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

SIMON SHAW1 and J. R. WHITEMAN1

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 $L^2$-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 $2 \times 2$ 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 $\Omega$ 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-
tic compressible body occupies \( \Omega \) and is acted upon by a system of body forces \( \mathbf{f} := (f_i(x, t))_{i=1}^n \) for \( \mathbf{x} := (x_i)_{i=1}^n \in \Omega \) 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 \( \mathbf{g} := (g_i(x, t))_{i=1}^n \) for \( \mathbf{x} \in \Gamma_N \) and \( t \in \mathcal{J} \). The unit outward directed normal vector to \( \Gamma_N \) is denoted by \( \mathbf{n} := (\hat{n}_i)_{i=1}^n \).

We use the function \( \mathbf{u} = (u_i)_{i=1}^n : \Omega \times \mathcal{J} \to \mathbb{R}^n \) to describe the displacement from equilibrium resulting from the action of the applied forces \( \mathbf{f} \) and \( \mathbf{g} \). This is the linear theory, wherein we assume that \( \mathbf{u} \) is “small” so as not to violate the assumption that \( \Omega \) 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 \( \mathbf{u} \equiv \mathbf{0} \) 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 \( dG(p):G(p) \) method is then analysed in §6 where an \( a \text{ posteriori} \) 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 \( a \text{ priori} \) 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 \( a \text{ posteriori} \) estimate for the error \( e := u - U \). The weak form and finite element approximation are as follows:

\[
(Au, v) = (f, v) \quad \forall v \in V \\
(AU, v) = (f, v) \quad \forall v \in V^h.
\]

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

\[
||D^p\chi|| \leq S||g|| \quad \text{and} \quad ||h^{-p}(\chi - \pi\chi)|| \leq C||D^p\chi||,
\]

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

\[
||e||^2 = (e, e) = (e, A^*\chi) = (Ae, \chi) = (r, \chi - \pi\chi) = (h^p_r, h^{-p}(\chi - \pi\chi)),
\]

\[
\leq CS||h^p_r|| ||e||, \quad \Rightarrow \quad ||u - U|| \leq CS||h^p_r||.
\]

This \( a \text{ posteriori} \) 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 \( ||u - U|| \leq \text{TOL} \).
March 24, 2003

VISCOELASTIC A POSTERIORI ESTIMATES

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,

\[
\|D\chi\| \leq S\|Dg\|,
\]

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

\[
\|u - U\|_w = \sup \frac{|(u - U, g)|}{\|g\|} = \sup \frac{|(hr, h^{-1}(\chi - \pi\chi))|}{\|Dg\|} \leq CS\|hr\|.
\]

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 \( \pi \) denoting the \( L_2 \) projection, the template is:

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

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 representation.)

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
SIMON SHAW AND J. R. WHITEMAN

March 24, 2003

[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 dGeG method to model a parabolic problem with a (known) time-dependent boundary.

Also, Aziz and Monk have used cGeG 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,

\[ Au(t) = L(t) + \int_0^t B(t-s)u(s)\,ds, \]

(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 \textit{a posteriori} error estimate. On the other hand, for the temporal stability the best we can hope for is \( ||D^p u|| \leq S||D^p L|| \) (see theorem 2 below). This suggests that to build the time step into the \textit{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:

\[ ||u - U||_{W^{r,s}(0,t_i)} \leq S(t_i) \left\| k^{s+m} \frac{d^{m+1}r}{dt^{m+1}} \right\|_{L_p(0,t_i)} \text{ for } 0 \leq s, m \leq 1, \]

(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 \textit{et al.} formulate a parallel solver in [9, 24]. This work, as well as the Bedivan-Fix approach, presents global space-time, \textit{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 \( i \in N(1, n) := \{1, \ldots, n\} \) that,

\[ -\sigma_{ij,j} = f_i \text{ in } \Omega \times J, \]

(3)

\[ u_i = 0 \text{ in } \Gamma_D \times J, \]

(4)

\[ \sigma_{ij} n_j = g_i \text{ in } \Gamma_N \times J. \]

(5)
Here and throughout, repeated indices imply summation and \( \sigma \) := \((\sigma_{ij})_{i,j=1}^{n} \) is the symmetric stress tensor. In the linear theory one derives the (small) strain tensor \( \varepsilon := (\varepsilon_{ij})_{i,j=1}^{n} \) from the displacement field using the relations,

\[
\varepsilon_{ij}(u) := \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right).
\]  

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

\[
\sigma_{ij}(u; x, t) = D_{ijkl}(x, 0)\varepsilon_{kl}(u(x, t)) - \int_{0}^{t} \frac{\partial D_{ijkl}(x, t - s)}{\partial s}\varepsilon_{kl}(u(x, s)) \, ds.
\]  

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

\[
D_{ijkl}(t) = D_{jikl}(t) = D_{ijlk}(t) \quad \text{but, in general,} \quad D_{ijkl}(t) \neq D_{klij}(t).
\]  

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 \( \mathbf{x} \) dependence. Also, the components of \( D \) can be assumed to be (a.e. in \( \Omega \)) functions of \( t \) that are smooth enough for their first time derivatives to be of class \( L^1(\mathcal{J}) \). In addition, since \( \mathbf{D}(0) \) measures instantaneous linear elastic response we follow Hooke's law and assume positive-definiteness:

\[
\gamma_{ij}\gamma_{kl}D_{ijkl}(0) > 0 \quad \text{a.e. in } \Omega
\]

for all non-zero symmetric second order tensors \( \gamma \).

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

\[
\sigma(u; t) = \sigma^E(u(t)) - \sigma^V(u; t, t),
\]

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

\[
\sigma_{ij}^{E}(u(t)) := D_{ijkl}(0)\varepsilon_{kl}(u(t)),
\]

\[
\sigma_{ij}^{V}(u; \tau, t) := \int_{0}^{\tau} \frac{\partial D_{ijkl}(t - s)}{\partial s}\varepsilon_{kl}(u(s)) \, ds \quad \text{for } 0 \leq \tau \leq t.
\]

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, \( H^{s}(\Omega) := H^{s}(\Omega)^n \), for \( s = 0, 1, 2, \ldots \), with inner products given by \( (w, v)_{s} := \sum_{i,j=1}^{n} (w_{ij}, v_{ij})_{H^{s}(\Omega)} \), for all \( w, v \in H^{s}(\Omega) \). These spaces have the natural norms \( \| \cdot \|_{s} := \sqrt{\langle \cdot, \cdot \rangle_{s}} \) and, of course, \( L_{2}(\Omega) \equiv H^{0}(\Omega) \). 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 \( \|v\|_{L_{p}(0, t; \mathcal{B})} := \| \|v(\cdot)\|_{\mathcal{B}} \|_{L_{p}(0, t)} \).

We also use the (symmetric second order) tensor-valued \( L_{2} \) space,

\[
L_{2}(\Omega) := \left\{ \gamma = (\gamma_{ij})_{i,j=1}^{n} : \gamma_{ij} = \gamma_{ji} \in L_{2}(\Omega) \forall i, j \in \mathbb{N}(1, n) \right\}.
\]

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

\[
H := \left\{ v \in H^{1}(\Omega) : v = 0 \text{ on } \Gamma_{D} \right\}.
\]
and (see e.g. [42] for details) arrive at the weak problem: find \( u \in L_p(\mathcal{J}; H) \) such that,

\[
A(u(t), v) = \langle L(t), v \rangle + \int_0^t B(t - s; u(s), v) \, ds \quad \forall v \in H, \text{ a.e. in } \mathcal{J}.
\]

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

\[
A(w, v) := \int_\Omega D_{ijkl}(0) \varepsilon_{kl}(w) \varepsilon_{ij}(v) \, d\Omega,
\]

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

for all \( w, v \in H \), and \( L : \mathcal{J} \to \text{“dual space”} \) is the time dependent linear form defined through,

\[
\langle L(t), v \rangle := \int_\Omega v \cdot f(t) \, d\Omega + \int_{\Gamma_N} v \cdot g(t) \, d\Gamma \quad \forall v \in H.
\]

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 \( \partial_t^n \), and also use primes as shorthand so that, for example, \( v'(t) = \partial_t v(t) \), \( v''(t) = \partial_t^2 v(t) \), \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 \( m \leq n \), and define \( \mathbb{N}(m, n) := \emptyset \) if \( m > n \).

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

\[
\| u \|_{W^m_p(a, b; H)} \leq (b - a)^{m-n} \| u \|_{W^m_p(a, b; H)} \quad \forall u \in W^m_p(a, b; H).
\]

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

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 \( r \geq 0 \) such that the following hold.

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

\[
|A(w, v)| \leq C\|w\|_H \|v\|_H, \quad \text{and} \quad A(v, v) \geq c\|v\|_H^2,
\]

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

\[
\|v\|_H := \sqrt{A(v, v)}
\]
for all \( \mathbf{v} \in H \), which (on \( H \)) is equivalent to the norm \( \| \cdot \|_1 \). Henceforth, by \( H \) we shall mean the subspace of \( H^1(\Omega) \) 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 \( \mathbf{Q} \) satisfies \( D_{ijkl} \in W^{r+1}_1(J; L_{\infty}(\Omega)) \). Then, the bilinear form \( B(t, \cdot, \cdot) \) is continuous and similar to \( A(\cdot, \cdot) \) in the sense that there exists \( \varphi \in L_1(J; [0, \infty)) \) such that

\[
\| B(t, \mathbf{w}, \mathbf{v}) \| \leq \varphi(t) \| \mathbf{w} \|_H \| \mathbf{v} \|_H.
\]
a.e. in \( J \) and for all \( \mathbf{w}, \mathbf{v} \in H \).

(iii) The linear form \( L \in W^r_p(J; H') \). Then,

\[
|\partial_t^s L(t, \mathbf{v})| \leq \| \partial_t^s L(t) \|_{H'} \| \mathbf{v} \|_H \quad \text{for } 0 \leq s \leq r,
\]
a.e. in \( J \) and for all \( \mathbf{v} \in H \).

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 \( f \in L_p(J; L_2(\Omega)) \) and \( g \in L_p(J; L_2(\Gamma_N)) \), 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 \textit{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 \( r = 0 \).

\textbf{Theorem 2 (data stability).} Let Assumptions 1 hold for some \( r \geq 0 \) with, in addition, \( L \in W^r_p(J; H') \). Then there exists a stability factor \( S : J \rightarrow [0, \infty) \) such that,

\[
\| \partial_t^s \mathbf{u} \|_{L_p(0,t; H)} \leq S(t) \| \partial_t^s L \|_{L_p(0,t; H')}
\]

for all \( t \in J \).

\textit{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,

\[
A(\mathbf{u}'(t), \mathbf{v}) = \langle L'(t), \mathbf{v} \rangle + \int_0^t B(t - s, \mathbf{u}'(s), \mathbf{v}) ds \quad \forall \mathbf{v} \in H.
\]

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. \( \square \)

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

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

$$\| \cdot \|_{W_p^{-m}(a,b;H')} := \sup \left\{ \int_a^b \langle \cdot, \mathbf{v}(t) \rangle dt : \mathbf{v} \in \dot{W}_q^m(a,b;H) \right\}, \quad \| \mathbf{v} \|_{W_p^m(a,b;H)} = 1. \quad (18)$$

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 $\mathcal{R}$; $H \rightarrow H'$ defined via the Riesz Representation Theorem through,

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

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 $m \geq 0$, via,

$$\| \cdot \|_{W_p^{-m}(a,b;H)} := \| \mathcal{R} \cdot \|_{W_p^{-m}(a,b;H')} = \sup \left\{ \int_a^b A(\cdot, \mathbf{v}(t)) dt : \mathbf{v} \in \dot{W}_q^m(a,b;H) \right\}, \quad \| \mathbf{v} \|_{\dot{W}_q^m(a,b;H)} = 1. \quad (20)$$

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 $\mathbf{u} \in L_p(\mathcal{J}; H)$ such that,

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

where $q$ is again conjugate to $p$ and,

$$a(\mathbf{w}, \mathbf{v}) := \int_0^T A(\mathbf{w}(t), \mathbf{v}(t)) dt - \int_0^T \int_0^t B(t - s; \mathbf{w}(s), \mathbf{v}(t)) ds dt, \quad (22)$$

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

for all $\mathbf{w} \in L_p(\mathcal{J}; H)$ and $\mathbf{v} \in L_q(\mathcal{J}; H)$. 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 $\chi \in L_q(\mathcal{J}; H)$ such that,

$$a^*(\chi, \mathbf{v}) = l^*(\mathbf{v}) \quad \forall \mathbf{v} \in L_p(\mathcal{J}; H), \quad (24)$$
where (compare (21)–(23)):

\[ a^*(\chi, v) := \int_0^T A(v(t), \chi(t)) \, dt - \int_0^T \int_t^T B(s - t; v(t), \chi(s)) \, ds \, dt, \]

\[ l^*(v) := \int_0^T \langle L^*(t), v(t) \rangle \, dt, \]  

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

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,

\[ ||\partial^m\chi||_{L_2(t,T;H)} \leq S(T-t)||\partial^m L^*||_{L_2(t,T;H)} \quad \forall m \in \mathbb{N}(0,r) \text{ and } \forall t \in \mathcal{J}. \]  

(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 \( \overline{\Omega} \) (in the usual way) a triangular/tetrahedral space-mesh of \( M_i \) elements and denote the domain \( \Omega \) with this mesh by \( \Omega_i \). Element \( j \) of \( \Omega_i \) will be denoted \( \Omega_{ij} \) (an open set) and we set,

\[ h_{ij} := \text{diam}(\Omega_{ij}), \]

and use \( h \) to denote the piecewise constant (on \( \Omega \times \mathcal{J} \) ) mesh function given by \( h|_{\Omega_j} := h_{ij} \). We also use \( h_t \) to denote the mesh function at times \( t \in \mathcal{J}_i \) given by \( h_t|_{\Omega_{ij}} := h_{ij} \), 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 \( \Omega_i \), the semidiscrete (spatial) finite element space,

\[ H_i := \left\{ v \in H \cap (C(\Omega))^n : v|_{\Omega_{ij}} \in P_{p}(\Omega_{ij})^n \text{ for each } \Omega_{ij} \subset \Omega_i \right\}. \]  

(28)

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

\[ \begin{align*}
V_r^i & := P_r(\mathcal{J}_i; H_i), \\
V_r & := \left\{ v \in L_{\infty}(\mathcal{J}; H) : v|_{\mathcal{J}_i} \in V_r^i \quad \forall i \in \mathbb{N}(1,N) \right\}.
\end{align*} \]

Here \( P_r(\mathcal{J}_i; H_i) \) 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 $\ell$ in the mesh on $\Omega$, we associate a unit normal $\mathbf{u}^\ell$. It is immaterial in which direction this points on internal edges/faces, but for $\ell \subset \partial \Omega$ we always take $\mathbf{u}^\ell = \mathbf{n}$, 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 $J_i$, by

\[
\|v\|_{L^2(\partial \Omega)} := \left( \sum_{\alpha_i \subset \partial \Omega} \|v\|^2_{L^2(\alpha_i)} \right)^{\frac{1}{2}},
\]

\[
\|v\|_{L^2(\bigcup \partial \Omega)} := \left( \sum_{\alpha_i \subset \partial \Omega} \|v\|^2_{L^2(\alpha_i)} \right)^{\frac{1}{2}}.
\]

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

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

(i) $f \in L^p(J; L^2(\Omega))$ and $g \in L^p(J; L^2(\Gamma_N))$, and the mesh can and will always be constructed so as to respect the Dirichlet–Neumann interface $\Gamma_D \cap \Gamma_N$.

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

(iii) Every space mesh is nondegenerate in that every element $\Omega_{ij}$ contains (resp. is contained by) a ball of radius $r_i$ (resp. $r_0$), and the ratio $r_i/r_0$ is bounded.

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

\[
\|\pi_i v\|_{L^2(\Omega)} \leq \Pi_0 \|v\|_{L^2(\Omega)}, \quad \text{and} \quad \|\pi_i v\|_H \leq \Pi_1 \|v\|_H,
\]

as well as the following error estimates:

\[
\|h^{-1}(w - \pi_i w)\|_{L^2(\Omega)} \leq \Pi_\Omega \|w\|_H,
\]

\[
\|h^{-\frac{1}{2}}(w - \pi_i w)\|_{L^2(\bigcup \partial \Omega)} \leq \Pi_\ell \|w\|_H,
\]

for all $w \in H$ and where $\Pi_0, \Pi_1, \Pi_\Omega$ and $\Pi_\ell$ are constants that are independent of $h$ and $w$, and which we assume can be taken independent of $i$ (i.e., of time). Occasionally we will also use the “global interpolator” $\pi$ defined piecewise by $\pi|_{J_i} = \pi_i$.

Note that the interpolation estimates do not require excessive regularity of $w$. For example, in [38, Theorems 3.1 and 4.1, and Equation 5.5] such interpolators are defined for “rough” functions $w \in H$ 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 $P_\Omega$ and $P_\Gamma$ be defined piecewise (on each $J_i$) by $P_{\Omega i} : L^2(\Omega) \to H_i$ and $P_{\Gamma i} : L^2(\Gamma_N) \to H_i|\Gamma_N$ where:

\[
(P_{\Omega i} w, v)_{L^2(\Omega)} = (w, v)_{L^2(\Omega)} \quad \forall w \in L^2(\Omega), v \in H_i,
\]

\[
(P_{\Gamma i} w, v)_{L^2(\Gamma_N)} = (w, v)_{L^2(\Gamma_N)} \quad \forall w \in L^2(\Gamma_N), v \in H_i.
\]

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 \( \mathbf{U} \in V_r \) such that,
\[
a(\mathbf{U}, \mathbf{v}) = l(\mathbf{v}) \quad \forall \mathbf{v} \in V_r,
\]
and subtracting this from (21) gives the fundamentally important Galerkin “orthogonality” relationship:
\[
a(\mathbf{u} - \mathbf{U}, \mathbf{v}) = 0 \quad \forall \mathbf{v} \in V_r.
\]
We recall from [42] the \textit{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 \( \Omega \):
- \( -\mathcal{D}_{ijkl}(t)\gamma_{ij}\gamma_{kl} \geq c_0\mathcal{D}_{ijkl}(0)\gamma_{ij}\gamma_{kl} \) for some constant \( c_0 = c_0(K) > 0 \), for all \( t \in [0, K] \) and all \( \gamma \in L_p(\Omega) \);
- no component \( \mathcal{D}_{ijkl}(t) \) 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 \( \mathbf{e} := \mathbf{u} - \mathbf{U} \) satisfies the \textit{a priori} error estimate,
\[
|\mathbf{u} - \mathbf{U}|_{L_\infty(\mathcal{J}; H)} \leq C(T) \left( \Pi_H \| hD^2 u \|_{L_\infty(\mathcal{J}; L_2(\Omega))} + \Pi_k \| k^{r+1} \mathbf{e}^{r+1} \|_{L_\infty(\mathcal{J}; H)} \right).
\]
Here \( C(T) \) is a discrete stability factor and \( \Pi_H, \Pi_k \) are constants. This estimate holds for \( r = 1 \) only if each \( k_i \) is small enough.

Our interest here is to generate an \textit{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”,
\[
H_{\mathcal{J}} := \left\{ \mathbf{v} \in L_q(\mathcal{J}; H) : \mathbf{v}|_{\mathcal{J}_i} \in L_q(\mathcal{J}_i; H_i) \forall i \in \mathbb{N}(1, N) \right\},
\]
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 \( \chi \) of the dual problem we now define \( \mathbf{\theta} \in L_q(\mathcal{J}; H) \) and \( \mathbf{\rho} \in H_{\mathcal{J}} \) in the following piecewise manner:
\[
\mathbf{\theta}|_{\mathcal{J}_i} \equiv \mathbf{\theta}_i := \mathbf{\chi} - \pi_i \mathbf{\chi} \in L_q(\mathcal{J}_i; H_i),
\]
\[
\mathbf{\rho}|_{\mathcal{J}_i} \equiv \mathbf{\rho}_i := \pi_i \mathbf{\chi} - P_r \pi_i \mathbf{\chi} \in L_q(\mathcal{J}_i; H_i),
\]
and with these we obtain a basic error representation formula.

**Lemma 6 (Error representation).** The Galerkin error \( \mathbf{e} := \mathbf{u} - \mathbf{U} \) satisfies
\[
\mathbf{l}^*(\mathbf{e}) = G(\mathbf{\theta}) + G(\mathbf{\rho}),
\]
where \( G(\cdot) := \| \cdot \| - a(\cdot, \cdot) \).

\textit{Proof.} In the dual problem (24) we take \( \mathbf{v} = \mathbf{e} \in L_p(\mathcal{J}; H) \), and use Lemma 3 with equations (37) and (21) to get:
\[
\mathbf{l}^*(\mathbf{e}) = a^*(\mathbf{\chi}, \mathbf{e}) = a(\mathbf{e}, \mathbf{\chi}) = a(\mathbf{e}, \mathbf{\chi} - P_r \pi \mathbf{\chi}) = a(\mathbf{u} - \mathbf{U}, \mathbf{\chi} - P_r \pi \mathbf{\chi}) = \mathbf{l}(\mathbf{\chi} - P_r \pi \mathbf{\chi}) - a(\mathbf{U}, \mathbf{\chi} - P_r \pi \mathbf{\chi}),
\]

since \( P_r \pi \mathbf{\chi} \in V_r \). Now, \( \mathbf{\chi} - P_r \pi \mathbf{\chi} = \mathbf{\theta} + \mathbf{\rho} \), and the result follows from the linearity of \( \mathbf{l}(\cdot) \) and \( a(\mathbf{U}, \cdot) \). \( \Box \)
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, \( \mathbf{U} \), into (7) or (equivalently) (9) to get, for \( t \in \mathcal{J}_t \),

\[
\mathbf{a}(\mathbf{U} ; t) = \mathbf{a}^V(\mathbf{U}(t)) - \int_{t_{i-1}}^t \mathbf{D}_i(t-s) \mathbf{G}(\mathbf{U}(s)) \, ds - \mathbf{a}^V(\mathbf{U} ; t_{i-1} , t) .
\]

These terms are spatially discontinuous and (if \( \mathbf{D} \) 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 5c-projection to represent the stress history on the current mesh by defining a new version of \( \mathbf{a}^V \) as \( \tilde{\mathbf{a}}^V \) via,

\[
(\tilde{\mathbf{a}}^V(\mathbf{U} ; t_{i-1} , t) , \mathbf{z})_{L^2(\Omega)} = (\mathbf{a}^V(\mathbf{U} ; t_{i-1} , t) , \mathbf{z})_{L^2(\Omega)} \tag{41}
\]

for all piecewise continuous tensor-valued degree \( p - 1 \) polynomials, \( \mathbf{z} \in L^2(\Omega) \).

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

\[
[\gamma_k(x)]_\ell := \pm \lim_{\varepsilon \to 0} (\gamma_k(x - \varepsilon \mathbf{u}^f) - \gamma_k(x + \varepsilon \mathbf{u}^f)) \quad \forall x \in \ell . \tag{42}
\]

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

\[
(\mathbf{w} , \mathbf{v})_\ell := \int_\ell \mathbf{w} \cdot \mathbf{v} \, d\ell \quad \text{with} \quad \| \cdot \|_\ell := \sqrt{(\cdot , \cdot)_\ell} .
\]

Note that the \( L^2 \) requirement on the viscous stress in the following lemma uses Assumption 4, part (i), and requires the stress projection described earlier in (41).

**Lemma 7 (spatial residual).** Suppose during \( \mathcal{J}_t \), for all \( i \in \mathbb{N}(1,N) \), that the discrete stress, defined by using \( \mathbf{U} \) in (9), satisfies \( \nabla \cdot \mathbf{a}(\mathbf{U} ; t)|_{\Omega_{in}} \in L^2(\Omega_{in}) \) for each \( \Omega_{in} \subset \Omega_t \) and
for every \( t \in J_t \). Then,

\[
G(\theta) = \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} \left( \sum_{\Omega_{ij} \subset \Omega_t} \int_{\Omega_{ij}} [f(t) + \nabla \cdot \sigma(U; t) \cdot \theta(t)] d\Omega + \oint_{\Gamma_N} [g(t) - \sigma(U; t) \cdot \vec{n}] \cdot \theta(t) d\Gamma + \sum_{\ell \subset \Gamma_t} \int_{\ell} [\sigma(U; t) \cdot \nu^{\ell} \cdot \theta(t)] d\ell \right) dt.
\]

Here, \( \nabla \cdot \sigma \) denotes the vector \((\sigma_{ij})_{p-1}^m \).

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

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

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

For each time level we define \( r|_{\Omega_i} = (r_k)_{k=1}^n \) by,

\[
r(U; t) := \begin{cases} 
\frac{1}{2} \left\| \sigma(U; t) \cdot \nu^{\ell} \right\|, & \text{for } \ell \subset \Omega_t, \\
\left| g(t) - \sigma(U; t) \cdot \vec{n} \right|, & \text{on } \Gamma_N, \\
0, & \text{on } \Gamma_D.
\end{cases}
\]

(43)

Recall that each \( \Omega_t \) is open, and so “\( \ell \subset \Omega_t \)” refers to all internal edges/faces.

Also, define \( \tilde{\sigma}(t) \in L_2(\Omega) \) by,

\[
\tilde{\sigma}(t)_{|_{\Omega_{ij}}} := \frac{\| r(U; t) \|_{L_2(\Omega_{ij})}}{\sqrt{h_{ij} \text{meas}(\Omega_{ij})}}
\]

(44)

for all \( \Omega_{ij} \subset \Omega_t \). Then,

\[
\left( \sum_{\Omega_{ij} \subset \Omega_t} h_{ij}^{\frac{1}{2}} \| r(U; t) \|_{L_2(\Omega_{ij})}^{2} \right) \frac{1}{\tilde{\sigma}} = \| h_{ij} \tilde{\sigma}(t) \|_{L_2(\Omega)}. \tag{45}
\]

Now we can state the bound.

**Lemma 8.** Let Assumptions 1 and 4 hold, and suppose that \( \chi|_{[t_1; t]} = 0 \) in the dual problem (24). Then,

\[
|G(\theta)| \leq E_{\Omega}(p, t_1; U) \| \chi \|_{L_q(0, t_1; H)},
\]

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

\[
E_{\Omega}(p, t_1; U) := \left( \sum_{i=1}^t \Pi_{\Omega_i} \left\| h_{ij} \| f + \nabla \cdot \sigma(U; \cdot) \|_{L_2(\Omega_i)} + \Pi_{\ell} \| h_{ij} \tilde{\sigma}(t) \|_{L_2(\ell)} \right\|^p \right)^{\frac{1}{p}},
\]

where \( t \) is the number of time steps.
while for $q = 1$,

$$
\mathcal{E}_2(\infty, t_i; U) := \max_{1 \leq i \leq t} \left\{ \left\| \Pi_{L_{1}} (f + \nabla \cdot \mathbf{g}(U; \cdot)) \right\|_{L_2(\Gamma_i)} + \Pi_{L_{1}} \| h_i \partial_t \|_{L_2(\Gamma_i)} \right\}.
$$

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

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

$$
\left| \sum_{\eta_n \subset \Omega_i} \int_{\Omega_i} \left[ f(t) + \nabla \cdot \mathbf{g}(U; t) \cdot \theta(t) \right] d\Omega \right|
\leq \sum_{\Omega_i \subset \mathcal{J}_i} \left\| h_{in} \left[ f(t) + \nabla \cdot \mathbf{g}(U; t) \right] \right\|_{L_2(\Omega_i)} \left\| h_{in}^{\frac{1}{2}} \theta(t) \right\|_{L_2(\Omega_i)},
$$

$$
\leq \Pi_{L_{1}} \left\| h_i \left[ f(t) + \nabla \cdot \mathbf{g}(U; t) \right] \right\|_{L_2(\Omega_i)} \left\| \chi(t) \right\|_{H},
$$

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

$$
\left| \sum_{\sigma_i \subset \mathcal{J}_i} \int_{\sigma_i} \left[ g(t) - \mathbf{g}(U; \cdot) \cdot \hat{n} \right] \cdot \theta(t) d\ell + \sum_{\sigma_i \subset \mathcal{J}_i} \int_{\sigma_i} \left[ \mathbf{g}(U; t) \cdot \nu^t \right] \cdot \theta(t) d\ell \right|
\leq \sum_{\sigma_i \subset \mathcal{J}_i} \int_{\sigma_i} \left| g(t) - \mathbf{g}(U; t) \cdot \hat{n} \right| | \theta(t) | d\ell + \sum_{\sigma_i \subset \mathcal{J}_i} \left\| \mathbf{g}(U; t) \cdot \nu^t \right\| | \theta(t) | d\ell,
$$

$$
= \sum_{\sigma_i \subset \mathcal{J}_i} \int_{\sigma_i} h_{in}^{\frac{1}{2}} | r(U; t) | h_{in}^{-\frac{1}{2}} | \theta(t) | d\ell,
$$

$$
\leq \Pi_{L_{1}} \left\| h_i \hat{\theta}(t) \right\|_{L_2(\Omega_i)} \left\| \chi(t) \right\|_{H},
$$

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 $\mathcal{E}_2$ 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”, $G(\rho)$, 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 $V_i^t$, 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 $P_r : L_1(\mathcal{J}_i; \mathcal{H}) \to P_r(\mathcal{J}_i; \mathcal{H})$ by,

$$
\int_{t_{i-1}}^{t_i} \left( w(t) - P_r w(t) \right) \phi(t) dt = 0 \quad \forall w \in L_1(\mathcal{J}_i; \mathcal{H}) \text{ and } \forall \phi \in P_r(\mathcal{J}_i).
$$

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 $H$, $L_2(\Gamma_N)$ and $L_2(\Omega)$. 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 $a \leq b$,
\[
\left\| \int_a^b v(t) \, dt \right\|_{\mathcal{H}} \leq \int_a^b \| v(t) \|_{\mathcal{H}} \, dt \quad \forall v \in L_1(a, b; \mathcal{H}),
\]
where we used the equality,
\[
\left( \int_{t_{i-1}}^{t_i} w(t) \, dt, v \right)_{\mathcal{H}} = \int_{t_{i-1}}^{t_i} (w(t), v)_{\mathcal{H}} \, dt \quad \forall w \in L_1(\mathcal{J}_i; \mathcal{H}), \, v \in \mathcal{H}.
\]

Our first estimates for this projection are concerned with stability in $L_2(\mathcal{J}_i; \mathcal{H})$ 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,
\[
\| v \|_{L_p(\mathcal{J}_i; \mathcal{H})} \leq C_r(p, q) k_i^{\frac{1}{p} - \frac{1}{2}} \| v \|_{L_q(\mathcal{J}_i; \mathcal{H})} \quad \forall v \in \mathbb{P}_r(\mathcal{J}_i),
\]
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 $r = 0, 1$,
\begin{enumerate}
  \item $\int_{t_{i-1}}^{t_i} (w(t) - P_r w(t), v)_{\mathcal{H}} \, dt = 0 \quad \forall w \in L_1(\mathcal{J}_i; \mathcal{H})$ and $v \in \mathbb{P}_r(\mathcal{J}_i)$,
  \item $\| P_r w \|_{L_q(\mathcal{J}_i; \mathcal{H})} \leq \| w \|_{L_q(\mathcal{J}_i; \mathcal{H})} \quad \forall w \in L_q(\mathcal{J}_i; \mathcal{H}),$
\end{enumerate}
where $\Xi_{q_0} = 1$ and $\Xi_q = C_1(q, 2) C_2 \frac{1}{q - 1}$, where the latter constants are from (48).

**Proof.** To prove (i) we use (46) to get,
\[
\int_{t_{i-1}}^{t_i} (w(t) - P_r w(t), v(\phi(t))_{\mathcal{H}} \, dt = \left( \int_{t_{i-1}}^{t_i} (w(t) - P_r w(t)) \phi(t) \, dt, v \right)_{\mathcal{H}} = 0,
\]
for all $\phi \in \mathbb{P}_r(\mathcal{J}_i)$ and for all $v \in \mathcal{H}$. This is (i) since the elements of $\mathbb{P}_r(\mathcal{J}_i)$ are linear combinations of the form $\sum q_i v_i$. 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 $v = \| P_0 w \|_{\mathcal{H}}^{-2} P_0 w$ in (i) to get,
\[
\| P_0 w \|_{L_q(\mathcal{J}_i; \mathcal{H})} \leq \| w \|_{L_q(\mathcal{J}_i; \mathcal{H})} \| P_0 w \|_{L_q(\mathcal{J}_i; \mathcal{H})}^{-1},
\]
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 $q$,
\[
\frac{k_i^{1/p - 1}}{C_1(q, 2) C_2} \| P_1 w \|_{L_q(\mathcal{J}_i; \mathcal{H})} \leq \| P_1 w \|_{L_2(\mathcal{J}_i; \mathcal{H})},
\]
\[
\leq \| w \|_{L_q(\mathcal{J}_i; \mathcal{H})} \| P_1 w \|_{L_p(\mathcal{J}_i; \mathcal{H})},
\]
\[
\leq C_1(p, q) k_i^{\frac{1}{p} - \frac{1}{2}} \| w \|_{L_q(\mathcal{J}_i; \mathcal{H})} \| P_1 w \|_{L_q(\mathcal{J}_i; \mathcal{H})}.
\]
Now, using $p = q/(q - 1)$ and $1/p - 1/q = 1 - 2/q$ then gives,

$$\|P_t w\|_{L^q(J_i; \mathcal{C})} \leq C_1 (q, 2)^2 C_1 \left(\frac{2}{q-1}, q\right) \|w\|_{L^q(J_i; \mathcal{C})},$$

as required.

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

**Lemma 12 (P₁-strong stability).** For $q \in [1, \infty]$,

$$\|\partial_t P_t w\|_{L^q(J_i; \mathcal{C})} \leq \frac{3}{2} \|\partial w\|_{L^q(J_i; \mathcal{C})} \quad \forall w \in W^1_q(J_i; \mathcal{C}).$$

**Proof.** From Lemma 11 we have,

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

for all $v \in P_t(J_i; \mathcal{C})$ and for all $w \in W^1_q(J_i; \mathcal{C})$. Obviously we can assume that $\partial_t P_t w \neq 0$ and so, for $q \in [1, \infty)$, we choose $v = 12k_i^2 (t - t_{i-1/2})/2 \partial_t P_t w \partial_t P_t w$, where $t_{i-1/2} := (t_{i-1} + t_i)/2$. (Note that $v \in P_t(J_i; \mathcal{C})$ because $\partial_t P_t w$ 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,

$$\int_{t_{i-1}}^{t_i} (t - t_{i-1/2}) (P_t w, \partial_t P_t w)_{\mathcal{C}} dt$$

$$= \frac{12}{k_i^4} \|\partial_t P_t w\|_{L^2(\mathcal{C})} \left[ \int_{t_{i-1}}^{t_i} (s - t_{i-1/2}) ds \right]_{t_{i-1}}^{t_i}$$

$$- \frac{12}{k_i^4} \|\partial_t P_t w\|_{L^2(\mathcal{C})} \left[ \int_{t_{i-1}}^{t_i} \left( \int_{t_{i-1}}^{t} (s - t_{i-1/2}) ds \right) \partial_t w, \partial_t P_t w \right]_{\mathcal{C}} dt,$$

$$= k_i \|\partial_t P_t w\|_{L^2(\mathcal{C})}.$$
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, \( \partial_t P_t w \) is constant on \( J_t \).

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

**Lemma 13.** For all \( w \in L_4(J_t; \mathcal{H}) \):

\[
\begin{align*}
(i) & \quad \int_{t_{i-1}}^{t_i} w(t) - P_t w(t) \, dt = 0 \quad \text{in } \mathcal{H} \quad \text{for } r = 0, 1, \\
(ii) & \quad \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} w(s) - P_t w(s) \, ds \, dt = 0 \quad \text{in } \mathcal{H},
\end{align*}
\]

for each time interval \( J_t \).

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

\[
0 = \int_{t_{i-1}}^{t_i} (w - P_t w) \phi \, dt = \phi(t) \int_{t_{i-1}}^{t} (w - P_t w) \, ds \bigg|_{t_{i-1}}^{t_i} - \int_{t_{i-1}}^{t_i} \phi'(t) \int_{t_{i-1}}^{t} (w - P_t w) \, ds \, dt.
\]

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

We now need error estimates for \( P_t \). As an aid in deriving these we invoke the following well known mean value theorem: if \( \phi \) is given by Lemma 11;

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

**Lemma 14** (\( P_t \)-error estimates). For \( r = 0 \) or 1:

\[
\begin{align*}
(i) & \quad \| w - P_t w \|_{L_4(J_t; \mathcal{H})} \leq (1 + \mathcal{Z}_{4r}) \| w \|_{L_4(J_t; \mathcal{H})}, \\
(ii) & \quad \| k_i^{t_{i-1}}(w - P_t w) \|_{L_4(J_t; \mathcal{H})} \leq \gamma_{t_{i-1}} \| \partial_t w \|_{L_4(J_t; \mathcal{H})} \quad \text{for } 0 \leq s \leq r + 1,
\end{align*}
\]

where \( \gamma_{t_{i-1}} = \gamma_{t_{i-2}} = 1, \gamma_{t_{i-1}} = \frac{1}{2}, \) and \( \gamma_{t_{i-1}} = (1 + \mathcal{Z}_{4r}) \).

**Proof.** The proof of \( (i) \) follows from the triangle inequality and Lemma 11 and the proof of \( (ii) \) when \( r = 0, s = 1 \) follows from this and Lemma 11.

To prove \( (ii) \) when \( r = 0, s = 1 \) we have that \( R_0 w \) is a constant (in time) and so solving (46) for \( R_0 w \) and using the fundamental theorem of calculus gives,

\[
w(t) - R_0 w(t) = w(t) - \frac{1}{k_i} \int_{t_{i-1}}^{t_i} \left[ w(t) + \int_t^{r} w'(\xi) \, d\xi \right] \, d\tau = - \frac{1}{k_i} \int_{t_{i-1}}^{t_i} \int_t^{r} w'(\xi) \, d\xi \, d\tau.
\]

Using (47) in this we obtain,

\[
\| w(t) - R_0 w(t) \|_{\mathcal{H}} \leq \int_{t_{i-1}}^{t_i} \| w'(s) \|_{\mathcal{H}} \, ds,
\]

and the result follows by taking the \( L_4(J_t) \) norm of both sides and using Hölder’s inequality.
For $s - 1 = r = 1$ we observe that $w \in C^1(J_\delta; H)$ and so, using (i) in Lemma 13 with the Mean Value Theorem, we get a $\tau_i \in (t_{i-1}, t_i)$ such that $w(\tau_i) = R_i w(t_{i-1})$. Now, set $t_{i-1}/2 := (t_{i-1} + t_i)/2$ and choose $\phi(t) = t - t_{i-1}/2 \in P_1(J_\delta)$ in (46). Integrating by parts, and noting that both boundary terms vanish, then gives,

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

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

$$Z_{t_{i-1}}^{t_i} \phi(t) \, dt,$$

such that

$$0 = \int_{t_{i-1}}^{t_i} (w(t) - R_i w(t)) = \int_{t_{i-1}}^{t_i} [w(\tau) - P_i w(\tau)] \, d\tau,$$

with the Mean Value Theorem, we get a

$$Z_{t_{i-1}}^{t_i} [w(\tau) - P_i w(\tau)] \, d\tau,$$

such that

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

since $P_i w \in P_1(J_\delta; H)$. Hence, using (47) again with Hölder’s inequality yields,

$$\|w(t) - R_i w(t)\|_{H^1} \leq k_i^{-1+\frac{1}{2}} \|w''\|_{L^1(J_{i;3})}.$$  

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

$$\|w(t) - P_i w(t)\|_{H^1} \leq \|\partial_t w\|_{L^1(J_{i;3})} + \|\partial_t P_i w\|_{L^1(J_{i;3})} \leq \frac{5}{2} k_i^{-\frac{1}{2}} \|\partial_t w\|_{L^1(J_{i;3})}.$$  

Now take the $L_q(J_{\delta})$ norm of both sides. \(\Box\)

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

$$G(\rho) = \sum_{i=1}^{N} \int_{t_i - 1}^{t_i} (f, \rho)_{L_2(\Omega)} + (g, \rho)_{L_2(\Gamma_N)} - (\xi(U; t), \xi(\rho))_{L_2(\Omega)} \, dt.$$  

We have three unlike terms and so we cannot combine them in a simple manner like 

$$\mathbf{f} + g - \mathbf{a}$$

to form a residual.

We can form a “proper” residual during $J_\delta$ by seeking a pseudo displacement $\tilde{W}_0 \in H_i$ such that $\delta \tilde{W}(t) = (\tilde{L}(t), v)$ for all $v \in H_i$. Then, using (10) we can define an equivalent stress $\xi^P(\tilde{W}(t))$. The “proper” residual is then $\xi^P(\tilde{W}(t)) - \xi(U; t)$. 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)),

\[
G(\rho) = \sum_{i=1}^{N} \int_{t_{i-1}}^{t_{i}} (P_{\alpha_{i}}f - P_{\rho}P_{\alpha_{i}}f, \rho)_{L_{2}(\Omega)} + \int_{t_{i-1}}^{t_{i}} (P_{\alpha_{i}}g - P_{\rho}P_{\alpha_{i}}g, \rho)_{L_{2}(\Gamma_{N})} - (\kappa(U, t) - P_{r}\kappa(U, t), \kappa(\rho))_{L_{2}(\Omega)} dt,
\]

where for \( w = (w_{i})_{i}^{n} \) we define \( P_{r}w \) in the natural way by \( (P_{r}w)_{i} := P_{r}w_{i} \). In the above, the inclusion of \( P_{\alpha_{i}} \) and \( P_{\rho} \) arises from orthogonality and is, therefore, optional.

Proof. Recalling from (40) that \( \rho = (I - P_{r})\pi_{1}\chi \in H_{i} \) on each \( J_{i} \) we (optionally) use (34) and obtain,

\[
\int_{t_{i-1}}^{t_{i}} (f, \rho)_{L_{2}(\Omega)} dt = \int_{t_{i-1}}^{t_{i}} (P_{\alpha_{i}}f - P_{\rho}P_{\alpha_{i}}f, \rho)_{L_{2}(\Omega)} dt,
\]

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 \( J_{i} \) to: a.e. in \( J_{i} \),

\[
|\partial^{s}(\pi_{1}w)|_{H} \leq \Pi_{1} |\partial^{s}w|_{H} \quad \forall w \in W_{s}^{r}(J_{i}; H_{i}) \quad \text{and for} \quad 0 \leq s \leq r + 1.
\]

Also, assume that \( \chi|_{(t_{i}, t_{i+1})} = 0 \) in the dual problem (24). Then,

\[
|G(\rho)| \leq \mathcal{E}_{J}(p, s, t_{i}; U)|\partial^{s} \chi|_{L_{2}(0, t_{i}; H)}
\]

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

\[
\mathcal{E}_{J}(p, s, t_{i}; U) := \left( \sum_{i=1}^{I} \Pi_{1}^{p} |\pi_{1}^{p}|^{2} \int_{t_{i-1}}^{t_{i}} \left| C_{H}^{i} \left| P_{\alpha_{i}}f(t) - P_{\rho}P_{\alpha_{i}}f(t) \right|_{L_{2}(\Omega)} + C_{T} \left| P_{\alpha_{i}}g(t) - P_{\rho}P_{\alpha_{i}}g(t) \right|_{L_{2}(\Gamma_{N})} \right|^{p} \right)^{\frac{1}{p}},
\]

and,

\[
\mathcal{E}_{J}(\infty, s, t_{i}; U) := \max_{1 \leq i \leq I} \left\{ \sum_{i=1}^{I} \Pi_{1} |\pi_{1}^{p}|^{2} \int_{t_{i-1}}^{t_{i}} \left| C_{H}^{i} \left| P_{\alpha_{i}}f(t) - P_{\rho}P_{\alpha_{i}}f(t) \right|_{L_{2}(\Omega)} + C_{T} \left| P_{\alpha_{i}}g(t) - P_{\rho}P_{\alpha_{i}}g(t) \right|_{L_{2}(\Gamma_{N})} \right| \right\},
\]

where, in these: the use of \( P_{\alpha_{i}} \) and \( P_{\rho} \) is optional; the constant \( \Pi_{1} \) is given above; \( \Pi_{1} \) comes from Lemma 14; and, \( C_{H}^{i}, C_{T}^{i} \) are defined for each \( J_{i} \) by,

\[
\begin{align*}
C_{H}^{i} & \text{ such that } |\kappa|_{L_{2}(\Omega)} \leq C_{H}^{i} |\kappa|_{H} \quad \forall \kappa \in H_{i}, \\
C_{T}^{i} & \text{ such that } |\partial^{s} \kappa|_{L_{2}(\Gamma_{N})} \leq C_{T}^{i} |\partial^{s} \kappa|_{H} \quad \forall \kappa \in H_{i}.
\end{align*}
\]
\(Pf\). We use Lemma 15. Firstly, note that the positive-definiteness (over symmetric second-order tensors) and symmetry of \(D(0)\) gives,

\[
|\langle \mathfrak{a} - P_r \mathfrak{a}, \mathfrak{a}(\rho) \rangle|_{L^2(\Omega)} = |\langle D^{-\frac{1}{2}}(0)(\mathfrak{a} - P_r \mathfrak{a}), D^{-\frac{1}{2}}(0)\mathfrak{a}(\rho) \rangle|_{L^2(\Omega)} \\
\leq \|D^{-\frac{1}{2}}(0)(\mathfrak{a} - P_r \mathfrak{a})\|_{L^2(\Omega)} \|\rho\|_H.
\]

Now, in each \(J_t\) we have \(\rho \in H_t\) for a.e. \(t \in J_t\), and so using the estimate above we have,

\[
\left| (P_{\alpha_1} f - P_r P_{\alpha_1} f, \rho)_{L^2(\Omega)} + (P_{\alpha_1} g - P_r P_{\alpha_1} g, \rho)_{L^2(\Omega_N)} \right| \\
\leq k_f \left( C_H^1 \|P_{\alpha_1} f - P_r P_{\alpha_1} f\|_{L^2(\Omega)} + C_T^1 \|P_{\alpha_1} g - P_r P_{\alpha_1} g\|_{L^2(\Omega)} \\
+ \|D^{-\frac{1}{2}}(0)(\mathfrak{a} - P_r \mathfrak{a})\|_{L^2(\Omega)} \right) \times k^{-s} \|\rho\|_H.
\]

Hölders inequality for integrals and then sums now yields,

\[
|G(\rho)| \leq \tilde{E}_{\alpha}(p, t; U) \times \|\nabla^3 \|k^{-s}(I - P_r) \pi \chi\|_{L^q(0, t; H)} ,
\]

and then Lemma 14 and our assumption on the interpolators \(\pi_i\) give,

\[
\|\nabla^3 \|k^{-s}(I - P_r) \pi \chi\|_{L^q(0, t; H)} \leq \Pi^{-1} \|\partial_t^\beta(\pi \chi)\|_{L^q(0, t; H)} \leq \|\partial_t^\beta \chi\|_{L^q(0, t; H)}.
\]

This completes the proof. \(\square\)

**Remark 17.** The constant \(C_H^1\) can be estimated by the square root of the least eigenvalue of the problem,

\[
A(w, v) = \lambda(w, v)_{L^2(\Omega)} \quad \forall v \in H_t,
\]

and \(C_T^1\) by the square root of the least eigenvalue of the problem,

\[
A(w, v) = \lambda(w, v)_{L^2(\Omega_N)} \quad \forall v \in H_t.
\]

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 \textit{a posteriori} Galerkin-error estimate.

6.4. \textit{A posteriori} error estimates. We begin with the theorem.

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

\[
\|u - U\|_{W^{-s}(0, t; H)} \leq S(t_j) \left( t_j^2 E_\Omega(p, t_j; U) + E_{\alpha}(p, s, t_j; U) \right),
\]

where \(S(t_j)\) is the stability factor from Theorem 2. Here, \(E_\Omega(\cdot, \cdot; \cdot)\) is from Lemma 8 and,

\[
E_{\alpha}(p, s, t_j; U) := \Pi_{\alpha} \Pi_1 \left( C_H^1 \|k^{s+m} P_{\alpha} \pi \chi f\|_{L^p(0, t_j; L^2(\Omega))} \\
+ C_T^1 \|k^{s+m} P_{\alpha} \pi \chi g\|_{L^p(0, t_j; L^2(\Omega_N))} \\
+ \|k^{s+m} D^{-\frac{1}{2}}(0) \pi \chi \mathfrak{a}(U; \cdot)\|_{L^p(0, t_j; L^2(\Omega))} \right).
\]

\[
E_{\alpha}(p, s, t_j; U) := \Pi_{\alpha} \Pi_1 \left( C_H^1 \|k^{s+m} P_{\alpha} \pi \chi f\|_{L^p(0, t_j; L^2(\Omega))} \\
+ C_T^1 \|k^{s+m} P_{\alpha} \pi \chi g\|_{L^p(0, t_j; L^2(\Omega_N))} \\
+ \|k^{s+m} D^{-\frac{1}{2}}(0) \pi \chi \mathfrak{a}(U; \cdot)\|_{L^p(0, t_j; L^2(\Omega))} \right).
\]
where \( \tau_{rm} \) and \( \tau_{rs} \) come from Lemma 14 and the inclusion of \( P_{\alpha} \) and \( P_{\tau} \) is optional.

Proof. In the dual problem, (24), take \( v = e := u - U \in L_p(J; H) \) and, for some \( s \in \mathbb{N}(0, r + 1) \) choose \( L^* \in \dot{W}_q^s(J; H') \) such that \( L^*[t_j, \tau] = 0 \) for an arbitrary time level \( t_j \). Then, by the stability estimate (27), we have \( \chi|_{[t_j, \tau]} = 0 \) and, moreover, by shifting—in the dual problem—the final time \( \tau \) backward to \( t_j \):

\[
\|\chi\|_{\dot{W}_q^s(0, t_j; H)} \leq S(t_j)\|L^*\|_{\dot{W}_q^s(0, t_j; H')},
\]

Lemmas 6, 8 and 16 give,

\[
|l^*(e)| \leq \mathcal{E}_\alpha(p, t_j; U)\|\chi\|_{L_q(0, t_j; H)} + \mathcal{E}_J(p, s, t_j; U)\|\chi\|_{\dot{W}_q^s(0, t_j; H)},
\]

and, by the equivalence of norms on \( \dot{W}_q^s(0, t_j; H) \), and the stability estimate given above,

\[
|l^*(e)| \leq S(t_j) \left( t_j^s \mathcal{E}_\alpha(p, t_j; U) + \mathcal{E}_J(p, s, t_j; U) \right)\|L^*\|_{\dot{W}_q^s(0, t_j; H')}.
\]

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

\[
\|e\|_{\dot{W}_q^s(0, t_j; H')} = \sup \left\{ \left\| \int_0^{t_j} \langle R, v \rangle \, dt \right\| : v \in \dot{W}_q^s(0, t_j; H) \right\} = S(t_j) \left( t_j^s \mathcal{E}_\alpha(p, t_j; U) + \mathcal{E}_J(p, s, t_j; U) \right)\|L^*\|_{\dot{W}_q^s(0, t_j; H')}.
\]

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

\[
\|P_{\alpha} f - P_{\alpha} \|_{L_p(J; L_2(\Omega))} + \|P_{\tau} g - P_{\tau} \|_{L_p(J; L_2(\Omega))} + \|Q_{\alpha}^* g - P_{\alpha} \|_{L_p(J; L_2(\Omega))} + \|Q_{\tau}^* g - P_{\tau} \|_{L_p(J; L_2(\Omega))} \leq \tau_{rm} k^m(\|Q_{\alpha}^* P_{\alpha} f \|_{L_p(J; L_2(\Omega))} + \|Q_{\tau}^* P_{\tau} g \|_{L_p(J; L_2(\Omega))} + \|Q_{\alpha}^* P_{\alpha} \|_{L_p(J; L_2(\Omega))} + \|Q_{\tau}^* P_{\tau} \|_{L_p(J; L_2(\Omega))}).
\]

The proof is completed by using each of (34) and (35) to obtain “\( \partial_1 P = P \partial_1 \)”. For example, applying \( \partial_1^m \) to both sides of (34) and taking the projection of \( \partial_1^m w \) gives:

\[
(\partial_1^m P_{\alpha}, w) = (\partial_1^m w, v) = (P_{\alpha}, \partial_1^m w, v)
\]

for all \( v \in H_i \). Choosing \( v = \partial_1^m P_{\alpha} w \) and using the Cauchy-Schwartz inequality then completes the proof. \( \square \)

Remark 19. In Theorem 18 we could take \( P_{\tau} = P_{\alpha} = \text{identity} \) 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 \( \Gamma_N = \emptyset \), so we have a Dirichlet problem. Our method of proof (for the \( \mathcal{E}_C \)) 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 \( \ell \) case would require a detailed stability analysis in order to obtain bounds on \( U' \): it is certain that this will introduce extra conditions on the data and time step.)

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

1. \( \| f_q \|_{L^2(\Omega)} + \| u_q - \pi_Q u \|_H \leq C \| f_q \|_{L^2(\Omega)} \)
2. \( \| f_q \|_{L^2(\Omega)} \leq C \| u \|_H \forall u \in H_q \),

where \( f_q \) is the discrete “second derivative” defined in (52). Then, there exists a positive constant \( C \) such that,

\[
\mathcal{E}_C(t; U(t)) \leq C \left( \| \mathbf{f} \|_{L^\infty(J; L^2(\Omega))} + \| \mathbf{D}^2 u \|_{L^\infty(J; L^2(\Omega))} + \left\| \frac{\partial^m \mathbf{f}}{\partial t^m} \right\|_{L^\infty(J; H)} \right).
\]

In the above \( m := r + 1 \) for approximation using the space \( V_r \).

The \( \mathcal{E}_C(\cdot) \) 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” \( \mathcal{E}_T \) as given in Theorem 18, since these are non-standard. The first two terms (involving \( \mathbf{f} \) and \( g \)) 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,

\[
\| U \|_{L^p(0,t; H)} \leq C_p \| L \|_{L^p(0,t; H^*)},
\]

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,

- \( \mathbf{D}^{-\frac{1}{2}}(0), \mathbf{D}'(0) \in L^\infty(\Omega) \) and \( \mathbf{D}'(0) \in L^p(J; L^\infty(\Omega)) \);
- The discrete stability estimate, (50), holds for some \( p \).
where \( p \in [1, \infty] \) and \( q \) are conjugate H"older indices. Then, there are constants \( C > 0 \) such that for \( p \in [1, \infty) \),
\[
\| k^{s+1} D^{-\frac{1}{2}}(0) \partial_t \sigma(U; \cdot) \|_{L_p(0,t; L_2(\Omega))}
\leq C \max_{1 \lesssim j \lesssim i} \left\{ k_j^{s+1} \left( \| f \|_{L_p(0,t; L_2(\Omega))} + \| g \|_{L_p(0,t; L_2(\Gamma_N))} \right) \right\},
\]
while for \( p = \infty \),
\[
\| k^{s+1} D^{-\frac{1}{2}}(0) \partial_t \sigma(U; \cdot) \|_{L_{\infty}(0,t; L_2(\Omega))}
\leq C \max_{1 \lesssim j \lesssim i} \left\{ k_j^{s+1} \left( \| f \|_{L_{\infty}(0,t; L_2(\Omega))} + \| g \|_{L_{\infty}(0,t; L_2(\Gamma_N))} \right) \right\}.
\]
Both of these hold for all \( i \in \mathbb{N}(1, N) \).

Proof. We obtain the discrete viscous stress, \( \sigma(U; t) \) by inserting \( U \) into (7), and from this we find \( \partial_t \sigma(U; t) \). In a given \( J_j \), this is,
\[
\partial_t \sigma(U; t) = \mathcal{D}(0) \epsilon(U'(t)) + \mathcal{D}'(0) \epsilon(U(t)) + \int_0^t \mathcal{D}_{st}(t - s) \epsilon(U(s)) \, ds,
\]
and (because \( r = 0 \)) we have \( U'(t) = 0 \). For ease of exposition it is convenient here to think of \( \sigma \) and \( \epsilon \) as vectors (not tensors), and \( \mathcal{D} \) as a matrix. The Euclidean norm will then be denoted \( \| \cdot \|_{\mathbb{R}} \). Hence,
\[
\| \partial_t \sigma(U; t) \|_{\mathbb{R}} \leq \| \mathcal{D}'(0) \|_{\mathbb{R}} \| \epsilon(U(t)) \|_{\mathbb{R}} + \int_0^t \| \mathcal{D}''(t - s) \|_{\mathbb{R}} \| \epsilon(U(s)) \|_{\mathbb{R}} \, ds.
\]
Taking \( L_2(\Omega) \) norms then gives,
\[
\| \partial_t \sigma(U; t) \|_{L_2(\Omega)} \leq C \| \mathcal{D}'(0) \|_{L_2(\Omega)} \| U(t) \|_{H} + C \int_0^t \| \mathcal{D}''(t - s) \|_{L_2(\Omega)} \| U(s) \|_{H} \, ds,
\]
for some constant \( C = C(\mathcal{D}) > 0 \) and where in both of these results we used variants of (47). Multiplying by \( k_j^{s+1} \) and taking \( L_p(J_j) \) norms now gives,
\[
\| k_j^{s+1} \partial_t \sigma(U; \cdot) \|_{L_p(J_j; L_2(\Omega))} \leq C k_j^{s+1} \left( \| \mathcal{D}'(0) \|_{L_2(\Omega)} \| U \|_{L_p(J_j; H)} + k_j^{\frac{1}{2}} \| \mathcal{D}'' \|_{L_q(J_j; L_2(\Omega))} \| U \|_{L_p(J_j; H)} \right).
\]
Now, absorbing the terms in \( \mathcal{D} \) into a generic constant, \( C \), we have,
\[
\| k^{s+1} D^{-\frac{1}{2}}(0) \partial_t \sigma(U; \cdot) \|_{L_p(0,t; L_2(\Omega))} \leq C \sum_{j=1}^i \| k_j^{s+1} \partial_t \sigma(U; \cdot) \|_{L_p(J_j; L_2(\Omega))},
\]
\[
\leq C \sum_{j=1}^i \left( \| U \|_{L_p(J_j; H)} + k_j^{\frac{1}{2}} \| U \|_{L_p(0,t; H)} \right)^p,
\]
\[
\leq 2^p C \sum_{j=1}^i \left( \| U \|_{L_p(J_j; H)} + k_j^{\frac{1}{2}} \| U \|_{L_p(0,t; H)} \right)^p,
\]
\[
\leq 2^p C \max_{1 \lesssim j \lesssim i} \left\{ \left( \| U \|_{L_p(0,t; H)} + t \| U \|_{L_p(0,t; H)} \right) \right\}^p,
\]
\[
\leq 2^p C \max_{1 \lesssim j \lesssim i} \left\{ \left( \| U \|_{L_p(0,t; H)} + t \| U \|_{L_p(0,t; H)} \right) \right\}^p.
\]
from the discrete stability estimate, (50). Finally we use,

\[ \|L\|_{L^p(0,t;H^j)} = \sup_{\theta \in H\setminus \{0\}} \frac{\|L(t,\theta)\|}{\|\theta\|_H} \leq C \|f\|_{L^2(\omega)} + C_\gamma \|g\|_{L^2(T_N)} \leq C(\|f\|_{L^p(0,t;L^2(\omega))} + \|g\|_{L^p(0,t;L^2(T_N))}), \]

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

For \( p = \infty \) we have,

\[ \|k^{s+1} L^{-\frac{1}{2}}(0) \partial_0 \varphi(U; \cdot)\|_{L^\infty(0,t;L^2(\omega))} \leq C \max_{1 \leq j \leq t} \left\{ \|k^{s+1} \partial_j \varphi(U; \cdot)\|_{L^\infty(\mathcal{J}_j;L^2(\omega))}, \right\}, \]

\[ \leq C \max_{1 \leq j \leq t} \left\{ k^{s+1} \|L\|_{L^\infty(0,t;H^j)} \right\}, \]

\[ \leq C \max_{1 \leq j \leq t} \left\{ k^{s+1} \left( \|f\|_{L^\infty(0,t;L^2(\omega))} + \|g\|_{L^\infty(0,t;L^2(T_N))} \right) \right\}, \]

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 \( \text{dG}(0)cG(1) \) scheme, under the previously indicated assumptions, there exists a constant \( C > 0 \) such that,

\[ \mathcal{E}_{\mathcal{J}}(\infty, s, 1, t; U) \leq C \left( \max_{1 \leq j \leq t} \left\{ k^{s+1} \left( \|f\|_{L^\infty(0,t;L^2(\omega))} + \|g\|_{L^\infty(0,t;L^2(T_N))} \right) \right\} \right), \]

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

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

\[ \mathcal{E}_{\mathcal{J}}(\infty, s, 1, t; U) \leq C_{\text{dis}} \left( \max_{1 \leq j \leq t} \left\{ k^{s+1} \|\partial_0 f\|_{L^\infty(\mathcal{J}_j;L^2(\omega))} \right\} \right), \]

since the projectors \( P_\alpha \) and \( P_{\Gamma} \) 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 $E_\tau$ as defined in Theorem 18 and $E_{\mathcal{T}}$ as defined in Lemma 8 we have for the $dG(0)/cG(1)$ approximation that,

\[
\|u - U\|_{W^{m,\infty}(0,t_i; H)} \leq S(t_i) \left( t_i^s E_\tau(\infty, t_i; U) + E_\tau(\infty, s, 1, t_i; U) \right),
\]

\[
= O\left( \|h\|_{L_{\infty}(\Omega \times \mathcal{T})} + \|k^{s+1}\|_{L_{\infty}(\mathcal{T})} \right),
\]

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 $\mathcal{D}$ 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 $\mathcal{D}$ 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 $f$ and $g$.

**$L_2(\Omega)$ estimator.** For problems in which $\partial \Omega$ is smooth and/or $\Omega$ is convex-polygonal, and the data are smooth, it makes sense to seek $L_2(\Omega)$ 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 $\mathcal{T}_n$, recall that we use $\Omega_\Delta$ to denote the space mesh on $\Omega$. Let $\Delta$ represent a generic triangle/tetrahedra in the mesh $\Omega_\Delta$ and set $h_\Delta := \text{diam}(\Delta)$. Following Eriksson and Johnson in [14, Remark 2.3] we let $N(\Delta)$ be the set of all triangles/tetrahedra sharing an edge/face with $\Delta$, and define a piecewise constant function in $L_2(\Omega)$ for each time slab $J_n$ by,

\[
D^\Delta_{h_\Delta} v \big|_\Delta := \max_{\nabla \in N(\Delta)} \left\{ \frac{\|\nabla v(P_\Delta) - k v(P_\nabla)\|_E}{\|P_\Delta - P_\nabla\|_E} \right\},
\]

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

\[
\frac{\|\nabla v(P_\Delta) - \nabla v(P_\nabla)\|_E}{\|P_\Delta - P_\nabla\|_E} \leq \frac{\|\nabla v(P_\Delta) - \nabla v(P_{\Delta'})\|_E}{\|P_\Delta - P_{\Delta'}\|_E} = \left| D^\Delta_{h_\Delta} v \right|_{\Delta'},
\]

for all $\nabla \in N(\Delta)$. Now, let $\ell$ be the edge/face common to both $\Delta$ and $\nabla$, then

\[
\|\nabla v(P_\Delta) - \nabla v(P_\nabla)\|_E = \|\nabla v\|_E,
\]
where $[\nabla v]_\ell$ is the $n$-vector $([v_{ij}]_\ell)_{j=1}^n$, and
\[
||[\nabla v]_\ell||_E \leq ||D_{h,q}^2 v|_\Delta|_{\ell} \max_{\ell \in \partial \Delta} \left\{ ||P_\Delta - P_{\partial \Delta}|_E \right\} \quad \forall \ell \in N(\Delta).
\]

We now assume that there exists a positive constant $\mu_q$ such that for all elements $\Delta$ in the mesh $\Omega_q$, we have $h_{\partial \Delta} / h_{\Delta} \leq \mu_q$ for all $\ell \in N(\Delta)$. Then:
\[
||P_\Delta - P_{\partial \Delta}|_E = O(h_{\Delta} + h_{\partial \Delta}) \leq (1 + \mu_q)ch_\Delta,
\]

and from this we infer that,
\[
||[\nabla v]_\ell||_E \leq C(\mu_q)h_\Delta \max_{\ell \in \partial \Delta} \left\{ ||P_\Delta - P_{\partial \Delta}|_E \right\} \quad \forall \ell \in \partial \Delta,
\]

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

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

**Lemma 24.** Let (53) hold and let $\{C_{kij}\}$ be a spatially constant tensor, then
\[
|\varepsilon_{ij}(U(t))|_\ell^2 \leq C \sum_{i=1}^n ||U_i||_E^2 \leq C \sum_{i=1}^n ||\nabla U_i||^2_E,
\]

for all $t \in T_q$, for all $\Delta$ in the mesh $\Omega_q$ and for all $\ell \in \partial \Delta$. The constant $C$ depends only on $\{C_{kij}\}$ and $\mu_q$.

*Proof.* Using (6) we get
\[
C_{kij}[\varepsilon_{ij}(U(t))]_\ell = \frac{C_{kij}}{2} [U_{ij}]_\ell + \frac{C_{kij}}{2} [U_{ji}]_\ell = \left( \frac{C_{kij} + C_{kji}}{2} \right) [U_{ij}]_\ell,
\]

which gives,
\[
\varepsilon_{ij}(U(t))|_\ell^2 \leq C \sum_{i=1}^n ||U_{ij}|_\ell^2 = C \sum_{i=1}^n ||\nabla U_i|_E^2,
\]

and the lemma follows from (53). \qed

We can now give the proof of Lemma 20.

*Proof of Lemma 20.* Since we are assuming that $\Gamma_N \neq \emptyset$ we have for the jump terms in $\varepsilon_{ij}(t; U(t))$, for the particular element $\Delta$ in the mesh $\Omega_q$ that,
\[
||h_{\Delta}^\frac{1}{2} r(U; t)||_{L^2(\partial \Delta)}^2 = \sum_{k=1}^n ||h_{\Delta}^\frac{1}{2} r_k(U; t)||_{L^2(\partial \Delta)}^2,
\]

and for $t \in T_q$,
\[
||h_{\Delta}^\frac{1}{2} r_k(U; t)||_{L^2(\partial \Delta)}^2 = h_{\Delta} \int_{\partial \Delta} |r_k(U; t)|^2 d\Gamma - \frac{h_{\Delta}}{4} \sum_{\ell \in \partial \Delta} \int_{\ell} \left( ||\varepsilon_{ij}(U; t) \cdot \nu^\ell||^2 d\ell + \right)
\]

where we recalled (43). We now have the following for $\Omega \subset \mathbb{R}^n$,
\[
||h_{\Delta}^\frac{1}{2} r_k(U; t)||_{L^2(\partial \Delta)}^2 = \frac{h_{\Delta}}{4} \sum_{\ell \in \partial \Delta} \int_{\ell} \nu^\ell D_{kij}(0)[\varepsilon_{ij}(U(t))]_\ell^2 - \int_0^t \left( D_{kij}(t-s) \frac{\partial [\varepsilon_{ij}(U(s))]_\ell}{\partial s} \right)^2 ds \]

\[
= \frac{c h_{\Delta}^n}{4} \sum_{\ell \in \partial \Delta} \int_{\ell} \left( D_{kij}(0)[\varepsilon_{ij}(U(t))]_\ell - \int_0^t D_{kij}(t-s) [\varepsilon_{ij}(U(s))]_\ell ds \right)^2 ds.
\]
where we used our strengthened assumption on $D(t)$ along with the fact that the strain jump $\varepsilon_{ij}(U)$ is constant on each edge $e$. From Lemma 24 we now obtain,

$$\|h_{\Delta}^{\frac{1}{2}} r_k(U; t)\|_{L^2(e)}^2 \leq C h_{\Delta}^2 \sum_{e \in \partial \Delta} \left( \int_{t=0}^t \|\varepsilon_{ij}(U(t))\|^2 dt \right)^2 + t \int_{t=0}^t \|\varepsilon_{ij}(U(s))\|^2 ds,
$$

$$\leq C h_{\Delta}^2 \sum_{e \in \partial \Delta} \sum_{i=1}^n \left( h_{\Delta}^2 \|D_{h,q}^2 U_i(t)\|_{L^2(e)}^2 + t \int_{t=0}^t h_{\Delta}^2 \|D_{h,q}^2 U_i(s)\|_{L^2(e)}^2 ds \right),
$$

$$\leq C h_{\Delta}^2 \left( \|D_{h,q}^2 U(t)\|_{L^2(\Delta)}^2 + t \int_{t=0}^t \|D_{h,q}^2 U(s)\|_{L^2(\Delta)}^2 ds \right),$$

where $C$ depends now on $D(t)$ and the geometry of $\Delta$. Performing the summation required in the definition of $c_2$ we now arrive at,

$$\sum_{\Omega_{q,j} \subset \Omega_q} \|h_{\Delta}^{\frac{1}{2}} r(U; t)\|_{L^2(\partial \Omega_{q,j})}^2 \leq C \sum_{\Omega_{q,j} \subset \Omega_q} \sum_{k=1}^n \left( \|h_{\Delta} D_{h,q}^2 U(t)\|_{L^2(\Omega_{q,j})}^2 + t \int_{t=0}^t \|h_{\Delta} D_{h,q}^2 U(s)\|_{L^2(\Omega_{q,j})}^2 ds \right),
$$

$$\leq C \|h_{\Delta} D_{h,q}^2 U(t)\|_{L^2(\Omega)}^2 + C t \int_{t=0}^t \|h_{\Delta} D_{h,q}^2 U(s)\|_{L^2(\Omega)}^2 ds.$$ 

From this it follows firstly that,

$$\left( \sum_{\Omega_{q,j} \subset \Omega_q} \|h_{\Delta}^{\frac{1}{2}} r(t; U(t))\|_{L^2(\partial \Omega_{q,j})} \right)^{\frac{1}{2}} \leq C \|h_{\Delta} D_{h,q}^2 U\|_{L^2({\mathcal J}^q)} + C t \|D_{h,q}^2 U\|_{L^2({\mathcal J}^q)},
$$

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

$$\|h_{\Delta} D_{h,q}^2 U\|_{L^2(\Omega)} \leq \|h_{\Delta} D_{h,q}^2 (U - \pi_q u)\|_{L^2(\Omega)} + \|h_{\Delta} D_{h,q}^2 \pi_q u\|_{L^2(\Omega)},
$$

$$\leq C \|U - \pi_q u\|_H + C \|h_{\Delta} D^2 u\|_{L^2(\Omega)},
$$

$$\leq C \|U - u\|_H + C \|h_{\Delta} D^2 u\|_{L^2(\Omega)}.
$$

Finally, we use this estimate with Theorem 5 to get

$$\left( \sum_{\Omega_{q,j} \subset \Omega_q} \|h_{\Delta}^{\frac{1}{2}} r(t; U(t))\|_{L^2(\partial \Omega_{q,j})} \right)^{\frac{1}{2}} \leq C \left( \|h D^2 u\|_{L^2({\mathcal J}; L^2(\Omega))^2} + \|\partial_{\Gamma^q} u\|_{L^2({\mathcal J}; H)} \right),$$

and Lemma 20 now follows from this.
REFERENCES

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


[44] ———, $L_2(0, t)$ 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.


