Ritz-Volterra Reconstructions and A Posteriori Error Analysis of Finite Element Method for Parabolic Integro-Differential Equations

G Murali Mohan Reddy* Rajen K. Sinha†

February 4, 2013

Abstract

We derive a posteriori error estimates for both semidiscrete and implicit fully discrete backward Euler method for linear parabolic integro-differential equations in a bounded convex polygonal or polyhedral domain. A novel space-time reconstruction operator is introduced, which is a generalization of the elliptic reconstruction operator [SIAM J. Numer. Anal., 41(2003), no. 4, pp. 1585–1594], and we call it as Ritz-Volterra reconstruction operator. The Ritz-Volterra reconstruction operator in conjunction with the linear approximation of the Volterra integral term are used in a crucial way to derive optimal order a posteriori error estimates in $L^\infty(L^2)$ and $L^2(H^1)$-norms. The related a posteriori error estimates for the Ritz-Volterra reconstruction error are also established. We allow only nested refinement of the space meshes for the fully discrete analysis.

Keywords. Parabolic integro-differential equation; finite element method; semidiscrete, fully discrete; optimal a posteriori error estimate.

1 Introduction

In this paper, we address successfully the problem of obtaining a posteriori error estimates for both semidiscrete and fully discrete approximations to the solutions of the initial-boundary value problems for the linear parabolic integro-differential equations (PIDE) of the form

$$
\begin{aligned}
(1) \qquad u_t(x,t) + \mathcal{A}u(x,t) &= \int_0^t \mathcal{B}(t,s)u(x,s)ds + f(x,t), \quad (x,t) \in \Omega \times J, \\
u(x,t) &= 0, \quad (x,t) \in \partial\Omega \times \bar{J}, \\
u(x,0) &= u_0(x), \quad x \in \Omega.
\end{aligned}
$$

*Department of Mathematics, Indian Institute of Technology Guwahati, Guwahati - 781039, India (mmreddy@iitg.ernet.in).

†Corresponding author: Department of Mathematics, Indian Institute of Technology Guwahati, Guwahati - 781039, India (rajen@iitg.ernet.in).

Here, $\Omega \subset \mathbb{R}^d, d \geq 1$ is a bounded convex polygonal or polyhedral domain with boundary $\partial\Omega$, $J = (0, T]$ with $T < \infty$ and $u_t(x,t) = \frac{\partial u}{\partial t}(x,t)$. Further, $\mathcal{A}$ is a self-adjoint, uniformly positive definite second-order linear elliptic partial differential operator of the form

$$\mathcal{A} = -\nabla \cdot (A\nabla u),$$

and the operator $\mathcal{B}(t,s)$ is of the form

$$\mathcal{B}(t,s) = -\nabla \cdot (B(t,s)\nabla u),$$

where "∇" denotes the spatial gradient and the coefficients matrices A and $B(t,s)$ are assumed to be in $L^\infty(\Omega)^{d\times d}$. Moreover, the initial value $u_0 = u_0(x)$ and the nonhomogeneous term f are assumed to be smooth for our purpose.

Such problems and variants of them arise in various applications, such as heat conduction in material with memory [10], the compression of poro-viscoelasticity media [11], nuclear reactor dynamics [12] and the epidemic phenomena in biology [5]. The existence, uniqueness and regularity results for above problems can be found in [22] and references therein.

A priori error estimates for such kind of problems and their variants are quite rich in the literature. We refer to [4, 14, 16, 20, 22] for optimal order a priori error estimates for semidiscrete scheme and [23] for fully discrete scheme. Although a wide range of articles related to a priori error estimates are available, a posteriori error estimation of such kind of problems is still wide open.

A posteriori error estimation is the basis for efficient adaptive meshing procedures designed to control and minimize the error. Over the last two decades, a posteriori error analysis for the finite element methods for partial differential equations has been an area of active research [1,6–8,13,15]. While much of interest has focussed on elliptic and parabolic problems, relatively less progress has been made in the direction of a posteriori error analysis of PIDE. For an overview and summary of current research activities in the later area we refer to the articles [18,19]. In the absence of the memory term, i.e., when $\mathcal{B}(t,s) = 0$, a posteriori error analysis for linear parabolic problems have been investigated by several authors [2,6,8,13,15,21] in recent years. In [8], the authors have derived quasi-optimal error estimates in $L^\infty(L^2(\Omega))$-norm via duality technique. Subsequently, optimal order estimates in $L^2(H^1(\Omega))$ and suboptimal estimates in $L^\infty(L^2(\Omega))$ norms are derived in [21]. In [13,15], the authors have used the elliptic reconstruction in combination with energy techniques to derive optimal order a posteriori error estimates for the heat equation in $L^\infty(L^2(\Omega))$-norm. Since then several authors have considered the elliptic reconstruction operator as an analytical tool to derive a posteriori error estimates in various norms for linear and nonlinear parabolic problems [9,13,15].

In this paper, we have derived first optimal order a posteriori bounds for PIDE in $L^\infty(L^2(\Omega))$-norm for the semidiscrete case and in $L^\infty(L^2(\Omega))$ and $L^2(H^1(\Omega))$ norms of the error for the practically more relevant backward Euler fully discrete scheme. The proof of a posteriori bounds for the semidiscrete and fully discrete analysis necessitates the careful introduction of a novel space-time reconstruction operator and we call it as Ritz-Volterra reconstruction operator. This Ritz-Volterra reconstruction may be thought of as a generalization of the elliptic reconstruction introduced earlier by the authors in [13,15] for parabolic problems. An attempt has been made in this exposition to carry over a posteriori error analysis of parabolic problems to PIDE. Due to the presence of the Volterra integral term in

(1) such an extension is not straightforward. The complications arises because the Volterra integral term memorizes the jumps over all element edges in all previous space meshes. We first prove optimal order a posteriori error estimates for the Ritz-Volterra reconstruction error. Then, introducing Ritz-Volterra reconstruction operator as an intermediate solution we derive a posteriori error estimates. For fully discrete analysis, we allow only nested refinement of the space meshes.

We organize the paper as follows. In Section 2, we introduce some standard notations and preliminary materials to be used in the subsequent sections. The Ritz-Volterra reconstruction operator in the context of semidiscrete scheme is introduced in Section 3. Section 4 is devoted to the related a posteriori error estimates for the reconstruction error and a posteriori error estimate for the parabolic error. Further, optimal order a posteriori error estimate in $L^\infty(L^2)$ norm is derived for the semidiscrete scheme. Finally, optimal order a posteriori error estimates for the implicit backward Euler fully discrete scheme are established in Section 5.

2 Notations and preliminaries

Given a Lebesgue measurable set $\omega \subset \mathbb{R}^2$, we denote by $L^p(\omega), 1 \le p \le +\infty$, the Lebesgue spaces with corresponding norms $\|\cdot\|_{L^p(\omega)}$. When $p = 2$, the space $L^2(\omega)$ is equipped with inner product $\langle\cdot,\cdot\rangle_\omega$ and the induced norm $\|\cdot\|_{L^2(\omega)}$. Whenever $\omega = \Omega$, we remove the subscripts of $\|.\|_{L^2(\omega)}$ and $\langle\cdot,\cdot\rangle_\omega$. Further, we shall use the standard notation for Sobolev spaces $W^{m,p}(\omega)$ with $1 \le p \le +\infty$. The norm on $W^{m,p}(\omega)$ is defined by

$$\|v\|_{m,p,\omega} = \left(\int_\omega \sum_{|\alpha|\le m} |D^\alpha v|^p dx \right)^{1/p}, \ 1 \le p < \infty$$

with the standard modification for $p = \infty$. When $p = 2$, we write $W^{m,2}(\Omega)$ by $H^m(\Omega)$ and denote the norm by $\|\cdot\|_m$. In particular, $H_0^1(\Omega)$ signifies the space of functions in $H^1(\Omega)$ that vanish on the boundary of Ω (boundary values are taken in the sense of traces).

Let $a(\cdot,\cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ be the bilinear form corresponding to the elliptic operator $\mathcal{A}$ defined by

$$a(v,\psi) := \langle A\nabla v, \nabla\psi\rangle, \quad \forall\, v,\psi \in H_0^1(\Omega).$$

Similarly, let $b(t,s;\cdot,\cdot)$ be the bilinear form corresponding to the operator $\mathcal{B}(t,s)$ defined on $H_0^1(\Omega) \times H_0^1(\Omega)$ by

$$b(t,s;v(s),\psi) := \langle B(t,s)\nabla v(s), \nabla\psi\rangle, \quad \forall\, v(s),\psi \in H_0^1(\Omega).$$

Let $b_t(t,s;\cdot,\cdot)$ and $b_s(t,s;\cdot,\cdot)$ be the bilinear forms obtained by differentiating the coefficient of $b(t,s;\cdot,\cdot)$ with respect to t and s, respectively.

We assume that the bilinear form $a(\cdot,\cdot)$ is continuous and coercive on $H_0^1(\Omega)$ i.e.,

$$|a(\psi,\phi)| \le \alpha\|\psi\|_1\|\phi\|_1 \quad \text{and} \quad a(\phi,\phi) \ge \beta\|\phi\|_1^2, \qquad \forall\, \psi,\phi \in H_0^1(\Omega), \tag{2}$$

with $\alpha, \beta \in \mathbb{R}^+$. Further, we assume that the bilinear forms $b(t,s;\cdot,\cdot)$ and $b_s(t,s;\cdot,\cdot)$ are continuous on $H_0^1(\Omega)$ i.e.,

$$|b(t,s;\psi(s),\phi)| \le \gamma\|\psi(s)\|_1\|\phi\|_1, \quad \forall\, \psi(s),\phi \in H_0^1(\Omega), \tag{3}$$

$$|b_s(t,s;\psi(s),\phi)| \le \gamma'\|\psi(s)\|_1\|\phi\|_1, \quad \forall\, \psi(s),\phi \in H_0^1(\Omega), \tag{4}$$

with $\gamma, \gamma' \in \mathbb{R}^+$.

The weak formulation of the problem (1.1) may be stated as follows: Find $u : \bar{J} \to H_0^1(\Omega)$ such that

$$\begin{aligned} \langle u_t, \phi\rangle + a(u,\phi) &= \int_0^t b(t,s;u(s),\phi)ds + \langle f,\phi\rangle, \quad \forall \phi \in H_0^1(\Omega), \quad t \in J, \\ u(0) &= u_0. \end{aligned} \tag{5}$$

3 Semidiscrete finite element approximations and Ritz-Volterra reconstruction

Let $h(x) = \text{diam}(K)$, where $K \in \mathcal{T}_h$ and $x \in K$ denotes positive piecewise constant mesh-size function corresponds to $\mathcal{T}_h = \{K\}$, a shape regular, conforming triangulation of Ω. Let $\mathcal{E}_h = \{E\}$ be the set of internal sides of $\mathcal{T}_h$. These internal sides are edges in $d = 2$ and faces in $d = 3$. The union of all internal sides $\bigcup_{E\in\mathcal{E}_h} E$ be denoted as $\sum_h$.

We associate the following finite element space corresponding to $\mathcal{T}_h$:

$$S_h = \{\chi \in H_0^1(\Omega) : \chi|_K \in \mathbb{P}_k(K), \text{ for all } K \in \mathcal{T}_h\},$$

where $\mathbb{P}_k$ is the space of polynomials of degree $\leq k$ with $k \in \mathbb{Z}^+$. The semidiscrete finite element approximation $u_h : \bar{J} \to S_h$ of u is defined by

$$\begin{aligned} \langle u_{h,t}, \chi\rangle + a(u_h,\chi) &= \int_0^t b(t,s;u_h(s),\chi)ds + \langle f,\chi\rangle, \quad \forall \chi \in S_h, \\ u_h(.,0) &= P_h^0 u_0, \end{aligned} \tag{6}$$

where $P_h^0 u_0$ is the L^2-projection of u_0 onto S_h.

Representation of the bilinear forms. For a function $v \in S_h$, following [13], the bilinear form $a(u, v)$ can be represented as

$$a(v,\phi) = \sum_{K\in\mathcal{T}_h} \langle -\text{div}(A\nabla v), \phi\rangle_K + \sum_{E\in\mathcal{E}_h} \langle J_1[v], \phi\rangle_E, \quad \forall \phi \in H_0^1(\Omega),$$

where $J_1[v]$ is the spatial jump of the field $A\nabla v$ across an element side $E \in \mathcal{E}_h$ defined as

$$J_1[v]|_E(x) = [A\nabla v]_E(x) := \lim_{\varepsilon\to 0}(A\nabla v(x+\varepsilon\nu_E) - A\nabla v(x-\varepsilon\nu_E)).\nu_E, \tag{7}$$

where ν_E is a unit normal vector to E at the point x. For $v \in S_h$, let $\mathcal{A}_{el}v$ be the regular part of the distribution $-\text{div}(A\nabla v)$, which is defined as a piecewise continuous function such that

$$\langle \mathcal{A}_{el}v, \phi\rangle = \sum_{K\in\mathcal{T}_h} \langle -\text{div}(A\nabla v), \phi\rangle, \quad \forall \phi \in H_0^1(\Omega).$$

Thus, we can represent our bilinear form $a(\cdot,\cdot)$ as

$$a(v,\phi) = \langle \mathcal{A}_{el}v, \phi\rangle + \langle J_1[v], \phi\rangle_{\sum_h}, \quad \forall \phi \in H_0^1(\Omega). \tag{8}$$

Similarly, one can represent the bilinear form $b(t,s;\cdot,\cdot)$ as

$$b(t,s;v(s),\phi) = \langle \mathcal{B}_{el}(t,s)v(s), \phi\rangle + \langle J_2[v(s)], \phi\rangle_{\sum_h}, \quad \forall \phi \in H_0^1(\Omega), \tag{9}$$

where $\mathcal{B}_{el}(t,s)v(s)$ is the regular part of the distribution $-\text{div}(B(t,s)\nabla v(s))$, which is defined as a piecewise continuous function such that

$$\langle \mathcal{B}_{el}(t,s)v(s), \phi\rangle = \sum_{K\in\mathcal{T}_h} \langle -\text{div}(B(t,s)\nabla v(s)), \phi\rangle, \quad \forall \phi \in H_0^1(\Omega), \tag{10}$$

and $J_2[v(s)]$ is the spatial jump of the field $-\text{div}(B(t,s)\nabla v(s))$ across an element side $E\in\mathcal{E}_h$ as defined in (7) with $B(t,s)$ replacing A.

For $s\in[0,t]$, following [18], we define the discrete operators $\mathcal{A}_h : H_0^1(\Omega)\to S_h$ and $\mathcal{B}_h(t,s) : H_0^1(\Omega)\to S_h$ by

$$\langle \mathcal{A}_h w, \chi\rangle = a(w,\chi) \quad \text{and} \quad \langle \mathcal{B}_h(t,s)w(s), \chi\rangle = b(t,s;w(s),\chi), \ \ \forall \chi\in S_h. \tag{11}$$

Recall from [14] the following Ritz-Volterra projection $W_h : \bar{J}\to S_h$ defined by

$$a(W_hu-u,\chi) = \int_0^t b(t,s;(W_hu-u)(s),\chi)ds, \quad \forall\chi\in S_h,\ t\in\bar{J}. \tag{12}$$

Definition 3.1 *(Elliptic reconstruction [13, 15]) For a given $v\in H_0^1(\Omega)$, we define the elliptic reconstruction operator $\mathcal{R} : [0,T]\to H_0^1(\Omega)$ associated with the bilinear form $a(\cdot,\cdot)$ and is given by*

$$a(\mathcal{R}v,\phi) = \langle \mathcal{A}_h v, \phi\rangle, \quad \forall\phi\in H_0^1(\Omega).$$

The following definition generalize the concept of elliptic reconstruction and we call it as Ritz-Volterra reconstruction.

Definition 3.2 *(Ritz-Volterra reconstruction) Define the time-dependent Ritz-Volterra reconstruction $R_w : [0,T]\to H_0^1$, which plays a crucial role in our error analysis, by*

$$a(R_wv,\phi) - \int_0^t b(t,s;R_wv(s),\phi)ds = \langle \mathcal{A}_h v,\phi\rangle - \int_0^t \langle \mathcal{B}_h(t,s)v(s),\phi\rangle ds, \tag{13}$$

for all $\phi\in H_0^1(\Omega)$.

The function R_wv is referred to as the Ritz-Volterra reconstruction of v. Note that in the absence of the memory term this definition is equivalent to the definition of the elliptic reconstruction operator above. Although, similar to the elliptic reconstruction we define the domain of definition of Ritz-Volterra reconstruction to be $H_0^1(\Omega)$ but we will use it effectively on the finite element spaces only. The wellposedness of the Ritz-Volterra reconstruction operator follows analogously to that of the elliptic reconstruction operator [15].

Remark. For $t\in\bar{J}$, an important property of the Ritz-Volterra reconstruction operator R_w is that for $v\in H_0^1(\Omega)$, $v-R_wv$ is orthogonal to S_h with respect to $a(\cdot,\cdot)-\int_0^t b(t,s;\cdot,\cdot)ds$, i.e.,

$$a(R_wv-v,\phi) - \int_0^t b(t,s;(R_wv-v)(s),\phi)ds = 0, \ \ \forall\phi\in S_h. \tag{14}$$

This property is known as Galerkin orthogonality and is important in the sense that it allows to obtain a posteriori error estimates.

4 Analysis for the semidiscrete scheme

In order to give a posteriori error bounds, we decompose the main error $e := u_h - u$ as follows:

$$e := \rho - \epsilon, \text{ where } \rho := R_w u_h - u, \quad \epsilon := R_w u_h - u_h.$$

Here, ϵ is referred to as the reconstruction error whereas the time approximation error information is conveyed by ρ, which will be referred to as the parabolic error. Unlike for the parabolic problem [13, 15], we don't have any a posteriori error estimators available in the literature to control the error ϵ. In this section, we first derive a posteriori error estimates for the reconstruction error (ϵ) which will then be used to obtain a posteriori error estimates for the spatially semidiscrete Galerkin approximations to the problem (1).

We now recall from [17] the following interpolation error estimates.

Lemma 4.1 *([17]) Let $\Pi_h : H_0^1(\Omega) \to S_h$ be the Clément-type interpolation operator. Then, for sufficiently smooth ψ and finite element polynomial space of degree l, there exist constants $C_{1,j}$ and $C_{2,j}$ depending only upon the shape-regularity of the family of triangulations such that for $j \leq l+1$*

$$\|h^{-j}(\psi - \Pi_h \psi)\| \leq C_{1,j} \|\psi\|_j,$$

and

$$\|h^{1/2-j}(\psi - \Pi_h \psi)\|_{\sum_h} \leq C_{2,j} \|\psi\|_j.$$

We shall use traditional residual type a posteriori error estimators.

Residual. Using the definitions of the discrete operators $\mathcal{A}_h$ and $\mathcal{B}_h(t,s)$ and the distributional form of semidiscrete equation (6), we have

$$\mathcal{A}_h u_h - \int_0^t \mathcal{B}_h(t,s) u_h(s) ds - \mathcal{A}_{el} u_h + \int_0^t \mathcal{B}_{el}(t,s) u_h(s) ds = \mathfrak{R}[u_h] + (f_h - f),$$

where $\mathfrak{R}[u_h] = f - u_{h,t} - \mathcal{A}_{el} u_h + \int_0^t \mathcal{B}_{el}(t,s) u_h(s) ds$ are the inner residuals and $f_h = P_h^0 f$. Further, we define

$$\mathfrak{J}[u_h] = J_1[u_h] - \int_0^t J_2[u_h(s)] ds$$

as the jump residuals.

Below, we shall derive a posteriori error estimates for Ritz-Volterra reconstruction error.

Lemma 4.2 *(Ritz-Volterra reconstruction error estimates)*
For any $v \in S_h$, the following estimates holds true:

$$\begin{aligned}
&\|(R_w v - v)(t)\|_1 \\
&\leq C_1 h \|\mathcal{A}_h v - \mathcal{A}_{el} v - \int_0^t \mathcal{B}_h(t,s) v(s) ds + \int_0^t \mathcal{B}_{el}(t,s) v(s) ds\| \\
&+ C_2 h^{1/2} \|J_1[v] - \int_0^t J_2[v(s)] ds\|_{\sum_h}
\end{aligned}$$

and

$$\begin{aligned}
&\|(R_w v - v)(t)\| \\
&\leq C_3 h^2 \|\mathcal{A}_h v - \mathcal{A}_{el} v - \int_0^t \mathcal{B}_h(t,s)v(s)ds + \int_0^t \mathcal{B}_{el}(t,s)v(s)ds\| \\
&+C_4 h^{3/2} \|J_1[v] - \int_0^t J_2[v(s)]ds\|_{\sum_h},
\end{aligned}$$

where $C_j, j = 1,2,3,4$ *are positive constants independent of the discretization parameter but depending upon the shape-regularity of the family of triangulations and the final time* T.

Proof. For all $\phi \in H_0^1(\Omega)$, using (8), (9) and (13) we have

$$\begin{aligned}
a(R_w v - v, \phi) \quad &- \quad \int_0^t b(t,s;(R_w v - v)(s), \phi)ds \\
&= \quad \langle \mathcal{A}_h v, \phi\rangle - \int_0^t \langle \mathcal{B}_h(t,s)v(s)ds, \phi\rangle - a(v,\phi) + \int_0^t b(t,s;v(s),\phi)ds \\
&= \quad \langle \mathcal{A}_h v - \int_0^t \mathcal{B}_h(t,s)v(s)ds - \mathcal{A}_{el} v + \int_0^t \mathcal{B}_{el}(t,s)v(s)ds, \phi\rangle \\
&\quad -\langle J_1[v] - \int_0^t J_2[v(s)]ds, \phi\rangle_{\sum_h}.
\end{aligned}$$

An application of the Galerkin orthogonality (14) yields

$$\begin{aligned}
a(R_w v - v, \phi) \quad &- \quad \int_0^t b(t,s;(R_w v - v)(s), \phi)ds \\
&= \quad \langle \mathcal{A}_h v - \int_0^t \mathcal{B}_h(t,s)v(s)ds - \mathcal{A}_{el} v + \int_0^t \mathcal{B}_{el}(t,s)v(s)ds, \phi - \Pi_h\phi\rangle \\
&\quad -\langle J_1[v] - \int_0^t J_2[v(s)]ds, \phi - \Pi_h\phi\rangle_{\sum_h}.
\end{aligned}$$

Now using Lemma 4.1 with $C_{i,1}, i = 1,2$ as interpolation constants, we obtain

$$\begin{aligned}
&|a(R_w v - v, \phi)| \\
&\leq C_{1,1} h \|\phi\|_1 \|\mathcal{A}_h v - \mathcal{A}_{el} v - \int_0^t \mathcal{B}_h(t,s)v(s)ds + \int_0^t \mathcal{B}_{el}(t,s)v(s)ds\| \\
&+C_{2,1} h^{1/2} \|\phi\|_1 \|J_1[v] - \int_0^t J_2[v(s)]ds\|_{\sum_h} + \int_0^t |b(t,s;(R_w v - v)(s), \phi)|ds.
\end{aligned}$$

Taking $\phi = R_w v - v$ and using (3), we have

$$\begin{aligned}
&\quad |a(R_w v - v, R_w v - v)| \\
&\leq \quad \|R_w v - v\|_1 \Big\{ C_{1,1} h \|\mathcal{A}_h v - \mathcal{A}_{el} v - \int_0^t \mathcal{B}_h(t,s)v(s)ds + \int_0^t \mathcal{B}_{el}(t,s)v(s)ds\| \\
&\quad +C_{2,1} h^{1/2} \|J_1[v] - \int_0^t J_2[v(s)]ds\|_{\sum_h} + \gamma \int_0^t \|(R_w v - v)(s)\|_1 ds \Big\}.
\end{aligned}$$

Now, coercivity property of $a(\cdot,\cdot)$ and an application of the Gronwall's lemma yield the first inequality with $C_i = C_{1,G}(T)C_{i,1}/\beta, i = 1,2$, where $C_{1,G}$ is a constant appear due to Gronwall's lemma.

The proof of L^2 error estimate will proceed by the duality technique. For $v \in S_h$, let $\psi \in H^2(\Omega) \cap H_0^1(\Omega)$ be the solution of

$$\begin{aligned} \mathcal{A}\psi &= R_w v - v \quad \text{in } \Omega, \\ \psi &= 0 \quad \text{on } \Omega, \end{aligned} \tag{15}$$

satisfying the following regularity estimate (Ω is convex) with the constant C_Ω depending on the domain Ω:

$$\|\psi\|_2 \leq C_\Omega \|R_w v - v\|. \tag{16}$$

Multiplying (15) by $R_w v - v$ and integrating over Ω and using Galerkin orthogonality (14), we obtain

$$\begin{aligned} \|R_w v - v\|^2 &= a(R_w v - v, \psi - \Pi_h \psi) + a(R_w v - v, \Pi_h \psi) \\ &= a(R_w v - v, \psi - \Pi_h \psi) - \int_0^t b(t, s; (R_w v - v)(s), \psi - \Pi_h \psi) ds \\ &\quad + \int_0^t b(t, s; (R_w v - v)(s), \psi) ds \\ &= \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3. \end{aligned}$$

Using (8), (9) and (13), we arrive at

$$\begin{aligned} \mathcal{I}_1 + \mathcal{I}_2 &= \langle \mathcal{A}_h v - \mathcal{A}_{el} v - \int_0^t \mathcal{B}_h(t, s) v(s) ds + \int_0^t \mathcal{B}_{el}(t, s) v(s) ds, \psi - \Pi_h \psi \rangle \\ &\quad - \langle J_1[v] - \int_0^t J_2[v(s)] ds, \psi - \Pi_h \psi \rangle_{\sum_h}. \end{aligned}$$

Now, using the fact

$$b(t, s; (R_w v - v)(s), \psi) := \langle (R_w v - v)(s), \mathcal{B}^*(t, s) \psi \rangle, \tag{17}$$

where $\mathcal{B}^*(t, s)$ is the formal adjoint of the operator $\mathcal{B}(t, s)$ and $\|\mathcal{B}^*(t, s)\psi\| \leq C_{\mathcal{B}_1^*} \|\psi\|_2$, we obtain

$$|\mathcal{I}_3| \leq C_{\mathcal{B}_1^*} \|\psi\|_2 \int_0^t \|(R_w v - v)(s)\| ds.$$

The above bounds on $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_3$, and an application of Lemma 4.1 with the interpolation constants as $C_{i,2}, i = 1, 2$, yields

$$\begin{aligned} \|R_w v - v\|^2 &\leq \|\psi\|_2 \Big\{ C_{1,2} h^2 \|\mathcal{A}_h v - \mathcal{A}_{el} v - \int_0^t \mathcal{B}_h(t, s) v(s) ds \\ &\quad + \int_0^t \mathcal{B}_{el}(t, s) v(s) ds\| + C_{2,2} h^{3/2} \|J_1[v] - \int_0^t J_2[v(s)] ds\|_{\sum_h} \\ &\quad + C_{\mathcal{B}_1^*} \int_0^t \|(R_w v - v)(s)\| ds \Big\}. \end{aligned}$$

And hence, with an aid of (16), we have

$$\begin{aligned} &\|R_w v - v\| \leq C_{1,2} C_\Omega h^2 \|\mathcal{A}_h v - \mathcal{A}_{el} v - \int_0^t \mathcal{B}_h(t, s) v(s) ds + \int_0^t \mathcal{B}_{el}(t, s) v(s) ds\| \\ &+ C_{2,2} C_\Omega h^{3/2} \|J_1[v] - \int_0^t J_2[v(s)] ds\|_{\sum_h} + C_\Omega C_{\mathcal{B}_1^*} \int_0^t \|(R_w v - v)(s)\| ds. \end{aligned}$$

Finally, an application of the Gronwall's lemma yield the desired estimate with $C_3 = C_{2,G}(T)C_{1,2}C_\Omega$ and $C_4 = C_{2,G}(T)C_{2,2}C_\Omega$, where $C_{2,G}$ is a constant appear due to Gronwall's lemma.

We now define the following error estimators.

$$\Theta_0(u_h(t)) := C_3h^2\|\mathfrak{R}[u_h(t)]\| + C_4h^{3/2}\|\mathfrak{J}[u_h(t)]\|, \tag{18}$$

$$Osc(g(t)) := \|g_h(t) - g(t)\|. \tag{19}$$

Here, (19) denotes the oscillations of g in L^2-norm where $g_h(t) = P_h^0 g(t)$.

The following lemma yields a bound for the time derivative of the reconstruction error.

Lemma 4.3 *For any $v \in S_h$, the following bound holds true in terms of the reconstruction error:*

$$\begin{aligned}\|(R_wv - v)_t\| \quad &\leq \quad C_5h^2\|\frac{d}{dt}\Big\{\mathcal{A}_hv - \mathcal{A}_{el}v - \int_0^t \mathcal{B}_h(t,s)v(s)ds + \int_0^t \mathcal{B}_{el}(t,s)v(s)ds\Big\}\| \\ &\quad +C_6h^{3/2}\|\frac{d}{dt}\Big\{J_1[v] - \int_0^t J_2[v(s)]ds\Big\}\|_{\sum_h} + C_7\int_0^t \|(R_wv - v)(s)\|ds \\ &\quad +C_8\|R_wv - v\|.\end{aligned}$$

In particular, for finite element solution u_h, the following a posteriori error bound holds true:

$$\begin{aligned}&\|(R_wu_h - u_h)_t\| \\ \leq\ & C_5h^2\|\frac{d}{dt}\Big\{\mathcal{A}_hu_h - \mathcal{A}_{el}u_h - \int_0^t \mathcal{B}_h(t,s)u_h(s)ds + \int_0^t \mathcal{B}_{el}(t,s)u_h(s)ds\Big\}\| \\ &+C_6h^{3/2}\|\frac{d}{dt}\Big\{J_1[u_h] - \int_0^t J_2[u_h(s)]ds\Big\}\|_{\sum_h} + C_7\int_0^t \Theta_0(u_h)(s)ds \\ &+C_7C_3\int_0^t h^2\ Osc(f(s))ds + C_8\Theta_0(u_h) + C_8C_3h^2\ Osc(f(t)),\end{aligned}$$

where $C_j, j = 5, 6, 7, 8$ are the positive constants independent of the discretization parameter but depending upon the shape-regularity of the family of triangulations and the final time T.

Proof. Differentiating (14) with respect to t, for all $\phi \in S_h$, we have

$$\begin{aligned}a((R_wv - v)_t, \phi) \quad &- \quad b(t,t;(R_wv - v)(t), \phi) \\ &- \quad \int_0^t b_t(t,s;(R_wv - v)(s), \phi)ds = 0.\end{aligned} \tag{20}$$

Consider the dual elliptic problem with the forcing function to be $(R_wv - v)_t$. For $v \in S_h$, let $\psi \in H^2(\Omega) \cap H_0^1(\Omega)$, be the solution of

$$\begin{aligned}\mathcal{A}\psi \quad &= \quad (R_wv - v)_t \quad \text{in } \Omega, \\ \psi \quad &= \quad 0 \quad \text{on } \Omega,\end{aligned} \tag{21}$$

satisfying the following regularity estimate (Ω is convex) with the constant $\bar{C}_\Omega$ depending on the domain Ω:

$$\|\psi\|_2 \leq \bar{C}_\Omega\|(R_wv - v)_t\|. \tag{22}$$

We first multiply (21) by $(R_wv-v)_t$ and integrate over Ω. Then, rearranging terms and using (20), we obtain

$$
\begin{aligned}
(23)\qquad & \|(R_wv-v)_t\|^2 \\
= \;& a((R_wv-v)_t,\psi-\Pi_h\psi)-\int_0^t b_t(t,s;(R_wv-v)(s),\psi-\Pi_h\psi)ds \\
& -b(t,t;(R_wv-v)(t),\psi-\Pi_h\psi)+\int_0^t b_t(t,s;(R_wv-v)(s),\psi)ds \\
& +b(t,t;(R_wv-v)(t),\psi):=\mathcal{J}_1+\mathcal{J}_2+\mathcal{J}_3+\mathcal{J}_4+\mathcal{J}_5.
\end{aligned}
$$

In order to handle the first three terms, arguing analogously as in the proof for the second inequality in Lemma 4.2, we have

$$
\begin{aligned}
& a(R_wv-v,\psi-\Pi_h\psi)-\int_0^t b(t,s;(R_wv-v)(s),\psi-\Pi_h\psi)ds \\
= \;& \langle \mathcal{A}_hv-\int_0^t \mathcal{B}_h(t,s)v(s)ds-\mathcal{A}_{el}v+\int_0^t \mathcal{B}_{el}(t,s)v(s)ds,\psi-\Pi_h\psi\rangle \\
& -\langle J_1[v]-\int_0^t J_2[v(s)]ds,\psi-\Pi_h\psi\rangle_{\sum_h}.
\end{aligned}
$$

We differentiate both sides of the above equation with respect to t. Then, use of Cauchy-Schwarz inequality and Lemma 4.1 with the interpolation constants $C_{i,2}$, $i=1,2$ leads to

$$
\begin{aligned}
& |a((R_wv-v)_t,\psi-\Pi_h\psi)-\int_0^t b_t(t,s;(R_wv-v)(s),\psi-\Pi_h\psi)ds \\
& -b(t,t;(R_wv-v)(t),\psi-\Pi_h\psi)| \\
= \;& |\langle \frac{d}{dt}\Big\{\mathcal{A}_hv-\mathcal{A}_{el}v-\int_0^t \mathcal{B}_h(t,s)v(s)ds+\int_0^t \mathcal{B}_{el}(t,s)v(s)ds\Big\},\psi-\Pi_h\psi\rangle \\
& -\langle \frac{d}{dt}\Big\{J_1[v]-\int_0^t J_2[v(s)]ds\Big\},\psi-\Pi_h\psi\rangle_{\sum_h}| \\
\leq \;& \|\psi\|_2\Big\{C_{1,2}h^2\|\frac{d}{dt}\Big\{\mathcal{A}_hv-\mathcal{A}_{el}v-\int_0^t \mathcal{B}_h(t,s)v(s)ds+\int_0^t \mathcal{B}_{el}(t,s)v(s)ds\Big\}\| \\
& +C_{2,2}h^{3/2}\|\frac{d}{dt}\Big\{J_1[v]-\int_0^t J_2[v(s)]ds\Big\}\|_{\sum_h}\Big\}.
\end{aligned}
$$

For the terms $\mathcal{J}_4$ and $\mathcal{J}_5$, use the fact

$$
b_t(t,s;(R_wv-v)(s),\psi) \quad := \quad \langle (R_wv-v)(s),\mathcal{B}_t^*(t,s)\psi\rangle,
$$

and (17) together with $\|\mathcal{B}_t^*(t,s)\psi\|\leq C_{\mathcal{B}_2^*}\|\psi\|_2$ and $\|\mathcal{B}^*(t,t)\psi\|\leq C_{\mathcal{B}_3^*}\|\psi\|_2$, where $\mathcal{B}_t^*(t,s)$ is obtained by differentiating the coefficient of the operator $\mathcal{B}^*(t,s)$ with respect to t, to obtain

$$
\begin{aligned}
& \|(R_wv-v)_t\|^2 \\
\leq \;& \|\psi\|_2\Big\{C_{1,2}h^2\|\frac{d}{dt}\Big\{\mathcal{A}_hv-\mathcal{A}_{el}v-\int_0^t \mathcal{B}_h(t,s)v(s)ds+\int_0^t \mathcal{B}_{el}(t,s)v(s)ds\Big\}\| \\
& +C_{2,2}h^{3/2}\|\frac{d}{dt}\Big\{J_1[v]-\int_0^t J_2[v(s)]ds\Big\}\|_{\sum_h} \\
& +C_{\mathcal{B}_2^*}\int_0^t \|(R_wv-v)(s)\|ds+C_{\mathcal{B}_3^*}\|R_wv-v\|\Big\}.
\end{aligned}
$$

Using (22), the desired estimate follows with constants $C_5 = C_{1,2}\bar{C}_\Omega$, $C_6 = C_{2,2}\bar{C}_\Omega$, $C_7 = C_{\mathcal{B}_2^*}\bar{C}_\Omega$ and $C_8 = C_{\mathcal{B}_3^*}\bar{C}_\Omega$. The second estimate follows immediately from Lemma 4.2, (18) and (19) with u_h replacing v.

The first two terms in the previous error estimate can be handled in the following way.

$$
\begin{aligned}
&\|\frac{d}{dt}\Big\{\mathcal{A}_h v - \mathcal{A}_{el} v - \int_0^t \mathcal{B}_h(t,s)v(s)ds + \int_0^t \mathcal{B}_{el}(t,s)v(s)ds\Big\}\| \\
&= \|\frac{d}{dt}[\mathfrak{R}[u_h] + (f_h - f)]\| = \|\mathfrak{R}_t[u_h]\| + Osc(f_t(t)),
\end{aligned}
$$

and

$$
\|\frac{d}{dt}\Big\{J_1[v] - \int_0^t J_2[v(s)]ds\Big\}\|_{\sum_h} = \|\mathfrak{J}_t[u_h]\|_{\sum_h}.
$$

Now, we define the estimator for the time derivative of the reconstruction error by

$$
\begin{aligned}
(24) \quad \Theta_{0,t}(u_h(t)) \quad := \quad & C_5 h^2 \|\mathfrak{R}_t[u_h(t)]\| + C_5 h^2\ Osc(f_t(t), L^2) + C_6 h^{3/2}\|\mathfrak{J}_t[u_h]\|_{\sum_h} \\
&+ C_7 \int_0^t \Theta_0(u_h)(s)ds + C_7 C_3 \int_0^t h^2\ Osc(f(s))ds \\
&+ C_8 \Theta_0(u_h) + C_8 C_3 h^2\ Osc(f(t)).
\end{aligned}
$$

We now derive a posteriori estimate for the parabolic error ρ in the following lemma.

Lemma 4.4 *The following estimates holds true for the parabolic error:*

$$
\|\rho(t)\| \le C_9 \left[\|\rho(0)\| + 2\int_0^t \{\Theta_{0,t}(u_h(t)) + Osc(f(t))\}\, ds \right],
$$

where C_9 is a positive constant independent of the discretization parameter but depending on the final time T.

Proof. Using (6) and the definition of the Ritz-Volterra reconstructions, we have the following error equation for $\rho(t)$

$$
\begin{aligned}
&\langle \rho_t, \phi\rangle + a(\rho, \phi) - \int_0^t b(t,s;\rho(s),\phi)ds \\
&= \langle R_w u_{h,t}, \phi\rangle + a(R_w u_h, \phi) - \int_0^t b(t,s; R_w u_h(s), \phi)ds - \langle f, \phi\rangle \\
&= \langle R_w u_{h,t}, \phi\rangle + \langle \mathcal{A}_h u_h - \int_0^t \mathcal{B}_h(t,s)u_h(s)ds, \phi\rangle - \langle f, \phi\rangle \\
(25) \qquad &= \langle \epsilon_t, \phi\rangle + \langle f_h - f, \phi\rangle, \qquad \mathrm{t} \in \bar{J}
\end{aligned}
$$

for all $\phi \in H_0^1(\Omega)$. Set $\phi = \rho$ in the error equation (25). Apply Cauchy-Schwarz inequality and Young's inequality together with (3) to obtain

$$
\begin{aligned}
(26) \qquad &\frac{1}{2}\frac{d}{dt}\|\rho\|^2 + a(\rho,\rho) = \langle \epsilon_t, \rho\rangle + \langle f_h - f, \rho\rangle + \int_0^t b(t,s;\rho(s),\rho)ds \\
&\le \frac{1}{2}\beta\|\rho\|_1^2 + \frac{\gamma^2}{2\beta}\left(\int_0^t \|\rho(s)\|_1 ds\right)^2 + \Big(\|\epsilon_t\| + \|f_h - f\|\Big)\|\rho\|.
\end{aligned}
$$

Integrate (26) from 0 to t. Then, use coercivity property of $a(\cdot,\cdot)$ and Cauchy-Schwarz inequality to obtain

$$\begin{aligned}\|\rho(t)\|^2 + \beta\int_0^t \|\rho\|_1^2 \quad &\leq \quad \|\rho(0)\|^2 + \frac{C'(T)\gamma^2}{\beta}\int_0^t\int_0^s \|\rho(\tau)\|_1^2 d\tau ds \\ &+2\int_0^t \Big(\|\epsilon_t\| + \|f_h - f\|\Big)\|\rho\| ds,\end{aligned}$$

where $C'(T)$ is a positive constant depending on the final time T. Applying Gronwall's lemma and letting $\|\rho(\bar{t})\| = \sup_{s\leq t}\|\rho(s)\|,\ 0\leq \bar{t}\leq t$, with the notations (19) and (24) yields the required estimate.

The semidiscrete a posteriori estimate in $L^\infty(L^2)$-norm is presented in the following theorem.

Theorem 4.5 *(Semidiscrete a posteriori error estimate) Let u and u_h satisfy (1) and (6), respectively. Then the following a posteriori error bound holds for $0\leq t\leq T$:*

$$\max_{0\leq t\leq T}\|(u-u_h)(t)\| \leq C_9\Bigg[\|u(0)-u_h(0)\| + \Theta_0(u_h(0)) + C_3 h^2 Osc(f(0))$$
$$+2\int_0^t \Bigg(\Theta_{0,t}(u_h(t)) + Osc(f(t))\Bigg)ds\Bigg] + \Theta_0(u_h(t)),$$

where C_9 is as defined in Lemma 4.4.

Proof. Choosing the Ritz-Volterra reconstruction $R_w u_h \in H_0^1(\Omega)$ as the comparison function, express the error as

$$e(t) = u_h(t) - u(t) = (R_w u_h(t) - u(t)) - (R_w u_h(t) - u_h(t)) = \rho(t) - \epsilon(t). \tag{27}$$

Also, we have

$$\|\rho(0)\| \leq \|u(0)-u_h(0)\| + \|R_w u_h(0) - u_h(0)\| = \|u(0)-u_h(0)\| + \|\epsilon(0)\|.$$

Combine Lemma 4.2 with Lemma 4.4 and (27) to complete the rest of the proof.

Remarks. (i) The a posteriori error estimator obtained in Theorem 4.5 generalizes the result of purely parabolic problem to parabolic integro-differential equation. In the absence of the memory term (i.e., $\mathcal{B}(t,s)=0$), our error estimator is similar to that for the parabolic problem [15].

(ii) Theorem 4.5 gives the dual a posteriori analogue of a priori error estimate for semidiscrete finite element approximations to PIDE (cf. [14]).

(iii) Note that Ritz-Volterra reconstruction operator defined by (13) is a partial right inverse of the Ritz-Volterra projection [14] defined by (12). Let $\hat{U}$ denote the Ritz-Volterra reconstruction of the finite element solution u_h, then by Galerkin orthogonality property (14)

$$\hat{U} = R_w u_h \Rightarrow W_h\hat{U} = u_h.$$

5 Analysis for the fully discrete scheme

In this section, we shall discuss a posteriori error bounds for the fully discrete Galerkin approximations to the PIDE (1) based on backward Euler method. While dealing with the fully discrete scheme, we use the same symbols introduced for the semidiscrete scheme by dropping the subscript index h and using the index n. Here, we denote the fully discrete finite element approximation by U as compared with the semi-discrete finite element approximation u_h.

In order to discretize in time, we introduce the partition $0 = t_0 < t_1 < \ldots < t_N = T$ of $[0, T]$. Let $I_n := (t_{n-1}, t_n]$ and we denote by $\tau_n := t_n - t_{n-1}$ the time steps. For $t = t_n,\ n \in [0 : N]$, we set $f^n(\cdot) = f(\cdot, t_n)$.

Let $h_n(x) = \text{diam}(K),$ where $K \in \mathcal{T}_n$ and $x \in K$ denotes the local mesh-size function corresponds to each given triangulation $\mathcal{T}_n$. Let $\mathcal{S}_n$ denotes the set of internal sides of $\mathcal{T}_n$ representing edges in $d = 2$ or faces in $d = 3$, and $\sum_n$ denotes the union of all internal sides $\cup_{E \in \mathcal{S}_n} E$.

Let $(\mathcal{T}_n)_{n\in[0:N]}$ be family of conforming triangulations of the domain Ω. Each triangulation $(\mathcal{T}_n)$, for $n \in [1 : N]$, is a refinement of a macro-triangulation $\mathcal{M}$ of the domain Ω that satisfies the same conformity and shape-regularity assumptions (cf. [3]) made on its refinements. We assume the following admissible criteria as mentioned in [13]:

1. The refined triangulation is conforming.
2. The shape-regularity of an arbitrary refinement depends only on the shape-regularity of the macro-triangulation $\mathcal{M}$.

We allow only nested refinement of the space meshes at each time level $t = t_n,\ n \in [0 : N]$ i.e., for $0 \leq j \leq i \leq N$, $\mathcal{S}_i \cap \mathcal{S}_j = \mathcal{S}_j$, (cf. [18]). The complications arises during mesh change because the Volterra integral term memorizes the jumps over element edges in all previous space meshes.

Now we associate with these triangulations the finite element spaces:

$$\mathbb{V}^n := \{\phi \in H_0^1(\Omega) :\ \phi|_K \in \mathbb{P}_l,\ \forall K \in \mathcal{T}_n\},$$

where $\mathbb{P}_l$ is the space of polynomials in d variables of degree at most $l \in \mathbb{Z}^+$.

Let σ^n be the quadrature rule used to discretize the Volterra integral term. To be consistent with the backward difference scheme, we use the left rectangular rule given by

$$\sigma^n(y) = \sum_{j=0}^{n-1} \tau_{j+1} y(t_j) \approx \int_0^{t_n} y(s)ds.$$

For a function $v \in \mathbb{V}^n$, the bilinear form $a(\cdot, \cdot)$ can be represented in the same way as in (8) i.e.,

$$a(v, \phi) = \langle \mathcal{A}_{el} v, \phi\rangle + \langle J_1[v], \phi\rangle_{\sum_n}, \quad \forall \phi \in H_0^1(\Omega).$$

But, the representation of the bilinear form $b(t_n; \cdot, \cdot)$ needs a little modification. For a function $v \in H_0^1(\Omega)$, we represent the bilinear form $b(t_n; \cdot, \cdot)$ as

$$\sigma^n(b(t_n; v, \phi)) = \langle \sigma^n(\mathcal{B}_{el} v), \phi\rangle + \langle \sigma^n(J_2[v]), \phi\rangle_{\sum_n}, \quad \forall \phi \in H_0^1(\Omega),$$

where $\mathcal{B}_{el}v$ and $J_2[v]$ have the usual meaning as in (10) and

$$\sigma^n(b(t_n;v,\phi)) = \langle\sigma^n(B(t_n)\nabla v),\nabla\phi\rangle = \langle\sum_{j=0}^{n-1}\tau_{j+1}B(t_n,t_j)\nabla v(t_j),\nabla\phi\rangle.$$

The discrete operators $\mathcal{A}^n$ and $\mathcal{B}^n(s)$ at $t=t_n$ for the fully discrete case are defined in the same way as in (11). Associated with $\mathcal{A}^n$, let $\mathcal{R}^n$ be the corresponding fully discrete elliptic reconstruction operator. Moreover, P_0^n denotes the L^2 projection operator into $\mathbb{V}^n$.

The backward Euler-Galerkin fully discrete scheme may be stated as follows: Given $U^0 = P_0^0 u(0)$, find $U^n \in \mathbb{V}^n, n\in[1:N]$ such that

$$\tau_n^{-1}\langle U^n - U^{n-1},\phi_n\rangle + a(U^n,\phi_n) = \sigma^n(b(t_n;U,\phi_n)) + \langle f^n,\phi_n\rangle,\ \ \forall\phi_n\in\mathbb{V}^n. \tag{28}$$

For all $t\in I_n$, we introduce the continuous, piecewise linear approximation in time defined by

$$U(t) := l_{n-1}(t)U^{n-1} + l_n(t)U^n,\ \text{for } n\in[1:N],$$

where $l_{n-1}(t)$ and $l_n(t)$ are functions defined by

$$l_n(t) := \frac{t-t_{n-1}}{\tau_n}\ \text{ and }\ l_{n-1}(t) := \frac{t_n - t}{\tau_n}.$$

In the context of fully discrete error analysis, we now define the Ritz-Volterra reconstruction operator $R_w^n : \mathbb{V}^n \to H_0^1$ by

$$a(R_w^n v,\phi) - \sigma^n(b(t_n;R_w v,\phi)) = \langle\mathcal{A}^n v,\phi\rangle - \langle\sigma^n(\mathcal{B}^n v),\phi\rangle,\ \ \forall\,\phi\in H_0^1(\Omega). \tag{29}$$

We shall use the following definitions in the subsequent analysis:

$$\partial U^n := \frac{U^n - U^{n-1}}{\tau_n},\qquad \bar\partial U^n := P_0^n\partial U^n = \frac{U^n - P_0^n U^{n-1}}{\tau_n},\quad \forall\, n\in[1:N]$$

and

$$\bar f^n := P_0^n f^n.$$

Further, Lemma 4.1 holds true for the fully discrete case with $\Pi^n : H_0^1(\Omega)\to\mathbb{V}^n$ as the Clément-type interpolation operator as introduced in [17].

Lemma 5.1 *(Ritz-Volterra reconstruction error estimates)*
For any $v\in\mathbb{V}^n$, the following estimates holds true:

$$\begin{aligned}\|R_w^n v - v\|_1 &\le C_1 h_n\|\mathcal{A}^n v - \mathcal{A}_{el}v - \sigma^n(\mathcal{B}^n v) + \sigma^n(\mathcal{B}_{el}v)\|\\ &\quad + C_2 h_n^{1/2}\|J_1[v] - \sigma^n(J_2[v])\|_{\sum_n},\\ \|R_w^n v - v\| &\le C_3 h_n^2\|\mathcal{A}^n v - \mathcal{A}_{el}v - \sigma^n(\mathcal{B}^n v) + \sigma^n(\mathcal{B}_{el}v)\|\\ &\quad + C_4 h_n^{3/2}\|J_1[v] - \sigma^n(J_2[v])\|_{\sum_n},\end{aligned}$$

where $C_j, j=1,2,3,4$ are the positive constants independent of the discretization parameters but depending upon the shape-regularity of the family of triangulations and the final time T.

Proof. The proof is same as that in the semidiscrete case. The discrete Gronwall's lemma [23] is used instead of the continuous one and the quadrature approximation is taken for the integral term.

For any $\phi \in H_0^1(\Omega)$, we have from (28)

$$\begin{aligned}
&\langle \bar{\partial} U^n + \mathcal{A}^n U^n - \sigma^n(\mathcal{B}^n U) - \bar{f}^n, \phi\rangle = \langle \bar{\partial} U^n + \mathcal{A}^n U^n - \sigma^n(\mathcal{B}^n U) - \bar{f}^n, P_0^n \phi\rangle \\
&= \langle \bar{\partial} U^n, P_0^n\phi\rangle + a(U^n, P_0^n\phi) - \sigma^n(b(t_n; U, P_0^n\phi)) - \langle \bar{f}^n, P_0^n\phi\rangle \\
&= \tau_n^{-1}\langle U^n - P_0^n U^{n-1}, P_0^n\phi\rangle + a(U^n, P_0^n\phi) - \sigma^n(b(t_n; U, P_0^n\phi)) - \langle \bar{f}^n, P_0^n\phi\rangle \\
&= \tau_n^{-1}\langle U^n - U^{n-1}, P_0^n\phi\rangle + a(U^n, P_0^n\phi) - \sigma^n(b(t_n; U, P_0^n\phi)) - \langle f^n, P_0^n\phi\rangle \\
&= 0.
\end{aligned}$$

The fully discrete scheme can be written in the following distributional form:

$$\bar{\partial} U^n + \mathcal{A}^n U^n(x) = \sigma^n(\mathcal{B}^n U(x)) + \bar{f}^n(x), \quad \forall x \in \Omega. \tag{30}$$

For the sake of convenience, we shall use the following shorthand notation

$$\omega(t) = R_w U(t), \quad \text{for } t \in I_n$$

to denote the Ritz-Volterra reconstruction of fully discrete solution $U(t)$. Now, associate $\omega(t),\ t \in I_n$ with the values ω^n and ω^{n-1} by

$$\omega(t) := l_{n-1}(t)\omega^{n-1} + l_n(t)\omega^n.$$

Lemma 5.2 *For each $n \in [1:N]$, and for each $\phi \in H_0^1(\Omega)$, we have the following parabolic error equation*

$$\begin{aligned}
\langle \rho_t, \phi\rangle \quad &+ \quad a(\rho, \phi) - \int_0^t b(t, s; \rho(s), \phi)ds \\
&= \quad \langle \epsilon_t, \phi\rangle + a(\omega - \omega^n, \phi) - \int_0^t b(t, s; \omega(s), \phi)ds \\
&\qquad + \sigma^n(b(t_n; \omega, \phi)) + \langle P_0^n f^n - f, \phi\rangle + \tau_n^{-1}\langle P_0^n U^{n-1} - U^{n-1}, \phi\rangle.
\end{aligned} \tag{31}$$

Proof. For $t \in I_n$, using (29), (5) and (30), we have $\forall \phi \in H_0^1(\Omega)$

$$\begin{aligned}
\langle \rho_t, \phi\rangle \quad &+ \quad a(\rho, \phi) - \int_0^t b(t, s; \rho(s), \phi)ds \\
&= \quad \langle \omega_t, \phi\rangle + a(\omega, \phi) - \int_0^t b(t, s; \omega(s), \phi)ds - \langle f, \phi\rangle \\
&= \quad \langle \omega_t, \phi\rangle + a(\omega, \phi) - \int_0^t b(t, s; \omega(s), \phi)ds - \langle f, \phi\rangle - \langle \partial U^n, \phi\rangle \\
&\qquad + \tau_n^{-1}\langle P_0^n U^{n-1} - U^{n-1}, \phi\rangle - a(\omega^n, \phi) + \sigma^n(b(t_n; \omega, \phi)) + \langle P_0^n f^n, \phi\rangle \\
&= \quad \langle \epsilon_t, \phi\rangle + a(\omega - \omega^n, \phi) - \int_0^t b(t, s; \omega(s), \phi)ds + \sigma^n(b(t_n; \omega, \phi)) \\
&\qquad + \langle P_0^n f^n - f, \phi\rangle + \tau_n^{-1}\langle P_0^n U^{n-1} - U^{n-1}, \phi\rangle,
\end{aligned}$$

where we have used the fact that $\partial U^n = U_t(t),\ \forall t \in I_n$.

Remark. One can observe that in the absence of time-discretization error and mesh change error (i.e., in the absence of the second, third, fourth and sixth terms in (31)) fully-discrete parabolic error equation (31) reduces to the parabolic error equation (25) for the semidiscrete scheme. This shows that space-time discretizations are properly adapted to the space discretizations.

Similar to the semidiscrete case, define the inner residual for $n \in [0:N]$ as:

$$\begin{aligned}\mathfrak{R}^n &:= \mathcal{A}_{el}U^n - \sigma^n(\mathcal{B}_{el}U) - \mathcal{A}^n U^n + \sigma^n(\mathcal{B}^n U) = \mathcal{A}_{el}U^n - \sigma^n(\mathcal{B}_{el}U) - \bar{f}^n + \bar{\partial}U^n,\\ \mathfrak{R}^0 &:= \mathcal{A}_{el}U^0 - \mathcal{A}^0 U^0,\end{aligned}$$

and the jump residual for $n \in [0:N]$ as

$$\begin{aligned}\mathfrak{J}^n &:= J_1[U^n] - \sigma^n(J_2[U]),\\ \mathfrak{J}^0 &:= J_1[U^0].\end{aligned} \tag{32}$$

The inner residual terms can also be written in the following form

$$< \mathfrak{R}^n, \phi >:= \sum_{K\in\mathcal{T}_n} \langle -\mathrm{div}(A\nabla U^n) + \sigma^n(\mathrm{div}(B(t_n)\nabla U)) - P_0^n f(t_n) + \frac{U^n - P_0^n U^{n-1}}{\tau_n}, \phi\rangle_K.$$

For the purpose of fully discrete analysis, we introduce the following estimators that are local in time.

For $n \in [0:N]$,

$$\alpha_n := C_1 h_n \|\mathfrak{R}^n\| + C_2 h_n^{1/2}\|\mathfrak{J}^n\|_{\sum_n}, \tag{33}$$

$$\beta_n := C_3 h_n^2 \|\mathfrak{R}^n\| + C_4 h_n^{3/2}\|\mathfrak{J}^n\|_{\sum_n}, \tag{34}$$

are Ritz-Volterra reconstruction error estimators.

For $n \in [0:N]$,

$$\begin{aligned}\xi_n := \frac{\gamma'}{2}\Big(&\hat{\tau}_{n-1}^2 \sum_{j=1}^{n-1}\{\alpha_j + \alpha_{j-1}\} + \hat{\tau}_n^2 \sum_{j=1}^{n}\{\alpha_j + \alpha_{j-1}\}\\ &+\hat{\tau}_{n-1}^2 \sum_{j=1}^{n-1}\Big\{\|U^j\|_1 + \|U^{j-1}\|_1\Big\} + \hat{\tau}_n^2 \sum_{j=1}^{n}\Big\{\|U^j\|_1 + \|U^{j-1}\|_1\Big\}\Big)\\ &+2\gamma\sqrt{1/3}\Big(\hat{\tau}_n \sum_{j=1}^{n}\{\alpha_j + \alpha_{j-1}\} + \hat{\tau}_n^2 \sum_{j=1}^{n}\|\partial U^j\|_1\Big)\end{aligned} \tag{35}$$

is the quadrature error estimator, where $\hat{\tau}_n = \max_{j=1}^n \tau_j$.

For $n \in [1:N]$,

$$\zeta_n := C_\Omega C_{10}\Big(\frac{\hat{\tau}_n}{\tau_n}\Big)\Big[h_n^2\|\partial\mathfrak{R}^n\| + h_n^{3/2}\|\partial\mathfrak{J}^n\|_{\Sigma_n} + \sum_{j=0}^{n-1}\beta_j\Big] \tag{36}$$

are the space and mesh modification error estimators.

$$\eta_n = \begin{cases} \frac{1}{2}\|\bar{f}^1 - \overline{\partial}U^1 - \mathcal{A}^0U^0\|, & \text{for } n = 1,\\ \frac{1}{2}\tau_n\|\partial(\bar{f}^n - \overline{\partial}U^n)\|, & \text{for } n \in [2:N]\end{cases} \tag{37}$$

are the time error estimators.

For $n \in [1:N]$,

$$\mu_n := C_{11} h_n \|(P_0^n - I)(f^n + \frac{U^{n-1}}{\tau_n})\| \tag{38}$$

are the data approximation and mesh modification error estimators, and for $n \in [1:N]$,

$$\lambda_n := \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \|f^n - f(t)\| dt \tag{39}$$

are the data oscillations error estimators.

The following lemma yields a bound for the parabolic error $\rho(t)$.

Lemma 5.3 *($L^\infty(L^2)$ and $L^2(H^1)$ a posteriori estimate for the parabolic error). For each $m \in [1:N]$, the following estimate holds:*

$$\left(\max_{[0,t_m]} \|\rho(t)\|^2 + \beta \int_0^{t_m} \|\rho(t)\|_1^2 dt \right)^{1/2} \leq \|\rho(t_0)\| + 2C(t_m)(\sigma_{1,m}^2 + \sigma_{2,m}^2)^{1/2},$$

where

$$\sigma_{1,m} = \sum_{n=1}^{m} (\zeta_n + \eta_n + \lambda_n)\tau_n,$$
$$\sigma_{2,m}^2 = \sum_{n=1}^{m} (\xi_n + \mu_n)^2 (\tau_n/\beta),$$

and $C(t_m)$ is a positive constant depends upon the time t_m.

Proof Set $\phi = \rho$ in the error equation (31) to obtain

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|\rho(t)\|^2 &+ a(\rho,\rho) = \int_0^t b(t,s;\rho(s),\rho)ds + \langle \epsilon_t, \rho \rangle \\ &+ a(\omega - \omega^n, \rho) - \int_0^t b(t,s;\omega(s),\rho)ds + \sigma^n(b(t_n;\omega,\rho)) \\ &+ \langle P_0^n f^n - f, \rho \rangle + \tau_n^{-1} \langle P_0^n U^{n-1} - U^{n-1}, \rho \rangle. \end{aligned}$$

Using (2), (3) and Young's inequality, we obtain

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|\rho(t)\|^2 &+ \frac{\beta}{2}\|\rho\|_1^2 \leq \frac{\gamma^2}{2\beta}\left(\int_0^t \|\rho(s)\|_1 ds \right)^2 + |\langle \epsilon_t, \rho \rangle| \\ &+ |a(\omega - \omega^n, \rho) - \int_0^t b(t,s;\omega(s),\rho)ds + \sigma^n(b(t_n;\omega,\rho))| \\ &+ |\langle P_0^n f^n - f, \rho \rangle| + |\tau_n^{-1} \langle P_0^n U^{n-1} - U^{n-1}, \rho \rangle|. \end{aligned}$$

Apply Cauchy-Schwarz inequality and then integrate from t_{n-1} to t_n. Summing over $n =$

$1:m$ and applying Gronwall's lemma it now leads to

$$\|\rho(t_m)\|^2 + \beta\int_0^{t_m} \|\rho(t)\|_1^2 dt - \|\rho(t_0)\|^2 \leq 2C(t_m)\Bigg[\sum_{n=1}^{m}\int_{t_{n-1}}^{t_n}\Bigg(|\langle\epsilon_t(t), \rho(t)\rangle|$$
$$+|a(\omega(t)-\omega^n, \rho(t)) - \int_0^t b(t,s;\omega(s),\rho(t))ds + \sigma^n(b(t_n;\omega,\rho(t))|$$
$$+|\langle P_0^n f^n - f^n + \tau_n^{-1}(P_0^n U^{n-1} - U^{n-1}), \rho(t)\rangle| + \langle f^n - f(t), \rho(t)\rangle|\Bigg)dt\Bigg]$$
$$:= 2C(t_m)\sum_{n=1}^{m}(\mathcal{I}_n^1 + \mathcal{I}_n^2 + \mathcal{I}_n^3 + \mathcal{I}_n^4) := 2C(t_m)\mathcal{I}_m,$$

where $\mathcal{I}_n^1$ denotes the spatial error, $\mathcal{I}_n^2$ the time discretization error, $\mathcal{I}_n^3$ the mesh change error, $\mathcal{I}_n^4$ the data oscillation error and $C(t_m)$ is a positive constant appeared due to the application of Gronwall's lemma which will depend on t_m.

Denote $t_m^* \in [0, t_m]$ the time for which

$$\max_{t\in[0,t_m]}\|\rho(t)\| = \|\rho(t_m^*)\| := \|\rho_*^m\|.$$

Hence, we have

$$\|\rho_*^m\|^2 + \beta\int_0^{t_m}\|\rho(t)\|_1^2 dt \leq \|\rho(t_0)\|^2 + 2C(t_m)\mathcal{I}_m.$$

Spatial error estimates. To estimate the term $\mathcal{I}_n^1$, for $n \in [1:N]$, we note that

$$\begin{aligned}\mathcal{I}_n^1 &= \int_{t_{n-1}}^{t_n} |\langle\epsilon_t(t), \rho(t)\rangle| dt \\ &= \tau_n^{-1}\int_{t_{n-1}}^{t_n} |\langle\omega^n - \omega^{n-1} - U^n + U^{n-1}, \rho(t)\rangle| dt. \qquad (40)\end{aligned}$$

Since $\omega^n - U^n$ is orthogonal to $\mathbb{V}^n$ with respect to $a(\cdot,\cdot) - \sigma^n(b(t_n;(\cdot),\cdot))$, the first term in the inner product is orthogonal to $\mathbb{V}^n \cap \mathbb{V}^{n-1}$. To give a simplified analysis, we exploit here orthogonality property of the Ritz-Volterra reconstructions under the nested refinement condition to introduce the Clément-type interpolation operator Π^n. We shall use duality technique to estimate (40).

For $t \in (0,T)$, let $\psi \in H^2(\Omega) \cap H_0^1(\Omega)$ be the solution of the following dual elliptic problem in the weak form

$$a(\chi, \psi(t)) = \langle\chi, \rho(t)\rangle,$$

satisfying the following regularity estimate:

$$\|\psi\|_2 \leq C_\Omega\|\rho\|, \ \forall\chi \in H_0^1(\Omega), \qquad (41)$$

where the constant C_Ω depending on the domain Ω. Now, using the definition of the Ritz-

Volterra reconstruction and making adjustment of the terms, we have

$$
\begin{aligned}
&\langle \omega^n - \omega^{n-1} - U^n + U^{n-1}, \rho(t)\rangle \\
= \;& a(\omega^n - \omega^{n-1} - U^n + U^{n-1}, \psi(t)) \\
= \;& a(\omega^n - \omega^{n-1} - U^n + U^{n-1}, \psi(t) - \Pi^n\psi(t)) \\
& + \sigma^n(b(t_n; \omega - U, \Pi^n\psi(t))) - \sigma^{n-1}(b(t_{n-1}; \omega - U, \Pi^n\psi(t))) \\
= \;& a(\omega^n - \omega^{n-1} - U^n + U^{n-1}, \psi(t) - \Pi^n\psi(t)) \\
& - \sigma^n(b(t_n; \omega - U, \psi(t) - \Pi^n\psi(t))) \\
& + \sigma^{n-1}(b(t_{n-1}; \omega - U, \psi(t) - \Pi^n\psi(t))) \\
& + \sigma^n(b(t_n; \omega - U, \psi(t))) - \sigma^{n-1}(b(t_{n-1}; \omega - U, \psi(t))) \\
= \;& \langle \mathcal{A}^n U^n - \sigma^n(\mathcal{B}^n U) - \mathcal{A}_{el} U^n + \sigma^n(\mathcal{B}_{el} U), \psi(t) - \Pi^n\psi(t)\rangle \\
& - \langle \mathcal{A}^{n-1} U^{n-1} - \sigma^{n-1}(\mathcal{B}^{n-1} U) - \mathcal{A}_{el} U^{n-1} + \sigma^n(\mathcal{B}_{el} U), \psi(t) - \Pi^n\psi(t)\rangle \\
& + \langle \sigma^n(J_2[U]) - J_1[U^n] - \sigma^{n-1}(J_2[U]) + J_1[U^{n-1}], \psi(t) - \Pi^n\psi(t)\rangle_{\Sigma_n} \\
& + \sigma^n(b(t_n; \omega - U, \psi(t))) - \sigma^{n-1}(b(t_{n-1}; \omega - U, \psi(t))).
\end{aligned}
$$

Using the distributional form of the fully discrete scheme, on each interval I_n, we have

$$
\begin{aligned}
& \mathcal{A}^n U^n - \mathcal{A}^{n-1} U^{n-1} + \sigma^{n-1}(\mathcal{B}^{n-1} U) - \sigma^n(\mathcal{B}^n U) \\
& + \mathcal{A}_{el} U^{n-1} - \mathcal{A}_{el} U^n + \sigma^n(\mathcal{B}_{el} U) - \sigma^n(\mathcal{B}_{el} U) \\
& = \mathfrak{R}^{n-1} - \mathfrak{R}^n = -\tau_n \partial \mathfrak{R}^n.
\end{aligned}
$$

Using (32) with $\mathfrak{J}^n - \mathfrak{J}^{n-1} = \tau_n \partial \mathfrak{J}^n$, we now obtain

$$
\begin{aligned}
& |\langle \omega^n - \omega^{n-1} - U^n + U^{n-1}, \rho(t)\rangle| \\
\leq \;& \tau_n \|\partial \mathfrak{R}^n\| \|\psi(t) - \Pi^n\psi(t)\| + \tau_n \|\partial \mathfrak{J}^n\|_{\Sigma_n} \|\psi(t) - \Pi^n\psi(t)\|_{\Sigma_n} \\
& + |\sigma^n(b(t_n; \omega - U, \psi(t))) - \sigma^{n-1}(b(t_{n-1}; \omega - U, \psi(t)))|.
\end{aligned} \tag{42}
$$

To handle the last term above, we use (17) by replacing $(R_w v - v)$ by $(\omega - U)$ and Cauchy-Schwarz inequality together with $\|\mathcal{B}^*(t_n, t_j)\psi\| \leq C_{\mathcal{B}_1^*}\|\psi\|_2$ for all $j \in [0 : n]$ to obtain

$$
\begin{aligned}
& |\sigma^n(b(t_n; \omega - U, \psi(t))) - \sigma^{n-1}(b(t_{n-1}; \omega - U, \psi(t)))| \\
& \leq |\langle \sum_{j=0}^{n-1} \tau_{j+1}(\omega - U)(t_j), \mathcal{B}^*(t_n, t_j)\psi(t)\rangle - \langle \sum_{j=0}^{n-2} \tau_{j+1}(\omega - U)(t_j), \mathcal{B}^*(t_{n-1}, t_j)\psi(t)\rangle| \\
& \leq \hat{\tau}_n \|\sum_{j=0}^{n-1} (\omega - U)(t_j)\| \|\mathcal{B}^*(t_n, t_j)\psi(t)\| + \hat{\tau}_{n-1} \|\sum_{j=0}^{n-2} (\omega - U)(t_j)\| \|\mathcal{B}^*(t_{n-1}, t_j)\psi(t)\| \\
& \leq C_{\mathcal{B}_1^*} \Big[\hat{\tau}_n \sum_{j=0}^{n-1} \beta_j + \hat{\tau}_{n-1} \sum_{j=0}^{n-2} \beta_j\Big] \|\psi(t)\|_2 \\
& \leq C_{\mathcal{B}_4^*} \hat{\tau}_n \Big[\sum_{j=0}^{n-1} \beta_j\Big] \|\psi(t)\|_2,
\end{aligned}
$$

where $C_{\mathcal{B}_4^*} = 2C_{\mathcal{B}_1^*}$. Using the above estimate in (42) and applying Lemma 4.1 with $C_{10} =$

$\max\Big(C_{\mathcal{B}_4^*}, 1\Big)$, we obtain

$$
\begin{aligned}
(43)\qquad & |\langle \omega^n - \omega^{n-1} - U^n + U^{n-1}, \rho(t)\rangle| \\
& \leq C_{10}\|\psi\|_2 \hat{\tau}_n \big[h_n^2\|\partial R^n\| + h_n^{3/2}\|\partial \mathfrak{J}^n\|_{\Sigma_n} + \sum_{j=0}^{n-1}\beta_j\big].
\end{aligned}
$$

Using (43) in (40) together with (41) yields

$$
\begin{aligned}
\mathcal{I}_n^1 &\leq C_{10}\tau_n^{-1}\int_{t_{n-1}}^{t_n}\|\psi(t)\|_2 dt\ \hat{\tau}_n\Big[h_n^2\|\partial\mathfrak{R}^n\| + h_n^{3/2}\|\partial\mathfrak{J}^n\|_{\Sigma_n} + \sum_{j=0}^{n-1}\beta_j\Big] \\
&\leq \max_{t\in I_n}\|\rho(t)\|\tau_n\zeta_n,
\end{aligned}
$$

where we have used (36). Summing from $n = 1 : m$ we obtain

$$
\sum_{n=1}^{m}\mathcal{I}_n^1 \leq \|\rho(t_*^m)\|\sum_{n=1}^{m}\tau_n\zeta_n.
$$

Remark. In the case of parabolic problem [13], the authors have introduced Clément-type interpolation operator $\hat{\Pi}^n : H_0^1(\Omega) \to \mathbb{V}^n \cap \mathbb{V}^{n-1}$ in order to handle spatial error term similar to (40). But, in our analysis due to presence of the quadrature term, one has to look back through all the previous time levels which will make the analysis much more complicated. So, the introduction to such kind of operator is avoided using nested refinement condition.

Time error estimates. In order to count the time discretization error, let

$$
\hat{\phi}(t) := \int_0^t B(t,s)\nabla\omega(s)ds.
$$

Then, for $t \in I_n$, we associate the integral vectors $\hat{\phi}(t_{n-1})$ and $\hat{\phi}(t_n)$ with $\hat{\phi}(t)$ as

$$
\hat{\phi}(t) := l_{n-1}(t)\hat{\phi}(t_{n-1}) + l_n(t)\hat{\phi}(t_n).
$$

Then

$$
\begin{aligned}
\int_0^t b(t,s;\omega(s),\rho(t))ds &= l_{n-1}(t)\int_0^{t_{n-1}} b(t_{n-1},s;\omega(s),\rho(t))ds \\
&+ l_n(t)\int_0^{t_n} b(t_n,s;\omega(s),\rho(t))ds.
\end{aligned}
$$

By the definition of discrete time extensions, we get

$$
\begin{aligned}
\mathcal{I}_n^2 &= \int_{t_{n-1}}^{t_n} |a(\omega(t)-\omega^n,\rho(t)) - \int_0^t b(t,s;\omega(s),\rho(t))ds + \sigma^n(b(t_n;\omega,\rho(t)))|dt \\
&= \int_{t_{n-1}}^{t_n} |a(l_{n-1}(t)\omega^{n-1} + l_n(t)\omega^n - \omega^n, \rho(t)) \\
&\quad -\Big[l_{n-1}(t)\int_0^{t_{n-1}} b(t_{n-1},s;\omega(s),\rho(t))ds \\
&\quad + l_n(t)\int_0^{t_n} b(t_n,s;\omega(s),\rho(t))ds - \sigma^n(b(t_n;\omega,\rho(t)))\Big]|dt
\end{aligned}
$$

$$\begin{aligned} = & \int_{t_{n-1}}^{t_n} |l_{n-1}(t)\Big\{a(\omega^{n-1},\rho(t)) - \sigma^{n-1}(b(t_{n-1};\omega,\rho(t)))\Big\} \\ & +(l_n(t)-1)\Big\{a(\omega^n,\rho(t)) - \sigma^n(b(t_n;\omega,\rho(t)))\Big\} \\ & +l_{n-1}(t)\Big\{\sigma^{n-1}(b(t_{n-1};\omega,\rho(t))) - \int_0^{t_{n-1}} b(t_{n-1},s;\omega(s),\rho(t))ds\Big\} \\ & +l_n(t)\Big\{\sigma^n(b(t_n;\omega,\rho(t))) - \int_0^{t_n} b(t_n,s;\omega(s),\rho(t))ds\Big\}|dt. \end{aligned}$$

Now, using the identity, $\frac{l_n(t)-1}{l_{n-1}(t)} = -1$, for $t \in I_n$, the definition of the Ritz-Volterra reconstructions for the fully discrete case and by adjusting some terms, we obtain

$$\begin{aligned} \mathcal{I}_n^2 = & \int_{t_{n-1}}^{t_n} |l_{n-1}(t)\Big\{\langle \mathcal{A}^{n-1}U^{n-1},\rho(t)\rangle - \langle \sigma^{n-1}(\mathcal{B}^{n-1}U),\rho(t)\rangle - \langle \mathcal{A}^nU^n,\rho(t)\rangle \\ & +\langle \sigma^n(\mathcal{B}^nU),\rho(t)\rangle\Big\} + l_{n-1}(t)\Big\{\sigma^{n-1}(b(t_{n-1};\omega-U,\rho(t))) \\ & -\int_0^{t_{n-1}} b(t_{n-1},s;\omega(s)-U(s),\rho(t))ds\Big\} + l_n(t)\Big\{\sigma^n(b(t_n;\omega-U,\rho(t))) \\ & -\int_0^{t_n} b(t_n,s;\omega(s)-U(s),\rho(t))ds\Big\}dt + l_{n-1}(t)\Big\{\sigma^{n-1}(b(t_{n-1};U,\rho(t))) \\ & -\int_0^{t_{n-1}} b(t_{n-1},s;U(s),\rho(t))ds\Big\} + l_n(t)\Big\{\sigma^n(b(t_n;U,\rho(t))) \\ & -\int_0^{t_n} b(t_n,s;U(s),\rho(t))ds\Big\}|dt. \end{aligned} \tag{44}$$

We know that if $\psi_{1n}(s) = (t_n - s)$, then we have

$$\int_{t_{n-1}}^{t_n} y(s)ds - \tau_n y(t_{n-1}) = \int_{t_{n-1}}^{t_n} \psi_{1n}(s)\frac{dy}{ds}ds. \tag{45}$$

By setting $y(s) = B(t_n,s)\nabla v(s),\ s \in I_n$, we have

$$\begin{aligned} \frac{dy}{ds} &= B_s(t_n,s)\nabla v(s) + B(t_n,s)\nabla v_s(s) \\ &= B_s(t_n,s)\nabla v(s) + B(t_n,s)\nabla\partial v^n \\ &= l_{n-1}(s)B_s(t_n,s)\nabla v^{n-1} + l_n(s)B_s(t_n,s)\nabla v^n + B(t_n,s)\nabla\partial v^n, \end{aligned}$$

where we have used the fact $v(s) = l_{n-1}(s)v^{n-1} + l_n(s)v^n$ and $v_s(s) = \partial v^n$. Using (45), the second term on the right hand side of (44) becomes

$$\begin{aligned} & \sigma^{n-1}(b(t_{n-1};\omega-U,\rho(t))) - \int_0^{t_{n-1}} b(t_{n-1},s;\omega(s)-U(s),\rho(t))ds \\ = & \langle -\sum_{j=1}^{n-1}\int_{t_{j-1}}^{t_j} \psi_{1j}(s)\frac{\partial\Big\{B(t_{n-1},s)\nabla(\omega-U)\Big\}}{\partial s}ds, \nabla\rho(t)\rangle \\ = & -\sum_{j=1}^{n-1}\int_{t_{j-1}}^{t_j} b_s(t_{n-1},s;\psi_{1j}(s)(\omega-U)(s),\rho(t))ds \\ & -\sum_{j=1}^{n-1}\int_{t_{j-1}}^{t_j} b(t_{n-1},s;\psi_{1j}(s)\partial(\omega^j-U^j),\rho(t))ds \end{aligned}$$

Thus, applying continuity of $b(t,s,\cdot,\cdot)$ and $b_s(t,s;\cdot,\cdot)$ together with the fact that $l_n(t) \leq 1,\ t \in I_n$ and $l_{n-1}(t) \leq 1,\ t \in I_n$, we have

$$
\begin{aligned}
&|\sigma^{n-1}(b(t_{n-1};\omega - U, \rho(t))) - \int_0^{t_{n-1}} b(t_{n-1}, s; \omega(s) - U(s), \rho(t))ds| \\
&\leq \Bigg[\frac{\gamma' \hat{\tau}_{n-1}^2}{2} \sum_{j=1}^{n-1} \Big\{ \|\omega^j - U^j\|_1 + \|\omega^{j-1} - U^{j-1}\|_1 \Big\} \\
&+ \gamma \hat{\tau}_{n-1} \sum_{j=1}^{n-1} \Big\{ \|\omega^j - U^j\|_1 + \|\omega^{j-1} - U^{j-1}\|_1 \Big\} \Bigg] \|\rho(t)\|_1.
\end{aligned}
$$

Similarly, for the other terms, we have

$$
\begin{aligned}
&|\sigma^{n}(b(t_{n};\omega - U, \rho(t))) - \int_0^{t_{n}} b(t_{n}, s; \omega(s) - U(s), \rho(t))ds| \\
&\leq \Bigg[\frac{\gamma' \hat{\tau}_{n}^2}{2} \sum_{j=1}^{n} \Big\{ \|\omega^j - U^j\|_1 + \|\omega^{j-1} - U^{j-1}\|_1 \Big\} \\
&+ \gamma \hat{\tau}_{n} \sum_{j=1}^{n} \Big\{ \|\omega^j - U^j\|_1 + \|\omega^{j-1} - U^{j-1}\|_1 \Big\} \Bigg] \|\rho(t)\|_1,
\end{aligned}
$$

$$
\begin{aligned}
&|\sigma^{n-1}(b(t_{n-1};U, \rho(t))) - \int_0^{t_{n-1}} b(t_{n-1}, s; U(s), \rho(t))ds| \\
&\leq \Bigg[\frac{\gamma' \hat{\tau}_{n-1}^2}{2} \sum_{j=1}^{n-1} \Big\{ \|U^j\|_1 + \|U^{j-1}\|_1 \Big\} + \gamma \hat{\tau}_{n-1} \sum_{j=1}^{n-1} \Big\{ \|U^j\|_1 + \|U^{j-1}\|_1 \Big\} \Bigg] \|\rho(t)\|_1
\end{aligned}
$$

and

$$
\begin{aligned}
&|\sigma^{n}(b(t_{n};U, \rho(t))) - \int_0^{t_{n}} b(t_{n}, s; U(s), \rho(t))ds| \\
&\leq \Bigg[\frac{\gamma' \hat{\tau}_{n}^2}{2} \sum_{j=1}^{n} \Big\{ \|U^j\|_1 + \|U^{j-1}\|_1 \Big\} + \gamma \hat{\tau}_{n} \sum_{j=1}^{n} \Big\{ \|U^j\|_1 + \|U^{j-1}\|_1 \Big\} \Bigg] \|\rho(t)\|_1,
\end{aligned}
$$

Using the above estimates in (44) and an application of Cauchy-Schwarz inequality yields

$$
\begin{aligned}
\mathcal{I}_n^2 \quad \leq \quad & \int_{t_{n-1}}^{t_n} l_{n-1}(t)\|\mathcal{A}^{n-1}U^{n-1} - \sigma^{n-1}(\mathcal{B}^{n-1}U) - \mathcal{A}^n U^n + \sigma^n(\mathcal{B}^n U)\|\|\rho(t)\|dt \\
& + \frac{\gamma'}{2}\int_{t_{n-1}}^{t_n} \Bigg[l_{n-1}(t)\hat{\tau}_{n-1}^2 \sum_{j=1}^{n-1} \Big\{ \|\omega^j - U^j\|_1 + \|\omega^{j-1} - U^{j-1}\|_1 \Big\} \\
& + l_n(t)\hat{\tau}_n^2 \sum_{j=1}^{n} \Big\{ \|\omega^j - U^j\|_1 + \|\omega^{j-1} - U^{j-1}\|_1 \Big\} \\
& + l_{n-1}(t)\hat{\tau}_{n-1}^2 \sum_{j=1}^{n-1} \Big\{ \|U^j\|_1 + \|U^{j-1}\|_1 \Big\} + l_n(t)\hat{\tau}_n^2 \sum_{j=1}^{n} \Big\{ \|U^j\| + \|U^{j-1}\|_1 \Big\} \Bigg] \|\rho(t)\|_1 dt \\
& + \gamma \int_{t_{n-1}}^{t_n} \Bigg[l_{n-1}(t)\hat{\tau}_{n-1} \sum_{j=1}^{n-1} \Big\{ \|\omega^j - U^j\|_1 + \|\omega^{j-1} - U^{j-1}\|_1 \Big\} \\
& + l_n(t)\hat{\tau}_n \sum_{j=1}^{n} \Big\{ \|\omega^j - U^j\|_1 + \|\omega^{j-1} - U^{j-1}\|_1 \Big\} \\
& + l_{n-1}(t)\hat{\tau}_{n-1} \sum_{j=1}^{n-1} \|U^j - U^{j-1}\|_1 + l_n(t)\hat{\tau}_n \sum_{j=1}^{n} \|U^j - U^{j-1}\|_1 \Bigg] \|\rho(t)\|_1 dt.
\end{aligned}
$$

With an aid of (33), we obtain

$$
\begin{aligned}
\mathcal{I}_n^2 \quad \leq \quad & 1/2 \; \tau_n \max_{t \in I_n} \|\rho(t)\| \; \|\mathcal{A}^{n-1}U^{n-1} - \sigma^{n-1}(\mathcal{B}^{n-1}U) - \mathcal{A}^n U^n + \sigma^n(\mathcal{B}^n U)\| \\
& + \frac{\gamma'}{2}\Bigg[\hat{\tau}_{n-1}^2 \Big(\int_{t_{n-1}}^{t_n} l_{n-1}^2(t)dt \Big)^{1/2} \sum_{j=1}^{n-1} \{\alpha_j + \alpha_{j-1}\} \\
& + \hat{\tau}_n^2 \Big(\int_{t_{n-1}}^{t_n} l_n^2(t)dt \Big)^{1/2} \sum_{j=1}^{n} \{\alpha_j + \alpha_{j-1}\} \\
& + \hat{\tau}_{n-1}^2 \Big(\int_{t_{n-1}}^{t_n} l_{n-1}^2(t)dt \Big)^{1/2} \sum_{j=1}^{n-1} \Big\{ \|U^j\|_1 + \|U^{j-1}\|_1 \Big\} \\
& + \hat{\tau}_n^2 \Big(\int_{t_{n-1}}^{t_n} l_n^2(t)dt \Big)^{1/2} \sum_{j=1}^{n} \Big\{ \|U^j\|_1 + \|U^{j-1}\|_1 \Big\} \Bigg] \Big(\int_{t_{n-1}}^{t_n} \|\rho(t)\|_1^2 dt \Big)^{1/2} \\
& + \gamma \Bigg[\hat{\tau}_{n-1} \Big(\int_{t_{n-1}}^{t_n} l_{n-1}^2(t)dt \Big)^{1/2} \sum_{j=1}^{n-1} \{\alpha_j + \alpha_{j-1}\} \\
& + \hat{\tau}_n \Big(\int_{t_{n-1}}^{t_n} l_n^2(t)dt \Big)^{1/2} \sum_{j=1}^{n} \{\alpha_j + \alpha_{j-1}\} \\
& + \Big(\int_{t_{n-1}}^{t_n} l_{n-1}^2(t)dt \Big)^{1/2} \hat{\tau}_{n-1} \sum_{j=1}^{n-1} \|U^j - U^{j-1}\|_1 \\
& + \Big(\int_{t_{n-1}}^{t_n} l_n^2(t)dt \Big)^{1/2} \hat{\tau}_n \sum_{j=1}^{n} \|U^j - U^{j-1}\|_1 \Bigg] \Big(\int_{t_{n-1}}^{t_n} \|\rho(t)\|_1^2 dt \Big)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
&\leq 1/2\ \tau_n \max_{t\in I_n}\|\rho(t)\|\ \|\mathcal{A}^{n-1}U^{n-1}-\sigma^{n-1}(\mathcal{B}^{n-1}U)-\mathcal{A}^nU^n+\sigma^n(\mathcal{B}^nU)\|\\
&\quad+\sqrt{1/3}\ \tau_n^{1/2}\Bigg[\frac{\gamma'}{2}\Bigg(\hat{\tau}_{n-1}^2\sum_{j=1}^{n-1}\{\alpha_j+\alpha_{j-1}\}+\hat{\tau}_n^2\sum_{j=1}^{n}\{\alpha_j+\alpha_{j-1}\}\\
&\quad+\hat{\tau}_{n-1}^2\sum_{j=1}^{n-1}\Big\{\|U^j\|_1+\|U^{j-1}\|_1\Big\}+\hat{\tau}_n^2\sum_{j=1}^{n}\Big\{\|U^j\|_1+\|U^{j-1}\|_1\Big\}\Bigg)\\
&\quad+\gamma\Bigg(\hat{\tau}_{n-1}\sum_{j=1}^{n-1}\{\alpha_j+\alpha_{j-1}\}+\hat{\tau}_n\sum_{j=1}^{n}\{\alpha_j+\alpha_{j-1}\}\\
&\quad+\hat{\tau}_{n-1}\sum_{j=1}^{n-1}\|U^j-U^{j-1}\|_1+\hat{\tau}_n\sum_{j=1}^{n}\|U^j-U^{j-1}\|_1\Bigg)\Bigg]\Bigg(\int_{t_{n-1}}^{t_n}\|\rho(t)\|_1^2dt\Bigg)^{1/2}\\
&\leq 1/2\ \tau_n \max_{t\in I_n}\|\rho(t)\|\ \|\mathcal{A}^{n-1}U^{n-1}-\sigma^{n-1}(\mathcal{B}^{n-1}U)-\mathcal{A}^nU^n+\sigma^n(\mathcal{B}^nU)\|\\
&\quad+\tau_n^{1/2}\Bigg[\frac{\gamma'}{2}\Bigg(\hat{\tau}_{n-1}^2\sum_{j=1}^{n-1}\{\alpha_j+\alpha_{j-1}\}+\hat{\tau}_n^2\sum_{j=1}^{n}\{\alpha_j+\alpha_{j-1}\}\\
&\quad+\hat{\tau}_{n-1}^2\sum_{j=1}^{n-1}\Big\{\|U^j\|_1+\|U^{j-1}\|_1\Big\}+\hat{\tau}_n^2\sum_{j=1}^{n}\Big\{\|U^j\|_1+\|U^{j-1}\|_1\Big\}\Bigg)\\
&\quad+2\gamma\sqrt{1/3}\Bigg(\hat{\tau}_n\sum_{j=1}^{n}\{\alpha_j+\alpha_{j-1}\}+\hat{\tau}_n^2\sum_{j=1}^{n}\|\partial U^j\|_1\Bigg)\Bigg]\Bigg(\int_{t_{n-1}}^{t_n}\|\rho(t)\|_1^2dt\Bigg)^{1/2}.
\end{aligned}
$$

In view of (35) and (37), we have

$$
\sum_{n=1}^{m}\mathcal{I}_n^2\leq\|\rho_*^m\|\sum_{n=1}^{m}\eta_n\tau_n+\sum_{n=1}^{m}\Bigg(\int_{t_{n-1}}^{t_n}\|\rho(t)\|_1^2dt\Bigg)^{1/2}\xi_n\tau_n^{1/2}.
$$

Mesh change estimates. The term $\mathcal{I}_n^3$ can be estimated using the orthogonality of the L^2-projection. Since $\mathbb{V}^n\subset\ker(P_0^n-I)$, we have

$$
\langle(P_0^n-I)(f^n+\tau_n^{-1}U^{n-1}),\phi_n\rangle=0,\ \ \forall\phi_n\in\mathbb{V}^n.
$$

Using Lemma 4.1 and Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
\mathcal{I}_n^3 &= \int_{t_{n-1}}^{t_n}|\langle(P_0^n-I)(f^n+\tau_n^{-1}U^{n-1}),\rho(t)-\Pi^n\rho(t)\rangle|dt\\
&\leq C_{11}h_n\int_{t_{n-1}}^{t_n}\|(P_0^n-I)(f^n+\tau_n^{-1}U^{n-1})\|\|\rho(t)\|_1dt\\
&\leq C_{11}h_n\tau_n^{1/2}\|(P_0^n-I)(f^n+\tau_n^{-1}U^{n-1})\|\Bigg(\int_{t_{n-1}}^{t_n}\|\rho(t)\|_1^2dt\Bigg)^{1/2}.
\end{aligned}
$$

Therefore, in view of (38) we have the following estimate,

$$
\sum_{n=1}^{m}\mathcal{I}_n^3\leq\sum_{n=1}^{m}\Bigg(\int_{t_{n-1}}^{t_n}\|\rho(t)\|_1^2dt\Bigg)^{1/2}\tau_n^{1/2}\mu_n.
$$

Data Oscillation estimates. We have

$$
\begin{aligned}
\mathcal{I}_n^4 &\leq \int_{t_{n-1}}^{t_n}\|f^n-f(t)\|\|\rho(t)\|dt\\
&\leq \Bigg(\max_{t\in I_n}\|\rho(t)\|\Bigg)\int_{t_{n-1}}^{t_n}\|f^n-f(t)\|dt.
\end{aligned}
$$

Thus, using (39) we obtain

$$\sum_{n=1}^{m} \mathcal{I}_n^4 \leq \|\rho_*^m\| \sum_{n=1}^{m} \lambda_n \tau_n.$$

Combining these estimates we arrive at

$$\begin{aligned}\|\rho_*^m\|^2 + \beta \int_0^{t_m} \|\rho(t)\|_1^2 dt \quad &\leq \quad \|\rho(t_0)\|^2 + 2C(t_m)\Bigg[\|\rho_*^m\| \sum_{n=1}^{m} (\zeta_n + \eta_n + \lambda_n)\tau_n \\ &+ \sum_{n=1}^{m} \left(\int_{t_{n-1}}^{t_n} \|\rho(t)\|_1^2 dt \right)^{1/2} \tau_n^{1/2} (\xi_n + \mu_n)\Bigg].\end{aligned}$$

To complete the proof of Lemma 5.3, we now use the following elementary fact. For $a = (a_0, a_1, \ldots, a_m)$, $b = (b_0, b_1, \ldots, b_m) \in \mathbb{R}^{m+1}$ and $c \in \mathbb{R}$, if

$$|a|^2 \leq c^2 + a.b,$$

then

$$|a| \leq |c| + |b|.$$

In particular for $n = [1 : m]$, taking

$$a_0 = \|\rho_*^m\|, \ \ a_n = \left(\beta \int_{t_{n-1}}^{t_n} \|\rho(t)\|_1^2 dt \right)^{1/2}, \quad c = \|\rho(t_0)\|,$$

$$b_0 = 2C(t_m) \sum_{n=1}^{m} (\zeta_n + \eta_n + \lambda_n)\tau_n, \quad b_n = 2C(t_m)(\tau_n/\beta)^{1/2}(\xi_n + \mu_n),$$

we obtained the required result.

The main results concerning fully discrete a posteriori error estimates in $L^\infty(L^2)$ and $L^2(H^1)$-norms are stated in the following theorem.

Theorem 5.4 *(Fully discrete a posteriori error estimates)*
For each $m \in [1 : N]$, the following error estimates hold:

$$\max_{[0,t_m]} \|u(t) - U(t)\| \leq \|R_w^0 U^0 - u(0)\| + \max_{n \in [0:m]} \beta_n + 2C(t_m)(\sigma_{1,m}^2 + \sigma_{1,m}^2)^{1/2},$$

$$\begin{aligned}\left(\int_0^{t_m} \|u(t) - U(t)\|_1^2 \right)^{1/2} \quad &\leq \quad \beta^{-1/2}\left[\|R_w^0 U^0 - u(0)\| + 2C(t_m)(\sigma_{1,m}^2 + \sigma_{2,m}^2)^{1/2}\right] \\ &+ \quad \left(\sum_{n=1}^{m} \tau_n \alpha_{n-1}^2 \right)^{1/2} + \left(\sum_{n=1}^{m} \tau_n \alpha_n^2 \right)^{1/2},\end{aligned}$$

where $\sigma_{1,m}^2$, $\sigma_{2,m}^2$ and $C(t_m)$ are defined as in Lemma 5.3.

Proof. We decompose the error with Ritz-Volterra reconstruction as an intermediate solution and obtain

$$\|u(t) - U(t)\| \leq \|\rho(t)\| + \|\epsilon(t)\|, \tag{46}$$

where $\rho(t) := \omega(t) - u(t)$ and $\epsilon(t) := \omega(t) - U(t)$.

Also, we know that for $t \in I_n$,

$$\|\epsilon(t)\| = \|l_{n-1}(t)\epsilon^{n-1} + l_n(t)\epsilon^n\| \le \max\Big(\|\epsilon^{n-1}\|, \|\epsilon^n\|\Big).$$

Therefore, for $t \in [0, t_m]$, using Lemma 5.1 and (34) we have

$$\|\epsilon(t)\| \le \max_{n\in[0,m]} \Big(\|\epsilon^{n-1}\|, \|\epsilon^n\|\Big) \le \max_{n\in[0,m]} \beta_n. \tag{47}$$

Then, the first estimate follows from (46), (47) and Lemma 5.3.

To prove the second estimate, using Lemma 5.1 and (33) we obtain

$$\begin{aligned}\left(\int_0^{t_m} \|\epsilon(t)\|_1^2\right)^{1/2} &= \left(\int_0^{t_m} \|l_{n-1}(t)\epsilon^{n-1} + l_n(t)\epsilon^n\|_1^2\right)^{1/2} \\ &\le \left(\sum_{n=1}^{m} \tau_n \alpha_{n-1}^2\right)^{1/2} + \left(\sum_{n=1}^{m} \tau_n \alpha_n^2\right)^{1/2}.\end{aligned}$$

The rest of the proof follows from Lemma 5.3.

Concluding Remarks. (i) It is known fact that a posteriori error estimators for parabolic problems in $L^\infty(L^2)$ and $L^2(H^1)$-norms are of optimal order [13]. Since PIDE (1) can be thought of as a perturbation to the parabolic problem, it is natural to expect that our a posteriori error estimators should reflect the contributions to the error from the approximation of the memory term. This fact can be easily observed through the estimator ξ_n which is of $O(\tau)$. Further, in the absence of the memory term (i.e., $\mathcal{B}(t,s) = 0$), the error estimators obtained in Theorem 5.4 are similar to that for the parabolic problems [13].

(ii) We know that the constants appearing in the a posteriori error bounds should be explicit or computable. For PIDE, the constants appeared in the bounds are time dependent due to the use of the Gronwall's lemma. However, the other constants (continuity constants, interpolation constants etc.) are computable. Since the final time T is finite, the constant appeared due to Gronwall's lemma will be at most $\exp(T)$ and thus, it is finite.

(iii) It is observed that the Ritz-Volterra projection is useful in apriori analysis for a wide range of (linear and nonlinear) parabolic and hyperbolic integro-differential problems. We strongly believe that, the Ritz-Volterra Reconstruction introduced in this paper, a counterpart of the Ritz-Volterra projection in the apriori analysis, can be appropriately modified to obtain estimators for a class of integro-differential problems.

(iv) The numerical computations of the proposed error bounds to study the behavior of the estimators through adaptive algorithms and to verify the optimality of our results is a challenging task which deserves attention and will be considered elsewhere. Also, we will address the problem of obtaining the a posteriori error estimates for the Crank-Nicolson scheme for parabolic integro-differential in near future.

References

[1] M. Ainsworth and J. T. Oden, *A posteriori error estimation in finite element analysis*, Wiley-Interscience, New York, 2000.

[2] G. Akrivis, C. Makridakis and R. H. Nochetto, *A posteriori error estimates for the Crank-Nicolson method for parabolic equations*, Math. Comp., 75 (2005), no. 254, pp. 511–531.

[3] S.C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, Springer-Verlag, New York, 2002.

[4] J. R. Cannon and Y.P. Lin, *A priori L^2 error estimates for finite-element methods for nonlinear diffusion equations with memory*, SIAM J. Numer. Anal., 27 (1990), pp. 595–607.

[5] V. Capasso, *Asymptotic stability for an integro-differential reaction-diffusion system*, J. Math. Anal. Appl., 103 (1984), pp. 575–588.

[6] J. de Frutos and J. Novo, *A posteriori error estimation with the p-version of the finite element method for nonlinear parabolic differential equations*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 4893–4904.

[7] K. Eriksson and C. Johnson, *An adaptive finite element method for linear Elliptic problem*, Math. Comp., 50 (1988), pp. 361–383.

[8] 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.

[9] J. de Frutos, B. García-Archilla and J. Novo, *A posteriori error estimates for fully discrete nonlinear parabolic problems*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 3462–3474.

[10] M.E. Gurtin and A. C. Pipkin, *A general theory of heat conduction with finite wave speeds*, Arch. Rational Mech. Anal., 31 (1968), pp. 113–126.

[11] G.J. Habetler and R.L. Schiffman, *A finite difference method for analysing the compression of poro-viscoelasticity media*, Comput., 6 (1970), pp. 342–348.

[12] W.E. Kastenberg and P.L. Chambre, *On the stability of nonlinear space dependent reactor kinetics*, Nucl. Sci. Eng., 31 (1968), pp. 67–79.

[13] O. Lakkis and C. Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp., 75 (2006), pp. 1627–1658.

[14] Y. Lin, V. Thomée and Lars B. Wahlbin *Ritz-Volterra projections to finite-element spaces and applications to integrodifferential and related equations*, SIAM J. Numer. Anal., 28 (1991), pp. 1047–1070.

[15] C. Makridakis and R. H. Nochetto, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 1585–1594.

[16] A.K. Pani and R.K. Sinha, *Error estimates for semidiscrete Galerkin approximation to a time dependent parabolic integro-differential equation with nonsmooth data*, Calcolo, 37 (2000), pp. 181–205.

[17] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[18] Simon Shaw and J.R. Whiteman, *Numerical solution of linear quasistatic hereditary viscoelasticity problems*, SIAM J. Numer. Anal., 38 (2000), pp. 80–97.

[19] Simon Shaw and J.R. Whiteman, *Negative norm error control for second-kind convolution Volterra equations*, Numer. Math., 85 (2000), pp. 329–341.

[20] V. Thomée and N.Y. Zhang, *Error estimates for semidiscrete finite element methods for parabolic integro-differential equations*, Math. Comp., 53 (1989), pp. 121–139.

[21] R. Verfürth, *A posteriori error estimates for finite element discretization of the heat equation*, Calcolo, 40 (2003), pp. 195–212.

[22] E.G. Yanik and G. Fairweather, *Finite element methods for parabolic and hyperbolic partial integro-differential equations*, Nonlinear Anal., 12 (1988), pp. 785–809.

[23] N. Y. Zhang, *On fully discrete Galerkin approximations for partial integro-differential equations of parabolic type*, Math. Comp., 60 (1993), pp. 133–166.