NUMERICAL SOLUTION OF LINEAR QUASISTATIC HEREDITARY VISCOELASTICITY PROBLEMS I: A PRIORI ESTIMATES*

SIMON SHAW† AND J. R. WHITEMAN†

January 8, 1999

Abstract. We give a space-time Galerkin finite element discretization of the linear quasistatic compressible viscoelasticity problem as described by an elliptic partial differential equation with a Volterra (memory) term. The discretization consists of a continuous piecewise linear approximation in space with a discontinuous piecewise constant or linear approximation in time. We derive an *a priori* maximum-energy Galerkin-error estimate by exploiting Galerkin "orthogonality" and the data-stability of a related discrete backward problem. We also describe the form of the numerical algorithm with particular emphasis on the computational advantage of modelling the viscoelastic relaxation functions with a Dirichlet-Prony series.

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

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

1. Introduction. For a positive real number T let $\mathcal{J} := [0, T]$ denote a time interval, and for $n \in \{1, 2, 3\}$ 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 viscoelastic compressible body $\mathcal{G}$ occupies Ω and is acted upon by a system of body forces $\boldsymbol{f} := \big(f_i(\boldsymbol{x}, t)\big)_{i=1}^n$ for $\boldsymbol{x} \in \Omega$ and $t \in \mathcal{J}$. Furthermore, we also suppose that the surface of $\mathcal{G}$ coincides with $\partial\Omega$, and there exists a time independent closed set $\Gamma_D \subseteq \partial\Omega$ of positive 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 $\boldsymbol{g} := \big(g_i(\boldsymbol{x}, t)\big)_{i=1}^n$ for $\boldsymbol{x} \in \Gamma_N$ and $t \in \mathcal{J}$. We denote the unit outward directed normal vector to Γ_N by $\widehat{\boldsymbol{n}} := (\widehat{n}_i)_{i=1}^n$.

As a consequence of the applied forces $\boldsymbol{f}$ and $\boldsymbol{g}$ the body will deform over time and the point $\boldsymbol{x}$ in the body will move to the point $\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}, t)$, where $\boldsymbol{u} = (u_i)_{i=1}^n$. Here the function $\boldsymbol{u} : \Omega \times \mathcal{J} \to \mathbb{R}^n$ describes the displacement of the body from its equilibrium configuration and is assumed to be small so as not to violate the assumption that Ω is time independent. This is the linear theory. We assume that $t = 0$ is a reference time such that $\boldsymbol{u} \equiv \boldsymbol{0}$ for all $t < 0$.

In the quasistatic theory of viscoelasticity one assumes that the inertia of $\mathcal{G}$ 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,

$$-\sigma_{ij,j} = f_i \quad \text{in } \Omega \times \mathcal{J}, \tag{1}$$

$$u_i = 0 \quad \text{in } \Gamma_D \times \mathcal{J}, \tag{2}$$

$$\sigma_{ij}\widehat{n}_j = g_i \quad \text{in } \Gamma_N \times \mathcal{J}. \tag{3}$$

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

*The research reported herein has been sponsored in part by the United States Army through its European Research Office, and in part by the Brunel University Research Committee: 96–97 QR funding scheme.

†BICOM, Brunel University, Uxbridge, UB8 3PH, U.K. (simon.shaw@brunel.ac.uk, john.whiteman@brunel.ac.uk, BICOM: http://www.brunel.ac.uk/~icsrbicm).

$\underline{\varepsilon} := (\varepsilon_{ij})_{i,j=1}^{n}$ from the displacement field using the relations,

$$\varepsilon_{ij}(\boldsymbol{u}) := \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right). \tag{4}$$

Recall that if $\mathcal{G}$ were linear elastic we could close this problem by introducing the Hooke's law constitutive relationship between stress and strain which states that,

$$\sigma_{ij} = D_{ijkl}\varepsilon_{kl}(\boldsymbol{u}) \quad \Longrightarrow \quad \underline{\boldsymbol{\sigma}} = \underline{\boldsymbol{D}}\,\underline{\boldsymbol{\varepsilon}}.$$

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

$$D_{ijkl} = D_{jikl}, \qquad D_{ijkl} = D_{ijlk}, \qquad \text{and} \qquad D_{ijkl} = D_{klij}. \tag{5}$$

However, viscoelastic materials possess memory and to model such a (linear) constitutive relationship it is necessary to introduce a time dependence into $\underline{\boldsymbol{D}}$ and augment Hooke's law with a memory term. In this way the stress is expressed as a linear functional of the strain which, for our problem, we write as,

$$\sigma_{ij}(\boldsymbol{x},t) = D_{ijkl}(\boldsymbol{x},0)\varepsilon_{kl}(\boldsymbol{u}(\boldsymbol{x},t)) - \int_0^t \frac{\partial D_{ijkl}(\boldsymbol{x},t-s)}{\partial s}\varepsilon_{kl}(\boldsymbol{u}(\boldsymbol{x},s))\,ds. \tag{6}$$

Now $\underline{\boldsymbol{D}}$ is a tensor of *stress relaxation functions* the components of which 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 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 $\underline{\boldsymbol{\gamma}}$. Here, and usually below, we omit the $\boldsymbol{x}$ dependence.

REMARK 1. *Note that in viscoelasticity the first two symmetries in* (5) *continue to hold for all time but the third,*

$$D_{ijkl}(t) = D_{klij}(t),$$

holds for all t only for isotropic materials. According to Lockett, [10, Equations 1.10 and 2.62]*, this symmetry holds in the more general anisotropic case only in the limits $t = 0$ and $t = \infty$. Note that this does not affect the symmetry of $\underline{\boldsymbol{\sigma}}$.*

In the simplest case we could have $\underline{\boldsymbol{D}}(\boldsymbol{x},t) = \varphi(t)\underline{\boldsymbol{D}}_0(\boldsymbol{x})$, where $\underline{\boldsymbol{D}}_0$ is a temporally constant tensor, and φ is a generic stress relaxation function which, from the *fading memory hypothesis*, can be assumed to be monotonically decreasing and positive for $t \geq 0$.

A popular form for modelling the stress relaxation functions is the Dirichlet-Prony series,

$$D_{ijkl}(t) = (D_{ijkl})_0 + \sum_{m=1}^{N_D}(D_{ijkl})_m e^{-\alpha_m t}, \tag{7}$$

for non-negative (and possibly $\boldsymbol{x}$ dependent) data $\{(D_{ijkl})_m\}_{m=0}^{N_D}$ and $\{\alpha_m\}_{m=1}^{N_D}$. Note that in general we ought to write $(\alpha_{ijkl})_m$ and $(N_{ijkl})_m$ instead of α_m and N_D, but we prefer not to so as to keep the notation manageable.

Viscoelastic materials can be distinguished as either solid or fluid depending upon the long-time behaviour of the stress relaxation functions. For example, if we again make the simplifying assumption that $\underline{\boldsymbol{D}}(t) = \varphi(t)\underline{\boldsymbol{D}}_0$, where $\underline{\boldsymbol{D}}_0$ is a temporally constant tensor and φ is a scalar function of time (these are so-called *synchronous* materials), and if

$$\lim_{t\to\infty} \varphi(t) = c > 0,$$

then the material is a solid and, for temporally constant loading the strain tensor $\underline{\boldsymbol{\varepsilon}}$ will tend toward a time independent limit. On the other hand if $c = 0$ then the material is a fluid and now at least some components of $\underline{\boldsymbol{\varepsilon}}$ will increase without bound as $t \to \infty$. This distinction has implications for the data-stability of $\boldsymbol{u}$ and we will refer back to it later in § 2 when we quote the relevant stability estimates.

For a more detailed account of viscoelasticity theory we refer to the many standard texts, and in particular to [6], [10] or [4].

In [12] and [11] this problem was discretized in space using the finite element method, and in time by replacing the history integral with the trapezoidal rule for numerical quadrature. *A priori* error estimates for the scheme were also given in the case of constant mesh width h and time step k. Such estimates and restrictions on the discretization are not useful for the numerical solution of *practical problems* which will usually demand graded and time dependent space-meshes, as well as variable time steps. The requirement of constant time step was in fact relaxed in [11], but of course *a priori* error estimates are in general not computable and so no error control or feedback is possible.

In this article we adopt the paradigm of adaptive space-time finite element approximation as pioneered and developed by Johnson *et al.* and discussed, for example, in [2]. In terms of space discretization our method is based upon that given in [9], which deals with the linear elasticity problem, and in terms of time discretization we extend the work presented in [14, 15, 16]. In the context of finite element approximation of pure-time Volterra problems we also mention the recent work in [1], which looks also at quadrature errors.

The advantage of using this methodology is that by applying the finite element method to approximate in space-time one can exploit the resulting Galerkin "orthogonality", along with the stability properties of associated dual problems, to derive rigorously founded *a priori* and *a posteriori* error estimates. *A posteriori* error estimates can be found in the companion papers [17] and [18], while in this article we derive an *a priori* maximum-energy Galerkin-error estimate (i.e. we measure the error in the $L_\infty(0, T)$-energy norm) which allows for graded space meshes and variable time steps. We also present data-stability estimates for the discrete problem, and discuss the numerical implementation of the scheme with particular regard to the advantage offered by using (7) to model the viscoelasticity.

This article is subdivided into seven sections, including this one, and contains two appendices. These are organized as follows.

Section 2. (Page 4) We give the weak formulation of the problem described earlier, state our baseline assumptions and quote the relevant data-stability estimates.

Section 3. (Page 7) Here we define the finite element approximation and include some additional and standard assumptions on both the data and the approximation properties of the finite element spaces.

Section 4. (Page 9) In this finite element paradigm the *a posteriori* estimate exploits a continuous dual backward problem, while the *a priori* error estimate uses a discrete backward problem. In this section we derive stability estimates for the discrete problem.

Section 5. (Page 16) Contains the *a priori* Galerkin energy-error estimate.

Section 6 (Page 19) Contains some details on the resulting numerical algorithms and their implementation.

Section 7. (Page 24) The concluding section contains a brief discussion of issues related to implementing the schemes.

Appendices A and B. (Page 25 onward) These appendices contain two of the longer technical proofs that we require as auxiliary results.

2. Weak formulation and stability. To give a weak formulation of this problem we need first to identify a test space of admissible functions defined on $\Omega \times \mathcal{J}$, and then integrate over $\Omega \times \mathcal{J}$ (by parts where necessary) the product of (1) with an arbitrary test function. However, it is convenient to proceed in two stages: first to integrate over Ω against a spatial test function only—yielding a "semi-weak" form of the problem; and then to extend the test space and integrate over $\mathcal{J}$. The reason for this is that it is easier to discuss data-stability in the context of the "semi-weak" problem.

To begin we define for $s = 0, 1, 2, \dots$ the product Hilbert spaces,

$$\Big(\boldsymbol{H}^s(\Omega), (\,\cdot\,,\,\cdot\,)_s\Big) := \Big(H^s(\Omega), (\,\cdot\,,\,\cdot\,)_{H^s(\Omega)}\Big)\times\cdots\times\Big(H^s(\Omega), (\,\cdot\,,\,\cdot\,)_{H^s(\Omega)}\Big) \qquad (n \text{ times}),$$

where the inner products are 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, $\boldsymbol{L}_2(\Omega) \equiv \boldsymbol{H}^0(\Omega)$. We also use the "trace space",

$$\boldsymbol{H}^{\frac{1}{2}}(\Gamma_N) := H^{\frac{1}{2}}(\Gamma_N) \times \cdots \times H^{\frac{1}{2}}(\Gamma_N) \qquad (n \text{ times}),$$

with norm $\|\cdot\|_{\frac{1}{2}}$. Also, and as is usual for time dependent problems, for a Banach space $(B, \|\cdot\|_B)$ we define the $L_p(0, t; B)$ norms by,

$$\|\boldsymbol{v}\|_{L_p(0,t;B)} := \left(\int_0^t \|\boldsymbol{v}(s)\|_B^p \, ds\right)^{\frac{1}{p}}$$

for $t \in \mathcal{J}$, and with the obvious "ess sup" modification when $p = \infty$.

Using the essential boundary condition (2) we now define the (spatial) test space,

$$H := \Big\{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) : \boldsymbol{v} = \boldsymbol{0} \text{ on } \Gamma_D\Big\}, \tag{8}$$

and form the Euclidean scalar product of equation (1) with an arbitrary $\boldsymbol{v} \in H$. Integrating by parts over Ω, using the symmetry of the stress tensor to see that $\sigma_{ij} v_{i,j} \equiv \sigma_{ij}\varepsilon_{ij}(\boldsymbol{v})$, then gives,

$$\int_\Omega \sigma_{ij}\varepsilon_{ij}(\boldsymbol{v})\, d\Omega = L(t; \boldsymbol{v}) := \int_\Omega \boldsymbol{v} \cdot \boldsymbol{f}(t)\, d\Omega + \oint_{\Gamma_N} \boldsymbol{v} \cdot \boldsymbol{g}(t)\, d\Gamma, \tag{9}$$

at a given $t \in \mathcal{J}$, and where we also used the traction boundary condition (3). Assuming for the moment regular enough $\boldsymbol{f}$ and $\boldsymbol{g}$ we now use (6) to substitute for the stress and arrive at the "semi-weak" problem: find $\boldsymbol{u} \in L_p(\mathcal{J}; H)$ such that,

$$A(\boldsymbol{u}(t), \boldsymbol{v}) = L(t; \boldsymbol{v}) + \int_0^t B(t, s; \boldsymbol{u}(s), \boldsymbol{v})\, ds \qquad \forall \boldsymbol{v} \in H, \text{ a.e. in } \mathcal{J}. \tag{10}$$

We call this problem "semi weak" because later we will integrate over $\mathcal{J}$ also to obtain a "fully weak" formulation. These terms are just labels and no mathematical significance should be attached to the "semi" and "fully" qualifiers. Here for the triangle,

$$\mathcal{T} := \big\{(t, s) \in \mathcal{J} \times \mathcal{J} : 0 \leq s \leq t \in \mathcal{J}\big\},$$

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

$$A(\boldsymbol{w}, \boldsymbol{v}) := \int_\Omega D_{ijkl}(0)\varepsilon_{kl}(\boldsymbol{w})\varepsilon_{ij}(\boldsymbol{v})\, d\Omega, \tag{11}$$

$$B(t, s; \boldsymbol{w}, \boldsymbol{v}) := \int_\Omega \frac{\partial D_{ijkl}(t-s)}{\partial s}\varepsilon_{kl}(\boldsymbol{w})\varepsilon_{ij}(\boldsymbol{v})\, d\Omega, \tag{12}$$

for all $\boldsymbol{w}$, $\boldsymbol{v} \in H$, and $L : \mathcal{J} \times H \to \mathbb{R}$ is a time dependent linear form as suggested by (9), and about which we give more details below. It is appropriate at this stage to state our basic assumptions for this problem.

ASSUMPTIONS 2 (**general assumptions**). *For some $p \in [1, \infty]$ we assume that the semi-weak problem* (10) *has a unique solution $\boldsymbol{u} \in L_p(\mathcal{J}; H)$ and has also the following properties.*

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

$$|A(\boldsymbol{w}, \boldsymbol{v})| \leq C\|\boldsymbol{w}\|_1\|\boldsymbol{v}\|_1 \qquad \text{and} \qquad A(\boldsymbol{v}, \boldsymbol{v}) \geq c\|\boldsymbol{v}\|_1^2,$$

for all $\boldsymbol{w}, \boldsymbol{v} \in H$ and where C and c are positive constants. Using these properties we define 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 $\|\cdot\|_1$.

(ii) *The bilinear form $B(t, s;\,\cdot\,, \cdot\,)$ 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*

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

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

(iii) *The linear form $L(t;\,\cdot\,)$ is continuous on H in the sense that there exists $F \in L_p(\mathcal{J})$ such that,*

$$|L(t; \boldsymbol{v})| \leq F(t)\|\boldsymbol{v}\|_H,$$

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

(iv) *The components of the stress relaxation tensor $\underline{\boldsymbol{D}}(t)$ are a.e. in Ω monotone-decreasing functions of t, with first time derivatives of class $L_1(\mathcal{J})$. We also assume* *that there is a $\widehat{k} > 0$ such that the following hold. Firstly, there exists a constant $c_0 > 0$ (possibly depending on $\widehat{k}$) such that,*

$$D_{ijkl}(t)\gamma_{ij}\gamma_{kl} \geq c_0 D_{ijkl}(0)\gamma_{ij}\gamma_{kl} \tag{13}$$

a.e. in Ω, for all $t \in [0, \widehat{k}]$ and for all symmetric second order tensors $\underline{\gamma}$. Secondly and for fixed $\boldsymbol{x}$, no component $D_{ijkl}(t)$ changes sign in $[0, \widehat{k}]$.

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 states that there is a constant K such that,

$$\int_\Omega \varepsilon_{ij}(\boldsymbol{v})\varepsilon_{ij}(\boldsymbol{v})\, d\Omega \geq K\|\boldsymbol{v}\|_1^2 \qquad \forall \boldsymbol{v} \in H.$$

In general this is a difficult result to establish, see for example [7] or [5], but can be safely assumed for our problem since we insist that $\text{meas}(\Gamma_D) > 0$, and thus H contains no nontrivial zero strain energy displacements (i.e. no rigid body motions). To motivate part (ii) we need only look back to equation (12) and use (7).

If we assume that $\boldsymbol{f} \in L_p(\mathcal{J}; \boldsymbol{L}_2(\Omega))$ and $\boldsymbol{g} \in L_p(\mathcal{J}; \boldsymbol{L}_2(\Gamma_N))$, then the definition of $L(t;\cdot)$ in (9) makes sense, and is in fact the representation we use below when deriving the error estimates. However, for the dual problem (30) it is convenient to take a more abstract approach. We note that $(H, A(\cdot,\cdot))$ is a Hilbert space with (topological) dual H', and with the norm on this dual space given by,

$$\|\cdot\|_{H'} := \sup\Big\{|\langle\cdot, \boldsymbol{v}\rangle_H| : \boldsymbol{v} \in H,\ \|\boldsymbol{v}\|_H = 1\Big\},$$

where the bilinear form $\langle\cdot,\cdot\rangle_H : H' \times H \to \mathbb{R}$ denotes the duality pairing between H and H'. Then $L(t;\boldsymbol{v}) \equiv \langle L(t), \boldsymbol{v}\rangle_H$. Now, to weaken the requirements on $\boldsymbol{f}$ we use $\boldsymbol{L}_2(\Omega)$ as a pivot space and define H^{-1} as the space of all linear functionals, $H \to \mathbb{R}$, bounded in the norm,

$$\|\cdot\|_{-1} := \sup\Big\{|\langle\cdot, \boldsymbol{v}\rangle| : \boldsymbol{v} \in H,\ \|\boldsymbol{v}\|_H = 1\Big\},$$

where $\langle\cdot,\cdot\rangle := (\cdot,\cdot)_0$. We now require only that,

$$\boldsymbol{f} \in L_p(\mathcal{J}; H^{-1}) \qquad \text{and} \qquad \boldsymbol{g} \in L_p(\mathcal{J}; (\boldsymbol{H}^{\frac{1}{2}}(\Gamma_N))'),$$

and the form $L(t)$ is then well defined by, for example,

$$L(t;\boldsymbol{v}) \equiv \langle L(t), \boldsymbol{v}\rangle_H := \langle \boldsymbol{f}(t), \boldsymbol{v}\rangle + \langle \boldsymbol{g}(t), \boldsymbol{v}\rangle_{\Gamma_N},$$

for all $\boldsymbol{v} \in H$, and where $\langle\cdot,\cdot\rangle_{\Gamma_N} := (\cdot,\cdot)_{L_2(\Gamma_N)} : (\boldsymbol{H}^{\frac{1}{2}}(\Gamma_N))' \times \boldsymbol{H}^{\frac{1}{2}}(\Gamma_N) \to \mathbb{R}$ is the duality pairing of the trace spaces. The function F in Assumptions 2, part (iii) is given by

$$F(t) := \|L(t)\|_{H'} \leq \|\boldsymbol{f}(t)\|_{-1} + C_T\, \|\boldsymbol{g}(t)\|_{-\frac{1}{2}},$$

where $\|\cdot\|_{-\frac{1}{2}}$ is the norm on $(\boldsymbol{H}^{\frac{1}{2}}(\Gamma_N))'$, and C_T is a trace constant such that,

$$|\langle \boldsymbol{g}(t), \boldsymbol{v}\rangle_{\Gamma_N}| \leq C_T\|\boldsymbol{g}(t)\|_{-\frac{1}{2}}\|\boldsymbol{v}\|_H \tag{14}$$

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

We have the following data-stability estimate for the problem (10).

THEOREM 3 (**data stability**). *Let Assumptions 2 hold, then there exists a function $S : \mathcal{J} \to [0, \infty)$ such that the basic data-stability estimate holds*:

$$\|\boldsymbol{u}\|_{L_p(0,t;H)} \leq S(t)\|L\|_{L_p(0,t;H')},$$

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

In the general case where ϕ in Assumptions 2, part (ii) is more or less arbitrary the basic estimate can be established (at least for $p = 1$ or $p = \infty$) by using Gronwall's inequality (for $p = 1$ see [15]), in which case $S(t)$ is exponentially large in t. However, for the viscoelasticity problem described above one can make more precise and physically reasonable assumptions on ϕ, motivated by (6) and (7), 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 meas$(\mathcal{J})$, and for a viscoelastic fluid (in the sense described earlier) we have

$$S(t) = O(1 + t).$$

These results are a generalization of the stability estimates given previously in [11], and full details can be found in [19].

Seeking a solution in the space $L_p(\mathcal{J}; H)$ seems natural for this problem since (10) contains no time derivatives. However, such functions do not in general possess pointwise values (in H) on $\mathcal{J}$ and so we define the "fully-weak" formulation of this problem as: find $\boldsymbol{u} \in L_p(\mathcal{J}; H)$ such that,

$$a(\boldsymbol{u}, \boldsymbol{v}) = l(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in L_q(\mathcal{J}; H), \tag{15}$$

where $q \in [1, \infty]$ is the conjugate Hölder index to p given by $p^{-1} + q^{-1} = 1$ (with the obvious interpretation if $p = 1$ or $p = \infty$), and:

$$a(\boldsymbol{u}, \boldsymbol{v}) := \int_0^T A(\boldsymbol{u}(t), \boldsymbol{v}(t))\, dt - \int_0^T \int_0^t B(t, s; \boldsymbol{u}(s), \boldsymbol{v}(t))\, ds dt, \tag{16}$$

$$l(\boldsymbol{v}) := \int_0^T L(t; \boldsymbol{v}(t))\, dt. \tag{17}$$

Note that in (6) 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.

3. Finite element approximation. Equation (15) is the starting point for the space-time finite element discretization of the problem, and to this we now turn. We firstly partition $\mathcal{J}$ into N time intervals $\{\mathcal{J}_i\}_{i=1}^N$, where $\mathcal{J}_i := (t_{i-1}, t_i)$, of lengths $k_i := t_i - t_{i-1} > 0$ and such that,

$$\mathcal{J} \equiv \overline{\mathcal{J}_1 \cup \mathcal{J}_2 \cup \cdots \cup \mathcal{J}_N},$$

so that $t_0 = 0$ and $t_N = T$. We use k to denote the piecewise constant function such that $k|_{\mathcal{J}_i} := k_i$.

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

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

and use h to denote the piecewise constant mesh function given by $h|_{\Omega_{ij}} := h_{ij}$. 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 $h|_{\mathcal{J}_i} := h_i$. We need to assume that arbitrary partitions of $\Omega \times \mathcal{J}$ of this nature always exist because in an adaptive solution the $\{\mathcal{J}_i\}$ and $\{\Omega_i\}$ are of course not known in advance.

Having now broken the prismatic domain $\Omega \times \mathcal{J}$ into the laminae (or "slabs") $\Omega \times \mathcal{J}_i$, indexed by the time levels $i \in \mathbb{N}(1, N)$, we define for each Ω_i the semidiscrete (spatial) finite element spaces,

$$H_i := \Big\{ \boldsymbol{v} \in H \cap \big(C(\overline{\Omega})\big)^n : \boldsymbol{v} \text{ is linear on } \Omega_{ij} \text{ for each } j \in \mathbb{N}(1, M_i) \Big\}.$$

The space-time finite element spaces are now given by V_r where,

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

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

We define also the set of internal edges (for triangular Ω_{ij} in $\mathbb{R}^2$), or faces (for tetrahedral Ω_{ij} in $\mathbb{R}^3$) in each Ω_i as,

$$\mathcal{L}_i := \Big\{ \ell \subset \Omega : \exists j \in \mathbb{N}(1, M_i) \text{ such that } \ell \text{ is an edge/face of } \Omega_{ij} \Big\},$$

and the set of edges, or faces, on the Neumann boundary Γ_N as,

$$\mathcal{F}_i := \Big\{ \ell \subset \Gamma_N : \exists j \in \mathbb{N}(1, M_i) \text{ such that } \ell \text{ is an edge/face of } \Omega_{ij} \Big\}.$$

We assume that there are respectively $N_{\mathcal{L}_i}$ and $N_{\mathcal{F}_i}$ such edges in the time interval $\mathcal{J}_i$. We also define,

$$\mathcal{H}_{ij} := \mathcal{L}_j \setminus (\mathcal{L}_i \cap \mathcal{L}_j) \qquad \text{for } 1 \leq j \leq i \leq N, \tag{18}$$

as the set of internal edges/faces belonging to $\mathcal{L}_j$ but not to $\mathcal{L}_i$. Note that $\mathcal{H}_{ii} \equiv \emptyset$, and that if we allow only nested refinements such that $H_{i-1} \subseteq H_i$ for $i \in \mathbb{N}(2, N)$, then $\mathcal{L}_i \cap \mathcal{L}_j = \mathcal{L}_j$ and $\mathcal{H}_{ij} \equiv \emptyset$.

Once we choose a value for r, which for us will be $r = 0$ or $r = 1$, we form the finite element approximation to (15) as: find $\boldsymbol{U} \in V_r$ such that,

$$a(\boldsymbol{U}, \boldsymbol{v}) = l(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in V_r, \tag{19}$$

and subtracting this from (15) gives the fundamentally important Galerkin "orthogonality" relationship:

$$a(\boldsymbol{u} - \boldsymbol{U}, \boldsymbol{v}) = 0 \qquad \forall \boldsymbol{v} \in V_r. \tag{20}$$

This property, when coupled to the data stability of an associated dual backward problem, is the basic building block in the error estimation technique developed by Johnson *et al.* in for example [2].

We need to make some additional assumptions concerning the regularity of the data and the approximation properties of the discretization. From now on we restrict our attention to the case where $p = \infty$ in Assumptions 2 since the norm $\|\cdot\|_{L_\infty(\mathcal{J};H)}$ appears to be the most convenient for measuring the discretization error.

ASSUMPTIONS 4 (**discretization assumptions**). *Let Assumptions* 2 *hold for* $p = \infty$. *Also let the following hold.*

(i) $\boldsymbol{f} \in L_\infty(\mathcal{J}; \boldsymbol{L}_2(\Omega))$ *and* $\boldsymbol{g} \in L_\infty(\mathcal{J}; \boldsymbol{L}_2(\Gamma_N))$.

(ii) *The components of* $\underline{\boldsymbol{D}}(t)$ *are bounded and piecewise constant in* Ω *for all* $t \in \mathcal{J}$. *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 the set* $\mathcal{S}_i \subset \Omega$ *of discontinuities of* $\underline{\boldsymbol{D}}(t)$ *always satisfies* $\mathcal{S}_i \subset \mathcal{L}_i$.

(iii) *Every space mesh is* nondegenerate *in that every element* Ω_{ij} *contains and is contained by balls of radii* r_{i} *and* r_{o} *respectively, 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*:

$$\|\pi_i \boldsymbol{w}\|_{\boldsymbol{L}_2(\Omega)} \leq \Pi_s \|\boldsymbol{w}\|_{\boldsymbol{L}_2(\Omega)}, \tag{21}$$

$$\|\pi_i \boldsymbol{w}\|_H \leq \widetilde{\Pi}_s \|\boldsymbol{w}\|_H, \tag{22}$$

as well as the following error estimates:

$$\left(\sum_{\Omega_{ij} \subset \overline{\Omega}_i} \left\| h_{ij}^{-1} (\boldsymbol{w} - \pi_i \boldsymbol{w}) \right\|_{\boldsymbol{L}_2(\Omega_{ij})}^2 \right)^{\frac{1}{2}} \leq \Pi_{\Omega_i} \|\boldsymbol{w}\|_H, \tag{23}$$

$$\left(\sum_{\Omega_{ij} \subset \overline{\Omega}_i} \left\| h_{ij}^{-\frac{1}{2}} (\boldsymbol{w} - \pi_i \boldsymbol{w}) \right\|_{\boldsymbol{L}_2(\partial\Omega_{ij})}^2 \right)^{\frac{1}{2}} \leq \Pi_\ell \|\boldsymbol{w}\|_H, \tag{24}$$

for all $\boldsymbol{w} \in H$ *and where* Π_s, $\widetilde{\Pi}_s$, Π_{Ω_i} *and* Π_ℓ *are "constants" which in general can vary with* i (*i.e. with time*) *but are independent of* h *and* $\boldsymbol{w}$. *Occasionally we will also use the "global interpolator"* π *defined piecewise by* $\pi|_{\mathcal{J}_i} := \pi_i$.

We could easily allow the case where $\underline{\boldsymbol{D}}$ is piecewise smooth in space but the piecewise constant case is more likely to arise in practical problems since most materials have piecewise constant properties. Also, this assumption simplifies the *a posteriori* error estimate developed in [17].

We now examine the data stability of the discrete solution.

4. data stability of the discrete problem. For each of the time slabs $\mathcal{J}_q$ we set $\boldsymbol{U}_q := \boldsymbol{U}|_{\mathcal{J}_q}$ and then for each q we may write,

$$\boldsymbol{U}_q(t) = \boldsymbol{U}_q^- \varphi_q^-(t) + \boldsymbol{U}_q^+ \varphi_q^+(t),$$

where: $\boldsymbol{U}_q^+, \boldsymbol{U}_q^- \in H_q$; and the $\varphi_q^+, \varphi_q^- \in \mathbb{P}_r(\mathcal{J}_q; \mathbb{R})$ are the temporal basis functions for $\mathcal{J}_q$. Note that in the case $r = 0$ we can write,

$$\boldsymbol{U}_q^+ \equiv \boldsymbol{U}_q^-, \qquad \varphi_q^-(t) := 1 \qquad \text{and} \qquad \varphi_q^+(t) := 0,$$

while for $r = 1$ we in general have $\boldsymbol{U}_q^+ \neq \boldsymbol{U}_q^-$,

$$\varphi_q^-(t) := 1 \qquad \text{and} \qquad \varphi_q^+(t) := \frac{t - t_{q-1}}{k_q},$$

although these choices are obviously not unique.

With this notation we now write the discrete problem (19) in the following way. For some $q \in \mathbb{N}(1, N)$, some $\boldsymbol{V} \in H_q$, and some $\varphi \in \mathbb{P}_r(\mathcal{J}_q; \mathbb{R})$ we take $\boldsymbol{v} = \varphi \boldsymbol{V}$ in the slab $\Omega \times \mathcal{J}_q$ and $\boldsymbol{v} = \boldsymbol{0}$ elsewhere, hence $\boldsymbol{v} \in V_r$ and the left hand side of equation (19) becomes

$$\begin{aligned} a(\boldsymbol{U}, \boldsymbol{v}) &= \int_{t_{q-1}}^{t_q} A(\boldsymbol{U}(t), \boldsymbol{v}(t))\, dt - \int_{t_{q-1}}^{t_q} \int_0^t B(t, s; \boldsymbol{U}(s), \boldsymbol{v}(t))\, ds dt, \\ &= \int_{t_{q-1}}^{t_q} \varphi(t) \left[A(\boldsymbol{U}_q(t), \boldsymbol{V}) - \int_{t_{q-1}}^{t} B(t, s; \boldsymbol{U}_q(s), \boldsymbol{V})\, ds \right] dt \\ &\qquad - \int_{t_{q-1}}^{t_q} \int_0^{t_{q-1}} B(t, s; \boldsymbol{U}(s), \boldsymbol{v}(t))\, ds dt. \end{aligned} \tag{25}$$

To simplify this further we take a closer look at the first integrand on the right hand side.

$$\begin{aligned} A(\boldsymbol{U}_q(t), \boldsymbol{V}) - \int_{t_{q-1}}^{t} B(t, s; \boldsymbol{U}_q(s), \boldsymbol{V})\, ds &= \varphi_q^+(t) A(\boldsymbol{U}_q^+, \boldsymbol{V}) + \varphi_q^-(t) A(\boldsymbol{U}_q^-, \boldsymbol{V}) \\ &\quad - \int_{t_{q-1}}^{t} \varphi_q^+(s) B(t, s; \boldsymbol{U}_q^+, \boldsymbol{V}) + \varphi_q^-(s) B(t, s; \boldsymbol{U}_q^-, \boldsymbol{V})\, ds, \\ = \varphi_q^+(t) A(\boldsymbol{U}_q^+, \boldsymbol{V}) &- \int_{t_{q-1}}^{t} \varphi_q^+(s) B(t, s; \boldsymbol{U}_q^+, \boldsymbol{V})\, ds \\ &+ \varphi_q^-(t) A(\boldsymbol{U}_q^-, \boldsymbol{V}) - \int_{t_{q-1}}^{t} \varphi_q^-(s) B(t, s; \boldsymbol{U}_q^-, \boldsymbol{V})\, ds. \end{aligned}$$

Working now from the definitions (11) and (12) we have first by a partial integration that,

$$\begin{aligned} &\varphi_q^+(t) A(\boldsymbol{U}_q^+, \boldsymbol{V}) - \int_{t_{q-1}}^{t} \varphi_q^+(s) B(t, s; \boldsymbol{U}_q^+, \boldsymbol{V})\, ds \\ &\quad = \int_\Omega \Bigg[\varphi_q^+(t) D_{ijkl}(0) \varepsilon_{kl}(\boldsymbol{U}_q^+) \varepsilon_{ij}(\boldsymbol{V}) \\ &\qquad\qquad - \int_{t_{q-1}}^{t} \varphi_q^+(s) \frac{\partial D_{ijkl}(t-s)}{\partial s} \varepsilon_{kl}(\boldsymbol{U}_q^+) \varepsilon_{ij}(\boldsymbol{V})\, ds \Bigg] d\Omega, \\ &\quad = \int_\Omega \Bigg[\varphi_q^+(t_{q-1}) D_{ijkl}(t - t_{q-1}) \varepsilon_{kl}(\boldsymbol{U}_q^+) \varepsilon_{ij}(\boldsymbol{V}) \\ &\qquad\qquad + \int_{t_{q-1}}^{t} \dot{\varphi}_q^+(s) D_{ijkl}(t-s) \varepsilon_{kl}(\boldsymbol{U}_q^+) \varepsilon_{ij}(\boldsymbol{V})\, ds \Bigg] d\Omega, \end{aligned}$$

with a similar result for the "$-$" terms and where $\dot{\varphi}_q^\pm := d\varphi_q^\pm / ds$.

Using these we find for $r = 0$ or $r = 1$ that,

$$
\begin{aligned}
A(\boldsymbol{U}_q(t), \boldsymbol{V}) &- \int_{t_{q-1}}^{t} B(t, s; \boldsymbol{U}_q(s), \boldsymbol{V})\, ds \\
&= \int_\Omega D_{ijkl}(t - t_{q-1})\varepsilon_{kl}(\varphi_q^+(t_{q-1})\boldsymbol{U}_q^+ + \varphi_q^-(t_{q-1})\boldsymbol{U}_q^-)\varepsilon_{ij}(\boldsymbol{V})\, d\Omega \\
&\qquad + \int_{t_{q-1}}^{t}\int_\Omega D_{ijkl}(t - s)\varepsilon_{kl}(\dot{\varphi}_q^+(s)\boldsymbol{U}_q^+ + \dot{\varphi}_q^-(s)\boldsymbol{U}_q^-)\varepsilon_{ij}(\boldsymbol{V})\, d\Omega ds, \\
&= A_0(t - t_{q-1}; \boldsymbol{U}_q(t_{q-1}), \boldsymbol{V}) + A_1(t, t_{q-1}; \dot{\boldsymbol{U}}_q, \boldsymbol{V}),
\end{aligned}
$$

where for all $\boldsymbol{w}_1, \boldsymbol{w}_2 \in H$:

$$
A_0(t - t_{q-1}; \boldsymbol{w}_1, \boldsymbol{w}_2) := \int_\Omega D_{ijkl}(t - t_{q-1})\varepsilon_{kl}(\boldsymbol{w}_1)\varepsilon_{ij}(\boldsymbol{w}_2)\, d\Omega, \tag{26}
$$

$$
A_1(t, t_{q-1}; \boldsymbol{w}_1, \boldsymbol{w}_2) := \int_\Omega \left(\int_{t_{q-1}}^{t} D_{ijkl}(t - s)\, ds\right)\varepsilon_{kl}(\boldsymbol{w}_1)\varepsilon_{ij}(\boldsymbol{w}_2)\, d\Omega, \tag{27}
$$

and we used the fact that $\dot{\boldsymbol{U}}_q =$ constant on $\mathcal{J}_q$. From these it follows that,

$$
\begin{aligned}
a(\boldsymbol{U}, \varphi(t)\boldsymbol{V}) &= \int_{t_{q-1}}^{t_q} A_0(t - t_{q-1}; \boldsymbol{U}_q(t_{q-1}), \varphi(t)\boldsymbol{V}) + A_1(t, t_{q-1}; \dot{\boldsymbol{U}}_q, \varphi(t)\boldsymbol{V})\, dt \\
&\qquad - \int_{t_{q-1}}^{t_q}\int_0^{t_{q-1}} B(t, s; \boldsymbol{U}(s), \varphi(t)\boldsymbol{V})\, dsdt.
\end{aligned}
$$

Hence for any $\boldsymbol{v} \in V_r$ of the form $\boldsymbol{v} = \boldsymbol{0}$ on $\mathcal{J} \setminus \mathcal{J}_q$ and,

$$
\boldsymbol{v}|_{\mathcal{J}_q} = \boldsymbol{V}^-\varphi_q^-(t) + \boldsymbol{V}^+\varphi_q^+(t),
$$

we use this result in (25) and get by linearity,

$$
\begin{aligned}
a(\boldsymbol{U}, \boldsymbol{v}) &= a(\boldsymbol{U}, \boldsymbol{V}^-\varphi_q^-(t)) + a(\boldsymbol{U}, \boldsymbol{V}^+\varphi_q^+(t)), \\
&= \int_{t_{q-1}}^{t_q} A_0(t - t_{q-1}; \boldsymbol{U}_q(t_{q-1}), \boldsymbol{v}(t)) + A_1(t, t_{q-1}; \dot{\boldsymbol{U}}_q, \boldsymbol{v}(t))\, dt \\
&\qquad - \int_{t_{q-1}}^{t_q}\int_0^{t_{q-1}} B(t, s; \boldsymbol{U}(s), \boldsymbol{v}(t))\, dsdt \qquad \forall \boldsymbol{v} \in \mathbb{P}_r(\mathcal{J}_q, H_q).
\end{aligned} \tag{28}
$$

We now give the following lemma for the case $r = 0$.

LEMMA 5 (**discrete coercivity for** $r = 0$). *Let Assumptions 2 hold, then for* $r = 0$ *we have,* $A_1(t, t_{q-1}; \dot{\boldsymbol{U}}_q, \cdot) \equiv 0$. *Moreover, if* $k_q \le \widehat{k}$ *(see Assumption 2, part (iv)), then for* $t \in \mathcal{J}_q$,

$$
A_0(t - t_{q-1}; \boldsymbol{V}, \boldsymbol{V}) \ge c_0\|\boldsymbol{V}\|_H^2 \qquad \forall \boldsymbol{V} \in H.
$$

This holds for all $q \in \mathbb{N}(1, N)$.

Proof. Since $r = 0$ we have $\dot{\boldsymbol{U}}_q = \boldsymbol{0}$ and so we clearly have $A_1(t, t_{q-1}; \dot{\boldsymbol{U}}_q, \boldsymbol{V}) \equiv 0$ for all $\boldsymbol{V} \in H$. Using (13) in Assumptions 2, part (iv) we also have,

$$
A_0(t - t_{q-1}; \boldsymbol{V}, \boldsymbol{V}) \ge c_0 A(\boldsymbol{V}, \boldsymbol{V}) = c_0\|\boldsymbol{V}\|_H^2,
$$

as required. ☐

The corresponding result for $r = 1$ seems a little more delicate and to simplify matters we will make further assumptions on $\underline{\boldsymbol{D}}$. But first we introduce a convenient lemma.

LEMMA 6. *For each q the mapping $\|\cdot\|_* : \mathbb{P}_r(\mathcal{J}_q; H_q) \to [0,\infty)$ defined by,*

$$\|\boldsymbol{V}_q\|_* := \|\varphi_q^+ \boldsymbol{V}_q^+\|_{L_\infty(\mathcal{J}_q;H)} + \|\varphi_q^- \boldsymbol{V}_q^-\|_{L_\infty(\mathcal{J}_q;H)},$$

is a norm on $\mathbb{P}_r(\mathcal{J}_q; H_q)$ equivalent to $\|\cdot\|_{L_\infty(\mathcal{J}_q;H)}$.

Proof. To check that $\|\cdot\|_*$ is indeed a norm we first note that,

$$\begin{aligned}\|\boldsymbol{V}_q\|_* &= \|\varphi_q^+\|_{L_\infty(\mathcal{J}_q)}\|\boldsymbol{V}_q^+\|_H + \|\varphi_q^-\|_{L_\infty(\mathcal{J}_q)}\|\boldsymbol{V}_q^-\|_H,\\ &= \|\boldsymbol{V}_q^+\|_H + \|\boldsymbol{V}_q^-\|_H.\end{aligned}$$

Therefore $\boldsymbol{V}_q = \boldsymbol{0} \Leftrightarrow \|\boldsymbol{V}_q\|_* = 0$, and it is easy to see also that $\|\alpha\boldsymbol{V}_q\|_* = |\alpha|\,\|\boldsymbol{V}_q\|_*$ for all real numbers α. Also,

$$\begin{aligned}\|\boldsymbol{V}_q + \boldsymbol{W}_q\|_* &= \|(\boldsymbol{V}+\boldsymbol{W})_q^+\|_H + \|(\boldsymbol{V}+\boldsymbol{W})_q^-\|_H,\\ &\le \|\boldsymbol{V}_q^+\|_H + \|\boldsymbol{V}_q^-\|_H + \|\boldsymbol{W}_q^+\|_H + \|\boldsymbol{W}_q^-\|_H = \|\boldsymbol{V}_q\|_* + \|\boldsymbol{W}_q\|_*.\end{aligned}$$

Norm equivalence follows because $\mathbb{P}_r(\mathcal{J}_q; H_q)$ is finite dimensional. ☐

LEMMA 7 (**discrete coercivity for** $r = 1$). *Let Assumptions 2 hold and moreover assume that the components of $\underline{\boldsymbol{D}}(t)$ are continuously differentiable functions of t, then there exists a constant $C > 0$ such that,*

$$\begin{aligned}A_0(t - t_{q-1}; \boldsymbol{U}_q(t_{q-1}), \boldsymbol{U}_q(t)) \quad &+ \quad A_1(t, t_{q-1}; \dot{\boldsymbol{U}}_q, \boldsymbol{U}_q(t))\\ &\qquad \ge c_0\|\boldsymbol{U}_q(t)\|_H^2 - Ck_q\|\boldsymbol{U}_q\|_{L_\infty(\mathcal{J}_q;H)}^2,\end{aligned}$$

where c_0 is the constant appearing in Lemma 5. Moreover, for k_q small enough,

$$\int_{t_{q-1}}^{t_q} A_0(t - t_{q-1}; \boldsymbol{U}_q(t_{q-1}), \boldsymbol{U}_q(t)) + A_1(t, t_{q-1}; \dot{\boldsymbol{U}}_q, \boldsymbol{U}_q(t))\, dt \ge c_1\|\boldsymbol{U}_q\|_{L_2(\mathcal{J}_q;H)}^2,$$

for a positive constant c_1 depending on k_q, $\underline{\boldsymbol{D}}$ and its time derivative, but not on $\boldsymbol{U}_q$.

Proof. Firstly we note that $\dot{\boldsymbol{U}}_q(t) = \dot{\varphi}_q^+(t)\boldsymbol{U}_q^+ = k_q^{-1}\boldsymbol{U}_q^+$, and also that due to our extra smoothness assumption on $\underline{\boldsymbol{D}}(t)$ there exist $\tau_{ijkl} \in [t_{q-1}, t]$ such that,

$$A_1(t, t_{q-1}; \dot{\boldsymbol{U}}_q, \boldsymbol{U}_q(t)) = \int_\Omega \left(\frac{t - t_{q-1}}{k_q}\right) D_{ijkl}(\tau_{ijkl} - t_{q-1})\varepsilon_{kl}(\boldsymbol{U}_q^+)\varepsilon_{ij}(\boldsymbol{U}_q(t))\, d\Omega,$$

where we used the mean value theorem for integrals and the "one sign" assumption in Assumption 2, part (iv). It is important to note that the symbol τ_{ijkl}, and the quantities derived from it below, should not be regarded as inducing tensor summation. With this we now have the following manipulation.

$$\begin{aligned}&A_0(t - t_{q-1}; \boldsymbol{U}_q(t_{q-1}), \boldsymbol{U}_q(t)) + A_1(t, t_{q-1}; \dot{\boldsymbol{U}}_q, \boldsymbol{U}_q(t))\\ &\quad = \int_\Omega \Big(D_{ijkl}(t - t_{q-1})\varepsilon_{kl}(\varphi_q^-(t)\boldsymbol{U}_q^-) + D_{ijkl}(\tau_{ijkl} - t_{q-1})\varepsilon_{kl}(\varphi_q^+(t)\boldsymbol{U}_q^+)\Big)\\ &\qquad\qquad\qquad\qquad\qquad\qquad \times \quad \varepsilon_{ij}(\boldsymbol{U}_q(t))\, d\Omega,\\ &\quad = A_0(t - t_{q-1}; \boldsymbol{U}_q(t), \boldsymbol{U}_q(t)) + \widetilde{A}(t, \{\tau_{ijkl}\}, t_{q-1}; \varphi_q^+(t)\boldsymbol{U}_q^+, \boldsymbol{U}_q(t)),\end{aligned}$$

where,

$$\begin{aligned}
&\widetilde{A}(t,\{\tau_{ijkl}\},t_{q-1};\varphi_q^+(t)\boldsymbol{U}_q^+,\boldsymbol{U}_q(t)) \\
&\quad := \int_\Omega \Big(D_{ijkl}(\tau_{ijkl}-t_{q-1}) - D_{ijkl}(t-t_{q-1}) \Big) \varepsilon_{kl}(\varphi_q^+(t)\boldsymbol{U}_q^+)\varepsilon_{ij}(\boldsymbol{U}_q(t))\,d\Omega.
\end{aligned}$$

Now, $0 \le \tau_{ijkl} - t_{q-1} \le t - t_{q-1}$ which implies that there exist $\xi_{ijkl} \in (\tau_{ijkl}, t)$ and $c_{ijkl} \in [0,1]$ such that,

$$D_{ijkl}(t-t_{q-1}) - D_{ijkl}(\tau_{ijkl}-t_{q-1}) = c_{ijkl}k_q \frac{\partial D_{ijkl}(\xi_{ijkl}-t_{q-1})}{\partial \xi_{ijkl}}$$

(again, summation is not implied here), and so,

$$\begin{aligned}
&|\widetilde{A}(t,\{\tau_{ijkl}\},t_{q-1};\varphi_q^+(t)\boldsymbol{U}_q^+,\boldsymbol{U}_q(t))| \\
&\qquad \le \frac{C(\underline{\boldsymbol{D}})k_q}{2}\Big(\|\varphi_q^+(t)\boldsymbol{U}_q^+\|_H^2 + \|\boldsymbol{U}_q(t)\|_H^2 \Big), \\
&\qquad \le \frac{C(\underline{\boldsymbol{D}})k_q}{2}\Big(\|\varphi_q^+\boldsymbol{U}_q^+\|_{L_\infty(\mathcal{J}_q;H)}^2 + \|\boldsymbol{U}_q\|_{L_\infty(\mathcal{J}_q;H)}^2 \Big),
\end{aligned}$$

where, with summation not implied,

$$C(\underline{\boldsymbol{D}}) := \max_{i,j,k,l} \left\{ c_{ijkl} \left\| \frac{\partial D_{ijkl}(\xi - t_{q-1})/\partial \xi}{D_{ijkl}(0)} \right\|_{L_\infty(\{t_{q-1}\le\xi\le t_q\};\Omega)} \right\}.$$

From these results and Lemma 5 it now follows that,

$$\begin{aligned}
&A_0(t-t_{q-1};\boldsymbol{U}_q(t_{q-1}),\boldsymbol{U}_q(t)) + A_1(t,t_{q-1};\dot{\boldsymbol{U}}_q,\boldsymbol{U}_q(t)) \\
&\qquad = A_0(t-t_{q-1};\boldsymbol{U}_q(t),\boldsymbol{U}_q(t)) + \widetilde{A}(t,\{\tau_{ijkl}\},t_{q-1};\varphi_q^+(t)\boldsymbol{U}_q^+,\boldsymbol{U}_q(t)), \\
&\qquad \ge c_0\|\boldsymbol{U}_q(t)\|_H^2 - \frac{C(\underline{\boldsymbol{D}})k_q}{2}\left(\|\varphi_q^+\boldsymbol{U}_q^+\|_{L_\infty(\mathcal{J}_q;H)}^2 + \|\boldsymbol{U}_q\|_{L_\infty(\mathcal{J}_q;H)}^2 \right), \\
&\qquad \ge c_0\|\boldsymbol{U}_q(t)\|_H^2 - Ck_q\left(\|\boldsymbol{U}_q\|_*^2 + \|\boldsymbol{U}_q\|_{L_\infty(\mathcal{J}_q;H)}^2 \right), \\
&\qquad \ge c_0\|\boldsymbol{U}_q(t)\|_H^2 - Ck_q\|\boldsymbol{U}_q\|_{L_\infty(\mathcal{J}_q;H)}^2,
\end{aligned}$$

as required for the first part of the lemma and where we used Lemma 6. To complete the proof we note that because $\boldsymbol{U}_q$ belongs to the finite dimensional space V_r, all norms are equivalent and so there exists a constant $\widetilde{C}$ such that,

$$Ck_q\|\boldsymbol{U}_q\|_{L_\infty(\mathcal{J}_q;H)}^2 \le \widetilde{C}\|\boldsymbol{U}_q\|_{L_2(\mathcal{J}_q;H)}^2,$$

and therefore,

$$\begin{aligned}
\int_{t_{q-1}}^{t_q} c_0\|\boldsymbol{U}_q(t)\|_H^2 - Ck_q\|\boldsymbol{U}_q\|_{L_\infty(\mathcal{J}_q;H)}^2\,dt \quad &\ge \quad c_0\|\boldsymbol{U}_q\|_{L_2(\mathcal{J}_q;H)}^2 - \widetilde{C}k_q\|\boldsymbol{U}_q\|_{L_2(\mathcal{J}_q;H)}^2 \\
&= \quad c_1\|\boldsymbol{U}_q\|_{L_2(\mathcal{J}_q;H)}^2,
\end{aligned}$$

where $c_1 > 0$ for small enough k_q. This completes the proof. □

With these discrete coercivity estimates established we can now return to equation (28) and obtain a data-stability estimate for the discrete problem (19). Our results here are unfortunately not as sharp as those given in [19] and quoted earlier in Theorem 3, and this is because a semi-weak formulation of the discrete problem has no meaning. For this reason we cannot apply the comparison equation technique used in [19], and moreover, we have not been able to find an alternative strategy for deriving such sharp estimates. The following theorem is thus a consequence of Gronwall's inequality.

THEOREM 8 (**discrete data stability**). *Let Assumptions* 2 *hold with part* (ii) *strengthened such that* $\phi \in L_2(\mathcal{J};[0,\infty))$. *Also, let there exist positive numbers* $\rho_2, \rho_3, \ldots, \rho_N$ *such that* $k_q \leq \rho_q k_{q-1}$ *for all* $q \in \mathbb{N}(2,N)$. *Then, there exists a stability constant* C_S *depending only on* ϕ, $\{\rho_i\}$ *and either* c_0 *or* c_1, *from Lemmas* 5 *or* 7 *respectively as* $r = 0$ *or* $r = 1$, *such that,*

$$\|\boldsymbol{U}\|_{L_2(0,t_q;H)} \leq C_S \|L\|_{L_2(0,t_q;H')}$$

for all $q \in \mathbb{N}(1,N)$. *Moreover, if we also have* $\phi \in L_\infty(\mathcal{J};[0,\infty))$, *there is another constant* (*also denoted by* C_S) *such that,*

$$\|\boldsymbol{U}\|_{L_1(0,t_q;H)} \leq C_S \|L\|_{L_1(0,t_q;H')}$$

for all $q \in \mathbb{N}(1,N)$. *This latter constant depends upon the constant* C_{L_1} *which is assumed such that* $k_q\|\boldsymbol{w}\|_{L_\infty(\mathcal{J}_q;H)} \leq C_{L_1}\|\boldsymbol{w}\|_{L_1(\mathcal{J}_q;H)}$ *for all* $\boldsymbol{w} \in \mathbb{P}_r(\mathcal{J}_q;H_q)$, *for* $r = 0$ *or* 1, *and for all* $q \in \mathbb{N}(1,N)$. *Finally, if* C_{L_∞} *is a constant such that* $C_{L_\infty} k_q \|\boldsymbol{w}\|^2_{L_\infty(\mathcal{J}_q;H)} \leq \|\boldsymbol{w}\|^2_{L_2(\mathcal{J}_q;H)}$ (*for all* $\boldsymbol{w} \in \mathbb{P}_r(\mathcal{J}_q;H_q)$, *for* $r = 0$ *or* 1, *and for all* $q \in \mathbb{N}(1,N)$), *then we also have another constant* C_S *such that,*

$$\|\boldsymbol{U}\|_{L_\infty(0,t_q;H)} \leq C_S \|L\|_{L_\infty(0,t_q;H')}.$$

Proof. In (19) we choose $\boldsymbol{v} \in V_r$ such that $\boldsymbol{v}|_{\mathcal{J}_q} := \boldsymbol{U}_q$ and $\boldsymbol{v} := \boldsymbol{0}$ elsewhere. Here $q \in \mathbb{N}(1,N)$ is arbitrary. Using (28) in (19) along with either Lemma 5 or 7 according as $r = 0$ or $r = 1$ we find that there is a constant c such that:

$$\begin{aligned}
c\|\boldsymbol{U}_q\|^2_{L_2(\mathcal{J}_q;H)} &\leq \left| \int_{t_{q-1}}^{t_q} \langle L(t), \boldsymbol{U}_q(t)\rangle_H \, dt \right| \\
&\qquad + \left| \int_{t_{q-1}}^{t_q} \sum_{m=1}^{q-1} \int_{t_{m-1}}^{t_m} B(t,s;\boldsymbol{U}_m(s),\boldsymbol{U}_q(t))\, dsdt \right|, \\
&\leq \int_{t_{q-1}}^{t_q} \|L(t)\|_{H'} \|\boldsymbol{U}_q(t)\|_H \, dt \\
&\qquad + \int_{t_{q-1}}^{t_q} \sum_{m=1}^{q-1} \int_{t_{m-1}}^{t_m} \phi(t-s)\|\boldsymbol{U}_m(s)\|_H \, ds \; \|\boldsymbol{U}_q(t)\|_H \, dt, \qquad (29)\\
&\leq \frac{1}{2\epsilon}\int_{t_{q-1}}^{t_q} \|L(t)\|^2_{H'} dt + \epsilon \int_{t_{q-1}}^{t_q} \|\boldsymbol{U}_q(t)\|^2_H \, dt \\
&\qquad + \frac{1}{2\epsilon}\int_{t_{q-1}}^{t_q} \left(\sum_{m=1}^{q-1} \int_{t_{m-1}}^{t_m} \phi(t-s)\|\boldsymbol{U}_m(s)\|_H \, ds \right)^2 dt
\end{aligned}$$

for all $\epsilon > 0$. Now, for the history term we have the following estimate:

$$\int_{t_{q-1}}^{t_q} \left(\sum_{m=1}^{q-1} \int_{t_{m-1}}^{t_m} \phi(t-s)\|\boldsymbol{U}_m(s)\|_H \, ds \right)^2 dt$$

$$\leq \int_{t_{q-1}}^{t_q} \left(\sum_{m=1}^{q-1} \int_{t_{m-1}}^{t_m} \phi^2(t-s)\, ds \right) \left(\sum_{m=1}^{q-1} \|\boldsymbol{U}_m\|^2_{L_2(\mathcal{J}_m;H)} \right) dt,$$
$$\leq \|\phi\|^2_{L_2(\mathcal{J})} k_q \|\boldsymbol{U}\|^2_{L_2(0,t_{q-1};H)}.$$

Choosing $\epsilon = c/2$ in the first estimate and using this one gives,

$$\frac{c}{2}\|\boldsymbol{U}_q\|^2_{L_2(\mathcal{J}_q;H)} \leq \frac{1}{c}\|L\|^2_{L_2(\mathcal{J}_q;H')} + \frac{1}{c}\|\phi\|^2_{L_2(\mathcal{J})} k_q \|\boldsymbol{U}\|^2_{L_2(0,t_{q-1};H)},$$

and now with $\alpha_j := \|\boldsymbol{U}\|^2_{L_2(0,t_j;H)}$ we sum over $q = 1, 2, \ldots, j \leq N$ and arrive at,

$$\alpha_j \equiv \|\boldsymbol{U}\|^2_{L_2(0,t_j;H)} \leq \frac{2}{c^2}\|L\|^2_{L_2(0,t_j;H')} + \frac{2}{c^2}\|\phi\|^2_{L_2(\mathcal{J})} \sum_{q=1}^{j} k_q \alpha_{q-1}.$$

Clearly, $\alpha_0 = 0$ and so we can simplify this to get,

$$\alpha_j \leq \frac{2}{c^2}\|L\|^2_{L_2(0,t_j;H')} + \frac{2}{c^2}\|\phi\|^2_{L_2(\mathcal{J})} \sum_{q=1}^{j-1} \rho_{q+1} k_q \alpha_q,$$

and Gronwall's lemma yields,

$$\|\boldsymbol{U}\|^2_{L_2(0,t_j;H)} \leq C_S^2 \|L\|^2_{L_2(0,t_j;H')},$$

for all $j \in \mathbb{N}(1, N)$ and where,

$$C_S^2 := \frac{2}{c^2} \exp\left(\frac{2}{c^2}\|\phi\|^2_{L_2(\mathcal{J})} \sum_{q=1}^{j-1} \rho_{q+1} k_q \right).$$

This proves the L_2 estimate. For the L_1 estimate we firstly obtain,

$$\int_{t_{q-1}}^{t_q} \sum_{m=1}^{q-1} \int_{t_{m-1}}^{t_m} \phi(t-s)\|\boldsymbol{U}_m(s)\|_H \, ds dt \leq \|\phi\|_{L_\infty(\mathcal{J})} k_q \|\boldsymbol{U}\|_{L_1(0,t_{q-1};H)},$$

and now returning to (29) and using this estimate along with our "equivalent norm" assumption yields,

$$\begin{aligned} c\|\boldsymbol{U}_q\|^2_{L_1(\mathcal{J}_q;H)} &\leq ck_q\|\boldsymbol{U}_q\|^2_{L_2(\mathcal{J}_q;H)}, \\ &\leq C_{L_1}\left(\|L\|_{L_1(\mathcal{J}_q;H')} + \|\phi\|_{L_\infty(\mathcal{J})} k_q \|\boldsymbol{U}\|_{L_1(0,t_{q-1};H)}\right)\|\boldsymbol{U}_q\|_{L_1(\mathcal{J}_q;H)}. \end{aligned}$$

Summing over $q = 1, 2, \ldots, j \leq N$ then gives,

$$\begin{aligned} \alpha_j := \|\boldsymbol{U}\|_{L_1(0,t_j;H)} &\leq \frac{C_{L_1}}{c}\|L\|_{L_1(0,t_j;H')} + \frac{C_{L_1}}{c}\|\phi\|_{L_\infty(\mathcal{J})} \sum_{q=1}^{j} k_q \alpha_{q-1}, \\ &\leq \frac{C_{L_1}}{c}\|L\|_{L_1(0,t_j;H')} + \frac{C_{L_1}}{c}\|\phi\|_{L_\infty(\mathcal{J})} \sum_{q=1}^{j-1} \rho_{q+1} k_q \alpha_q, \end{aligned}$$

and applying Gronwall's inequality once again we get,

$$\|\boldsymbol{U}\|_{L_1(0,t_j;H)} \leq C_S \|L\|_{L_1(0,t_j;H')}$$

for all $j \in \mathbb{N}(1,N)$ and where this time we set,

$$C_S := \frac{C_{L_1}}{c} \exp\left(\frac{C_{L_1}}{c} \|\phi\|_{L_\infty(\mathcal{J})} \sum_{q=1}^{j-1} \rho_{q+1} k_q \right).$$

Finally, to prove the L_∞ estimate we again return to (29) and use our "inverse estimate" to obtain this time,

$$\|\boldsymbol{U}_q\|_{L_\infty(\mathcal{J}_q;H)} \leq C\|L\|_{L_\infty(\mathcal{J}_q;H')} + C\|\phi\|_{L_\infty(\mathcal{J})} \sum_{i=1}^{q-1} k_i \|\boldsymbol{U}_i\|_{L_\infty(\mathcal{J}_i;H)},$$

one more application of Gronwall's lemma then completes the proof. □

We can move on now and derive our *a priori* Galerkin energy-error estimate.

5. *A priori* error estimate. In the methodology of Johnson *et al.* a continuous dual backward problem is used to derive the *a posteriori* error estimates and a discrete backward problem is used for deriving the *a priori* error estimates. In our case the appropriate discrete backward problem is: find $\boldsymbol{X} \in V_r$ such that

$$a^*(\boldsymbol{X}, \boldsymbol{v}) = l^*(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in V_r, \tag{30}$$

where

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

for some $L^* \in L_1(\mathcal{J}; H')$, and where (compare (15)):

$$\begin{aligned} a^*(\boldsymbol{\chi}, \boldsymbol{v}) &:= \int_0^T A(\boldsymbol{\chi}(t), \boldsymbol{v}(t))\, dt - \int_0^T \int_t^T B^*(s,t; \boldsymbol{\chi}(s), \boldsymbol{v}(t))\, ds dt, \\ B^*(s,t;\boldsymbol{\chi}(s), \boldsymbol{v}(t)) &:= B(s,t; \boldsymbol{v}(t), \boldsymbol{\chi}(s)), \end{aligned}$$

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 (15). (Note that due to Remark 1 the dual problem may have a slightly different character to (15).)

LEMMA 9 (**duality**). *$a^*(\boldsymbol{w}, \boldsymbol{v}) = a(\boldsymbol{v}, \boldsymbol{w})$ $\forall \boldsymbol{v} \in L_\infty(\mathcal{J}; H)$ and $\forall \boldsymbol{w} \in L_1(\mathcal{J}; H)$.*

In [17] we demonstrate how a simple change of variables can cast a continuous backward problem into a continuous forward problem, and that this allows the data stability estimate of Theorem 3 to be applied to the backward problem. Clearly the same technique will work also for the discrete backward problem and so from Theorem 8 we immediately get the following discrete backward data-stability bound. There exists a positive constant C_S such that

$$\|\boldsymbol{X}\|_{L_1(\mathcal{J};H)} \leq C_S \|L^*\|_{L_1(\mathcal{J};H')}. \tag{31}$$

We begin now with a simple lemma.

LEMMA 10 (**continuity of** $a(\,\cdot\,,\,\cdot\,)$). *Let Assumptions 2 hold, then the bilinear form $a(\,\cdot\,,\,\cdot\,)$ is bounded in the following sense:*

$$|a(\boldsymbol{w}, \boldsymbol{v})| \leq (1 + \|\phi\|_{L_1(\mathcal{J})}) \|\boldsymbol{w}\|_{L_p(\mathcal{J};H)} \|\boldsymbol{v}\|_{L_q(\mathcal{J};H)},$$

for all $\boldsymbol{w} \in L_p(\mathcal{J};H)$*, for all* $\boldsymbol{v} \in L_q(\mathcal{J};H)$*, and for all conjugate Hölder indices* $1 \le p,q \le \infty$.

Proof. Using (16) with Assumptions 2 we use Hölder's inequality twice: first for Laplace-type convolution, and then in a standard way. This gives:

$$\begin{aligned}
|a(\boldsymbol{w},\boldsymbol{v})| &= \left| \int_0^T A(\boldsymbol{w},\boldsymbol{v}) - \int_0^t B(t,s;\boldsymbol{w}(s),\boldsymbol{v}(t))\,dsdt \right|, \\
&\le \int_0^T \|\boldsymbol{w}\|_H \|\boldsymbol{v}\|_H\,dt + \int_0^T \int_0^t \phi(t-s)\|\boldsymbol{w}(s)\|_H\,ds\ \|\boldsymbol{v}\|_H\,dt, \\
&\le \|\boldsymbol{w}\|_{L_p(\mathcal{J};H)}\|\boldsymbol{v}\|_{L_q(\mathcal{J};H)} + \|\phi\|_{L_1(\mathcal{J})}\|\boldsymbol{w}\|_{L_p(\mathcal{J};H)}\|\boldsymbol{v}\|_{L_q(\mathcal{J};H)},
\end{aligned}$$

as required. □

The next step is to define the space of "semidiscrete dual functions",

$$H_{\mathcal{J}} := \Big\{ \boldsymbol{v} \in L_1(\mathcal{J};H) : \boldsymbol{v}|_{\mathcal{J}_i} \in L_1(\mathcal{J}_i;H_i)\ \forall i \in \mathbb{N}(1,N) \Big\}, \tag{32}$$

and let $P_r : H_{\mathcal{J}} \to V_r$ denote the L_2-energy projection of the semidiscrete space $H_{\mathcal{J}}$ onto the finite element space V_r. To be precise, for $\boldsymbol{w} \in H_{\mathcal{J}}$ we obtain $P_r\boldsymbol{w}$ by solving the variational problem:

$$\int_0^T A(P_r\boldsymbol{w},\boldsymbol{v})\,dt = \int_0^T A(\boldsymbol{w},\boldsymbol{v})\,dt \qquad \forall \boldsymbol{v} \in V_r, \tag{33}$$

we then have the following stability estimate for this projection.

LEMMA 11 (P_r **stability**). *For* $r = 0$ *and* $r = 1$ *there are constants* C_{P_r} *such that*:

$$\|P_r\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)} \le C_{P_r}\|\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)} \qquad \forall \boldsymbol{w} \in H_{\mathcal{J}},$$

and for all $i \in \mathbb{N}(1,N)$*. Suitable values are,* $C_{P_0} = 1$ *and* $C_{P_1} = \sqrt{50} \approx 7.1$.

Proof. For $r = 0$ we choose an $i \in \mathbb{N}(1,N)$ and take $\boldsymbol{v} = \boldsymbol{0}$ everywhere in $\mathcal{J}$ except in $\mathcal{J}_i$ where we choose $\boldsymbol{v} = P_0\boldsymbol{w}$. Hence $\boldsymbol{v} \in V_0$ and:

$$\begin{aligned}
\|P_0\boldsymbol{w}\|^2_{L_1(\mathcal{J}_i;H)} &= k_i \int_{t_{i-1}}^{t_i} A(P_0\boldsymbol{w},P_0\boldsymbol{w})\,dt = k_i \int_{t_{i-1}}^{t_i} A(\boldsymbol{w},P_0\boldsymbol{w})\,dt, \\
&\le k_i \int_{t_{i-1}}^{t_i} \|\boldsymbol{w}\|_H\,\|P_0\boldsymbol{w}\|_H\,dt = \|P_0\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)}\,\|\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)},
\end{aligned}$$

as required. For $r = 1$ see Appendix A. □

We also need some approximation estimates for P_r, and to begin we introduce the Ritz (elliptic) projection $R\boldsymbol{u}$ of $\boldsymbol{u}$, defined piecewise by $R_q : H \to H_q$ where,

$$A(R_q\boldsymbol{u},\boldsymbol{v}) = A(\boldsymbol{u},\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in H_q. \tag{34}$$

LEMMA 12. *Assume that the interpolators* $\{\pi_q\}$ *given in Assumptions 4 satisfy the estimate* $\|D_t\pi\boldsymbol{\chi}\|_H \le \Pi_{P_r}\|D_t\boldsymbol{\chi}\|_H$ *a.e. in* $\mathcal{J}$ *for a positive constant* Π_{P_r}*, then for the* L_2*-energy projection just described in* (33) *we have the following error estimate,*

$$\|k_q^{-1}(I-P_r)\pi_q\boldsymbol{\chi}\|_{L_p(\mathcal{J}_q;H)} \le \Pi_{P_r}\|D_t\boldsymbol{\chi}\|_{L_p(\mathcal{J}_q;H)} \qquad \forall q \in \mathbb{N}(1,N). \tag{35}$$

Moreover, if the Ritz projection defined in (34) *satisfies* $\|D_t^{r+1}R\boldsymbol{u}\|_H \le \|D_t^{r+1}\boldsymbol{u}\|_H$ *a.e. in* $\mathcal{J}$ *for* $r=0$ *or* $r=1$, *then we also have,*

$$\|(I-P_r)R\boldsymbol{u}\|_{L_p(\mathcal{J};H)} \le \Pi_k \left\|k^{r+1}D_t^{r+1}\boldsymbol{u}\right\|_{L_p(\mathcal{J};H)}, \tag{36}$$

where $\Pi_k = 4^r$ *and in both of these estimates* $p \in [1,\infty]$.

Proof. See Appendix B. □

Now we can state the *a priori* error estimate.

THEOREM 13 (***A priori* Galerkin energy-error estimate**). *Let Assumptions 2 and 4 hold. Further, let the backward stability estimate* (31) *hold* (*i.e.* $\underline{\boldsymbol{D}}(t)$ *is smooth enough in the case* $r=1$) *and assume the following.*

(i) *The solution* $\boldsymbol{u}$ *to* (15) *satisfies*

$$\boldsymbol{u} \in L_\infty(\mathcal{J};\boldsymbol{H}^2(\Omega)) \cap W_\infty^{r+1}(\mathcal{J};H).$$

(ii) *The spatial discretization is such that the Ritz projection* $R\boldsymbol{u}$ *of* $\boldsymbol{u}$ *defined previously in* (34) *satisfies the standard error estimate,*

$$\|\boldsymbol{u}-R\boldsymbol{u}\|_{L_\infty(\mathcal{J};H)} \le \Pi_h \|hD^2\boldsymbol{u}\|_{L_\infty(\mathcal{J};\boldsymbol{L}_2(\Omega))},$$

for a positive constant Π_h (*see for example* [3]).

Then for approximation in V_r, *for* $r=0,1$, *the Galerkin error* $\boldsymbol{e} := \boldsymbol{u}-\boldsymbol{U}$ *satisfies the a priori error estimate,*

$$\|\boldsymbol{u}-\boldsymbol{U}\|_{L_\infty(\mathcal{J};H)} \le C(T)\left(\Pi_h \left\|hD^2\boldsymbol{u}\right\|_{L_\infty(\mathcal{J};\boldsymbol{L}_2(\Omega))} + \Pi_k \left\|k^{r+1}\frac{\partial^{r+1}\boldsymbol{u}}{\partial t^{r+1}}\right\|_{L_\infty(\mathcal{J};H)}\right),$$

where:

$$C(T) := 1 + C_S(1+\|\phi\|_{L_1(\mathcal{J})});$$

and, C_S *is the stability constant in* (31). *This estimate holds for* $r=1$ *only if each* k_q *is small enough in the sense required by Lemma 7.*

Proof. Recall P_r as defined in (33), then for every $q \in \mathbb{N}(1,N)$ define $\boldsymbol{E} \in V_r$ by $\boldsymbol{E}|_{\mathcal{J}_q} := \boldsymbol{U} - P_rR_q\boldsymbol{u}$ and take $\boldsymbol{v} = \boldsymbol{E}$ in the discrete backward problem (30). Using Lemma 9 and the Galerkin orthogonality (20) we then get,

$$\begin{aligned} l^*(\boldsymbol{E}) &= a^*(\boldsymbol{X},\boldsymbol{E}) = a(\boldsymbol{E},\boldsymbol{X}) = a(\boldsymbol{E},\boldsymbol{X}) + a(\boldsymbol{e},\boldsymbol{X}),\\ &= a(\boldsymbol{u}-P_rR\boldsymbol{u},\boldsymbol{X}). \end{aligned}$$

By Lemma 10 and the discrete backward data-stability estimate (31) we now use this representation to get,

$$\begin{aligned} |l^*(\boldsymbol{E})| &\le (1+\|\phi\|_{L_1(\mathcal{J})})\|\boldsymbol{u}-P_rR\boldsymbol{u}\|_{L_\infty(\mathcal{J};H)}\|\boldsymbol{X}\|_{L_1(\mathcal{J};H)},\\ &\le C_S(1+\|\phi\|_{L_1(\mathcal{J})})\|\boldsymbol{u}-P_rR\boldsymbol{u}\|_{L_\infty(\mathcal{J};H)}\|L^*\|_{L_1(\mathcal{J};H')}, \end{aligned}$$

and it remains to interpret the left hand side.

From Assumptions 2, part (i) we see that $(H, A(\,\cdot\,,\,\cdot\,))$ is a Hilbert space and so we may use the Riesz theorem with $\boldsymbol{E} \in L_\infty(\mathcal{J};H)$ to uniquely define $w_{\boldsymbol{E}} \in L_\infty(\mathcal{J};H')$ via

$$A(\boldsymbol{E},\,\cdot\,) = \langle w_{\boldsymbol{E}},\,\cdot\,\rangle_H \qquad \text{a.e. in } \mathcal{J}.$$

Moreover, $\|w_{\boldsymbol{E}}\|_{H'} = \|\boldsymbol{E}\|_H$ a.e. We now use this observation to derive a bound for the right hand side of the dual problem. This is the key to providing an L_∞-energy (as opposed to $L_\infty(L_2)$) error estimate. In the discrete backward problem (30) we take $L^* := zw_{\boldsymbol{E}}$ where $z \in L_1(\mathcal{J})$ is arbitrary, then $L^* \in L_1(\mathcal{J}; H')$ and satisfies $\|L^*\|_{L_1(\mathcal{J};H')} \leq \|z\|_{L_1(\mathcal{J})}\|\boldsymbol{E}\|_{L_\infty(\mathcal{J};H)}$ and, moreover,

$$\sup\left\{|l^*(\boldsymbol{E})| : L^* = zw_{\boldsymbol{E}},\ \|z\|_{L_1(\mathcal{J})} = 1\right\} = \|\boldsymbol{E}\|^2_{L_\infty(\mathcal{J};H)}.$$

To see this we note that,

$$l^*(\boldsymbol{E}) = \int_0^T \langle L^*, \boldsymbol{E}\rangle_H\, dt = \int_0^T \langle w_{\boldsymbol{E}}, \boldsymbol{E}\rangle_H z\, dt = \int_0^T \|\boldsymbol{E}\|_H^2 z\, dt,$$

and from this we get,

$$\big\|\|\boldsymbol{E}\|_H^2\big\|_{L_\infty(\mathcal{J})} = \sup\left\{\left|\int_0^T \|\boldsymbol{E}\|_H^2 z\, dt\right| : \|z\|_{L_1(\mathcal{J})} = 1\right\},$$

since $L_\infty(\mathcal{J})$ is dual to $L_1(\mathcal{J})$. Using these in our original inequality we obtain,

$$\|\boldsymbol{E}\|^2_{L_\infty(\mathcal{J};H)} = \sup_z\{|l^*(\boldsymbol{E})|\} \leq C_S(1 + \|\phi\|_{L_1(\mathcal{J})})\|\boldsymbol{u} - P_rR\boldsymbol{u}\|_{L_\infty(\mathcal{J};H)}\|\boldsymbol{E}\|_{L_\infty(\mathcal{J};H)},$$

which implies,

$$\|\boldsymbol{E}\|_{L_\infty(\mathcal{J};H)} \leq C_S(1 + \|\phi\|_{L_1(\mathcal{J})})\|\boldsymbol{u} - P_rR\boldsymbol{u}\|_{L_\infty(\mathcal{J};H)}.$$

Applying the triangle inequality we get,

$$\begin{aligned} \|\boldsymbol{u} - \boldsymbol{U}\|_{L_\infty(\mathcal{J};H)} &\leq \|\boldsymbol{u} - P_rR\boldsymbol{u}\|_{L_\infty(\mathcal{J};H)} + \|P_rR\boldsymbol{u} - \boldsymbol{U}\|_{L_\infty(\mathcal{J};H)}, \\ &\leq \Big(1 + C_S(1 + \|\phi\|_{L_1(\mathcal{J})})\Big)\|\boldsymbol{u} - P_rR\boldsymbol{u}\|_{L_\infty(\mathcal{J};H)}, \end{aligned}$$

and the theorem now follows by using

$$\|\boldsymbol{u} - P_rR\boldsymbol{u}\|_{L_\infty(\mathcal{J};H)} \leq \|\boldsymbol{u} - R\boldsymbol{u}\|_{L_\infty(\mathcal{J};H)} + \|R\boldsymbol{u} - P_rR\boldsymbol{u}\|_{L_\infty(\mathcal{J};H)}$$

along with Lemma 12 and our assumption for R. □

6. Numerical implementation. In this section we give some details of the practical implementation of the numerical algorithms represented by (19) in the two cases $r = 0$ and $r = 1$. We will see that the case $r = 0$ yields a system of second-kind Volterra equations for the piecewise constant (in time) vectors of nodal displacements, while the $r = 1$ scheme results in a pair of simultaneous systems of Volterra equations for the two nodal solution vectors at each time level. For this section we recall the notation introduced earlier at the beginning of § 4.

First we choose a time slab $\mathcal{J}_q$ and note that since $\boldsymbol{U}$ is only piecewise continuous in time we may pose the discrete problem (19) on this time level in the following way. Given the solution history $\{\boldsymbol{U}_p\}_{p=1}^{q-1}$, find $\boldsymbol{U}_q := \boldsymbol{U}|_{\mathcal{J}_q}$ such that,

$$a(\boldsymbol{U}, \boldsymbol{v}) = l(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in \mathbb{P}_r(\mathcal{J}_q; H_q).$$

Using (28) this problem may be written more explicitly as,

$$\int_{t_{q-1}}^{t_q} A_0(t-t_{q-1};\boldsymbol{U}_q(t_{q-1}),\boldsymbol{v}(t)) + A_1(t,t_{q-1};\dot{\boldsymbol{U}}_q,\boldsymbol{v}(t))\,dt = \int_{t_{q-1}}^{t_q} L(t;\boldsymbol{v}(t))\,dt$$

$$+\int_{t_{q-1}}^{t_q}\sum_{p=1}^{q-1}\int_{t_{p-1}}^{t_p} B(t,s;\boldsymbol{U}_p(s),\boldsymbol{v}(t))\,dsdt \qquad \forall \boldsymbol{v}\in\mathbb{P}_r(\mathcal{J}_q;H_q). \tag{37}$$

Let $\{\psi_{qi}\}_{i=1}^{I_q}$ be the canonical piecewise linear basis relative to the mesh Ω_q. Then, taking $\Omega\subset\mathbb{R}^3$ for example, we may construct bases for each of the H_q as $\{\boldsymbol{\psi}_{qi}\}_{i=1}^{3I_q}$ where:

$$\boldsymbol{\psi}_{qi} = \begin{cases} (\psi_{qi}\ 0\ 0)^{\mathrm{T}} & \text{for } i=1,4,7,\ldots,I_{q-2},\\ (0\ \psi_{qi}\ 0)^{\mathrm{T}} & \text{for } i=2,5,8,\ldots,I_{q-1},\\ (0\ 0\ \psi_{qi})^{\mathrm{T}} & \text{for } i=3,6,9,\ldots,I_q. \end{cases}$$

Introducing the nodal degrees of freedom $\{\mathsf{U}_{qi}(t)\}_{i=1}^{3I_q}$ we then write (in the usual way),

$$\boldsymbol{U}_q(\boldsymbol{x},t) = \sum_{i=1}^{3I_q}\mathsf{U}_{qi}(t)\boldsymbol{\psi}_{qi}(\boldsymbol{x}).$$

Let $\varphi_q\in\mathbb{P}_r(\mathcal{J}_q;\mathbb{R})$ and in (37) take $\boldsymbol{v}$ to be $\varphi_q\boldsymbol{\psi}_{qj}$ for each $j\in\mathbb{N}(1,3I_q)$ in turn. This gives

$$\int_{t_{q-1}}^{t_q}\mathbf{A}_q^{(0)}(t)\mathbf{U}_q(t_{q-1})\varphi_q(t)\,dt + \int_{t_{q-1}}^{t_q}\mathbf{A}_q^{(1)}(t)\dot{\mathbf{U}}_q\varphi_q(t)\,dt$$

$$= \mathbf{F}_q + \int_{t_{q-1}}^{t_q}\sum_{p=1}^{q-1}\int_{t_{p-1}}^{t_p}\mathbf{B}_{qp}(t,s)\mathbf{U}_p(s)\varphi_q(t)\,dsdt, \tag{38}$$

where:

$$\begin{aligned}\mathbf{A}_q^{(0)}(t) &:= \big(A_0(t-t_{q-1};\boldsymbol{\psi}_{qi},\boldsymbol{\psi}_{qj})\big)_{i,j=1}^{3I_q},\\ \mathbf{A}_q^{(1)}(t) &:= \big(A_1(t,t_{q-1};\boldsymbol{\psi}_{qi},\boldsymbol{\psi}_{qj})\big)_{i,j=1}^{3I_q},\end{aligned}$$

are square "stiffness matrices", and,

$$\mathbf{F}_q := \left(\int_{t_{q-1}}^{t_q}\varphi_q(t)L(t;\boldsymbol{\psi}_{qi})\,dt\right)_{i=1}^{3I_q}$$

is a "load vector". In addition, for $t\in\mathcal{J}_q$ and $s\in\mathcal{J}_p$,

$$\mathbf{B}_{qp}(t,s) := \big(B(t,s;\boldsymbol{\psi}_{pi},\boldsymbol{\psi}_{qj})\big)_{i=1,j=1}^{i=3I_p,j=3I_q} \tag{39}$$

is in general a non-square matrix which effectively maps previous solutions onto the current approximating space H_q. This term is potentially very expensive to calculate in the non-nested refinement (or de-refinement) case where $H_p\not\subseteq H_q$, since it is not trivial to represent a function $\boldsymbol{W} \in H_p$ as an "equivalent" function $\boldsymbol{W} \in H_q$ (for $p < q$). Note that we are using an upright bold sans-serif font (as in A) for matrices and vectors resulting from spatial semi-discretization, and an upright bold roman font (as in $\mathbf{F}$) for those resulting from full space-time discretization.

However, in the case where we allow only nested refinements of the space mesh, functions in H_p can be easily and exactly represented in H_q (for $p < q$) by a straightforward interpolation of the extra degrees of freedom onto the extra nodes on the mesh Ω_q. This restriction can result in substantial cost savings in the resulting algorithm when the relaxation functions are in the form of a Dirichlet-Prony series as given by (7). We give details below.

The case $r = 0$. For the case of piecewise constant temporal approximation using the finite element space V_0 we have $\dot{\boldsymbol{U}}_q = \mathbf{0}$ on $\mathcal{J}_q$ and we take $\varphi_q(t) = 1$ in (38). In addition, $\mathsf{U}_q(t_{q-1}) = \mathsf{U}_q =$ a constant vector, and so the numerical scheme simplifies considerably to become: for each $q \in \mathbb{N}(1, N)$ in turn, find $\mathbf{U}_q \equiv \mathsf{U}_q$ such that,

$$\mathbf{A}_q^{(0)} \mathbf{U}_q = \mathbf{F}_q + \sum_{p=1}^{q-1} \mathbf{B}_{qp} \mathbf{U}_p, \tag{40}$$

where,

$$\begin{aligned} \mathbf{A}_q^{(0)} &:= \int_{t_{q-1}}^{t_q} \mathsf{A}_q^{(0)}(t)\, dt, \\ \mathbf{B}_{qp} &:= \int_{t_{q-1}}^{t_q} \int_{t_{p-1}}^{t_p} \mathsf{B}_{qp}(t, s)\, ds dt, \end{aligned}$$

and $\mathbf{F}_q$ is as given previously with $\varphi_q(t) := 1$. Note that the number of operations required to compute $\mathbf{U}_q$ is in general proportional to q^2. This is a severe limitation and we now demonstrate how in an important special case it can be overcome.

In the case of nested refinements we have $H_p \subseteq H_q$ for all $p < q$ and so

$$\boldsymbol{U}_p \in H_p \quad \Longrightarrow \quad \boldsymbol{U}_p \in H_q,$$

which means that we may describe $\boldsymbol{U}_p$ using the current basis $\{\boldsymbol{\psi}_{qi}\}_{i=1}^{3I_q}$ for H_q and construct $\mathsf{B}_{qp}(t, s)$ as a square matrix. Interpolating the extra nodal values into $\mathbf{U}_p$ we then see that the vector,

$$\mathbf{W}_p := \mathbf{B}_{qp} \mathbf{U}_p,$$

describes a function $\boldsymbol{W}_q \in H_q$, and trivially also (again by interpolation) a function $\boldsymbol{W}_{q+1} \in H_{q+1}$, since $H_q \subseteq H_{q+1}$.

Combining this restriction on the space meshes with the Dirichlet-Prony series representation of the relaxation functions now results in the sum on the right of (40) effectively collapsing, taking the operation count of order $O(q^2)$ with it to leave a count of only order $O(q)$. To see this we make the simplifying assumption in (6) that there is a single stress relaxation function $\phi(t)$ such that

$$D_{ijkl}(\boldsymbol{x}, t) \equiv \phi(t) D_{ijkl}(\boldsymbol{x}),$$

where $\{D_{ijkl}(\boldsymbol{x})\}$ on the right now represents a temporally constant tensor and,

$$\phi(t) := \phi_0 + \sum_{m=1}^{N_\phi} \phi_m \exp(-\alpha_m t),$$

and is normalized so that $\phi(0) = 1$. Such viscoelastic materials are sometimes described as *synchronous*.

Recalling the definitions (11) and (12) it is clear that

$$B(t, s; \boldsymbol{w}, \boldsymbol{v}) = \phi_s(t-s) A(\boldsymbol{w}, \boldsymbol{v}) = \sum_{m=1}^{N_\phi} \alpha_m \phi_m e^{-\alpha_m (t-s)} A(\boldsymbol{w}, \boldsymbol{v}),$$

and from this it follows that (since $H_p \subseteq H_q$)

$$\mathsf{B}_{qp}(t,s) = \Big(A(\boldsymbol{\psi}_{qi}, \boldsymbol{\psi}_{qj})\Big)_{i,j=1}^{3I_q} \sum_{m=1}^{N_\phi} \alpha_m \phi_m e^{-\alpha_m (t-s)}.$$

However, $A(\,\cdot\,,\,\cdot\,) \equiv A_0(0;\,\cdot\,,\,\cdot\,)$ and so

$$\mathsf{B}_{qp}(t,s) = \mathsf{A}_q^{(0)}(t_{q-1}) \sum_{m=1}^{N_\phi} \alpha_m \phi_m e^{-\alpha_m (t-s)},$$

which implies that

$$\begin{aligned} \mathbf{B}_{qp} &= \mathsf{A}_q^{(0)}(t_{q-1}) \int_{t_{q-1}}^{t_q} \int_{t_{p-1}}^{t_p} \sum_{m=1}^{N_\phi} \alpha_m \phi_m e^{-\alpha_m (t-s)} \, ds dt, \\ &= \mathsf{A}_q^{(0)}(t_{q-1}) \sum_{m=1}^{N_\phi} \beta_m(q,p), \end{aligned}$$

where we define

$$\beta_m(q,p) := \frac{\phi_m}{\alpha_m} (1 - e^{-\alpha_m k_q})(e^{\alpha_m k_p} - 1) e^{-\alpha_m (t_{q-1} - t_{p-1})}.$$

The point of these manipulations is as follows. Set

$$\mathbf{W}_{qp}^{(m)} := \sum_{p=1}^{q-1} \beta_m(q,p) \mathsf{A}_q^{(0)}(t_{q-1}) \mathbf{U}_p,$$

and then the sum on the right hand side of (40) becomes

$$\sum_{p=1}^{q-1} \mathbf{B}_{qp} \mathbf{U}_p \equiv \sum_{m=1}^{N_\phi} \mathbf{W}_{qp}^{(m)}.$$

Suppose each of the vectors in $\{\mathbf{W}_{qp}^{(m)}\}_{m=1}^{N_\phi}$ are available when we solve for $\mathbf{U}_q$ on the time slab $\mathcal{J}_q$, then to advance the algorithm forward to $\mathcal{J}_{q+1}$ we first note the simple update,

$$\beta_m(q+1, p) = \left(\frac{1 - e^{-\alpha_m k_{q+1}}}{e^{\alpha_m k_q} - 1}\right) \beta_m(q,p),$$

and then interpolate the $\mathbf{W}_{qp}^{(m)}$ on to the new space mesh Ω_{q+1}, and thus into the new approximating space H_{q+1}. In effect this gives us the vectors,

$$\mathbf{W}_{qp}^{(m)} = \sum_{p=1}^{q-1} \beta_m(q,p)\mathsf{A}_{q+1}^{(0)}(t_q)\mathbf{U}_p,$$

for $m = 1, \ldots, N_\phi$. Now, by definition,

$$\sum_{p=1}^{q} \mathbf{B}_{(q+1)p}\mathbf{U}_p = \sum_{m=1}^{N_\phi} \mathbf{W}_{(q+1)p}^{(m)},$$

where,

$$\begin{aligned}
\mathbf{W}_{(q+1)p}^{(m)} &= \beta_m(q+1,q)\mathsf{A}_{q+1}^{(0)}(t_q)\mathbf{U}_q + \sum_{p=1}^{q-1} \beta_m(q+1,p)\mathsf{A}_{q+1}^{(0)}(t_q)\mathbf{U}_p, \\
&= \beta_m(q+1,q)\mathsf{A}_{q+1}^{(0)}(t_q)\mathbf{U}_q + \left(\frac{1-e^{-\alpha_m k_{q+1}}}{e^{\alpha_m k_q}-1}\right) \sum_{p=1}^{q-1} \beta_m(q,p)\mathsf{A}_{q+1}^{(0)}(t_q)\mathbf{U}_p.
\end{aligned}$$

Therefore, once we have augmented $\mathbf{W}_{qp}^{(m)}$ by interpolating the new degrees of freedom in the mesh Ω_{q+1}, we can obtain $\mathbf{W}_{(q+1)p}^{(m)}$ simply by scaling $\mathbf{W}_{qp}^{(m)}$ by a quantity depending on the two most recent time steps, and adding on a contribution from the previous solution:

$$\mathbf{W}_{(q+1)p}^{(m)} = \beta_m(q+1,q)\mathsf{A}_{q+1}^{(0)}(t_q)\mathbf{U}_q + \left(\frac{1-e^{-\alpha_m k_{q+1}}}{e^{\alpha_m k_q}-1}\right)\mathbf{W}_{qp}^{(m)}.$$

At the next time step the algorithm (40) becomes:

$$\mathbf{A}_{q+1}^{(0)}\mathbf{U}_{q+1} = \mathbf{F}_{q+1} + \sum_{m=1}^{N_\phi} \mathbf{W}_{(q+1)p}^{(m)}.$$

With more care over notation this scheme can be extended to asynchronous viscoelastic materials (as described by (7)), and so we see that the restriction of nested mesh refinements coupled with a Dirichlet-Prony series form of the relaxation functions results in a much faster algorithm due to the history summation in (40) being effectively eliminated.

The case $r = 1$. We now rather more briefly consider the form of the numerical algorithm when using the temporally piecewise linear approximating space V_1. We again begin with (38), but this time we have

$$\boldsymbol{U}_q(t) = \boldsymbol{U}_q^-\varphi_q^-(t) + \boldsymbol{U}_q^+\varphi_q^+(t)$$

(recall the notation at the beginning of § 4), and in (38) take $\varphi_q(t)$ as $\varphi_q^-(t)$ and $\varphi_q^+(t)$ in turn. This gives two systems of linear equations in the two unknown vectors $\mathbf{U}_q^-$ and $\mathbf{U}_q^+$:

$$\begin{aligned}
&\int_{t_{q-1}}^{t_q} \mathsf{A}_q^{(0)}(t)\mathsf{U}_q^-\varphi_q^\pm(t)\,dt + \frac{1}{k_q}\int_{t_{q-1}}^{t_q} \mathsf{A}_q^{(1)}(t)\mathsf{U}_q^+\varphi_q^\pm(t)\,dt \\
&\qquad = \mathbf{F}_q^\pm + \sum_{p=1}^{q-1}\int_{t_{q-1}}^{t_q}\int_{t_{p-1}}^{t_p} \left[\mathsf{B}_{qp}(t,s)\mathsf{U}_p^-\varphi_p^-(s) + \mathsf{B}_{qp}(t,s)\mathsf{U}_p^+\varphi_p^+(s)\right]\varphi_q^\pm(t)\,dsdt,
\end{aligned}$$

where now

$$\mathbf{F}_q^{\pm} := \left(\int_{t_{q-1}}^{t_q} \varphi_q^{\pm}(t) L(t; \boldsymbol{\psi}_{qi})\, dt \right)_{i=1}^{3I_q}.$$

Define

$$\begin{aligned}
\mathbf{A}_q^{(1,1)} &:= \int_{t_{q-1}}^{t_q} \mathsf{A}_q^{(0)}(t)\varphi_q^{-}(t)\, dt,\\
\mathbf{A}_q^{(1,2)} &:= k_q^{-1} \int_{t_{q-1}}^{t_q} \mathsf{A}_q^{(1)}(t)\varphi_q^{-}(t)\, dt,\\
\mathbf{A}_q^{(2,1)} &:= \int_{t_{q-1}}^{t_q} \mathsf{A}_q^{(0)}(t)\varphi_q^{+}(t)\, dt,\\
\mathbf{A}_q^{(2,2)} &:= k_q^{-1} \int_{t_{q-1}}^{t_q} \mathsf{A}_q^{(1)}(t)\varphi_q^{+}(t)\, dt,
\end{aligned}$$

and then the discrete scheme takes the form of a pair of simultaneous systems of linear equations:

$$\begin{aligned}
\mathbf{A}_q^{(1,1)}\mathbf{U}_q^{-} + \mathbf{A}_q^{(1,2)}\mathbf{U}_q^{+} &= \widetilde{\mathbf{F}}_q^{-},\\
\mathbf{A}_q^{(2,1)}\mathbf{U}_q^{-} + \mathbf{A}_q^{(2,2)}\mathbf{U}_q^{+} &= \widetilde{\mathbf{F}}_q^{+},
\end{aligned}$$

where

$$\widetilde{\mathbf{F}}_q^{\pm} := \mathbf{F}_q^{\pm} + \sum_{p=1}^{q-1} \int_{t_{q-1}}^{t_q} \int_{t_{p-1}}^{t_p} \mathsf{B}_{qp}(t,s) \left[\mathsf{U}_p^{-}\varphi_p^{-}(s) + \mathsf{U}_p^{+}\varphi_p^{+}(s) \right] \varphi_q^{\pm}(t)\, dsdt.$$

For the case where only nested space-mesh refinements are permitted, and where the relaxation functions are in the form of a Dirichlet-Prony series, one could again exploit the recursion property of the exponential function to collapse the history sum contained in the two right hand sides. The details are more involved and so we leave them for another time.

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

Variational crimes. In practice the relaxation functions are usually 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 $\boldsymbol{g}$. This should be fairly straightforward and an example of such an analysis is given in the context of a model pure-time problem in [13].

Non-nested refinements. In the discussion in § 6 we pointed out the difficulties in implementation that can arise when the spatial approximating spaces are not nested in the sense that $H_{q-1} \not\subseteq H_q$ for $q = 2, 3, \ldots, N$. This is a particularly severe restriction when a singular feature of the solution attenuates over time and therefore makes it desirable to de-refine. We therefore suggest three possible ways in which this nested mesh requirement can be overcome.

(i) At every time level $\mathcal{J}_q$ we use the $\boldsymbol{L}_2(\Omega)$ projection to generate the map $H_{q-1} \to H_q$, which is then used to update the entire solution history onto the current space mesh. This will considerably simplify the construction of the (square) matrix $\mathbf{B}_{qp}(t,s)$ in (39) since one could substitute the current basis $\{\boldsymbol{\psi}_{qj}\}_j$ for $\{\boldsymbol{\psi}_{pi}\}_i$. The projection could be simplified by "lumping the mass".

(ii) Every space mesh Ω_q should be an instance or de-refinement of an *a priori* determined parent mesh. This should provide a reasonably simple mesh data structure to allow the integrations required in (39) to be carried out efficiently over the elements in the parent mesh. The results can then be "scattered" to the appropriate nodes on the current mesh. With this scheme one could also allow nested refinements of the parent mesh.

(iii) A third possibility is to keep the space-mesh constant through the time stepping and then use an *a posteriori* error estimate (see e.g. [17]) to refine it only when the error becomes unacceptably large. The algorithm is then halted and restarted with this adapted mesh, and this iteration continued until the final time T is reached without need for further refinement. The efficiency of this method will depend heavily upon the quality of the *a posteriori* estimator and the initial mesh.

Finally, let us remark that a non-smooth temporal finite element discretization as discussed in this article is not appropriate for hereditary viscoelasticity problems when the viscoelastic functions are not characterized by the Dirichlet-Prony series (7). This is because of the quadratic operation count induced by the persistence of the history term in (38). Note that the sparse quadrature rules developed in [20] will not apply to this non-smooth approximation. On the other hand, if one accepts the Dirichlet-Prony series representation then it could be that the internal variable formulation of A. Johnson, see for example [8], will provide an equal—if not preferable—alternative to the Volterra formulation described above. This is currently under investigation by the authors in collaboration with A. Johnson.

Appendix A. Proof of Lemma 11 for $r = 1$. For the case $r = 1$ we again choose $\boldsymbol{v} = \boldsymbol{0}$ everywhere except in $\mathcal{J}_i$, then $\boldsymbol{v} \in V_r$ and,

$$\int_{t_{i-1}}^{t_i} A\big(P_1\boldsymbol{w}, \boldsymbol{v}\big)\,dt = \int_{t_{i-1}}^{t_i} A\big(\boldsymbol{w}, \boldsymbol{v}\big)\,dt. \tag{41}$$

For convenience we write $\boldsymbol{W} := P_1\boldsymbol{w}$ as,

$$\boldsymbol{W} = \boldsymbol{W}_0\phi_0(t) + \boldsymbol{W}_1\phi_1(t),$$

where $\boldsymbol{W}_0$, $\boldsymbol{W}_1 \in H_i$, and the ϕ_0, $\phi_1 \in \mathbb{P}_1(\mathcal{J}_i;\mathbb{R})$ are temporal basis functions given by,

$$\phi_0(t) := 1 \qquad \text{and} \qquad \phi_1(t) := \frac{t - t_{i-1}}{k_i}.$$

For $\boldsymbol{V} \in H_i$ we take $\boldsymbol{v} = \boldsymbol{V}\phi_j$ for $j = 0, 1$, then from (41) we have,

$$\int_{t_{i-1}}^{t_i} A\big(\boldsymbol{W}_0\phi_0 + \boldsymbol{W}_1\phi_1, \boldsymbol{V}\phi_0\big)\,dt = \int_{t_{i-1}}^{t_i} A(\boldsymbol{w}, \boldsymbol{V}\phi_0)\,dt,$$

$$\int_{t_{i-1}}^{t_i} A\big(\boldsymbol{W}_0\phi_0 + \boldsymbol{W}_1\phi_1, \boldsymbol{V}\phi_1\big)\,dt = \int_{t_{i-1}}^{t_i} A(\boldsymbol{w}, \boldsymbol{V}\phi_1)\,dt,$$

for all $\boldsymbol{V} \in H_i$. These can be written in the simpler form:

$$\begin{aligned} A(\boldsymbol{W}_0, \boldsymbol{V}) + \frac{1}{2}A(\boldsymbol{W}_1, \boldsymbol{V}) &= F_0(\boldsymbol{V}) := \frac{1}{k_i}\int_{t_{i-1}}^{t_i} A(\boldsymbol{w}, \boldsymbol{V})\,dt, \\ \frac{1}{2}A(\boldsymbol{W}_0, \boldsymbol{V}) + \frac{1}{3}A(\boldsymbol{W}_1, \boldsymbol{V}) &= F_1(\boldsymbol{V}) := \frac{1}{k_i}\int_{t_{i-1}}^{t_i} A(\boldsymbol{w}, \boldsymbol{V})\phi_1\,dt. \end{aligned}$$

Solving these equations gives,

$$A(\boldsymbol{W}_0, \boldsymbol{V}) = 4F_0(\boldsymbol{V}) - 6F_1(\boldsymbol{V}) \qquad \text{and} \qquad A(\boldsymbol{W}_1, \boldsymbol{V}) = 12F_1(\boldsymbol{V}) - 6F_0(\boldsymbol{V}),$$

for all $\boldsymbol{V} \in H_i$ and it follows that,

$$\|\boldsymbol{W}_0\|_H^2 = 4F_0(\boldsymbol{W}_0) - 6F_1(\boldsymbol{W}_0) \qquad \text{and} \qquad \|\boldsymbol{W}_1\|_H^2 = 12F_1(\boldsymbol{W}_1) - 6F_0(\boldsymbol{W}_1).$$

We may now obtain our estimate for $\|P_1\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)}$ by direct calculation. Firstly:

$$\begin{aligned} \|P_1\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)} &= \int_{t_{i-1}}^{t_i} \|\boldsymbol{W}_0\phi_0 + \boldsymbol{W}_1\phi_1\|_H\,dt \le \int_{t_{i-1}}^{t_i} \phi_0\|\boldsymbol{W}_0\|_H + \phi_1\|\boldsymbol{W}_1\|_H\,dt, \\ &= k_i\|\boldsymbol{W}_0\|_H + \frac{1}{2}k_i\|\boldsymbol{W}_1\|_H. \end{aligned}$$

Squaring and using the inequality $ab \le a^2/2\epsilon + \epsilon b^2/2$ now gives,

$$\|P_1\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)}^2 \le \left(1 + \frac{1}{2\epsilon}\right)k_i^2\|\boldsymbol{W}_0\|_H^2 + \left(\frac{1}{4} + \frac{\epsilon}{2}\right)k_i^2\|\boldsymbol{W}_1\|_H^2 \tag{42}$$

for all $\epsilon > 0$. Now, using our previous results for the H norms of $\boldsymbol{W}_0$ and $\boldsymbol{W}_1$ we have:

$$\begin{aligned} k_i\|\boldsymbol{W}_0\|_H^2 &= \int_{t_{i-1}}^{t_i} A(\boldsymbol{w}, 4\boldsymbol{W}_0 - 6\phi_1\boldsymbol{W}_0)\,dt \le 4\|\boldsymbol{W}_0\|_H\,\|\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)}, \\ k_i\|\boldsymbol{W}_1\|_H^2 &= \int_{t_{i-1}}^{t_i} A(\boldsymbol{w}, 12\phi_1\boldsymbol{W}_1 - 6\boldsymbol{W}_1)\,dt \le 6\|\boldsymbol{W}_1\|_H\,\|\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)}, \end{aligned}$$

and from these it follows that,

$$\begin{aligned} k_i\|\boldsymbol{W}_0\|_H^2 &\le \frac{2}{\varepsilon_1}\|\boldsymbol{W}_0\|_H^2 + 2\varepsilon_1\|\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)}^2, \\ &\qquad \Longrightarrow \quad \|\boldsymbol{W}_0\|_H^2 \le \frac{2\varepsilon_1^2}{\varepsilon_1 k_i - 2}\|\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)}^2, \\ k_i\|\boldsymbol{W}_1\|_H^2 &\le \frac{3}{\varepsilon_2}\|\boldsymbol{W}_1\|_H^2 + 3\varepsilon_2\|\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)}^2, \\ &\qquad \Longrightarrow \quad \|\boldsymbol{W}_1\|_H^2 \le \frac{3\varepsilon_2^2}{\varepsilon_2 k_i - 3}\|\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)}^2, \end{aligned}$$

for all $\varepsilon_1 > 2/k_i$ and all $\varepsilon_2 > 3/k_i$. Using these in (42) then gives:

$$\|P_1\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)}^2 \le \left[\left(1 + \frac{1}{2\epsilon}\right)\frac{2\varepsilon_1^2 k_i^2}{\varepsilon_1 k_i - 2} + \left(\frac{1}{4} + \frac{\epsilon}{2}\right)\frac{3\varepsilon_2^2 k_i^2}{\varepsilon_2 k_i - 3}\right]\|\boldsymbol{w}\|_{L_1(\mathcal{J}_i;H)}^2.$$

To minimize the coefficient we choose $\varepsilon_1 = 4/k_i$, $\varepsilon_2 = 6/k_i$ and $\epsilon = 8/9$, thus:

$$\|P_1\boldsymbol{w}\|^2_{L_1(\mathcal{J}_i;H)} \le 50\|\boldsymbol{w}\|^2_{L_1(\mathcal{J}_i;H)},$$

and this completes the proof.

Appendix B. Proof of Lemma 12. Before giving the proof we first establish two preparatory lemmas, the first dealing with the case $r = 0$ and the second with the case $r = 1$. Recall also the definition of $H_{\mathcal{J}}$ in (32).

LEMMA 14 (**the case** $r = 0$). *Let $\boldsymbol{w} \in W^1_1(\mathcal{J};H) \cap H_{\mathcal{J}}$, then*

$$\|\boldsymbol{w}(s) - P_0\boldsymbol{w}(s)\|_H \le \int_{t_{i-1}}^{t_i} \|D_t\boldsymbol{w}\|_H\,dt,$$

for all $s \in \mathcal{J}_i$ and for all $i \in \mathbb{N}(1,N)$.

Proof. For convenience we write $\boldsymbol{W}(s) := P_0\boldsymbol{w}(s)$ for $s \in \mathcal{J}_i$ and $i \in \mathbb{N}(1,N)$ arbitrary. Then, from (33) we have

$$A(\boldsymbol{W}(s),\boldsymbol{V}) = \frac{1}{k_i}\int_{t_{i-1}}^{t_i} A(\boldsymbol{w}(s),\boldsymbol{V})\,dt \qquad \forall \boldsymbol{V} \in H_i,$$

where we chose $\boldsymbol{v} \in V_0$ such that $\boldsymbol{v} = \boldsymbol{0}$ on $\mathcal{J} \setminus \mathcal{J}_i$ and $\boldsymbol{v} = \boldsymbol{V} \in H_i$ on $\mathcal{J}_i$. The fundamental theorem of calculus then gives,

$$\begin{aligned}
A(\boldsymbol{w}(s)-\boldsymbol{W}(s),\boldsymbol{V}) &= A(\boldsymbol{w}(s),\boldsymbol{V}) - \frac{1}{k_i}\int_{t_{i-1}}^{t_i} A(\boldsymbol{w}(t),\boldsymbol{V})\,dt,\\
&= A(\boldsymbol{w}(s),\boldsymbol{V}) - \frac{1}{k_i}\int_{t_{i-1}}^{t_i} A\left(\boldsymbol{w}(s) + \int_s^t \boldsymbol{w}'(\xi)\,d\xi, \boldsymbol{V}\right)\,dt,\\
&= -\frac{1}{k_i}\int_{t_{i-1}}^{t_i}\int_s^t A(\boldsymbol{w}'(\xi),\boldsymbol{V})\,d\xi dt.
\end{aligned}$$

Now, for any $s \in \mathcal{J}_i$ we choose $\boldsymbol{V} := \boldsymbol{w}(s) - \boldsymbol{W}(s) \in H_i$ and arrive at,

$$\|\boldsymbol{w}(s)-\boldsymbol{W}(s)\|^2_H \le \|\boldsymbol{w}(s)-\boldsymbol{W}(s)\|_H \int_{t_{i-1}}^{t_i} \|\boldsymbol{w}'\|_H\,dt,$$

from which the lemma follows. ☐

LEMMA 15 (**the case** $r = 1$). *Let $\boldsymbol{w} \in W^2_1(\mathcal{J};H) \cap H_{\mathcal{J}}$, then*

$$\|\boldsymbol{w}(s) - P_1\boldsymbol{w}(s)\|_H \le 4k_i \int_{t_{i-1}}^{t_i} \|D^2_t\boldsymbol{w}\|_H\,dt,$$

for all $s \in \mathcal{J}_i$ and for all $i \in \mathbb{N}(1,N)$.

Proof. Using the notation and results of Appendix A we write $\boldsymbol{W} := P_1\boldsymbol{w}|_{\mathcal{J}_i}$, for any $i \in \mathbb{N}(1,N)$, where

$$\boldsymbol{W}(t) = \boldsymbol{W}_0\phi_0(t) + \boldsymbol{W}_1\phi_1(t),$$

$\boldsymbol{W}_0$, $\boldsymbol{W}_1 \in H_i$ and with $\phi_0(t) := 1$ and $\phi_1(t) := (t - t_{i-1})/k_i$ in $\mathcal{J}_i$. Then, from Appendix A we have,

$$\begin{aligned}
A(\boldsymbol{W}_0,\boldsymbol{V}) &= \frac{4}{k_i}\int_{t_{i-1}}^{t_i} A(\boldsymbol{w},\boldsymbol{V})\,dt - \frac{6}{k_i}\int_{t_{i-1}}^{t_i} A(\boldsymbol{w},\boldsymbol{V})\phi_1\,dt,\\
A(\boldsymbol{W}_1,\boldsymbol{V}) &= \frac{12}{k_i}\int_{t_{i-1}}^{t_i} A(\boldsymbol{w},\boldsymbol{V})\phi_1\,dt - \frac{6}{k_i}\int_{t_{i-1}}^{t_i} A(\boldsymbol{w},\boldsymbol{V})\,dt,
\end{aligned}$$

for all $\boldsymbol{V} \in H_i$. Hence,

$$A(\boldsymbol{w}(s) - \boldsymbol{W}(s), \boldsymbol{V}) = A(\boldsymbol{w}(s), \boldsymbol{V}) - A(\boldsymbol{W}_0, \boldsymbol{V}) - A(\boldsymbol{W}_1, \boldsymbol{V})\phi_1(s),$$
$$= A(\boldsymbol{w}(s), \boldsymbol{V}) + \frac{1}{k_i}\int_{t_{i-1}}^{t_i} (6\phi_1(t) + 6\phi_1(s) - 12\phi_1(t)\phi_1(s) - 4)A(\boldsymbol{w}(t), \boldsymbol{V})\,dt.$$

Now we substitute for $\boldsymbol{w}(t)$ in the integrand using

$$\boldsymbol{w}(t) = \boldsymbol{w}(s) + (t-s)\boldsymbol{w}'(s) + \int_s^t (t-\xi)\boldsymbol{w}''(\xi)\,d\xi,$$

and arrive at the following:

$$\frac{1}{k_i}\int_{t_{i-1}}^{t_i} (6\phi_1(t) + 6\phi_1(s) - 12\phi_1(t)\phi_1(s) - 4)A(\boldsymbol{w}(t), \boldsymbol{V})\,dt$$
$$= A(\boldsymbol{w}(s), \boldsymbol{V})\frac{1}{k_i}\int_{t_{i-1}}^{t_i} (6\phi_1(t) + 6\phi_1(s) - 12\phi_1(t)\phi_1(s) - 4)\,dt$$
$$+A(\boldsymbol{w}'(s), \boldsymbol{V})\frac{1}{k_i}\int_{t_{i-1}}^{t_i} (6\phi_1(t) + 6\phi_1(s) - 12\phi_1(t)\phi_1(s) - 4)(t-s)\,dt$$
$$+\frac{1}{k_i}\int_{t_{i-1}}^{t_i} (6\phi_1(t) + 6\phi_1(s) - 12\phi_1(t)\phi_1(s) - 4)\int_s^t (t-\xi)A(\boldsymbol{w}''(\xi), \boldsymbol{V})\,d\xi\,dt.$$

Straightforward (but rather tedious) calculations now show that the first integral on the right evaluates to -1 while the second vanishes, and using these in our previous result gives,

$$A(\boldsymbol{w}(s) - \boldsymbol{W}(s), \boldsymbol{V})$$
$$= \frac{1}{k_i}\int_{t_{i-1}}^{t_i} (6\phi_1(t) + 6\phi_1(s) - 12\phi_1(t)\phi_1(s) - 4)\int_s^t (t-\xi)A(\boldsymbol{w}''(\xi), \boldsymbol{V})\,d\xi\,dt.$$

It is easy to check that

$$|6\phi_1(t) + 6\phi_1(s) - 12\phi_1(t)\phi_1(s) - 4| \leq 4 \qquad \text{for } (t,s) \in \mathcal{J}_i \times \mathcal{J}_i,$$

and from this it follows that

$$|A(\boldsymbol{w}(s) - \boldsymbol{W}(s), \boldsymbol{V})| \leq 4k_i\|\boldsymbol{V}\|_H \int_{t_{i-1}}^{t_i} \|\boldsymbol{w}''\|_H\,d\xi.$$

Finally we choose $\boldsymbol{V} := \boldsymbol{w}(s) - \boldsymbol{W}(s) \in H_i$ for fixed $s \in \mathcal{J}_i$ and this completes the proof. □

Proof of Lemma 12.
To prove (35) we choose $\boldsymbol{w} := \pi_i\boldsymbol{\chi}$ in Lemma 14 and obtain,

$$\|(I - P_0)\pi_i\boldsymbol{\chi}\|_H \leq \int_{t_{i-1}}^{t_i} \|D_t\pi_i\boldsymbol{\chi}\|_H\,dt \leq \Pi_{P_r}\int_{t_{i-1}}^{t_i} \|D_t\boldsymbol{\chi}\|_H\,dt,$$

from which the desired result follows.

To prove (36) we choose $\boldsymbol{w} := R_i\boldsymbol{u}$ in either Lemma 14 or 15 and obtain,

$$\|(I - P_r)R\boldsymbol{u}\|_H \leq 4^r k_i^r \int_{t_{i-1}}^{t_i} \|D_t^{r+1}R\boldsymbol{u}\|_H\,dt \leq 4^r k_i^r \int_{t_{i-1}}^{t_i} \|D_t^{r+1}\boldsymbol{u}\|_H\,dt.$$

From this we get,

$$\|(I-P_r)R\boldsymbol{u}\|^p_{L_p(\mathcal{J}_i;H)} \leq 4^{pr}\int_{t_{i-1}}^{t_i} \|k_i^{r+1}D_t^{r+1}\boldsymbol{u}\|^p_H \, dt,$$

from which the result follows (with obvious modification in the case $p=\infty$).

This completes the proof of Lemma 12.

REFERENCES

[1] D. M. Bedivan and G. J. Fix, *Analysis of finite element approximation and quadrature of Volterra integral equations*, Numer. Methods Partial Differential Equations, 13 (1997), pp. 663—672.

[2] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Introduction to adaptive methods for differential equations*, Acta Numerica, (1995), pp. 105—158.

[3] 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.

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

[5] K. O. Friedrichs, *On the boundary-value problems of the theory of elasticity and Korn's inequality*, Annals of Mathematics, 48 (1947), pp. 267—297.

[6] J. M. Golden and G. A. C. Graham, *Boundary value problems in linear viscoelasticity*, Springer-Verlag, 1988.

[7] C. O. Horgan, *Korn's inequalities and their applications in continuum mechanics*, SIAM Review, 37 (1995), pp. 491—511.

[8] 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.

[9] C. Johnson and P. Hansbo, *Adaptive finite element methods in computational mechanics*, Comput. Methods Appl. Mech. Engrg., 101 (1992), pp. 143—181.

[10] F. J. Lockett, *Nonlinear viscoelastic solids*, Academic Press, 1972.

[11] S. Shaw, M. K. Warby, and J. R. Whiteman, *Error estimates with sharp constants for a fading memory Volterra problem in linear solid viscoelasticity*, SIAM J. Numer. Anal., 34 (1997), pp. 1237—1254. (See also, `http://www.brunel.ac.uk/~icsrbicm`).

[12] S. Shaw, M. K. Warby, J. R. Whiteman, C. Dawson, and M. F. Wheeler, *Numerical techniques for the treatment of quasistatic viscoelastic stress problems in linear isotropic solids*, Comput. Methods Appl. Mech. Engrg., 118 (1994), pp. 211—237.

[13] S. Shaw and J. R. Whiteman, *Backward Euler and Crank-Nicolson finite element variants with rational adaptivity and* a posteriori *error estimates for an integrodifferential equation.* Submitted to Math. Comp. (see `http://www.brunel.ac.uk/~icsrbicm`), 1996.

[14] ———, *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.

[15] ———, *Towards adaptive finite element schemes for partial differential Volterra equation solvers*, Advances in Computational Mathematics, 6 (1996), pp. 309—323.

[16] ———, *Towards robust adaptive finite element methods for partial differential Volterra equation problems arising in viscoelasticity theory*, in The Mathematics of Finite Elements and Applications. mafelap 1996, J. R. Whiteman, ed., Wiley, Chichester, 1997, pp. 55—80.

[17] ———, *Numerical solution of linear quasistatic hereditary viscoelasticity problems* II: *a posteriori estimates.* Submitted to *SIAM J. Numer. Anal*; BICOM Technical Report 98/3, see `http://www.brunel.ac.uk/~icsrbicm`, 1998.

[18] ———, *Numerical solution of linear quasistatic hereditary viscoelasticity problems* III: *a posteriori estimates in a weak norm.* BICOM Technical Report 98/4, see `http://www.brunel.ac.uk/~icsrbicm`, 1998.

[19] ———, *Optimal long-time $L_p(0,T)$ data stability and semidiscrete error estimates for the Volterra formulation of the linear quasistatic viscoelasticity problem.* Submitted to *Numer. Math*; BICOM Tech. Rep. 98/7 see: `http://www.brunel.ac.uk/~icsrbicm.`, 1998.

[20] I. H. Sloan and V. Thomée, *Time discretization of an integro-differential equation of parabolic type*, SIAM J. Numer. Anal., 23 (1986), pp. 1052—1061.