_\beta} d\tau d\sigma ds \right| \\ & \lesssim \int_0^t \left(\int_{\Gamma_D} \sigma |\dot{\xi}_\varphi(s)|^2 d\sigma \right)^{1/2} \|\xi_\varphi(s)\|_{1,h} ds, \end{aligned}$$

and

$$\begin{aligned} & \left| \int_0^t \int_{\Gamma_{\text{int}}} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d \left\langle A_{i\alpha j\beta} (\nabla W(s) + \tau \nabla(\varphi(s) - W(s))) \nu_\alpha \frac{\partial \xi_{\varphi,j}}{\partial x_\beta} dx \right\rangle [\dot{\xi}_{\varphi,i}] d\tau d\sigma ds \right| \\ & \lesssim \int_0^t \left(\int_{\Gamma_{\text{int}}} \sigma |[\dot{\xi}_\varphi(s)]|^2 d\sigma \right)^{1/2} \|\xi_\varphi(s)\|_{1,h} ds. \end{aligned}$$

Substituting these into (5.5) yields

$$\begin{aligned} & \|\dot{\xi}_\varphi(t)\|_{L^2(\Omega)}^2 + \frac{1}{2} \min(1, M_1) \|\xi_\varphi(t)\|_{1,h}^2 + \int_{\Gamma_D} \sigma |\xi_\varphi(t)|^2 ds + \int_{\Gamma_{\text{int}}} \sigma |[\xi_\varphi(t)]|^2 ds \\ & \quad + 2 \int_0^t \int_{\Gamma_{\text{int}}} \sigma |[\dot{\xi}_\varphi]|^2 ds d\tau + 2 \int_0^t \int_{\Gamma_D} \sigma |\dot{\xi}_\varphi|^2 ds d\tau \tag{5.6} \\ & \lesssim 2 \int_0^t |(\ddot{\eta}(\tau), \dot{\xi}_\varphi(\tau))| d\tau + 2 \int_0^t \int_{\Gamma_{\text{int}}} \sigma |[\dot{\eta}]| |[\dot{\xi}_\varphi]| ds d\tau + 2 \int_0^t \int_{\Gamma_D} \sigma |\dot{\eta}| |\dot{\xi}_\varphi| ds d\tau \\ & \quad + \int_0^t \|\xi_\varphi(\tau)\|_{1,h}^2 d\tau + \int_0^t \left(\int_{\Gamma_D} \sigma |\dot{\xi}_\varphi(\tau)|^2 ds \right)^{1/2} \|\xi_\varphi(\tau)\|_{1,h} d\tau \\ & \quad + \int_0^t \left(\int_{\Gamma_{\text{int}}} \sigma |[\dot{\xi}_\varphi(\tau)]|^2 ds \right)^{1/2} \|\xi_\varphi(\tau)\|_{1,h} d\tau. \end{aligned}$$

Using ε -inequality $ab \leq \frac{\varepsilon}{2} a^2 + \frac{1}{2\varepsilon} b^2$, $a, b \geq 0$, in the first three and last two terms on the right-hand side, followed by an application of Gronwall's lemma, we deduce that

$$\begin{aligned} & \|\dot{\xi}_\varphi(t)\|_{L^2(\Omega)}^2 + \|\xi_\varphi(t)\|_{1,h}^2 + \int_0^t \int_{\Gamma_{\text{int}}} \sigma |[\dot{\xi}_\varphi]|^2 ds d\tau + \int_0^t \int_{\Gamma_D} \sigma |\dot{\xi}_\varphi|^2 ds d\tau \\ & \lesssim \left(\int_0^t \|\ddot{\eta}(\tau)\|_{L^2(\Omega)}^2 d\tau + \int_0^t \int_{\Gamma_{\text{int}}} \sigma |[\dot{\eta}]|^2 ds d\tau + \int_0^t \int_{\Gamma_D} \sigma |\dot{\eta}|^2 ds d\tau \right). \tag{5.7} \end{aligned}$$

Substituting the bounds from Section 6 on the relevant norms of $\dot{\eta}$ and $\ddot{\eta}$ into the right-hand side of (5.7), we deduce that for an appropriate choice of $C_*(u)$, \mathcal{N} maps the set \mathcal{J} into itself.

Remark 3 Since our strategy for proving that \mathcal{N} maps \mathcal{J} into itself was very similar to the one presented for the case of the quasilinear elliptic problem considered earlier, we were more concise here than in the corresponding discussion for the elliptic problem. In particular, unlike our detailed analysis in the case of the elliptic problem where we made a deliberate effort to carefully track the constants in the bounds so as to be able to explicitly specify the value of the constant C_* featuring in the definition of the set \mathcal{J} , here, for the sake of brevity, we had refrained from doing so. As a matter of fact, the corresponding constant C_* can be found in an identical manner as in the case of the elliptic problem. ◀

Next we prove that \mathcal{N} is a contraction of \mathcal{J} in the norm $\|\cdot\|_Y$. For this purpose, consider $u_\varphi = \mathcal{N}(\varphi) \in \mathcal{J}$ and $u_\psi = \mathcal{N}(\psi) \in \mathcal{J}$ defined analogously. That is, $u_\varphi = \mathcal{N}(\varphi)$ is defined as the solution to the following linear problem:

$$\begin{aligned} & (\ddot{u}_\varphi - \ddot{W}, v) + \int_0^1 \tilde{b}(W(t) + \tau(\varphi(t) - W(t)); u_\varphi - W, v) \, d\tau \\ & + \int_{\Gamma_{\text{int}}} \sigma[\dot{u}_\varphi - \dot{W}] \cdot [v] \, ds + \int_{\Gamma_D} \sigma(\dot{u}_\varphi - \dot{W}) \cdot v \, ds \\ & = (\ddot{\eta}, v) + \int_{\Gamma_{\text{int}}} \sigma[\dot{\eta}] \cdot [v] \, ds + \int_{\Gamma_D} \sigma \dot{\eta} \cdot v \, ds \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \end{aligned}$$

with $u_\varphi(0) = u_{\text{DG}}^0$, $\dot{u}_\varphi(0) = u_{\text{DG}}^1$. Similarly, $u_\psi = \mathcal{N}(\psi)$ is defined as the solution to the following linear problem:

$$\begin{aligned} & (\ddot{u}_\psi - \ddot{W}, v) + \int_0^1 \tilde{b}(W(t) + \tau(\psi(t) - W(t)); u_\psi - W, v) \, d\tau \\ & + \int_{\Gamma_{\text{int}}} \sigma[\dot{u}_\psi - \dot{W}] \cdot [v] \, ds + \int_{\Gamma_D} \sigma(\dot{u}_\psi - \dot{W}) \cdot v \, ds \\ & = (\ddot{\eta}, v) + \int_{\Gamma_{\text{int}}} \sigma[\dot{\eta}] \cdot [v] \, ds + \int_{\Gamma_D} \sigma \dot{\eta} \cdot v \, ds \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \end{aligned}$$

with $u_\psi(0) = u_{\text{DG}}^0$, $\dot{u}_\psi(0) = u_{\text{DG}}^1$. Thus, by subtracting,

$$\begin{aligned} & (\ddot{u}_\varphi - \ddot{u}_\psi, v) + \int_{\Gamma_{\text{int}}} \sigma[\dot{u}_\varphi - \dot{u}_\psi] \cdot [v] \, ds + \int_{\Gamma_D} \sigma(\dot{u}_\varphi - \dot{u}_\psi) \cdot v \, ds \\ & + \int_0^1 \tilde{b}(W(t) + \tau(\varphi(t) - W(t)); u_\varphi - W, v) \, d\tau \\ & - \int_0^1 \tilde{b}(W(t) + \tau(\psi(t) - W(t)); u_\psi - W, v) \, d\tau = 0 \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \end{aligned}$$

with $(u_\varphi - u_\psi)(0) = 0$ and $(\dot{u}_\varphi - \dot{u}_\psi)(0) = 0$. Equivalently,

$$\begin{aligned} & (\ddot{u}_\varphi - \ddot{u}_\psi, v) + \int_{\Gamma_{\text{int}}} \sigma[\dot{u}_\varphi - \dot{u}_\psi] \cdot [v] \, ds + \int_{\Gamma_D} \sigma(\dot{u}_\varphi - \dot{u}_\psi) \cdot v \, ds \\ & + \int_0^1 \tilde{b}(W(t) + \tau(\varphi(t) - W(t)); u_\varphi - u_\psi, v) \, d\tau \\ & = \int_0^1 \tilde{b}(W(t) + \tau(\varphi(t) - W(t)); W - u_\psi, v) \, d\tau \\ & \quad - \int_0^1 \tilde{b}(W(t) + \tau(\psi(t) - W(t)); W - u_\psi, v) \, d\tau \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \end{aligned}$$

with $(u_\varphi - u_\psi)(0) = 0$ and $(\dot{u}_\varphi - \dot{u}_\psi)(0) = 0$. On taking $v = \dot{u}_\varphi - \dot{u}_\psi$, we have that

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|\dot{u}_\varphi - \dot{u}_\psi\|_{L^2(\Omega)}^2 + \int_{\Gamma_{\text{int}}} \sigma \|\dot{u}_\varphi - \dot{u}_\psi\|^2 \, ds + \int_{\Gamma_D} \sigma |\dot{u}_\varphi - \dot{u}_\psi|^2 \, ds \\ & + \int_0^1 \tilde{b}(W(t) + \tau(\varphi(t) - W(t)); u_\varphi - u_\psi, \dot{u}_\varphi - \dot{u}_\psi) \, d\tau \\ & = \int_0^1 \tilde{b}(W(t) + \tau(\varphi(t) - W(t)); W - u_\psi, \dot{u}_\varphi - \dot{u}_\psi) \, d\tau \\ & \quad - \int_0^1 \tilde{b}(W(t) + \tau(\psi(t) - W(t)); W - u_\psi, \dot{u}_\varphi - \dot{u}_\psi) \, d\tau. \quad (5.8) \end{aligned}$$

In order to proceed, it is helpful to write out in full the expression

$$\tilde{b}(W(t) + \tau(\varphi(t) - W(t)); u_\varphi - u_\psi, \dot{u}_\varphi - \dot{u}_\psi)$$

appearing on the left-hand side of (5.8). The definition (3.4) of \tilde{b} indicates that the expression

$$\tilde{b}(W(t) + \tau(\varphi(t) - W(t)); u_\varphi - u_\psi, \dot{u}_\varphi - \dot{u}_\psi)$$

consists of five terms. The second and third terms of \tilde{b} are transferred to the right-hand side of (5.8). For the first term of the expression \tilde{b} which has been retained on the left-hand side of (5.8) we proceed as follows. We recall the symmetry

$$A_{i\alpha j\beta} = A_{j\beta i\alpha}, \quad i, \alpha, j, \beta = 1, \dots, d,$$

of the fourth-order tensor $A(t)$, employ the identity

$$(A(t)w, \dot{w}) = \frac{1}{2} \frac{d}{dt} (A(t)w, w) - \frac{1}{2} (\dot{A}(t)w, w), \quad (5.9)$$

where

$$\dot{A}(t) = \frac{d}{dt} A(t),$$

we retain the term corresponding to the first term on the right-hand side of (5.9) on the left-hand side of (5.8), and move the term corresponding to the second term on the right-hand of (5.9) to the right-hand side of (5.8).

We deduce thereby that the following identity holds:

$$\begin{aligned}
& \frac{1}{2} \frac{d}{dt} \left[\|\dot{u}_\varphi - \dot{u}_\psi\|_{L^2(\Omega)}^2 \right. \\
& \quad + \int_{\Gamma_D} \sigma |u_\varphi - u_\psi|^2 ds \\
& \quad + \int_{\Gamma_{\text{int}}} \sigma \llbracket u_\varphi - u_\psi \rrbracket^2 ds \\
& \quad + \int_{\Omega} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta} (\nabla W(t) + \tau \nabla(\varphi(t) - W(t))) \\
& \quad \quad \times \partial_{x_\alpha} (u_\varphi - u_\psi)_i \partial_{x_\beta} (u_\varphi - u_\psi)_j d\tau dx \Big] \\
& \quad + \int_{\Gamma_{\text{int}}} \sigma \llbracket \dot{u}_\varphi - \dot{u}_\psi \rrbracket^2 ds + \int_{\Gamma_D} \sigma |\dot{u}_\varphi - \dot{u}_\psi|^2 ds \\
& = \frac{1}{2} \int_{\Omega} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d \left\{ \frac{d}{dt} [A_{i\alpha j\beta} (\nabla W(t) + \tau \nabla(\varphi(t) - W(t)))] \right\} \\
& \quad \quad \times \partial_{x_\alpha} (u_\varphi - u_\psi)_i \partial_{x_\beta} (u_\varphi - u_\psi)_j d\tau dx \\
& \quad + \frac{1}{2} \int_{\Gamma_D} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta} (\nabla W(t) + \tau \nabla(\varphi(t) - W(t))) \\
& \quad \quad \times (\dot{u}_\varphi - \dot{u}_\psi)_i \nu_\alpha \partial_{x_\beta} (u_\varphi - u_\psi)_j d\tau ds \\
& \quad + \frac{1}{2} \int_{\Gamma_{\text{int}}} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d \langle A_{i\alpha j\beta} (\nabla W(t) + \tau \nabla(\varphi(t) - W(t))) \nu_\alpha \partial_{x_\beta} (u_\varphi - u_\psi)_j \rangle \\
& \quad \quad \times \llbracket (\dot{u}_\varphi - \dot{u}_\psi)_i \rrbracket d\tau ds \\
& \quad + \int_0^1 \tilde{b}(W(t) + \tau(\varphi(t) - W(t)); W - u_\psi, \dot{u}_\varphi - \dot{u}_\psi) d\tau \\
& \quad - \int_0^1 \tilde{b}(W(t) + \tau(\psi(t) - W(t)); W - u_\psi, \dot{u}_\varphi - \dot{u}_\psi) d\tau.
\end{aligned}$$

Next, we integrate both sides of the last identity with respect to the variable t , note that

$$(u_\varphi - u_\psi)(0) = 0$$

and

$$(\dot{u}_\varphi - \dot{u}_\psi)(0) = 0,$$

and we multiply the resulting identity by 2.

Hence we deduce that

$$\begin{aligned}
& \|\dot{u}_\varphi(t) - \dot{u}_\psi(t)\|_{L^2(\Omega)}^2 + \int_{\Gamma_D} \sigma |u_\varphi(t) - u_\psi(t)|^2 ds + \int_{\Gamma_{\text{int}}} \sigma \|\llbracket u_\varphi(t) - u_\psi(t) \rrbracket\|^2 ds \\
& + \int_{\Omega} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta} (\nabla W(t) + \tau \nabla(\varphi(t) - W(t))) \\
& \quad \times \partial_{x_\alpha}(u_\varphi - u_\psi)_i \partial_{x_\beta}(u_\varphi - u_\psi)_j d\tau dx \\
& + 2 \int_0^t \int_{\Gamma_{\text{int}}} \sigma \|\dot{u}_\varphi - \dot{u}_\psi\|^2 ds d\tau + 2 \int_0^t \int_{\Gamma_D} \sigma |\dot{u}_\varphi - \dot{u}_\psi|^2 ds d\tau \\
& = \int_0^t \int_{\Omega} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d \left\{ \frac{d}{dt} [A_{i\alpha j\beta} (\nabla W(s) + \tau \nabla(\varphi(s) - W(s)))] \right\} \\
& \quad \times \partial_{x_\alpha}(u_\varphi - u_\psi)_i \partial_{x_\beta}(u_\varphi - u_\psi)_j d\tau dx ds \\
& + \int_0^t \int_{\Gamma_D} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta} (\nabla W(s) + \tau \nabla(\varphi(s) - W(s))) \\
& \quad \times (\dot{u}_\varphi - \dot{u}_\psi)_i \nu_\alpha \partial_{x_\beta}(u_\varphi - u_\psi)_j d\tau d\sigma ds \\
& + \int_0^t \int_{\Gamma_{\text{int}}} \int_0^1 \sum_{i,\alpha,j,\beta=1}^d \langle A_{i\alpha j\beta} (\nabla W(s) + \tau \nabla(\varphi(s) - W(s))) \nu_\alpha \partial_{x_\beta}(u_\varphi - u_\psi)_j \rangle \\
& \quad \times \|\llbracket (\dot{u}_\varphi - \dot{u}_\psi)_i \rrbracket\| d\tau d\sigma ds \\
& + \int_0^t \int_0^1 \int_{\Omega} \sum_{i,\alpha,j,\beta=1}^d [A_{i\alpha j\beta} (\nabla W(s) + \tau \nabla(\varphi(s) - W(s))) \\
& \quad - A_{i\alpha j\beta} (\nabla W(s) + \tau \nabla(\psi(s) - W(s)))] (\partial_{x_\alpha}(\dot{u}_\varphi - \dot{u}_\psi)_i) (\partial_{x_\beta}(W - u_\psi)_j) dx d\tau ds \\
& - \int_0^t \int_0^1 \int_{\Gamma_D} \sum_{i,\alpha,j,\beta=1}^d [A_{i\alpha j\beta} (\nabla W(s) + \tau \nabla(\varphi(s) - W(s))) \\
& \quad - A_{i\alpha j\beta} (\nabla W(s) + \tau \nabla(\psi(s) - W(s)))] (\dot{u}_\varphi - \dot{u}_\psi)_i \nu_\alpha (\partial_{x_\beta}(W - u_\psi)_j) d\sigma d\tau ds \\
& - \int_0^t \int_0^1 \int_{\Gamma_{\text{int}}} \sum_{i,\alpha,j,\beta=1}^d \left\langle [A_{i\alpha j\beta} (\nabla W(s) + t \nabla(\varphi(s) - W(s))) \right. \\
& \quad \left. - A_{i\alpha j\beta} (\nabla W(s) + t \nabla(\psi(s) - W(s)))] \nu_\alpha \partial_{x_\beta}((W - u_\psi)_j) \right\rangle \|\llbracket (\dot{u}_\varphi - \dot{u}_\psi)_i \rrbracket\| d\sigma d\tau ds \\
& \equiv T_1 + \dots + T_6.
\end{aligned}$$

Now, we proceed to bound each of the terms T_1, \dots, T_6 appearing on the right-hand side of the last identity. We begin by noting that

$$T_1 \lesssim \int_0^t \|u_\varphi - u_\psi\|_{1,h}^2 ds.$$

Further,

$$T_2 \lesssim \int_0^t \|\sigma^{1/2}(\dot{u}_\varphi - \dot{u}_\psi)\|_{L^2(\Gamma_D)} \|u_\varphi - u_\psi\|_{1,h} \, ds,$$

and

$$T_3 \lesssim \int_0^t \|\sigma^{1/2}[\![\dot{u}_\varphi - \dot{u}_\psi]\!]\|_{L^2(\Gamma_{\text{int}})} \|u_\varphi - u_\psi\|_{1,h} \, ds.$$

In addition, using that

$$\max_{t \in [0, T]} \|\nabla u_\varphi(t) - \nabla W(t)\|_{L^2(\Omega)} \leq C_*(u)h^r \text{ and } \max_{t \in [0, T]} \|\nabla u_\psi(t) - \nabla W(t)\|_{L^2(\Omega)} \leq C_*(u)h^r,$$

with $(d/2) + 1 < r \leq \min(m, p)$, and recalling our earlier notation

$$|L| = \left(\sum_{i, \alpha, j, \beta=1}^d L_{i\alpha j\beta}^2 \right)^{1/2},$$

we have, using the first of the three inverse inequalities stated in (2.1), together with our hypothesis that the family $\{\mathcal{T}_h\}_{h>0}$ is quasiuniform,

$$\begin{aligned} T_4 &\leq |L| \int_0^t \|\nabla u_\psi - \nabla W\|_{L^\infty(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\nabla \dot{u}_\varphi - \nabla \dot{u}_\psi\|_{L^2(\Omega)} \, ds \\ &\lesssim |L| \int_0^t h^{-d/2-1} \|\nabla u_\psi - \nabla W\|_{L^2(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\dot{u}_\varphi - \dot{u}_\psi\|_{L^2(\Omega)} \, ds \\ &\lesssim h^{r-d/2-1} \int_0^t \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\dot{u}_\varphi - \dot{u}_\psi\|_{L^2(\Omega)} \, ds. \end{aligned}$$

For T_5 , we have, completely analogously,

$$\begin{aligned} T_5 &\lesssim |L| \int_0^t h^{-d/2} \|\nabla u_\varphi - \nabla W\|_{L^2(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}(\dot{u}_\varphi - \dot{u}_\psi)\|_{L^2(\Gamma_D)} \, ds \\ &\lesssim h^{r-d/2-1} \int_0^t \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}(\dot{u}_\varphi - \dot{u}_\psi)\|_{L^2(\Gamma_D)} \, ds. \end{aligned}$$

For T_6 , an identical argument gives that

$$\begin{aligned} T_6 &\lesssim |L| \int_0^t h^{-d/2} \|\nabla u_\varphi - \nabla W\|_{L^2(\Omega)} \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}[\![\dot{u}_\varphi - \dot{u}_\psi]\!]\|_{L^2(\Gamma_{\text{int}})} \, ds \\ &\lesssim h^{r-d/2-1} \int_0^t \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}[\![\dot{u}_\varphi - \dot{u}_\psi]\!]\|_{L^2(\Gamma_{\text{int}})} \, ds. \end{aligned}$$

Hence,

$$\begin{aligned}
& \|\dot{u}_\varphi(t) - \dot{u}_\psi(t)\|_{L^2(\Omega)}^2 + \|\sigma^{1/2}(u_\varphi(t) - u_\psi(t))\|_{L^2(\Gamma_D)}^2 + \|\sigma^{1/2}[\![u_\varphi(t) - u_\psi(t)]\!]\|_{L^2(\Gamma_{\text{int}})}^2 \\
& + \int_0^1 \int_\Omega \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta}(\nabla W(t) + \tau \nabla(\varphi(t) - W(t))) \partial_{x_\alpha}(u_\varphi - u_\psi)_i \partial_{x_\beta}(u_\varphi - u_\psi)_j \, d\tau \, dx \\
& + 2 \int_0^t \|\sigma^{1/2}[\![\dot{u}_\varphi - \dot{u}_\psi]\!]\|_{L^2(\Gamma_{\text{int}})}^2 \, ds + 2 \int_0^t \|\sigma^{1/2}(\dot{u}_\varphi - \dot{u}_\psi)\|_{L^2(\Gamma_D)}^2 \, ds \\
& \lesssim \int_0^t \|u_\varphi - u_\psi\|_{1,h}^2 \, ds \\
& + \int_0^t \|\sigma^{1/2}(\dot{u}_\varphi - \dot{u}_\psi)\|_{L^2(\Gamma_D)} \|u_\varphi - u_\psi\|_{1,h} \, ds \\
& + \int_0^t \|\sigma^{1/2}[\![\dot{u}_\varphi - \dot{u}_\psi]\!]\|_{L^2(\Gamma_{\text{int}})} \|u_\varphi - u_\psi\|_{1,h} \, ds \\
& + h^{r-d/2-1} \int_0^t \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\dot{u}_\varphi - \dot{u}_\psi\|_{L^2(\Omega)} \, ds \\
& + h^{r-d/2-1} \int_0^t \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}(\dot{u}_\varphi - \dot{u}_\psi)\|_{L^2(\Gamma_D)} \, ds \\
& + h^{r-d/2-1} \int_0^t \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}[\![\dot{u}_\varphi - \dot{u}_\psi]\!]\|_{L^2(\Gamma_{\text{int}})} \, ds.
\end{aligned}$$

Assuming that $\alpha \geq K_\delta^2 M_1^{-1} C_3 C_d$ and $h \in (0, h_0]$ we deduce (cf. (4.5)) that

$$\begin{aligned}
& \|\dot{u}_\varphi(t) - \dot{u}_\psi(t)\|_{L^2(\Omega)}^2 + \|u_\varphi(t) - u_\psi(t)\|_{1,h}^2 \\
& + \int_0^t \|\sigma^{1/2}[\![\dot{u}_\varphi - \dot{u}_\psi]\!]\|_{L^2(\Gamma_{\text{int}})}^2 \, ds + \int_0^t \|\sigma^{1/2}(\dot{u}_\varphi - \dot{u}_\psi)\|_{L^2(\Gamma_D)}^2 \, ds \\
& \lesssim \left[\int_0^t \|u_\varphi - u_\psi\|_{1,h}^2 \, ds \right. \\
& + \int_0^t \|\sigma^{1/2}(\dot{u}_\varphi - \dot{u}_\psi)\|_{L^2(\Gamma_D)} \|u_\varphi - u_\psi\|_{1,h} \, ds \\
& + \int_0^t \|\sigma^{1/2}[\![\dot{u}_\varphi - \dot{u}_\psi]\!]\|_{L^2(\Gamma_{\text{int}})} \|u_\varphi - u_\psi\|_{1,h} \, ds \\
& + h^{r-d/2-1} \int_0^t \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\dot{u}_\varphi - \dot{u}_\psi\|_{L^2(\Omega)} \, ds \\
& + h^{r-d/2-1} \int_0^t \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}(\dot{u}_\varphi - \dot{u}_\psi)\|_{L^2(\Gamma_D)} \, ds \\
& \left. + h^{r-d/2-1} \int_0^t \|\nabla \varphi - \nabla \psi\|_{L^2(\Omega)} \|\sigma^{1/2}[\![\dot{u}_\varphi - \dot{u}_\psi]\!]\|_{L^2(\Gamma_{\text{int}})} \, ds \right].
\end{aligned}$$

Thus, by applying the ε -inequality $ab \leq \frac{\varepsilon}{2}a^2 + \frac{1}{2\varepsilon}b^2$, with ε sufficiently small, to the second, third, fourth, fifth and sixth term on the right-hand side, we deduce from Gronwall's

lemma that

$$\begin{aligned}
& \|\dot{u}_\varphi(t) - \dot{u}_\psi(t)\|_{L^2(\Omega)}^2 + \|u_\varphi(t) - u_\psi(t)\|_{1,h}^2 \\
& + \int_0^t \|\sigma^{1/2}[\dot{u}_\varphi - \dot{u}_\psi]\|_{L^2(\Gamma_{\text{int}})}^2 ds + \int_0^t \|\sigma^{1/2}(\dot{u}_\varphi - \dot{u}_\psi)\|_{L^2(\Gamma_D)}^2 ds \\
& \lesssim h^{2(r-d/2-1)} \int_0^t \|\nabla\varphi - \nabla\psi\|_{L^2(\Omega)}^2 ds \lesssim h^{2(r-d/2-1)} \max_{t \in [0,T]} \|\nabla\varphi(t) - \nabla\psi(t)\|_{L^2(\Omega)}^2,
\end{aligned}$$

which, in turn, implies that, for h sufficiently small, \mathcal{N} is a contraction of \mathcal{J} into itself in the norm $\|\cdot\|_Y$. Therefore, by Banach's fixed point theorem for h sufficiently small \mathcal{N} has a unique fixed point, $u_{\text{DG}} \in \mathcal{J}$, the semidiscrete discontinuous Galerkin finite element approximation to u defined by (5.2). In other words, for h sufficiently small,

$$\begin{aligned}
\max_{t \in [0,T]} \left(\|\dot{u}_{\text{DG}}(t) - \dot{W}(t)\|_{L^2(\Omega)} + \|u_{\text{DG}}(t) - W(t)\|_{1,h} \right) & \leq C_*(u)h^r, \\
(d/2) + 1 & < r \leq \min(m, p).
\end{aligned}$$

To complete the convergence analysis, we need to derive bounds on $u - W$, $\dot{u} - \dot{W}$ and $\ddot{u} - \ddot{W}$ in the $\|\cdot\|_{1,h}$ norm. These are given in the next section, where it is shown in particular that, for h sufficiently small,

$$\max_{t \in [0,T]} \left(\|\dot{u}(t) - \dot{W}(t)\|_{L^2(\Omega)} + \|u(t) - W(t)\|_{1,h} \right) \lesssim h^r, \quad (d/2) + 1 < r \leq \min(m, p).$$

Combining the last two bounds we then deduce, for h sufficiently small, that

$$\max_{t \in [0,T]} \left(\|\dot{u}(t) - \dot{u}_{\text{DG}}(t)\|_{L^2(\Omega)} + \|u(t) - u_{\text{DG}}(t)\|_{1,h} \right) \lesssim h^r, \quad (d/2) + 1 < r \leq \min(m, p),$$

which is the desired optimal convergence estimate.

6 Bounds on the nonlinear projection error

The purpose of this section is to derive the required bounds on the error between a function u and its nonlinear elliptic projection W .

6.1 Bounds on $\eta = u - W$

The main idea for bounding $\|u - W\|_{1,h}$ is very similar to the derivation of the error bound presented in Section 4 for the quasilinear elliptic problem. The starting point of the analysis in Section 4 was the Galerkin orthogonality property

$$B(u_{\text{DG}}, v) = B(u, v) \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}),$$

satisfied by $u_{\text{DG}} \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ which gave

$$B(u_{\text{DG}}, v) - B(\Pi_h u, v) = B(u, v) - B(\Pi_h u, v) \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).$$

Hence, we deduced the existence and uniqueness of $u_{\text{DG}} \in \mathcal{J} \subset S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ and that, for sufficiently small $h \in (0, 1]$,

$$\|u_{\text{DG}} - \Pi_h u\|_{1,h} \leq C_* h^r \|u\|_{H^{r+1}(\Omega)}, \quad d/2 < r \leq \min(m, p).$$

Since, for each $t \in [0, T]$, $W(t) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ is defined by

$$B(W(t), v) = B(u(t), v) \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}),$$

which gives

$$B(W(t), v) - B(\Pi_h u(t), v) = B(u(t), v) - B(\Pi_h u(t), v) \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \quad (6.1)$$

an identical argument yields the existence and uniqueness of $W(t) \in \mathcal{J} \subset S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ and

$$\|W(t) - \Pi_h u(t)\|_{1,h} \leq C_* h^r \|u(t)\|_{H^{r+1}(\Omega)}, \quad d/2 < r \leq \min(m, p), \quad t \in [0, T].$$

Since, by the approximation properties of Π_h , in the $\|\cdot\|_{1,h}$ norm

$$\|u(t) - \Pi_h u(t)\|_{1,h} \lesssim h^r \|u(t)\|_{H^{r+1}(\Omega)}, \quad d/2 < r \leq \min(m, p), \quad t \in [0, T],$$

it follows from the triangle inequality that

$$\|W(t) - u(t)\|_{1,h} \lesssim h^r \|u(t)\|_{H^{r+1}(\Omega)}, \quad d/2 < r \leq \min(m, p), \quad t \in [0, T].$$

6.2 Bounds on $\dot{\eta} = \dot{u} - \dot{W}$

We begin by establishing the differentiability of the mapping $t \mapsto W(t)$. Suppose that $U \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ and $t \in [0, T]$. The mapping $V \mapsto B(U, V) - B(u(t), V)$ is a bounded linear functional on $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$; hence, by the Riesz representation theorem, there exists a unique (Riesz representer) $\mathcal{B}(t, U) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ such that

$$(\mathcal{B}(t, U), V) = B(U, V) - B(u(t), V).$$

This defines the (nonlinear) mapping

$$\mathcal{B} : (t, U) \in [0, T] \times S^p(\Omega, \mathcal{T}_h, \mathbf{F}) \mapsto \mathcal{B}(t, U) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).$$

It follows from the results of Section 4.2 that the derivative of $(t, U) \mapsto \mathcal{B}(t, U)$ with respect to U exists and is invertible for any $t \in [0, T]$. Note, furthermore, that $\mathcal{B}(t, W(t)) = 0$. Since $t \mapsto u(t)$ is differentiable, it follows that $(t, U) \mapsto \mathcal{B}(t, U)$ is differentiable in a neighbourhood of $(t_0, W(t_0))$ for any $t_0 \in (0, T)$. We then deduce from the implicit function theorem that $t \mapsto W(t)$ is differentiable in $(0, T)$.

Next, we bound $\|\dot{W}(t) - \dot{u}(t)\|_{1,h}$ and $\|\dot{W}(t) - \dot{u}(t)\|_{L^2(\Omega)}$. We begin by noting that, according to the definition of $W(t)$,

$$\begin{aligned} & \int_0^1 \tilde{b}(\Pi_h u(t) + \tau(W(t) - \Pi_h u(t)); W(t) - \Pi_h u(t), v) \, d\tau \\ &= \int_0^1 \tilde{b}(\Pi_h u(t) + \tau(u(t) - \Pi_h u(t)); u(t) - \Pi_h u(t), v) \, d\tau \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \end{aligned}$$

After differentiation with respect to t , we obtain

$$\begin{aligned}
& \int_0^1 \tilde{b}(\Pi_h u(t) + \tau(W(t) - \Pi_h u(t)); \dot{W}(t) - \Pi_h \dot{u}(t), v) \, d\tau \\
& + \int_0^1 \int_{\Omega} \sum_{i,\alpha,j,\beta,k,\gamma=1}^d \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{k\gamma}} (\nabla \Pi_h u(t) + \tau \nabla (W(t) - \Pi_h u(t))) \\
& \quad \times \partial_{x_\gamma} (\Pi_h \dot{u}(t) + \tau (\dot{W}(t) - \Pi_h \dot{u}(t)))_k \partial_{x_\alpha} v_i \partial_{x_\beta} (W - \Pi_h u)_j \, dx \, d\tau \\
& - \int_0^1 \int_{\Gamma_D} \sum_{i,\alpha,j,\beta,k,\gamma=1}^d \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{k\gamma}} (\nabla \Pi_h u(t) + \tau \nabla (W(t) - \Pi_h u(t))) \\
& \quad \times \partial_{x_\gamma} (\Pi_h \dot{u}(t) + \tau (\dot{W}(t) - \Pi_h \dot{u}(t)))_k v_i \nu_\alpha \partial_{x_\beta} (W - \Pi_h u)_j \, ds \, d\tau \\
& - \int_0^1 \int_{\Gamma_{\text{int}}} \sum_{i,\alpha,j,\beta,k,\gamma=1}^d \left\langle \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{k\gamma}} (\nabla \Pi_h u(t) + \tau \nabla (W(t) - \Pi_h u(t))) \right. \\
& \quad \left. \times \partial_{x_\gamma} (\Pi_h \dot{u}(t) + \tau (\dot{W}(t) - \Pi_h \dot{u}(t)))_k \nu_\alpha \partial_{x_\beta} (W - \Pi_h u)_j \right\rangle \llbracket v_i \rrbracket \, ds \, d\tau \\
& = \int_0^1 \tilde{b}(\Pi_h u(t) + \tau(u(t) - \Pi_h u(t)); \dot{u}(t) - \Pi_h \dot{u}(t), v) \, d\tau \\
& + \int_0^1 \int_{\Omega} \sum_{i,\alpha,j,\beta,k,\gamma=1}^d \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{k\gamma}} (\nabla \Pi_h u(t) + \tau \nabla (u(t) - \Pi_h u(t))) \\
& \quad \times \partial_{x_\gamma} (\Pi_h \dot{u}(t) + \tau (\dot{u}(t) - \Pi_h \dot{u}(t)))_k \partial_{x_\alpha} v_i \partial_{x_\beta} (u - \Pi_h u)_j \, dx \, d\tau \\
& - \int_0^1 \int_{\Gamma_D} \sum_{i,\alpha,j,\beta,k,\gamma=1}^d \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{k\gamma}} (\nabla \Pi_h u(t) + \tau \nabla (u(t) - \Pi_h u(t))) \\
& \quad \times \partial_{x_\gamma} (\Pi_h \dot{u}(t) + \tau (\dot{u}(t) - \Pi_h \dot{u}(t)))_k v_i \nu_\alpha \partial_{x_\beta} (u - \Pi_h u)_j \, ds \, d\tau \\
& - \int_0^1 \int_{\Gamma_{\text{int}}} \sum_{i,\alpha,j,\beta,k,\gamma=1}^d \left\langle \frac{\partial A_{i\alpha j\beta}}{\partial \eta_{k\gamma}} (\nabla \Pi_h u(t) + \tau \nabla (u(t) - \Pi_h u(t))) \right. \\
& \quad \left. \times \partial_{x_\gamma} (\Pi_h \dot{u}(t) + \tau (\dot{u}(t) - \Pi_h \dot{u}(t)))_k \nu_\alpha \partial_{x_\beta} (u - \Pi_h u)_j \right\rangle \llbracket v_i \rrbracket \, ds \, d\tau.
\end{aligned}$$

Of the four terms appearing on the left-hand side of this inequality, only the first will be retained on the left-hand side. The remaining three terms are moved across to the right-hand side, resulting in seven terms T_1, \dots, T_7 , starting with the existing four terms on the right-hand side. We take $v(t) = W(t) - \Pi_h u(t)$, and estimate each of the terms T_1, \dots, T_7 ; below, $(d/2) + 1 < r \leq \min(m, p)$.

$$T_1 \lesssim h^r \|\dot{u}(t)\|_{H^{r+1}(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h}.$$

Further,

$$\begin{aligned}
T_2 & \lesssim (\|\nabla \Pi_h \dot{u}(t)\|_{L^\infty(\Omega)} + \|\nabla (\dot{u}(t) - \Pi_h \dot{u}(t))\|_{L^\infty(\Omega)}) \\
& \quad \times \|\nabla (u(t) - \Pi_h u(t))\|_{L^2(\Omega)} \|\nabla (\dot{W}(t) - \Pi_h \dot{u}(t))\|_{1,h} \\
& \lesssim h^r \|u(t)\|_{H^{r+1}(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h}.
\end{aligned}$$

$$\begin{aligned}
T_3 &\lesssim (\|\nabla \Pi_h \dot{u}(t)\|_{L^\infty(\Omega)} + \|\nabla(\dot{u}(t) - \Pi_h \dot{u}(t))\|_{L^\infty(\Omega)}) \\
&\quad \times \|\sigma^{-1/2} \nabla(u(t) - \Pi_h u(t))\|_{L^2(\Gamma_D)} \|\sigma^{1/2} (\dot{W}(t) - \Pi_h \dot{u}(t))\|_{L^2(\Gamma_D)} \\
&\lesssim h^r \|u(t)\|_{H^{r+1}(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
T_4 &\lesssim (\|\nabla \Pi_h \dot{u}(t)\|_{L^\infty(\Omega)} + \|\nabla(\dot{u}(t) - \Pi_h \dot{u}(t))\|_{L^\infty(\Omega)}) \\
&\quad \times \|\sigma^{-1/2} \langle |\nabla(u(t) - \Pi_h u(t))| \rangle\|_{L^2(\Gamma_{\text{int}})} \|\sigma^{1/2} [\dot{W}(t) - \Pi_h \dot{u}(t)]\|_{L^2(\Gamma_{\text{int}})} \\
&\lesssim h^r \|u(t)\|_{H^{r+1}(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h}.
\end{aligned}$$

Next, we bound T_5 , T_6 and T_7 . For T_5 , using the bound $\|W(t) - \Pi_h u(t)\|_{1,h} \leq C_* h^r \|u(t)\|_{H^{r+1}(\Omega)}$, we deduce that

$$\begin{aligned}
T_5 &\lesssim (\|\nabla \Pi_h \dot{u}(t)\|_{L^\infty(\Omega)} + \|\nabla(\dot{W}(t) - \Pi_h \dot{u}(t))\|_{L^\infty(\Omega)}) \\
&\quad \times \|\nabla(W(t) - \Pi_h u(t))\|_{L^2(\Omega)} \|\nabla(\dot{W}(t) - \Pi_h \dot{u}(t))\|_{1,h} \\
&\lesssim h^r \|u(t)\|_{H^{r+1}(\Omega)} (\|\nabla \Pi_h \dot{u}(t)\|_{L^\infty(\Omega)} + \|\nabla(\dot{W}(t) - \Pi_h \dot{u}(t))\|_{L^\infty(\Omega)}) \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h} \\
&\lesssim h^r \|u(t)\|_{H^{r+1}(\Omega)} (\|\nabla \dot{u}(t)\|_{L^\infty(\Omega)} + \|\nabla(\Pi_h \dot{u}(t) - \dot{u}(t))\|_{L^\infty(\Omega)}) \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h} \\
&\quad + h^{r-d/2} \|u(t)\|_{H^{r+1}(\Omega)} \|\nabla(\dot{W}(t) - \Pi_h \dot{u}(t))\|_{L^2(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h} \\
&\lesssim h^r \|u(t)\|_{H^{r+1}(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h} \\
&\quad + h^{r-d/2} \|u(t)\|_{H^{r+1}(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h}^2.
\end{aligned}$$

Further, by means of an inverse inequality, $\|\sigma^{-1/2} \nabla(W(t) - \Pi_h u(t))\|_{L^2(\Gamma_D)}$ can be bounded in terms of $\|\nabla(W(t) - \Pi_h u(t))\|_{L^2(\Omega)}$ which, in turn, is bounded by $\|W(t) - \Pi_h u(t)\|_{1,h} \leq C_* h^r \|u(t)\|_{H^{r+1}(\Omega)}$; therefore, proceeding as in the case of T_5 , we have

$$\begin{aligned}
T_6 &\lesssim (\|\nabla \Pi_h \dot{u}(t)\|_{L^\infty(\Omega)} + \|\nabla(\dot{W}(t) - \Pi_h \dot{u}(t))\|_{L^\infty(\Omega)}) \\
&\quad \times \|\sigma^{-1/2} \nabla(W(t) - \Pi_h u(t))\|_{L^2(\Gamma_D)} \|\sigma^{1/2} \nabla(\dot{W}(t) - \Pi_h \dot{u}(t))\|_{L^2(\Gamma_D)} \\
&\lesssim h^r \|u(t)\|_{H^{r+1}(\Omega)} (\|\nabla \Pi_h \dot{u}(t)\|_{L^\infty(\Omega)} + \|\nabla(\dot{W}(t) - \Pi_h \dot{u}(t))\|_{L^\infty(\Omega)}) \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h} \\
&\lesssim h^r \|u(t)\|_{H^{r+1}(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h} \\
&\quad + h^{r-d/2} \|u(t)\|_{H^{r+1}(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h}^2.
\end{aligned}$$

In the same way,

$$T_7 \lesssim h^r \|u(t)\|_{H^{r+1}(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h} + h^{r-d/2} \|u(t)\|_{H^{r+1}(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h}^2.$$

Now, selecting $\alpha \geq K_\delta^2 M_1^{-1} C_3 C_d$ and $h \in (0, \min(h_0, h_1)]$ as before (see (4.5)), and combining the bounds on the terms T_1, \dots, T_7 ,

$$\begin{aligned}
\frac{1}{2} \min(1, M_1) \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h}^2 &\lesssim h^r (\|u(t)\|_{H^{r+1}(\Omega)} + \|\dot{u}(t)\|_{H^{r+1}(\Omega)}) \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h} \\
&\quad + h^{r-d/2} \|u(t)\|_{H^{r+1}(\Omega)} \|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h}^2.
\end{aligned}$$

Since $r > (d/2) + 1 > d/2$, there exists $h_2 \in (0, \min(h_0, h_1)]$ such that, for all $h \in (0, h_2]$, the coefficient of $\|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h}^2$ on the right-hand side is less than or equal to $\frac{1}{4} \min(1, M_1)$. Therefore, the second term on the right-hand side can be absorbed into the left-hand side. Hence,

$$\|\dot{W}(t) - \Pi_h \dot{u}(t)\|_{1,h} \lesssim h^r (\|u(t)\|_{H^{r+1}(\Omega)} + \|\dot{u}(t)\|_{H^{r+1}(\Omega)}), \quad \begin{cases} (d/2) + 1 < r \leq \min(m, p), \\ h \in (0, h_2]. \end{cases}$$

Since, by the approximation properties of Π_h ,

$$\|\dot{u}(t) - \Pi_h \dot{u}(t)\|_{1,h} \lesssim h^r \|\dot{u}(t)\|_{H^{r+1}(\Omega)}, \quad (d/2) + 1 < r \leq \min(m, p), \quad h \in (0, h_2],$$

by the triangle inequality we have

$$\|\dot{u}(t) - \dot{W}(t)\|_{1,h} \lesssim h^r (\|u(t)\|_{H^{r+1}(\Omega)} + \|\dot{u}(t)\|_{H^{r+1}(\Omega)}), \quad \begin{cases} (d/2) + 1 < r \leq \min(m, p), \\ h \in (0, h_2]. \end{cases}$$

Finally, by the broken Poincaré–Friedrichs inequality (see [3]), we deduce that

$$\|\dot{W}(t) - \dot{u}(t)\|_{L^2(\Omega)} \lesssim h^r (\|u(t)\|_{H^{r+1}(\Omega)} + \|\dot{u}(t)\|_{H^{r+1}(\Omega)}), \quad \begin{cases} (d/2) + 1 < r \leq \min(m, p), \\ h \in (0, h_2]. \end{cases}$$

The last inequality is not of optimal order, but it is sufficiently sharp for our purposes.

6.3 Bounds on $\ddot{\eta} = \ddot{u} - \ddot{W}$

By proceeding in an identical manner as in the previous section we find that $\dot{W}(t)$ is differentiable and we get, for $(d/2) + 1 < r \leq \min(m, p)$, that

$$\|\ddot{W}(t) - \Pi_h \ddot{u}(t)\|_{1,h} \lesssim h^r (\|u(t)\|_{H^{r+1}(\Omega)} + \|\dot{u}(t)\|_{H^{r+1}(\Omega)} + \|\ddot{u}(t)\|_{H^{r+1}(\Omega)}), \quad h \in (0, h_2].$$

Invoking, once again, the approximation properties of Π_h , we deduce from the triangle inequality that, for $(d/2) + 1 < r \leq \min(m, p)$,

$$\|\ddot{W}(t) - \ddot{u}(t)\|_{1,h} \lesssim h^r (\|u(t)\|_{H^{r+1}(\Omega)} + \|\dot{u}(t)\|_{H^{r+1}(\Omega)} + \|\ddot{u}(t)\|_{H^{r+1}(\Omega)}), \quad h \in (0, h_2].$$

Finally, by the broken Poincaré–Friedrichs inequality (see [3]), we have, for $(d/2) + 1 < r \leq \min(m, p)$,

$$\|\ddot{W}(t) - \ddot{u}(t)\|_{L^2(\Omega)} \lesssim h^r (\|u(t)\|_{H^{r+1}(\Omega)} + \|\dot{u}(t)\|_{H^{r+1}(\Omega)} + \|\ddot{u}(t)\|_{H^{r+1}(\Omega)}), \quad h \in (0, h_2].$$

Again, the last inequality is not of optimal order, but it is sufficiently sharp for our purposes.

7 Extension to strongly elliptic systems

When the mapping $\eta \mapsto S(\eta)$ is not strongly monotone (cf. (1.5)), but instead only satisfies the weaker Legendre–Hadamard strong ellipticity condition (1.6) then a few modifications need to be made in the convergence analysis. First of all, we note that the bilinear form \tilde{b} , defined in Section 3 can no longer be expected to be coercive. Indeed, if $x \mapsto (a_{i\alpha j\beta}(x))$ is an arbitrary continuous tensor-valued map satisfying the Legendre–Hadamard condition, that is there exists a constant $M_1 > 0$ such that

$$\sum_{i,\alpha,j,\beta=1}^d a_{i\alpha j\beta}(x) \zeta_i \zeta_j \xi_\alpha \xi_\beta \geq M_1 |\zeta|^2 |\xi|^2 \quad \forall \zeta, \xi \in \mathbb{R}^d, \quad x \in \overline{\Omega},$$

then the most we can expect is that Gårding’s inequality,

$$\sum_{i,\alpha,j,\beta=1}^d \int_{\Omega} a_{i\alpha j\beta}(x) \partial_{x_\alpha} v_i \partial_{x_\beta} v_j \, dx \geq \frac{1}{2} M_1 \|\nabla v\|_{L^2(\Omega)}^2 - c_1 \|v\|_{L^2(\Omega)}^2, \quad (7.1)$$

holds (see Theorem 6.5.1 on p.253 of Morrey’s monograph [12]). Even this weaker inequality is, to the best of our knowledge, known only for $v \in H_0^1(\Omega)^d$. Hence, we shall assume throughout this section that $\Gamma_N = \emptyset$.

Note furthermore, that the constant c_1 in (7.1) depends strongly on the variation of the functions $a_{i\alpha j\beta}$, $i, \alpha, j, \beta = 1, \dots, d$. Thus, to extend the convergence analysis to this case we first need to prove a broken version of (7.1) which holds for

$$a_{i\alpha j\beta}(x) = A_{i\alpha j\beta}(\nabla \varphi(x)), \quad i, \alpha, j, \beta = 1, \dots, d, \quad x \in \overline{\Omega},$$

uniformly for φ ranging through $\mathcal{J} \subset \mathcal{M}_\delta$ (cf. Section 5). Such a result will be given in Section 7.2. Its proof relies on a recovery operator, linking discontinuous piecewise polynomial functions to their continuous “relatives”, which we construct in Lemma 4.

7.1 Recovery operator for discontinuous piecewise polynomial functions

Our first step in the proof of a broken Gårding inequality is the construction of a recovery operator, which connects each discontinuous piecewise polynomial function to a continuous relative. A similar technique was used in [4] to prove broken versions of Korn’s inequalities, the main difference being that the recovery operator used there is not sufficient for our purposes; thus we construct an extended version in Lemma 4 below.

As in [4], we shall prove the result for regular simplicial meshes and reduce the case of meshes consisting of parallelepipeds to simplicial meshes (by refining each d -dimensional parallelepiped in the mesh into d -dimensional simplices) and irregular meshes to regular meshes, by proceeding as follows. If $d = 2$ we prove below that this reduction is always possible for 1-irregular (cf. Proposition 2) meshes without affecting the quality of the mesh too much. The three-dimensional case is significantly more difficult, however.

Section 6 in [4] and Section 7 in [3] provide excellent discussions of the technical details of such a reduction. Here, we shall simply assume that for each $h > 0$ there exists a regular simplicial mesh $\tilde{\mathcal{T}}_h$ such that the closure of each element in \mathcal{T}_h is a union of closures of elements of $\tilde{\mathcal{T}}_h$, and that there exist positive constants θ and C , independent of h , such that the smallest angle between any two edges in $\tilde{\mathcal{T}}_h$ is greater than or equal to θ and $h/\min_{\kappa \in \tilde{\mathcal{T}}_h} h_\kappa \leq C$. We shall call such a family $\{\mathcal{T}_h\}_{h>0}$ *uniformly simplicially reducible*. Proposition 2 shows that quasiuniform families of 1-irregular meshes in two dimensions satisfy this property, while Proposition 3 gives an important example of a class of families of meshes in three dimensions which are uniformly simplicially reducible.

Proposition 2 *Let $d = 2$ and let \mathcal{T}_h be a quasiuniform mesh consisting only of triangles and parallelograms. Assume, further, that \mathcal{T}_h is 1-irregular (i.e., each open edge of any one $\kappa \in \mathcal{T}_h$ contains at most one hanging node, which we assume to be its midpoint), and that the smallest angle of any one $\kappa \in \mathcal{T}_h$ exceeds a fixed constant $\theta > 0$. Then, there exists a regular, simplicial refinement $\tilde{\mathcal{T}}_h$ of \mathcal{T}_h whose smallest angle is at least $c_1\theta$, where c_1 is a universal constant. Moreover, if $h/\min_{\kappa \in \mathcal{T}_h} h_\kappa \leq C$, then $h/\min_{\kappa \in \tilde{\mathcal{T}}_h} h_\kappa \leq c_2$, where c_2 is a constant which depends only on θ and on C .*

Proof First, assume that \mathcal{T}_h is a 1-irregular, simplicial mesh. Let $\kappa \in \mathcal{T}_h$ be an arbitrary

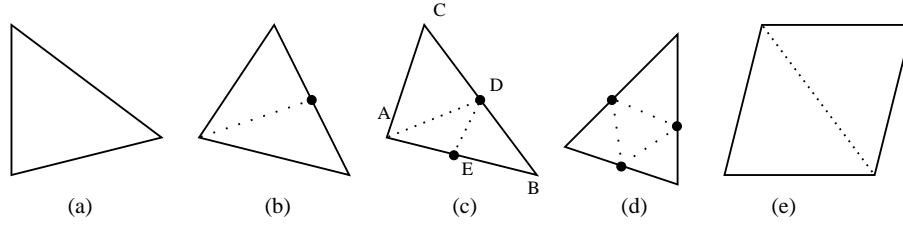


Figure 2: Refining a mesh to remove hanging nodes.

triangular element. Depending on the structure of its hanging nodes, we refine it according to Figure 2 (b)–(d), which yields a set of elements $\{\kappa_1, \dots, \kappa_m\}$ that make up all of κ .

We transform the element, together with its refined submesh to the reference element $\hat{\kappa}$. Since for the reference element there is only a finite number of possible cases, the resulting minimal angle $\hat{\theta}$ and shortest edge length \hat{h} must be constants independent of h and θ . Upon transforming back to κ , quasiuniformity guarantees that all angles in the submesh of κ will be larger than a fixed constant times $\hat{\theta}$, and all edges will be longer than another constant times \hat{h} .

If there are parallelograms present in the mesh, we can bisect each according to Figure 2 (e), to obtain two triangles. The resulting mesh will still be quasiuniform and 1-irregular. To estimate the new minimum angle we can use the same argument as for the bisection of triangles, since a parallelogram is the affine image of the reference square. Now we can proceed as above to remove the hanging nodes. ■

Proposition 3 *Let $d = 3$ and let $\{\mathcal{T}_h\}_{h>0}$ be a family of meshes containing either tetrahedra or parallelepipeds, which are obtained by hierarchical refinement of a given regular mesh \mathcal{T} according to the following rules:*

- Each simplex is refined into eight smaller simplices, by putting exactly one additional vertex on each edge (e.g. at the midpoint of the edge; cf. Figure 3 (a)).
- Each parallelepiped is refined into eight parallelepipeds of the same shape by adding exactly one vertex to each edge (e.g. at the midpoint of the edge), and one vertex to each face (e.g. at the midpoint of the face (cf. Figure 3 (b))).

If the family $\{\mathcal{T}_h\}_{h>0}$ is quasiuniform then it is uniformly simplicially reducible.

Proof Let θ be the smallest angle of any one element in \mathcal{T} and define

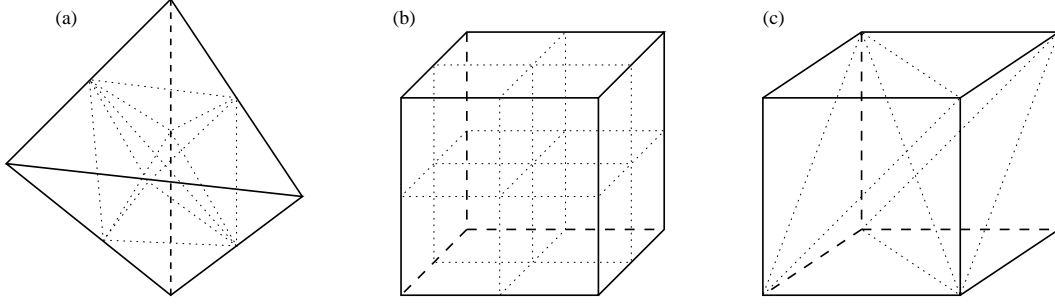


Figure 3: Hierarchical refinement of a tetrahedron into eight tetrahedra (a), of a parallelepiped into eight parallelepipeds (b), and of a parallelepiped into six tetrahedra (c).

$$C = \max_{\kappa \in \mathcal{T}} h_{\kappa} / \min_{\kappa \in \mathcal{T}} h_{\kappa} \quad \text{and} \quad C_h = h / \min_{\kappa \in \mathcal{T}_h} h_{\kappa}.$$

To each element $\kappa \in \mathcal{T}_h$ we can associate a level $\ell_{\kappa} \in \mathbb{N}$ which is the integer denoting the number of refinements performed to obtain κ . Let $\ell_h = \max_{\kappa \in \mathcal{T}_h} \ell_{\kappa}$. Then, we define $\tilde{\mathcal{T}}_h$ by refining \mathcal{T} successively ℓ_h times according to the rules stated in the proposition. Clearly, the angles in $\tilde{\mathcal{T}}_h$ are bounded by θ and $\max_{\kappa \in \tilde{\mathcal{T}}_h} h_{\kappa} / \min_{\kappa \in \tilde{\mathcal{T}}_h} h_{\kappa} = C$. Thus, we can estimate

$$\max_{\kappa \in \tilde{\mathcal{T}}_h} h_{\kappa} \leq C \min_{\kappa \in \tilde{\mathcal{T}}_h} h_{\kappa} \leq C \min_{\kappa \in \mathcal{T}_h} h_{\kappa} \leq CC_h h.$$

If the family $\{\mathcal{T}_h\}_{h>0}$ is quasiuniform, then there exists a constant \hat{C} such that $\sup_{h>0} C_h \leq \hat{C}$, and consequently $\{\mathcal{T}_h\}_{h>0}$ is uniformly simplicially reducible, if the meshes are simplicial. If they contain parallelepipeds, then we can refine each parallelepiped into six tetrahedra, without creating hanging nodes or hanging edges (cf. Figure 3 (c)). ■

Lemma 4 Let $d \in \{2, 3\}$ and let $\{\mathcal{T}_h\}_{h>0}$ be a quasiuniform, uniformly simplicially reducible family of partitions of $\bar{\Omega}$. Then, there exists a constant C_r , which is independent of h , and linear operators $\mathcal{R}: S^p(\Omega, \mathcal{T}_h, \mathbf{F}) \rightarrow W^{1,\infty}(\Omega)^d$ and $\mathcal{R}_0: S^p(\Omega, \mathcal{T}_h, \mathbf{F}) \rightarrow W_0^{1,\infty}(\Omega)^d$ such that for all $u \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ and $k \in \{0, 1\}$,

$$|u - \mathcal{R}u|_{k, \mathcal{T}_h}^2 \leq C_r \int_{\Gamma_{\text{int}}} h^{1-2k} |\llbracket u \rrbracket|^2 ds, \quad (7.2)$$

$$|u - \mathcal{R}_0 u|_{k, \mathcal{T}_h}^2 \leq C_r \left(\int_{\Gamma_{\text{int}}} h^{1-2k} |\llbracket u \rrbracket|^2 ds + \int_{\partial\Omega} h^{1-2k} |u|^2 ds \right). \quad (7.3)$$

Proof Using the assumption that the family of meshes is uniformly simplicially reducible we can assume, without loss of generality that the family $\{\mathcal{T}_h\}_{h>0}$ is a regular, simplicial mesh family. Any $u \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ will exist also on the refined finite element space over the mesh $\tilde{\mathcal{T}}_h$, but will have no additional jumps introduced across the faces which have been added during the refinement. The minimum angle and constant of quasiuniformity are both independent of h . Upon possibly increasing the polynomial degree, we may furthermore assume that the polynomial space on each element κ is \mathcal{P}_κ .

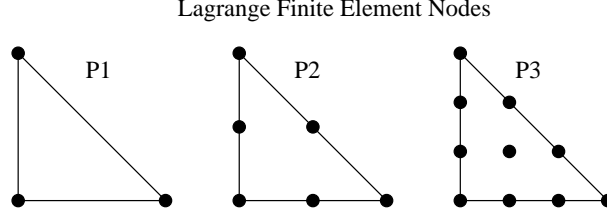


Figure 4: Nodes of several Lagrange finite elements.

For this proof, we need to introduce two additional sets. Let \mathcal{N}_h be the set of vertices in \mathcal{T}_h , and, in three dimensions only, let \mathcal{G}_h be the set of edges.

We associate with each element $\kappa \in \mathcal{T}_h$ its usual nodal basis $\{\phi_{\kappa,i}, i = 1, \dots, N_\kappa\}$. The two-dimensional case is shown in Figure 4. Let N_κ be the number of basis functions for an element with polynomial degree p . If $d = 2$, then $N_\kappa = p(p+1)/2$; if $d = 3$ then $N_\kappa = p(p+1)(p+2)/6$. The basis on the reference element is denoted by $\{\phi_i, i = 1, \dots, N_\kappa\}$. The basis functions are divided into vertex, face and bubble functions for $d = 2$, and vertex, edge, face, and bubble functions for $d = 3$, which are defined as follows. *Vertex functions* are nodal basis functions corresponding to the $d+1$ vertices of an element. In particular, a vertex function vanishes along the closed face opposite the respective vertex. *Edge functions* exist in three dimensions only and if $p \geq 2$. They correspond to those nodes of a simplex which lie on an open 1-dimensional edge in three dimensions. In particular, they vanish on all closed 2-dimensional faces which do not intersect the respective open edge. *Face functions* exist in both two and three dimensions; they are nodal basis functions which correspond to nodes that are contained in an open $(d-1)$ -dimensional face; in particular they vanish on the closure of the union of the remaining $(d-1)$ -dimensional faces. They exist only if $p \geq 2$ in two dimensions and if $p \geq 3$ in three dimensions. Finally, *bubble functions* are nodal basis functions which correspond to nodes in the interior of the element, i.e., they vanish on the entire boundary of κ . Bubble functions are contained in the finite element space only if $p \geq 3$ in two dimensions and if $p \geq 4$ in three dimensions. To avoid having to distinguish between various special cases, without loss of generality we shall assume that $p \geq 4$ throughout this proof. The proof is easily adjusted to cover the cases of $p = 1, 2, 3$ which are excluded here.

If $u \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ then, for $x \in \bar{\kappa}$,

$$u(x) = \sum_{i=1}^{N_\kappa} U_{\kappa,i} \phi_i(x).$$

Let $e \in \mathcal{E}_h$; then we can associate a nodal basis $\{\psi_{e,i}, i = 1, \dots, N_e\}$ with e and correspondingly a nodal basis $\{\psi_i, i = 1, \dots, N_e\}$ on the reference face \hat{e} which is the reference simplex in dimension $d-1$. Clearly, $N_e = p+1$ if $d = 2$ and $N_e = p(p+1)/2$ if $d = 3$. If κ_e and κ'_e are

the neighbours of e in \mathcal{T}_h then the corresponding nodal values are denoted by $U_{e,i}$ and $U'_{e,i}$ and we set $\llbracket U \rrbracket_{e,i} = U_{e,i} - U'_{e,i}$. Thus, the jumps are given by

$$\llbracket u \rrbracket(s) = \sum_{i=1}^{N_e} \llbracket U \rrbracket_{e,i} \psi_{e,i}(s).$$

Due to norm-equivalence in a finite-dimensional linear space, there exist constants $c_0, c_1 > 0$ such that

$$c_0 \sum_{i=1}^{N_e} |\llbracket U \rrbracket_{e,i}|^2 \leq \int_{\hat{e}} \left| \sum_{i=1}^{N_e} \llbracket U \rrbracket_{e,i} \psi_i(\hat{s}) \right|^2 d\hat{s} \leq c_1 \sum_{i=1}^{N_e} |\llbracket U \rrbracket_{e,i}|^2,$$

and, by transforming the integral from \hat{e} to e ,

$$c_0 \sum_{i=1}^{N_e} |\llbracket U \rrbracket_{e,i}|^2 \leq ((d-1)|e|)^{-1} \int_e |\llbracket u \rrbracket|^2(s) ds \leq c_1 \sum_{i=1}^{N_e} |\llbracket U \rrbracket_{e,i}|^2,$$

where $|e|$ is the $(d-1)$ -dimensional surface area of e . Using the quasiuniformity of the mesh, we may write

$$\sum_{i=1}^{N_e} |\llbracket U \rrbracket_{e,i}|^2 \approx \int_e h^{1-d} |\llbracket u \rrbracket|^2(s) ds. \quad (7.4)$$

Let $z \in \mathcal{N}_h$ be a vertex of the mesh. We define $T(z) = \{\kappa \in \mathcal{T}_h : z \in \bar{\kappa}\}$. For each $\kappa \in T(z)$, let $U_{z,\kappa}$ be the nodal value of u in $\bar{\kappa}$ at the node z . If $d = 3$, for any edge $\gamma \in \mathcal{G}_h$, let $T(\gamma) = \{\kappa \in \mathcal{T}_h : \gamma \subset \bar{\kappa}\}$. The nodal value of the i th node on γ inside the element κ will be denoted by $U_{\gamma,\kappa,i}$.

We now turn to the construction of the recovery function. Let v be the continuous finite element function, which has the nodal values $V_{\kappa,i}$ on each element $\kappa \in \mathcal{T}_h$, $V_{e,i}$ on each face $e \in \mathcal{E}_h$, $V_{\gamma,i}$ on each edge $\gamma \in \mathcal{G}_h$ (if $d = 3$) and V_z at each vertex $z \in \mathcal{N}_h$. For each $\kappa \in \mathcal{T}_h$ and all indices i corresponding to bubble functions, we define $V_{\kappa,i} = U_{\kappa,i}$. For each face $e \in \mathcal{E}_h$, and all indices i which correspond to shape functions vanishing on the boundary of the edge (edge bubbles), we set $V_{e,i} = \frac{1}{2}(U_{e,i} + U'_{e,i})$. Using (7.4), we easily obtain

$$\sum_i |V_{e,i} - U_{e,i}|^2 = \frac{1}{4} \sum_i |U'_{e,i} - U_{e,i}|^2 \lesssim \int_e h^{1-d} |\llbracket u \rrbracket|^2 ds, \quad (7.5)$$

where the sum is taken over all relatively interior nodes on the edge e .

Next, for each $z \in \mathcal{N}_h$ we set

$$V_z = \frac{1}{\#T(z)} \sum_{\kappa \in T(z)} U_{z,\kappa}.$$

We equip the set $T(z)$ with the equivalence relation \sim , where $\kappa \sim \kappa'$ if κ and κ' share a face of \mathcal{T}_h . The resulting graph is connected and therefore, for any two elements $\kappa \neq \kappa' \in T(z)$ there exists a path $(\kappa_1, \dots, \kappa_m)$ from κ to κ' . If any $\kappa_i \in T(z)$ appears twice in the path, then it has a loop which we can remove from it. Therefore, for the shortest path, we have $m \leq \#T(z)$.

Let e_j be the edge which joins κ_{j-1} and κ_j . We can now estimate

$$\begin{aligned}
|U_{z,\kappa} - V_z|^2 &= \left| \frac{1}{\#T(z)} \sum_{\kappa' \in T(z)} (U_{z,\kappa} - U_{z,\kappa'}) \right|^2 \\
&\lesssim \sum_{\kappa' \in T(z)} |U_{z,\kappa} - U_{z,\kappa'}|^2 \\
&\lesssim \sum_{\kappa' \in T(z)} \sum_{j=2}^m |U_{z,\kappa_j} - U_{z,\kappa_{j-1}}|^2 \\
&\lesssim \sum_{\kappa' \in T(z)} \sum_{j=2}^m \int_{e_j} h^{1-d} |\llbracket u \rrbracket|^2 ds,
\end{aligned}$$

where we used (7.4) in the last estimate. Upon defining $\tilde{e}(z) = \bigcup \{e \in \mathcal{E}_h : z \in \bar{e}\}$ we obtain

$$|U_{z,\kappa} - V_z|^2 \lesssim \int_{\tilde{e}(z)} h^{1-d} |\llbracket u \rrbracket|^2 ds. \quad (7.6)$$

If $d = 3$ then each edge γ can be treated in an analogous manner. For each relatively interior node i , we set

$$V_{\gamma,i} = \frac{1}{\#T(\gamma)} \sum_{\kappa \in T(\gamma)} U_{\gamma,\kappa,i},$$

and we obtain

$$|U_{\gamma,\kappa,i} - V_{\gamma,i}|^2 \lesssim \int_{\tilde{e}(\gamma)} h^{1-2k} |\llbracket u \rrbracket|^2 ds, \quad (7.7)$$

where $\tilde{e}(\gamma) = \bigcup \{e \in \mathcal{E}_h : \gamma \subset \bar{e}\}$.

Next, we estimate the (semi-)norms of the shape functions. First, we have

$$\int_{\kappa} |\phi_{\kappa,i}|^2 dx \approx |\kappa| \int_{\hat{\kappa}} |\phi_i|^2 d\hat{x} \approx h^d \|\phi_i\|_{L^2(\hat{\kappa})}^2.$$

For the H^1 -seminorm, let J be the Jacobi matrix of the coordinate transformation from $\hat{\kappa}$ to κ . Then, $\nabla \phi_{\kappa,i} \circ F_{\kappa} = J^{-1} \nabla \phi_i$. The shape-regularity of κ implies that $|J^{-1}| \approx h^{-1}$, and therefore

$$\int_{\kappa} |\nabla \phi_{\kappa,i}|^2 dx \lesssim |\kappa| |J^{-1}|^2 \int_{\hat{\kappa}} |\nabla \phi_i|^2 d\hat{x} \lesssim h^{d-2} |\nabla \phi_i|_{H^1(\hat{\kappa})}^2.$$

Thus, we have shown that

$$|\phi_{\kappa,i}|_{H^k(\kappa)}^2 \lesssim h^{d-2k} |\phi_i|_{H^k(\hat{\kappa})}^2, \quad k \in \{0, 1\}. \quad (7.8)$$

To compute the error committed upon replacing u by v , using (7.8), (7.5), (7.6), and possibly (7.7), we estimate

$$\begin{aligned}
|u - v|_{H^k(\kappa)}^2 &\lesssim \sum_{i=1}^{N_{\kappa}} |U_{\kappa,i} - V_{\kappa,i}|^2 |\phi_{\kappa,i}|_{H^k(\kappa)}^2 \\
&\lesssim \sum_{e \in \mathcal{E}_h, e \subset \bar{\kappa}} \int_e h^{1-2k} |\llbracket u \rrbracket|^2 ds + \sum_{\gamma \in \mathcal{G}_h, \gamma \subset \bar{\kappa}} \int_{\tilde{e}(\gamma)} h^{1-2k} |\llbracket u \rrbracket|^2 ds + \sum_{z \in \mathcal{N}_h, z \in \bar{\kappa}} \int_{\tilde{e}(z)} h^{1-2k} |\llbracket u \rrbracket|^2 ds \\
&\lesssim \int_{\tilde{e}(\kappa)} h^{1-2k} |\llbracket u \rrbracket|^2 ds,
\end{aligned}$$

where $\tilde{e}(\kappa) = \bigcup\{e' \in \mathcal{E}_h : e \cap \bar{\kappa} \neq \emptyset\}$. The minimum angle condition restricts the number of elements which may share a common edge or vertex. Thus, summing over all elements in the triangulation gives (7.2).

To obtain the $W_0^{1,\infty}$ -version of the inequality, let v_0 be a finite element function on the same mesh, with the same nodal values as v except that those which lie on the boundary are set to zero. The estimate (7.3) then follows using exactly the same procedure as above. One could also deduce (7.3) by embedding \mathcal{T}_h into a larger mesh and setting $u = 0$ outside Ω . ■

Remark 4 Lemma 4 is the ‘minimal formulation’ for our purposes in terms of generality. The proof shows that the same result can be obtained, for example, for regular meshes which are not necessarily quasiuniform and for all $k \in \mathbb{R}_{\geq 0}$. More interesting is the question whether the constant C_r may be chosen independently of the polynomial degree. ◀

7.2 A uniform broken Gårding inequality

We begin by stating a slightly refined version of Gårding’s inequality.

Lemma 5 *Let $u \in C([0, T]; C^{1,\hat{\alpha}}(\bar{\Omega})^d)$ for some $\hat{\alpha} \in (0, 1)$, and suppose that the Legendre–Hadamard condition (1.6) holds. Then, there exists $\delta_0 > 0$, and a constant $c_1 > 0$ such that for $0 \leq t \leq T$ and for any matrix-valued function $\Phi \in L^\infty(\Omega)^{d \times d}$ satisfying $\|\nabla u(t, \cdot) - \Phi\|_{L^\infty(\Omega)} \leq \delta_0$ we have*

$$\sum_{i,\alpha,j,\beta=1}^d \int_{\Omega} A_{i\alpha j\beta}(\Phi(x)) \partial_{x_\alpha} v_i \partial_{x_\beta} v_j \, dx \geq \frac{1}{2} c_0 \|\nabla v\|_{L^2(\Omega)}^2 - c_1 \|v\|_{L^2(\Omega)}^2 \quad \forall v \in H_0^1(\Omega)^d.$$

Proof The result follows upon noting that the constant c_1 depends only on the partition of Ω , which is used in the proof and can be chosen independently of Φ if $\|\nabla u(t, \cdot) - \Phi\|_{L^\infty(\Omega)}$ is sufficiently small.

Recall the definition of K_δ in (1.3) and note that $\delta \mapsto K_\delta$ is non-decreasing, i.e., we can assume without loss of generality that $K_\delta \leq K$ for some $K > 0$. Furthermore, we define

$$K'_\delta = \max_{\substack{\eta, \eta' \in \mathcal{M}_\delta \\ \eta \neq \eta'}} \frac{|A(\eta) - A(\eta')|}{|\eta - \eta'|} < \infty,$$

which is also a non-decreasing function of δ . Note also that if δ_0 is sufficiently small then $\Phi \in \mathcal{A}_\delta$.

Next, note that since $u \in C([0, T], C^{1,\hat{\alpha}}(\bar{\Omega})^d)$, there exists a constant independent of x and t such that

$$|\nabla u(t, x) - \nabla u(t, y)| \leq C|x - y|^{\hat{\alpha}} \quad \forall x, y \in \bar{\Omega}, \quad \forall t \in [0, T].$$

Thus, any open ball B with center $x_B \in \Omega$ and radius $R < (\varepsilon/(2K'_\delta C))^{1/\hat{\alpha}}$ satisfies

$$\sup_{t \in [0, T]} \sup_{y \in B \cap \Omega} |\nabla u(t, x_B) - \nabla u(t, y)| \leq \varepsilon/(2K'_\delta), \quad (7.9)$$

where ε is a positive constant which we shall fix later. It is now easy to see that there exists a finite cover \mathcal{B} of Ω containing only such balls.

Now, assume that $\|\Phi - \nabla u(t, \cdot)\|_{L^\infty(\Omega)} \leq \varepsilon/(4K'_\delta)$ for some $t \in [0, T]$. Then, for all $B \in \mathcal{B}$ we can estimate

$$\begin{aligned} |A(\Phi(x_B)) - A(\Phi(y))| &\leq K'_\delta |\Phi(x_B) - \Phi(y)| \\ &\leq 2K'_\delta \|\Phi - \nabla u(t, \cdot)\|_{L^\infty(\Omega)} + K'_\delta |\nabla u(t, x_B) - \nabla u(t, y)| \\ &\leq \varepsilon/2 + \varepsilon/2 = \varepsilon, \end{aligned}$$

leading to

$$\sup_{y \in B \cap \Omega} |A(\Phi(x_B)) - A(\Phi(y))| \leq \varepsilon \quad (7.10)$$

for all $B \in \mathcal{B}$.

From this point on we can continue with the usual proof of Gårding's inequality, following for example Theorem 6.5.1 in [12]. Let $\{\zeta_B\}$ be a partition of unity associated with \mathcal{B} , i.e.

$$\zeta_B \geq 0, \quad \zeta_B \in C_0^\infty(\mathbb{R}^d), \quad \sum_{B \in \mathcal{B}} \zeta_B^2(x) = 1, \quad \text{and} \quad \text{supp} \zeta_B \subset B.$$

For notational convenience set $A_{i\alpha j\beta}(x) = A_{i\alpha j\beta}(\Phi(x))$ and $A_{i\alpha j\beta}^B = A_{i\alpha j\beta}(\Phi(x_B))$. For each $B \in \mathcal{B}$ and each $v \in H_0^1(\Omega)^d$, it follows from a Parseval's identity that

$$\sum_{i,\alpha,j,\beta=1}^d \int_B A_{i\alpha j\beta}^B \partial_{x_\alpha}(\zeta_B v_i) \partial_{x_\beta}(\zeta_B v_j) \, dx \geq M_1 \|\nabla(\zeta_B v)\|_{L^2(B)}^2. \quad (7.11)$$

Repeated application of the ε -inequality and the use of (7.10) gives

$$\begin{aligned} \sum_{i,\alpha,j,\beta=1}^d \int_\Omega A_{i\alpha j\beta} \partial_{x_\alpha} v_i \partial_{x_\beta} v_j \, dx &= \sum_{i,\alpha,j,\beta=1}^d \int_\Omega A_{i\alpha j\beta} \left(\sum_{B \in \mathcal{B}} \zeta_B^2 \right) \partial_{x_\alpha} v_i \partial_{x_\beta} v_j \, dx \\ &= \sum_{i,\alpha,j,\beta=1}^d \sum_{B \in \mathcal{B}} \int_B A_{i\alpha j\beta} [\partial_{x_\alpha}(\zeta_B v_i) - v_i \partial_{x_\alpha} \zeta_B] [\partial_{x_\beta}(\zeta_B v_j) - v_j \partial_{x_\beta} \zeta_B] \, dx \\ &\geq \sum_{B \in \mathcal{B}} \left\{ \sum_{i,\alpha,j,\beta=1}^d \int_B A_{i\alpha j\beta} \partial_{x_\alpha}(\zeta_B v_i) \partial_{x_\beta}(\zeta_B v_j) \, dx \right. \\ &\quad \left. - 2K_\delta \|\nabla \zeta_B\|_{L^\infty(B)} \|\nabla(\zeta_B v)\|_{L^2(B)} \|v\|_{L^2(B)} - K_\delta \|\nabla \zeta_B\|_{L^\infty(B)}^2 \|v\|_{L^2(B)}^2 \right\} \\ &\geq \sum_{B \in \mathcal{B}} \left\{ \sum_{i,\alpha,j,\beta=1}^d \int_B A_{i\alpha j\beta}^B \partial_{x_\alpha}(\zeta_B v_i) \partial_{x_\beta}(\zeta_B v_j) \, dx - \varepsilon \|\nabla(\zeta_B v)\|_{L^2(B)}^2 \right. \\ &\quad \left. - \varepsilon \|\nabla(\zeta_B v)\|_{L^2(B)}^2 - c_B \|v\|_{L^2(B)}^2 \right\}, \end{aligned}$$

where c_B depends on K_δ , on ε and on ζ_B . Combining this estimate with (7.11), we obtain

$$\begin{aligned}
\sum_{i,\alpha,j,\beta=1}^d \int_{\Omega} A_{i\alpha j\beta} \partial_{x_\alpha} v_i \partial_{x_\beta} v_j \, dx &\geq \sum_{B \in \mathcal{B}} \left\{ (M_1 - 2\varepsilon) \|\nabla(\zeta_B v)\|_{L^2(B)}^2 - c_B \|v\|_{L^2(B)}^2 \right\} \\
&\geq (M_1 - 2\varepsilon) \sum_{B \in \mathcal{B}} \int_B \left[\zeta_B^2 |\nabla v|^2 - 2\zeta_B |v| |\nabla \zeta_B| |\nabla v| + |v|^2 |\nabla \zeta_B|^2 \right] dx - c'_1 \|v\|_{L^2(\Omega)}^2 \\
&\geq (M_1 - 3\varepsilon) \int_{\Omega} \left(\sum_{B \in \mathcal{B}} \zeta_B^2 \right) |\nabla v|^2 \, dx - c_1 \|v\|_{L^2(\Omega)}^2 \\
&= (M_1 - 3\varepsilon) \|\nabla v\|_{L^2(\Omega)}^2 - c_1 \|v\|_{L^2(\Omega)}^2,
\end{aligned}$$

where c'_1 and c_1 depend on M_1 and c_B and ζ_B . Setting $\varepsilon = M_1/6$ gives the desired result. \blacksquare

Theorem 6 (Uniform broken Gårding inequality) *Let $d \in \{2, 3\}$ and let $\{\mathcal{T}_h\}_{h>0}$ be a quasiuniform, uniformly simplicially reducible family of partitions of Ω . Let $u \in C([0, T], C^{1,\hat{\alpha}}(\Omega)^d)$ for some $\hat{\alpha} \in (0, 1)$ and assume that (1.6) holds. Then, there exist constants $\delta_0, c_1, c_2 > 0$, independent of h , such that for each $\Phi \in L^\infty(\Omega)^{d \times d}$ for which there exists $t \in [0, T]$ such that $\|\nabla u(t, \cdot) - \Phi\|_{L^\infty(\Omega)} \leq \delta_0$ we have*

$$\begin{aligned}
\sum_{i,\alpha,j,\beta=1}^d \int_{\Omega} A_{i\alpha j\beta}(\Phi(x)) \partial_{x_\alpha} v_i \partial_{x_\beta} v_j \, dx &\geq \frac{1}{4} M_1 |v|_{1,\mathcal{T}_h}^2 - c_1 \|v\|_{L^2(\Omega)}^2 \\
&\quad - c_2 \int_{\Gamma_{\text{int}}} h^{-1} \|\llbracket v \rrbracket\|^2 \, ds - c_2 \int_{\partial\Omega} h^{-1} |v|^2 \, ds \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).
\end{aligned}$$

Proof Let $v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ and let $w = \mathcal{R}_0 v$ be its continuous reconstruction with homogeneous boundary values, according to Lemma 4. For notational convenience, we set

$$a(\varphi, \psi) = \sum_{i,\alpha,j,\beta=1}^d \int_{\Omega} A_{i\alpha j\beta}(\Phi(x)) \partial_{x_\alpha} \varphi_i \partial_{x_\beta} \psi_j \, dx.$$

Using Lemma 5, we have

$$\begin{aligned}
a(v, v) &= a(w, w) - a(w - v, w - v) - a(w - v, v) - a(v, w - v) \\
&\geq \frac{1}{2} M_1 \|\nabla w\|_{L^2(\Omega)}^2 - c'_1 \|w\|_{L^2(\Omega)}^2 - K_\delta \|\nabla v - \nabla w\|_{L^2(\Omega)}^2 - 2K_\delta \|\nabla v - \nabla w\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} \\
&\geq \frac{1}{2} M_1 \|\nabla v\|_{L^2(\Omega)}^2 - c'_1 \|v\|_{L^2(\Omega)}^2 - (K_\delta - M_1/2) \|\nabla v - \nabla w\|_{L^2(\Omega)}^2 \\
&\quad - 2(K_\delta + M_1/2) \|\nabla v - \nabla w\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} - 2c'_1 \|v - w\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} - c'_1 \|v - w\|_{L^2(\Omega)}^2.
\end{aligned}$$

Applying ε -inequalities to the fourth and fifth term on the right-hand side, we obtain constants c_1 and c'_2 which are independent of h such that

$$a(v, v) \geq \frac{1}{4} M_1 |v|_{1,\mathcal{T}_h}^2 - c_1 \|v\|_{L^2(\Omega)}^2 - c'_2 \|v - w\|_{1,\mathcal{T}_h}^2.$$

Employing Lemma 4, we obtain the desired result. \blacksquare

7.3 Convergence analysis in the case of strongly elliptic systems

Given the uniform broken Gårding inequality, Theorem 6, only minor changes need to be made in order to extend the proof of Section 5 to the case where, instead of assuming that S is uniformly monotone, we only require that ∇S satisfies the, weaker, Legendre–Hadamard condition.

To this end, we now define the nonlinear projection $W(t)$ by the equation

$$B(W(t), v) + c(W(t), v) = B(u(t), v) + c(u(t), v) \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \quad (7.12)$$

where $c = M_1/4 + c_1$, M_1 and c_1 being the constants in Theorem 6. The error analysis for $W(t)$ then proceeds just as in Section 6, without any significant changes, and the function $W(t)$ defined by (7.12) has the same approximation properties as the function $W(t)$ considered in Section 6. In the following, we only discuss the coercivity of the corresponding bilinear form briefly.

7.3.1 Coercivity of $\tilde{b} + c(\cdot, \cdot)$

The proof of coercivity is the only step in extending the proof of Section 4 which requires further remarks. If $\varphi \in \mathcal{J}$, then $\|\varphi - \Pi_h u\|_{1, \mathcal{T}_h} \leq C_* h^r$ and therefore there exists $h_1 > 0$ such that $\|\nabla \varphi - \nabla u\|_{L^\infty(\Omega)} \lesssim h^{r-d/2} \leq \delta_0$ for all $h \in (0, h_1]$. Recalling the analysis of Section 4.2, we estimate

$$\int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); v, v) \, d\tau + c(v, v) \geq T_1 + \dots + T_5,$$

where, now,

$$\begin{aligned} T_1 &= \sum_{i, \alpha, j, \beta=1}^d \int_0^1 \int_{\Omega} A_{i\alpha j\beta} (\nabla \Pi_h u + \tau \nabla (\varphi - \Pi_h u)) \partial_{x_\alpha} v_i \partial_{x_\beta} v_j \, dx \, d\tau + c\|v\|_{L^2(\Omega)}^2 \\ &\geq \frac{1}{4} M_1 (\|\nabla v\|_{L^2(\Omega)}^2 + \|v\|_{L^2(\Omega)}^2) - c_2 \int_{\Gamma_{\text{int}} \cup \Gamma_D} h^{-1} \llbracket v \rrbracket^2 \, dx \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \end{aligned}$$

The terms T_2 and T_3 are treated in exactly the same way as in Section 4. By choosing the factor α in the definition of the penalty parameter sufficiently large, the last term above can be combined with T_4 and T_5 to give a coercive bilinear form. Instead of (4.5), we obtain

$$\int_0^1 \tilde{b}(\Pi_h u + \tau(\varphi - \Pi_h u); v, v) \, d\tau + c(v, v) \geq \frac{1}{8} \min(1, M_1) \|v\|_{1,h}^2 \quad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).$$

7.3.2 Convergence analysis

In the hyperbolic case, the modification of the nonlinear projection will introduce additional terms in the error analysis which we discuss briefly. In essence, all additional terms are of a lower order and can be therefore controlled by Gronwall's inequality.

The starting point of the analysis is (5.4). In order to be able to obtain (5.5), we add the term

$$\frac{1}{2} \frac{d}{dt} [c(\xi_\varphi, \xi_\varphi)]$$

to both sides of the equality. As a consequence, the coercivity estimate from Section 7.3.1 and thus an analogue of the left-hand side of (5.5) can be obtained. On the right-hand side of (5.5) we will have the additional term

$$c \int_0^t (\xi_\varphi, \dot{\xi}_\varphi) \, ds \leq \frac{c}{2} \int_0^t \left[\|\xi_\varphi\|_{L^2(\Omega)}^2 + \|\dot{\xi}_\varphi\|_{L^2(\Omega)}^2 \right] \, ds,$$

which does not change the structure of (5.6) and (5.7). For the proof of contractivity, the same modification can be used.

Remark 5 It is a remarkable feature of the discontinuous Galerkin formulation that in the modifications of this section, we never had to take boundary values into account. Since boundary conditions are imposed weakly, it was not necessary to modify the reconstruction operator \mathcal{R}_0 from Lemma 4 for inhomogeneous or even time-dependent boundary values. ◀

8 Conclusions

We derived optimal-order convergence estimates in the broken H^1 norm for discontinuous Galerkin finite element approximations to second-order quasilinear elliptic and hyperbolic systems of partial differential equations, using piecewise polynomials of degree $p > d/2$ in the elliptic case, and of degree $p > d/2 + 1$ in the (spatially semidiscrete) hyperbolic case, where d is the spatial dimension of the problem. In the physically relevant cases of $d = 2$ and $d = 3$ these correspond to assuming that $p \geq 2$ and $p \geq 3$, respectively. The main contribution of the paper is that these optimal-order, $\mathcal{O}(h^p)$, convergence rates have been proved without assuming that the nonlinear coefficient $S(\nabla u)$ appearing in the principal part of the operator is globally Lipschitz continuous with respect to ∇u and, in the hyperbolic case, only assuming the Legendre–Hadamard condition. Although, for technical reasons, the cases $p = 1$ (for $d = 2$) and $p = 1, 2$ (for $d = 3$) have been excluded from our analysis of the DGFEM approximation of the nonlinear hyperbolic problem, we believe that the methods considered remain optimally convergent in the energy norm in these cases as well; certainly, this is true for the nonlinear elliptic problem in the special case when the nonlinearity $\eta \mapsto S(\eta)$ is globally Lipschitz continuous and uniformly monotone (see, [9]). We note in connection with our analysis for the nonlinear elliptic system that Brouwer’s fixed point theorem can be used instead of Banach’s fixed point theorem to establish the existence of a fixed point, since \mathcal{J} is a compact convex set in a finite-dimensional linear space. Of course, in the case of the hyperbolic problem the corresponding set \mathcal{J} is infinite-dimensional, so Brouwer’s fixed point theorem is then not applicable; thus, for the sake of coherence of presentation we chose to use Banach’s fixed point theorem both in the elliptic and the hyperbolic case.

We had, quite consciously, omitted the analysis of the quasilinear elliptic system (1.5) assuming the Legendre–Hadamard condition only: in this case, the bilinear form \tilde{b} arising through linearization is not only nonsymmetric and indefinite, but also adjoint-inconsistent (in the sense of [1]), so the argument of Schatz [13] concerning the error analysis of classical conforming Galerkin finite element methods for indefinite elliptic problems does not (or, at least, does not obviously) extend to this case. We expect that by altering our discontinuous Galerkin discretization so as to ensure that \tilde{b} is symmetric and adjoint-consistent this difficulty can be overcome; however, we have not, so far, considered this problem in any particular detail.

We note that all of our results can be straightforwardly extended to quasilinear elliptic and hyperbolic partial differential equations where $S(\nabla u)$ is replaced by $S(u, \nabla u)$ under the same hypotheses; the presence of the lower-order nonlinearity causes no additional technical difficulties.

As our key objective here was to understand the analysis of discontinuous Galerkin approximations of locally Lipschitz spatial nonlinearities in quasilinear elliptic and hyperbolic systems, we did not discuss fully discrete discontinuous Galerkin finite element approximations of quasilinear hyperbolic problems. The convergence analysis of fully discrete schemes can be carried out using very similar theoretical tools to those presented here. We refer to the paper of Makridakis [11], for example, to the corresponding analysis in the case of spatially H_0^1 -conforming finite element methods which may serve as a starting point for further analytical considerations in that direction.

References

- [1] D.N. ARNOLD, F. BREZZI, B. COCKBURN, AND L.D. MARINI, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Num. Anal. **39** (2002), 1749–1779.
- [2] D. BRAESS, *Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, 1997.
- [3] S. BRENNER, *Poincaré–Friedrichs inequalities for piecewise H^1 functions*, SIAM J. Numer. Anal. **41** (2003), 306–324.
- [4] S. BRENNER, *Korn’s inequalities for piecewise H^1 vector fields*, Math. Comp. **73** (2004), no. 247, 1067–1087.
- [5] V.C. CHEN AND W. VON WAHL, *Das Rand-Anfangswertproblem für quasilineare Wellengleichungen in Sobolev-räumen niedriger Ordnung*, J. Reine Angew. Math. **337** (1982), 77–112.
- [6] B. COCKBURN, G.E. KARNIADAKIS, AND C.-W. SHU, *The development of discontinuous Galerkin methods*. In: B. Cockburn, G.E. Karniadakis, and C.-W. Shu (Eds.), *Discontinuous Galerkin Finite Element Methods*, Lecture Notes in Computational Science and Engineering 11, Springer-Verlag 2000.

- [7] C.M. DAFERMOS AND W. HRUSA, *Energy methods for quasilinear hyperbolic initial-boundary value problems. Applications to elastodynamics*, Arch. Rational Mech. Anal. **87** (1985), 267–292.
- [8] M.E. GURTIN, *An Introduction to Continuum Mechanics*, Academic Press, New York, 1981.
- [9] P. HOUSTON, J. ROBSON, AND E. SÜLI, *Discontinuous Galerkin finite element approximation of quasilinear elliptic boundary-value problems I*, IMA J. of Numer. Anal. **25** (2005), 726–749.
- [10] P. HOUSTON, CH. SCHWAB, AND E. SÜLI, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal. **39** (2002), 2133–2163.
- [11] CH. B. MAKRIDAKIS, *Finite element approximations of nonlinear elastic waves*, Math. Comp. **61** (1993), 569–594.
- [12] C.B. MORREY, *Multiple Integrals in the Calculus of Variations*, Springer-Verlag, Berlin, Heidelberg, New York, 1966.
- [13] A.H. SCHATZ, *An observation concerning Ritz–Galerkin methods with indefinite bilinear forms*. Math. Comp. **28**(128) (1974), 959–962.
- [14] S. SUN, *Discontinuous Galerkin Methods for Reactive Transport in Porous Media*. Ph.D. thesis, The University of Texas at Austin, 2003.