CHALMERS | UNIVERSITY OF GOTHENBURG

PREPRINT 2009:17

A Continuous Space-Time Finite Element Method for an Integro-Differential Equation Modeling Dynamic Fractional Order Viscoelasticity

FARDIN SAEDPANAH

Department of Mathematical Sciences
Division of Mathematics
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Göteborg Sweden 2009

Preprint 2009:17

A Continuous Space-Time Finite Element Method for an Integro-Differential Equation Modeling Dynamic Fractional Order Viscoelasticity

Fardin Saedpanah

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Göteborg, Sweden
Göteborg, April 2009

Preprint 2009:17
ISSN 1652-9715

Matematiska vetenskaper
Göteborg 2009

A CONTINUOUS SPACE-TIME FINITE ELEMENT METHOD FOR AN INTEGRO-DIFFERENTIAL EQUATION MODELING DYNAMIC FRACTIONAL ORDER VISCOELASTICITY

FARDIN SAEDPANAH

ABSTRACT. A fractional order integro-differential equation with a weakly singular kernel is considered. A weak form is formulated, and stability of the primal and the dual problem is studied. The continuous Galerkin method of degree one is formulated, and optimal a priori error estimate is obtained by duality argument. A posteriori error representation based on space-time cells is presented such that it can be used for adaptive strategies based on dual weighted residual methods. Some global a posteriori error estimates are also proved.

1. INTRODUCTION

We study an initial-boundary value problem, modeling dynamic fractional order viscoelasticity of the form (we use '$\cdot$' to denote the time derivative),

$$
\begin{aligned}
&\rho\ddot{u}(x,t) - \nabla\cdot\sigma_0(u;x,t) \\
&\qquad + \int_0^t \beta(t-s)\nabla\cdot\sigma_0(u;x,s)\,ds = f(x,t) && \text{in } \Omega\times(0,T),\\
&u(x,t) = 0 && \text{on } \Gamma_{\mathrm{D}}\times(0,T),\\
&\sigma(u;x,t)\cdot n = g(x,t) && \text{on } \Gamma_{\mathrm{N}}\times(0,T),\\
&u(x,0) = u^0(x), \quad \dot{u}(x,0) = v^0(x) && \text{in } \Omega.
\end{aligned}
\tag{1.1}
$$

Here u is the displacement vector, ρ is the (constant) mass density, f and g represent, respectively, the volume and surface loads. The stress $\sigma = \sigma(u;x,t)$ is determined by

$$\sigma(t) = \sigma_0(t) - \int_0^t \beta(t-s)\sigma_0(s)\,ds,$$

with

$$\sigma_0(t) = 2\mu_0\epsilon(t) + \lambda_0\,\mathrm{Tr}(\epsilon(t))\mathrm{I},$$

where I is the identity operator, ϵ is the strain which is defined by $\epsilon = \frac{1}{2}(\nabla u + (\nabla u)^T)$, and $\mu_0, \lambda_0 > 0$ are elastic constants of Lamé type. The kernel is defined

Date: April 12, 2009.

1991 *Mathematics Subject Classification.* 65M60, 45K05.

Key words and phrases. fractional order viscoelasticity, Galerkin approximation, weakly singular kernel, stability, a priori estimate, a posteriori estimate.

by:

$$\begin{aligned}\beta(t) &= -\gamma \frac{d}{dt} E_\alpha\Big(-\Big(\frac{t}{\tau}\Big)^\alpha\Big) = \gamma \frac{\alpha}{\tau}\Big(\frac{t}{\tau}\Big)^{-1+\alpha} E'_\alpha\Big(-\Big(\frac{t}{\tau}\Big)^\alpha\Big) \\ &\approx Ct^{-1+\alpha}, \quad t \to 0,\end{aligned} \tag{1.2}$$

which means that it is weakly singular with the properties

$$\begin{aligned}&\beta(t) \geq 0, \\ &\|\beta\|_{L_1(\mathbb{R}^+)} = \int_0^\infty \beta(t)\,dt = \gamma\Big(E_\alpha(0) - E_\alpha(\infty)\Big) = \gamma < 1,\end{aligned} \tag{1.3}$$

where $\gamma \in (0,1)$ is a constant and $\tau > 0$ is the relaxation time. Here E_α is the Mittag-Leffler function of order $\alpha \in (0,1)$ and defined by,

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(1+n\alpha)}.$$

There is an extensive literature on theoretical and numerical analysis of integro-differential equations modeling linear and fractional order viscoelasticity, see, e.g., [1], [7], [8], [9], [10], [12], and [13]. Existence, uniqueness and regularity of solution of a problem in the form of (1.1) has been studied in [7], [11]. A posteriori analysis of temporal finite element approximation of a parabolic type problem and discontinuous Galerkin finite element approximation of a quasi-static ($\rho\ddot{u} \approx 0$) linear viscoelasticity problem has been studied, respectively, in [1] and [13]. The present work extends previous works, e.g., [1], [2], and [13].

The outline of this paper is as follows. In §2 we define a weak form of (1.1) and the corresponding dual (adjoint) problem, and we study the stability. In §3 we formulate a continuous Galerkin method of degree one and we obtain stability estimates. Then in §4 we obtain an optimal a priori error estimate. We present an a posteriori error representation based on the dual weighted residual method in §5, and we prove some global a posteriori error estimates.

2. Weak formulation and stability

We let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded polygonal domain with boundary $\Gamma = \Gamma_\mathrm{D} \cup \Gamma_\mathrm{N}$ where Γ_D and Γ_N are disjoint and $\mathrm{meas}(\Gamma_\mathrm{D}) \neq 0$. We introduce the function spaces $H = L_2(\Omega)^d$, $H_{\Gamma_\mathrm{N}} = L_2(\Gamma_\mathrm{N})^d$, and $V = \{v \in H^1(\Omega)^d : v|_{\Gamma_\mathrm{D}} = 0\}$. We denote the norms in H and H_{Γ_N} by $\|\cdot\|$ and $\|\cdot\|_{\Gamma_\mathrm{N}}$, respectively. We also define a bilinear form (with the usual summation convention)

$$a(v,w) = \int_\Omega \big(2\mu_0 \epsilon_{ij}(v)\epsilon_{ij}(w) + \lambda_0 \epsilon_{ii}(v)\epsilon_{jj}(w)\big)\,dx, \quad v, w \in V, \tag{2.1}$$

which is coercive on V, and we equip V with the inner product $a(\cdot,\cdot)$ and norm $\|v\|_V^2 = a(v,v)$. We define $Au = -\nabla \cdot \sigma_0(u)$, which is a selfadjoint, positive definite, unbounded linear operator, with $\mathcal{D}(A) = H^2(\Omega)^d \cap V$, and we use the norms $\|v\|_s = \|A^{s/2}v\|$.

We use a "velocity-displacement" formulation of (1.1) which is obtained by introducing a new velocity variable. Henceforth we use the new variables $u_1 = u$, $u_2 = \dot{u}$ and $u = (u_1, u_2)$ the pair of vector valued functions. Now we define the bilinear and linear forms $\mathcal{A}:\mathcal{V}\times\mathcal{W}\to\mathbb{R}$, $\mathcal{A}^*_\tau:\mathcal{W}^*\times\mathcal{V}^*\to\mathbb{R}$, $F:\mathcal{W}\to\mathbb{R}$, $J_\tau:\mathcal{W}^*\to\mathbb{R}$, for $\tau\in\mathbb{R}^{\geq 0}$, by

$$\begin{aligned}
\mathcal{A}(u,w) &= \int_0^T \Big\{(\dot u_1,w_1)-(u_2,w_1)+\rho(\dot u_2,w_2)+a(u_1,w_2)\\
&\qquad -\int_0^t \beta(t-s)a\big(u_1(s),w_2\big)\,ds\Big\}\,dt\\
&\quad + \big(u_1(0),w_1(0)\big)+\rho\big(u_2(0),w_2(0)\big),\\
\mathcal{A}^*_\tau(w,z) &= \int_\tau^T \Big\{-(w_1,\dot z_1)+a(w_1,z_2)-\int_t^T \beta(s-t)a\big(w_1,z_2(s)\big)\,ds\\
&\qquad -\rho(w_2,\dot z_2)-(w_2,z_1)\Big\}\,dt+\big(w_1(T),z_1(T)\big)+\rho\big(w_2(T),z_2(T)\big),\\
F(w) &= \int_0^T \Big\{(f,w_2)+(g,w_2)_{\Gamma_{\mathrm{N}}}\Big\}\,dt+\big(u^0,w_1(0)\big)+\rho\big(v^0,w_2(0)\big),\\
J_\tau(w) &= \int_\tau^T \Big\{(w_1,j_1)+(w_2,j_2)\Big\}\,dt+\big(w_1(T),z_1^T\big)+\rho(w_2(T),z_2^T),
\end{aligned}$$

where j_1, j_2 and z_1^T, z_2^T represent, respectively, the load terms and the initial data of the dual (adjoint) problem. In case of $\tau=0$, we use the notation $\mathcal{A}^*$, J for short. Here

$$\begin{aligned}
\mathcal{V} &= H^1\big((0,T);V\big)\times H^1\big((0,T);H\big),\\
\mathcal{V}^* &= H^1\big((0,T);H\big)\times H^1\big((0,T);V\big),\\
\mathcal{W} &= \big\{w=(w_1,w_2): w\in L_2\big((0,T);H\big)\times L_2\big((0,T);V\big),\\
&\qquad\qquad w_i \text{ are right continuous in time}\big\},\\
\mathcal{W}^* &= \big\{w=(w_1,w_2): w\in L_2\big((0,T);V\big)\times L_2\big((0,T);H\big),\\
&\qquad\qquad w_i \text{ are left continuous in time}\big\},
\end{aligned} \tag{2.2}$$

and we note that $\mathcal{V}\subset\mathcal{W}^*$, $\mathcal{V}^*\subset\mathcal{W}$.

The variational formulation analogue to (1.1) is then to find $u\in\mathcal{V}$ such that,

$$\mathcal{A}(u,w)=F(w),\quad \forall w\in\mathcal{W}. \tag{2.3}$$

Here the definition of the velocity $u_2=\dot u_1$ is enforced in the L_2 sense, and the initial data are placed in the bilinear form in a weak sense. A variant is used in [7] where the velocity has been enforced in the H^1 sense, without placing the initial data in the bilinear form. We also note that the initial data are retained by the choice of the function space $\mathcal{W}$, that consists of right continuous functions with respect to time.

To obtain the dual (adjoint) problem we note that $\mathcal{A}^*$ is the adjoint form of $\mathcal{A}$. Indeed, integrating by parts with respect to time in $\mathcal{A}$, then changing the order of integrals in the convolution term as well as changing the role of the variables s,t, we have,

$$\mathcal{A}(v,w)=\mathcal{A}^*(v,w),\quad \forall v\in\mathcal{V},\ w\in\mathcal{V}^*. \tag{2.4}$$

Then the variational formulation of the dual problem is to find $z\in\mathcal{V}^*$ such that,

$$\mathcal{A}^*(w,z)=J(w),\quad \forall w\in\mathcal{W}^*, \tag{2.5}$$

that is a weak formulation of

$$\rho \ddot{z}_2 + Az_2 - \int_t^T \beta(s-t) Az_2(s)\,ds = j_1 - \frac{\partial}{\partial t} j_2,$$

with initial data z_1^T, z_2^T, and function $j = (j_1, j_2)$ that is defined by $J(w) = \int_0^T (w, j)\,dt$.

In the analysis below we use a positive type kernel ξ. Indeed, we recall (1.2), (1.3) and we define the function

$$\xi(t) = \gamma - \int_0^t \beta(s)\,ds = \int_t^\infty \beta(s)\,ds = \gamma E_\alpha(t), \tag{2.6}$$

and it is easy to see that

$$D_t \xi(t) = -\beta(t) < 0, \quad \xi(0) = \gamma, \quad \lim_{t\to\infty} \xi(t) = 0, \quad 0 < \xi(t) \leq \gamma. \tag{2.7}$$

Besides, ξ is a completely monotone function, that is,

$$(-1)^j D_t^j \xi(t) \geq 0, \quad t \in (0, \infty),\, j \in \mathbb{N},$$

since the Mittag-Leffler function E_α, $\alpha \in [0, 1]$ is completely monotone, see, e.g., [5]. Consequently, an important property of ξ is that, it is a positive type kernel, that is, it is continuous and, for any $T \geq 0$, satisfies

$$\int_0^T \int_0^t \xi(t-s)\phi(t)\phi(s)\,ds\,dt \geq 0, \quad \forall \phi \in \mathcal{C}([0, T]). \tag{2.8}$$

Theorem 1. *Let u be the solution of* (2.3) *with sufficiently smooth data u^0, v^0, f, g. Then for $l \in \mathbb{R}, T > 0$, we have the identity,*

$$\begin{aligned}
&\rho \|u_2(T)\|_l^2 + \tilde{\gamma} \|u_1(T)\|_{l+1}^2 + 2\int_0^T \int_0^t \xi(t-s) a\big(\dot{u}(s), A^l \dot{u}(t)\big)\,ds\,dt \\
&\quad = \rho \|v^0\|_l^2 + (1+\gamma)\|u^0\|_{l+1}^2 + 2\int_0^T \big(f(t), A^l \dot{u}_1(t)\big)\,dt \\
&\quad + 2\int_0^T \big(g(t), A^l \dot{u}_1(t)\big)_{\Gamma_N}\,dt \\
&\quad - 2\int_0^T \beta(t) a\big(u^0, A^l u_1(t)\big)\,dt - 2\xi(T) a\big(u^0, A^l u_1(T)\big),
\end{aligned} \tag{2.9}$$

where $\tilde{\gamma} = 1 - \gamma$. Moreover, with $\Gamma_N = \varnothing$ or $\Gamma_N \neq \varnothing$, $g = 0$, we have,

$$\|u_2(T)\|_l + \|u_1(T)\|_{l+1} \leq C\Big\{ \|v^0\|_l + \|u^0\|_{l+1} + \int_0^T \|f(t)\|_l\,dt \Big\}, \tag{2.10}$$

for some $C = C(\rho, \gamma, T)$. And with $\Gamma_N \neq \varnothing$, $g \neq 0$, $l = 0$, we have the estimate,

$$\begin{aligned}
&\|u_2(T)\| + \|u_1(T)\|_1 \\
&\qquad \leq C\big\{ \|v^0\| + \|u^0\|_1 + \|f\|_{L_1((0,t);H)} + \|g\|_{W_1^1((0,t);H^{\Gamma_N})} \big\},
\end{aligned} \tag{2.11}$$

for some $C = C(\Omega, \rho, \gamma, T)$.

Proof. Since u is a solution of (2.3), we obviously have $u_2 = \dot u_1$. We recall $\tilde\gamma = 1-\gamma$, and from (2.7) we have that $\beta(t-s) = D_s\xi(t-s)$, $\xi(0)=\gamma$. These and partial integration in the convolution term in $\mathcal{A}$, yield

$$\begin{aligned}\mathcal{A}(u,w) = \int_0^T \Big\{\rho(\dot u_2, w_2) + \tilde\gamma a(u_1, w_2) + \int_0^t \xi(t-s) a\big(\dot u_1(s), w_2\big)\, ds \\ + \xi(t) a(u^0, w_2)\Big\}\, dt + \big(u_1(0), w_1(0)\big) + \rho\big(u_2(0), w_2(0)\big).\end{aligned} \tag{2.12}$$

Putting this with $w=(w_1,w_2) = (A^{l+1}u_1, A^l u_2) = (A^{l+1}u_1, A^l \dot u_1)$ in (2.3) we obtain,

$$\begin{aligned}&\rho\|u_2(T)\|_l^2 + \tilde\gamma\|u_1(T)\|_{l+1}^2 + 2\int_0^T\int_0^t \xi(t-s) a\big(\dot u_1(s), A^l\dot u_1(t)\big)\, ds\, dt \\ &\quad = \rho\|v^0\|_l^2 + \tilde\gamma\|u^0\|_{l+1}^2 + 2\int_0^T (f, A^l\dot u_1)\, dt + 2\int_0^T (g, A^l \dot u_1)_{\Gamma_{\mathrm{N}}}\, dt \\ &\quad - 2\int_0^T \xi(t) a\big(u^0, A^l\dot u_1(t)\big)\, dt. \\ &\quad = \rho\|v^0\|_l^2 + \tilde\gamma\|u^0\|_{l+1}^2 + 2\int_0^T (f, A^l\dot u_1)\, dt + 2\int_0^T (g, A^l \dot u_1)_{\Gamma_{\mathrm{N}}}\, dt \\ &\quad - 2\int_0^T \beta(t) a\big(u^0, A^l u_1(t)\big)\, dt - 2\xi(T) a\big(u^0, A^l u_1(T)\big) + 2\gamma\|u^0\|_{l+1}^2,\end{aligned}$$

where for the last equality we used partial integration and $D_t\xi(t) = -\beta(t)$. This, having $\tilde\gamma = 1-\gamma$, implies the identity (2.9).

Now we prove the stimates (2.10) and (2.11). First, in (2.9), we use (2.8), integration by parts in the term with the surface load g, and the Cauchy-Schwarz inequality and conclude,

$$\begin{aligned}&\rho\|u_2(T)\|_l^2 + \tilde\gamma\|u_1(T)\|_{l+1}^2 \\ &\quad \le \rho\|v^0\|_l^2 + (1+\gamma)\|u^0\|_{l+1}^2 \\ &\quad + 2/C_1 \max_{0\le t\le T}\|\dot u_1(t)\|_l^2 + 2C_1\Big(\int_0^T \|f(t)\|_l\, dt\Big)^2 \\ &\quad - 2\int_0^T \big(\dot g(t), A^l u_1(t)\big)_{\Gamma_{\mathrm{N}}}\, dt + 2\big(g(T), A^l u_1(T)\big)_{\Gamma_{\mathrm{N}}} - 2\big(g(0), A^l u^0\big)_{\Gamma_{\mathrm{N}}} \\ &\quad + 2\|u^0\|_{l+1}\max_{0\le t\le T}\|u_1(t)\|_{l+1}\int_0^T \beta(t)\, dt \\ &\quad + 2\gamma/C_2\|u^0\|_{l+1}^2 + 2\gamma C_2\|u_1(T)\|_{l+1}^2.\end{aligned}$$

This with $\Gamma_{\mathrm{N}} = \varnothing$ or $\Gamma_{\mathrm{N}} \ne \varnothing$, $g=0$, considering $\|\beta\|_{L_1(\mathbb{R}^+)} = \gamma$, implies (2.10). But for the case that the surface load $g\ne 0$ ($\Gamma_{\mathrm{N}}\ne\varnothing$), we need to restrict to $l=0$, due to the trace theorem. That is, in this case with $l=0$ and using $\|u_1(t)\|_{\Gamma_{\mathrm{N}}} \le C(\Omega)\|u_1(t)\|_1$ by the trace theorem, in a standard way, we have the estimate

$$\begin{aligned}\|u_2(T)\| + \|u_1(T)\|_1 \le C\big\{\|v^0\| + \|u^0\|_1 + \|f\|_{L_1((0,t);H)} \\ + \|\dot g\|_{L_1((0,T);H^{\Gamma_{\mathrm{N}}})} + \|g\|_{L_\infty((0,T);H^{\Gamma_{\mathrm{N}}})}\big\}.\end{aligned}$$

This and the fact that $\|g\|_{L_\infty((0,T);H^{\Gamma_N})} \leq C\|g\|_{W_1^1((0,T);H^{\Gamma_N})}$, by Sobolev inequality, imply (2.11). □

Remark 1. An identity, slightly different from (2.9) has been presented in [2], by using a function $w(t,s) = u(t) - u(t-s)$ that must belong to a weighted L_2-space, introduced in [6], see also [7] where the abstract framework by [6] has been, specifically, applied to this problem. The proof presented here avoids using function w and seems to be simple and straightforward.

Remark 2. An important property, used here is that the kernel is of positive type. This means that the technique presented here can be applied to the problems with positive type kernels. For example, with $g = 0$, we could consider problems with positive type kernels and of the form

$$\rho\ddot{u} + Au - \int_0^t \beta(t-s)Bu(s)\,ds = f,$$

where B is a selfadjoint, positive definite linear operator such that, for some suitable constant C,

$$(Bv, w) \leq C(Av, w), \quad \forall v, w \in \mathcal{D}(A).$$

For example, with kernel β defined in (1.2), we should have $1 - \gamma C > 0$. Then a similar argument can be applied with $l = 0$, and also with $l \neq 0$ provided B and A be comutative.

Theorem 2. *Let z be the solution of the dual problem* (2.5) *with sufficiently smooth data z_1^T, z_2^T, j_1, j_2. Then for $l \in \mathbb{R}$, $0 < t < T$, we have the identity,*

$$\begin{aligned}
&\|z_1(t)\|_l^2 + \rho\tilde{\gamma}\|z_2(t)\|_{l+1}^2 + 2\rho\int_t^T\int_r^T \xi(s-r)a(A^l\dot{z}_2(r), \dot{z}_2(s))\,ds\,dr \\
&\quad = \|z_1^T\|_l^2 + \rho(1+\gamma)\|z_2^T\|_{l+1}^2 + 2\int_t^T \Big\{(A^l z_1, j_1) + (A^{l+1}z_2, j_2)\Big\}\,dr \\
&\quad\quad - 2\rho\int_t^T \beta(T-r)a(A^l z_2(r), z_2^T)\,dr \\
&\quad\quad - 2\rho\xi(T-t)a(A^l z_2(t), z_2^T).
\end{aligned} \tag{2.13}$$

where $\tilde{\gamma} = 1 - \gamma$. Moreover, for some constant $C = C(\rho, \gamma, T)$, we have stability estimates

$$\|z_1(t)\|_l + \|z_2(t)\|_{l+1} \leq C\Big\{\|z_1^T\|_l + \|z_2^T\|_{l+1} + \int_t^T \Big(\|j_1\|_l + \|j_2\|_{l+1}\Big)\,dr\Big\}. \tag{2.14}$$

Proof. Since z is a solution of (2.5), we obviously have $z_1 = -\rho\dot{z}_2$ and, for $t \in [0, T)$, z satisfies

$$\mathcal{A}_t^*(w, z) = J_t(w), \quad \forall v \in \mathcal{W}^*. \tag{2.15}$$

We recall $\tilde{\gamma} = 1 - \gamma$, and from (2.7) we have $\beta(s-t) = -D_s\xi(s-t)$, $\xi(0) = \gamma$. Then, by partial integration with respect to time in the convolution term in $\mathcal{A}_t^*$, we obtain

$$\begin{aligned}
\mathcal{A}_t^*(w, z) = \int_t^T \Big\{ &-(w_1, \dot{z}_1) + \tilde{\gamma}a(w_1, z_2) - \int_r^T \xi(s-r)a\big(w_1, \dot{z}_2(s)\big)\,ds \\
&+ \xi(T-r)a\big(w_1, z_2(T)\big)\Big\}\,dr \\
&+ \big(w_1(T), z_1(T)\big) + \rho\big(w_2(T), z_2(T)\big).
\end{aligned} \tag{2.16}$$

Putting this with $w = (w_1, w_2) = (A^l z_1, A^{l+1} z_2) = (-\rho A^l \dot{z}_2, A^{l+1} z_2)$ in (2.15) we have

$$\int_t^T \Big\{ -\frac{1}{2} D_t \|z_1(r)\|_l^2 - \frac{1}{2}\rho\tilde{\gamma} D_t \|z_2(r)\|_{l+1}^2 + \rho \int_r^T \xi(s-r) a\big(A^l \dot{z}_2(r), \dot{z}_2(s)\big)\, ds$$
$$- \rho \xi(T-r) a\big(A^l \dot{z}_2(r), z_2^T\big) \Big\}\, dr = \int_t^T \Big\{ (A^l z_1, j_1) + (A^{l+1} z_2, j_2) \Big\}\, dr,$$

that implies

$$\|z_1(t)\|_l^2 + \rho\tilde{\gamma} \|z_2(t)\|_{l+1}^2 + 2\rho \int_t^T \int_r^T \xi(s-r) a\big(A^l \dot{z}_2(r), \dot{z}_2(s)\big)\, ds\, dr$$
$$= \|z_1^T\|_l^2 + \rho\tilde{\gamma} \|z_2^T\|_{l+1}^2 + 2 \int_t^T \Big\{ (A^l z_1, j_1) + (A^{l+1} z_2, j_2) \Big\}\, dr$$
$$+ 2\rho \int_t^T \xi(T-r) a\big(A^l \dot{z}_2(r), z_2^T\big)\, dr.$$

Now integration by parts in the last term, having $D_r \xi(T-r) = \beta(T-r)$ by (2.7), gives the identity (2.13). Then, using (2.8) in (2.13), similar to the proof of (2.10), in a standard way, we conclude the inequality (2.14), and this completes the proof. □

3. The continuous Galerkin method

Let $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n < \cdots < t_N = T$ be a partition of the time interval $[0, T]$. To each discrete time level t_n we associate a triangulation $\mathcal{T}_h^n$ of the polygonal domain Ω with the mesh function,

$$h_n(x) = h_K = \mathrm{diam}(K), \quad x \in K,\ K \in \mathcal{T}_h^n, \tag{3.1}$$

and a finite element space V_h^n consisting of continuous piecewise linear polynomials. For each time subinterval $I_n = (t_{n-1}, t_n)$ of length $k_n = t_n - t_{n-1}$, we define intermediate triangulaion $\bar{\mathcal{T}}_h^n$ which is composed of mutually finest meshes of the neighboring meshes $\mathcal{T}_h^n$, $\mathcal{T}_h^{n-1}$ defined at discrete time levels t_n, t_{n-1}, respectively. The mesh function $\bar{h}_n$ is then defined by

$$\bar{h}_n(x) = \bar{h}_K = \mathrm{diam}(K), \quad x \in K,\ K \in \bar{\mathcal{T}}_h^n. \tag{3.2}$$

Correspondingly, we define the finite element spaces $\bar{V}_h^n$ consisting of continuous piecewise linear polynomials. This construction is used in order to allow continuity in time of the trial functions when the meshes change with time. Hence we obtain a decomposition of each time slab $\Omega^n = \Omega \times I_n$ into space-time cells $K^n = K \times I_n$, $K \in \bar{\mathcal{T}}_h^n$ (prisms, for example, in case of $\Omega \subset \mathbb{R}^2$). We note the difference between the mesh functions h_n and $\bar{h}_n$, and this is important in our a posteriori error analysis.

The trial and test function spaces for the discrete form are, respectively:

$$\begin{aligned}\mathcal{V}_{hk} = \Big\{ U = (U_1, U_2) : &\ U \text{ continuous in } \Omega \times [0, T],\\ &\ U(x,t)|_{I_n} \text{ linear in } t,\\ &\ U(\cdot, t_n) \in (V_h^n)^2, U(\cdot,t)|_{I_n} \in (\bar{V}_h^n)^2 \Big\},\\ \mathcal{W}_{hk} = \Big\{ V = (V_1, V_2) : &\ V(\cdot, t) \text{ continuous in } \Omega,\\ &\ V(\cdot,t)|_{I_n} \in (V_h^n)^2,\\ &\ V(x,t)|_{I_n} \text{ piecewise constant in } t \Big\}.\end{aligned} \tag{3.3}$$

We note that global continuity of the trail functions in $\mathcal{V}_{hk}$ requires the use of '*hanging nodes*' if the spatial mesh changes across a time level t_n. We allow one hanging node per edge or face.

Remark 3. If we do not change the spatial mesh or just refine the spatial mesh from one time level to the next one, i.e.,

$$V_h^{n-1} \subset V_h^n, \quad n = 1, \ldots, N, \tag{3.4}$$

then we have $\bar{V}_h^n = V_h^n$.

In the construction of $\mathcal{V}_{hk}$ we have associated the triangulation $\mathcal{T}_h^n$ with discrete time levels instead of the time slabs Ω^n, and in the interior of time slabs we let U be from the union of the finite element spaces defined on the triangulations at the two adjacent time levels. This construction is necessary to allow for trial functions that are continuous also at the discrete time leveles even if grids change between time steps. Associating triangulation with time slabs instead of time levels would yield a variant scheme which includes jump terms due to discontinuity at discrete time leveles, when coarsening happens. This means that there are extra degrees of freedom that one might use suitable projections for transfering solution at the time levels t_n, see [7].

The continuous Galerkin method, based on the variational fomulation (2.3), is to find $U \in \mathcal{V}_{hk}$ such that,

$$\mathcal{A}(U, W) = F(W), \quad \forall\, W \in \mathcal{W}_{hk}. \tag{3.5}$$

The Galerkin orthogonality, with $u = (u_1, u_2)$ being the exact solution of (2.3), is then,

$$\mathcal{A}(U - u, W) = 0, \quad \forall\, W \in \mathcal{W}_{hk}. \tag{3.6}$$

Similarly the continuous Galerkin method, based on the dual variational formulation (2.5), is to find $Z \in \mathcal{V}_{hk}$ such that,

$$\mathcal{A}^*(W, Z) = J(W), \quad \forall\, W \in \mathcal{W}_{hk}. \tag{3.7}$$

Then, Z also satisfies, for $n = 0, 1, \ldots, N-1$,

$$\mathcal{A}^*_{t_n}(W, Z) = J_{t_n}(W), \quad \forall\, W \in \mathcal{W}_{hk}. \tag{3.8}$$

We notice that, rather than using the dual formulation of the discrete problem (3.5), we formulated the same finite element method for the continuous dual problem (2.5).

From (3.5) we can recover the time stepping scheme,

$$
\begin{aligned}
&\int_{I_n} \Big\{ (\dot U_1, W_1) - (U_2, W_1) \Big\}\, dt = 0,\\
&\int_{I_n} \Big\{ \rho(\dot U_2, W_2) + a(U_1, W_2) - \int_0^t \beta(t-s) a\big(U_1(s), W_2(t)\big)\, ds \Big\}\, dt \\
&\qquad\qquad = \int_{I_n} \big\{ (f, W_2)\, dt + (g, W_2)_{\Gamma_N} \big\}\, dt, \quad \forall\, W_1, W_2 \in \mathcal{W}_{hk},\\
&U_1(0) = u_h^0, \quad U_2(0) = v_h^0,
\end{aligned}
\tag{3.9}
$$

for suitable choice of $u_h^0, v_h^0 \in V_h^0$ as approximations of the initial data u^0, v^0. Here, as a natural choice, we have

$$
u_h^0 = \mathcal{P}_h u^0, \quad v_h^0 = \mathcal{P}_h v^0. \tag{3.10}
$$

Typical functions $U = (U_1, U_2) \in \mathcal{V}_{hk}$, $W = (W_1, W_2) \in \mathcal{W}_{hk}$ are as follows:

$$
\begin{aligned}
U_i(x, t_n) &= U_i^n(x) = \sum_{j=1}^{m_n} U_{i,j}^n \varphi_j^n(x),\\
U_i(x,t)|_{I_n} &= \psi_{n-1}(t) U_i^{n-1}(x) + \psi_n(t) U_i^n(x),\\
W_i(x,t)|_{I_n} &= \sum_{j=1}^{m_n} W_{i,j}^n \varphi_j^n(x),
\end{aligned}
\tag{3.11}
$$

where m_n is the number of degrees of freedom in $\mathcal{T}_h^n$, $\{\varphi_j^n(x)\}_{j=1}^{m_n}$ are the nodal basis functions for V_h^n defined on triangulation $\mathcal{T}_h^n$, and $\psi_n(t)$ is the nodal basis function defined at time level t_n. Hence (3.9) yields

$$
\begin{aligned}
&(U_1^n - U_1^{n-1}, W_1) - \frac{k_n}{2}\big(U_2^n + U_2^{n-1}, W_1\big) = 0,\\
&\rho(U_2^n - U_2^{n-1}, W_2) + \frac{k_n}{2}(U_1^n + U_1^{n-1}, W_2)\\
&\qquad - \sum_{l=1}^{n} a(U_1^{l-1}, W_2) \int_{I_n} \int_{t_{l-1}}^{t \wedge t_l} \beta(t-s) \psi_{l-1}(s)\, ds\, dt\\
&\qquad - \sum_{l=1}^{n} a(U_1^{l}, W_2) \int_{I_n} \int_{t_{l-1}}^{t \wedge t_l} \beta(t-s) \psi_{l}(s)\, ds\, dt\\
&\qquad\qquad = \int_{I_n} \big\{ (f, W_2)\, dt + (g, W_2)_{\Gamma_N} \big\}\, dt, \quad \forall\, W_1, W_2 \in V_h^n,\\
&U_1^0 = u_h^0, \quad U_2^0 = v_h^0.
\end{aligned}
$$

This implies the discrete linear system,

$$
\begin{aligned}
&M^n \tilde U_1^n - \frac{k_n}{2} M^n \tilde U_2^n = M^{n-1,n} \tilde U_1^{n-1} + \frac{k_n}{2} M^{n-1,n} \tilde U_2^{n-1},\\
&\rho M^n \tilde U_2^n + (\frac{k_n}{2} - \omega_{n,n}^-) S^n \tilde U_1^n = \rho M^{n-1,n} \tilde U_2^{n-1} + (-\frac{k_n}{2} + \omega_{n,n-1}) S^{n-1,n} \tilde U_1^{n-1}\\
&\qquad\qquad + S^{0,n} \tilde U_1^0 \omega_{n,0}^+ + \sum_{l=1}^{n-2} \omega_{n,l} S^{l,n} \tilde U_1^l + B^n,\\
&\tilde U_1^0 = u_h^0, \quad \tilde U_2^0 = v_h^0,
\end{aligned}
$$

where

$$
\begin{aligned}
&\omega_{n,0}^{+}=\int_{I_n}\int_0^{t\wedge t_1}\beta(t-s)\psi_l(s)\,ds\,dt,\quad \omega_{n,n}^{-}=\int_{I_n}\int_{t_{n-1}}^{t}\beta(t-s)\psi_l(s)\,ds\,dt,\\
&\omega_{n,l}=\int_{I_n}\int_{t_{l-1}}^{t\wedge t_{l+1}}\beta(t-s)\psi_l(s)\,ds\,dt,\\
&B^n=(B_j^n)_j=\Big(\int_{I_n}\big\{(f,\varphi_j)+(g,\varphi_j)_{\Gamma_N}\big\}\,dt\Big)_j,\\
&M^n=(M_{ij}^n)_{ij}=\big((\varphi_i^n,\varphi_j^n)\big)_{ij},\quad M^{n-1,n}=(M_{ij}^{n-1,n})_{ij}=\big((\varphi_i^{n-1},\varphi_j^n)\big)_{ij},\\
&S^{l,n}=(S_{ij}^{l,n})_{ij}=\big(a(\varphi_i^l,\varphi_j^n)\big)_{ij},
\end{aligned}
$$

and $\tilde{U}_i^n=(U_{i,j}^n)_{j=1}^{m_n}$ with $U_{i,j}^n$ introduced in (3.11).

We define the orthogonal projections $\mathcal{R}_{h,n}:V\to V_h^n$, $\mathcal{P}_{h,n}:H\to V_h^n$ and $\mathcal{P}_{k,n}:L_2(I_n)^d\to\mathbb{P}_0^d(I_n)$, respectively, by

$$
\begin{aligned}
a(\mathcal{R}_{h,n}v-v,\chi)&=0,\quad \forall v\in V,\,\chi\in V_h^n,\\
(\mathcal{P}_{h,n}v-v,\chi)&=0,\quad \forall v\in H,\,\chi\in V_h^n,\\
\int_{I_n}(\mathcal{P}_{k,n}v-v)\cdot\psi\,dt&=0,\quad \forall v\in L_2(I_n)^d,\,\psi\in\mathbb{P}_0^d(I_n)\,,
\end{aligned}
\tag{3.12}
$$

with $\mathbb{P}_0^d$ denoting the set of all vector-valued constant polynomials. Correspondingly, we define $\mathcal{R}_h v$, $\mathcal{P}_h v$ and $\mathcal{P}_k v$ for $t\in I_n\,(n=1,\cdots,N)$, by $(\mathcal{R}_h v)(t)=\mathcal{R}_{h,n}v(t)$, $(\mathcal{P}_h v)(t)=\mathcal{P}_{h,n}v(t)$, and $\mathcal{P}_k v=\mathcal{P}_{k,n}(v|_{I_n})$.

Remark 4. In the case of assumption (3.4), by Remark 3 and the definition of the L_2-projection $\mathcal{P}_k$, we have $\dot{V}$, $\mathcal{P}_k V\in\mathcal{W}_{hk}$, for any $V\in\mathcal{V}_{hk}$.

We introduce the linear operator $A_{n,r}:V_h^r\to V_h^n$ by

$$a(v_r,w_n)=(A_{n,r}v_r,w_n),\quad \forall v_r\in V_h^r,\,w_n\in V_h^n.$$

We set $A_n=A_{n,n}$, with discrete norms

$$\|v_n\|_{h,l}=\|A_n^{l/2}v_n\|=\sqrt{(v_n,A_n^l v_n)},\quad v_n\in V_h^n \text{ and } l\in\mathbb{R}\,,$$

and A_h so that $A_h v=A_n v$ for $v\in V_h^n$. We use $\bar{A}_h$ when it acts on $\bar{V}_h^n$. For later use in our error analysis we note that $\mathcal{P}_h A=A_h\mathcal{R}_h$.

Theorem 3. *Let Z be the solution of* (3.7) *with sufficiently smooth data z_1^T, z_2^T, j_1, j_2. Further, we assume* (3.4). *Then for $l \in \mathbb{R}$, we have the identity,*

$$
\begin{aligned}
&\|Z_1(t_n)\|_{h,l}^2 + \rho\tilde{\gamma}\|Z_2(t_n)\|_{h,l+1}^2 + 2\rho\int_{t_n}^{T}\int_{t}^{T}\xi(s-t)a\big(A_h^l\dot{Z}_2(t),\dot{Z}_2(s)\big)\,ds\,dt\\
&\quad= \|Z_1(T)\|_{h,l}^2 + \rho(1+\gamma)\|Z_2(T)\|_{h,l+1}^2\\
&\qquad+ 2\int_{t_n}^{T}(A_h^l Z_1,\mathcal{P}_k\mathcal{P}_h j_1)\,dt + 2\int_{t_n}^{T}(A_h^{l+1}Z_2,\mathcal{P}_k\mathcal{P}_h j_2)\,dt\\
&\qquad- 2\int_{t_n}^{T}\int_{t}^{T}\beta(s-t)a\big(A_h^l\mathcal{P}_k\mathcal{P}_h j_2, Z_2(s)\big)\,ds\,dt\\
&\qquad- 2\rho\int_{t_n}^{T}\beta(T-t)a\big(A_h^l Z_2(t), Z_2(T)\big)\,dt\\
&\qquad- 2\rho\xi(T-t_n)a\big(A_h^l Z_2(t_n), Z_2(T)\big).
\end{aligned}
\tag{3.13}
$$

where $\tilde{\gamma} = 1-\gamma$. Moreover, for some constant $C = C(\rho,\gamma,T)$, we have stability estimate

$$
\begin{aligned}
\|Z_1(t_n)\|_{h,l} + \|Z_2(t_n)\|_{h,l+1} \le C\Big\{&\|\mathcal{P}_h z_1^T\|_{h,l} + \|\mathcal{P}_h z_2^T\|_{h,l+1}\\
&+ \int_{t_n}^{T}\Big(\|\mathcal{P}_h j_1\|_{h,l} + \|\mathcal{P}_h j_2\|_{h,l+1}\Big)\,dt\Big\}.
\end{aligned}
\tag{3.14}
$$

Proof. The solution Z of (3.7) also satisfies (3.8), for $n = N-1,\ldots,1,0$. Then recalling Remark 4 for the assumption (3.4), we obviously have,

$$
\mathcal{P}_k Z_1 = -\rho\dot{Z}_2 - \mathcal{P}_k\mathcal{P}_h j_2. \tag{3.15}
$$

Using this in (3.8) and recalling the initial data $Z_i(T) = \mathcal{P}_h z_i^T,\ i = 1,2$, we obtain

$$
\begin{aligned}
\int_{t_n}^{T}\Big\{-(W_1,\dot{Z}_1) + a(W_1,Z_2) - \int_{t}^{T}\beta(s-t)a\big(W_1,Z_2(s)\big)\,ds\Big\}\,dt&\\
+ \big(W_1(T),\mathcal{P}_h z_1^T\big) + \rho\big(W_2(T),\mathcal{P}_h z_2^T\big)&\\
= \int_{t_n}^{T}(W_1,j_1)\,dt + \big(W_1(T),z_1^T\big) + \rho\big(W_2(T),z_2^T\big)&.
\end{aligned}
$$

The terms concerning the initial data are canceled by the definition of the orthogonal projection $\mathcal{P}_h$. Besides, for the convolution term we recall $\beta(s-t) = -D_s\xi(s-t)$ from (2.7) and then partial integration yields,

$$
\begin{aligned}
-\int_{t_n}^{T}\int_{t}^{T}\beta(s-t)a\big(W_1,Z_2(s)\big)\,ds\,dt = &-\int_{t_n}^{T}\int_{t}^{T}\xi(s-t)a\big(W_1,\dot{Z}_2(s)\big)\,ds\,dt\\
&+\int_{t_n}^{T}\xi(T-t)a\big(W_1,Z_2(T)\big)\,dt\\
&-\gamma\int_{t_n}^{T}a\big(W_1,Z_2(t)\big)\,dt.
\end{aligned}
$$

These and $\tilde{\gamma} = 1 - \gamma$ imply that the solution Z satisfies,

$$\int_{t_n}^{T} \Big\{ -(W_1, \dot{Z}_1) + \tilde{\gamma} a(W_1, Z_2) - \int_t^T \xi(s-t) a\big(W_1, \dot{Z}_2(s)\big)\, ds + \xi(T-t) a(W_1, Z_2(T)) \Big\}\, dt = \int_{t_n}^{T} (W_1, \mathcal{P}_h j_1)\, dt.$$

Now we set $W_1 = A_h^l \mathcal{P}_k Z_1$, and we have

$$\begin{aligned} &\int_{t_n}^{T} \Big\{ -(A_h^l \mathcal{P}_k Z_1, \dot{Z}_1) + \tilde{\gamma} a(A_h^l \mathcal{P}_k Z_1, Z_2) \\ &\quad - \int_t^T \xi(s-t) a\big(A_h^l \mathcal{P}_k Z_1, \dot{Z}_2(s)\big)\, ds \\ &\quad + \xi(T-t) a\big(A_h^l \mathcal{P}_k Z_1, Z_2(T)\big) \Big\}\, dt = \int_{t_n}^{T} (A_h^l \mathcal{P}_k Z_1, \mathcal{P}_h j_1)\, dt. \end{aligned} \tag{3.16}$$

We study the four terms at the left side of the above equation. For the first term we have

$$\begin{aligned} \int_{t_n}^{T} -(A_h^l \mathcal{P}_k Z_1, \dot{Z}_1)\, dt &= -\frac{1}{2} \int_{t_n}^{T} D_t \|Z_1(t)\|_{h,l}^2\, dt \\ &= -\frac{1}{2} \|Z_1(T)\|_{h,l}^2 + \frac{1}{2} \|Z_1(t_n)\|_{h,l}^2. \end{aligned} \tag{3.17}$$

With (3.15) we can write the second term as

$$\begin{aligned} \tilde{\gamma} \int_{t_n}^{T} a(A_h^l \mathcal{P}_k Z_1, Z_2)\, dt &= -\rho \tilde{\gamma} \int_{t_n}^{T} a(A_h^l \dot{Z}_2, Z_2)\, dt \\ &\quad - \tilde{\gamma} \int_{t_n}^{T} a(A_h^l \mathcal{P}_k \mathcal{P}_h j_2, Z_2)\, dt \\ &= -\frac{\rho \tilde{\gamma}}{2} \|Z_2(T)\|_{h,l+1}^2 + \frac{\rho \tilde{\gamma}}{2} \|Z_2(t_n)\|_{h,l+1}^2 \\ &\quad - \tilde{\gamma} \int_{t_n}^{T} a(A_h^l \mathcal{P}_k \mathcal{P}_h j_2, Z_2)\, dt. \end{aligned} \tag{3.18}$$

For the third term in (3.16), by virtue of (3.15) and integration by parts, we obtain

$$
\begin{aligned}
-\int_{t_n}^{T}\int_{t}^{T}\xi(s-t)a\big(A_h^l\mathcal{P}_kZ_1,\dot Z_2(s)\big)\,ds\,dt \\
&= \rho\int_{t_n}^{T}\int_{t}^{T}\xi(s-t)a\big(A_h^l\dot Z_2(t),\dot Z_2(s)\big)\,ds\,dt\\
&\quad+\int_{t_n}^{T}\int_{t}^{T}\xi(s-t)a\big(A_h^l\mathcal{P}_k\mathcal{P}_hj_2,\dot Z_2(s)\big)\,ds\,dt\\
&= \rho\int_{t_n}^{T}\int_{t}^{T}\xi(s-t)a\big(A_h^l\dot Z_2(t),\dot Z_2(s)\big)\,ds\,dt\\
&\quad+\int_{t_n}^{T}\int_{t}^{T}\beta(s-t)a\big(A_h^l\mathcal{P}_k\mathcal{P}_hj_2,Z_2(s)\big)\,ds\,dt\\
&\quad+\int_{t_n}^{T}\xi(T-t)a(A_h^l\mathcal{P}_k\mathcal{P}_hj_2,Z_2(T))\,dt\\
&\quad-\gamma\int_{t_n}^{T}a\big(A_h^l\mathcal{P}_k\mathcal{P}_hj_2,Z_2(t)\big)\,dt.
\end{aligned}
\tag{3.19}
$$

Finally, for the last term at the left side of (3.16), we use (3.15) and integration by parts to have

$$
\begin{aligned}
\int_{t_n}^{T}\xi(T-t)a\big(A_h^l\mathcal{P}_kZ_1,Z_2(T)\big)\,dt &= -\rho\int_{t_n}^{T}\xi(T-t)a\big(A_h^l\dot Z_2(t),Z_2(T)\big)\,dt\\
&\quad-\int_{t_n}^{T}\xi(T-t)a\big(A_h^l\mathcal{P}_k\mathcal{P}_hj_2,Z_2(T)\big)\,dt\\
&= \rho\int_{t_n}^{T}\beta(T-t)a\big(A_h^lZ_2(t),Z_2(T)\big)\,dt\\
&\quad-\rho\gamma\|Z_2(T)\|_{h,l+1}^2\\
&\quad+\rho\xi(T-t_n)a\big(A_h^lZ_2(t_n),Z_2(T)\big)\\
&\quad-\int_{t_n}^{T}\xi(T-t)a\big(A_h^l\mathcal{P}_k\mathcal{P}_hj_2,Z_2(T)\big)\,dt.
\end{aligned}
\tag{3.20}
$$

Putting (3.17)–(3.20) in (3.16) we conclude the identity (3.13).

Now we prove the estimate (3.14). We recall, from (2.8), that ξ is a positive type kernel. Then, using the Cauchy-Schwarz inequality in (3.13), and $\|\beta\|_{L_1(\mathbb{R}^+)} =$

$\gamma,\ \xi(t)\leq\gamma$, we get

$$
\begin{aligned}
&\|Z_1(t_n)\|_{h,l}^2+\rho\tilde{\gamma}\|Z_2(t_n)\|_{h,l+1}^2\\
&\quad\leq \|Z_1(T)\|_{h,l}^2+\rho(1+\gamma)\|Z_2(T)\|_{h,l+1}^2\\
&\qquad +2\int_{t_n}^T (A_h^l Z_1,\mathcal{P}_k\mathcal{P}_h j_1)\,dt+2\int_{t_n}^T (A_h^{l+1} Z_2,\mathcal{P}_k\mathcal{P}_h j_2)\,dt\\
&\qquad -2\int_{t_n}^T\int_t^T \beta(s-t)a\big(A_h^l\mathcal{P}_k\mathcal{P}_h j_2, Z_2(s)\big)\,ds\,dt\\
&\qquad -2\rho\int_{t_n}^T \beta(T-t)a\big(A_h^l Z_2(t),Z_2(T)\big)\,dt\\
&\qquad -2\rho\xi(T-t_n)a\big(A_h^l Z_2(t_n),Z_2(T)\big)\\
&\quad\leq \|Z_1(T)\|_{h,l}^2+\rho(1+\gamma)\|Z_2(T)\|_{h,l+1}^2\\
&\qquad +C_1\max_{t_n\leq t\leq T}\|Z_1\|_{h,l}^2+1/C_1\Big(\int_{t_n}^T\|\mathcal{P}_k\mathcal{P}_h j_1\|_{h,l}\,dt\Big)^2\\
&\qquad +C_2\max_{t_n\leq t\leq T}\|Z_2\|_{h,l+1}^2+1/C_2\Big(\int_{t_n}^T\|\mathcal{P}_k\mathcal{P}_h j_2\|_{h,l+1}\,dt\Big)^2\\
&\qquad +C_3\|Z_2(T)\|_{h,l+1}^2+1/C_3\max_{t_n\leq t\leq T}\|Z_2\|_{h,l+1}^2\\
&\qquad +C_4\|Z_2(T)\|_{h,l+1}^2+1/C_4\|Z_2(t_n)\|_{h,l+1}^2.
\end{aligned}
$$

Using that, for piecewise linear functions, we have

$$\max_{[0,T]}|U_i|\leq\max_{0\leq n\leq N}|U_i(t_n)|, \tag{3.21}$$

and

$$\int_0^T|\mathcal{P}_k f|\,dt\leq\int_0^T|f|\,dt, \tag{3.22}$$

and that the above inequality holds for arbitrary N, in a standard way, we conclude the estimate inequality (3.14). Now the proof is complete. □

4. A priori error estimates

We define the standard interpolant I_k with $I_k v$ belong to the space of continuous piecewise linear polynomials, and

$$I_k v(t_n)=v(t_n),\quad n=0,\,1,\,\cdots,\,N. \tag{4.1}$$

By standard arguments in approximation theory we see that, for $q=0,\,1$,

$$\int_0^T\|I_k v-v\|_i\,dt\leq Ck^{q+1}\int_0^T\|D_t^{q+1}v\|_i\,dt,\quad\text{for } i=0,\,1, \tag{4.2}$$

where $k=\max_{1\leq n\leq N}k_n$.

We assume the elliptic regularity estimate $\|v\|_2\leq C\|Av\|,\ \forall\, v\in\mathcal{D}(A)$, so that the following error estimates for the Ritz projection (3.12), hold true

$$\|\mathcal{R}_h v-v\|\leq Ch^s\|v\|_s\,,\quad\forall v\in H^s\cap V,\quad s=1,\,2. \tag{4.3}$$

Hence we must specialize to the pure Dirichlet boundary condition and a convex polygonal domain. We note that the energy norm $\|\cdot\|_V$ is equinalent to $\|\cdot\|_1$ on V.

Theorem 4. *Assume that $\Gamma_N = \varnothing$, Ω is a convex polygonal domain, and (3.4). Let u and U be the solutions of (2.3) and (3.5). Then, with $e = U - u$ and $C = C(\rho, \gamma, T)$, we have*

$$\|e_1(T)\| \leq Ch^2\Big(\|u^0\|_2 + \|u_1(T)\|_2 + \int_0^T \|\dot{u}_1\|_2\,dt\Big) + Ck^2\int_0^T \big(\|\ddot{u}_2\| + \|\ddot{u}_1\|_1\big)\,dt.$$

Proof. We recall Remark 4 for the assumption (3.4). We set $e = U - u = \theta + \eta + \omega$ with

$$\begin{aligned}
\theta_1 &= U_1 - I_k\mathcal{R}_h u_1, \quad &\eta_1 &= (I_k - I)\mathcal{R}_h u_1, \quad &\omega_1 &= (\mathcal{R}_h - I)u_1,\\
\theta_2 &= U_2 - I_k\mathcal{P}_h u_2, \quad &\eta_2 &= (I_k - I)\mathcal{P}_h u_2, \quad &\omega_2 &= (\mathcal{P}_h - I)u_2.
\end{aligned}$$

Now, putting $W = \mathcal{P}_k\theta$ in (3.7) we have

$$J(\mathcal{P}_k\theta) = \mathcal{A}^*(\mathcal{P}_k\theta, Z),$$

where by definition

$$\begin{aligned}
J(\mathcal{P}_k\theta) = &\int_0^T \big\{(\mathcal{P}_k\theta_1, j_1) + (\mathcal{P}_k\theta_2, j_2)\big\}\,dt\\
&+ \big((\mathcal{P}_k\theta_1)(T), z_1^T\big) + \rho\big((\mathcal{P}_k\theta_2)(T), z_2^T\big),
\end{aligned}$$

and by partial integration

$$\begin{aligned}
\mathcal{A}^*(\mathcal{P}_k\theta, Z) = \mathcal{A}(\theta, \mathcal{P}_k Z) &+ \big((\mathcal{P}_k\theta_1)(T), Z_1(T)\big) - \big(\theta_1(T), Z_1(T)\big)\\
&+ \rho\big((\mathcal{P}_k\theta_2)(T), Z_2(T)\big) - \rho\big(\theta_2(T), Z_2(T)\big).
\end{aligned}$$

We set $j_1 = j_2 = 0$ and $z_2^T = 0$, $\mathcal{P}_h z_1^T = \theta_1(T)$, and we recall that $Z_i(T) = \mathcal{P}_h z_i^T$, $i = 1, 2$. Hence using the definition of the orthogonal projection $\mathcal{P}_h$ we have

$$\|\theta_1(T)\|^2 = \mathcal{A}(\theta, \mathcal{P}_k Z),$$

that, using $\theta = e - \eta - \omega$ and the Galerkin orthogonality (3.6), implies,

$$\begin{aligned}
\|\theta_1(T)\|^2 &= -\mathcal{A}(\eta, \mathcal{P}_k Z) - \mathcal{A}(\omega, \mathcal{P}_k Z)\\
&= \int_0^T \Big\{ - (\dot{\eta}_1, \mathcal{P}_k Z_1) + (\eta_2, \mathcal{P}_k Z_1) + \rho(\dot{\eta}_2, \mathcal{P}_k Z_2) - a(\eta_1, \mathcal{P}_k Z_2)\\
&\qquad + \int_0^t \beta(t-s)a\big(\eta_1(s), \mathcal{P}_k Z_2\big)\,ds\Big\}\,dt\\
&\quad - \big(\eta_1(0), \mathcal{P}_k Z_1(0)\big) - \rho\big(\eta_2(0), \mathcal{P}_k Z_2(0)\big)\\
&\quad + \int_0^T \Big\{ - (\dot{\omega}_1, \mathcal{P}_k Z_1) + (\omega_2, \mathcal{P}_k Z_1) + \rho(\dot{\omega}_2, \mathcal{P}_k Z_2) - a(\omega_1, \mathcal{P}_k Z_2)\\
&\qquad + \int_0^t \beta(t-s)a\big(\omega_1(s), \mathcal{P}_k Z_2\big)\,ds\Big\}\,dt\\
&\quad - \big(\omega_1(0), \mathcal{P}_k Z_1(0)\big) - \rho\big(\omega_2(0), \mathcal{P}_k Z_2(0)\big).
\end{aligned}$$

By the definition of η, that indicates the interpolation error, terms including $\dot{\eta}_i$, $\eta_i(0)$ vanish. We also use the definition of ω, that indicates the projection error, and we conclude

$$\|\theta_1(T)\|^2 = \int_0^T \Big\{ (\eta_2, \mathcal{P}_k Z_1) - a(\eta_1, \mathcal{P}_k Z_2) + \int_0^t \beta(t-s) a\big(\eta_1(s), \mathcal{P}_k Z_2\big)\, ds \Big\}\, dt - \int_0^T (\dot{\omega}_1, \mathcal{P}_k Z_1)\, dt - \big(\omega_1(0), \mathcal{P}_k Z_1(0)\big),$$

that by the Cauchy-Schwarz inequality implies

$$\begin{aligned} \|\theta_1(T)\|^2 \leq C_1 \max_{0\leq t\leq T} \|\mathcal{P}_k Z_1\|^2 + 1/C_1 \Big(\int_0^T \|\eta_2\|\, dt \Big)^2 \\ + C_2 \max_{0\leq t\leq T} \|\mathcal{P}_k Z_2\|_1^2 + 1/C_2 \Big(\int_0^T \|\eta_1\|_1\, dt \Big)^2 \\ + C_3 \max_{0\leq t\leq T} \|\mathcal{P}_k Z_2\|_1^2 + 1/C_3 \Big(\int_0^T \big(\beta * \|\eta_1\|_1\big)(t)\, dt \Big)^2 \\ + C_4 \max_{0\leq t\leq T} \|\mathcal{P}_k Z_1\|^2 + 1/C_4 \Big(\int_0^T \|\dot{\omega}_1\|\, dt \Big)^2 \\ + C_5 \|\mathcal{P}_k Z_1(0)\|^2 + 1/C_5 \|\omega_1(0)\|^2. \end{aligned}$$

Using (3.21), $\|\beta\|_{L_1(\mathbb{R}^+)} = \gamma$, and the stability estimate (3.14) with $l = 0$, in a standard way, we have

$$\|\theta_1(T)\| \leq C \Big\{ \|\omega_1(0)\| + \int_0^T \Big(\|\eta_2\| + \|\eta_1\|_1 + \|\dot{\omega}_1\| \Big)\, dt \Big\}.$$

Recalling $\theta_1(T) = e_1(T) - \omega_1(T)$, and stability of the projections $\mathcal{P}_h$, $\mathcal{R}_h$ with respect to $\|\cdot\|, \|\cdot\|_1$, respectively, we have

$$\begin{aligned} \|\theta_1(T)\| \leq C \Big\{ \|(\mathcal{R}_h - I)u^0\| + \|(\mathcal{R}_h - I)u_1(T)\| \\ + \int_0^T \Big(\|(I_k - I)u_2\| + \|(I_k - I)u_1\|_1 + \|(\mathcal{R}_h - I)\dot{u}_1\| \Big)\, dt \Big\}. \end{aligned}$$

This completes the proof by (4.2), (4.3). □

5. A posteriori error estimates

Having certain regularity on the data, i.e., initial data u^0, v^0 and the force terms f, g, there are still two types of limitation for higher global regularity of a weak solution of (1.1). One is due to the mixed Dirichlet-Neumann boundary condition. This type of boundary condition are natural in practice, and a pure Dirichlet boundary condition can not be realistic in applications. Other limitation is the singularity of the convolution kernel β. This means that even with the pure Dirichlet boundary condition, higher regularity of a weak solution is limited, see [7], [11], though with smoother kernels we can get higher regularity. Besides, the stability and a priori error estimates presented in Theorem 3 and Theorem 4 do not admit adaptive meshes. These, and other general motivations such as no practical use of a priori error estimates, call for adaptive meshes based on a posteriori error analysis.

Here a space-time cellwise error representation is given. The main framework is adapted from [3], and a general linear goal functional $J(\cdot)$ is used. This error representation can be used for goal-oriented adaptive strategies based on dual weighted residual method. For more details on dual weighted residual method and its practical aspects for differential equations, see [3] and references therein.

Theorem 5. *Let u and U be the solutions of* (2.3) *and* (3.5), *and $J(\cdot)$ the linear functional defined in* §2. *Then, with $e = U - u$, we have the error representation*

$$J(e) = \sum_{K\in\mathcal{T}_h^0} \Theta_{0,K} + \sum_{n=1}^{N} \sum_{K\in\bar{\mathcal{T}}_h^n} \sum_{i=1}^{6} \Theta_{i,K}^n, \tag{5.1}$$

where, with $z_{hk} \in \mathcal{W}_{hk}$ being an approximation of the dual solution z and $E_{hk}z = z_{hk} - z$ being the error operator,

$$\begin{aligned}
\Theta_{0,K} &= \big(U_1(0) - u^0, E_{hk}z_1(0)\big)_K + \rho\big(U_2(0) - v^0, E_{hk}z_2(0)\big)_K,\\
\Theta_{1,K}^n &= (\dot{U}_1 - U_2, E_{hk}z_1)_{K^n}, \quad \Theta_{2,K}^n = (\rho\dot{U}_2 - f, E_{hk}z_2)_{K^n},\\
\Theta_{3,K}^n &= (r_h, E_{hk}z_2)_{\partial K^n}, \qquad \Theta_{4,K}^n = (g_h - g, E_{hk}z_2)_{\partial K^n},\\
\Theta_{5,K}^n &= -\Big(r_h, \int_s^T \beta(t-s)E_{hk}z_2(t)\,dt\Big)_{\partial K^n},\\
\Theta_{6,K}^n &= -\Big(g_h, \int_s^T \beta(t-s)E_{hk}z_2(t)\,dt\Big)_{\partial K^n}.
\end{aligned} \tag{5.2}$$

Here $K^n = K \times I_n$ and $\partial K^n = \partial K \times I_n$ are the space-time cells, and r_h, g_h are defined below, (5.6), (5.7), *respectively.*

It should be noticed that ∂K^n is not the boundary of K^n.

Proof. Using the identity (2.4) and the Galerkin orthogonality (3.6) we have,

$$\begin{aligned}
J(e) &= \mathcal{A}^*(e,z) = \mathcal{A}(e,z) = \mathcal{A}(e, E_{hk}z)\\
&= \mathcal{A}(U, E_{hk}z) - F(E_{hk}z) = \mathrm{R}(U; E_{hk}z),
\end{aligned} \tag{5.3}$$

where $\mathrm{R}(U;\cdot)$ is the residual of the Galerkin approximation U as a functional on the solution space $\mathcal{V}^*$. Then by the definition of $\mathcal{A}$, F we have

$$\begin{aligned}
J(e) = {}& \big(U_1(0) - u^0, E_{hk}z_1(0)\big) + \rho\big(U_2(0) - v^0, E_{hk}z_2(0)\big)\\
&+ \int_0^T \Big\{(\dot{U}_1, E_{hk}z_1) - (U_2, E_{hk}z_1) + \rho(\dot{U}_2, E_{hk}z_2) + a(U_1, E_{hk}z_2)\\
&\qquad - \int_0^t \beta(t-s)a\big(U_1(s), E_{hk}z_2(t)\big)\,ds\Big\}\,dt\\
&- \int_0^T \Big\{(f, E_{hk}z_2) + (g, E_{hk}z_2)_{\Gamma_{\mathrm{N}}}\Big\}\,dt.
\end{aligned} \tag{5.4}$$

Now, by partial integration with respect to the space variable, we obtain

$$
\begin{aligned}
\int_0^T a(U_1, z_2)\,dt &= \sum_{n=1}^N \int_{I_n} \sum_{K\in\bar{\mathcal{T}}_h^n} a(U_1, z_2)_K\,dt \\
&= \sum_{n=1}^N \int_{I_n} \sum_{K\in\bar{\mathcal{T}}_h^n} \big(\sigma_0(U_1)\cdot n, z_2\big)_{\partial K}\,dt \\
&= \sum_{n=1}^N \int_{I_n} \Big\{ \sum_{E\in\mathcal{E}_I^n} \big(-[\sigma_0(U_1)\cdot n], z_2\big)_E \\
&\qquad + \sum_{E\in\mathcal{E}_{\Gamma_{\mathrm{N}}}^n} \big(\sigma_0(U_1)\cdot n, z_2\big)_E \Big\}\,dt \\
&= \sum_{n=1}^N \int_{I_n} \sum_{K\in\bar{\mathcal{T}}_h^n} \big\{(r_h, z_2)_{\partial K} + (g_h, z_2)_{\partial K}\big\}\,dt \\
&= \sum_{n=1}^N \sum_{K\in\bar{\mathcal{T}}_h^n} \big\{(r_h, z_2)_{\partial K^n} + (g_h, z_2)_{\partial K^n}\big\}
\end{aligned}
\tag{5.5}
$$

where $\mathcal{E}_I^n$, $\mathcal{E}_{\Gamma_{\mathrm{N}}}^n$ are, respectively, the sets of the interior edges and the edges on the Neumann boundary, corresponding to the triangulation $\bar{\mathcal{T}}_h^n$. Here r_h are the residuals representing the jumps of the normal derivatives $\sigma_0(U_1)\cdot n$, and determined by,

$$
r_h|_\Gamma = \begin{cases} -\frac{1}{2}[\sigma_0(U_1)\cdot n] & \text{if } \Gamma\subset\partial K\setminus\partial\Omega, \\ 0 & \text{if } \Gamma\subset\partial\Omega, \end{cases}
\tag{5.6}
$$

and g_h is the contribution from the Neumann boundary defined as

$$
g_h|_\Gamma = \begin{cases} \sigma_0(U_1)\cdot n & \text{if } \Gamma\subset\partial K\cap\Gamma_{\mathrm{N}}, \\ 0 & \text{if } \Gamma\subset\partial\Omega. \end{cases}
\tag{5.7}
$$

For the convolution term in (5.4) we first change the order of the time integrals, then similar to (5.5), we have,

$$
\begin{aligned}
\int_0^T \int_0^t & \beta(t-s) a\big(U_1(s), z_2(t)\big)\,ds\,dt \\
&= \int_0^T \int_s^T \beta(t-s) a\big(U_1(s), z_2(t)\big)\,dt\,ds \\
&= \sum_{n=1}^N \int_{I_n} \sum_{K\in\bar{\mathcal{T}}_h^n} a\Big(U_1(s), \int_s^T \beta(t-s) z_2(t)\,dt\Big)_K\,ds \\
&= \sum_{n=1}^N \sum_{K\in\bar{\mathcal{T}}_h^n} \Big\{ \Big(r_h, \int_s^T \beta(t-s) z_2(t)\,dt\Big)_{\partial K^n} \\
&\qquad + \Big(g_h, \int_s^T \beta(t-s) z_2(t)\,dt\Big)_{\partial K^n} \Big\}.
\end{aligned}
\tag{5.8}
$$

Now, using (5.5), (5.8) and space-time cellwise representation of the other terms in (5.4) we conclude the error representation (5.1). □

The error representation (5.1) leads us to the weighted a posteriori estimate,

$$J(e) \le \sum_{K\in\mathcal{T}_h^0} \mathrm{R}_{0,K}\omega_{0,K} + \sum_{n=1}^{N}\sum_{K\in\bar{\mathcal{T}}_h^n}\sum_{i=1}^{3} \mathrm{R}_{i,K}^n\omega_{i,K}^n, \tag{5.9}$$

with the residuals and weights defined as

$$\begin{aligned}
\mathrm{R}_{0,K} &= \big(\|U_1(0)-u^0\|_K^2 + \|U_2(0)-v^0\|_K^2\big)^{1/2},\\
\omega_{0,K} &= \big(\|E_{hk}z_1(0)\|_K^2 + \|E_{hk}z_2(0)\|_K^2\big)^{1/2},\\
\mathrm{R}_{1,K}^n &= \|\dot{U}_1 - U_2\|_{K^n}, \quad \omega_{1,K}^n = \|E_{hk}z_1\|_{K^n},\\
\mathrm{R}_{2,K}^n &= \Big(\|\rho\dot{U}_2 - f\|_{K^n}^2 + 2\bar{h}_K^{-1}\|r_h\|_{\partial K^n}^2\Big)^{1/2},\\
\omega_{2,K}^n &= \Bigg(\|E_{hk}z_2\|_{K^n}^2 + \bar{h}_K\|E_{hk}z_2\|_{\partial K^n}^2\\
&\qquad + \bar{h}_K\Big\|\int_s^T \beta(t-s)E_{hk}z_2(t)\,dt\Big\|_{\partial K^n}^2\Bigg)^{1/2},\\
\mathrm{R}_{3,K}^n &= \Big(\bar{h}_K^{-1}\|g_h - g\|_{\partial K^n}^2 + \bar{h}_K^{-1}\|g_h\|_{\partial K^n}^2\Big)^{1/2},\\
\omega_{3,K}^n &= \Big(\bar{h}_K\|E_{hk}z_2\|_{\partial K^n}^2 + \bar{h}_K\Big\|\int_s^T \beta(t-s)E_{hk}z_2(t)\,dt\Big\|_{\partial K^n}^2\Big)^{1/2}.
\end{aligned}$$

In order to evaluate the a posteriori error representation (5.1) or the a posteriori estimate (5.9), we need information about the continuous dual solution z. Such information has to be obtained either through a priori analysis in form of bounds for z in certain Sobolev norms or through computation by solving the dual problem numerically. In this context we provide information through a priori analysis and we leave the investigation on the second case to a latter work.

In the following, the target functional $J(\cdot)$ will be the global L_2-norm of the approximation displacement $u_1 = u_1(x,t)$. We first present a weighted global a posteriori error estimate, using global L_2-projections $\mathcal{P}_k, \mathcal{P}_h$ defined in (3.12), and error estimates of $\mathcal{P}_h$ in a weighted L_2-norm.

We recall the weighted global error estimates of the L_2-projection $\mathcal{P}_h$ (3.12), see [4]. First we recall some notation. Let $\mathcal{T}$ be a given triangulation with mesh function h, and for any simplex $K\in\mathcal{T}$, ρ_K denote the radius of the largest ball contained in the closure of K, that is $\bar{K}$. A family $\mathcal{F}$ of triangulations $\mathcal{T}$ is called non-degenerate, if there exsist a constant c_0 such that we have

$$c_0 = \max_{\mathcal{T}\in\mathcal{F}}\max_{K\in\mathcal{T}}\frac{h_K}{\rho_K}.$$

Let $S_K = \{K'\in\mathcal{T} : \bar{K}'\cap\bar{K}\neq\varnothing\}$ and $\delta_{\mathcal{T}}$ be a measure for the given triangulation $\mathcal{T}$ defined by

$$\delta_{\mathcal{T}} = \max_{K\in\mathcal{T}}\max_{K'\in S_K}|1 - h_{K'}^2/h_K^2|.$$

We define a measure, $\delta_{\mathcal{F}}$, for a given family $\mathcal{F}$, by

$$\delta_{\mathcal{F}} = \max_{\mathcal{T}\in\mathcal{F}}\delta_{\mathcal{T}}. \tag{5.10}$$

We define the error operators E_{hk}, E_h, and E_k by

$$E_{hk}v = (\mathcal{P}_k\mathcal{P}_h - I)v, \quad E_h v = (\mathcal{P}_h - I)v, \quad E_k v = (\mathcal{P}_k - I)v, \tag{5.11}$$

and we note that

$$E_{hk} = E_h + E_k\mathcal{P}_h. \tag{5.12}$$

Lemma 1. *Assume that the family $\mathcal{F}$ of traingulations $\mathcal{T}$ be non-degenerate. Then for sufficiently small $\delta_{\mathcal{F}}$, there exists a constant C such that for any triangulation $\mathcal{T} \in \mathcal{F}$ we have, for all $v \in H^2$,*

$$\|h^{-s}E_h v\| \leq C\|\nabla^s v\|, \quad s = 1,\, 2,\ \forall v \in H^s, \tag{5.13}$$

$$\|h^{-1}\nabla E_h v\| \leq C\|\nabla^2 v\|, \quad \forall v \in H^2, \tag{5.14}$$

where '∇' denotes the usual gradient.

For more details on the practical aspects of $\delta_{\mathcal{F}}$, see [4].

For the next theorem we recall the mesh functions h_n, $\bar{h}_n$ from (3.1), (3.2), and we define the notations

$$\bar{h}_{min,n} = \min_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K, \quad h_{max,n} = \max_{K\in\mathcal{T}_h^n} h_K.$$

Theorem 6. *Let u be the solutions of* (2.3)*, and U be the solution of* (3.5) *with a non-degenerate family $\mathcal{F}_h$ of triangulations $\mathcal{T}_h^n$, $n = 0, 1, \ldots, N$, with sufficiently small $\delta_{\mathcal{F}_h}$, such that the weighted global error estimates* (5.13) *and* (5.14) *hold. Then, with $e = U - u$, we have the weighted a posteriori error estimate*

$$\begin{aligned}
\|e_1(T)\| \leq C\Big\{ & \|h_0\big(U_1(0) - u^0\big)\| + \|h_0^2\big(U_2(0) - v^0\big)\| \\
& + \sum_{n=1}^{N}\int_{I_n}\Big\{\|h_n(\dot{U}_1 - U_2)\| + \|h_n^2(\rho\dot{U}_2 - f)\| \\
& + (\zeta_n + \zeta_{n,N})\Big(\sum_{K\in\bar{\mathcal{T}}_h^n}\bar{h}_K^3\|r_h\|_{\partial K}^2\Big)^{1/2} \\
& + \zeta_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n}\bar{h}_K^3\|g_h - g\|_{\partial K}^2\Big)^{1/2} + \zeta_{n,N}\Big(\sum_{K\in\bar{\mathcal{T}}_h^n}\bar{h}_K^3\|g_h\|_{\partial K}^2\Big)^{1/2} \\
& + k_n\|\dot{U}_1 - U_2\| + k_n\|E_k\bar{A}_h U_1\| \\
& + k_n\Big\|E_k\int_0^t\beta(t-s)\bar{A}_h U_1(s)\,ds\Big\| \\
& + k_n\|E_k f\| + k_n\Big(\sum_{K\in\mathcal{T}_h^n}h_K^{-1}\|E_k g\|_{\partial K}^2\Big)^{1/2}\Big\}\,dt\Big\},
\end{aligned} \tag{5.15}$$

where

$$\zeta_n = \bar{h}_{min,n}^{-2}h_{max,n}^2, \quad \zeta_{n,N} = \bar{h}_{min,n}^{-2}\max_{n\leq j\leq N}h_{max,j}^2. \tag{5.16}$$

Proof. Let $z \in \mathcal{V}^*$ be the solution of the dual problem (2.5). From the definition of the L_2 projections $\mathcal{P}_k, \mathcal{P}_h$ in (3.12) and the test space $\mathcal{W}_{hk}$ in (3.3) we have $\mathcal{P}_k\mathcal{P}_h z \in \mathcal{W}_{hk}$. Therefore, using (5.3) and the error operators (5.11) we have,

$$J(e) = \mathrm{R}(U; E_{hk}z) = \mathrm{R}(U; E_h z) + \mathrm{R}(U; E_k\mathcal{P}_h z), \tag{5.17}$$

where we used (5.12). We study the two terms at the right side of this equation.

For the first term we can write,

$$
\begin{aligned}
\mathrm{R}(U;E_h z) &= \big(U_1(0)-u^0, E_h z_1(0)\big) + \rho\big(U_2(0)-v^0, E_h z_2(0)\big) \\
&\quad + \sum_{n=1}^{N} \int_{I_n} \Big\{ \big(\dot U_1 - U_2, E_h z_1\big) + \big(\rho \dot U_2 - f, E_h z_2\big) \\
&\quad + a(U_1, E_h z_2) - (g, E_h z_2)_{\Gamma_{\mathrm N}} \\
&\quad - \int_0^t \beta(t-s) a\big(U_1(s), E_h z_2\big)\, ds \Big\}\, dt.
\end{aligned}
$$

Now by partial integration in space, similar to (5.5), (5.8), we have

$$
\begin{aligned}
\mathrm{R}(U;E_h z) &= \sum_{K\in\mathcal{T}_h^0} \big\{ \big(U_1(0)-u^0, E_h z_1(0)\big)_K + \rho\big(U_2(0)-v^0, E_h z_2(0)\big)_K \big\} \\
&\quad + \sum_{n=1}^{N} \int_{I_n} \sum_{K\in\bar{\mathcal{T}}_h^n} \Big\{ \big(\dot U_1 - U_2, E_h z_1\big)_K + (\rho \dot U_2 - f, E_h z_2)_K \\
&\quad + (r_h, E_h z_2)_{\partial K} + (g_h - g, E_h z_2)_{\partial K} \\
&\quad - \Big(r_h, \int_t^T \beta(s-t) E_h z_2(s)\, ds \Big)_{\partial K} \\
&\quad - \Big(g_h, \int_t^T \beta(s-t) E_h z_2(s)\, ds \Big)_{\partial K} \Big\}\, dt \\
&= \sum_{i=1}^{2} \mathcal{I}_i + \sum_{i=1}^{6} \mathcal{II}_i.
\end{aligned} \tag{5.18}
$$

We then, for each term, use the Cauchy-Schwarz inequality twice. First on the local elements $K, \partial K$, to obtain local L_2-norms, and then on the sum over the elements to obtain global norms such that the weighted global error estimates (5.13), (5.14) can be used. For $\mathcal{I}_1$ we have

$$
\mathcal{I}_1 \le \|h_0(U_1(0)-u^0)\| \|h_0^{-1} E_h z_1(0)\| \le C\|h_0(U_1(0)-u^0)\| \|\nabla z_1(0)\|, \tag{5.19}
$$

and in a similar way we have,

$$
\mathcal{I}_2 \le C\|h_0^2(U_2(0)-v^0)\| \|\nabla^2 z_2(0)\|. \tag{5.20}
$$

For the next term, using the error estimate (5.13), we have,

$$
\begin{aligned}
\mathcal{II}_1 &= \sum_{n=1}^{N} \int_{I_n} \sum_{K\in\mathcal{T}_h^n} \big(\dot U_1 - U_2, E_h z_1\big)_K\, dt \\
&\le \sum_{n=1}^{N} \int_{I_n} \|h_n(\dot U_1 - U_2)\| \|h_n^{-1} E_h z_1\|\, dt \\
&\le C \max_{[0,T]} \|\nabla z_1(t)\| \sum_{n=1}^{N} \int_{I_n} \|h_n(\dot U_1 - U_2)\|\, dt,
\end{aligned} \tag{5.21}
$$

and similarly we obtain,

$$\mathcal{II}_2 \leq C \max_{[0,T]} \|\nabla^2 z_2(t)\| \sum_{n=1}^N \int_{I_n} \|h_n^2(\rho \dot{U}_2 - f)\| \, dt. \tag{5.22}$$

For $\mathcal{II}_3$, we first have,

$$\mathcal{II}_3 \leq \sum_{n=1}^N \int_{I_n} \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h(t)\|_{\partial K}^2 \Big)^{1/2} \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \|E_h z_2(t)\|_{\partial K}^2 \Big)^{1/2} dt.$$

Then by a scaled trace inequality and the weighted global error estimates (5.13), (5.14), we obtain

$$\begin{aligned}\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \|E_h z_2\|_{\partial K}^2 &\leq C \sum_{K \in \bar{\mathcal{T}}_h^n} \big\{ \bar{h}_K^{-4} \|E_h z_2\|_K^2 + \bar{h}_K^{-2} \|\nabla E_h z_2\|_K^2 \big\} \\ &\leq C \big\{ \bar{h}_{min,n}^{-4} h_{max,n}^4 \|h_n^{-2} E_h z_2\|^2 \\ &\qquad + \bar{h}_{min,n}^{-2} h_{max,n}^2 \|h_n^{-1} \nabla E_h z_2\|^2 \big\} \\ &\leq C \bar{h}_{min,n}^{-4} h_{max,n}^4 \|\nabla^2 z_2\|^2.\end{aligned}$$

These imply the estimate

$$\mathcal{II}_3 \leq C \max_{[0,T]} \|\nabla^2 z_2(t)\| \sum_{n=1}^N \int_{I_n} \bar{h}_{min,n}^{-2} h_{max,n}^2 \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h\|_{\partial K}^2 \Big)^{1/2} dt. \tag{5.23}$$

In a similar way, we have

$$\mathcal{II}_4 \leq C \max_{[0,T]} \|\nabla^2 z_2(t)\| \sum_{n=1}^N \int_{I_n} \bar{h}_{min,n}^{-2} h_{max,n}^2 \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|g_h - g\|_{\partial K}^2 \Big)^{1/2} dt. \tag{5.24}$$

Finally we study $\mathcal{II}_5$ and a similar result will hold for $\mathcal{II}_6$. To this end, first we note that,

$$\begin{aligned}\mathcal{II}_5 \leq \sum_{n=1}^N \int_{I_n} &\Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h(t)\|_{\partial K}^2 \Big)^{1/2} \\ &\times \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \Big\| \int_t^T \beta(s-t) E_h z_2(s) \, ds \Big\|_{\partial K}^2 \Big)^{1/2} dt.\end{aligned}$$

Then, using Minkowski's inequality, the Cuachy-Schwarz inequality, and the fact that $\|\beta\|_{L_1(\mathbb{R}^+)} = \gamma$, we have

$$\begin{aligned}
&\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-3}\Big\| \int_t^T \beta(s-t) E_h z_2(s)\,ds\Big\|_{\partial K}^2 \\
&\qquad \le \sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-3}\Big(\int_t^T \beta(s-t)\|E_h z_2(s)\|_{\partial K}\Big)^2\,ds \\
&\qquad \le \sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-3} \int_t^T \beta(s-t)\,ds \int_t^T \beta(s-t)\|E_h z_2(s)\|_{\partial K}^2\,ds \\
&\qquad \le \gamma \int_t^T \beta(s-t) \sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-3}\|E_h z_2(s)\|_{\partial K}^2\,ds,
\end{aligned}$$

that using a scaled trace inequality and the error estimates (5.13), (5.14), we have

$$\begin{aligned}
&\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-3}\Big\| \int_t^T \beta(s-t) E_h z_2(s)\,ds\Big\|_{\partial K}^2 \\
&\qquad \le C\int_t^T \beta(s-t) \sum_{K\in\bar{\mathcal{T}}_h^n} \big\{\bar{h}_K^{-4}\|E_h z_2(s)\|_K^2 + \bar{h}_K^{-2}\|\nabla E_h z_2(s)\|_K^2\big\}\,ds \\
&\qquad \le C\sum_{j=n}^N \int_{t\vee t_{j-1}}^{t_j} \beta(s-t)\big\{\bar{h}_{min,n}^{-4} h_{max,j}^4 \|h_j^{-2} E_h z_2(s)\|^2 \\
&\qquad\quad + \bar{h}_{min,n}^{-2} h_{max,j}^2 \|h_j^{-1}\nabla E_h z_2(s)\|^2\big\}\,ds \\
&\qquad \le C\sum_{j=n}^N \int_{t\vee t_{j-1}}^{t_j} \beta(s-t)\bar{h}_{min,n}^{-4} h_{max,j}^4 \|\nabla^2 z_2(s)\|^2\,ds \\
&\qquad \le C\gamma \max_{[t,T]} \|\nabla^2 z_2(s)\|^2 \bar{h}_{min,n}^{-4} \max_{n\le j\le N} h_{max,j}^4.
\end{aligned}$$

Hence we obtain,

$$\begin{aligned}
\mathcal{II}_5 &\le C\max_{[0,T]} \|\nabla^2 z_2(t)\| \\
&\quad\times \sum_{n=1}^N \int_{I_n} \bar{h}_{min,n}^{-2} \max_{n\le j\le N} h_{max,j}^2 \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h\|_{\partial K}^2\Big)^{1/2}\,dt.
\end{aligned} \tag{5.25}$$

A similar estimate for $\mathcal{II}_6$ holds, that is,

$$\begin{aligned}
\mathcal{II}_6 &\le C\max_{[0,T]} \|\nabla^2 z_2(t)\| \\
&\quad\times \sum_{n=1}^N \int_{I_n} \bar{h}_{min,n}^{-2} \max_{n\le j\le N} h_{max,j}^2 \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|g_h\|_{\partial K}^2\Big)^{1/2}\,dt.
\end{aligned} \tag{5.26}$$

Putting (5.19)–(5.26) in (5.18) we conclude,

$$
\begin{aligned}
\mathrm{R}(U;E_hz) \leq C\max\big(\max_{[0,T]}\|\nabla^2 z_2(t)\|,\max_{[0,T]}\|\nabla z_1(t)\|\big)\\
&\times\Big\{\|h_0\big(U_1(0)-u^0\big)\|+\|h_0^2\big(U_2(0)-v^0\big)\|\\
&+\sum_{n=1}^N\int_{I_n}\Big\{\|h_n(\dot U_1-U_2)\|+\|h_n^2(\rho\dot U_2-f)\|\\
&+\bar h_{min,n}^{-2}h_{max,n}^2\Big(\sum_{K\in\bar{\mathcal T}_h^n}\bar h_K^3\|r_h\|_{\partial K}^2\Big)^{1/2}\\
&+\bar h_{min,n}^{-2}h_{max,n}^2\Big(\sum_{K\in\bar{\mathcal T}_h^n}\bar h_K^3\|g_h-g\|_{\partial K}^2\Big)^{1/2}\\
&+\bar h_{min,n}^{-2}\max_{n\leq j\leq N}h_{max,j}^2\Big(\sum_{K\in\bar{\mathcal T}_h^n}\bar h_K^3\|r_h\|_{\partial K}^2\Big)^{1/2}\\
&+\bar h_{min,n}^{-2}\max_{n\leq j\leq N}h_{max,j}^2\Big(\sum_{K\in\bar{\mathcal T}_h^n}\bar h_K^3\|g_h\|_{\partial K}^2\Big)^{1/2}\Big\}\,dt\Big\}.
\end{aligned}
\tag{5.27}
$$

Now we study the second term in (5.17), that is,

$$
\begin{aligned}
\mathrm{R}(U;E_k\mathcal P_hz)&=\mathrm{R}(U;E_k\mathcal P_hz)\pm\int_0^T\big\{(\mathcal P_kf,E_k\mathcal P_hz_2)+(\mathcal P_kg,E_k\mathcal P_hz_2)_{\Gamma_\mathrm{N}}\big\}\,dt\\
&=\big(U_1(0)-u^0,E_k\mathcal P_hz_1(0)\big)+\rho\big(U_2(0)-v^0,E_k\mathcal P_hz_2(0)\big)\\
&\quad+\sum_{n=1}^N\int_{I_n}\Big\{\big(\dot U_1-U_2,E_k\mathcal P_hz_1\big)\\
&\qquad+a(U_1,E_k\mathcal P_hz_2)-\int_0^t\beta(t-s)a\big(U_1(s),E_k\mathcal P_hz_2\big)\,ds\\
&\qquad+\big(\rho\dot U_2-\mathcal P_kf,E_k\mathcal P_hz_2\big)-(\mathcal P_kg,E_k\mathcal P_hz_2)_{\Gamma_\mathrm{N}}\\
&\qquad+(E_kf,E_k\mathcal P_hz_2)+(E_kg,E_k\mathcal P_hz_2)_{\Gamma_\mathrm{N}}\Big\}\,dt.
\end{aligned}
$$

Recalling the initial condition $U_i(0)=\mathcal P_hu_i(0)$, $i=1,2$, the first two terms on the right side vanish. Besides, from the second equation of (3.9) we have, for $W\in V_h^n$,

$$
\begin{aligned}
&\int_{I_n}\big\{\rho(\dot U_2,W)-(\mathcal P_kf,W)-(\mathcal P_kg,W)\big\}\,dt\\
&\qquad=-\int_{I_n}\Big\{a(\mathcal P_kU_1,W)-a\Big(\mathcal P_k\int_0^t\beta(t-s)U_1(s)\,ds,W\Big)\Big\}\,dt.
\end{aligned}
$$

Hence, we conclude

$$
\begin{aligned}
\mathrm{R}(U;E_k\mathcal{P}_hz) = \sum_{n=1}^{N}\int_{I_n}\Big\{&\big(\dot{U}_1-U_2,E_k\mathcal{P}_hz_1\big)\\
&-a(E_kU_1,E_k\mathcal{P}_hz_2)+a\Big(E_k\int_0^t\beta(t-s)U_1(s)\,ds,E_k\mathcal{P}_hz_2\Big)\\
&+(E_kf,E_k\mathcal{P}_hz_2)+(E_kg,E_k\mathcal{P}_hz_2)_{\Gamma_\mathrm{N}}\Big\}\,dt.
\end{aligned}
\tag{5.28}
$$

For the last term we have,

$$
\begin{aligned}
&\sum_{n=1}^{N}\int_{I_n}(E_kg,E_k\mathcal{P}_hz_2)_{\Gamma_\mathrm{N}}\,dt\\
&\qquad\le\sum_{n=1}^{N}\int_{I_n}\Big(\sum_{K\in\mathcal{T}_h^n}h_K^{-1}\|E_kg\|_{\partial K}^2\Big)^{1/2}\Big(\sum_{K\in\mathcal{T}_h^n}h_K\|E_k\mathcal{P}_hz_2\|_{\partial K}^2\Big)^{1/2}.
\end{aligned}
$$

By a scaled trace inequality and local inverse inequality we have,

$$
\begin{aligned}
\sum_{K\in\bar{\mathcal{T}}_h^n}h_K\|E_k\mathcal{P}_hz_2\|_{\partial K}^2 &\le C\sum_{K\in\mathcal{T}_h^n}\big\{\|E_k\mathcal{P}_hz_2\|_K^2+h_K^2\|\nabla E_k\mathcal{P}_hz_2\|_K^2\big\}\\
&\le C\sum_{K\in\mathcal{T}_h^n}\big\{\|E_k\mathcal{P}_hz_2\|_K^2+\|E_k\mathcal{P}_hz_2\|_K^2\big\}\\
&= C\|E_k\mathcal{P}_hz_2\|^2.
\end{aligned}
$$

Hence,

$$
\sum_{n=1}^{N}\int_{I_n}(E_kg,E_k\mathcal{P}_hz_2)_{\Gamma_\mathrm{N}}\,dt\le C\Big(\sum_{K\in\mathcal{T}_h^n}h_K^{-1}\|E_kg\|_{\partial K}^2\Big)^{1/2}\|E_k\mathcal{P}_hz_2\|.
$$

Considering this in (5.28) and using the Cauchy-Schwarz inequality we have

$$
\begin{aligned}
\mathrm{R}(U;E_k\mathcal{P}_hz)\le C\sum_{n=1}^{N}\int_{I_n}\Big\{&\|\dot{U}_1-U_2\|\|E_k\mathcal{P}_hz_1\|+\|E_k\bar{A}_hU_1\|\|E_k\mathcal{P}_hz_2\|\\
&+\Big\|E_k\int_0^t\beta(t-s)\bar{A}_hU_1(s)\,ds\Big\|\|E_k\mathcal{P}_hz_2\|\\
&+\|E_kf\|\|E_k\mathcal{P}_hz_2\|+\Big(\sum_{K\in\mathcal{T}_h^n}h_K^{-1}\|E_kg\|_{\partial K}^2\Big)^{1/2}\|E_k\mathcal{P}_hz_2\|\Big\}.
\end{aligned}
$$

This, L_2-stability of the L_2-projection $\mathcal{P}_h$, and a standard error estimation of the error operator E_k, conclude

$$
\begin{aligned}
\mathrm{R}(U; E_k\mathcal{P}_h z) \leq C \max\big(\max_{[0,T]}\|\dot z_1(t)\|, \max_{[0,T]}\|\dot z_2(t)\|\big)& \\
\times \sum_{n=1}^N \int_{I_n} \Big\{ k_n\|\dot U_1 - U_2\| + k_n\|E_k\bar A_h U_1\|& \\
+ k_n\big\|E_k\int_0^t \beta(t-s)\bar A_h U_1(s)\,ds\big\| + k_n\|E_k f\|& \\
+ k_n\Big(\sum_{K\in\mathcal{T}_h^n} h_K^{-1}\|E_k g\|_{\partial K}^2\Big)^{1/2}\Big\}\,dt.&
\end{aligned}
\tag{5.29}
$$

We now set $j_1 = j_2 = z_2^T = 0$ and $z_1^T = A^{-1/2}e_1(T)$. Then, putting (5.27) and (5.29) in (5.17), using the stability estimates (2.14) and a standard argument, we conclude the a posteriori error estimate (5.15), and this completes the proof. ☐

Remark 5. We note that for the error estimate (5.15) there are two types of restriction on the triangulations; One by ζ_n, $\zeta_{n,N}$, that measures the quasiuniformity of the family of triangulation, and the other by $\delta_{\mathcal{F}_h}$, that measure the regularity of the family of triangulations in a slightly different sense. Although maybe not explicitly, but ζ_n, $\zeta_{n,N}$ and $\delta_{\mathcal{F}_h}$ can be related. In practice we use finitely many triangulations, that means quasiuniformity holds, though possibly with big ζ_n, $\zeta_{n,N}$. This means that we still can use the a posteriori error estimate (5.15). But when $\delta_{\mathcal{F}_h}$ is not sufficiently small, the error estimate (5.15) does not hold. This calls for using local interpolants instead of global L_2-projection $\mathcal{P}_h$. In the next theorem we present an a posteriori error estimate using interpolation, linear in space and constant in time on each space-time cell. We also note that a possible, but not neccesserily optimal, way of ignoring these limitations could be using global error estimates of $\mathcal{P}_h$ with global mesh size h_{max}. That is a posteriori error estimate in the form, with $k(t) = k_n$ for $t \in I_n$,

$$
\begin{aligned}
\|e_1(T)\| \leq{}& C\big\{h_{max,0}\|U_1(0)-u^0\| + h_{max,0}^2\|U_2(0)-v^0\|\big\} \\
&+ Ch_{max}\int_0^T \|\dot U_1 - U_2\|\,dt \\
&+ Ch_{max}^2\int_0^T \Big\{\|\rho\dot U_2 - f\| + \|\tilde r_h\| + \|\tilde g_h - \tilde g\|_{\Gamma_{\mathrm{N}}} + \|\tilde g_h\|_{\Gamma_{\mathrm{N}}}\Big\}\,dt \\
&+ \int_0^T k\Big\{\|\dot U_1 - U_2\| + \|E_k\bar A_h U_1\| \\
&\quad + \Big\|E_k\int_0^t \beta(t-s)\bar A_h U_1(s)\,ds\Big\| + \|E_k f\| + \|E_k g\|_{\Gamma_{\mathrm{N}}}\Big\}\,dt,
\end{aligned}
$$

where $\tilde r_h|_K = h_K^{-1}\max_{\partial K}|r_h|$, $\tilde g_h|_K = h_K^{-1/2}|g_h|$, and $\tilde g|_K = h_K^{-1/2}|g|$.

We recall the decomposition of the space-time slab $\Omega^n = \Omega\times I_n$ into cells $K^n = K\times I_n$, $K\in\bar{\mathcal{T}}_h^n$. Let I_{hk} be the standard interpolant, such that $I_{hk}v|_{K^n}$ be linear in space and constant in time. We define the error operator E_{hk} by $E_{hk}v = (I_{hk}-I)v$.

A variant of the Bramble-Hilbert lemma then implies the error estimates,

$$\|E_{hk}v\|_{K^n} \le C\big(h_K^r\|\nabla^r v\|_{\tilde{K}^n} + k_n\|\dot{v}\|_{\tilde{K}^n}\big), \quad r = 1,\, 2, \tag{5.30}$$

$$\|\nabla E_{hk}v\|_{K^n} \le C\big(h_K^2\|\nabla^2 v\|_{\tilde{K}^n} + k_n\|\nabla\dot{v}\|_{\tilde{K}^n}\big), \tag{5.31}$$

where $\tilde{K}$ is a patch of space cells suitably chosen around K.

We recall the mesh function $\bar{h}_n$ from (3.2), and we define the notation

$$\bar{h}_{max,n} = \max_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K.$$

We will also use the fact that

$$\|v\|_{\Omega^n}^2 = \int_{I_n} \|v(t)\|^2\, dt \le k_n \max_{I_n} \|v(t)\|. \tag{5.32}$$

Theorem 7. *Let u and U be the solutions of* (2.3) *and* (3.5). *Then, with $e = U - u$, we have the weighted a posteriori error estimate*

(5.33)

$$\begin{aligned}
&\|e_1(T)\| \\
&\le C\Big\{\|h_0\big(U_1(0)-u^0\big)\| + \|h_0^2\big(U_2(0)-v^0\big)\|\Big\} \\
&\quad + C\sum_{n=1}^{N} k_n^{1/2}\bigg\{\|\bar{h}_n(\dot{U}_1 - U_2)\|_{\Omega^n} + k_n\|\dot{U}_1 - U_2\|_{\Omega^n} \\
&\quad + \|\bar{h}_n^2(\rho\dot{U}_2 - f)\|_{\Omega^n} + k_n\|\rho\dot{U}_2 - f\|_{\Omega^n} \\
&\quad + \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3\|r_h\|_{\partial K^n}^2\Big)^{1/2} + \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3\|g_h - g\|_{\partial K^n}^2\Big)^{1/2} \\
&\quad + k_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-1}\|r_h\|_{\partial K^n}^2\Big)^{1/2} + k_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-1}\|g_h - g\|_{\partial K^n}^2\Big)^{1/2} \\
&\quad + k_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K\|r_h\|_{\partial K^n}^2\Big)^{1/2} + k_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K\|g_h - g\|_{\partial K^n}^2\Big)^{1/2} \\
&\quad + \big(\bar{h}_{max,n}^2 + k_n\big)\bigg(\int_{I_n}\bigg(\sum_{j=1}^{n}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j} \bar{h}_K^{-1}\|(\beta * r_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\bigg)^2 dt\bigg)^{1/2} \\
&\quad + \big(\bar{h}_{max,n}^2 + k_n\big)\bigg(\int_{I_n}\bigg(\sum_{j=1}^{n}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j} \bar{h}_K^{-1}\|(\beta * g_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\bigg)^2 dt\bigg)^{1/2} \\
&\quad + \big(\bar{h}_{max,n} + k_n\big) \\
&\qquad\times\bigg(\int_{I_n}\bigg(\sum_{j=1}^{n}\bar{h}_{max,j}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j} \bar{h}_K^{-1}\|(\beta * r_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\bigg)^2 dt\bigg)^{1/2} \\
&\quad + \big(\bar{h}_{max,n} + k_n\big) \\
&\qquad\times\bigg(\int_{I_n}\bigg(\sum_{j=1}^{n}\bar{h}_{max,j}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j} \bar{h}_K^{-1}\|(\beta * g_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\bigg)^2 dt\bigg)^{1/2}\bigg\}.
\end{aligned}$$

Proof. We write the error representation (5.1) as

$$J(e)=\sum_{K\in\mathcal{T}_h^0}\Theta_{0,K}+\sum_{i=1}^{6}\sum_{n=1}^{N}\sum_{K\in\bar{\mathcal{T}}_h^n}\Theta_{i,K}^n=I_0+\sum_{i=1}^{6}I_i. \tag{5.34}$$

First we estimate I_0. To this end, recalling $\Theta_{0,K}$ from (5.2), we use the Cauchy-Schwarz inequality and the interpolation error estimate (5.30) to obtain,

$$\begin{aligned}\sum_{K\in\mathcal{T}_h^0}\big(U_1(0)-u^0,E_{hk}z_1(0)\big)_K&\le\sum_{K\in\mathcal{T}_h^0}\|U_1(0)-u^0\|_K\|E_{hk}z_1(0)\|_K\\&\le C\sum_{K\in\mathcal{T}_h^0}\|U_1(0)-u^0\|_K h_K\|\nabla z_1(0)\|_K\\&\le C\|\nabla z_1(0)\|\Big(\sum_{K\in\mathcal{T}_h^0}h_K^2\|U_1(0)-u^0\|_K^2\Big)^{1/2}\\&=C\|\nabla z_1(0)\|\|h_0\big(U_1(0)-u^0\big)\|.\end{aligned}$$

Similarly we have

$$\sum_{K\in\mathcal{T}_h^0}\rho\big(U_2(0)-v^0,E_{hk}z_2(0)\big)_K\le C\|\nabla^2z_2(0)\|\|h_0^2\big(U_2(0)-v^0\big)\|.$$

From these two estimates we conclude

$$\begin{aligned}I_0&\le C\max(\|\nabla z_1(0)\|,\|\nabla^2z_2(0)\|)\\&\quad\times\big\{\|h_0\big(U_1(0)-u^0\big)\|+\|h_0^2\big(U_2(0)-v^0\big)\|\big\}.\end{aligned} \tag{5.35}$$

For the next term, using the Cauchy-Schwarz inequality and the error estimate (5.30), we have

$$\begin{aligned}I_1&\le\sum_{n=1}^{N}\sum_{K\in\bar{\mathcal{T}}_h^n}\|\dot U_1-U_2\|_{K^n}\|E_{hk}z_1\|_{K^n}\\&\le C\sum_{n=1}^{N}\sum_{K\in\bar{\mathcal{T}}_h^n}\|\dot U_1-U_2\|_{K^n}\big(\bar h_K\|\nabla z_1\|_{\tilde K^n}+k_n\|\dot z_1\|_{\tilde K^n}\big)\\&\le C\sum_{n=1}^{N}\Big(\sum_{K\in\bar{\mathcal{T}}_h^n}\bar h_K^2\|\dot U_1-U_2\|_{K^n}^2\Big)^{1/2}\Big(\sum_{K\in\bar{\mathcal{T}}_h^n}\|\nabla z_1\|_{\tilde K^n}^2\Big)^{1/2}\\&\quad+C\sum_{n=1}^{N}k_n\Big(\sum_{K\in\bar{\mathcal{T}}_h^n}\|\dot U_1-U_2\|_{K^n}^2\Big)^{1/2}\Big(\sum_{K\in\bar{\mathcal{T}}_h^n}\|\dot z_1\|_{\tilde K^n}^2\Big)^{1/2}\\&=C\sum_{n=1}^{N}\Big\{\|\bar h_n(\dot U_1-U_2)\|_{\Omega^n}\|\nabla z_1\|_{\Omega^n}+k_n\|\dot U_1-U_2\|_{\Omega^n}\|\dot z_1\|_{\Omega^n}\Big\},\end{aligned}$$

that using (5.32) we have

$$\begin{aligned}I_1&\le C\max\Big(\max_{[0,T]}\|\nabla z_1(t)\|,\max_{[0,T]}\|\dot z_1(t)\|\Big)\\&\quad\times\sum_{n=1}^{N}k_n^{1/2}\big\{\|\bar h_n(\dot U_1-U_2)\|_{\Omega^n}+k_n\|\dot U_1-U_2\|_{\Omega^n}\big\}.\end{aligned} \tag{5.36}$$

In the same way we obtain

$$
\begin{aligned}
I_2 \leq C \max\Big(\max_{[0,T]} \|\nabla^2 z_2(t)\|, \max_{[0,T]} \|\dot{z}_2(t)\|\Big) \\
\times \sum_{n=1}^{N} k_n^{1/2} \big\{ \|\bar{h}_n^2(\rho \dot{U}_2 - f)\|_{\Omega^n} + k_n \|\rho \dot{U}_2 - f\|_{\Omega^n}\big\}.
\end{aligned} \tag{5.37}
$$

Now for I_3, we use the Cauchy-Schwarz inequality, a trace inequality, and the error estimates (5.30), (5.31) to obtain,

$$
\begin{aligned}
I_3 &\leq \sum_{n=1}^{N} \sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-1/2} \|r_h\|_{\partial K^n} \bar{h}_K^{1/2} \|E_{hk} z_2\|_{\partial K^n} \\
&\leq C \sum_{n=1}^{N} \sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-1/2} \|r_h\|_{\partial K^n} \big\{ \|E_{hk} z_2\|_{K^n} + \bar{h}_K \|\nabla E_{hk} z_1\|_{K^n} \big\} \\
&\leq C \sum_{n=1}^{N} \sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-1/2} \|r_h\|_{\partial K^n} \\
&\quad \times \big\{ 2\bar{h}_K^2 \|\nabla^2 z_2\|_{\tilde{K}^n} + k_n \|\dot{z}_2\|_{\tilde{K}^n} + \bar{h}_K k_n \|\nabla \dot{z}_2\|_{\tilde{K}^n} \big\} \\
&\leq C \sum_{n=1}^{N} \bigg\{ \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h\|_{\partial K^n}^2 \Big)^{1/2} \|\nabla^2 z_2\|_{\Omega^n} \\
&\qquad + k_n \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-1} \|r_h\|_{\partial K^n}^2 \Big)^{1/2} \|\dot{z}_2\|_{\Omega^n} \\
&\qquad + k_n \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K \|r_h\|_{\partial K^n}^2 \Big)^{1/2} \|\nabla \dot{z}_2\|_{\Omega^n} \bigg\},
\end{aligned}
$$

that using (5.32) we have

$$
\begin{aligned}
I_3 \leq C \max\Big(\max_{[0,T]} \|\nabla^2 z_2(t)\|, \max_{[0,T]} \|\dot{z}_2(t)\|, \max_{[0,T]} \|\nabla \dot{z}_2(t)\|\Big) \\
\times \sum_{n=1}^{N} k_n^{1/2} \bigg\{ \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|r_h\|_{\partial K^n}^2 \Big)^{1/2} + k_n \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K^{-1} \|r_h\|_{\partial K^n}^2 \Big)^{1/2} \\
+ k_n \Big(\sum_{K \in \bar{\mathcal{T}}_h^n} \bar{h}_K \|r_h\|_{\partial K^n}^2 \Big)^{1/2} \bigg\}.
\end{aligned} \tag{5.38}
$$

And similarly

$$
\begin{aligned}
I_4 \leq C \max\Big(\max_{[0,T]} \|\nabla^2 z_2(t)\|, \max_{[0,T]} \|\dot{z}_2(t)\|, \max_{[0,T]} \|\nabla \dot{z}_2(t)\|\Big) \\
\times \sum_{n=1}^{N} k_n^{1/2} \Big\{ \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^3 \|g_h - g\|_{\partial K^n}^2 \Big)^{1/2} \\
+ k_n \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K^{-1} \|g_h - g\|_{\partial K^n}^2 \Big)^{1/2} \\
+ k_n \Big(\sum_{K\in\bar{\mathcal{T}}_h^n} \bar{h}_K \|g_h - g\|_{\partial K^n}^2 \Big)^{1/2} \Big\}.
\end{aligned}
\tag{5.39}
$$

Finally we study I_5, I_6 which include the convolution terms. We find an estimate for I_5 and a similar argument holds for I_6. First, recalling the definition of $\Theta_{5,K}^n$ from (5.2), we can write I_5 as,

$$
I_5 = \sum_{n=1}^{N} \sum_{K\in\bar{\mathcal{T}}_h^n} \Theta_{5,K}^n = \sum_{n=1}^{N} \int_{I_n} \int_s^T \beta(t-s) \sum_{K\in\bar{\mathcal{T}}_h^n} \big(r_h(s), E_{hk} z_2(t)\big)_{\partial K} \, dt \, ds.
$$

Then we change the order of the time integrals and we obtain,

$$
\begin{aligned}
I_5 &= \sum_{n=1}^{N} \int_{I_n} \sum_{j=1}^{n} \int_{t_{j-1}}^{t\wedge t_j} \beta(t-s) \sum_{K\in\bar{\mathcal{T}}_h^j} \big(r_h(s), E_{hk} z_2(t)\big)_{\partial K} \, ds \, dt \\
&= \sum_{n=1}^{N} \int_{I_n} \sum_{j=1}^{n} \sum_{K\in\bar{\mathcal{T}}_h^j} \big((\beta * r_h)_j(t), E_{hk} z_2(t)\big)_{\partial K} \, dt,
\end{aligned}
$$

where

$$
(\beta * v)_j(t) = \int_{t_{j-1}}^{t\wedge t_j} \beta(t-s) v(s) \, ds.
$$

Now by a trace inequality and then Cauchy-Schwarz inequality in the sum over the triangles, and further Cuachy-Schwarz inequality in the integral over I_n, we have,

$$
\begin{aligned}
I_5 &\le C\sum_{n=1}^{N}\int_{I_n}\sum_{j=1}^{n}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j}\bar{h}_K^{-1}\|(\beta*r_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\\
&\qquad\times\big\{\|E_{hk}z_2(t)\|+\bar{h}_{max,j}\|\nabla E_{hk}z_2(t)\|\big\}\,dt\\
&= C\sum_{n=1}^{N}\int_{I_n}\|E_{hk}z_2(t)\|\sum_{j=1}^{n}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j}\bar{h}_K^{-1}\|(\beta*r_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\,dt\\
&\quad+C\sum_{n=1}^{N}\int_{I_n}\|\nabla E_{hk}z_2(t)\|\sum_{j=1}^{n}\bar{h}_{max,j}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j}\bar{h}_K^{-1}\|(\beta*r_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\,dt\\
&\le C\sum_{n=1}^{N}\Big(\int_{I_n}\|E_{hk}z_2(t)\|^2\,dt\Big)^{1/2}\\
&\qquad\times\bigg(\int_{I_n}\bigg(\sum_{j=1}^{n}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j}\bar{h}_K^{-1}\|(\beta*r_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\bigg)^2\,dt\bigg)^{1/2}\\
&\quad+C\sum_{n=1}^{N}\Big(\int_{I_n}\|\nabla E_{hk}z_2(t)\|^2\,dt\Big)^{1/2}\\
&\qquad\times\bigg(\int_{I_n}\bigg(\sum_{j=1}^{n}\bar{h}_{max,j}\Big(\sum_{K\in\bar{\mathcal{T}}_h^j}\bar{h}_K^{-1}\|(\beta*r_h)_j(t)\|_{\partial K}^2\Big)^{1/2}\bigg)^2\,dt\bigg)^{1/2}.
\end{aligned}
$$

Since by the error estimate (5.30) we have

$$
\begin{aligned}
\int_{I_n}\|E_{hk}z_2(t)\|^2\,dt &= \int_{I_n}\sum_{K\in\bar{\mathcal{T}}_h^n}\|E_{hk}z_2(t)\|_K^2\,dt=\sum_{K\in\bar{\mathcal{T}}_h^n}\|E_{hk}z_2(t)\|_{K^n}^2\\
&\le C\sum_{K\in\bar{\mathcal{T}}_h^n}\big(\bar{h}_K^4\|\nabla^2 z_2\|_{\tilde{K}^n}^2+k_n^2\|\dot{z}_2\|_{\tilde{K}^n}^2\big)\\
&\le C\big(\bar{h}_{max,n}^4\|\nabla^2 z_2\|_{\Omega^n}^2+k_n^2\|\dot{z}_2\|_{\Omega^n}^2\big),
\end{aligned}
$$

and similarly by (5.31)

$$
\begin{aligned}
\int_{I_n}\|\nabla E_{hk}z_2(t)\|^2\,dt &= \int_{I_n}\sum_{K\in\bar{\mathcal{T}}_h^n}\|\nabla E_{hk}z_2(t)\|_K^2\,dt=\sum_{K\in\bar{\mathcal{T}}_h^n}\|\nabla E_{hk}z_2(t)\|_{K^n}^2\\
&\le C\sum_{K\in\bar{\mathcal{T}}_h^n}\big(\bar{h}_K^2\|\nabla^2 z_2\|_{\tilde{K}^n}^2+k_n^2\|\nabla\dot{z}_2\|_{\tilde{K}^n}^2\big)\\
&\le C\big(\bar{h}_{max,n}^2\|\nabla^2 z_2\|_{\Omega^n}^2+k_n^2\|\nabla\dot{z}_2\|_{\Omega^n}^2\big),
\end{aligned}
$$

then, recalling (5.32), we conclude the estimate

$$
\begin{aligned}
I_5 \leq C \max \Big(\max_{[0,T]} \|\nabla^2 z_2(t)\|, \max_{[0,T]} \|\dot{z}_2(t)\|, \max_{[0,T]} \|\nabla \dot{z}_2(t)\| \Big) \\
\times \Bigg\{ \sum_{n=1}^{N} k_n^{1/2} \big(\bar{h}_{max,n}^2 + k_n \big) \\
\times \Bigg(\int_{I_n} \Bigg(\sum_{j=1}^{n} \Big(\sum_{K \in \bar{\mathcal{T}}_h^j} \bar{h}_K^{-1} \|(\beta * r_h)_j(t)\|_{\partial K}^2 \Big)^{1/2} \Bigg)^2 dt \Bigg)^{1/2} \\
+ C \sum_{n=1}^{N} k_n^{1/2} \big(\bar{h}_{max,n} + k_n \big) \\
\times \Bigg(\int_{I_n} \Bigg(\sum_{j=1}^{n} \bar{h}_{max,j} \Big(\sum_{K \in \bar{\mathcal{T}}_h^j} \bar{h}_K^{-1} \|(\beta * r_h)_j(t)\|_{\partial K}^2 \Big)^{1/2} \Bigg)^2 dt \Bigg)^{1/2} \Bigg\}.
\end{aligned}
\tag{5.40}
$$

The same estimate holds for I_6 with r_h be replaced by g_h.

We now set $j_1 = j_2 = z_2^T = 0$ and $z_1^T = A^{-1/2} e_1(T)$. Then, putting (5.35)-(5.40), and the counterpart of (5.40) for I_6, in (5.1), using the stability estimates (2.14) and a standard argument, we conclude the a posteriori error estimate (5.33), and this completes the proof. □

Remark 6. We can compute

$$
\begin{aligned}
\|r_h\|_{\partial K^n} &= \Big(\int_{I_n} \|\psi_{n-1}(t) r_h(t_{n-1}) + \psi_n(t) r_h(t_n)\|_{\partial K}^2 \, dt \Big)^{1/2} \\
&\leq \frac{k_n^{1/2}}{\sqrt{3}} \Big(\|r_h(t_{n-1})\|_{\partial K}^2 + \|r_h(t_n)\|_{\partial K}^2 \Big)^{1/2} \\
&\leq \sqrt{\frac{2}{3}} k_n^{1/2} \big(\|r_h(t_{n-1})\|_{\partial K} + \|r_h(t_n)\|_{\partial K} \big).
\end{aligned}
$$

Remark 7. We note that the last a posteriori error estimate presented in (5.33), does not have the restrictions that were mentioned in Remark 5.

References

1. K. Adolfsson, M. Enelund, and S. Larsson, *Space-time discretization of an integro-differential equation modeling quasi-static fractional-order viscoelasticity*, J. Vib. Control **14** (2008), 1631–1649.
2. K. Adolfsson, M. Enelund, S. Larsson, and M. Racheva, *Discretization of integro-differential equations modeling dynamic fractional order viscoelasticity*, LNCS **3743** (2006), 76–83.
3. W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, 2003.
4. M. Boman, *Estimates for the L_2-projection onto continuous finite element spaces in a weighted L_p-norm*, BIT **46** (2006), 249–260.
5. A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions. Vol. III*, Robert E. Krieger Publishing Co. Inc., 1981, Based on notes left by Harry Bateman, Reprint of the 1955 original.
6. R. H. Fabiano and K. Ito, *Semigroup theory and numerical approximation for equations in linear viscoelasticity*, SIAM J. Math. Anal. **21** (1990), 374–393.
7. S. Larsson and F. Saedpanah, *The continuous Galerkin method for an integro-differential equation modeling dynamic fractional order viscoelasticity*, IMA J. Numer. Anal., to appear.

8. W. McLean, V. Thomée, and L. B. Wahlbin, *Discretization with variable time steps of an evolution equation with a positive-type memory term*, J. of Comput. Appl. Math. **69** (1996), 49–69.
9. A. K. Pani, V. Thomée, and L. B. Wahlbin, *Numerical methods for hyperbolic and parabolic integro-differential equations*, J. Integral Equations Appl. **4** (1992), 533–584.
10. B. Riviére, S. Shaw, and J. R. Whiteman, *Discontinuous Galerkin finite element methods for dynamic linear solid viscoelasticity problems*, Numer. Methods Partial Differential Equations **23** (2007), 1149–1166.
11. F. Saedpanah, *Existence and uniqueness of the solution of an integro-differential equation with weakly singular kernel*, Preprint 2009:16, Dept. of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg.
12. S. Shaw and J. R. Whiteman, *Applications and numerical analysis of partial differential Volterra equations: a brief survey*, Comput. Methods Appl. Mech. Engrg. **150** (1997), 397–409.
13. ———, *A posteriori error estimates for space-time finite element approximation of quasistatic hereditary linear viscoelasticity problems*, Comput. Methods Appl. Mech. Engrg. **193** (2004), 5551–5572.

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Göteborg, Sweden
E-mail address: fardin@chalmers.se