SIAM J. NUMER. ANAL.
Vol. 45, No. 4, pp. 1370–1397

DISCONTINUOUS GALERKIN FINITE ELEMENT APPROXIMATION OF NONLINEAR SECOND-ORDER ELLIPTIC AND HYPERBOLIC SYSTEMS*

CHRISTOPH ORTNER† AND ENDRE SÜLI†

Abstract. We develop the convergence analysis of discontinuous Galerkin finite element approximations to symmetric second-order quasi-linear elliptic and hyperbolic systems of partial differential equations in divergence form in a bounded spatial domain in $\mathbb{R}^d$, subject to mixed Dirichlet–Neumann boundary conditions. Optimal-order asymptotic bounds are derived on the discretization error in each case without requiring the global Lipschitz continuity or uniform monotonicity of the stress tensor. Instead, only local smoothness and a Gårding inequality are used in the analysis.

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

AMS subject classifications. 65M60, 74S05, 74H15

DOI. 10.1137/06067119X

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 quasi-linear elliptic and hyperbolic systems. Nonlinear elasticity is a particularly fertile source of equations of this type, and our results are phrased with this particular application area in mind, although the ideas and techniques developed are valid generally, provided the structural hypotheses on the nonlinearity assumed herein are satisfied.

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 quasi-linear elliptic partial differential equations. We shall then state the corresponding dynamic problem, which is a mixed initial-boundary-value problem for a second-order quasi-linear 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 $u : \overline{\Omega} \to \mathbb{R}^d$ such that u is a stationary point of the energy functional

$$J : v \mapsto J(v) := \int_\Omega [W(\nabla v(x)) - f(x) \cdot v(x)]\, \mathrm{d}x - \int_{\Gamma_\mathrm{N}} g_\mathrm{N}(s) \cdot v(s)\, \mathrm{d}s, \tag{1.1}$$

defined over the set of all (sufficiently smooth) d-component vector functions v on $\overline{\Omega}$ satisfying the boundary condition $v = g_\mathrm{D}$ on Γ_D, where $\Gamma_\mathrm{D} \subset \Gamma = \partial\Omega$ has positive

*Received by the editors October 1, 2006; accepted for publication (in revised form) March 22, 2007; published electronically July 11, 2007. 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).
http://www.siam.org/journals/sinum/45-4/67119.html
†University of Oxford, Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, United Kingdom (Christoph.Ortner@comlab.ox.ac.uk, Endre.Suli@comlab.ox.ac.uk).

$(d-1)$-dimensional surface measure $\mathcal{H}^{d-1}(\Gamma_{\mathrm{D}})$, $\Gamma_{\mathrm{N}} = \Gamma \setminus \Gamma_{\mathrm{D}}$, $W \in \mathrm{C}^4(\mathbb{R}^{d\times d};\mathbb{R})$ is the stored energy function, $f \in \mathrm{L}^2(\Omega)^d$ is a given body force, and $g_{\mathrm{N}} \in \mathrm{L}^2(\Gamma_{\mathrm{N}})^d$. Let us define the Piola–Kirchhoff stress tensor S as the gradient of W, that is,

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

and let

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

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:

$$-\sum_{\alpha=1}^{d} \partial_{x_\alpha} S_{i\alpha}(\nabla u(x)) = f_i(x), \qquad i = 1, \dots, d, \qquad x \in \Omega, \tag{1.2}$$

subject to the boundary conditions

$$u = g_{\mathrm{D}} \quad \text{on } \Gamma_{\mathrm{D}} \qquad \text{and} \qquad S(\nabla u)\nu = g_{\mathrm{N}} \quad \text{on } \Gamma_{\mathrm{N}}, \tag{1.3}$$

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 $\partial_{x_\alpha} = \partial/\partial x_\alpha$. We note that, except in section 7, we do not use the fact that (1.2) is an Euler–Lagrange equation but only require the symmetry of the tensor $A_{i\alpha j\beta}$.

The weak formulation of the boundary-value problem (1.2), (1.3) is posed as follows: Find the function $u \in \mathrm{H}^1_{\mathrm{D},g_{\mathrm{D}}}(\Omega)^d = \{v \in \mathrm{H}^1(\Omega)^d : v|_{\Gamma_{\mathrm{D}}} = g_{\mathrm{D}}\}$ such that

$$\int_\Omega S(\nabla u) : \nabla v \, \mathrm{d}x = \int_\Omega f \cdot v \, \mathrm{d}x + \int_{\Gamma_{\mathrm{N}}} g_{\mathrm{N}} \cdot v \, \mathrm{d}s \qquad \forall v \in \mathrm{H}^1_{\mathrm{D},0}(\Omega)^d.$$

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

For future reference we also define the bilinear form $a(\Phi;\cdot,\cdot)$, $\Phi \in \mathrm{L}^\infty(\Omega)^{d\times d}$, by

$$a(\Phi; v, w) := \sum_{i,\alpha,j,\beta=1}^{d} \int_\Omega A_{i\alpha j\beta}(\Phi)\partial_{x_\alpha} v_i \partial_{x_\beta} w_j \, \mathrm{d}x \qquad \forall v, w \in \mathrm{H}^1_{\mathrm{D},0}(\Omega)^d. \tag{1.4}$$

Formally at least, $a(\nabla u;\cdot,\cdot)$ defines the hessian of J at u; more generally, we shall consider $a(\Phi;\cdot,\cdot)$ for Φ in a certain neighborhood of ∇u which we shall now define.

For $\delta > 0$, let

$$\mathcal{Z}_\delta := \left\{\Phi \in \mathrm{C}_{\mathrm{pw}}(\overline{\Omega})^{d\times d} : \|\Phi - \nabla u\|_{\mathrm{L}^\infty(\Omega)} \le \delta\right\}, \tag{1.5}$$

where $\mathrm{C}_{\mathrm{pw}}(\overline{\Omega})$ denotes the set of bounded piecewise continuous functions defined on $\overline{\Omega}$. The set $\mathcal{Z}_\delta$ will be required in the convergence analysis of the finite element method: We will show that, for sufficiently small h, it contains the *piecewise gradients* (relative to the finite element subdivision $\mathcal{T}_h$ of the computational domain Ω) of discontinuous

Galerkin finite element approximations to u. Their point values must therefore be contained in the set

$$\mathcal{M}_\delta := \operatorname{conv}\big\{\eta \in \mathbb{R}^{d\times d} : \inf_{x\in\Omega} |\eta - \nabla u(x)| \le \delta\big\},$$

where $|\cdot|$ denotes the Frobenius norm on $\mathbb{R}^{d\times d}$ defined, for $\eta \in \mathbb{R}^{d\times d}$, by $|\eta| = (\eta : \eta)^{1/2}$. Clearly, as it is the convex hull of a closed and bounded set, $\mathcal{M}_\delta$ is itself closed, bounded, and, of course, convex.

We note here that we do not require S to be globally Lipschitz continuous, but we will use the local Lipschitz constant of S in $\mathcal{M}_\delta$, defined by

$$K_\delta := \sup_{\eta\in\mathcal{M}_\delta} \left(\sum_{i,\alpha,j,\beta=1}^{d} |A_{i\alpha j\beta}(\eta)|^2 \right)^{1/2}, \tag{1.6}$$

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

$$L_\delta := \sup_{\eta,\sigma\in\mathcal{M}_\delta,\eta\neq\sigma} |\eta-\sigma|^{-1} \left(\sum_{i,\alpha,j,\beta=1}^{d} |A_{i\alpha j\beta}(\eta) - A_{i\alpha j\beta}(\sigma)|^2 \right)^{1/2}. \tag{1.7}$$

Since, for every $\delta > 0$, the set $\mathcal{M}_\delta$ is compact in $\mathbb{R}^{d\times d}$ and $A \in \mathrm{C}^2(\mathcal{M}_\delta)^{d\times d\times d\times d}$, it follows that K_δ and L_δ are finite.

We shall also consider the dynamic counterpart of the boundary-value problem (1.2), (1.3)—the initial-boundary-value problem for the second-order nonlinear evolution equation

$$\partial_t^2 u_i - \sum_{\alpha=1}^{d} \partial_{x_\alpha} S_{i\alpha}(\nabla u) = f_i(t,x), \qquad i = 1,\dots,d, \qquad x \in \Omega, \quad t \in (0,T], \tag{1.8}$$

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 [11, 1], for example. Suitable generalizations of the sets $\mathcal{M}_\delta$ and $\mathcal{Z}_\delta$ for the hyperbolic case are given in section 6.

We now formulate our structural hypotheses on the stress tensor S. For most constitutive laws in solid mechanics and many other applications, the mapping $\eta \mapsto S(\eta)$ satisfies the *axiom of frame indifference*, that is,

$$S(F - \mathrm{id}) = S(QF - \mathrm{id}) \qquad \forall Q \in SO(d), \quad \forall F \in \mathbb{R}^{d\times d}, \tag{1.9}$$

where id is the $d \times d$ identity matrix and $SO(d)$ is the group of special orthogonal $d \times d$ matrices. Note that the form of (1.9) is slighly nonstandard, as our partial differential equation is formulated in terms of displacement rather than deformation. If S satisfies (1.9), then, except in trivial cases, S cannot be monotone; for a detailed discussion of this point, we refer to pages 490–491 in the monograph of Antman [1]. Hence, the *uniform monotonicity* condition which hypothesizes the existence of a real number $M_1 > 0$ such that

$$(S(F) - S(G)) : (F - G) \ge M_1 |F - G|^2 \qquad \forall F, G \in \mathbb{R}^{d\times d}, \tag{1.10}$$

which is commonly assumed in the analysis of finite element approximations to quasi-linear elliptic problems, is inappropriate in the context of nonlinear elasticity and needs to be relaxed in order to cover physically meaningful problems.

In fact, the condition (1.10) can be relaxed in several ways in order to capture the physics while still recovering some of the theory available in the uniformly elliptic setting which stems from the uniform monotonicity condition (1.10). It is reasonable, for example, to assume that a metastable state of the elastic energy functional (1.1) is not merely a critical point satisfying the Euler–Lagrange equation but that the hessian of J is positive definite at this point. Thus, in the static case, we shall replace (1.10) by the following condition, which requires the existence of a real number $M_1 = M_1(u) > 0$ such that

$$a(\nabla u; v, v) \geq M_1 \|\nabla v\|^2_{\mathrm{L}^2(\Omega)} \qquad \forall v \in \mathrm{H}^1_{\mathrm{D},0}(\Omega)^d. \tag{1.11}$$

Similarly, for the dynamic case, it was shown in [8] that, if S satisfies the strong Legendre–Hadamard condition

$$\sum_{i,\alpha,j,\beta=1}^{d} A_{i\alpha j\beta}(\eta)\zeta_i\zeta_j\xi_\alpha\xi_\beta \geq M_1|\zeta|^2|\xi|^2 \qquad \forall \zeta, \xi \in \mathbb{R}^d, \quad \forall \eta \in \mathbb{R}^{d\times d}, \tag{1.12}$$

for some constant $M_1 > 0$, then a smooth solution to (1.8) is guaranteed to exist locally in time, subject to given initial conditions and the same boundary conditions as in the static case (at least when $\Gamma_\mathrm{N} = \emptyset$ and $g_\mathrm{D} = 0$). Condition (1.12) is satisfied by most constitutive laws for elastic materials. In this case, the semilinear form a defined in (1.4) satisfies the following Gårding inequality: For any $\varphi \in \mathrm{C}^1(\overline{\Omega})^d$, there exists $M_0 = M_0(\varphi) \geq 0$ such that

$$a(\nabla\varphi; v, v) \geq \tfrac{1}{2} M_1 \|\nabla v\|^2_{\mathrm{L}^2(\Omega)} - M_0(\varphi)\|v\|^2_{\mathrm{L}^2(\Omega)} \qquad \forall v \in \mathrm{H}^1_0(\Omega)^d; \tag{1.13}$$

cf. Theorem 6.5.1 on p. 253 in [16]. Even this weaker inequality is, to the best of our knowledge, known only for $v \in \mathrm{H}^1_0(\Omega)^d$. As we shall see, (1.13) is sufficient for the convergence analysis in the dynamic case.

In the case of classical conforming finite element methods based on finite-dimensional subspaces of $\mathrm{H}^1_{\mathrm{D},0}(\Omega)^d$ or $\mathrm{H}^1_0(\Omega)^d$, as the case may be, consisting of continuous piecewise polynomial functions of degree $p \geq 1$ defined over a family of subdivisions $\{\mathcal{T}_h\}_{h>0}$ of the computational domain Ω, the inequalities (1.11) and (1.13) will automatically hold in such subspaces. Discontinuous Galerkin finite element methods which are the focus of this paper are, however, built over finite-dimensional spaces consisting of discontinuous piecewise polynomial functions defined on Ω, which are, clearly, *not* contained in $\mathrm{H}^1(\Omega)^d$, let alone $\mathrm{H}^1_{\mathrm{D},0}(\Omega)^d$ or $\mathrm{H}^1_0(\Omega)^d$. As a matter of fact, both (1.11) and (1.13) are global conditions and, unlike uniform monotonicity (1.10), do not automatically translate to the space $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$, defined in section 2, of discontinuous piecewise polynomial functions of degree p on $\mathcal{T}_h$. Thus, in section 3, we shall derive the “broken” versions of these inequalities which hold over $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$. To the best of our knowledge, the analysis of discontinuous Galerkin finite element approximations to second-order quasi-linear 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 [7] for a comprehensive historical survey of the field and [2, 13] 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 (1.2), (1.3). In section 3, we derive broken Gårding inequalities to aid us in our subsequent analysis. In section 4 we develop the linearization of the semilinear form appearing in the definition of the finite element method. In section 5 we perform the convergence analysis of the discontinuous Galerkin finite element approximation of the elliptic boundary-value problem (1.2), (1.3) under hypothesis (1.11). We note, in particular, that our analysis does not assume the global Lipschitz continuity of the functions $S_{i\alpha}$, $i, \alpha = 1, \ldots, d$, with respect to ∇u, nor do we explicitly require the uniform monotonicity condition (1.10). Building on the work of Makridakis [15] for classical conforming methods, in section 6 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 quasi-linear hyperbolic equations of the form (1.8). This analysis requires a nonlinear projection operator whose approximation properties are analyzed, closely following section 5, in Appendix A. Extensions of our analysis to fully discrete approximations of the hyperbolic problem would proceed along the same lines as in [15] in the case of conforming methods; thus, we do not consider these here. In section 7 we show how our framework can be used to derive optimal error estimates for discontinuous Galerkin finite element methods other than the formulation which we have adopted in this paper.

2. Finite element spaces. For $h \in (0, 1]$, 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 $h_\kappa = \mathrm{diam}(\kappa)$. Each $\kappa \in \mathcal{T}_h$ is assumed to be the image of the open reference simplex under a bijective affine mapping or of the open unit hypercube under a bilinear mapping, denoted by F_κ. We shall denote either master element by $\hat{\kappa}$.

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. We collect the F_κ in the vector $\mathbf{F} = \{F_\kappa : \kappa \in \mathcal{T}_h\}$ and consider, for $p \geq 1$, the finite element space

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

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

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. Figure 2.1), $\mathcal{T}_h$ may be

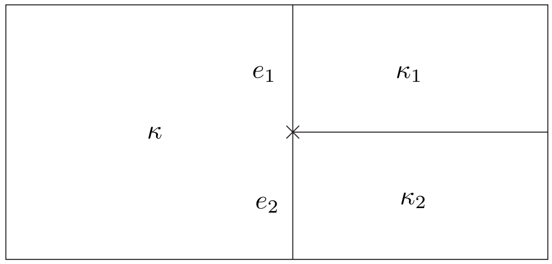

FIG. 2.1. *Hanging node* $\times$ *and faces* $e_1, e_2 \in \mathcal{E}_{\rm int}$.

irregular, and therefore $\mathcal{E}$ will be understood to contain the smallest common $(d-1)$-dimensional open faces of neighboring elements. Further, we denote by $\mathcal{E}_{\rm int}$ the set of all e in $\mathcal{E}$ that are contained in Ω, we let $\Gamma_{\rm int} = \{x \in \Omega : x \in e \text{ for some } e \in \mathcal{E}_{\rm int}\}$, and we introduce the set $\mathcal{E}_{\rm D}$ of $(d-1)$-dimensional boundary faces contained in the subset $\Gamma_{\rm D}$ 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 $\Gamma_{\rm D}$ or $\Gamma_{\rm N}$. Given $e \in \mathcal{E}$, we define $h_e := \mathrm{diam}(e)$.

In the convergence analyses of the discontinuous Galerkin finite element approximations to the partial differential equations considered here, we shall adopt the following hypotheses on the family $\{\mathcal{T}_h\}_{h>0}$, the first of which controls the number of hanging nodes which any one element may have, the second is the standard quasi-uniformity assumption, while the third is a technical condition on the lowest polynomial degree which our analysis admits. **H2** and **H3** are required in order to deduce, by the use of inverse inequalities from bounds in a broken H^1 norm, that the elementwise gradient of the numerical solution lies in $\mathcal{Z}_\delta$. Finally, the fourth hypothesis is required for the definition of the *continuous reconstruction operator* in section 3. We assume that the assumptions **H1**–**H4** hold throughout the remainder of the article.

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 $\kappa \in \mathcal{T}_h$,

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

H2. The family of subdivisions $\{\mathcal{T}_h\}_{h>0}$ is *quasi-uniform*; 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 (1.2) the polynomial degree $p > d/2$, and in the case of the hyperbolic problem (1.8) the polynomial degree $p > (d/2) + 1$ (viz. $p \geq 2$ for $d = 2, 3$, and $p \geq 3$ for $d = 2, 3$, respectively).

H4. The family of subdivisions $\{\mathcal{T}_h\}_{h>0}$ is *uniformly simplicially reducible*; i.e., for each $h > 0$ there exists a regular (no hanging nodes) 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 such 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 $h / \min_{\kappa \in \tilde{\mathcal{T}}_h} h_\kappa \leq C$.

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

$$\begin{aligned} \|\nabla w\|_{\mathrm{L}^\infty(\kappa)} &\le \frac{C_3}{h_\kappa^{d/2}}\|\nabla w\|_{\mathrm{L}^2(\kappa)}, \quad & \|\nabla w\|^2_{\mathrm{L}^2(\kappa)} &\le \frac{C_3}{h_\kappa^2}\|w\|^2_{\mathrm{L}^2(\kappa)}, \\ \|w\|^2_{\mathrm{L}^2(e)} &\le \frac{C_3}{h_e}\|w\|^2_{\mathrm{L}^2(\kappa)}, & \|\nabla w\|^2_{\mathrm{L}^2(e)} &\le \frac{C_3}{h_e}\|\nabla w\|^2_{\mathrm{L}^2(\kappa)}, \end{aligned} \tag{2.1}$$

for all $w \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$. In the case of the first two inverse inequalities 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 elementwise weak derivative (called the broken derivative) and the elementwise weak gradient (called the 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. *Let the* broken Sobolev space $\mathrm{H}^1(\Omega, \mathcal{T}_h)$ *be defined by*

$$\mathrm{H}^1(\Omega, \mathcal{T}_h) := \left\{ v \in \mathrm{L}^2(\Omega) : v|_\kappa \in \mathrm{H}^1(\kappa) \quad \forall \kappa \in \mathcal{T}_h \right\}.$$

For $v \in \mathrm{H}^1(\Omega, \mathcal{T}_h)$, we use ∇v to denote the piecewise weak gradient of v (relative to $\mathcal{T}_h$), i.e.,

$$\nabla v(x) := \nabla v|_\kappa(x) \qquad \forall x \in \kappa, \quad \forall \kappa \in \mathcal{T}_h,$$

where, on the right-hand side, $\nabla v|_\kappa$ denotes the weak gradient of $v|_\kappa \in \mathrm{H}^1(\kappa)$. The broken partial derivative $\partial_{x_j} v_i = \partial v_i / \partial x_j$ of $v \in \mathrm{H}^1(\Omega, \mathcal{T}_h)^d$ is the (i, j) component of its broken gradient ∇v.

For each $e \in \mathcal{E}_{\mathrm{int}}$ 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 $v \in \mathrm{H}^1(\Omega, \mathcal{T}_h)^d$ across e and the *mean value* of v on e by

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

respectively. If $e \in \mathcal{E}_{\mathrm{D}}$ is a face on the Dirichlet boundary, contained in the boundary $\partial\kappa$ of an element $\kappa \in \mathcal{T}_h$, it is also customary to define

$$[\![v]\!]_e := v|_{\partial\kappa \cap e} \quad \text{and} \quad \langle v \rangle_e := v|_{\partial\kappa \cap e}.$$

These definitions will enable us to condense our notation. 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, $i > j$.

Suppose that σ is a positive, piecewise constant function defined on $\Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{int}}$ (to be defined below). We equip the space $\mathrm{H}^1(\Omega, \mathcal{T}_h)$ with the *broken Sobolev norm* $\|\cdot\|_{1,h}$ defined by

$$\|v\|_{1,h} := \left(\int_\Omega |\nabla v|^2 \,\mathrm{d}x + \int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} \sigma |[\![v]\!]|^2 \,\mathrm{d}s \right)^{1/2}.$$

For the definition of the discontinuous Galerkin method, we introduce the semilinear form

$$
\begin{aligned}
B(w,v) := \int_\Omega S(\nabla w) : \nabla v \,\mathrm{d}x - \int_{\Gamma_\mathrm{D}} S(\nabla w)\nu \cdot v \,\mathrm{d}s - \int_{\Gamma_\mathrm{int}} \langle S(\nabla w)\nu\rangle \cdot [\![v]\!] \,\mathrm{d}s \\
+ \int_{\Gamma_\mathrm{int}\cup\Gamma_\mathrm{D}} \sigma [\![w]\!] \cdot [\![v]\!] \,\mathrm{d}s, \qquad w \in \mathrm{C}^1(\overline{\Omega}, \mathcal{T}_h)^d, \quad v \in \mathrm{H}^1(\Omega,\mathcal{T}_h)^d,
\end{aligned}
\tag{2.2}
$$

and the linear functional

$$
\ell(v) := \int_\Omega f \cdot v \,\mathrm{d}x + \int_{\Gamma_\mathrm{D}} \sigma g_\mathrm{D} \cdot v \,\mathrm{d}s + \int_{\Gamma_\mathrm{N}} g_\mathrm{N} \cdot v \,\mathrm{d}s, \qquad v \in \mathrm{H}^1(\Omega,\mathcal{T}_h)^d. \tag{2.3}
$$

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

$$
\sigma|_e := \sigma_e = \frac{\alpha}{h_e} \qquad \text{for } e \subset \Gamma_\mathrm{int} \cup \Gamma_\mathrm{D}. \tag{2.4}
$$

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

The discontinuous Galerkin finite element approximation of problem (1.2), (1.3) is posed as follows: Find $u_\mathrm{DG} \in S^p(\Omega,\mathcal{T}_h,\mathbf{F})$ such that

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

If the problem were linear, our discretization would correspond to the incomplete interior penalty method (see, for example, [9, 18]).

3. Broken Gårding inequality. The proofs of the broken versions of the Gårding inequalities (1.11) and (1.13) rely on the construction of a recovery operator, which connects each discontinuous piecewise polynomial function from $S^p(\Omega,\mathcal{T}_h,\mathbf{F})$ to a continuous relative. Such an operator has been used previously in similar contexts, for example, by Karakashian and Pascal [14] for deriving residual-based a posteriori error estimates and by Brenner [4] for the proof of broken Korn inequalities.

Here we follow the construction used by Karakashian and Pascal [14], though we will slightly reformulate their result. By our hypothesis **H4**, the family of meshes $(\mathcal{T}_h)_{h>0}$ is uniformly simplicially reducible, meaning that, for each h there exists a regular simplicial mesh $\widetilde{\mathcal{T}}_h$ which refines $\mathcal{T}_h$. For example, quasi-uniform families of 1-irregular meshes in two dimensions satisfy this property (cf. Figure 3.1 and Proposition 2 in [17]). Another important class are quasiuniform quadrilateral meshes obtained by hierarchical refinement (cf. Proposition 3 in [17]). For such families of meshes, we have the following result. For a proof we refer to Theorems 2.2 and 2.3 in [14] or section 7.1 in [17].

LEMMA 3.1. *There exists a constant C_r, independent of h, and a linear operator $\mathcal{R}\colon S^p(\Omega,\mathcal{T}_h,\mathbf{F}) \to \mathrm{H}^1_{\mathrm{D},0}(\Omega)^d$ such that, for all $u \in S^p(\Omega,\mathcal{T}_h,\mathbf{F})$ and $k \in \{0,1\}$,*

$$
\|\nabla^k(u - \mathcal{R}u)\|_{\mathrm{L}^2(\Omega)} \le C_r \int_{\Gamma_\mathrm{int}\cup\Gamma_\mathrm{D}} h^{1-2k} |[\![u]\!]|^2 \,\mathrm{d}s, \tag{3.1}
$$

where $\nabla^0 = \mathrm{id}$ and $\nabla^1 = \nabla$.

Lemma 3.1 provides a link between discontinuous piecewise polynomial functions and functions in $\mathrm{H}^1_{\mathrm{D},0}(\Omega)^d$. Thus, to establish a broken Gårding inequality, we replace the test function v by its continuous representative $\mathcal{R}v$ and estimate the error

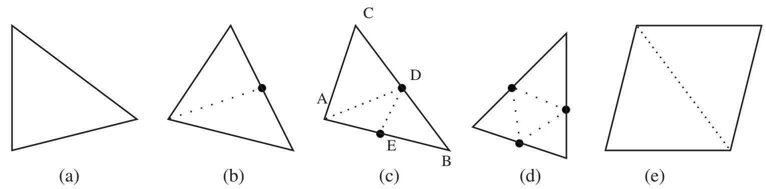

FIG. 3.1. *Refinement of triangular elements in the presence of hanging nodes in order to obtain the mesh $\tilde{\mathcal{T}}_h$ featured in hypothesis* **H4**.

committed in doing so in terms of the jumps of v. This procedure yields the following result.

LEMMA 3.2. *Let $u \in \mathrm{C}^1(\overline{\Omega})^d$ be such that the following Gårding inequality holds:*

$$a(\nabla u; v, v) \geq M_1 \|\nabla v\|^2_{\mathrm{L}^2(\Omega)} - M_0 \|v\|^2_{\mathrm{L}^2(\Omega)} \qquad \forall v \in \mathrm{H}^1_{\mathrm{D},0}(\Omega)^d, \tag{3.2}$$

where $M_1 > 0$ and $M_0 \geq 0$. Assume furthermore that $\delta \leq M_1/(4L_\delta)$. Then, for all $\Phi \in \mathcal{Z}_\delta$ and $h \leq 1$, the following broken Gårding inequality holds:

$$\begin{aligned} a(\Phi; v, v) \geq{} & \tfrac{1}{2} M_1 \|\nabla v\|^2_{\mathrm{L}^2(\Omega)} - 2M_0 \|v\|^2_{\mathrm{L}^2(\Omega)} \\ & - C_1 \int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} h^{-1} |[\![v]\!]|^2 \, \mathrm{d}s \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \end{aligned} \tag{3.3}$$

where $C_1 = C_1(M_0, M_1, K_\delta, C_r)$ is independent of h.

Proof. Note that the definition (1.6) of K_δ implies that

$$a(\nabla u; v, w) \leq K_\delta \|\nabla v\|_{\mathrm{L}^2(\Omega)} \|\nabla w\|_{\mathrm{L}^2(\Omega)} \qquad \forall v, w \in \mathrm{H}^1(\Omega, \mathcal{T}_h)^d.$$

Step 1. We begin by assuming that $\Phi = \nabla u$. In this case, we then have that

$$\begin{aligned} a(\nabla u; v, v) &= a(\nabla u; \mathcal{R}v, \mathcal{R}v) + a(\nabla u; v - \mathcal{R}v, v - \mathcal{R}v) + 2a(\nabla u; v - \mathcal{R}v, \mathcal{R}v) \\ &\geq M_1 \|\nabla \mathcal{R}v\|^2_{\mathrm{L}^2(\Omega)} - M_0 \|\mathcal{R}v\|^2_{\mathrm{L}^2(\Omega)} - K_\delta \|\nabla v - \nabla \mathcal{R}v\|^2_{\mathrm{L}^2(\Omega)} \\ &\quad - 2K_\delta \|\nabla v - \nabla \mathcal{R}v\|_{\mathrm{L}^2(\Omega)} \|\nabla \mathcal{R}v\|_{\mathrm{L}^2(\Omega)} \\ &\geq M_1 \|\nabla v + (\nabla \mathcal{R}v - \nabla v)\|^2_{\mathrm{L}^2(\Omega)} - 2M_0 \|v\|^2_{\mathrm{L}^2(\Omega)} - 2M_0 \|\mathcal{R}v - v\|^2_{\mathrm{L}^2(\Omega)} \\ &\quad - 3K_\delta \|\nabla v - \nabla \mathcal{R}v\|^2_{\mathrm{L}^2(\Omega)} - 2K_\delta \|\nabla v - \nabla \mathcal{R}v\|_{\mathrm{L}^2(\Omega)} \|\nabla v\|_{\mathrm{L}^2(\Omega)}. \end{aligned}$$

Using the inverse triangle inequality in the first term on the right-hand side of the last inequality gives

$$\begin{aligned} a(\nabla u; v, v) \geq{} & M_1 (\|\nabla v\|^2_{\mathrm{L}^2(\Omega)} - 2\|\nabla v\|_{\mathrm{L}^2(\Omega)} \|\nabla v - \nabla \mathcal{R}v\|_{\mathrm{L}^2(\Omega)} + \|\nabla v - \nabla \mathcal{R}v\|^2_{\mathrm{L}^2(\Omega)}) \\ & -2M_0 \|v\|^2_{\mathrm{L}^2(\Omega)} - 2M_0 \|\mathcal{R}v - v\|^2_{\mathrm{L}^2(\Omega)} \\ & -3K_\delta \|\nabla v - \nabla \mathcal{R}v\|^2_{\mathrm{L}^2(\Omega)} - 2K_\delta \|\nabla v - \nabla \mathcal{R}v\|_{\mathrm{L}^2(\Omega)} \|\nabla v\|_{\mathrm{L}^2(\Omega)}. \end{aligned}$$

We use the ε-inequality, $ab \leq \frac{\varepsilon}{2} a^2 + \frac{1}{2\varepsilon} b^2$, $\varepsilon > 0$, twice, with $\varepsilon = \varepsilon_1 > 0$ and $\varepsilon = \varepsilon_2 > 0$, to obtain

$$\begin{aligned} a(\nabla u; v, v) \geq{} & (M_1 - \varepsilon_1 M_1 - \varepsilon_2 K_\delta) \|\nabla v\|^2_{\mathrm{L}^2(\Omega)} - 2M_0 \|v\|^2_{\mathrm{L}^2(\Omega)} - 2M_0 \|v - \mathcal{R}v\|^2_{\mathrm{L}^2(\Omega)} \\ & -(3K_\delta - M_1 + \varepsilon_1^{-1} M_1 + \varepsilon_2^{-1} K_\delta) \|\nabla v - \nabla \mathcal{R}v\|^2_{\mathrm{L}^2(\Omega)}. \end{aligned} \tag{3.4}$$

Step 2. Next, for each $\Phi \in \mathcal{Z}$, and $v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$, we can use the Lipschitz condition (1.7), which immediately implies that

$$a(\Phi; v, v) \geq a(\nabla u; v, v) - L_\delta \|\nabla u - \Phi\|_{\mathrm{L}^\infty(\Omega)} \|\nabla v\|^2_{\mathrm{L}^2(\Omega)}.$$

As, by hypothesis, $\|\nabla u - \Phi\|_{\mathrm{L}^\infty(\Omega)} \leq \delta$, with $\delta \leq M_1/(4L_\delta)$, it is straightforward to choose ε_1 and ε_2 in (3.4) and to apply (3.1) in order to obtain (3.3). $\square$

4. Linearization. Before embarking on the analysis of the discontinuous Galerkin finite element method (2.5), we prove some auxiliary results about its linearization. We begin by noting that for any $\eta, \zeta \in \mathbb{R}^{d\times d}$ we have that

$$\begin{aligned} S_{i\alpha}(\eta) - S_{i\alpha}(\zeta) &= \sum_{j,\beta=1}^{d} (\eta_{j\beta} - \zeta_{j\beta}) \int_0^1 \frac{\partial S_{i\alpha}}{\partial \eta_{j\beta}}(\tau\eta + (1-\tau)\zeta)\,\mathrm{d}\tau \\ &= \sum_{j,\beta=1}^{d} (\eta_{j\beta} - \zeta_{j\beta}) \int_0^1 A_{i\alpha j\beta}(\tau\eta + (1-\tau)\zeta)\,\mathrm{d}\tau. \end{aligned} \tag{4.1}$$

Let $\mathrm{C}^1(\overline{\Omega}, \mathcal{T}_h)^d$ denote the space of all d-component piecewise C^1 functions, relative to the subdivision $\mathcal{T}_h$, defined on $\overline{\Omega}$. Taking (4.1) as a starting point, a straightforward computation shows that for any $w_i \in \mathrm{C}^1(\overline{\Omega}, \mathcal{T}_h)^d$, $i = 1, 2$, we have that

$$B(w_1, v) - B(w_2, v) = \int_0^1 \tilde{b}(w_2 + \tau(w_1 - w_2); w_1 - w_2, v)\,\mathrm{d}\tau \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \tag{4.2}$$

where, for $\varphi \in \mathrm{C}^1(\overline{\Omega}, \mathcal{T}_h)^d$, $\tilde{b}(\varphi; \cdot, \cdot)$ is the bilinear form defined by

$$\begin{aligned} \tilde{b}(\varphi; v, w) := &\int_\Omega \sum_{i,\alpha,j,\beta=1}^{d} A_{i\alpha j\beta}(\nabla\varphi) \frac{\partial w_i}{\partial x_\alpha} \frac{\partial v_j}{\partial x_\beta}\,\mathrm{d}x - \int_{\Gamma_{\mathrm{D}}} \sum_{i,\alpha,j,\beta=1}^{d} A_{i\alpha j\beta}(\nabla\varphi) w_i \nu_\alpha \frac{\partial v_j}{\partial x_\beta}\,\mathrm{d}s \\ &- \int_{\Gamma_{\mathrm{int}}} \sum_{i,\alpha,j,\beta=1}^{d} \left\langle A_{i\alpha j\beta}(\nabla\varphi)\nu_\alpha \frac{\partial v_j}{\partial x_\beta} \right\rangle [\![w_i]\!]\,\mathrm{d}s + \int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma [\![v]\!] \cdot [\![w]\!]\,\mathrm{d}s. \end{aligned}$$

In the next section, we shall use $\tilde{b}$ to perform a convergence analysis of the method (2.5), where the Gårding inequality (3.2) and the local Lipschitz continuity of $\tilde{b}$ w.r.t. its first argument are crucial. We prove these three results in the following three lemmas.

LEMMA 4.1. *Suppose that $u \in \mathrm{C}^1(\overline{\Omega})^d$ satisfies the Gårding inequality* (3.2). *Then there exists $\alpha_0 > 0$, independent of h, such that for all $\alpha \geq \alpha_0$, for all $h \in (0, 1]$, and for all $\varphi \in \mathrm{C}^1(\overline{\Omega}, \mathcal{T}_h)^d$, with $\nabla\varphi \in \mathcal{Z}_\delta$,*

$$\tilde{b}(\varphi; v, v) \geq \tilde{M}_1 \|v\|^2_{1,h} - 2M_0 \|v\|^2_{\mathrm{L}^2(\Omega)} \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \tag{4.3}$$

where $\tilde{M}_1 := \frac{1}{4}\min(1, M_1)$.

Proof. For $\nabla\varphi \in \mathcal{Z}_\delta$ fixed and $v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ we consider

$$\begin{aligned} \tilde{b}(\varphi; v, v) = &\int_\Omega \sum_{i,\alpha,j,\beta=1}^{d} A_{i\alpha j\beta}(\nabla\varphi) \frac{\partial v_i}{\partial x_\alpha} \frac{\partial v_j}{\partial x_\beta}\,\mathrm{d}x - \int_{\Gamma_{\mathrm{D}}} \sum_{i,\alpha,j,\beta=1}^{d} A_{i\alpha j\beta}(\nabla\varphi) v_i \nu_\alpha \frac{\partial v_j}{\partial x_\beta}\,\mathrm{d}s \\ &- \int_{\Gamma_{\mathrm{int}}} \sum_{i,\alpha,j,\beta=1}^{d} \left\langle A_{i\alpha j\beta}(\nabla\varphi)\nu_\alpha \frac{\partial v_j}{\partial x_\beta} \right\rangle [\![v_i]\!]\,\mathrm{d}s + \int_{\Gamma_{\mathrm{D}}} \sigma |v|^2\,\mathrm{d}s + \int_{\Gamma_{\mathrm{int}}} \sigma |[\![v]\!]|^2\,\mathrm{d}s \\ \equiv\; &T_1 + T_2 + T_3 + T_4 + T_5. \end{aligned}$$

Lemma 3.2 implies that

$$T_1 \geq \tfrac{1}{2} M_1 \|\nabla v\|^2_{\mathrm{L}^2(\Omega)} - 2M_0 \|v\|^2_{\mathrm{L}^2(\Omega)} - C_1 \int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} h^{-1} |[\![v]\!]|^2 \, \mathrm{d}s,$$

where C_1 is independent of h and φ.

Next, we bound T_2. Since we assumed that $\nabla\varphi \in \mathcal{Z}_\delta$, it follows that $\nabla\varphi(x) \in \mathcal{M}_\delta$ for a.e. $x \in \Omega$. Hence,

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

where K_δ is defined in (1.6). Using the third of the inverse inequalities (2.1) and recalling the definition of the penalty parameter σ_e on $e \subset \Gamma_{\mathrm{D}}$, we have that

$$|\mathrm{T}_2| \leq K_\delta (C_3 \alpha^{-1} 2d)^{1/2} \left(\int_\Omega |\nabla v|^2 \, \mathrm{d}x \right)^{1/2} \left(\int_{\Gamma_{\mathrm{D}}} \sigma |v|^2 \, \mathrm{d}s \right)^{1/2},$$

where $2d$ stands for the maximum number of faces any one element may have on Γ_{D}.

Analogously,

$$|\mathrm{T}_3| \leq K_\delta \int_{\Gamma_{\mathrm{int}}} \langle |\nabla v| \rangle \, |[\![v]\!]| \, \mathrm{d}s \leq K_\delta \left(\int_{\Gamma_{\mathrm{int}}} \sigma^{-1} \langle |\nabla v| \rangle^2 \, \mathrm{d}s \right)^{1/2} \left(\int_{\Gamma_{\mathrm{int}}} \sigma [\![v]\!]^2 \, \mathrm{d}s \right)^{1/2}.$$

Let us note that

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

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

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

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

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

Thanks to our assumption **H1** 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 that

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

and therefore

$$|\mathrm{T}_3| \le K_\delta (C_3\alpha^{-1}c_d)^{1/2} \left(\int_\Omega |\nabla v|^2\,\mathrm{d}x\right)^{1/2} \left(\int_{\Gamma_{\mathrm{int}}} \sigma [\![v]\!]^2\,\mathrm{d}s\right)^{1/2}. \tag{4.4}$$

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

$$\begin{aligned}\int_0^1 \tilde{b}(\varphi; v, v)\,\mathrm{d}\tau \ \ge\ & \tfrac{1}{2}M_1\|\nabla v\|^2_{\mathrm{L}^2(\Omega)} - 2M_0\|v\|^2_{\mathrm{L}^2(\Omega)} + \int_{\Gamma_\mathrm{D}} \sigma|v|^2\,\mathrm{d}s + \int_{\Gamma_{\mathrm{int}}} \sigma|[\![v]\!]|^2\,\mathrm{d}s\\ & - K_\delta(C_3\alpha^{-1}2d)^{1/2}\left(\int_\Omega |\nabla v|^2\,\mathrm{d}x\right)^{1/2}\left(\int_{\Gamma_\mathrm{D}} \sigma|v|^2\,\mathrm{d}s\right)^{1/2}\\ & - K_\delta(C_3\alpha^{-1}c_d)^{1/2}\left(\int_\Omega |\nabla v|^2\,\mathrm{d}x\right)^{1/2}\left(\int_{\Gamma_{\mathrm{int}}} \sigma|[\![v]\!]|^2\,\mathrm{d}s\right)^{1/2}.\end{aligned}$$

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

$$\int_0^1 \tilde{b}(\varphi; v, v)\,\mathrm{d}\tau \ge \frac{M_1}{2}\left(1 - \frac{K_\delta^2 C_3 C_d}{2M_1\alpha}\right)\int_\kappa |\nabla v|^2\,\mathrm{d}x + \frac{1}{2}\int_{\Gamma_{\mathrm{int}}\cup\Gamma_\mathrm{D}} \sigma[\![v]\!]^2\,\mathrm{d}s - 2M_0\|v\|^2_{\mathrm{L}^2}.$$

On selecting α such that $\alpha \ge K_\delta^2 M_1^{-1} C_3 C_d \equiv \alpha_0$, we deduce that, for all $h \in (0,1]$, (4.3) holds. ◻

LEMMA 4.2. *For each $\delta > 0$ there exists a constant $\tilde{K}_\delta$ depending only on K_δ, C_3, and c_d such that, for all $\varphi \in \mathrm{C}^1(\Omega, \mathcal{T}_h)^d$ with $\nabla\varphi \in \mathcal{Z}_\delta$,*

$$\left|\tilde{b}(\varphi; v, w)\right| \le \tilde{K}_\delta \|v\|_{1,h}\|w\|_{1,h} \qquad \forall v, w \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).$$

Proof. By the definition of K_δ, for $\nabla\varphi \in \mathcal{Z}_\delta$, we have that

$$\sum_{i,\alpha,j,\beta=1}^{d} \int_\Omega |A_{i\alpha j\beta}(\nabla\varphi)||\partial_{x_\alpha} v_i||\partial_{x_\beta} w_j|\,\mathrm{d}x \le K_\delta\|\nabla v\|_{\mathrm{L}^2(\Omega)}\|\nabla w\|_{\mathrm{L}^2(\Omega)},$$

and, using also the fourth inverse inequality from (2.1),

$$\begin{aligned}\sum_{i,\alpha,j,\beta=1}^{d} \int_{\Gamma_\mathrm{D}} |A_{i\alpha j\beta}(\nabla\varphi)||w_i||\nu_\alpha||\partial_{x_\beta} v_j|\,\mathrm{d}s &\le K_\delta \int_{\Gamma_\mathrm{D}} |w||\nabla v|\,\mathrm{d}s\\ &\le\ K_\delta\left[\int_{\Gamma_\mathrm{D}} \sigma|w|^2\,\mathrm{d}s\right]^{1/2}\left[\int_{\Gamma_\mathrm{D}} \sigma^{-1}|\nabla v|^2\,\mathrm{d}s\right]^{1/2}\\ &\le\ K_\delta C(C_3, c_d)\|\nabla v\|_{\mathrm{L}^2(\Omega)}\|\sigma^{1/2}w\|_{\mathrm{L}^2(\Gamma_\mathrm{D})}.\end{aligned}$$

Using a similar argument, we can deduce that

$$\sum_{i,\alpha,j,\beta=1}^{d} \int_{\Gamma_{\mathrm{int}}} \left|\left\langle A_{i\alpha j\beta}(\nabla\varphi)\nu_\alpha\partial_{x_\beta} v\right\rangle\right|\left|[\![w_i]\!]\right|\,\mathrm{d}s \le K_\delta C(C_3, c_d)\|\nabla v\|_{\mathrm{L}^2(\Omega)}\|\sigma^{1/2}[\![w]\!]\|_{\mathrm{L}^2(\Gamma_{\mathrm{int}})}.$$

The result follows by inserting these three estimates into the definition of $\tilde{b}$. ◻

LEMMA 4.3. *For every $\delta > 0$ there exists a constant $\tilde{L}_\delta$, depending only on L_δ, C_3, and c_d such that, for all $\varphi, \psi \in \mathrm{C}^1(\Omega, \mathcal{T}_h)^d$ with $\nabla\varphi, \nabla\psi \in \mathcal{Z}_\delta$,*

$$\left|\tilde{b}(\varphi; v, w) - \tilde{b}(\psi; v, w)\right| \leq \tilde{L}_\delta \|\nabla\varphi - \nabla\psi\|_{\mathrm{L}^\infty(\Omega)} \|\nabla v\|_{\mathrm{L}^2(\Omega)} \|w\|_{1,h} \quad \forall v, w \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).$$

Proof. The proof follows precisely that of Lemma 4.2. Using the fact that the integrands in $\tilde{b}$ are linear in the tensor, we can replace $A_{i\alpha j\beta}(\nabla\varphi)$ by $(A_{i\alpha j\beta}(\nabla\varphi) - A_{i\alpha j\beta}(\nabla\psi))$ and use the Lipschitz condition (1.7) instead of the bound (1.6). Furthermore, the penalty terms cancel each other out, which gives $\|\nabla v\|_{\mathrm{L}^2(\Omega)}$ instead of $\|v\|_{1,h}$; see [17] for additional details. □

5. The elliptic case. Throughout this section, we assume that $u \in \mathrm{H}^{m+1}(\Omega)^d$, with $m > d/2$, is a solution of (1.2), (1.3), satisfying the Gårding inequality (1.11); in our analysis of the discontinuous Galerkin finite element approximation to the corresponding hyperbolic problem (1.8), we shall suppose that the weaker inequality (3.2) holds.

The convergence analysis will be based on Banach's fixed point theorem. We begin by constructing a nonlinear mapping whose unique fixed point in a neighborhood of u is the numerical solution u_{DG}. For this purpose, 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 $(\Pi_h u)|_\kappa := \Pi_p^{\hat{\kappa}}(u|_\kappa \circ F_\kappa) \in \mathcal{R}_p(\hat{\kappa})$, where $\Pi_p^{\hat{\kappa}}(u|_\kappa \circ F_\kappa)$ is the classical finite element interpolant of $u|_\kappa \circ F_\kappa$ from $\mathcal{R}_p(\hat{\kappa})$. We can take $w_1 = u_{\mathrm{DG}}$ and $w_2 = \Pi_h u$ in the identity (4.2) above, which gives

$$B(u_{\mathrm{DG}}, v) - B(\Pi_h u, v) = \int_0^1 \tilde{b}(\Pi_h u + \tau(u_{\mathrm{DG}} - \Pi_h u); u_{\mathrm{DG}} - \Pi_h u, v)\,\mathrm{d}\tau \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).$$

Let us write

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

Note that since $u \in \mathrm{C}^1(\overline{\Omega})^d \cap \mathrm{H}^2(\Omega)^d$, we have that $B(u, v) = \ell(v)$ for all v in $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$. Hence,

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

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

$$\int_0^1 \tilde{b}(\Pi_h u + \tau(u_{\mathrm{DG}} - \Pi_h u); u_{\mathrm{DG}} - \Pi_h u, v)\,\mathrm{d}\tau = \int_0^1 \tilde{b}(\Pi_h u + \tau(u - \Pi_h u); u - \Pi_h u, v)\,\mathrm{d}\tau. \tag{5.1}$$

Upon defining the bilinear form $\tilde{B}(\varphi; \cdot, \cdot)$ by

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

we may rewrite (5.1) as

$$\tilde{B}(u_{\mathrm{DG}}; u_{\mathrm{DG}} - \Pi_h u, v) = \tilde{B}(u; u - \Pi_h u, v) \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \tag{5.2}$$

Lemmas 4.1–4.3 immediately imply that

$$\tilde{B}(\varphi; v, v) \geq \tilde{M}_1 \|v\|_{1,h}^2, \tag{5.3}$$

$$\left|\tilde{B}(\varphi; v, w)\right| \leq \tilde{K}_\delta \|v\|_{1,h} \|w\|_{1,h}, \quad \text{and} \tag{5.4}$$

$$\left|\tilde{B}(\varphi; v, w) - \tilde{B}(\psi; v, w)\right| \leq \tilde{L}_\delta \|\nabla v\|_{\mathrm{L}^2(\Omega)} \|w\|_{1,h} \tag{5.5}$$

for all $v, w \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ and $\varphi, \psi \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ such that $\nabla\varphi, \nabla\psi \in \mathcal{Z}_\delta$.

Let us recall our hypotheses that $u \in \mathrm{H}^{m+1}(\Omega)^d$, 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 $\mathrm{H}^1(\Omega, \mathcal{T}_h)^d$:

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

where C_* is a fixed positive constant, independent of h, whose actual value will be fixed below. 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 $\mathrm{H}^1(\Omega, \mathcal{T}_h)^d$ in the topology induced by the norm $\|\cdot\|_{1,h}$. Finally, we note that for each $v \in \mathcal{J}$, using the first inverse inequality in (2.1) and the approximation properties of Π_h (see, for example, [6]), we have that

$$\begin{aligned} \|\nabla v - \nabla u\|_{\mathrm{L}^\infty(\Omega)} &\leq \|\nabla v - \nabla \Pi_h u\|_{\mathrm{L}^\infty(\Omega)} + \|\nabla \Pi_h u - \nabla u\|_{\mathrm{L}^\infty(\Omega)} \\ &\leq C_* C_3 h^{r-d/2} \|u\|_{\mathrm{H}^{r+1}(\Omega)} + \|\nabla \Pi_h u - \nabla u\|_{\mathrm{L}^\infty(\Omega)} \\ &\leq C_* C_3 h^{r-d/2} \|u\|_{\mathrm{H}^{r+1}(\Omega)} + C_5 h^{r-d/2} \|u\|_{\mathrm{H}^{r+1}(\Omega)}. \end{aligned}$$

Since $r > d/2$ by hypothesis, given $\delta > 0$, there exists $h_0 \in (0,1]$ such that, for all $h \in (0, h_0]$,

$$\varphi \in \mathcal{J} \;\Rightarrow\; \nabla \varphi \in \mathcal{Z}_\delta. \tag{5.6}$$

Motivated by the form of (5.2), we define the fixed point mapping $\mathcal{N}$ on $\mathcal{J}$ as follows. Given $\varphi \in \mathcal{J}$, we denote by $\mathcal{N}(\varphi) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ the solution to the following linear variational problem: Find $\mathcal{N}(\varphi) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ such that

$$\tilde{B}(\varphi; \mathcal{N}(\varphi) - \Pi_h u, v) = \tilde{B}(u; u - \Pi_h u, v) \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \tag{5.7}$$

Equivalently, we can restate this as follows: Find $\mathcal{N}(\varphi) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ such that

$$\tilde{B}(\varphi; \mathcal{N}(\varphi), v) = \tilde{B}(u; u - \Pi_h u, v) + \tilde{B}(\varphi; \Pi_h u, v) \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).$$

Since $S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ is a finite-dimensional linear space, the existence and uniqueness of a solution $\mathcal{N}(\varphi) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ to problem (5.7) follows immediately from (5.3).

To prove that $\mathcal{N}$ maps $\mathcal{J}$ into itself, we test (5.7) with $v = \mathcal{N}(\varphi) - \Pi_h u$ and use (5.3) and (5.4) to obtain

$$\begin{aligned} \tilde{M}_1 \|\mathcal{N}(\varphi) - \Pi_h u\|_{1,h}^2 &\leq \tilde{B}(\varphi; \mathcal{N}(\varphi) - \Pi_h u, \mathcal{N}(\varphi) - \Pi_h u) \\ &= \tilde{B}(u; u - \Pi_h u, \mathcal{N}(\varphi) - \Pi_h u) \\ &\leq \tilde{K}_\delta \|u - \Pi_h u\|_{1,h} \|\mathcal{N}(\varphi) - \Pi_h u\|_{1,h}. \end{aligned}$$

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

$$\|\mathcal{N}(\varphi) - \Pi_h u\|_{1,h} \leq \tilde{M}_1^{-1} \tilde{K}_\delta C_6 h^r \|u\|_{\mathrm{H}^{r+1}(\Omega)}.$$

If we define $C_* = \tilde{M}_1^{-1} \tilde{C}_\delta C_6$, then $\mathcal{N}$ indeed maps $\mathcal{J}$ into itself. Note that, while h_0 depends on C_*, the constant C_* does not depend on h_0, and hence this seemingly implicit construction of C_* is correct.

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

$$\begin{aligned} &\tilde{B}(\varphi; \mathcal{N}(\varphi) - \Pi_h u, v) = \tilde{B}(u; u - \Pi_h u, v) \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}) \;\text{ and} \\ &\tilde{B}(\psi; \mathcal{N}(\psi) - \Pi_h u, v) = \tilde{B}(u; u - \Pi_h u, v) \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \end{aligned}$$

Upon subtracting the second line from the first, choosing $v = \mathcal{N}(\varphi) - \mathcal{N}(\psi)$, and using (5.3) and (5.5), we deduce that

$$\begin{aligned}\tilde{M}_1 \|\mathcal{N}(\varphi) - \mathcal{N}(\psi)\|_{1,h}^2 &\le \tilde{B}(\varphi; \mathcal{N}(\varphi) - \mathcal{N}(\psi), \mathcal{N}(\varphi) - \mathcal{N}(\psi)) \\ &= \tilde{B}(\psi; \mathcal{N}(\psi) - \Pi_h u, \mathcal{N}(\varphi) - \mathcal{N}(\psi)) - \tilde{B}(\varphi; \mathcal{N}(\psi) - \Pi_h u, \mathcal{N}(\varphi) - \mathcal{N}(\psi)) \\ &\le \tilde{L}_\delta \|\nabla\psi - \nabla\varphi\|_{\mathrm{L}^\infty(\Omega)} \|\mathcal{N}(\psi) - \Pi_h u\|_{1,h} \|\mathcal{N}(\varphi) - \mathcal{N}(\psi)\|_{1,h}.\end{aligned}$$

Using the first inverse inequality in (2.1), and the fact that $\mathcal{N}(\psi) \in \mathcal{J}$, we have that

$$\|\mathcal{N}(\varphi) - \mathcal{N}(\psi)\|_{1,h} \le \tilde{M}_1^{-1} \tilde{L}_\delta C_3 C_* h^{r-d/2} \|\nabla\varphi - \nabla\psi\|_{\mathrm{L}^2(\Omega)} \le c(h) \|\varphi - \psi\|_{1,h},$$

where $c(h) = \tilde{M}^{-1} \tilde{L}_\delta C_3 C_* h^{r-d/2}$. Since $r > d/2$ by hypothesis **H3**, there exists a positive constant $h_1 \in (0, 1]$ such that $c(h) < 1$. 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

$$\|u_{\mathrm{DG}} - \Pi_h u\|_{1,h} \le C_* h^r \|u\|_{\mathrm{H}^{r+1}(\Omega)}, \qquad d/2 < r \le \min(m, p); \tag{5.8}$$

furthermore, $\nabla u_{\mathrm{DG}} \in \mathcal{Z}_\delta$ 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 $a \le Cb$ 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

$$\|u - \Pi_h u\|_{1,h} \le C_6 h^r \|u\|_{\mathrm{H}^{r+1}(\Omega)}, \qquad d/2 < r \le \min(m, p), \tag{5.9}$$

we deduce from (5.8) and (5.9) via the triangle inequality that, for all $h \in (0, \min(h_0, h_1)]$,

$$\|u - u_{\mathrm{DG}}\|_{1,h} \lesssim h^r \|u\|_{\mathrm{H}^{r+1}(\Omega)}, \qquad d/2 < r \le \min(m, p), \tag{5.10}$$

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

6. The hyperbolic problem. Now consider the hyperbolic problem

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

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

$$u(t, x) = g_{\mathrm{D}}(t, x) \quad \text{on } (0, T] \times \Gamma_{\mathrm{D}} \qquad \text{and} \qquad S(\nabla u(t, x))\nu = g_{\mathrm{N}}(t, x) \quad \text{on } (0, T] \times \Gamma_{\mathrm{N}}. \tag{6.1}$$

Since g_{D} and g_{N} now depend on t, so does the linear functional, which we denote by $\ell(t, \cdot)$ and is otherwise defined as in (2.3).

We refer to [8, 5] for theoretical results concerning the existence of a unique local (in time) solution to (6.1), subject to the given initial conditions, in the special case of a homogeneous Dirichlet boundary condition on Γ.

It will be assumed throughout that

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

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. We shall further suppose that, for all $t \in [0,T]$, $u(t,\cdot)$ satisfies the Gårding inequality (3.2) for some $M_0 \geq 0$ and $M_1 > 0$, both independent of t. If one assumes the uniform monotonicity condition (1.10), then this is always true with $M_0 = 0$. If, on the other hand, one adopts the (considerably weaker) strong Legendre–Hadamard condition (1.12) and $\Gamma_{\mathrm{D}} = \Gamma$, then the Gårding inequality (3.2) holds with $M_1 > 0$ for some $M_0 \geq 0$ which may depend on $u(t)$; however, since $u \in \mathrm{C}^2([0,T] \times \overline{\Omega})$ by the Sobolev embedding theorem, M_0 can be chosen independent of t; cf. Theorem 6.5.1 on p. 253 of Morrey [16].

As in the elliptic case, let $\mathcal{M}_\delta$ be defined by

$$\mathcal{M}_\delta := \mathrm{conv}\big\{\eta \in \mathbb{R}^{d\times d} : \inf_{x\in\Omega, t\in[0,T]} |\eta - \nabla u(t,x)| \leq \delta\big\},$$

and define the constants K_δ and L_δ by the formulas (1.6) and (1.7). The set $\mathcal{Z}_\delta$ is now given by

$$\mathcal{Z}_\delta := \big\{\Phi \in \mathrm{C}_{\mathrm{pw}}(\overline{\Omega})^{d\times d} : \min_{t\in[0,T]} \|\Phi - \nabla u(t)\|_{\mathrm{L}^\infty(\Omega)} \leq \delta\big\}.$$

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

$$(\ddot{u}_{\mathrm{DG}}, v) + B(u_{\mathrm{DG}}, v) + \int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma [\![\dot{u}_{\mathrm{DG}}]\!] \cdot [\![v]\!] \,\mathrm{d}s = \ell(t,v) + \int_{\Gamma_{\mathrm{D}}} \sigma \dot{g}_{\mathrm{DG}} \cdot v \,\mathrm{d}s \tag{6.2}$$

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

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

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

We highlight the presence of the last term on the left-hand side and the second term on the right-hand side of (6.2) 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. In order to highlight this point further, note that, in an energy analysis of the discontinuous Galerkin approximation (2.5) to the elliptic problem (1.2), (1.3), the natural choice of test function is $v = u_{\mathrm{DG}}$, while in the case of (6.2) it is $v = \dot{u}_{\mathrm{DG}}$, which, in turn, motivates the inclusion of the additional terms in (6.2) compared to the elliptic case.

Let $M_0 \geq 0$ be the constant from (3.2). We denote by $W(t) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ the nonlinear projection of $u(t)$ defined by

$$B(W(t),v) + 2M_0(W(t),v) = B(u(t),v) + 2M_0(u(t),v) \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \qquad 0 \leq t \leq T,$$

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

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

The existence, uniqueness, approximation properties, and differentiability with respect to t of $W(t)$ are established in Appendix A, in Lemma A.1. For the sake of simplicity of presentation, we choose $u_{\mathrm{DG}}^0 = W(0)$ and $u_{\mathrm{DG}}^1 = \dot{W}(0)$ 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 [15], 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.

6.1. Definition of the fixed point map. We decompose

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

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,

$$\begin{aligned}(\ddot{\xi}, v) + B(u_{\mathrm{DG}}, v) - B(W, v) &+ \int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} \sigma [\![\dot{\xi}]\!] \cdot [\![v]\!] \,\mathrm{d}s \\ &= (\ddot{\eta}, v) - 2M_0(\eta, v) + \int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} \sigma [\![\dot{\eta}]\!] \cdot [\![v]\!] \,\mathrm{d}s \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).\end{aligned}$$

Upon linearization of the term $B(u_{\mathrm{DG}}, v) - B(W, v)$, in terms of our earlier notation, we have that

$$\begin{aligned}(\ddot{\xi}, v) + \int_0^1 \tilde{b}(W + \tau(u_{\mathrm{DG}} - W); u_{\mathrm{DG}} - W, v) \,\mathrm{d}\tau &+ \int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} \sigma [\![\dot{\xi}]\!] \cdot [\![v]\!] \,\mathrm{d}s \\ &= (\ddot{\eta}, v) - 2M_0(\eta, v) + \int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} \sigma [\![\dot{\eta}]\!] \cdot [\![v]\!] \,\mathrm{d}s \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}).\end{aligned} \tag{6.3}$$

As in the case of the elliptic problem, we can simplify the notation considerably by defining the bilinear form $\tilde{B}(t, \varphi; \cdot, \cdot)$ by

$$\tilde{B}(t, \varphi; v, w) := \int_0^1 \tilde{b}(W(t) + \tau(\varphi - W(t)); v, w) \,\mathrm{d}t$$

and the linear functional $\rho(t; \cdot)$ by

$$\rho(t; v) := (\ddot{\eta}, v) - 2M_0(\eta, v) + \int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} \sigma [\![\dot{\eta}]\!] \cdot [\![v]\!] \,\mathrm{d}s,$$

which allows us to rewrite (6.3) as

$$(\ddot{\xi}, v) + \tilde{B}(t, u_{\mathrm{DG}}(t); \xi, v) + \int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} \sigma [\![\dot{\xi}]\!] \cdot [\![v]\!] \,\mathrm{d}s = \rho(t; v) \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \tag{6.4}$$

We consider the set $\mathcal{J} \subset \mathrm{C}^1([0, T]; S^p(\Omega, \mathcal{T}_h, \mathbf{F})) \equiv Y$ defined by

$$\begin{aligned}\mathcal{J} := \big\{ \psi \in Y \,:\, &\|\psi - W\|_Y \\ &:= \max_{t \in [0,T]} \Big(\|\psi(t) - W(t)\|_{1,h} + \|\dot{\psi}(t) - \dot{W}(t)\|_{\mathrm{L}^2(\Omega)} \Big) \le C_*(u) h^r \big\},\end{aligned}$$

where $C_*(u)$ is a positive constant and $(d/2) + 1 < r \le \min(m, p)$. As in the elliptic case, by the first inverse inequality in (2.1), there exists $h_0 > 0$ such that, for all $h \in (0, h_0]$,

$$\psi \in \mathcal{J} \ \Rightarrow\ \nabla \psi(t) \in \mathcal{Z}_\delta \ \forall\, t \in [0, T]. \tag{6.5}$$

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

Now, motivated by the form of (6.4) and the definition of ξ, similarly as in the case of the elliptic problem, we are led to the following definition of the fixed point map $\mathcal{N}$ on $\mathcal{J}$: If $\varphi \in \mathcal{J}$, the image $u_\varphi = \mathcal{N}(\varphi) \in \mathrm{C}^2([0,T]; S^p(\Omega, \mathcal{T}_h, \mathbf{F}))$ is defined as the solution to the following linear problem:

$$(\ddot{u}_\varphi - \ddot{W}, v) + \tilde{B}(t, \varphi(t); u_\varphi - W, v) + \int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} \sigma [\![\dot{u}_\varphi - \dot{W}]\!] \cdot [\![v]\!] \, \mathrm{d}s = \rho(t; v) \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \tag{6.6}$$

with $u_\varphi(0) = u^0_{\mathrm{DG}}$, $\dot{u}_\varphi(0) = u^1_{\mathrm{DG}}$. Clearly, this variational form can be rewritten as an explicit linear ordinary differential equation for u_φ, and hence $\mathcal{N}$ is well-defined. Our objective now is to show, via Banach's fixed point theorem, that the nonlinear mapping $\varphi \in \mathcal{J} \mapsto \mathcal{N}(\varphi)$ has a unique fixed point $u_{\mathrm{DG}} \in \mathcal{J}$.

6.2. Auxiliary results. In the analysis of the linear problem (6.6), it will be crucial to replace a term of the form

$$\tilde{B}(t, \varphi(t); \xi(t), \dot{\xi}(t))$$

by a total derivative. Since $\tilde{B}(t, \varphi(t); \cdot, \cdot)$ is not symmetric in its last two arguments, we split $\tilde{B}$ into a symmetric term and a remainder which can be controlled:

$$\tilde{B}(t, \varphi(t); v, w) = \tilde{B}^{(S)}(t, \varphi(t); v, w) + \tilde{B}^{(A)}(t, \varphi(t); v, w), \tag{6.7}$$

where

$$\tilde{B}^{(S)}(t, \varphi(t); v, w) := \int_0^1 \int_\Omega \sum_{i,\alpha,j,\beta=1}^{d} A^\tau_{i\alpha j\beta} \partial_{x_\alpha} w_i \partial_{x_\beta} v_j \, \mathrm{d}x \, \mathrm{d}\tau + \int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} \sigma [\![v]\!] \cdot [\![w]\!] \, \mathrm{d}s,$$

and

$$\tilde{B}^{(A)}(t, \varphi(t); v, w) := -\int_0^1 \sum_{i,\alpha,j,\beta=1}^{d} \left[\int_{\Gamma_{\mathrm{int}}} \left\langle A^\tau_{i\alpha j\beta} \nu_\alpha \partial_{x_\beta} v_j \right\rangle [\![w_i]\!] \, \mathrm{d}s + \int_{\Gamma_{\mathrm{D}}} A^\tau_{i\alpha j\beta} \nu_\alpha w_i \partial_{x_\beta} v_j \, \mathrm{d}s \right] \mathrm{d}\tau,$$

where $A^\tau_{i\alpha j\beta} := A_{i\alpha j\beta}(\nabla W(t) + \tau(\nabla\varphi(t) - \nabla W(t)))$. Note that $\tilde{B}^{(A)}(t, \varphi; \cdot, \cdot)$ is not skew-symmetric but asymmetric, i.e., simply, *not symmetric*.

Following the proof of Lemma 4.1 closely, we obtain for all $\varphi \in \mathcal{J}$ and for all $\alpha \ge \alpha_0$, where α_0 is as in Lemma 4.1,

$$\tilde{B}^{(S)}(t, \varphi(t); v, v) \ \ge \ \tfrac{1}{2} M_1 \|v\|^2_{1,h} - 2M_0 \|v\|^2_{\mathrm{L}^2(\Omega)} \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \tag{6.8}$$

and

$$\left| \tilde{B}^{(A)}(t, \varphi(t); v, w) \right| \ \lesssim \ \|\nabla v\|_{\mathrm{L}^2(\Omega)} \left(\int_{\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}}} \sigma |[\![w]\!]|^2 \, \mathrm{d}s \right)^{1/2} \qquad \forall v, w \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \tag{6.9}$$

In addition, we shall require an estimate on the expression

$$\tilde{B}^{(S)}_t(t, \varphi(t); v, w) := \int_0^1 \int_\Omega \sum_{i,\alpha,j,\beta=1}^{d} \Big[\frac{\mathrm{d}}{\mathrm{d}t} A^\tau_{i\alpha j\beta} \Big] \partial_{x_\alpha} w_i \partial_{x_\beta} v_j \, \mathrm{d}x \, \mathrm{d}\tau, \qquad v, w \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}),$$

where $A^\tau_{i\alpha j\beta}$ is as defined above. Upon setting

$$K'_\delta := \operatorname*{ess.sup}_{x\in\Omega,\tau\in[0,1]} \left(\sum_{i,\alpha,j,\beta=1}^{d} \left| \frac{\mathrm{d}}{\mathrm{d}t} A^\tau_{i\alpha j\beta} \right|^2 \right)^{1/2},$$

we deduce that

$$\left| \tilde{B}^{(S)}_t(t,\varphi(t);v,w) \right| \le K'_\delta \|\nabla v\|_{\mathrm{L}^2(\Omega)} \|\nabla w\|_{\mathrm{L}^2(\Omega)}.$$

To estimate K'_δ, consider

$$\begin{aligned} K'_\delta &= \operatorname*{ess.sup}_{x\in\Omega,\tau\in[0,1]} \left(\sum_{i,\alpha,j,\beta=1}^{d} \left| \nabla A_{i\alpha j\beta}(\nabla\psi(t,x)) \right|^2 |\nabla\dot{W}(t) + \tau(\nabla\dot{\varphi}(t) - \nabla\dot{W}(t))|^2 \right)^{1/2} \\ &\le L_\delta \big(\|\nabla\dot{W}(t)\|_{\mathrm{L}^\infty(\Omega)} + \|\nabla\dot{\varphi}(t) - \nabla\dot{W}(t)\|_{\mathrm{L}^\infty(\Omega)} \big). \end{aligned}$$

As $\varphi \in \mathcal{J}$, the first of the inverse inequalities (2.1), the bound (A.2), and the definition of the set $\mathcal{J}$ yield

$$K'_\delta \lesssim \|\nabla\dot{W}(t)\|_{\mathrm{L}^\infty(\Omega)} + h^{-d/2}\|\nabla\dot{\varphi}(t) - \nabla\dot{W}(t)\|_{\mathrm{L}^2(\Omega)} \lesssim 1 + h^{r-d/2}.$$

Combining these estimates and recalling that, by hypothesis $r > (d/2) + 1$ and, a fortiori, $r > d/2$, we obtain for all $\varphi \in \mathcal{J}$, for all $t \in [0,T]$, and for all $h \in (0,1]$

$$\left| \tilde{B}^{(S)}_t(t,\varphi(t);v,w) \right| \lesssim \|\nabla v\|_{\mathrm{L}^2(\Omega)} \|\nabla w\|_{\mathrm{L}^2(\Omega)} \qquad \forall v,w \in S^p(\Omega,\mathcal{T}_h,\mathbf{F}). \tag{6.10}$$

Finally, we shall require an estimate on the right-hand side $\rho(t;v)$ in (6.4). A straightforward computation gives

$$\begin{aligned} |\rho(t,v)| &\le \left(2\|\ddot{\eta}\|^2_{\mathrm{L}^2(\Omega)} + 8M_0^2\|\eta\|^2_{\mathrm{L}^2(\Omega)} + \int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma|[\![\dot{\eta}]\!]|^2\,\mathrm{d}s \right)^{1/2} \\ &\qquad \left(\|v\|^2_{\mathrm{L}^2(\Omega)} + \int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma|[\![v]\!]|^2\,\mathrm{d}s \right)^{1/2} \\ &\lesssim h^r \left(\|v\|^2_{\mathrm{L}^2(\Omega)} + \int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma|[\![v]\!]|^2\,\mathrm{d}s \right)^{1/2}, \end{aligned} \tag{6.11}$$

where we used (A.2) and (A.7) to bound the different norms of η.

6.3. Convergence analysis. For the sake of notational simplicity, we define

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

Testing (6.6) with $v = \dot{\xi}_\varphi$, and using the decomposition (6.7), we deduce that

$$(\ddot{\xi}_\varphi, \dot{\xi}_\varphi) + \tilde{B}^{(S)}(t,\varphi(t);\xi_\varphi,\dot{\xi}_\varphi) + \int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma|[\![\dot{\xi}_\varphi]\!]|^2\,\mathrm{d}s \;=\; \rho(t;\dot{\xi}_\varphi) - \tilde{B}^{(A)}(t,\varphi(t);\xi_\varphi,\dot{\xi}_\varphi),$$

which can be rewritten as

$$\begin{aligned} &\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\Big[\|\dot{\xi}_\varphi\|^2_{\mathrm{L}^2(\Omega)} + \tilde{B}^{(S)}(t,\varphi(t);\xi_\varphi,\xi_\varphi) \Big] + \int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma|[\![\dot{\xi}_\varphi]\!]|^2\,\mathrm{d}s \\ &\qquad = \rho(t;\dot{\xi}_\varphi) - \tilde{B}^{(A)}(t,\varphi(t);\xi_\varphi,\dot{\xi}_\varphi) - \frac{1}{2}\tilde{B}^{(S)}_t(t,\varphi;\xi_\varphi,\xi_\varphi). \end{aligned} \tag{6.12}$$

On noting that $\xi_\varphi(0) = 0$ and $\dot{\xi}_\varphi(0) = 0$, integrating the above identity in t, and multiplying by 2, we deduce from (6.8) that, for $\alpha \geq \alpha_0$ and $h \in (0, h_0]$,

$$\|\dot{\xi}_\varphi(t)\|^2_{\mathrm{L}^2(\Omega)} + \tfrac{1}{2} M_1 \|\xi_\varphi(t)\|^2_{1,h} - 2M_0 \|\xi_\varphi(t)\|^2_{\mathrm{L}^2(\Omega)} + 2\int_0^t \int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma |[\![\dot{\xi}_\varphi(\tau)]\!]|^2 \,\mathrm{d}s\,\mathrm{d}\tau$$
$$\leq \int_0^t \Big[2|\rho(\tau;\dot{\xi}_\varphi(\tau))| + 2\big|\tilde{B}^{(A)}(\tau,\varphi(\tau);\xi_\varphi(\tau),\dot{\xi}_\varphi(\tau))\big| + \big|\tilde{B}_t^{(S)}(\tau;\varphi(\tau);\xi_\varphi,\xi_\varphi)\big| \Big]\,\mathrm{d}\tau. \tag{6.13}$$

Next we estimate the terms on the right-hand side, using (6.11), (6.9), and (6.10). Transferring the term $2M_0\|\xi_\varphi(t)\|^2_{\mathrm{L}^2(\Omega)}$ to the right-hand side, we obtain

$$\|\dot{\xi}_\varphi(t)\|^2_{\mathrm{L}^2(\Omega)} + \tfrac{1}{2} M_1 \|\xi_\varphi(t)\|^2_{1,h} + 2\int_0^t \int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma |[\![\dot{\xi}_\varphi(\tau)]\!]|^2 \,\mathrm{d}s\,\mathrm{d}\tau$$
$$\lesssim \|\xi_\varphi(t)\|^2_{\mathrm{L}^2(\Omega)} + h^r \int_0^t \Big[\|\dot{\xi}_\varphi(\tau)\|^2_{\mathrm{L}^2(\Omega)} + \int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma |[\![\dot{\xi}_\varphi(\tau)]\!]|^2 \,\mathrm{d}s \Big]^{1/2} \mathrm{d}\tau$$
$$+ \int_0^t \|\nabla \xi_\varphi(\tau)\|_{\mathrm{L}^2(\Omega)} \Big[\int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma |[\![\dot{\xi}_\varphi(\tau)]\!]|^2 \,\mathrm{d}s \Big]^{1/2} \mathrm{d}\tau + \int_0^t \|\nabla \xi_\varphi(\tau)\|^2_{\mathrm{L}^2(\Omega)} \,\mathrm{d}\tau. \tag{6.14}$$

Using $\xi_\varphi(0) = 0$, the first term on the right-hand side can be estimated by

$$\|\xi_\varphi(t)\|^2_{\mathrm{L}^2(\Omega)} = \left\| \int_0^t \dot{\xi}_\varphi(\tau)\,\mathrm{d}\tau \right\|^2_{\mathrm{L}^2(\Omega)} \leq T \int_0^t \|\dot{\xi}_\varphi(\tau)\|^2_{\mathrm{L}^2(\Omega)} \,\mathrm{d}\tau.$$

Terms containing integrals over $[0,t] \times (\Gamma_{\mathrm{int}} \cup \Gamma_{\mathrm{D}})$ in (6.14) can be absorbed into the third term on the left-hand side of (6.14) by apply the ε-inequality with sufficiently small ε (but independent of h). After normalization, we obtain

$$\|\dot{\xi}_\varphi(t)\|^2_{\mathrm{L}^2(\Omega)} + \|\xi_\varphi(t)\|^2_{1,h} + \int_0^t \int_{\Gamma_{\mathrm{int}}\cup\Gamma_{\mathrm{D}}} \sigma |[\![\dot{\xi}_\varphi(\tau)]\!]|^2 \,\mathrm{d}s\,\mathrm{d}\tau$$
$$\lesssim h^{2r} + \int_0^t \Big[\|\dot{\xi}_\varphi(\tau)\|^2_{\mathrm{L}^2(\Omega)} + \|\nabla \xi_\varphi(\tau)\|^2_{\mathrm{L}^2(\Omega)} \Big] \mathrm{d}\tau. \tag{6.15}$$

Hence, an application of Gronwall's lemma gives

$$\max_{t\in[0,T]} \|\mathcal{N}(\varphi)(t) - W(t)\|_Y \lesssim h^r,$$

which allows us to deduce the existence of a constant $C_* = C_*(u)$, independent of h, such that, for $h \leq h_0$, $\mathcal{N}$ maps $\mathcal{J}$ into itself.

Remark 1. 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 quasi-linear 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_* featured in the definition of the set $\mathcal{J}$, here, for the sake of brevity, we 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. Setting $\xi_\varphi = u_\varphi - W$ and $\xi_\psi = u_\psi - W$, we have that

$$(\ddot{\xi}_\varphi, v) + \tilde{B}(t, \varphi; \xi_\varphi, v) + \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma [\![\dot{\xi}_\varphi]\!] \cdot [\![v]\!] \,\mathrm{d}s = \rho(t; v), \quad \text{and}$$
$$(\ddot{\xi}_\psi, v) + \tilde{B}(t, \psi; \xi_\psi, v) + \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma [\![\dot{\xi}_\psi]\!] \cdot [\![v]\!] \,\mathrm{d}s = \rho(t; v) \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}),$$

subject to $\xi_\varphi(0) = \xi_\psi(0) = 0$ and $\dot{\xi}_\varphi(0) = \dot{\xi}_\psi(0) = 0$. By subtracting the second line from the first line, and testing with

$$v = \dot{\xi}_\varphi - \dot{\xi}_\psi = \dot{u}_\varphi - \dot{u}_\psi \equiv \dot{e},$$

where $e = u_\varphi - u_\psi$, we obtain

$$(\ddot{e}, \dot{e}) + \tilde{B}(t, \varphi; e, \dot{e}) + \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma |[\![\dot{e}]\!]|^2 \,\mathrm{d}s = \tilde{B}(t, \psi; \xi_\psi, \dot{e}) - \tilde{B}(t, \varphi; \xi_\psi, \dot{e}).$$

By virtue of Lemma 4.3,

$$\big|\tilde{B}(t, \psi; \xi_\psi, \dot{e}) - \tilde{B}(t, \varphi; \xi_\psi, \dot{e})\big| \lesssim \|\nabla\varphi - \nabla\psi\|_{\mathrm{L}^\infty(\Omega)} \|\xi_\psi\|_{1,h} \|\dot{e}\|_{1,h}.$$

Thus, by using the same procedure as in the proof of the inclusion $\mathcal{N}(\mathcal{J}) \subset \mathcal{J}$, we obtain

$$\begin{aligned}
&\|\dot{e}(t)\|^2_{\mathrm{L}^2(\Omega)} + \|e(t)\|^2_{1,h} + \int_0^t \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma |[\![\dot{e}(\tau)]\!]|^2 \,\mathrm{d}s \,\mathrm{d}\tau \\
&\quad\lesssim \int_0^t \Big[\|\dot{e}(\tau)\|^2_{\mathrm{L}^2(\Omega)} + \|\nabla e(\tau)\|^2_{\mathrm{L}^2(\Omega)}\Big] \,\mathrm{d}\tau \\
&\qquad + \int_0^t \|\nabla\varphi(\tau) - \nabla\psi(\tau)\|_{\mathrm{L}^\infty(\Omega)} \|\xi_\psi(\tau)\|_{1,h} \|\dot{e}(\tau)\|_{1,h} \,\mathrm{d}\tau.
\end{aligned} \tag{6.16}$$

As $u_\psi \in \mathcal{J}$, we have $\max_{t\in[0,T]} \|\xi_\psi(t)\|_{1,h} \le C_* h^r$, and, by the first inequality in (2.1), we also have that

$$\|\nabla\varphi(\tau) - \nabla\psi(\tau)\|_{\mathrm{L}^\infty(\Omega)} \lesssim h^{-d/2} \|\nabla\varphi(\tau) - \nabla\psi(\tau)\|_{\mathrm{L}^2(\Omega)}.$$

The only term on the right-hand side of (6.16) which cannot be directly controlled by any of the terms featured on the left-hand side of (6.16) is $\|\dot{e}(\tau)\|_{1,h}$. Employing the second inverse inequality in (2.1), we handle this term as follows:

$$\begin{aligned}
\|\dot{e}(\tau)\|^2_{1,h} &= \|\nabla\dot{e}(\tau)\|^2_{\mathrm{L}^2(\Omega)} + \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma |[\![\dot{e}(\tau)]\!]|^2 \,\mathrm{d}s \\
&\lesssim h^{-2} \|\dot{e}(\tau)\|^2_{\mathrm{L}^2(\Omega)} + \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma |[\![\dot{e}(\tau)]\!]|^2 \,\mathrm{d}s.
\end{aligned}$$

Inserting these bounds into (6.16), we obtain

$$\begin{aligned}
&\|\dot{e}(t)\|^2_{\mathrm{L}^2(\Omega)} + \|e(t)\|^2_{1,h} + \int_0^t \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma |[\![\dot{e}(\tau)]\!]|^2 \,\mathrm{d}s \,\mathrm{d}\tau \\
&\quad\lesssim \int_0^t \Big[\|\dot{e}(\tau)\|^2_{\mathrm{L}^2(\Omega)} + \|e(\tau)\|^2_{1,h}\Big] \,\mathrm{d}\tau + h^{r-d/2-1} \int_0^t \|\nabla\varphi(\tau) - \nabla\psi(\tau)\|_{\mathrm{L}^2(\Omega)} \\
&\quad\times \bigg[\|\dot{e}(\tau)\|^2_{\mathrm{L}^2(\Omega)} + h \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma |[\![\dot{e}(\tau)]\!]|^2 \,\mathrm{d}s\bigg]^{1/2} \,\mathrm{d}\tau.
\end{aligned}$$

Thus, by applying to the two last terms on the right-hand side of the ε-inequality $ab \le \frac{\varepsilon}{2}a^2 + \frac{1}{2\varepsilon}b^2$, with $\varepsilon > 0$ sufficiently small, we deduce from Gronwall's lemma that

$$\begin{aligned}\|\dot u_\varphi(t) - \dot u_\psi(t)\|^2_{\mathrm{L}^2(\Omega)} + \|u_\varphi(t) - u_\psi(t)\|^2_{1,h} &\lesssim h^{2(r-d/2-1)} \int_0^t \|\nabla\varphi(\tau) - \nabla\psi(\tau)\|^2_{\mathrm{L}^2(\Omega)}\,\mathrm{d}\tau \\ &\lesssim h^{2(r-d/2-1)}\|\varphi - \psi\|^2_Y,\end{aligned}$$

and thereby

$$\|\mathcal{N}(\varphi) - \mathcal{N}(\psi)\|_Y \lesssim h^{r-d/2-1}\|\varphi - \psi\|_Y \qquad \forall \varphi, \psi \in \mathcal{J},$$

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, $u_{\mathrm{DG}} \in \mathcal{J}$, the semidiscrete discontinuous Galerkin finite element approximation to u defined by (6.2). In other words, for h sufficiently small,

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

Combining the last bound with (A.1) and (A.7) we then deduce, for h sufficiently small, that

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

which is the desired optimal convergence estimate.

7. Extensions to other methods. It is straightforward to extend our error analysis to different discontinuous finite element methods. Note, for example, that in the elliptic case only Lemmas 4.1–4.3 are method-dependent. Once they are established, the remaining analysis is independent of the particular form of discretization used. We shall demonstrate this through the example of the discontinuous Galerkin finite element method (DGFEM) of Eyck and Lew [10], which is a particularly attractive candidate for variational problems since it is defined via a discrete energy principle.

The idea is to use the lifting operator introduced in [2] to find a gradient representation for the jumps across element interfaces to define a *discontinuous Galerkin (DG) gradient operator*. More precisely, for $v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$, let

$$\nabla_{\mathrm{DG}} v = \nabla v + \mathbf{R}(v),$$

where $\mathbf{R} : S^p(\Omega, \mathcal{T}_h, \mathbf{F}) \to \mathrm{C}_{\mathrm{pw}}(\overline{\Omega})^{d\times d}$ is defined by

$$\int_\Omega \mathbf{R}(v) : F \,\mathrm{d}x = -\int_{\Gamma_{\mathrm{int}}} [\![v]\!] \cdot \langle F\,\nu_{\mathrm{int}}\rangle \,\mathrm{d}s \qquad \forall F \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})^d.$$

We shall also use $\nabla^{i\alpha}_{\mathrm{DG}}$ to denote the (i,α) component of ∇_{DG}. It is straightforward to show that $\mathbf{R}$ is a bounded operator; more precisely,

$$\|\mathbf{R}(v)\|_{\mathrm{L}^2(\Omega)} \le C_L \left(\int_{\Gamma_{\mathrm{int}}} \sigma |[\![v]\!]|^2 \,\mathrm{d}x\right)^{1/2} \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}), \tag{7.1}$$

where C_L is independent of h. Using the definition of the DG gradient, we define the discrete functional $J_h : S^p(\Omega, \mathcal{T}_h, \mathbf{F}) \to \mathbb{R}$ by

$$J_h(v) = \int_\Omega \left[W(\nabla_{\rm DG} v) - f\cdot v\right]\,{\rm d}x - \int_{\Gamma_{\rm N}} g_{\rm N}\cdot v\,{\rm d}s + \int_{\Gamma_{\rm int}} \sigma|[\![v]\!]|^2\,{\rm d}s + \int_{\Gamma_{\rm D}} \sigma|v-g_{\rm D}|^2\,{\rm d}s,$$

as an approximation to the functional J defined in (1.1). The resulting DGFEM for (1.2) is simply the Euler–Lagrange equation $\delta J_h(u_{\rm DG}) = 0$, where δJ_h, the first variation of J_h, is given by

$$\delta J_h(\varphi; v) = \int_\Omega \sum_{i,\alpha=1}^d S_{i\alpha}(\nabla_{\rm DG}\varphi)\nabla_{\rm DG}^{i,\alpha} v\,{\rm d}x + \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma[\![\varphi]\!]\cdot[\![v]\!]\,{\rm d}s - \ell(v),$$

where $\ell(v)$ is defined as in (2.3). Since $\mathbf{R}(u) = 0$ if u is continuous on $\overline{\Omega}$, the method is consistent. Similarly, the second variation of J_h is defined by

$$\delta^2 J_h(\varphi; v, w) = \int_\Omega \sum_{i,\alpha,j,\beta=1}^d A_{i\alpha j\beta}(\nabla_{\rm DG}\varphi)\nabla_{\rm DG}^{i,\alpha} v\ \nabla_{\rm DG}^{j,\beta} w\,{\rm d}x + \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma[\![v]\!]\cdot[\![w]\!]\,{\rm d}s.$$

Suppose that $u \in {\rm C}^1(\overline{\Omega})$ satisfies (3.2). While Lemma 3.2 cannot be applied directly, it is nevertheless straightforward to modify its proof to obtain for $h \le 1$, $\alpha \ge \alpha_0 = \alpha_0(C_r, K_\delta, M_1, M_0)$, and for all $\varphi \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ such that $\|\nabla_{\rm DG}\varphi - \nabla u\|_{{\rm L}^\infty(\Omega)} \le \delta \le M_1/(4L_\delta)$

$$\delta^2 J_h(\varphi; v, v) \ge \tfrac{1}{2} M_1 \|v\|_{1,h}^2 - 2M_0\|v\|_{{\rm L}^2(\Omega)}^2 \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \tag{7.2}$$

The boundedness and Lipschitz continuity of $\varphi \mapsto \delta^2 J_h(\varphi; \cdot, \cdot)$ over the set of all φ such that $\nabla_{\rm DG}\varphi \in \mathcal{M}_\delta$ can be obtained precisely as in Lemma 4.2 and 4.3. Using (7.1) we can again deduce that for $h \le h_0$

$$\varphi \in \mathcal{J} \ \Rightarrow\ \|\nabla_{\rm DG}\varphi - \nabla u\|_{{\rm L}^\infty(\Omega)} \le \delta,$$

and thus, the convergence analysis of section 5 can be repeated verbatim to obtain the existence of a solution $u_{\rm DG}$ to $\delta J_h(u_{\rm DG}; v) = 0$ for all $v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$, satisfying the optimal-order error estimate (5.10).

The analysis in the hyperbolic case can be generalized just as easily. The DGFEM based on the energy principle outlined above reads: For $t \in (0, T]$ find $u_{\rm DG}(t, \cdot) \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$ such that

$$(\ddot{u}_{\rm DG}, v) + \delta J_h(u_{\rm DG}; v) + \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma[\![\dot{u}_{\rm DG}]\!]\cdot[\![v]\!]\,{\rm d}s = \int_{\Gamma_{\rm D}} \sigma\dot{g}_{\rm D}\cdot v\,{\rm d}s \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \tag{7.3}$$

Upon defining

$$\tilde{B}(t, \varphi(t); v, w) = \int_0^1 \delta^2 J_h(W(t) + \tau(\varphi(t) - W(t)); v, w)\,{\rm d}\tau,$$

the analysis proceeds almost exactly as in section 6. The only difference now is that, since $\tilde{B}(t, \varphi(t); \cdot, \cdot)$ is symmetric, we do not have to split it into a symmetric and an asymmetric part. Instead of (6.12), we will obtain

$$\frac{1}{2}\frac{\rm d}{{\rm d}t}\left[\|\dot{\xi}_\varphi\|_{{\rm L}^2}^2 + \tilde{B}(t, \varphi(t); \xi_\varphi, \xi_\varphi)\right] + \int_{\Gamma_{\rm int}\cup\Gamma_{\rm D}} \sigma|[\![\dot{\xi}_\varphi]\!]|^2\,{\rm d}s = \rho(t; \dot{\xi}_\varphi) - \frac{1}{2}\tilde{B}_t(t, \varphi(t); \xi_\varphi, \xi_\varphi).$$

From the Lipschitz continuity of $\varphi \mapsto \delta^2 J_h(\varphi;\cdot,\cdot)$ on $\mathcal{J}$, we immediately obtain the bound on $\tilde{B}_t$ equivalent to (6.10), and we can thus proceed as in section 6.3 to prove the existence of a solution to (7.3) and an optimal error bound, identical to the one we had previously established.

8. Conclusions. We derived optimal-order convergence estimates in the broken H^1 norm for discontinuous Galerkin finite element approximations to second-order quasi-linear 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. These technical restrictions were also present in the work of Makridakis [15], whose techniques we have employed here. They occur, since we have used the inverse estimate (2.1) in order to obtain L^∞ bounds for elements of the set $\mathcal{J}$ defined, respectively, in sections 5 and 6.1, which in turn are required to obtain the uniform Gårding inequality of Lemma 3.2. However, we have reason to believe that the methods considered remain optimally convergent in the energy norm in these excluded 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 [12]). The same statement would also follow immediately if one could prove directly, without involving the first inverse inequality in (2.1), that ∇u_{DG} is sufficiently close to ∇u in the L^∞-norm.

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 or uniformly monotone (cf. (1.10)); instead, we assumed only local Lipschitz continuity of S and the Gårding inequality (3.2).

The main body of the paper was devoted to an analysis of the incomplete interior penalty method [9, 18]. However, we have demonstrated in section 7, where we showed how to extend all results to the variational DGFEM of Eyck and Lew [10], that the framework which we had developed should apply to virtually any discontinuous Galerkin discretization of the quasi-linear elliptic and hyperbolic equations considered. The crucial step is a proof of the coercivity estimate (5.3), using (a variation of) the broken Gårding inequality, stated in Lemma 3.2.

We note that all of our results can be straightforwardly extended to quasi-linear 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 quasi-linear elliptic and hyperbolic systems, we did not discuss fully discrete discontinuous Galerkin finite element approximations of quasi-linear 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 [15], for example, for the corresponding analysis in the case of spatially H^1_0-conforming finite element methods which may serve as a starting point for further analytical considerations in that direction.

Appendix A. 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.

Lemma A.1. *Let* $u \in \mathrm{C}^2([0,T];\mathrm{H}^{m+1}(\Omega)^d)$, $m > d/2+1$, *satisfy* (3.2) *with constants* $M_1 > 0$ *and* $M_0 \ge 0$ *which are independent of* t. *Suppose also that the family* $\{\mathcal{T}_h\}_{h>0}$ *satisfies* **H1**–**H4** *of section* 2. *Then there exists* $h_0 > 0$ *such that for* $h \le h_0$ *there exists a solution* $W(t) \in S^p(\Omega,\mathcal{T}_h,\mathbf{F})$ *to the nonlinear equation*

$$B(W(t);v) + 2M_0(W(t),v) = B(u(t);v) + 2M_0(u(t),v) \qquad \forall v \in S^p(\Omega,\mathcal{T}_h,\mathbf{F}).$$

Furthermore, $t \mapsto W(t)$ *is twice differentiable in* $[0,T]$ *and satisfies*

$$\|u(t) - W(t)\|_{1,h} \le C_p h^r, \tag{A.1}$$

$$\|\dot u(t) - \dot W(t)\|_{1,h} \le C_p'' h^r, \quad \text{and} \tag{A.2}$$

$$\|\ddot u(t) - \ddot W(t)\|_{1,h} \le C_p'' h^r, \tag{A.3}$$

where C_p, C_p', *and* C_p'' *are constants independent of* h *and* t.

We skip the proof of existence of $W(t)$ and of the bound (A.1) which can be established by identical arguments to those in section 5 (see [17] for details). The proofs of (A.2) and (A.3) are given in the following two sections.

A.1. Bounds on $\dot{u} - \dot{W}$. Having established the existence of the nonlinear projection $W(t)$ of $u(t)$ for $t \in [0,T]$, we next prove 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 $V \mapsto B(U,V) - B(u(t),V) + 2M_0(U-u(t),V)$ 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

$$(\mathcal{B}(t,U),V) = B(U,V) - B(u(t),V) + 2M_0(U-u(t),V) \qquad \forall V \in S^p(\Omega,\mathcal{T}_h,\mathbf{F}).$$

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 linearization process in section 4 and from Lemma 4.1 that the derivative of $(t,U) \mapsto \mathcal{B}(t,U)$ with respect to U, evaluated at $U = W(t)$, 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 neighborhood 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 $(0,T)$.

Set

$$u(t) - W(t) = (u(t) - \Pi_h u(t)) - (W(t) - \Pi_h u(t)) \equiv \eta - \xi.$$

We begin by noting that, according to the definition of $W(t)$,

$$\tilde B(t,W(t);\xi,v) + 2M_0(\xi,v) = \tilde B(t,u(t);\eta,v) + 2M_0(\eta,v) \qquad \forall v \in S^p(\Omega,\mathcal{T}_h,\mathbf{F}),$$

where

$$\tilde B(t,\varphi;v,w) = \int_0^t \tilde b(\Pi_h u(t) + \tau(\varphi - \Pi_h u(t));v,w)\,\mathrm{d}\tau.$$

After differentiation with respect to t, we obtain

$$\begin{aligned} \tilde B(t,W(t);\dot\xi(t),v) + 2M_0(\dot\xi(t),v) &= \tilde B(t,u(t);\dot\eta(t),v) - \tilde B_t(t,W(t);\xi(t),v) \\ &\quad + \tilde B_t(t,u(t);\eta(t),v), \end{aligned} \tag{A.4}$$

where

$$\tilde{B}_t(t,\varphi(t);v,w) = \frac{\mathrm{d}}{\mathrm{d}t}\tilde{B}(t,\varphi(t);v,w) = \int_0^1 \int_\Omega \sum_{i,\alpha,j,\beta=1}^{d} \Big[\frac{\mathrm{d}}{\mathrm{d}t} A^\tau_{i\alpha j\beta}\Big] \partial_{x_\alpha} w_i \partial_{x_\beta} v_j \,\mathrm{d}x\,\mathrm{d}\tau$$

for $v, w \in S^p(\Omega, \mathcal{T}_h, \mathbf{F})$, and $A^\tau_{i\alpha j\beta}$ is as before.

Arguing as in the proof of the bound (6.10), we obtain

$$\big|\tilde{B}_t(t,\varphi(t);v,w)\big| \lesssim \big(\|\nabla\Pi_h \dot{u}(t)\|_{\mathrm{L}^\infty(\Omega)} + \|\nabla\dot{\varphi}(t) - \nabla\Pi_h\dot{u}(t)\|_{\mathrm{L}^\infty(\Omega)}\big)\|v\|_{1,h}\|w\|_{1,h} \tag{A.5}$$

for all $\varphi \in \mathcal{Z}_\delta$. We note that

$$\|\nabla\Pi_h\dot{u}(t)\|_{\mathrm{L}^\infty(\Omega)} \le \|\nabla\dot{u}(t)\|_{\mathrm{L}^\infty(\Omega)} + C_5 h^{r-d/2}\|\dot{u}\|_{\mathrm{H}^{r+1}(\Omega)} \lesssim 1,$$

where, in the last inequality, we made use of hypothesis **H3** whereby $r > (d/2)+1$, and, a fortiori, $r > d/2$.

In order to bound the last two terms in (A.4) we shall need to consider two specific choices of φ in (A.5): $\varphi = W$ and $\varphi = u$. For the case of $\varphi = W$ in (A.5), we shall use the following bound, which results on applying the first inverse inequality in (2.1):

$$\begin{aligned}\|\nabla\dot{W}(t) - \nabla\Pi_h\dot{u}(t)\|_{\mathrm{L}^\infty(\Omega)} &\lesssim h^{-d/2}\|\nabla\dot{W}(t) - \nabla\Pi_h\dot{u}(t)\|_{\mathrm{L}^2(\Omega)}\\ &\lesssim h^{-d/2}\|\dot{W}(t) - \Pi_h\dot{u}(t)\|_{1,h} \approx h^{-d/2}\|\dot{\xi}\|_{1,h}.\end{aligned}$$

On the other hand, for the case of $\varphi = u$, we shall use the bound

$$\|\nabla\dot{u}(t) - \nabla\Pi_h\dot{u}(t)\|_{\mathrm{L}^\infty(\Omega)} \lesssim C_5 h^{r-d/2}\|\dot{u}(t)\|_{\mathrm{H}^1(\Omega)}.$$

Thus, we obtain

$$\begin{aligned}\big|\tilde{B}_t(t,W(t);\xi,v)\big| &\lesssim \Big(1 + h^{-d/2}\|\dot{\xi}\|_{1,h}\Big)\|\xi\|_{1,h}\|v\|_{1,h} \quad \text{and}\\ \big|\tilde{B}_t(t,u(t);\eta,v)\big| &\lesssim \|\eta\|_{1,h}\|v\|_{1,h}.\end{aligned}$$

Upon testing (A.4) with $v = \dot{\xi}(t)$ and using (5.4) on the first term on its right-hand side, we obtain

$$\begin{aligned}\|\dot{\xi}(t)\|^2_{1,h} &\lesssim \|\dot{\eta}\|_{1,h}\|\dot{\xi}\|_{1,h} + \|\eta\|_{1,h}\|\dot{\xi}\|_{1,h} + \|\xi\|_{1,h}\|\dot{\xi}\|_{1,h} + h^{-d/2}\|\xi\|_{1,h}\|\dot{\xi}\|^2_{1,h}\\ &\lesssim h^r\|\dot{\xi}\|_{1,h} + h^{r-d/2}\|\dot{\xi}\|^2_{1,h},\end{aligned}$$

where we also used the approximation properties of Π_h and estimate (A.1). Since $r > d/2$, there exists $h_2 \in (0, \min(h_0,h_1)]$ such that for $h \in (0,h_2]$ the coefficient of $\|\dot{\xi}\|^2_{1,h}$ on the right-hand side is less than or equal to $\frac{1}{2}$. We can therefore bring this term to the left-hand side and divide by $\frac{1}{2}\|\dot{\xi}\|_{1,h}$ to finally obtain

$$\|\dot{\xi}\|_{1,h} = \|\dot{W}(t) - \Pi_h\dot{u}(t)\|_{1,h} \lesssim h^r$$

for $(d/2)+1 < r \le \min(m,p)$, from which (A.2) follows immediately on invoking the approximation properties of Π_h.

A.2. Bounds on $\ddot{\boldsymbol{\eta}} = \ddot{\boldsymbol{u}} - \ddot{\boldsymbol{W}}$. By proceeding in an identical manner as in the previous section we find that the mapping $t \mapsto \dot{W}(t)$ is differentiable on $(0, T)$, and we get, for $(d/2) + 1 < r \le \min(m, p)$, that

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

Invoking, once again, the approximation properties of Π_h, we deduce from the triangle inequality that, for $(d/2) + 1 < r \le \min(m, p)$, (A.3) holds.

Technically, the only additional step in this argument in comparison with that in the previous section is to establish a bound, similar to (A.5), on the term

$$\tilde{B}_{tt}(t, \varphi(t); v, w) = \frac{\mathrm{d}^2}{\mathrm{d}t^2} \tilde{B}(t, \varphi(t); v, w)$$

for $\varphi \in \mathcal{Z}_\delta$. Here we require a uniform bound on the fourth derivative of W, i.e., on the second derivatives

$$\frac{\partial^2}{\partial \eta_{\gamma k} \partial \eta_{\iota m}} A_{i\alpha j\beta}(\eta)$$

for $\eta \in \mathcal{M}_\delta$ and can otherwise argue similarly as in the proof of (6.10); hence our assumption $W \in \mathrm{C}^4(\mathbb{R}^{d\times d}; \mathbb{R})$ on the regularity of the stored energy function W was adopted in the introductory section of the paper.

A.3. $\mathbf{L^2}$-bounds. Since, by hypothesis **H4**, the family $\{\mathcal{T}_h\}_{h>0}$ is uniformly simplicially reducible (cf. also section 3), the broken Friedrichs inequality (cf. [3]) implies the existence of a positive constant C, independent of h, such that

$$\|v\|^2_{\mathrm{L}^2(\Omega)} \le C \|v\|^2_{1,h} \qquad \forall v \in S^p(\Omega, \mathcal{T}_h, \mathbf{F}). \tag{A.6}$$

Here the constant C depends only on certain shape-regularity properties of the family $\{\mathcal{T}_h\}_{h>0}$, the penalty parameter α, and the Friedrichs constant for $\mathrm{H}^1_{\mathrm{D},0}(\Omega)$.

In fact, (A.6) can also be obtained from Lemma 3.1, in which case the corresponding constant C would depend on the constant C_r, the penalty parameter α, and the Friedrichs constant for $\mathrm{H}^1_{\mathrm{D},0}(\Omega)$.

Either way, on applying (A.6) to (A.1)–(A.3), we obtain

$$\|W(t) - u(t)\|_{\mathrm{L}^2(\Omega)} + \|\dot{W}(t) - u(t)\|_{\mathrm{L}^2(\Omega)} + \|\ddot{W}(t) - \ddot{u}(t)\|_{\mathrm{L}^2(\Omega)} \lesssim h^r. \tag{A.7}$$

While this bound is not optimal (the optimal rate would be $r + 1$ rather than r), it is entirely adequate for the purposes of deriving an optimal bound on $u - u_{\mathrm{DG}}$ in the *energy norm* $\|\cdot\|_Y$.

REFERENCES

[1] S. S. Antman, *Nonlinear Problems of Elasticity*, Appl. Math. Sci. 107, 2nd ed., Springer, New York, 2005.
[2] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2001), pp. 1749–1779.
[3] S. C. Brenner, *Poincaré-Friedrichs inequalities for piecewise H^1 functions*, SIAM J. Numer. Anal., 41 (2003), pp. 306–324.
[4] S. C. Brenner, *Korn's inequalities for piecewise H^1 vector fields*, Math. Comp., 73 (2004), pp. 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), pp. 77–112.

[6] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, in Stud. Math. Appl. 4, North-Holland, Amsterdam, 1978.
[7] B. Cockburn, G. E. Karniadakis, and C.-W. Shu, *The development of discontinuous Galerkin methods*, in Discontinuous Galerkin Methods (Newport, RI, 1999), Lect. Notes Comput. Sci. Eng. 11, Springer, Berlin, 2000, pp. 3–50.
[8] C. M. Dafermos and W. J. Hrusa, *Energy methods for quasilinear hyperbolic initial-boundary value problems. Applications to elastodynamics*, Arch. Ration. Mech. Anal., 87 (1985), pp. 267–292.
[9] C. Dawson, S. Sun, and M. F. Wheeler, *Compatible algorithms for coupled flow and transport*, Comput. Methods Appl. Mech. Engrg., 193 (2004), pp. 2565–2580.
[10] T. Eyck and A. Lew, *Discontinuous Galerkin methods for nonlinear elasticity*, Int. J. Numer. Meth. Engrg., 67 (2006), pp. 1204–1243.
[11] M. E. Gurtin, *An Introduction to Continuum Mechanics*, Math. Sci. Eng. 158, Academic Press, New York, 1981.
[12] P. Houston, J. Robson, and E. Süli, *Discontinuous Galerkin finite element approximation of quasilinear elliptic boundary value problems.* I. *The scalar case*, IMA J. Numer. Anal., 25 (2005), pp. 726–749.
[13] P. Houston, C. Schwab, and E. Süli, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal., 39 (2002), pp. 2133–2163.
[14] O. A. Karakashian and F. Pascal, *A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 2374–2399.
[15] C. G. Makridakis, *Finite element approximations of nonlinear elastic waves*, Math. Comp., 61 (1993), pp. 569–594.
[16] C. B. Morrey, Jr., *Multiple integrals in the calculus of variations*, Grundlehren Math. Wi. 130, Springer, New York, 1966.
[17] C. Ortner and E. Süli, *Discontinuous Galerkin Finite Element Approximation of Nonlinear Second-Order Elliptic and Hyperbolic Systems*, Technical report NA-06/05, Oxford University Computing Laboratory, London, 2006.
[18] S. Sun, *Discontinuous Galerkin Methods for Reactive Transport in Porous Media*, Ph.D. thesis, The University of Texas at Austin, 2003.