A POSTERIORI ERROR ESTIMATES FOR SPACE-TIME FINITE ELEMENT APPROXIMATION OF QUASISTATIC HEREDITARY LINEAR VISCOELASTICITY PROBLEMS*

SIMON SHAW† AND J. R. WHITEMAN†

March 24, 2003

Abstract. We give a space-time Galerkin finite element discretisation of the quasistatic compressible linear viscoelasticity problem as described by an elliptic partial differential equation with a fading memory Volterra integral. The numerical scheme consists of a continuous Galerkin approximation in space based on piecewise polynomials of degree $p > 0$ (cG(p)), with a discontinuous Galerkin piecewise constant (dG(0)) or linear (dG(1)) approximation in time.

A posteriori Galerkin-error estimates are derived by exploiting the Galerkin framework and optimal stability estimates for a related dual backward problem. The *a posteriori* error estimates are quite flexible: strong L_p-energy norms of the errors are estimated using time derivatives of the residual terms when the data are smooth, while weak-energy norms are used when the data are nonsmooth (in time).

We also give upper bounds on the dG(0)cG(1) *a posteriori* error estimates which indicate optimality. However, a complete analysis is not given.

Key words. Viscoelasticity, adaptivity, finite element method, discontinuous Galerkin method, *a posteriori* error estimates.

AMS subject classifications. 73F15, 45D05, 65M60

1. Introduction. The purpose of this paper is to derive *a posteriori* error estimates for space-time finite element approximations to the quasistatic hereditary linear viscoelasticity problem as described below. Spatial discretisation is effected using a standard Galerkin finite element method based on continuous piecewise polynomials of degree $p > 0$. We abbreviate the scheme to "cG(p)".

The time discretisation is carried out also by the (Galerkin) finite element method and we approximate in time with discontinuous piecewise polynomials of degree $r = 0$ or $r = 1$, which we refer to as "dG(r)". For $p > 0$ the fully discrete schemes are then abbreviated to "dG(r)cG(p)", for $r = 0, 1$.

The dG(0)cG(p) scheme generates a numerical algorithm which requires, for its implementation, relatively simple modifications to cG(p) linear elasticity software. The dG(1)cG(p) scheme, on the other hand, leads to a 2×2 block system at each time level and is less trivially inserted into existing software. For this reason it seems appropriate to also consider a cG(1)cG(p) finite element approximation—we leave this for another time.

The companion paper to this is [42] where we give *a priori* error estimates for dG(r)cG(1). For detailed accounts of viscoelasticity theory we refer to the many standard texts, and in particular to [20], [31] or [18]. Our motivation is the study of damping applications in engineering in which viscoelastic polymers play an important role, see [27, 28] and their references.

For a positive real number T let $\mathcal{J} := [0, T]$ denote a time interval, and for $n \in \{2, 3\}$ let Ω be a time-independent open bounded domain in $\mathbb{R}^n$ with polygonal/polyhedral boundary $\partial\Omega$. We suppose that the interior of a (linear) viscoelas-

*The authors would like to acknowledge the support of the US Army through its **Army Research Office** (grant # DAAD19–00–1–0421) and **European Research Office** (contract # N68171–97–M–5763), as well as the UK's **Engineering and Physical Sciences Research Council** (GR/R10844/01 and GR/M38070).

†`{simon.shaw,john.whiteman}@brunel.ac.uk`, BICOM, `www.brunel.ac.uk/~icsrbicm`, Brunel University, Uxbridge, UB8 3PH, U.K.

tic compressible body occupies Ω and is acted upon by a system of body forces $\boldsymbol{f} := (f_i(\boldsymbol{x},t))_{i=1}^n$ for $\boldsymbol{x} := (x_i)_{i=1}^n \in \Omega$ and $t \in \mathcal{J}$. Furthermore, we also suppose that the surface of the body coincides with $\partial\Omega$, and there exists a time independent closed set $\Gamma_D \subseteq \partial\Omega$ of positive surface measure on which the body is rigidly fixed in space and time. On the open (and possibly empty) set $\Gamma_N := \partial\Omega \setminus \Gamma_D$ there is prescribed a system of surface tractions $\boldsymbol{g} := (g_i(\boldsymbol{x},t))_{i=1}^n$ for $\boldsymbol{x} \in \Gamma_N$ and $t \in \mathcal{J}$. The unit outward directed normal vector to Γ_N is denoted by $\widehat{\boldsymbol{n}} := (\widehat{n}_i)_{i=1}^n$.

We use the function $\boldsymbol{u} = (u_i)_{i=1}^n : \Omega \times \mathcal{J} \to \mathbb{R}^n$ to describe the displacement from equilibrium resulting from the action of the applied forces $\boldsymbol{f}$ and $\boldsymbol{g}$. This is the linear theory, wherein we assume that $\boldsymbol{u}$ is "small" so as not to violate the assumption that Ω is time independent, and the deformation can be adequately described by the "small strain" tensor, given below in (6). We assume also that $t = 0$ is a reference time such that $\boldsymbol{u} \equiv \boldsymbol{0}$ for all $t < 0$.

The analysis that follows is an application of the so-called *Johnson paradigm*, see [12], and can be regarded as an extension of the linear elasticity results given in [30]. The next section, § 2, gives a short overview of this technique so that, after describing the viscoelasticity problem in § 3, we can place our work into the broader context. Section 4 then deals with the weak formulation of the problem, gives our basic assumptions and sets up the background necessary to the basic finite element discretisation that follows in § 5. The dG(r)cG(p) method is then analysed in § 6 where an *a posteriori* error estimate is given along with some sharpness bounds. We then finish with some concluding remarks in § 7 as well as an appendix: Appendix 7.

Numerical results are not included since we are currently developing code for both the problem described below, and the dynamic problem that results when the inertia term is retained.

2. The "Johnson paradigm". This "paradigm" is essentially a generalisation of the Aubin-Nitsche duality technique frequently employed in *a priori* error estimation. Borrowing heavily from [12], it can be explained as follows.

Consider the problem: find $u \in V$ such that $Au = f$. Having determined a finite element approximation $U \in V^h \subset V$ to u it is natural then to seek an *a posteriori* estimate for the error $e := u - U$. The weak form and finite element approximation are as follows:

$$(Au, v) = (f, v) \quad \forall v \in V \qquad \text{and} \qquad (AU, v) = (f, v) \quad \forall v \in V^h.$$

Hence, by Galerkin "orthogonality", $(r, e) = 0$ for all $v \in V^h$ where $r = Ae$ is the residual. Let χ solve the (continuous) dual problem $(v, A^*\chi) = (v, g)\ \forall v \in V$, for given g, and assume strong stability of derivative order p in the dual problem and a corresponding V^h-approximation property:

$$\|D^p\chi\| \leqslant S\|g\| \qquad \text{and} \qquad \|h^{-p}(\chi - \pi\chi)\| \leqslant C\|D^p\chi\|,$$

where $\pi\chi \in V^h$ is a suitable interpolant. Thus, taking $v = g = e$ in the dual problem gives:

$$\begin{aligned}
\|e\|^2 &= (e, e) = (e, A^*\chi) = (Ae, \chi) = (r, \chi - \pi\chi) = (h^p r, h^{-p}(\chi - \pi\chi)), \\
&\leqslant CS\|h^p r\|\,\|e\|, \\
\Longrightarrow \|u - U\| &\leqslant CS\|h^p r\|.
\end{aligned}$$

This *a posteriori* error bound is computable and, moreover, because it explicitly contains the discretisation parameter, "h", it can be used as the basis of an adaptive algorithm wherein we seek U such that $\|u - U\| \leqslant \mathrm{TOL}$.

This template, or *algorithm*, for *a posteriori* error analysis has been applied widely. See [12, 13] for an overview and for specific examples see Johnson and Hansbo, [30], for linear elasticity, the series by Eriksson *et al.*, [14, 15]..., for parabolic problems, Asadzadeh and Eriksson, [2], for boundary integral equations, Estep and French, [16, 17], for ordinary differential equations, Johnson, [29], for the wave equation and Süli and Houston, [45], for first order hyperbolic equations.

In each case *except the last* the strong stability estimate is available almost as a natural consequence of the underlying equation. For example, for ODEs one has $A = \frac{d}{dt}$ and so can expect $p = 1$, while for second-order elliptic problems, $A = -\nabla^2$, $p = 1$ comes automatically by standard energy arguments, and $p = 2$ can arise from elliptic regularity. For the heat and wave equations we may need to use different p-values for the space and time discretisation, but we nevertheless get $p > 0$.

On the other hand, in the last case Süli and Houston had only $p = 0$—but they did have the *pseudo strong stability* estimate,

$$\|D\chi\| \leqslant S\|Dg\|,$$

and their approach was to estimate the error in a weak norm:

$$\|u - U\|_- = \sup \frac{|(u - U, g)|}{\|g\|_+} = \sup \frac{|(hr, h^{-1}(\chi - \pi\chi))|}{\|Dg\|} \leqslant CS\|hr\|.$$

Although the underlying problems are very different, we encounter exactly this difficulty with the viscoelasticity problem set out in the next section. We also estimate the (temporal) error in a weak norm—just as above—but take an alternative approach as well: differentiate the residual. In this case, with π denoting the L_2 projection, the template is:

$$\|u - U\| = \sup \frac{|(r, \chi - \pi\chi)|}{\|g\|} = \sup \frac{|(r - \pi r, \chi - \pi\chi)|}{\|g\|} \leqslant CS\|hr'\|,$$

and is related to Estep and French's approach to cG approximations to ODEs, [17]. Lastly, let us also mention the, related, "Rannacher paradigm".

The underlying functional-analytic framework to the templates above is sufficiently flexible to allow a moderate choice of norms to use for estimating the error but, in practice, it is not possible for the user to select "any" norm. Even if it were, in practical problems it is often some linear *functional* (or output) of the error that is of most interest to the user (e.g. moisture uptake, air drag). Rannacher's technique, see e.g. [37], runs as follows.

To control $|l(u) - l(U)| = |l(e)|$ construct the dual problem $(v, A^*\chi) = l(v)$ and take $v = e$. Then,

$$|l(u) - l(U)| = |(r, \chi - \pi\chi)|,$$

The right hand side is now *computed* by "solving" (in reality, approximating) the dual problem so that $\chi - \pi\chi$ is known. (We can take $\pi\chi = 0$ in this error represenation.)

To close this section let us just mention that the idea of finite element methods in time is not new. In the 1970s the continuous Galerkin method was used for ODEs by Hulme in [23, 22] and this was followed up in the early 1980s with discontinuous approximation by Delfour, Hager and Trochu in [11].

Applications to space-time problems, however, pre-date these. The earliest references seem to go back to the 1960s with Oden, [33], and Fried, [19] (see also

[33, 34, 35, 36] for other early references). In the 1970s Zienkiewicz considered the interpretation of classical schemes in the Galerkin framework in [48], and Jamet, [25], used the dGcG method to model a parabolic problem with a (known) time-dependent boundary.

Also, Aziz and Monk have used cGcG for the heat equation in [4] with Aziz and Lui following with [3] for the forward-backward heat equation. Babuška and Janik in [5, 6] have considered the *hp*-version, in time, for the heat equation.

The viscoelasticity problem we consider below can be thought of as a second-kind Volterra equation,

$$Au(t) = L(t) + \int_0^t B(t-s)u(s)\,ds, \tag{1}$$

taking values in a Hilbert space, and so before moving on we will try to describe the main point of this paper in the context of a prototype (pure-time) problem. Recalling the discussion in § 2, we can guess that the presence of the elliptic operator, A, in the above means that we have basic energy stability (with $p = 1$) on the spatial derivatives, and therefore the mesh width, "h", appears in our *a posteriori* error estimate. On the other hand, for the temporal stability the best we can hope for is $\|D^p u\| \leqslant S\|D^p L\|$ (see theorem 2 below). This suggests that to build the time step into the *a posteriori* error estimate we need to measure the error either in a weak norm, or differentiate the residual (or both). In the context of the pure-time prototype,

$$u(t) = f(t) + \int_0^t \phi(t-s)u(s)\,ds,$$

(see [39, 41, 44]) we then expect the general result:

$$\|u-U\|_{W_p^{-s}(0,t_i)} \leqslant S(t_i)\left\| k^{s+m}\frac{d^m r}{dt^m}\right\|_{L_p(0,t_i)} \qquad \text{for } 0 \leqslant s,m \leqslant 1, \tag{2}$$

where: U is a dG(0) approximation to u; r is the residual; and, k is the time step function. (Note: r is in general discontinuous and so the norm on the right is a broken norm.)

Apart from the references given above there are also other approaches to the finite element discretisation of Volterra equations. For example, Bedivan and Fix in [8] describe a cG(1) Galerkin approximation to a scalar problem and follow this in [7] with a least squares finite element method. Also, with a specific application to viscoelasticity problems Buch *et al.* formulate a parallel solver in [9, 24]. This work, as well as the Bedivan-Fix approach, presents global space-time, *one-shot* solvers, as opposed to time stepping schemes, and are, in this sense, suited to the "Rannacher paradigm" of adaptivity as discussed earlier in § 2. For a brief survey of classical discretisations of Volterra equations we refer to [40], and for numerical viscoelasticity based on internal variables we recommend [26, 28].

3. Problem and background. In the quasistatic theory of viscoelasticity one assumes that the inertia of the body is negligible, and then Newton's second law of motion with boundary conditions gives for each $i \in \mathbb{N}(1,n) := \{1,\ldots,n\}$ that,

$$-\sigma_{ij,j} = f_i \qquad \text{in } \Omega \times \mathcal{J}, \tag{3}$$

$$u_i = 0 \qquad \text{in } \Gamma_D \times \mathcal{J}, \tag{4}$$

$$\sigma_{ij}\widehat{n}_j = g_i \qquad \text{in } \Gamma_N \times \mathcal{J}. \tag{5}$$

Here and throughout, repeated indices imply summation and $\underline{\boldsymbol{\sigma}} := (\sigma_{ij})_{i,j=1}^n$ is the symmetric stress tensor. In the linear theory one derives the (small) strain tensor $\underline{\boldsymbol{\varepsilon}} := (\varepsilon_{ij})_{i,j=1}^n$ from the displacement field using the relations,

$$\varepsilon_{ij}(\boldsymbol{u}) := \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right). \tag{6}$$

We close this problem by introducing the following linear hereditary viscoelastic constitutive relationship between stress and strain,

$$\sigma_{ij}(\boldsymbol{u};\boldsymbol{x},t) = D_{ijkl}(\boldsymbol{x},0)\varepsilon_{kl}(\boldsymbol{u}(\boldsymbol{x},t)) - \int_0^t \frac{\partial D_{ijkl}(\boldsymbol{x},t-s)}{\partial s}\varepsilon_{kl}(\boldsymbol{u}(\boldsymbol{x},s))\,ds. \tag{7}$$

Here $\underline{\boldsymbol{D}}(\boldsymbol{x},t) := \big(D_{ijkl}(\boldsymbol{x},t)\big)_{i,j,k,l=1}^n$ is a fourth-order stress relaxation tensor satisfying the following symmetries:

$$D_{ijkl}(t) = D_{jikl}(t) = D_{ijlk}(t) \qquad \text{but, in general,} \qquad D_{ijkl}(t) \neq D_{klij}(t). \tag{8}$$

However, we do have $D_{ijkl}(t) = D_{klij}(t)$ for $t = 0$ and $t = \infty$ in general, and for all t for isotropic materials (see e.g. [31, equations (1.10), (2.62)]). Here, and usually below, we omit the $\boldsymbol{x}$ dependence. Also, the components of $\underline{\boldsymbol{D}}$ can be assumed to be (a.e. in Ω) functions of t that are smooth enough for their first time derivatives to be of class $L_1(\mathcal{J})$. In addition, since $\underline{\boldsymbol{D}}(0)$ measures instantaneous linear elastic response we follow Hooke's law and assume positive-definiteness:

$$\gamma_{ij}\gamma_{kl}D_{ijkl}(0) > 0 \qquad \text{a.e. in } \Omega$$

for all non-zero symmetric second order tensors $\underline{\boldsymbol{\gamma}}$.

We note that the constitutive relationship (7) can also be written as,

$$\underline{\boldsymbol{\sigma}}(\boldsymbol{u};t) = \underline{\boldsymbol{\sigma}}^E(\boldsymbol{u}(t)) - \underline{\boldsymbol{\sigma}}^V(\boldsymbol{u};t,t), \tag{9}$$

where the elastic (E) and viscous (V) stresses are given by:

$$\sigma_{ij}^E(\boldsymbol{u}(t)) := D_{ijkl}(0)\varepsilon_{kl}(\boldsymbol{u}(t)), \tag{10}$$

$$\sigma_{ij}^V(\boldsymbol{u};\tau,t) := \int_0^\tau \frac{\partial D_{ijkl}}{\partial s}(t-s)\varepsilon_{kl}(\boldsymbol{u}(s))\,ds \qquad \text{for } 0 \leqslant \tau \leqslant t. \tag{11}$$

This form is useful for implementation.

4. Weak formulation and preliminaries. To give a weak formulation of this problem we first define the product Hilbert spaces, $\boldsymbol{H}^s(\Omega) := H^s(\Omega)^n$, for $s = 0, 1, 2, \ldots$, with inner products given by $(\boldsymbol{w},\boldsymbol{v})_s := \sum_{i=1}^n (w_i, v_i)_{H^s(\Omega)}$, for all $\boldsymbol{w}$, $\boldsymbol{v} \in \boldsymbol{H}^s(\Omega)$. These spaces have the natural norms $\|\cdot\|_s := \sqrt{(\cdot\,,\cdot)_s}$ and, of course, $\boldsymbol{L}_2(\Omega) \equiv \boldsymbol{H}^0(\Omega)$. Also, and as is usual for time dependent problems, for a Banach space $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ we define the $L_p(0,t;\mathcal{B})$ norms by, $\|\boldsymbol{v}\|_{L_p(0,t;\mathcal{B})} := \big\|\, \|\boldsymbol{v}(\,\cdot\,)\|_{\mathcal{B}}\, \big\|_{L_p(0,t)}$. We also use the (symmetric second order) tensor-valued L_2 space,

$$\underline{\boldsymbol{L}}_2(\Omega) := \left\{\underline{\boldsymbol{\gamma}} = (\gamma_{ij})_{i,j=1}^n : \gamma_{ij} = \gamma_{ji} \in L_2(\Omega)\ \forall i,j \in \mathbb{N}(1,n)\right\}.$$

Using the essential boundary condition (4) we now define the (spatial) test space,

$$H := \left\{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) : \boldsymbol{v} = \boldsymbol{0} \text{ on } \Gamma_D\right\}, \tag{12}$$

and (see e.g. [42] for details) arrive at the weak problem: find $\boldsymbol{u} \in L_p(\mathcal{J}; H)$ such that,

$$A(\boldsymbol{u}(t), \boldsymbol{v}) = \langle L(t), \boldsymbol{v}\rangle + \int_0^t B(t-s; \boldsymbol{u}(s), \boldsymbol{v})\, ds \qquad \forall \boldsymbol{v} \in H, \text{ a.e. in } \mathcal{J}. \tag{13}$$

Here the bilinear forms $A : H \times H \to \mathbb{R}$ and $B : \mathcal{J} \times H \times H \to \mathbb{R}$ are defined by,

$$A(\boldsymbol{w}, \boldsymbol{v}) := \int_\Omega D_{ijkl}(0)\varepsilon_{kl}(\boldsymbol{w})\varepsilon_{ij}(\boldsymbol{v})\, d\Omega, \tag{14}$$

$$B(t-s; \boldsymbol{w}, \boldsymbol{v}) := \int_\Omega \frac{\partial D_{ijkl}(t-s)}{\partial s}\varepsilon_{kl}(\boldsymbol{w})\varepsilon_{ij}(\boldsymbol{v})\, d\Omega, \tag{15}$$

for all $\boldsymbol{w}, \boldsymbol{v} \in H$, and $L : \mathcal{J} \to$ "dual space" is the time dependent linear form defined through,

$$\langle L(t), \boldsymbol{v}\rangle := \int_\Omega \boldsymbol{v} \cdot \boldsymbol{f}(t)\, d\Omega + \oint_{\Gamma_N} \boldsymbol{v} \cdot \boldsymbol{g}(t)\, d\Gamma \qquad \forall \boldsymbol{v} \in H. \tag{16}$$

The "dual space" and duality brackets used above will be defined properly in part (i) of Assumptions 1 below, and the reason for introducing them given after those assumptions. However, before we come to these we need some more notation.

Apart from using subscripts in the usual way we also abbreviate n-fold partial time differentiation by ∂_t^n, and also use primes as shorthand so that, for example, $v'(t) = \partial_t v(t) := \partial_t^1 v(t)$, $v''(t) = \partial_t^2 v(t), \ldots$. Also, and as already indicated above (3), we denote subsets of $\mathbb{Z}$ by $\mathbb{N}(m, n) := \{m, m+1, \ldots, n-1, n\}$, for $m \leqslant n$, and define $\mathbb{N}(m, n) := \varnothing$ if $m > n$.

Recall now that for a bounded interval $(a, b) \subset \mathbb{R}$ (with $a < b$ of course), the Sobolev space $W_p^m(a, b; H)$ (for $m = 0, 1, 2, \ldots$) contains all Lebesgue measurable functions $v\colon (a, b) \to H$ such that $\partial_t^r v \in L_p(a, b; H)$, for $r \in \mathbb{N}(0, m)$. The norm on $W_p^m(a, b; H)$ is given by $(\sum_{n=0}^m \|\partial_t^n \cdot\|^p_{L_p(a,b;H)})^{\frac{1}{p}}$ (with the obvious modification if $p = \infty$). The subspaces $\mathring{W}_p^m(a, b; H)$ contain all such functions with vanishing boundary traces: $\partial_t^r v(a) = \partial_t^r v(b) = 0$, for $r \in \mathbb{N}(0, m-1)$. We norm $\mathring{W}_p^m(a, b; H)$ by $\|\cdot\|_{\mathring{W}_p^m(a,b;H)} := \|\partial_t^m \cdot\|_{L_p(a,b;H)}$ and, using the boundary conditions, it follows from the fundamental theorem of calculus and the Hölder and arithmetic-geometric mean inequalities that for $0 \leqslant n \leqslant m$,

$$\|v\|_{\mathring{W}_p^n(a,b;H)} \leqslant (b-a)^{m-n}\|v\|_{\mathring{W}_p^m(a,b;H)} \qquad \forall v \in \mathring{W}_p^m(a, b; H). \tag{17}$$

We recall that when $v \in W_1^{m+1}(a, b; H)$ we can always take $\partial_t^m v \in C([a, b]; H)$ for all $m \geqslant 0$, and we adopt the standard convention that $\mathring{W}_p^0 \equiv W_p^0 \equiv L_p$.

It is appropriate at this stage to state our basic assumptions for this problem.

ASSUMPTIONS 1 (**general assumptions**). *There is a $p \in (1, \infty]$ and integer $r \geqslant 0$ such that the following hold.*

(i) *The symmetric bilinear form $A(\,\cdot\,,\,\cdot\,)$ is continuous and coercive on H in the respective senses*:

$$|A(\boldsymbol{w}, \boldsymbol{v})| \leqslant C\|\boldsymbol{w}\|_1\|\boldsymbol{v}\|_1 \qquad \textit{and} \qquad A(\boldsymbol{v}, \boldsymbol{v}) \geqslant c\|\boldsymbol{v}\|_1^2,$$

for all $\boldsymbol{w}, \boldsymbol{v} \in H$ and where C and c are positive constants. Hence, $A(\,\cdot\,,\,\cdot\,)$ is a scalar product on H and induces the energy norm,

$$\|\boldsymbol{v}\|_H := \sqrt{A(\boldsymbol{v}, \boldsymbol{v})}$$

for all $\boldsymbol{v} \in H$, which (on H) is equivalent to the norm $\|\cdot\|_1$. Henceforth, by H we shall mean the subspace of $\boldsymbol{H}^1(\Omega)$ as defined above in (12)*, but equipped with this energy scalar product and norm. We denote the dual space by H' (as used in* (16)*) and the duality pairing between H and H' by $\langle \cdot , \cdot \rangle$.*

(ii) *Each component of $\underline{\boldsymbol{D}}$ satisfies $D_{ijkl} \in W_1^{r+1}(\mathcal{J}; L_\infty(\Omega))$. Then, the bilinear form $B(t; \cdot , \cdot)$ is continuous and similar to $A(\cdot , \cdot)$ in the sense that there exists $\phi \in L_1\big(\mathcal{J}; [0, \infty)\big)$ such that*

$$|B(t; \boldsymbol{w}, \boldsymbol{v})| \leqslant \phi(t)\|\boldsymbol{w}\|_H \|\boldsymbol{v}\|_H,$$

a.e. in $\mathcal{J}$ and for all $\boldsymbol{w}, \boldsymbol{v} \in H$.

(iii) *The linear form $L \in W_p^r(\mathcal{J}; H')$. Then,*

$$|\langle \partial_t^s L(t), \boldsymbol{v}\rangle| \leqslant \|\partial_t^s L(t)\|_{H'}\|\boldsymbol{v}\|_H \qquad \text{for } 0 \leqslant s \leqslant r,$$

a.e. in $\mathcal{J}$ and for all $\boldsymbol{v} \in H$.

Note the introduction of the energy norm $\|\cdot\|_H$ in part (i), and its subsequent reappearance in parts (ii) and (iii). This H-coercivity is a consequence of *Korn's inequality* which we assume holds since we insist that $\text{meas}(\Gamma_D) > 0$ (and so H contains no rigid body motions). To motivate part (ii) we need only look back to equation (15).

Below we assume that $\boldsymbol{f} \in L_p(\mathcal{J}; \boldsymbol{L}_2(\Omega))$ and $\boldsymbol{g} \in L_p(\mathcal{J}; \boldsymbol{L}_2(\Gamma_N))$, which means that the definition of the functional $L(t)$ in (16) is not really necessary. However, for the dual problem, (24), it is convenient to take a more abstract approach since the stability estimates that follow (which play a crucial role in the *a posteriori* error analysis) can be written more accurately using H'.

A proof of the following estimate for the problem (13), under assumptions that are reasonable for linear viscoelasticity, is given in [43] for the case $r = 0$.

THEOREM 2 (**data stability**). *Let Assumptions* 1 *hold for some $r \geqslant 0$ with, in addition, $L \in \mathring{W}_p^r(\mathcal{J}; H')$. Then there exists a* stability factor *$S : \mathcal{J} \to [0, \infty)$ such that,*

$$\|\partial_t^r \boldsymbol{u}\|_{L_p(0,t;H)} \leqslant S(t)\|\partial_t^r L\|_{L_p(0,t;H')}$$

for all $t \in \mathcal{J}$.

Proof. Accepting from [43] that the estimate holds for $r = 0$ we have only to verify the results for $r > 0$. Assuming smooth functions we can differentiate both sides of (13) and integrate by parts in the Volterra integral (noting that the boundary terms vanish) to get,

$$A(\boldsymbol{u}'(t), \boldsymbol{v}) = \langle L'(t), \boldsymbol{v}\rangle + \int_0^t B(t - s; \boldsymbol{u}'(s), \boldsymbol{v})\, dt \qquad \forall \boldsymbol{v} \in H.$$

This is of precisely the same form as (13) and so we can use the basic form of the estimate (with $r = 0$) to obtain the estimate for $r = 1$. Continuing in this way proves the result for all positive r in the case of smooth functions. The theorem follows by a standard density argument. □

In the general case the stability factor $S(t)$ is derived from Gronwall's inequality and is exponentially large in t. However, for the viscoelasticity problem described above one can make more precise and physically reasonable assumptions on $\underline{\boldsymbol{D}}$, motivated by viscoelastic *fading memory*, and use a more sensitive comparison theorem to establish sharper estimates for $S(t)$. In particular, for a viscoelastic solid we have

$S(t) = O(1)$, independent of $\mathrm{meas}(\mathcal{J})$, and for a viscoelastic fluid (in the sense described by Golden and Graham in [20]) we have $S(t) = O(1+t)$. Full details can be found in [43].

Negative (weak) Sobolev norms are defined for $p \in (1,\infty]$ by,

$$\|\cdot\|_{W_p^{-m}(a,b;H')} := \sup\left\{ \left| \int_a^b \langle\,\cdot\,,\boldsymbol{v}(t)\rangle\,dt \right| : \begin{array}{c} \boldsymbol{v} \in \mathring{W}_q^m(a,b;H) \\ \|\boldsymbol{v}\|_{\mathring{W}_q^m(a,b;H)} = 1 \end{array} \right\}, \tag{18}$$

for $m = 0, 1, 2, \ldots$, and where $q \in [1,\infty)$ is the conjugate Hölder index to p. We can also arrive at a *weak-strong* norm by invoking the Riesz map $\mathcal{R}\colon H \to H'$ defined via the Riesz Representation Theorem through,

$$\langle \mathcal{R}\boldsymbol{w}, \boldsymbol{v}\rangle = A(\boldsymbol{w},\boldsymbol{v}) \qquad \forall \boldsymbol{w},\boldsymbol{v} \in H. \tag{19}$$

Using the fact that $\mathcal{R}$ is a bijective isometry we now define a weak-strong (or a *negative-positive*) norm, for $p \in (1,\infty]$ and $m \geqslant 0$, via,

$$\begin{aligned} \|\cdot\|_{W_p^{-m}(a,b;H)} &:= \|\mathcal{R}\,\cdot\,\|_{W_p^{-m}(a,b;H')} \\ &= \sup\left\{ \left| \int_a^b A(\,\cdot\,,\boldsymbol{v}(t))\,dt \right| : \begin{array}{c} \boldsymbol{v} \in \mathring{W}_q^m(a,b;H) \\ \|\boldsymbol{v}\|_{\mathring{W}_q^m(a,b;H)} = 1 \end{array} \right\}. \end{aligned} \tag{20}$$

The point is that H and not H' now appears on the left, and so an adaptive algorithm based on estimating the error in this norm will produce strong error control in that it monitors the error in H rather than the weaker H' (see Theorem 18 below).

Towards constructing a finite element approximation of (13) we first allow the test functions to be time dependent, and then integrate over $\mathcal{J}$. This results in an alternative statement of the problem as: find $\boldsymbol{u} \in L_p(\mathcal{J};H)$ such that,

$$a(\boldsymbol{u},\boldsymbol{v}) = l(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in L_q(\mathcal{J};H), \tag{21}$$

where q is again conjugate to p and,

$$a(\boldsymbol{w},\boldsymbol{v}) := \int_0^T A(\boldsymbol{w}(t),\boldsymbol{v}(t))\,dt - \int_0^T\int_0^t B(t-s;\boldsymbol{w}(s),\boldsymbol{v}(t))\,dsdt, \tag{22}$$

$$l(\boldsymbol{v}) := \int_0^T \langle L(t),\boldsymbol{v}(t)\rangle\,dt, \tag{23}$$

for all $\boldsymbol{w} \in L_p(\mathcal{J};H)$ and $\boldsymbol{v} \in L_q(\mathcal{J};H)$. Note that in (7), for example, we also use the symbol "l" as an integer index; a similar clash of notation will also occur below where we use k to denote time steps. Since the contexts are so different no confusion should arise.

The structure behind the proof of the *a posteriori* error estimates as derived later is similar to that of "Nitsche's lift", in that one first defines a continuous dual problem in order to derive an error representation formula, and then uses the stability properties of this problem and approximation-error estimates to arrive at the error bound. We define the continuous dual backward problem to be: find $\boldsymbol{\chi} \in L_q(\mathcal{J};H)$ such that,

$$a^*(\boldsymbol{\chi},\boldsymbol{v}) = l^*(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in L_p(\mathcal{J};H), \tag{24}$$

where (compare (21)–(23)):

$$a^*(\boldsymbol{\chi}, \boldsymbol{v}) := \int_0^T A(\boldsymbol{v}(t), \boldsymbol{\chi}(t))\,dt - \int_0^T \int_t^T B(s-t; \boldsymbol{v}(t), \boldsymbol{\chi}(s))\,dsdt, \tag{25}$$

$$l^*(\boldsymbol{v}) := \int_0^T \langle L^*(t), \boldsymbol{v}(t)\rangle\,dt, \tag{26}$$

for some $L^* \in L_q(\mathcal{J}; H')$. (Note that due to the last of equation (8) the dual problem may have a slightly different character to (21).) Using the symmetry of the bilinear form $A(\cdot\,,\,\cdot)$ and interchanging the order of time integration, we easily arrive at the following relationship between this dual problem and (21).

LEMMA 3 (**duality**). *$a^*(\boldsymbol{w}, \boldsymbol{v}) = a(\boldsymbol{v}, \boldsymbol{w})$ $\forall \boldsymbol{v} \in L_p(\mathcal{J}; H)$ and $\forall \boldsymbol{w} \in L_q(\mathcal{J}; H)$.*

To determine the data-stability properties of (24) we observe that it is no more than a forward problem in reversed time. Hence, with an appropriate change to the time variables, we conclude that if Theorem 2 holds in the context of the dual problem (24) (i.e. with L replaced by L^* and p replaced by its conjugate q) then,

$$\|\partial_t^m \boldsymbol{\chi}\|_{L_q(t,T;H)} \leqslant S(T-t)\|\partial_t^m L^*\|_{L_q(t,T;H')} \qquad \forall m \in \mathbb{N}(0,r) \text{ and } \forall t \in \mathcal{J}. \tag{27}$$

We now describe the finite element approximation.

5. Finite element preliminaries. The weak form, equation (21), is the starting point for the space-time finite element discretisation of the problem. To effect this we firstly discretise $\mathcal{J}$ into discrete times $0 = t_0 < t_1 < \cdots < t_N = T$, and then define the time intervals $\mathcal{J}_i := [t_{i-1}, t_i]$, and time steps, $k_i := t_i - t_{i-1} > 0$. We use $k \in L_\infty(\mathcal{J})$ to denote the piecewise constant function such that $k|_{\mathcal{J}_i} := k_i$.

During each $\mathcal{J}_i$ we construct on $\bar{\Omega}$ (in the usual way) a triangular/tetrahedral space-mesh of M_i elements and denote the domain Ω with this mesh by Ω_i. Element j of Ω_i will be denoted Ω_{ij} (an open set) and we set,

$$h_{ij} := \operatorname{diam}(\Omega_{ij}),$$

and use h to denote the piecewise constant (on $\bar{\Omega} \times \mathcal{J}$) mesh function given by $h|_{\Omega_{ij}\times\mathcal{J}_i} := h_{ij}$. We also use h_i to denote the mesh function at times $t \in \mathcal{J}_i$ given by $h_i|_{\Omega_{ij}} := h_{ij}$, and use these notations to see that $h|_{\mathcal{J}_i} := h_i$.

During each $\mathcal{J}_i$ we define, relative to the mesh on Ω_i, the semidiscrete (spatial) finite element space,

$$H_i := \Big\{\boldsymbol{v} \in H \cap \big(C(\bar{\Omega})\big)^n : \boldsymbol{v}|_{\Omega_{ij}} \in \mathbb{P}_p(\Omega_{ij})^n \text{ for each } \Omega_{ij} \subset \Omega_i\Big\}. \tag{28}$$

For space-time finite element approximation we also define, for $r = 0$ or $r = 1$, the fully discrete finite element spaces:

$$V_r^i := \mathbb{P}_r(\mathcal{J}_i; H_i),$$
$$V_r := \Big\{\boldsymbol{v} \in L_\infty(\mathcal{J}; H) : \boldsymbol{v}|_{\mathcal{J}_i} \in V_r^i \quad \forall i \in \mathbb{N}(1,N)\Big\}.$$

Here $\mathbb{P}_r(\mathcal{J}_i; H_i)$ is the vector space of polynomials of degree at most r defined on $\mathcal{J}_i$ and with coefficients in H_i. Note that our approximating functions in V_r are continuous in space but, in general, temporally discontinuous at the knots $\{t_i\}_{i=1}^{N-1}$. These discontinuities allow the space-meshes to change with time and this is the basis of the dG(r)cG(p) scheme.

For each edge/face ℓ in the mesh on Ω_i we associate a unit normal $\boldsymbol{\nu}^\ell$. It is immaterial in which direction this points on internal edges/faces, but for $\ell \subset \partial\Omega$ we always take $\boldsymbol{\nu}^\ell = \widehat{\boldsymbol{n}}$, the unit outward.

Next we need to make some additional assumptions concerning the regularity of the data and the approximation properties of the spatial discretisation. To ease the notation both below and later on (e.g. Lemma 8) we introduce broken norms, defined for each $\mathcal{J}_i$, by,

$$\|\!|\boldsymbol{v}|\!\|_{\boldsymbol{L}_2(\Omega_i)} := \left(\sum_{\Omega_{ij} \subset \Omega_i} \|\boldsymbol{v}\|^2_{\boldsymbol{L}_2(\Omega_{ij})} \right)^{\frac{1}{2}}, \tag{29}$$

$$\|\!|\boldsymbol{v}|\!\|_{\boldsymbol{L}_2(\cup_j \partial\Omega_{ij})} := \left(\sum_{\Omega_{ij} \subset \Omega_i} \|\boldsymbol{v}\|^2_{\boldsymbol{L}_2(\partial\Omega_{ij})} \right)^{\frac{1}{2}}. \tag{30}$$

(Note that the second is, in general, only a seminorm for $\boldsymbol{v}$.)

ASSUMPTIONS 4 (**discretisation assumptions**). *In addition to Assumptions* 1 *we also assume the following.*

(i) $\boldsymbol{f} \in L_p(\mathcal{J}; \boldsymbol{L}_2(\Omega))$ *and* $\boldsymbol{g} \in L_p(\mathcal{J}; \boldsymbol{L}_2(\Gamma_N))$, *and the mesh can and will always be constructed so as to respect the Dirichlet–Neumann interface* $\Gamma_D \cap \overline{\Gamma}_N$.

(ii) *The components of* $\underline{\boldsymbol{D}}(t)$ *are bounded and piecewise smooth on* Ω *for each* $t \in \mathcal{J}$. *Furthermore, for each* $i \in \mathbb{N}(1, N)$ *and for* $t \in \mathcal{J}_i$, *we can and do choose the mesh on* Ω_i *such that it respects the spatial discontinuities of* $\underline{\boldsymbol{D}}(t)$.

(iii) *Every space mesh is* nondegenerate *in that every element* Ω_{ij} *contains* (*resp. is contained by*) *a ball of radius* r_{i} (*resp.* r_{o}), *and the ratio* $r_{\mathrm{o}}/r_{\mathrm{i}}$ *is bounded.*

(iv) *Corresponding to the time slabs* $\{\mathcal{J}_i\}_{i=1}^N$ *there exists a collection* $\{\pi_i\}_{i=1}^N$ *of interpolators* $\pi_i : H \to H_i$ *for which the following stability estimates hold*:

$$\|\pi_i \boldsymbol{w}\|_{\boldsymbol{L}_2(\Omega)} \leqslant \Pi_0 \|\boldsymbol{w}\|_{\boldsymbol{L}_2(\Omega)} \qquad \textit{and} \qquad \|\pi_i \boldsymbol{w}\|_H \leqslant \Pi_1 \|\boldsymbol{w}\|_H, \tag{31}$$

as well as the following error estimates:

$$\|\!|h^{-1}(\boldsymbol{w} - \pi_i \boldsymbol{w})|\!\|_{\boldsymbol{L}_2(\Omega_i)} \leqslant \Pi_\Omega \, \|\boldsymbol{w}\|_H \tag{32}$$

$$\|\!|h^{-\frac{1}{2}}(\boldsymbol{w} - \pi_i \boldsymbol{w})|\!\|_{\boldsymbol{L}_2(\cup_j \partial\Omega_{ij})} \leqslant \Pi_\ell \, \|\boldsymbol{w}\|_H, \tag{33}$$

for all $\boldsymbol{w} \in H$ *and where* Π_0, Π_1, Π_Ω *and* Π_ℓ *are constants that are independent of* h *and* $\boldsymbol{w}$, *and which we assume can be taken independent of* i (*i.e. of time*). *Occasionally we will also use the "global interpolator"* π *defined piecewise by* $\pi|_{\mathcal{J}_i} := \pi_i$.

Note that the interpolation estimates do not require excessive regularity of $\boldsymbol{w}$. For example, in [38, Theorems 3.1 and 4.1, and Equation 5.5] such interpolators are defined for "rough" functions $\boldsymbol{w} \in H$ and estimates of the type assumed above are given. In terms of estimating the constants in interpolation-error estimates we refer also to [10, Exercise 3.1.2] and also to the methods used in [21, 30, 32].

Later we will also use the following L_2 projections. Let $\boldsymbol{P}_\Omega$ and $\boldsymbol{P}_\Gamma$ be defined piecewise (on each $\mathcal{J}_i$) by $\boldsymbol{P}_{\Omega_i} : \boldsymbol{L}_2(\Omega) \to H_i$ and $\boldsymbol{P}_{\Gamma_i} : \boldsymbol{L}_2(\Gamma_N) \to H_i|_{\Gamma_N}$ where:

$$(\boldsymbol{P}_{\Omega_i} \boldsymbol{w}, \boldsymbol{v})_{\boldsymbol{L}_2(\Omega)} = (\boldsymbol{w}, \boldsymbol{v})_{\boldsymbol{L}_2(\Omega)} \qquad \forall \boldsymbol{w} \in \boldsymbol{L}_2(\Omega),\ \boldsymbol{v} \in H_i, \tag{34}$$

$$(\boldsymbol{P}_{\Gamma_i} \boldsymbol{w}, \boldsymbol{v})_{\boldsymbol{L}_2(\Gamma_N)} = (\boldsymbol{w}, \boldsymbol{v})_{\boldsymbol{L}_2(\Gamma_N)} \qquad \forall \boldsymbol{w} \in \boldsymbol{L}_2(\Gamma_N),\ \boldsymbol{v} \in H_i. \tag{35}$$

We now move on to the finite element discretisation and *a posteriori* error analysis.

6. The dG(r)cG(p) finite element method. For $r = 0$ or 1 we form the finite element approximation to (21) as: find $\boldsymbol{U} \in V_r$ such that,

$$a(\boldsymbol{U}, \boldsymbol{v}) = l(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in V_r, \tag{36}$$

and subtracting this from (21) gives the fundamentally important Galerkin "orthogonality" relationship:

$$a(\boldsymbol{u} - \boldsymbol{U}, \boldsymbol{v}) = 0 \qquad \forall \boldsymbol{v} \in V_r. \tag{37}$$

We recall from [42] the *a priori* error estimate. (This is of course only a summary statement.)

THEOREM 5 (***A priori*** **energy-error estimate**). *If $\exists K > 0$ such that a.e. in Ω:*

- *$D_{ijkl}(t)\gamma_{ij}\gamma_{kl} \geqslant c_0 D_{ijkl}(0)\gamma_{ij}\gamma_{kl}$ for some constant $c_0 = c_0(K) > 0$, for all $t \in [0, K]$ and all $\underline{\gamma} \in \underline{\boldsymbol{L}}_2(\Omega)$;*
- *no component $D_{ijkl}(t)$ changes sign in $[0, K]$,*

then, under standard assumptions, for dG(r)cG(1) approximation in V_r, for $r = 0, 1$, the Galerkin error $\boldsymbol{e} := \boldsymbol{u} - \boldsymbol{U}$ satisfies the a priori error estimate,

$$\|\boldsymbol{u} - \boldsymbol{U}\|_{L_\infty(\mathcal{J};H)} \leqslant C(T) \left(\Pi_h \left\| h D^2_{\boldsymbol{x}} \boldsymbol{u} \right\|_{L_\infty(\mathcal{J};\boldsymbol{L}_2(\Omega))} + \Pi_k \left\| k^{r+1} \partial_t^{r+1} \boldsymbol{u} \right\|_{L_\infty(\mathcal{J};H)} \right).$$

Here $C(T)$ is a discrete stability factor and Π_h, Π_k are constants. This estimate holds for $r = 1$ only if each k_i is small enough.

Our interest here is to generate an *a posteriori* error estimate and our first step toward this is to derive an error representation formula in terms of the dual problem, (24). For this we need some notation.

Define the space of "semidiscrete dual functions",

$$H_{\mathcal{J}} := \left\{ \boldsymbol{v} \in L_q(\mathcal{J}; H) : \boldsymbol{v}|_{\mathcal{J}_i} \in L_q(\mathcal{J}_i; H_i) \ \forall i \in \mathbb{N}(1, N) \right\}, \tag{38}$$

and let $P_r : H_{\mathcal{J}} \to V_r$ denote a map which we will specify later in Subsection 6.2. With regard to the solution $\boldsymbol{\chi}$ of the dual problem we now define $\boldsymbol{\theta} \in L_q(\mathcal{J}; H)$ and $\boldsymbol{\rho} \in H_{\mathcal{J}}$ in the following piecewise manner:

$$\boldsymbol{\theta}|_{\mathcal{J}_i} \equiv \boldsymbol{\theta}_i := \boldsymbol{\chi} - \pi_i \boldsymbol{\chi} \in L_q(\mathcal{J}_i; H), \tag{39}$$

$$\boldsymbol{\rho}|_{\mathcal{J}_i} \equiv \boldsymbol{\rho}_i := \pi_i \boldsymbol{\chi} - P_r \pi_i \boldsymbol{\chi} \in L_q(\mathcal{J}_i; H_i), \tag{40}$$

and with these we obtain a basic error representation formula.

LEMMA 6 (**error representation**). *The Galerkin error $\boldsymbol{e} := \boldsymbol{u} - \boldsymbol{U}$ satisfies*

$$l^*(\boldsymbol{e}) = G(\boldsymbol{\theta}) + G(\boldsymbol{\rho}),$$

where $G(\,\cdot\,) := l(\cdot) - a(\boldsymbol{U}, \,\cdot\,)$.

Proof. In the dual problem (24) we take $\boldsymbol{v} = \boldsymbol{e} \in L_p(\mathcal{J}; H)$, and use Lemma 3 with equations (37) and (21) to get:

$$\begin{aligned} l^*(\boldsymbol{e}) &= a^*(\boldsymbol{\chi}, \boldsymbol{e}) = a(\boldsymbol{e}, \boldsymbol{\chi}) = a(\boldsymbol{e}, \boldsymbol{\chi} - P_r\pi\boldsymbol{\chi}) = a(\boldsymbol{u} - \boldsymbol{U}, \boldsymbol{\chi} - P_r\pi\boldsymbol{\chi}), \\ &= l(\boldsymbol{\chi} - P_r\pi\boldsymbol{\chi}) - a(\boldsymbol{U}, \boldsymbol{\chi} - P_r\pi\boldsymbol{\chi}), \end{aligned}$$

since $P_r\pi\boldsymbol{\chi} \in V_r$. Now, $\boldsymbol{\chi} - P_r\pi\boldsymbol{\chi} = \boldsymbol{\theta} + \boldsymbol{\rho}$, and the result follows from the linearity of $l(\,\cdot\,)$ and $a(\boldsymbol{U}, \,\cdot\,)$. ◻

In an adaptive scheme we want to retain the freedom to refine and de-refine the space mesh, as necessary, throughout the time stepping. This is problematic in terms of computing stress residuals since they consist of discrete-stress divergences on elements, and discrete-stress jumps over element edges (see Lemma 7 below). Since (7) implies that the discrete stress will contain contributions from all previous space meshes, these residual terms quickly become impractical to compute. Our remedy for this is as follows.

The discrete stress is found by inserting the approximate displacement, $\boldsymbol{U}$, into (7) or (equivalently) (9) to get, for $t \in \mathcal{J}_i$,

$$\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t) = \underline{\boldsymbol{\sigma}}^E(\boldsymbol{U}(t)) - \int_{t_{i-1}}^{t} \underline{\boldsymbol{D}}_s(t-s)\underline{\boldsymbol{\varepsilon}}(\boldsymbol{U}(s))\,ds - \underline{\boldsymbol{\sigma}}^V(\boldsymbol{U};t_{i-1},t).$$

These terms are spatially discontinuous and (if $\underline{\boldsymbol{D}}$ is piecewise constant) piecewise polynomial of degree $p-1$. The first two terms on the right arise from the displacement on the current mesh but the last term contains contributions from all previous meshes. To eliminate this "mesh history" we use the $\underline{\boldsymbol{L}}_2(\Omega)$ peojection to represent the stress history on the current mesh by defining a new version of $\underline{\boldsymbol{\sigma}}^V$ as $\tilde{\underline{\boldsymbol{\sigma}}}^V$ via,

$$(\tilde{\underline{\boldsymbol{\sigma}}}^V(\boldsymbol{U};t_{i-1},t),\underline{\boldsymbol{\varepsilon}})_{\underline{\boldsymbol{L}}_2(\Omega)} = (\underline{\boldsymbol{\sigma}}^V(\boldsymbol{U};t_{i-1},t),\underline{\boldsymbol{\varepsilon}})_{\underline{\boldsymbol{L}}_2(\Omega)} \tag{41}$$

for all piecewise continuous tensor-valued degree $p-1$ polynomials, $\underline{\boldsymbol{\varepsilon}} \in \underline{\boldsymbol{L}}_2(\Omega)$. Note that this projection can be localised to each spatial element and is therefore cheap to compute. Note also that, in principle, no relationship whatsoever between temporally adjacent space meshes need be assumed but, in practice, some kind of parent-child element data structure will result in a significantly simpler implementation.

6.1. Bounds for $G(\boldsymbol{\theta})$. Our next task is to derive some preliminary bounds on $G(\boldsymbol{\theta})$ and $G(\boldsymbol{\rho})$. We leave $G(\boldsymbol{\rho})$ until later in Subsection 6.3 and concern ourselves here with $G(\boldsymbol{\theta})$, but first we need some standard notation. Suppose, during $\mathcal{J}_i$, that the edge/face ℓ is shared by the elements Ω_{im} and Ω_{in} and consider an arbitrary element $\boldsymbol{\gamma} := (\gamma_k)_{k=1}^n \in H$. We define the jump in $\boldsymbol{\gamma}$ across ℓ as the vector $[\![\boldsymbol{\gamma}]\!]_\ell$ with components given by,

$$[\![\gamma_k(\boldsymbol{x})]\!]_\ell := \pm\lim_{\epsilon\downarrow 0}\big(\gamma_k(\boldsymbol{x}-\epsilon\boldsymbol{\nu}^\ell) - \gamma_k(\boldsymbol{x}+\epsilon\boldsymbol{\nu}^\ell)\big) \qquad \forall \boldsymbol{x} \in \ell. \tag{42}$$

where, to avoid elaborate notation later on, we use "$\pm$" to indicate that the sign of this jump quantity is of no interest at all.

We denote surface integrals on the element boundaries by

$$(\boldsymbol{w},\boldsymbol{v})_\ell := \int_\ell \boldsymbol{w}\cdot\boldsymbol{v}\,d\ell \qquad \text{with} \qquad \|\cdot\|_\ell := \sqrt{(\,\cdot\,,\,\cdot\,)_\ell}.$$

Note that the $\boldsymbol{L}_2$ requirement on the viscous stress in the following lemma uses Assumption 4, part (ii), and requires the stress projection described earlier in (41).

LEMMA 7 (**spatial residual**). *Suppose during $\mathcal{J}_i$, for all $i \in \mathbb{N}(1,N)$, that the discrete stress, defined by using $\boldsymbol{U}$ in* (9)*, satisfies $\nabla\cdot\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)|_{\Omega_{in}} \in \boldsymbol{L}_2(\Omega_{in})$ for each $\Omega_{in} \subset \Omega_i$ and* *for every $t \in \mathcal{J}_i$. Then,*

$$G(\boldsymbol{\theta}) = \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} \Bigg(\sum_{\Omega_{in} \subset \Omega_i} \int_{\Omega_{in}} \big[\boldsymbol{f}(t) + \nabla \cdot \sigma(\boldsymbol{U};t)\big] \cdot \boldsymbol{\theta}(t)\, d\Omega$$
$$+ \oint_{\Gamma_N} \big[\boldsymbol{g}(t) - \underline{\boldsymbol{\sigma}}(\boldsymbol{U};t) \cdot \widehat{\boldsymbol{n}}\big] \cdot \boldsymbol{\theta}(t)\, d\Gamma + \sum_{\ell \subset \Omega_i} \int_{\ell} [\![\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t) \cdot \boldsymbol{\nu}^{\ell}]\!]_{\ell} \cdot \boldsymbol{\theta}(t)\, d\ell \Bigg)\, dt.$$

Here, $\nabla \cdot \underline{\boldsymbol{\sigma}}$ denotes the vector $(\sigma_{ij,j})_{i=1}^{n}$.

Proof. Using (16), (10) and (11) in (13) with the definition of $G(\boldsymbol{\theta})$ from Lemma 6, noting (21)–(23), and that (from (9)) $\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t) = \underline{\boldsymbol{\sigma}}^E(\boldsymbol{U}(t)) - \underline{\boldsymbol{\sigma}}^V(\boldsymbol{U};t,t)$, we have,

$$G(\boldsymbol{\theta}) = \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} \left(\int_{\Omega} \boldsymbol{f} \cdot \boldsymbol{\theta}\, d\Omega + \oint_{\Gamma_D} \boldsymbol{g} \cdot \boldsymbol{\theta}\, d\Gamma - \int_{\Omega} \sigma_{mn}(\boldsymbol{U};t) \varepsilon_{mn}(\boldsymbol{\theta})\, d\Omega \right) dt.$$

Now consider each time integrand (i.e. for each i) as a sum over the elements $\Omega_{im} \subset \Omega_i$ and integrate by parts elementwise in the standard way. ∎

For each time level we define $\boldsymbol{r}|_{\mathcal{J}_i} = (r_k)_{k=1}^{n}$ by,

$$\boldsymbol{r}(\boldsymbol{U};t) := \begin{cases} \frac{1}{2} \Big| [\![\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t) \cdot \boldsymbol{\nu}^{\ell}]\!]_{\ell} \Big|, & \text{for } \ell \subset \Omega_i, \\ \Big| \boldsymbol{g}(t) - \underline{\boldsymbol{\sigma}}(\boldsymbol{U};t) \cdot \widehat{\boldsymbol{n}} \Big|, & \text{on } \Gamma_N, \\ 0, & \text{on } \Gamma_D. \end{cases} \tag{43}$$

Recall that each Ω_i is open, and so "$\ell \subset \Omega_i$" refers to all internal edges/faces.

Also, define $\mathcal{J}(t) \in L_2(\Omega)$ by,

$$\mathcal{J}(t)|_{\Omega_{ij}} := \frac{\|\boldsymbol{r}(\boldsymbol{U};t)\|_{\boldsymbol{L}_2(\partial\Omega_{ij})}}{\sqrt{h_{ij} \operatorname{meas}(\Omega_{ij})}} \tag{44}$$

for all $\Omega_{ij} \in \Omega_i$. Then,

$$\left(\sum_{\Omega_{ij} \subset \Omega_i} \left\| h_{ij}^{\frac{1}{2}} \boldsymbol{r}(\boldsymbol{U};t) \right\|^2_{\boldsymbol{L}_2(\partial\Omega_{ij})} \right)^{\frac{1}{2}} = \|h_i \mathcal{J}(t)\|_{L_2(\Omega)}. \tag{45}$$

Now we can state the bound.

LEMMA 8. *Let Assumptions* 1 *and* 4 *hold, and suppose that $\boldsymbol{\chi}|_{(t_I,T]} = \boldsymbol{0}$ in the dual problem* (24). *Then,*

$$|G(\boldsymbol{\theta})| \leqslant \mathcal{E}_{\Omega}(p, t_I; \boldsymbol{U}) \|\boldsymbol{\chi}\|_{L_q(0,t_I;H)},$$

where q is the conjugate Hölder index to p and, for $p \in [1, \infty)$,

$$\mathcal{E}_{\Omega}(p, t_I; \boldsymbol{U}) := \left(\sum_{i=1}^{I} \Big\| \Pi_{\Omega} \big\| h_i [\boldsymbol{f} + \nabla \cdot \underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)] \big\|_{\boldsymbol{L}_2(\Omega_i)} + \Pi_{\ell} \|h_i \mathcal{J}\|_{L_2(\Omega)} \Big\|^p_{L_p(\mathcal{J}_i)} \right)^{\frac{1}{p}},$$

while for $q = 1$,

$$\mathcal{E}_\Omega(\infty, t_I; \boldsymbol{U}) := \max_{1\leqslant i\leqslant I}\left\{\left\|\Pi_\Omega \left\|\!\left| h_i[\boldsymbol{f} + \nabla\cdot\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)]\right|\!\right\|_{\boldsymbol{L}_2(\Omega_i)} + \Pi_\ell \|h_i \partial\|_{L_2(\Omega)}\right\|_{L_\infty(\mathcal{J}_i)}\right\}.$$

The broken norms are given by (29) *and the constants come from* (32) *and* (33).

Proof. We use Lemma 7. Firstly, for $t \in \mathcal{J}_i$,

$$\begin{aligned}
\left|\sum_{\Omega_{in}\subset\Omega_i}\int_{\Omega_{in}} [\boldsymbol{f}(t) + \nabla\cdot\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)]\cdot\boldsymbol{\theta}(t)\,d\Omega\right| \qquad&\\
\leqslant \sum_{\Omega_{in}\subset\Omega_i} \|h_{in}[\boldsymbol{f}(t) + \nabla\cdot\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)]\|_{\boldsymbol{L}_2(\Omega_{in})}\,\|h_{in}^{-1}\boldsymbol{\theta}(t)\|_{\boldsymbol{L}_2(\Omega_{in})},&\\
\leqslant \left\|\!\left| h_i[\boldsymbol{f}(t) + \nabla\cdot\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)]\right|\!\right\|_{\boldsymbol{L}_2(\Omega_i)}\|h_i^{-1}\boldsymbol{\theta}(t)\|_{\boldsymbol{L}_2(\Omega_i)},&\\
\leqslant \Pi_\Omega\left\|\!\left| h_i[\boldsymbol{f}(t) + \nabla\cdot\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)]\right|\!\right\|_{\boldsymbol{L}_2(\Omega_i)}\|\boldsymbol{\chi}(t)\|_H,&
\end{aligned}$$

where we used the interpolation estimate (32). To deal with the stress jumps we let $|\cdot|$ denote the Euclidean norm and notice that $\boldsymbol{r} = \boldsymbol{\theta} = \boldsymbol{0}$ on Γ_D. Then,

$$\begin{aligned}
\left|\sum_{\ell\subset\Gamma_N}\int_\ell [\boldsymbol{g}(t) - \underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)\cdot\hat{\boldsymbol{n}}]\cdot\boldsymbol{\theta}(t)\,d\ell + \sum_{\ell\subset\Omega_i}\int_\ell [\![\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)\cdot\boldsymbol{\nu}^\ell]\!]\cdot\boldsymbol{\theta}(t)\,d\ell\right| \qquad&\\
\leqslant \sum_{\ell\subset\Gamma_N}\int_\ell \left|\boldsymbol{g}(t) - \underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)\cdot\hat{\boldsymbol{n}}\right|\,|\boldsymbol{\theta}(t)|\,d\ell + \sum_{\ell\subset\Omega_i}\int_\ell \left|[\![\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)\cdot\boldsymbol{\nu}^\ell]\!]\right|\,|\boldsymbol{\theta}(t)|\,d\ell,&\\
= \sum_{\Omega_{in}\subset\Omega_i}\oint_{\partial\Omega_{in}} h_{in}^{\frac12}|\boldsymbol{r}(\boldsymbol{U};t)|\,|h_{in}^{-\frac12}\boldsymbol{\theta}(t)|\,d\Gamma,&\\
\leqslant \Pi_\ell\|h_i\partial(t)\|_{L_2(\Omega)}\|\boldsymbol{\chi}(t)\|_H,&
\end{aligned}$$

where we used (33) and (45). The lemma now follows by invoking Hölder's inequality twice: first for each time integral over $\mathcal{J}_i$, and then for the resulting sum. ◻

The term $\mathcal{E}_\Omega$ introduced in Lemma 8 will appear in the *a posteriori* error estimate, and we note that it explicitly contains the meshwidth "h". This is due to the strong energy stability of the spatial derivatives (recall the discussion in § 2). The situation regarding the "temporal residual", $G(\boldsymbol{\rho})$, is not so straightforward, and before we consider this term we need to study the approximating properties of the discontinuous polynomials used in dG(r).

6.2. Discontinuous L_2-in-time projection. Our goal in this section is to define and analyse a discontinuous L_2-in-time projection in terms of $L_2(\mathcal{J}_i)$ projections onto V_r^i, for each time interval $\mathcal{J}_i$. In the following the choice of time interval, $\mathcal{J}_i$, is arbitrary (for $i \in \mathbb{N}(1, N)$).

DEFINITION 9 (**discontinuous L_2 projection**). *Let $(\mathcal{H}, (\cdot,\cdot)_\mathcal{H})$ be a Hilbert space. For $r = 0$ or 1 we define the map $P_r\colon L_1(\mathcal{J}_i;\mathcal{H}) \to \mathbb{P}_r(\mathcal{J}_i;\mathcal{H})$ by,*

$$\int_{t_{i-1}}^{t_i}\big(w(t) - P_r w(t)\big)\phi(t)\,dt = 0 \qquad \forall w \in L_1(\mathcal{J}_i;\mathcal{H}) \text{ and } \forall\phi\in\mathbb{P}_r(\mathcal{J}_i). \tag{46}$$

Equality in the above takes place in $\mathcal{H}$, and we note that with this definition the action of the map P_r is effectively independent of $\mathcal{H}$.

Later we will take $\mathcal{H}$ to be any of H, $\boldsymbol{L}_2(\Gamma_N)$ and $\underline{\boldsymbol{L}}_2(\Omega)$. In this subsection we collect together various results for this projection that will be important later. Full proofs are included because we are not aware of a reference which we can cite. Note first that for $a \leqslant b$,

$$\left\| \int_a^b v(t)\, dt \right\|_{\mathcal{H}} \leqslant \int_a^b \|v(t)\|_{\mathcal{H}}\, dt \qquad \forall v \in L_1(a,b;\mathcal{H}), \tag{47}$$

where we used the equality,

$$\left(\int_{t_{i-1}}^{t_i} w(t)\, dt, v \right)_{\mathcal{H}} = \int_{t_{i-1}}^{t_i} \big(w(t), v\big)_{\mathcal{H}}\, dt \qquad \forall w \in L_1(\mathcal{J}_i;\mathcal{H}),\ v \in \mathcal{H}.$$

Our first estimates for this projection are concerned with stability in $L_q(\mathcal{J}_i;\mathcal{H})$ but, beforehand, we need the following equivalence-of-norms result which is a consequence of $\mathbb{P}_r(\mathcal{J}_i)$ being finite dimensional.

LEMMA 10 (**equivalence of norms**). *There are constants $C_r(p,q)$ such that,*

$$\|v\|_{L_p(\mathcal{J}_i;\mathcal{H})} \leqslant C_r(p,q) k_i^{\frac{1}{p}-\frac{1}{q}} \|v\|_{L_q(\mathcal{J}_i;\mathcal{H})} \qquad \forall v \in \mathbb{P}_r(\mathcal{J}_i;\mathcal{H}), \tag{48}$$

and for $p,q \in [1,\infty]$. When $r=0$ we have $C_0(p,q)=1$ for all p, q.

Proof. We give the proof for $r=0$ only. Since v is then time-independent in each $\mathcal{J}_i$ we have easily that $k_i^{-1/p}\|v\|_{L_p(\mathcal{J}_i;\mathcal{H})} = \|v\|_{\mathcal{H}}$. The lemma follows from this. ☐

LEMMA 11 (P_r**-stability**). *For $r = 0, 1$,*

(i) $$\int_{t_{i-1}}^{t_i} (w(t) - P_r w(t), v)_{\mathcal{H}}\, dt = 0 \qquad \forall w \in L_1(\mathcal{J}_i;\mathcal{H}) \text{ and } \forall v \in \mathbb{P}_r(\mathcal{J}_i;\mathcal{H}),$$

(ii) $$\|P_r w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \leqslant \beth_{qr} \|w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \qquad \forall w \in L_q(\mathcal{J}_i;\mathcal{H}),$$

where $\beth_{q0} = 1$ and $\beth_{q1} = C_1(q,2)^2 C_1(\frac{q}{q-1}, q)$ where the latter constants are from (48).

Proof. To prove (i) we use (46) to get,

$$\int_{t_{i-1}}^{t_i} (w(t) - P_r w(t), v\phi(t))_{\mathcal{H}}\, dt = \Big(\int_{t_{i-1}}^{t_i} (w(t) - P_r w(t))\phi(t)\, dt, v \Big)_{\mathcal{H}} = 0,$$

for all $\phi \in \mathbb{P}_r(\mathcal{J}_i)$ and for all $v \in \mathcal{H}$. This is (i) since the elements of $\mathbb{P}_r(\mathcal{J}_i;\mathcal{H})$ are linear combinations of the form $\sum_i \phi_i(t) v_i$.

To prove (ii) for $r=0$ we assume that $\|P_0 w\|_{\mathcal{H}} \neq 0$ (otherwise the estimate is obvious). Since $P_0 w$ is constant in time we may take $v = \|P_0 w\|_{\mathcal{H}}^{q-2} P_0 w$ in (i) to get,

$$\|P_0 w\|_{L_q(\mathcal{J}_i;\mathcal{H})}^q \leqslant \|w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \|P_0 w\|_{L_q(\mathcal{J}_i;\mathcal{H})}^{q-1},$$

by Hölder's inequality. This is equivalent to (i) since $\|P_0 w\|_{\mathcal{H}} \neq 0$.

To prove (ii) for $r=1$ we take $v = P_1 w \in \mathbb{P}_1(\mathcal{J}_i;\mathcal{H})$ in (i). Then, two applications of (48) give, with p Hölder conjugate to q,

$$\begin{aligned}
\frac{k_i^{-2\left(\frac{1}{q}-\frac{1}{2}\right)}}{C_1(q,2)^2} \|P_1 w\|_{L_q(\mathcal{J}_i;\mathcal{H})}^2 &\leqslant \|P_1 w\|_{L_2(\mathcal{J}_i;\mathcal{H})}^2, \\
&\leqslant \|w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \|P_1 w\|_{L_p(\mathcal{J}_i;\mathcal{H})}, \\
&\leqslant C_1(p,q) k_i^{\frac{1}{p}-\frac{1}{q}} \|w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \|P_1 w\|_{L_q(\mathcal{J}_i;\mathcal{H})}.
\end{aligned}$$

Now, using $p = q/(q-1)$ and $1/p - 1/q = 1 - 2/q$ then gives,

$$\|P_1 w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \leqslant C_1(q,2)^2 C_1(\tfrac{q}{q-1}, q)\|w\|_{L_q(\mathcal{J}_i;\mathcal{H})},$$

as required. $\square$

When $r = 1$ we also have a stability estimate on the time derivative.

LEMMA 12 (**P_1-strong stability**). *For $q \in [1, \infty]$,*

$$\|\partial_t P_1 w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \leqslant \frac{3}{2}\|\partial_t w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \qquad \forall w \in W_q^1(\mathcal{J}_i;\mathcal{H}).$$

Proof. From Lemma 11 we have,

$$\int_{t_{i-1}}^{t_i} (P_1 w, v)_{\mathcal{H}}\, dt = \int_{t_{i-1}}^{t_i} (w, v)_{\mathcal{H}}\, dt$$

for all $v \in \mathbb{P}_1(\mathcal{J}_i;\mathcal{H})$ and for all $w \in W_q^1(\mathcal{J}_i;\mathcal{H})$. Obviously we can assume that $\partial_t P_1 w \neq 0$ and so, for $q \in [1,\infty)$, we choose $v = 12k_i^{-2}(t - t_{i-1/2})\|\partial_t P_1 w\|_{\mathcal{H}}^{q-2}\, \partial_t P_1 w$, where $t_{i-1/2} := (t_{i-1} + t_i)/2$. (Note that $v \in \mathbb{P}_1(\mathcal{J}_i;\mathcal{H})$ because $\partial_t P_1 w$ is constant in time.) We examine the left and right hand sides of the above equality separately. For the left hand side we have, by partial integration, that,

$$\begin{aligned}
&\frac{12}{k_i^2}\|\partial_t P_1 w\|_{\mathcal{H}}^{q-2} \int_{t_{i-1}}^{t_i} (t - t_{i-1/2})(P_1 w, \partial_t P_1 w)_{\mathcal{H}}\, dt \\
&\quad = \left[\frac{12}{k_i^2}\|\partial_t P_1 w\|_{\mathcal{H}}^{q-2}(P_1 w, \partial_t P_1 w)_{\mathcal{H}} \int_{t_{i-1}}^{t} (s - t_{i-1/2})\, ds\right]_{t_{i-1}}^{t_i} \\
&\qquad - \frac{12}{k_i^2}\|\partial_t P_1 w\|_{\mathcal{H}}^{q-2}(\partial_t P_1 w, \partial_t P_1 w)_{\mathcal{H}} \int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{t} (s - t_{i-1/2})\, ds dt, \\
&\quad = k_i\|\partial_t P_1 w\|_{\mathcal{H}}^{q} \;=\; \|\partial_t P_1 w\|_{L_q(\mathcal{J}_i;\mathcal{H})}^{q}.
\end{aligned}$$

Now, for the right hand side we again integrate by parts and get:

$$\begin{aligned}
&\frac{12}{k_i^2}\|\partial_t P_1 w\|_{\mathcal{H}}^{q-2} \int_{t_{i-1}}^{t_i} (t - t_{i-1/2})(w, \partial_t P_1 w)_{\mathcal{H}}\, dt \\
&\quad = \left[\frac{12}{k_i^2}\|\partial_t P_1 w\|_{\mathcal{H}}^{q-2}(w, \partial_t P_1 w)_{\mathcal{H}} \int_{t_{i-1}}^{t} (s - t_{i-1/2})\, ds\right]_{t_{i-1}}^{t_i} \\
&\qquad - \frac{12}{k_i^2}\|\partial_t P_1 w\|_{\mathcal{H}}^{q-2} \int_{t_{i-1}}^{t_i} \left[\int_{t_{i-1}}^{t} (s - t_{i-1/2})\, ds\right] (\partial_t w, \partial_t P_1 w)_{\mathcal{H}} dt, \\
&\quad = \frac{12}{k_i^2}\|\partial_t P_1 w\|_{\mathcal{H}}^{q-2} \int_{t_{i-1}}^{t_i} \left[\frac{k_i^2}{8} - \frac{(t - t_{i-1/2})^2}{2}\right] (\partial_t w, \partial_t P_1 w)_{\mathcal{H}} dt, \\
&\quad \leqslant \frac{3}{2}\|\partial_t P_1 w\|_{\mathcal{H}}^{q-2}\|\partial_t P_1 w\|_{L_p(\mathcal{J}_i;\mathcal{H})}\|\partial_t w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \\
&\quad = \frac{3}{2}\left(k_i^{\frac{1}{q}}\|\partial_t P_1 w\|_{\mathcal{H}}\right)^{q-1} \|\partial_t w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \;=\; \frac{3}{2}\|\partial_t P_1 w\|_{L_q(\mathcal{J}_i;\mathcal{H})}^{q-1}\|\partial_t w\|_{L_q(\mathcal{J}_i;\mathcal{H})}.
\end{aligned}$$

For $q \in [1, \infty)$ the lemma follows from this. For $q = \infty$ the lemma follows from the $q = 1$ case by using Hölder's inequality on the right and noting that, on the left, $\partial_t P_1 w$ is constant on $\mathcal{J}_i$. □

Our next result is concerned with certain integral preserving properties of the projection.

LEMMA 13. *For all $w \in L_1(\mathcal{J}_i; \mathcal{H})$:*

$$\text{(i)} \qquad \int_{t_{i-1}}^{t_i} w(t) - P_r w(t)\, dt = 0 \text{ (in } \mathcal{H}) \qquad \textit{for } r = 0, 1,$$

$$\text{(ii)} \qquad \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} w(s) - P_1 w(s)\, ds dt = 0 \text{ (in } \mathcal{H}),$$

for each time interval $\mathcal{J}_i$.

Proof. For (i) take $\phi = 1$ in (46). For (ii) partially integrate in (46) to get,

$$0 = \int_{t_{i-1}}^{t_i} (w - P_1 w)\phi\, dt = \phi(t) \int_{t_{i-1}}^{t} (w - P_1 w)\, ds \Big|_{t_{i-1}}^{t_i} - \int_{t_{i-1}}^{t_i} \phi'(t) \int_{t_{i-1}}^{t} (w - P_1 w)\, ds\, dt.$$

This is (ii) since $\phi' \in \mathbb{R}$ is arbitrary and the boundary terms vanish. □

We now need error estimates for P_r. As an aid in deriving these we invoke the following well known mean value theorem: if $v \in C(\mathcal{J}_i)$ then,

$$\int_{t_{i-1}}^{t_i} v(t)\, dt = k_i v(\xi) \qquad \text{for some} \qquad \xi \in (t_{i-1}, t_i).$$

LEMMA 14 (**P_r-error estimates**). *For $r = 0$ or 1:*

– *for all $w \in L_q(\mathcal{J}_i; \mathcal{H})$,*

$$\text{(i)} \quad \|w - P_r w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \leqslant (1 + \beth_{qr}) \|w\|_{L_q(\mathcal{J}_i;\mathcal{H})},$$

with $\beth_{qr}$ given by Lemma 11;

– *if $w \in W_q^{r+1}(\mathcal{J}_i; \mathcal{H})$ and $q \in [1, \infty]$, then,*

$$\text{(ii)} \quad \|k_i^{-s}(w - P_r w)\|_{L_q(\mathcal{J}_i;\mathcal{H})} \leqslant \daleth_{rs} \|\partial_t^s w\|_{L_q(\mathcal{J}_i;\mathcal{H})} \qquad \textit{for} \qquad 0 \leqslant s \leqslant r + 1,$$

where $\daleth_{01} = \daleth_{12} = 1$, $\daleth_{11} = \frac{5}{2}$ and $\daleth_{r0} = (1 + \beth_{qr})$.

Proof. The proof of (i) follows from the triangle inequality and Lemma 11 and the proof of (ii) when $r = s = 0$ follows from this and Lemma 11.

To prove (ii) when $r = 0$, $s = 1$ we have that $P_0 w$ is a constant (in time) and so solving (46) for $P_0 w$ and using the fundamental theorem of calculus gives,

$$w(t) - P_0 w(t) = w(t) - \frac{1}{k_i} \int_{t_{i-1}}^{t_i} \left[w(t) + \int_t^{\tau} w'(\xi)\, d\xi \right] d\tau = -\frac{1}{k_i} \int_{t_{i-1}}^{t_i} \int_t^{\tau} w'(\xi)\, d\xi d\tau.$$

Using (47) in this we obtain,

$$\|w(t) - P_0 w(t)\|_{\mathcal{H}} \leqslant \int_{t_{i-1}}^{t_i} \|w'(s)\|_{\mathcal{H}}\, ds,$$

and the result follows by taking the $L_q(\mathcal{J}_i)$ norm of both sides and using Hölder's inequality.

For $s-1=r=1$ we observe that $w \in C^1(\mathcal{J}_i;\mathcal{H})$ and so, using (i) in Lemma 13 with the Mean Value Theorem, we get a $\tau_1 \in (t_{i-1},t_i)$ such that $w(\tau_1) = P_1w(\tau_1)$. Now, set $t_{i-1/2} := (t_{i-1}+t_i)/2$ and choose $\phi(\tau) = \tau - t_{i-1/2} \in \mathbb{P}_1(\mathcal{J}_i)$ in (46). Integrating by parts, and noting that both boundary terms vanish, then gives,

$$0 = \int_{t_{i-1}}^{t_i} (w - P_1w)\phi \, dt = \int_{t_{i-1}}^{t_i} \left[-\int_{t_{i-1}}^{t} \phi(\tau)\, d\tau \right] (w - P_1w)'(t)\, dt.$$

The bracketed part of the integrand on the right is a negative quadratic, with roots at t_{i-1} and t_i, and so another application of the Mean Value Theorem gives a $\tau_2 \in (t_{i-1},t_i)$ such that $w'(\tau_2) = (P_1w)'(\tau_2)$. Using these τ_i we apply the fundamental theorem of calculus twice to get,

$$\begin{aligned}
[w(t) - P_1w(t)] &= [w(\tau_1) - P_1w(\tau_1)] + \int_{\tau_1}^{t} [w(\xi) - P_1w(\xi)]'\, d\xi, \qquad (\star)\\
&= (t-\tau_1)[w(\tau_2) - P_1w(\tau_2)]' + \int_{\tau_1}^{t}\int_{\tau_2}^{\xi} [w(\eta) - P_1w(\eta)]''\, d\eta d\xi,\\
&= \int_{\tau_1}^{t}\int_{\tau_2}^{\xi} w''(\eta)\, d\eta d\xi,
\end{aligned}$$

since $P_1w \in \mathbb{P}_1(\mathcal{J}_i;\mathcal{H})$. Hence, using (47) again with Hölder's inequality yields,

$$\|w(t) - P_1w(t)\|_{\mathcal{H}} \leqslant k_i^{1+\frac{1}{p}} \|w''\|_{L_q(\mathcal{J}_i;\mathcal{H})}.$$

The proof is completed by taking the $L_q(\mathcal{J}_i)$ norm of both sides. For the case $s=1$, $r=1$ we return to $(\star)$ and use Lemma 12 to get,

$$\|w(t) - P_1w(t)\|_{\mathcal{H}} \leqslant \|\partial_t w\|_{L_1(\mathcal{J}_i;\mathcal{H})} + \|\partial_t P_1 w\|_{L_1(\mathcal{J}_i;\mathcal{H})}, \leqslant \frac{5}{2} k_i^{\frac{1}{p}} \|\partial_t w\|_{L_q(\mathcal{J}_i;\mathcal{H})}.$$

Now take the $L_q(\mathcal{J}_i)$ norm of both sides. □

6.3. Bounds for $G(\boldsymbol{\rho})$. Using the piecewise L_2 projection from the previous subsection we can now give a preliminary bound on $G(\boldsymbol{\rho})$, as defined in Lemma 6. Note first that by using (9)—(11) in the definitions (22) and (23) of $l(\,\cdot\,)$ and $a(\boldsymbol{U},\,\cdot\,)$ means that we may write,

$$G(\boldsymbol{\rho}) = \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} (\boldsymbol{f},\boldsymbol{\rho})_{\boldsymbol{L}_2(\Omega)} + (\boldsymbol{g},\boldsymbol{\rho})_{\boldsymbol{L}_2(\Gamma_N)} - (\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t), \underline{\boldsymbol{\varepsilon}}(\boldsymbol{\rho}))_{\underline{\boldsymbol{L}}_2(\Omega)}\, dt.$$

We have three unlike terms and so we cannot combine them in a simple manner like "$\boldsymbol{f}+\boldsymbol{g}-\underline{\boldsymbol{\sigma}}$" to form a residual.

We can form a "proper" residual during $\mathcal{J}_i$ by seeking a pseudo displacement $\boldsymbol{W} \in H_i$ such that $A(\boldsymbol{W}(t),\boldsymbol{v}) = \langle L(t),\boldsymbol{v}\rangle$ for all $\boldsymbol{v} \in H_i$. Then, using (10) we can define an equivalent stress $\underline{\boldsymbol{\sigma}}^P(\boldsymbol{W}(t))$. The "proper" residual is then $\underline{\boldsymbol{\sigma}}^P(\boldsymbol{W}(t)) - \underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)$. The drawback here is that a linear elasticity system has to be solved in order to compute the residual (and, hence, the error estimate), and so the approach is impractical. Instead, we explore the idea of "differentiating the residual" (compare (2)).

LEMMA 15. *For the dG(r)cG(p) scheme we have (recall* (34) *and* (35)),

$$G(\boldsymbol{\rho}) = \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} (\boldsymbol{P}_{\Omega_i}\boldsymbol{f} - P_r\boldsymbol{P}_{\Omega_i}\boldsymbol{f}, \boldsymbol{\rho})_{\boldsymbol{L}_2(\Omega)}$$
$$+ (\boldsymbol{P}_{\Gamma_i}\boldsymbol{g} - P_r\boldsymbol{P}_{\Gamma_i}\boldsymbol{g}, \boldsymbol{\rho})_{\boldsymbol{L}_2(\Gamma_N)}$$
$$- (\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t) - P_r\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t), \underline{\boldsymbol{\varepsilon}}(\boldsymbol{\rho}))_{\underline{\boldsymbol{L}}_2(\Omega)}\,dt,$$

where for $\boldsymbol{w} = (w_i)_{i=1}^n$ we define $P_r\boldsymbol{w}$ in the natural way by $(P_r\boldsymbol{w})_i := P_r w_i$. In the above, the inclusion of $\boldsymbol{P}_\Omega$ and $\boldsymbol{P}_\Gamma$ arises from orthogonality and is, therefore, optional.

Proof. Recalling from (40) that $\boldsymbol{\rho} = (I - P_r)\pi_i\boldsymbol{\chi} \in H_i$ on each $\mathcal{J}_i$ we (optionally) use (34) and obtain,

$$\int_{t_{i-1}}^{t_i} (\boldsymbol{f}, \boldsymbol{\rho})_{\boldsymbol{L}_2(\Omega)}\,dt = \int_{t_{i-1}}^{t_i} (\boldsymbol{P}_{\Omega_i}\boldsymbol{f} - P_r\boldsymbol{P}_{\Omega_i}\boldsymbol{f}, \boldsymbol{\rho})_{\boldsymbol{L}_2(\Omega)}\,dt,$$

from the orthogonality built into (46). Now apply exactly the same process to the traction and stress terms. □

We can now state the counterpart to Lemma 8.

LEMMA 16 (**temporal residual**). *Let Assumptions* 1*, for $r = 0$ or* 1*, and* 4 *hold with the second of* (31) *strengthened, for each $\mathcal{J}_i$, to: a.e. in $\mathcal{J}_i$,*

$$\|\partial_t^s(\pi_i\boldsymbol{w})\|_H \leqslant \Pi_1\|\partial_t^s\boldsymbol{w}\|_H \qquad \forall\boldsymbol{w} \in W_q^s(\mathcal{J}_i; H_i) \quad \textit{and for} \quad 0 \leqslant s \leqslant r+1.$$

Also, assume that $\boldsymbol{\chi}|_{(t_I,T)} = \boldsymbol{0}$ in the dual problem (24)*. Then,*

$$|G(\boldsymbol{\rho})| \leqslant \hat{\mathcal{E}}_{\mathcal{J}}(p, s, t_I; \boldsymbol{U})\|\partial_t^s\boldsymbol{\chi}\|_{L_q(0,t_I;H)},$$

where, for $p \in [1, \infty)$,

$$\hat{\mathcal{E}}_{\mathcal{J}}(p, s, t_I; \boldsymbol{U}) := \Bigg(\sum_{i=1}^{I} \daleth_{rs}^p k_i^{ps}\Pi_1^p\Bigg\| C_H^i\big\|\boldsymbol{P}_{\Omega_i}\boldsymbol{f}(t) - P_r\boldsymbol{P}_{\Omega_i}\boldsymbol{f}(t)\big\|_{\boldsymbol{L}_2(\Omega)}$$
$$+ C_T^i\big\|\boldsymbol{P}_{\Gamma_i}\boldsymbol{g}(t) - P_r\boldsymbol{P}_{\Gamma_i}\boldsymbol{g}(t)\big\|_{\boldsymbol{L}_2(\Gamma_N)}$$
$$+ \big\|\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)(\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t) - P_r\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t))\big\|_{\underline{\boldsymbol{L}}_2(\Omega)}\Bigg\|_{L_p(\mathcal{J}_i)}^p\Bigg)^{\frac{1}{p}},$$

and,

$$\hat{\mathcal{E}}_{\mathcal{J}}(\infty, s, t_I; \boldsymbol{U}) := \max_{1\leqslant i\leqslant I}\Bigg\{\daleth_{rs}k_i^s\Pi_1\Bigg\| C_H^i\big\|\boldsymbol{P}_{\Omega_i}\boldsymbol{f}(t) - P_r\boldsymbol{P}_{\Omega_i}\boldsymbol{f}(t)\big\|_{\boldsymbol{L}_2(\Omega)}$$
$$+ C_T^i\big\|\boldsymbol{P}_{\Gamma_i}\boldsymbol{g}(t) - P_r\boldsymbol{P}_{\Gamma_i}\boldsymbol{g}(t)\big\|_{\boldsymbol{L}_2(\Gamma_N)}$$
$$+ \big\|\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)(\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t) - P_r\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t))\big\|_{\underline{\boldsymbol{L}}_2(\Omega)}\Bigg\|_{L_\infty(\mathcal{J}_i)}\Bigg\},$$

where, in these: the use of $\boldsymbol{P}_\Omega$ and $\boldsymbol{P}_\Gamma$ is optional; the constant Π_1 is given above; $\daleth_{rs}$ comes from Lemma 14*; and, C_H, C_T are defined for each $\mathcal{J}_i$ by,*

$$C_H^i \textit{ such that} \qquad \|\boldsymbol{v}\|_{\boldsymbol{L}_2(\Omega)} \leqslant C_H^i\|\boldsymbol{v}\|_H \quad \forall\boldsymbol{v} \in H_i,$$
$$C_T^i \textit{ such that} \qquad \|\boldsymbol{v}\|_{\boldsymbol{L}_2(\Gamma_N)} \leqslant C_T^i\|\boldsymbol{v}\|_H \quad \forall\boldsymbol{v} \in H_i.$$

Proof. We use Lemma 15. Firstly, note that the positive-definiteness (over symmetric second-order tensors) and symmetry of $\underline{\boldsymbol{D}}(0)$ gives,

$$\begin{aligned}|(\underline{\boldsymbol{\sigma}} - P_r\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\varepsilon}}(\boldsymbol{\rho}))_{\underline{\boldsymbol{L}}_2(\Omega)}| &= |(\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)(\underline{\boldsymbol{\sigma}} - P_r\underline{\boldsymbol{\sigma}}), \underline{\boldsymbol{D}}^{\frac{1}{2}}(0)\underline{\boldsymbol{\varepsilon}}(\boldsymbol{\rho}))_{\underline{\boldsymbol{L}}_2(\Omega)}| \\ &\leqslant \|\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)(\underline{\boldsymbol{\sigma}} - P_r\underline{\boldsymbol{\sigma}})\|_{\underline{\boldsymbol{L}}_2(\Omega)}\|\boldsymbol{\rho}\|_H.\end{aligned}$$

Now, in each $\mathcal{J}_i$ we have $\boldsymbol{\rho} \in H_i$ for a.e. $t \in \mathcal{J}_i$, and so using the estimate above we have,

$$\begin{aligned}&\Big|(\boldsymbol{P}_{\Omega_i}\boldsymbol{f} - P_r\boldsymbol{P}_{\Omega_i}\boldsymbol{f}, \boldsymbol{\rho})_{\boldsymbol{L}_2(\Omega)} + (\boldsymbol{P}_{\Gamma_i}\boldsymbol{g} - P_r\boldsymbol{P}_{\Gamma_i}\boldsymbol{g}, \boldsymbol{\rho})_{\boldsymbol{L}_2(\Gamma_N)} + (\underline{\boldsymbol{\sigma}} - P_r\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\varepsilon}}(\boldsymbol{\rho}))_{\underline{\boldsymbol{L}}_2(\Omega)}\Big| \\ &\quad\leqslant k_i^s\Big(C_H^i\|\boldsymbol{P}_{\Omega_i}\boldsymbol{f} - P_r\boldsymbol{P}_{\Omega_i}\boldsymbol{f}\|_{\boldsymbol{L}_2(\Omega)} + C_T^i\|\boldsymbol{P}_{\Gamma_i}\boldsymbol{g} - P_r\boldsymbol{P}_{\Gamma_i}\boldsymbol{g}\|_{\boldsymbol{L}_2(\Gamma_N)} \\ &\qquad + \|\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)(\underline{\boldsymbol{\sigma}} - P_r\underline{\boldsymbol{\sigma}})\|_{\underline{\boldsymbol{L}}_2(\Omega)}\Big) \times k_i^{-s}\|\boldsymbol{\rho}\|_H.\end{aligned}$$

Hölders inequality for integrals and then sums now yields,

$$|G(\boldsymbol{\rho})| \leqslant \hat{\mathcal{E}}_{\mathcal{J}}(p, t_I; \boldsymbol{U}) \times \daleth_{rs}^{-1}\Pi_1^{-1}\|k^{-s}(I - P_r)\pi_i\boldsymbol{\chi}\|_{L_q(0,t_I;H)},$$

and then Lemma 14 and our assumption on the interpolators π_i give,

$$\daleth_{rs}^{-1}\Pi_1^{-1}\|k^{-s}(I - P_r)\pi\boldsymbol{\chi}\|_{L_q(0,t_I;H)} \leqslant \Pi_1^{-1}\|\partial_t^s(\pi\boldsymbol{\chi})\|_{L_q(0,t_I;H)} \leqslant \|\partial_t^s\boldsymbol{\chi}\|_{L_q(0,t_I;H)}.$$

This completes the proof. ☐

REMARK 17. *The constant C_H^i can be estimated by the square root of the least eigenvalue of the problem,*

$$A(\boldsymbol{w}, \boldsymbol{v}) = \lambda(\boldsymbol{w}, \boldsymbol{v})_{\boldsymbol{L}_2(\Omega)} \qquad \forall \boldsymbol{v} \in H_i,$$

and C_T^i by the square root of the least eigenvalue of the problem,

$$A(\boldsymbol{w}, \boldsymbol{v}) = \lambda(\boldsymbol{w}, \boldsymbol{v})_{\boldsymbol{L}_2(\Gamma_N)} \qquad \forall \boldsymbol{v} \in H_i.$$

Within the context of a time-stepping scheme these computations are comparably inexpensive if we approximate the eigenvalues once only.

Now we can derive our *a posteriori* Galerkin-error estimate.

6.4. *A posteriori* error estimates. We begin with the theorem.

THEOREM 18 (***a posteriori* Galerkin energy-error estimate for** $dG(r)cG(p)$). *Let Assumptions* 1 *and* 4 *hold, and also let Assumptions* 1 *hold in the context of the dual problem* (24). *Then: for $0 \leqslant m, s \leqslant r+1$, $p > 1$ and each $j \in \mathbb{N}(1, N)$,*

$$\|\boldsymbol{u} - \boldsymbol{U}\|_{W_p^{-s}(0,t_j;H)} \leqslant S(t_j)\Big(t_j^s\mathcal{E}_\Omega(p, t_j; \boldsymbol{U}) + \mathcal{E}_{\mathcal{J}}(p, s, m, t_j; \boldsymbol{U})\Big),$$

where $S(t_j)$ is the stability factor from Theorem 2. Here, $\mathcal{E}_\Omega(\cdot,\cdot;\cdot)$ is from Lemma 8 and,

$$\begin{aligned}\mathcal{E}_{\mathcal{J}}(p, s, m, t_j; \boldsymbol{U}) := \daleth_{rm}\daleth_{rs}\Pi_1\Big(&C_H\|k^{s+m}\boldsymbol{P}_\Omega\partial_t^m\boldsymbol{f}\|_{L_p(0,t_j;\boldsymbol{L}_2(\Omega))} \\ &+ C_T\|k^{s+m}\boldsymbol{P}_\Gamma\partial_t^m\boldsymbol{g}\|_{L_p(0,t_j;\boldsymbol{L}_2(\Gamma_N))} \\ &+ \|k^{s+m}\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)\partial_t^m\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\,\cdot)\|_{L_p(0,t_j;\underline{\boldsymbol{L}}_2(\Omega))}\Big),\end{aligned}$$

where $\daleth_{rm}$ and $\daleth_{rs}$ come from Lemma 14 and the inclusion of $\boldsymbol{P}_\Omega$ and $\boldsymbol{P}_\Gamma$ is optional.

Proof. In the dual problem, (24), take $\boldsymbol{v} = \boldsymbol{e} := \boldsymbol{u} - \boldsymbol{U} \in L_p(\mathcal{J}; H)$ and, for some $s \in \mathbb{N}(0, r+1)$ choose $L^* \in \mathring{W}^s_q(\mathcal{J}; H')$ such that $L^*|_{(t_j,T]} = 0$ for an arbitrary time level t_j. Then, by the stability estimate (27), we have $\boldsymbol{\chi}|_{(t_j,T]} = \boldsymbol{0}$ and, moreover, by shifting—in the dual problem—the final time T backward to t_j:

$$\|\boldsymbol{\chi}\|_{\mathring{W}^s_q(0,t_j;H)} \leqslant S(t_j)\|L^*\|_{\mathring{W}^s_q(0,t_j;H')}.$$

Lemmas 6, 8 and 16 give,

$$|l^*(\boldsymbol{e})| \leqslant \mathcal{E}_\Omega(p, t_j; \boldsymbol{U})\|\boldsymbol{\chi}\|_{L_q(0,t_j;H)} + \hat{\mathcal{E}}_\mathcal{J}(p, s, t_j; \boldsymbol{U})\|\boldsymbol{\chi}\|_{\mathring{W}^s_q(0,t_j;H)},$$

and, by the equivalence of norms on $\mathring{W}^s_q(0, t_j; H)$, and the stability estimate given above,

$$|l^*(\boldsymbol{e})| \leqslant S(t_j)\Big(t_j^s\mathcal{E}_\Omega(p, t_j; \boldsymbol{U}) + \hat{\mathcal{E}}_\mathcal{J}(p, s, t_j; \boldsymbol{U})\Big)\|L^*\|_{\mathring{W}^s_q(0,t_j;H')}.$$

Now, using the "weak-strong" norm, (20), the symmetry $\langle \mathcal{R}\boldsymbol{e}, \boldsymbol{v}\rangle = \langle \mathcal{R}\boldsymbol{v}, \boldsymbol{e}\rangle$ in (19), and the definition of l^* from (26) we have, by making the correspondence $\mathcal{R}\boldsymbol{v} = L^*$, that,

$$\begin{aligned}
\|\boldsymbol{e}\|_{W_p^{-s}(0,t_j;H)} &= \sup\left\{\left|\int_0^{t_j}\langle \mathcal{R}\boldsymbol{e}, \boldsymbol{v}\rangle\,dt\right| : \begin{array}{c}\boldsymbol{v} \in \mathring{W}^s_q(0,t_j;H)\\ \|\boldsymbol{v}\|_{\mathring{W}^s_q(0,t_j;H)} = 1\end{array}\right\},\\
&= \sup\left\{\left|\int_0^{t_j}\langle L^*, \boldsymbol{e}\rangle\,dt\right| : \begin{array}{c}L^* \in \mathring{W}^s_q(0,t_j;H')\\ \|L^*\|_{\mathring{W}^s_q(0,t_j;H')} = 1\end{array}\right\},\\
&\leqslant S(t_j)\Big(t_j^s\mathcal{E}_\Omega(p, t_j; \boldsymbol{U}) + \hat{\mathcal{E}}_\mathcal{J}(p, s, t_j; \boldsymbol{U})\Big).
\end{aligned}$$

Next we examine the term $\mathcal{E}_\mathcal{J}(\cdot)$. When $m = 0$ we set P_r to be the zero map and obtain the theorem, while for $m > 0$ we use Lemma 14 and the triangle inequality to get,

$$\begin{aligned}
\|\boldsymbol{P}_{\Omega_i}\boldsymbol{f} - P_r\boldsymbol{P}_{\Omega_i}\boldsymbol{f}\|_{L_p(\mathcal{J}_i;\boldsymbol{L}_2(\Omega))} &+ \|\boldsymbol{P}_{\Gamma_i}\boldsymbol{g} - P_r\boldsymbol{P}_{\Gamma_i}\boldsymbol{g}\|_{L_p(\mathcal{J}_i;\boldsymbol{L}_2(\Gamma_N))}\\
&+ \|\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)\big(\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot) - P_r\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)\big)\|_{L_p(\mathcal{J}_i;\underline{\boldsymbol{L}}_2(\Omega))}\\
\leqslant \daleth_{rm}k_i^m\Big(&\|\partial_t^m\boldsymbol{P}_{\Omega_i}\boldsymbol{f}\|_{L_p(\mathcal{J}_i;\boldsymbol{L}_2(\Omega))} + \|\partial_t^m\boldsymbol{P}_{\Gamma_i}\boldsymbol{g}\|_{L_p(\mathcal{J}_i;\boldsymbol{L}_2(\Gamma_N))}\\
&+ \|\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)\partial_t^m\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)\|_{L_p(\mathcal{J}_i;\underline{\boldsymbol{L}}_2(\Omega))}\Big).
\end{aligned}$$

The proof is completed by using each of (34) and (35) to obtain "$\partial_t\boldsymbol{P} = \boldsymbol{P}\partial_t$". For example, applying ∂_t^m to both sides of (34) and taking the projection of $\partial_t^m\boldsymbol{w}$ gives:

$$(\partial_t^m\boldsymbol{P}_{\Omega_i}\boldsymbol{w}, \boldsymbol{v}) = (\partial_t^m\boldsymbol{w}, \boldsymbol{v}) = (\boldsymbol{P}_{\Omega_i}\partial_t^m\boldsymbol{w}, \boldsymbol{v}) \tag{49}$$

for all $\boldsymbol{v} \in H_i$. Choosing $\boldsymbol{v} = \partial_t^m\boldsymbol{P}_{\Omega_i}\boldsymbol{w}$ and using the Cauchy-Schwartz inequality then completes the proof. ☐

REMARK 19. *In Theorem 18 we could take $\boldsymbol{P}_\Gamma = \boldsymbol{P}_\Omega =$ identity because these maps did not need to be introduced in the proof of Lemma 15. This would simplify implementation.*

Note the degree of flexibility Theorem 18 affords: when the data are smooth full advantage can be taken to achieve error control in the strong L_p-energy norm. On the other hand, even for non-differentiable data, error control is still possible but at the price of estimation in a weak-energy norm.

Of course, these comments must be predicated on upper bounds on the residuals which demonstrate that they are sharp in the sense that they are of the same order as the error itself. This is the subject of the next subsection.

6.5. Upper bounds on the residuals. Our results in this section concern the sharpness of the *a posteriori* error estimate given in Theorem 18. Our goal is to show that the terms on the right of this error estimate yield an optimal *a priori* error estimate and thus can be used as the basis of an efficient adaptive algorithm.

For brevity we make the simplifying assumption that $\Gamma_N = \varnothing$, so we have a Dirichlet problem. Our method of proof (for the $\mathcal{E}_\Omega$ term) and assumptions on the approximation properties of the finite element spaces follows closely that used by Eriksson and Johnson in [14], and is given in the following Lemma. The proof, along with other technical assumptions, can be found in Appendix 7. For a different approach to estimating this type of explicit residual-based *a posteriori* error estimate see Verfürth [46], and also Ainsworth and Oden [1].

Below we consider only the dG(0)cG(1) scheme. (The $r = 1$ case would require a detailed stability analysis in order to obtain bounds on $\boldsymbol{U}'(t)$: it is certain that this will introduce extra conditions on the data and time step.)

LEMMA 20 (**Bound on** $\mathcal{E}_\Omega$). *Let the assumptions already stated continue to hold with Assumption* 4*, part* (ii) *strengthened so that the components of* $\underline{\boldsymbol{D}}(t)$ *are spatially constant for every* $t \in \mathcal{J}$*. Further, assume that each mesh* Ω_q *is constructed so that the following interpolation and inverse estimates hold*:

$$\text{(i) } \|h_q D^2_{h,q}\pi_q \boldsymbol{u}\|_{\boldsymbol{L}_2(\Omega)} + \|\boldsymbol{u} - \pi_q \boldsymbol{u}\|_H \leqslant C_\pi \|h_q D^2 \boldsymbol{u}\|_{\boldsymbol{L}_2(\Omega)},$$
$$\text{(ii) } \|h_q D^2_{h,q}\boldsymbol{v}\|_{\boldsymbol{L}_2(\Omega)} \leqslant C_q \|\boldsymbol{v}\|_H \qquad \forall \boldsymbol{v} \in H_q,$$

where $D^2_{h,q}$ *is the discrete "second derivative" defined in* (52)*. Then, there exists a positive constant* C *such that,*

$$\mathcal{E}_\Omega(t; \boldsymbol{U}(t)) \leqslant C \left(\|h\boldsymbol{f}\|_{L_\infty(\mathcal{J};\boldsymbol{L}_2(\Omega))} + \|hD^2\boldsymbol{u}\|_{L_\infty(\mathcal{J};\boldsymbol{L}_2(\Omega))} + \left\| k^m \frac{\partial^m \boldsymbol{u}}{\partial t^m} \right\|_{L_\infty(\mathcal{J};H)} \right).$$

In the above $m := r + 1$ *for approximation using the space* V_r*.*

The $\mathcal{E}_\Omega(\cdot)$ residual is quite standard in the *a posteriori* error analysis of elliptic problems and that is why we do not dwell on it. It is of greater interest here to examine the "temporal residuals" $\mathcal{E}_\mathcal{J}$ as given in Theorem 18, since these are non-standard. The first two terms (involving $\boldsymbol{f}$ and $\boldsymbol{g}$) are not the issue, but the third term does require further study.

To prepare, we first recall the following discrete stability estimate from [42, Theorem 6]: under not-too-restrictive assumptions on the data and discretisation there are positive constants, C_p, such that,

$$\|\boldsymbol{U}\|_{L_p(0,t_j;H)} \leqslant C_p \|L\|_{L_p(0,t_j;H')}, \tag{50}$$

for $p = 1, 2, \infty$ and all $j \in \mathbb{N}(1, N)$. Now, a preliminary lemma.

LEMMA 21. *In addition to the Assumptions made for Theorem* 18 *assume further that,*

– $\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)$, $\underline{\boldsymbol{D}}'(0) \in \underline{\boldsymbol{L}}_\infty(\Omega)$ *and* $\underline{\boldsymbol{D}}'' \in L_q(\mathcal{J}; \underline{\boldsymbol{L}}_\infty(\Omega))$;
– *The discrete stability estimate,* (50)*, holds for some* p;

where $p \in [1,\infty]$ and q are conjugate Hölder indices. Then, there are constants $C > 0$ such that for $p \in [1,\infty)$,

$$\begin{aligned}\|k^{s+1}\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)\|_{L_p(0,t_i;\underline{\boldsymbol{L}}_2(\Omega))}\\ \leqslant C \max_{1\leqslant j\leqslant i}\{k_j^{s+1}\}\Big(\|\boldsymbol{f}\|_{L_p(0,t_i;\boldsymbol{L}_2(\Omega))} + \|\boldsymbol{g}\|_{L_p(0,t_i;\boldsymbol{L}_2(\Gamma_N))}\Big),\end{aligned}$$

while for $p = \infty$,

$$\begin{aligned}\|k^{s+1}\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)\|_{L_\infty(0,t_i;\underline{\boldsymbol{L}}_2(\Omega))}\\ \leqslant C \max_{1\leqslant j\leqslant i}\Big\{k_j^{s+1}\Big(\|\boldsymbol{f}\|_{L_\infty(0,t_i;\boldsymbol{L}_2(\Omega))} + \|\boldsymbol{g}\|_{L_\infty(0,t_i;\boldsymbol{L}_2(\Gamma_N))}\Big)\Big\}.\end{aligned}$$

Both of these hold for all $i \in \mathbb{N}(1,N)$.

Proof. We obtain the discrete viscous stress, $\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)$ by inserting $\boldsymbol{U}$ into (7), and from this we find $\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)$. In a given $\mathcal{J}_j$ this is,

$$\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t) = \underline{\boldsymbol{D}}(0)\underline{\boldsymbol{\varepsilon}}(\boldsymbol{U}'(t)) + \underline{\boldsymbol{D}}'(0)\underline{\boldsymbol{\varepsilon}}(\boldsymbol{U}(t)) + \int_0^t \underline{\boldsymbol{D}}_{ss}(t-s)\underline{\boldsymbol{\varepsilon}}(\boldsymbol{U}(s))\,ds, \qquad (51)$$

and (because $r = 0$) we have $\boldsymbol{U}'(t) = \mathbf{0}$. For ease of exposition it is convenient here to think of $\underline{\boldsymbol{\sigma}}$ and $\underline{\boldsymbol{\varepsilon}}$ as vectors (not tensors), and $\underline{\boldsymbol{D}}$ as a matrix. The Euclidean norm will then be denoted $\|\cdot\|_{\mathbb{E}}$. Hence,

$$\|\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)\|_{\mathbb{E}} \leqslant \|\underline{\boldsymbol{D}}'(0)\|_{\mathbb{E}}\|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{U}(t))\|_{\mathbb{E}} + \int_0^t \|\underline{\boldsymbol{D}}''(t-s)\|_{\mathbb{E}}\|\underline{\boldsymbol{\varepsilon}}(\boldsymbol{U}(s))\|_{\mathbb{E}}\,ds.$$

Taking $\boldsymbol{L}_2(\Omega)$ norms then gives,

$$\|\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)\|_{\underline{\boldsymbol{L}}_2(\Omega)} \leqslant C\|\underline{\boldsymbol{D}}'(0)\|_{\underline{\boldsymbol{L}}_\infty(\Omega)}\|\boldsymbol{U}(t)\|_H + C\int_0^t \|\underline{\boldsymbol{D}}''(t-s)\|_{\underline{\boldsymbol{L}}_\infty(\Omega)}\|\boldsymbol{U}(s)\|_H\,ds,$$

for some constant $C = C(\underline{\boldsymbol{D}}) > 0$ and where in both of these results we used variants of (47). Multiplying by k_j^{s+1} and taking $L_p(\mathcal{J}_j)$ norms now gives,

$$\begin{aligned}\|k_j^{s+1}\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)\|_{L_p(\mathcal{J}_j;\underline{\boldsymbol{L}}_2(\Omega))}\\ \leqslant Ck_j^{s+1}\Big(\|\underline{\boldsymbol{D}}'(0)\|_{L_\infty(\Omega)}\|\boldsymbol{U}\|_{L_p(\mathcal{J}_j;H)} + k_j^{\frac{1}{p}}\|\underline{\boldsymbol{D}}''\|_{L_q(0,t_j;\underline{\boldsymbol{L}}_\infty(\Omega))}\|\boldsymbol{U}\|_{L_p(0,t_j;H)}\Big).\end{aligned}$$

Now, absorbing the terms in $\underline{\boldsymbol{D}}$ into a generic constant, C, we have,

$$\begin{aligned}\|k^{s+1}\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)\|^p_{L_p(0,t_i;\underline{\boldsymbol{L}}_2(\Omega))} &\leqslant C\sum_{j=1}^{i}\|k_j^{s+1}\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)\|^p_{L_p(\mathcal{J}_j;\underline{\boldsymbol{L}}_2(\Omega))},\\ &\leqslant C\sum_{j=1}^{i}k_j^{(s+1)p}\Big(\|\boldsymbol{U}\|_{L_p(\mathcal{J}_j;H)} + k_j^{\frac{1}{p}}\|\boldsymbol{U}\|_{L_p(0,t_j;H)}\Big)^p,\\ &\leqslant 2^pC\sum_{j=1}^{i}k_j^{(s+1)p}\Big(\|\boldsymbol{U}\|^p_{L_p(\mathcal{J}_j;H)} + k_j\|\boldsymbol{U}\|^p_{L_p(0,t_j;H)}\Big),\\ &\leqslant 2^pC\max_{1\leqslant j\leqslant i}\Big\{k_j^{(s+1)p}\Big(\|\boldsymbol{U}\|^p_{L_p(0,t_i;H)} + t_i\|\boldsymbol{U}\|^p_{L_p(0,t_i;H)}\Big)\Big\},\\ &\leqslant 2^pC(T)C_p^p\max_{1\leqslant j\leqslant i}\{k_j^{(s+1)p}\}\|L\|^p_{L_p(0,t_i;H')},\end{aligned}$$

from the discrete stability estimate, (50). Finally we use,

$$\begin{aligned}\|L\|_{L_p(0,t_i;H')} &= \left\| \sup_{\boldsymbol{v}\in H\setminus\{0\}} \frac{|\langle L(t),\boldsymbol{v}\rangle|}{\|\boldsymbol{v}\|_H} \right\|_{L_p(0,t_i)}, \\ &\leqslant \left\| C\|\boldsymbol{f}\|_{\boldsymbol{L}_2(\Omega)} + C_\gamma\|\boldsymbol{g}\|_{\boldsymbol{L}_2(\Gamma_N)} \right\|_{L_p(0,t_i)}, \\ &\leqslant C\Big(\|\boldsymbol{f}\|_{L_p(0,t_i;\boldsymbol{L}_2(\Omega))} + \|\boldsymbol{g}\|_{L_p(0,t_i;\boldsymbol{L}_2(\Gamma_N))}\Big),\end{aligned}$$

and this completes the proof for $p \in [1,\infty)$.

For $p=\infty$ we have,

$$\begin{aligned}\|k^{s+1}\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)\|_{L_\infty(0,t_i;\underline{\boldsymbol{L}}_2(\Omega))} &\leqslant C\max_{1\leqslant j\leqslant i}\left\{ \|k_j^{s+1}\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)\|_{L_\infty(\mathcal{J}_j;\underline{\boldsymbol{L}}_2(\Omega))}, \right\} \\ &\leqslant C\max_{1\leqslant j\leqslant i}\left\{ k_j^{s+1}\Big(\|\boldsymbol{U}\|_{L_\infty(\mathcal{J}_j;H)} + \|\boldsymbol{U}\|_{L_\infty(0,t_j;H)}\Big)\right\}, \\ &\leqslant C\max_{1\leqslant j\leqslant i}\left\{ k_j^{s+1}\|L\|_{L_\infty(0,t_j;H')}\right\}, \\ &\leqslant C\max_{1\leqslant j\leqslant i}\left\{ k_j^{s+1}\Big(\|\boldsymbol{f}\|_{L_\infty(0,t_j;\boldsymbol{L}_2(\Omega))} + \|\boldsymbol{g}\|_{L_\infty(0,t_j;\boldsymbol{L}_2(\Gamma_N))}\Big)\right\},\end{aligned}$$

as required. □

We can now use Lemma 21 to give an upper bound on the temporal residual $\mathcal{E}_{\mathcal{J}}(\cdot)$. Since the cases $p=\infty$, $r=0$ are the most useful from a practical point of view, we restrict attention to these values

LEMMA 22 (**Bound on $\mathcal{E}_{\mathcal{J}}$**). *For the dG(0)cG(1) scheme, under the previously indicated assumptions, there exists a constant $C>0$ such that,*

$$\begin{aligned}\mathcal{E}_{\mathcal{J}}(\infty,s,1,t_i;\boldsymbol{U}) \leqslant C\Bigg(&\max_{1\leqslant j\leqslant i}\left\{ k_j^{s+1}\Big(\|\boldsymbol{f}\|_{L_\infty(0,t_j;\boldsymbol{L}_2(\Omega))} + \|\boldsymbol{g}\|_{L_\infty(0,t_j;\boldsymbol{L}_2(\Gamma_N))}\Big)\right\} \\ &+ \max_{1\leqslant j\leqslant i}\left\{ k_j^{s+1}\|\partial_t\boldsymbol{f}\|_{L_\infty(\mathcal{J}_j;\boldsymbol{L}_2(\Omega))}\right\} \\ &+ \max_{1\leqslant j\leqslant i}\left\{ k_j^{s+1}\|\partial_t\boldsymbol{g}\|_{L_\infty(\mathcal{J}_j;\boldsymbol{L}_2(\Gamma_N))}\right\}\Bigg),\end{aligned}$$

for all $i\in\mathbb{N}(1,N)$.

Proof. Using the definition given in Theorem 18 we have,

$$\begin{aligned}&\mathcal{E}_{\mathcal{J}}(\infty,s,1,t_i;\boldsymbol{U}) \\ &\quad\leqslant C\urcorner_{0s}\urcorner_{01}\Pi_1\Bigg(\max_{1\leqslant j\leqslant i}\left\{ k_j^{s+1}\|\partial_t\boldsymbol{f}\|_{L_\infty(\mathcal{J}_j;\boldsymbol{L}_2(\Omega))}\right\} \\ &\qquad+ \max_{1\leqslant j\leqslant i}\left\{ k_j^{s+1}\|\partial_t\boldsymbol{g}\|_{L_\infty(\mathcal{J}_j;\boldsymbol{L}_2(\Gamma_N))}\right\} \\ &\qquad+ \|k^{s+1}\underline{\boldsymbol{D}}^{-\frac{1}{2}}(0)\partial_t\underline{\boldsymbol{\sigma}}(\boldsymbol{U};\cdot)\|_{L_\infty(0,t_j;\underline{\boldsymbol{L}}_2(\Omega))}\Bigg),\end{aligned}$$

since the projectors $\boldsymbol{P}_\Omega$ and $\boldsymbol{P}_\Gamma$ have norms bounded by unity. The proof is completed by using Lemma 21. □

Putting together Lemmas 20, and 22 together with Theorem 18 we have the following.

THEOREM 23. *For $\mathcal{E}_{\mathcal{J}}$ as defined in Theorem* 18 *and $\mathcal{E}_{\Omega}$ as defined in Lemma* 8 *we have for the dG(0)cG(1) approximation that,*

$$\|\boldsymbol{u}-\boldsymbol{U}\|_{W_\infty^{-s}(0,t_i;H)} \leqslant S(t_i)\Big(t_i^s\mathcal{E}_\Omega(\infty,t_i;\boldsymbol{U})+\mathcal{E}_{\mathcal{J}}(\infty,s,1,t_i;\boldsymbol{U})\Big),$$
$$= O\Big(\|h\|_{L_\infty(\Omega\times\mathcal{J})}+\|k^{s+1}\|_{L_\infty(\mathcal{J})}\Big),$$

for all $i \in \mathbb{N}(1,N)$ and where the explicit form of the rightmost term is given by combining the bounds in Lemmas 20 *and* 22*.*

This crude result shows that the *a posteriori* error estimate implied by Theorem 18 furnishes, up to a multiplicative constant, an optimal reflection of the error in the case $p=\infty, r=0$.

7. Closure. In this closing section we outline a few points regarding the interpretation and implementation of the foregoing material.

History storage. In general the entire solution history must be stored in order to be able to evaluate the Volterra integral for the stress (see (7)). However, in linear viscoelasticity it is common to represent the time dependence of $\underline{\boldsymbol{D}}$ as a *Prony series*—a linear combination of decaying exponentials. In this case recurrence relationships can be derived which means only one "history" vector need be stored for each term in the sum.

Variational crimes. In practice the relaxation functions in $\underline{\boldsymbol{D}}$ are simple enough for the inner products etc. to be evaluated exactly. However, in general, special attention will be required for the non-Galerkin quadrature errors introduced when integrating the load terms involving $\boldsymbol{f}$ and $\boldsymbol{g}$.

$\boldsymbol{L}_2(\Omega)$ estimator. For problems in which $\partial\Omega$ is smooth and/or Ω is convex-polygonal, and the data are smooth, it makes sense to seek $\boldsymbol{L}_2(\Omega)$ error estimates. These can be obtained for the problem considered above by using the "operator stability" estimates given in [43].

Appendix A. Proof of Lemma 20. For each time level $\mathcal{J}_q$, recall that we use Ω_q to denote the space mesh on Ω. Let $\triangle$ represent a generic triangle/tetrahedra in the mesh Ω_q and set $h_\triangle := \mathrm{diam}(\triangle)$. Following Eriksson and Johnson in [14, Remark 2.3] we let $N(\triangle)$ be the set of all triangles/tetrahedra sharing an edge/face with $\triangle$, and define a piecewise constant function in $L_2(\Omega)$ for each time slab $\mathcal{J}_q$ by,

$$D^2_{h,q}v\big|_\triangle := \max_{\triangledown\in N(\triangle)}\left\{\frac{\|\nabla v(\boldsymbol{P}_\triangle)-\nabla v(\boldsymbol{P}_\triangledown)\|_E}{\|\boldsymbol{P}_\triangle-\boldsymbol{P}_\triangledown\|_E}\right\}, \tag{52}$$

for all v that are continuous and piecewise linear with respect to the mesh Ω_q; where $\boldsymbol{P}_\triangle$ and $\boldsymbol{P}_\triangledown$ denote respectively the centres of gravity of $\triangle$ and $\triangledown$; and, where $\|\cdot\|_E$ denotes the Euclidean norm on $\mathbb{R}^n$. Hence, there exists $\triangle' \in N(\triangle)$ such that,

$$\frac{\|\nabla v(\boldsymbol{P}_\triangle)-\nabla v(\boldsymbol{P}_\triangledown)\|_E}{\|\boldsymbol{P}_\triangle-\boldsymbol{P}_\triangledown\|_E} \leqslant \frac{\|\nabla v(\boldsymbol{P}_\triangle)-\nabla v(\boldsymbol{P}_{\triangle'})\|_E}{\|\boldsymbol{P}_\triangle-\boldsymbol{P}_{\triangle'}\|_E} = \big|D^2_{h,q}v\big|_\triangle\big|,$$

for all $\triangledown \in N(\triangle)$. Now, let ℓ be the edge/face common to both $\triangle$ and $\triangledown$, then

$$\|\nabla v(\boldsymbol{P}_\triangle)-\nabla v(\boldsymbol{P}_\triangledown)\|_E = \|[\![\nabla v]\!]_\ell\|_E\,,$$

where $[\![\nabla v]\!]_\ell$ is the n-vector $([\![v_{,j}]\!]_\ell)_{j=1}^n$, and

$$\|[\![\nabla v]\!]_\ell\|_E \leqslant \left|D^2_{h,q}v\big|_\triangle\right| \max_{\ell\in\partial\triangle}\left\{\|\boldsymbol{P}_\triangle - \boldsymbol{P}_\triangledown\|_E\right\} \qquad \forall\triangledown \in N(\triangle).$$

We now assume that there exists a positive constant μ_q such that for all elements $\triangle$ in the mesh Ω_q, we have $h_\triangledown/h_\triangle \leqslant \mu_q$ for all $\triangledown \in N(\triangle)$. Then:

$$\|\boldsymbol{P}_\triangle - \boldsymbol{P}_\triangledown\|_E = O(h_\triangle + h_\triangledown) \leqslant (1+\mu_q)Ch_\triangle,$$

and from this we infer that,

$$\|[\![\nabla v]\!]_\ell\|_E \leqslant C(\mu_q)h_\triangle\left|D^2_{h,q}v\big|_\triangle\right| \qquad \forall\ell\in\partial\triangle, \tag{53}$$

for each $\triangle \in \Omega_q$ and for each time level $q \in \mathbb{N}(1,N)$.

Our first result is a straightforward bound on the jump terms $[\![\varepsilon_{ij}(\boldsymbol{U})]\!]_\ell$.

LEMMA 24. *Let* (53) *hold and let* $\{C_{kij}\}$ *be a spatially constant tensor, then*

$$|C_{kij}[\![\varepsilon_{ij}(\boldsymbol{U}(t))]\!]_\ell|^2 \leqslant C\sum_{i=1}^n h_\triangle^2\left|D^2_{h,q}U_i(t)\big|_\triangle\right|^2$$

for all $t\in\mathcal{J}_q$, *for all* $\triangle$ *in the mesh* Ω_q *and for all* $\ell\in\partial\triangle$. *The constant* C *depends only on* $\{C_{kij}\}$ *and* μ_q.

Proof. Using (6) we get

$$C_{kij}[\![\varepsilon_{ij}(\boldsymbol{U}(t))]\!]_\ell = \frac{C_{kij}}{2}[\![U_{i,j}]\!]_\ell + \frac{C_{kij}}{2}[\![U_{j,i}]\!]_\ell = \left(\frac{C_{kij}+C_{kji}}{2}\right)[\![U_{i,j}]\!]_\ell,$$

which gives,

$$|C_{kij}[\![\varepsilon_{ij}(\boldsymbol{U}(t))]\!]_\ell|^2 \leqslant C\sum_{i,j=1}^n |[\![U_{i,j}]\!]_\ell|^2 = C\sum_{i=1}^n \|[\![\nabla U_i]\!]_\ell\|_E^2,$$

and the lemma follows from (53). $\square$

We can now give the proof of Lemma 20.

Proof. **Of Lemma 20**. Since we are assuming that $\Gamma_N = \varnothing$ we have for the jump terms in $\mathcal{E}_\Omega(t;\boldsymbol{U}(t))$, for the particular element $\triangle$ in the mesh Ω_q that,

$$\|h_\triangle^{\frac12}\boldsymbol{r}(\boldsymbol{U};t)\|^2_{\boldsymbol{L}_2(\partial\triangle)} = \sum_{k=1}^n \|h_\triangle^{\frac12}r_k(\boldsymbol{U};t)\|^2_{\boldsymbol{L}_2(\partial\triangle)},$$

and for $t\in\mathcal{J}_q$,

$$\|h_\triangle^{\frac12}r_k(\boldsymbol{U};t)\|^2_{\boldsymbol{L}_2(\partial\triangle)} = h_\triangle\oint_{\partial\triangle}|r_k(\boldsymbol{U};t)|^2 d\Gamma = \frac{h_\triangle}{4}\sum_{\ell\in\partial\triangle}\int_\ell\left|[\![\underline{\boldsymbol{\sigma}}(\boldsymbol{U};t)\cdot\boldsymbol{\nu}^\ell]\!]_\ell\right|^2 d\ell,$$

where we recalled (43). We now have the following for $\Omega\subset\mathbb{R}^n$,

$$\begin{aligned}\|h_\triangle^{\frac12}r_k(\boldsymbol{U};t)\|^2_{\boldsymbol{L}_2(\partial\triangle)} &= \frac{h_\triangle}{4}\sum_{\ell\in\partial\triangle}\int_\ell\bigg|\boldsymbol{\nu}_l^\ell D_{klij}(0)[\![\varepsilon_{ij}(\boldsymbol{U}(t))]\!]_\ell \\ &\qquad\qquad - \int_0^t \boldsymbol{\nu}_l^\ell\frac{\partial D_{klij}(t-s)}{\partial s}[\![\varepsilon_{ij}(\boldsymbol{U}(s))]\!]_\ell\,ds\bigg|^2 d\ell, \\ &= \frac{ch_\triangle^n}{4}\sum_{\ell\in\partial\triangle}\left|d^{(0)}_{kij}[\![\varepsilon_{ij}(\boldsymbol{U}(t))]\!]_\ell - \int_0^t d^{(1)}_{kij}(t-s)[\![\varepsilon_{ij}(\boldsymbol{U}(s))]\!]_\ell\,ds\right|^2,\end{aligned}$$

where we used our strengthened assumption on $\underline{\boldsymbol{D}}(t)$ along with the fact that the strain jump $[\![\varepsilon_{ij}(\boldsymbol{U})]\!]_\ell$ is constant on each edge ℓ. From Lemma 24 we now obtain,

$$
\begin{aligned}
&\|h_\triangle^{\frac{1}{2}} r_k(\boldsymbol{U};t)\|^2_{\boldsymbol{L}_2(\partial\triangle)} \\
&\quad \leqslant C h_\triangle^n \sum_{\ell\in\partial\triangle} \left[\left| d^{(0)}_{kij} [\![\varepsilon_{ij}(\boldsymbol{U}(t))]\!]_\ell \right|^2 + t\int_0^t \left| d^{(1)}_{kij}(t-s) [\![\varepsilon_{ij}(\boldsymbol{U}(s))]\!]_\ell \right|^2 ds \right], \\
&\quad \leqslant C h_\triangle^n \sum_{\ell\in\partial\triangle} \sum_{i=1}^n \left[h_\triangle^2 |D^2_{h,q} U_i(t)|_\triangle|^2 + t\int_0^t h_\triangle^2 |D^2_{h,q} U_i(s)|_\triangle|^2\, ds \right], \\
&\quad \leqslant C h_\triangle^2 \left(\|D^2_{h,q}\boldsymbol{U}(t)\|^2_{\boldsymbol{L}_2(\triangle)} + t\int_0^t \|D^2_{h,q}\boldsymbol{U}(s)\|^2_{\boldsymbol{L}_2(\triangle)}\, ds \right),
\end{aligned}
$$

where C depends now on $\underline{\boldsymbol{D}}(t)$ and the geometry of $\triangle$. Performing the summation required in the definition of $\mathcal{E}_\Omega$ we now arrive at,

$$
\begin{aligned}
&\sum_{\Omega_{qj}\subset\Omega_q} \|h_{qj}^{\frac{1}{2}} \boldsymbol{r}(\boldsymbol{U};t)\|^2_{\boldsymbol{L}_2(\partial\Omega_{qj})} \\
&\quad \leqslant C \sum_{\Omega_{qj}\subset\Omega_q} \sum_{k=1}^n \left(\|h_{qj} D^2_{h,q}\boldsymbol{U}(t)\|^2_{\boldsymbol{L}_2(\Omega_{qj})} + t\int_0^t \|h_{qj} D^2_{h,q}\boldsymbol{U}(s)\|^2_{\boldsymbol{L}_2(\Omega_{qj})}\, ds \right), \\
&\quad \leqslant C \|h_q D^2_{h,q}\boldsymbol{U}(t)\|^2_{\boldsymbol{L}_2(\Omega)} + Ct\int_0^t \|h_q D^2_{h,q}\boldsymbol{U}(s)\|^2_{\boldsymbol{L}_2(\Omega)}\, ds.
\end{aligned}
$$

From this it follows firstly that,

$$
\begin{aligned}
&\left\| \left(\sum_{\Omega_{qj}\subset\overline{\Omega}_q} \|h_q^{\frac{1}{2}} \boldsymbol{r}(t;\boldsymbol{U}(t))\|^2_{\boldsymbol{L}_2(\partial\Omega_{qj})} \right)^{\frac{1}{2}} \right\|_{L_\infty(\mathcal{J}_q)} \\
&\qquad \leqslant C\|h_q D^2_{h,q}\boldsymbol{U}\|_{L_\infty(\mathcal{J}_q;\boldsymbol{L}_2(\Omega))} + Ct_q \|h_q D^2_{h,q}\boldsymbol{U}\|_{L_\infty(0,t_q;\boldsymbol{L}_2(\Omega))}, \\
&\qquad \leqslant C\|h_q D^2_{h,q}\boldsymbol{U}\|_{L_\infty(0,t_q;\boldsymbol{L}_2(\Omega))},
\end{aligned}
$$

and then, by our assumptions (i) and (ii),

$$
\begin{aligned}
\|h_q D^2_{h,q}\boldsymbol{U}\|_{\boldsymbol{L}_2(\Omega)} &\leqslant \|h_q D^2_{h,q}(\boldsymbol{U}-\pi_q\boldsymbol{u})\|_{\boldsymbol{L}_2(\Omega)} + \|h_q D^2_{h,q}\pi_q\boldsymbol{u}\|_{\boldsymbol{L}_2(\Omega)}, \\
&\leqslant C_q\|\boldsymbol{U}-\pi_q\boldsymbol{u}\|_H + C_\pi\|h_q D^2\boldsymbol{u}\|_{\boldsymbol{L}_2(\Omega)}, \\
&\leqslant C_q\|\boldsymbol{U}-\boldsymbol{u}\|_H + C_q\|\boldsymbol{u}-\pi_q\boldsymbol{u}\|_H + C_\pi\|h_q D^2\boldsymbol{u}\|_{\boldsymbol{L}_2(\Omega)}, \\
&\leqslant C_q\|\boldsymbol{U}-\boldsymbol{u}\|_H + C\|h_q D^2\boldsymbol{u}\|_{\boldsymbol{L}_2(\Omega)}.
\end{aligned}
$$

Finally, we use this estimate with Theorem 5 to get

$$
\begin{aligned}
&\left\| \left(\sum_{\Omega_{qj}\subset\overline{\Omega}_q} \|h_q^{\frac{1}{2}} \boldsymbol{r}(t;\boldsymbol{U}(t))\|^2_{\boldsymbol{L}_2(\partial\Omega_{qj})} \right)^{\frac{1}{2}} \right\|_{L_\infty(\mathcal{J}_q)} \\
&\qquad\qquad \leqslant C\left(\|hD^2\boldsymbol{u}\|_{L_\infty(\mathcal{J};\boldsymbol{L}_2(\Omega))} + \left\| k^{r+1}\frac{\partial^{r+1}\boldsymbol{u}}{\partial t^{r+1}} \right\|_{L_\infty(\mathcal{J};H)} \right),
\end{aligned}
$$

and Lemma 20 now follows from this. $\square$

REFERENCES

[1] M. AINSWORTH AND J. T. ODEN, *A posteriori error estimation in finite element analysis*, Comput. Methods Appl. Mech. Engrg., 142 (1997), pp. 1—88.

[2] M. ASADZADEH AND K. ERIKSSON, *On adaptive finite element methods for Fredholm integral equations of the second kind*, SIAM J. Numer. Anal., 31 (1994), pp. 831—855.

[3] A. K. AZIZ AND J.-L. LIU, *A Galerkin method for the forward-backward heat equation*, Math. Comp., 56 (1991), pp. 35—44.

[4] A. K. AZIZ AND P. MONK, *Continuous finite elements in space and time for the heat equation*, Math. Comp., 52 (1989), pp. 255—274.

[5] I. BABUŠKA AND T. JANIK, *The h-p version of the finite element method for parabolic equations. Part I: the p-version in time*, Numer. Methods Partial Differential Equations, 5 (1989), pp. 363—399.

[6] ———, *The h-p version of the finite element method for parabolic equations. Part II: the h-p version in time*, Numer. Methods Partial Differential Equations, 6 (1990), pp. 343—369.

[7] D. M. BEDIVAN AND G. J. FIX, *Least squares methods for Volterra equations and generalizations*. University of Texas at Arlington, Dept. Mathematics, Texas, 76019-0408.

[8] ———, *Analysis of finite element approximation and quadrature of Volterra integral equations*, Numer. Methods Partial Differential Equations, 13 (1997), pp. 663—672.

[9] M. BUCH, A. IDESMAN, R. NIEKAMP, AND E. STEIN, *Finite elements in space and time for parallel computing of viscoelastic deformation*, Computational Mechanics, 24 (1999), pp. 386—395.

[10] P. G. CIARLET, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

[11] M. DELFOUR, W. HAGER, AND F. TROCHU, *Discontinuous Galerkin methods for ordinary differential equations*, Math. Comp., 36 (1981), pp. 455—473.

[12] K. ERIKSSON, D. ESTEP, P. HANSBO, AND C. JOHNSON, *Introduction to adaptive methods for differential equations*, Acta Numerica, (1995), pp. 105—158. Cambridge University Press.

[13] ———, *Computational differential equations*, Cambridge University Press, 1996.

[14] K. ERIKSSON AND C. JOHNSON, *Adaptive finite element methods for parabolic problems. I: a linear model problem*, SIAM J. Numer. Anal., 28 (1991), pp. 43—77.

[15] ———, *Adaptive finite element methods for parabolic problems. II: optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal., 32 (1995), pp. 706—740.

[16] D. ESTEP, *A posteriori error bounds and global error control for approximation of ordinary differential equations*, SIAM J. Numer. Anal., 32 (1995), pp. 1—48.

[17] D. ESTEP AND D. FRENCH, *Global error control for the continuous Galerkin finite element method for ordinary differential equations*, M^2AN, 28 (1994), pp. 815—852.

[18] J. D. FERRY, *Viscoelastic properties of polymers*, John Wiley and Sons Inc., 1970.

[19] I. FRIED, *Finite element analysis of time dependent phenomena*, tech. rep., Institut für Statik und Dynamik Luft-Und Raumfahrtkonstruktionen, Universität Stuttgart, 1968.

[20] J. M. GOLDEN AND G. A. C. GRAHAM, *Boundary value problems in linear viscoelasticity*, Springer-Verlag, 1988.

[21] D. C. HANDSCOMBE, *Error of linear interpolation on a triangle*, tech. rep., Oxford University Computing Laboratory, 1995. Tech. Rep. NA 95/09.

[22] B. L. HULME, *Discrete Galerkin and related one-step methods for ordinary differential equations*, Math. Comp., 26 (1972), pp. 881—891.

[23] ———, *One-step piecewise polynomial Galerkin methods for initial value problems*, Math. Comp., 26 (1972), pp. 415—426.

[24] A. IDESMAN, R. NIEKAMP, AND E. STEIN, *Finite element methods in space and time for generalized viscoelastic Maxwell model*, Journal of Computational Mechanics, 27 (2001), pp. 49—60.

[25] P. JAMET, *Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain*, SIAM J. Numer. Anal., 15 (1978), pp. 912—928.

[26] A. R. JOHNSON, *Modeling viscoelastic materials using internal variables*, The Shock and Vibration Digest, 31 (1999), pp. 91—100.

[27] ———, *A viscoelastic hybrid shell finite element*, in Whiteman [47], pp. 87—96.

[28] A. R. JOHNSON AND A. TESSLER, *A viscoelastic high order beam finite element*, in The Mathematics of Finite Elements and Applications. MAFELAP 1996, J. R. Whiteman, ed., Wiley, Chichester, 1997, pp. 333—345.

[29] C. JOHNSON, *Discontinuous Galerkin finite element methods for second order hyperbolic problems*, Comp. Meth. Appl. Mech. Eng., 107 (1993), pp. 117—129.

[30] C. JOHNSON AND P. HANSBO, *Adaptive finite element methods in computational mechanics*, Comput. Methods Appl. Mech. Engrg., 101 (1992), pp. 143—181.

[31] F. J. LOCKETT, *Nonlinear viscoelastic solids*, Academic Press, 1972.

[32] M. J. LUDWIG, *Finite element error estimation and adaptivity for problems of elasticity*, PhD thesis, Brunel University, England, 1998. (See `www.brunel.ac.uk/~icsrbicm`).

[33] J. T. ODEN, *A general theory of finite elements II: Applications*, Internat. J. Numer. Methods. Engrg., 1 (1969), pp. 247—259.

[34] J. T. ODEN AND L. C. WELLFORD JR., *Discontinuous finite element approximations for the analysis of acceleration waves in elastic solids*, in The mathematics of finite elements and applications, II. MAFELAP, 1975., J. R. Whiteman, ed., Academic Press, London, 1976, pp. 269—284.

[35] ———, *A theory of discontinuous finite element Galerkin approximations of shock waves in nonlinear elastic solids. Part I: Variational theory*, Comput. Methods Appl. Mech. Engrg., 8 (1976), pp. 1—16.

[36] ———, *A theory of discontinuous finite element Galerkin approximations of shock waves in nonlinear elastic solids. Part II: Accuracy and convergence*, Comput. Methods Appl. Mech. Engrg., 8 (1976), pp. 17—36.

[37] R. RANNACHER, *The dual-weighted-residual method for error control and mesh adaption in finite element methods*, in Whiteman [47], pp. 97—116.

[38] L. R. SCOTT AND S. ZHANG, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483—493.

[39] S. SHAW AND J. R. WHITEMAN, *Discontinuous Galerkin method with a posteriori $L_p(0,t_i)$ error estimate for second-kind Volterra problems*, Numer. Math., 74 (1996), pp. 361—383.

[40] ———, *Applications and numerical analysis of partial differential Volterra equations: a brief survey*, Comput. Methods Appl. Mech. Engrg., 150 (1997), pp. 397—409.

[41] ———, *Negative norm error control for second-kind convolution Volterra equations*, Numer. Math., 85 (2000), pp. 329–341.

[42] ———, *Numerical solution of linear quasistatic hereditary viscoelasticity problems*, SIAM J. Numer. Anal, 38 (2000), pp. 80—97.

[43] ———, *Optimal long-time $L_p(0,T)$ data stability and semidiscrete error estimates for the Volterra formulation of the linear quasistatic viscoelasticity problem*, Numer. Math., 88 (2001), pp. 743—770. (BICOM Tech. Rep. 98/7 see: `www.brunel.ac.uk/~icsrbicm`).

[44] ———, *$L_p(0,t)$ error control using the derivative of the residual for a finite element approximation of a second-kind Volterra equation*. BICOM TR03/1, `www.brunel.ac.uk/~icsrbicm`, 2003.

[45] E. SÜLI AND P. HOUSTON, *Finite element methods for hyperbolic problems: a posteriori error analysis and adaptivity*, tech. rep., Numerical Analysis Group, Oxford University Computing Laboratory, OX1 3QD, England, 1996. 96/09.

[46] R. VERFÜRTH, *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Wiley-Teubner, 1996.

[47] J. R. WHITEMAN, ed., *The mathematics of finite elements and applications, X. MAFELAP 1999.*, Elsevier, Amsterdam, 2000.

[48] O. C. ZIENKIEWICZ, *A new look at the Newmark, Houbolt and other time stepping formulas. A weighted residual approach*, Earthquake engineering and structural dynamics, 5 (1977), pp. 413—418.