MIXED FINITE ELEMENT APPROXIMATIONS OF PARABOLIC INTEGRO-DIFFERENTIAL EQUATIONS WITH NONSMOOTH INITIAL DATA

RAJEN K. SINHA* RICHARD E. EWING† AND RAYTCHO D. LAZAROV‡

November 16, 2004

Abstract

We analyze the semidiscrete mixed finite element methods for parabolic integro-differential equations which arise in the modeling of nonlocal reactive flows in porous media. A priori L^2 error estimates for pressure and velocity are obtained with both smooth and nonsmooth initial data. More precisely, a mixed Ritz-Volterra projection, introduced earlier by Ewing *et. al.* in [SIAM J. Numer. Anal., 40 (2002), pp.1538-1560], is used to derive optimal L^2-error estimates for problems with initial data in $H^2 \cap H_0^1$. In addition, for homogeneous equations we derive optimal L^2-error estimates for initial data just in L^2. Here we use elementary energy technique and duality argument.

Key words. Parabolic integro-differential equation, mixed finite element method, semidiscrete, optimal error estimate, smooth and nonsmooth initial data.

AMS Subject Classifications. 65M12, 65M60, 65N40.

1 Introduction

In this paper, we consider mixed finite element approximations to the following initial-boundary value problem of the form

$$\begin{aligned} u_t - \nabla \cdot (A\nabla u) &= -\int_0^t \nabla \cdot (B(t,s)\nabla u(s))ds + f(x,t) \ \text{ in } \ \Omega \times J, \\ u &= 0 \ \text{ on } \ \partial\Omega \times J, \\ u(\cdot,0) &= u_0 \ \text{ in } \ \Omega, \end{aligned} \tag{1.1}$$

where Ω is a bounded domain in $\mathbb{R}^d$ $(d = 2,3)$ with smooth boundary $\partial\Omega$, $J = (0,T]$, $T < \infty$ and $u_t = \partial u/\partial t$, $A = \{a_{ij}(x)\}$ and $B(t,s) = \{b_{ij}(x;t,s)\}$ are two $d \times d$ matrices with smooth coefficients. Here, by ∇u we denote the gradient of a scalar function u and by $\nabla \cdot \sigma$ we denote the divergence of the vector function σ. Further, we assume that A is positive definite uniformly in Ω. The nonhomogeneous term $f = f(x,t)$ is assumed to be smooth. Equations of the above type arise naturally in many applications, such as in nonlocal reactive flows in porous media (cf. Cushman and Glinn [6] and Dagan [7]) and heat conduction through materials with memory (cf. Renardy *et*

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

†Department of Mathematics and Institute for Scientific Computation, Texas A&M University, College Station, TX 77843-3404, USA (richard-ewing@tamu.edu).

‡Department of Mathematics and ISC, Texas A&M University, College Station, TX 77843-3404, USA (lazarov@math.tamu.edu).

al. [19]). Flows of this type, called NonFickian flows (cf. Ewing *et al.* [10]), exhibit mixing length growth.

Now we give brief summary of the works regarding numerical methods for this type of problem using finite elements. Finite element approximation schemes of the problem (1.1) with smooth and nonsmooth initial data have been developed and studied quite intensively in the last decade (cf. [3, 4, 13, 15, 16, 17, 21]). The construction and the analysis of the proposed schemes use the standard tools of the finite element method and the Ritz-Volterra projection, introduced by Cannon and Lin in [3].

In [21], Thomeé and Zhang have studied this type of problem for both smooth and nonsmooth initial data. In particular, for a homogeneous equation with nonsmooth initial data, an optimal-order L^2-error estimate is proved via a semigroup theoretic approach. Subsequently, using energy method for the homogeneous equation Pani and Peterson in [16] showed convergence of the finite element approximations of order $O(t^{-1}h^2)$ in L^2-norm and $O\left(t^{-1}h^2\log(\frac{1}{h})\right)$ in L^∞-norm, when the initial data u_0 is in $H_0^1(\Omega)\cap H^2(\Omega)$. Recently, in [17], Pani and Sinha have carried over the analysis of Luskin and Rannacher [14] for parabolic equations (i.e. equation (1.1) with $B(t,s)=0$) to finite element approximations of time dependent integro-differential equation of parabolic type. They have proved optimal-order error estimates by an energy technique and a duality argument for the homogeneous equation with both smooth and nonsmooth initial data.

Often the problem (1.1) is reformulated by introducing a new dependent varialbe

$$\sigma(t) = A\nabla u - \int_0^t B(t,s)\nabla u(s)ds, \tag{1.2}$$

which in flow in porous media has a meaning of velocity field (or if properly scaled, mass flux). Then the equation $u_t - \nabla\cdot\sigma = f$ expresses a mass balance in any subdomain of Ω. The finite element method for this setting, called mixed formulation, gives direct approximation of the velocity field and the pressure at the same time, while maintaining the underlying local mass conservation. This property makes the mixed formulation more favorable for certain applications. In recent years, the analysis of mixed finite element method for such problems has been investigated in [9, 12, 10]. While the authors of [9] have discussed the general setting of the problem, the formulation and analysis described in [12] are valid for only a special case, namely, when the operator $B(t,s)$ is proportional to the operator A. More recently, Ewing et al. [10] have studied the problem (1.1) with when A depends on time and have derived sharp error estimates in L^2-norm for the velocity field and pressure. The analysis uses Ritz-Volterra projection instead of the mixed Ritz projection used earlier in [9]. In addition, local L^2 superconvergence for the velocity along the Gauss lines and for the pressure at the Gauss point are also derived for the mixed finite element method. In all these papers error estimates are obtained assuming high regularity on the solution which in turn demands high regularity on the initial function and the boundary of the domain.

It is well known that the solutions of a homogeneous linear parabolic equation have the so-called *smoothing property* (cf. [20]). That is, the solution is sufficiently smooth for positive time t, even when the initial data are not. Optimal error estimates for the pressure and the velocity by mixed finite element method of parabolic problems for smooth and nonsmooth initial data were derived in [5]. The results in [5] use the smoothing property of the parabolic equation to obtaine also superconvergence results for mixed finite element methods with nonsmooth data. Unfortunately, unlike parabolic equations, parabolic integro-differential equations have a limited smoothing property; e.g., when $u_0 \in L^2(\Omega)$ the solution can not have higher regularity than $H^2(\Omega)$, a fact established by Thomeé and Zhang in [21]. Further, the mathematical difficulty associated with the analysis of numerical approximations to the solution of (1.1) lies on the integral term when added to standard parabolic equations. Since (1.1) is an integral perturbation of a parabolic equations, it is natural to examine how far the mixed methods for parabolic problems [5] can be extended to the integro-differential equations.

The aim of this paper is to study the convergence of the approximate solutions of (1.1) by mixed finite element methods. First, we establish an optimal rate $O(h^2)$ in L^2-norm for the "smooth" case, namely, when the initial function $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$. These results rely on a mixed Ritz-Volterra projection introduced in [10] (instead of the standard Ritz projection). Here we were able to improve the results of [10] by reducing the smoothness of the initial data u_0 from H^3 to H^2.

The main goal of the paper, optimal estimates for nonsmooth data, namely, $u_0 \in L^2(\Omega)$, are considered in the last section of the paper. The first new result is the estimate (5.1) where we establish and optimal $O(ht^{-1})$ convergence rate of the velocity field and the same convergence rate for the pressure as well (see, Remark 5.2). However the convergence rate in the pressure is suboptimal. Unlike [5], our analysis does not use semigroup theoretic approach and is based only on relatively simple energy technique and duality argument. Unfortunately, we were not able to derive optimal estimate for the pressure in the generality of problem (1.1).

An optimal error estimate for the pressure is established for a class of problems when $A = a(t)I$ and $B = b(t)I$ and a and b are independent of the spacial variable x. In this case, we were able to apply duality argument and to show optimal convergence rate $O(h^2t^{-1})$ for the pressure.

The paper is organized as follows. In Section 2, we give the mixed setting of the problem (1.1) and prove some a priori estimates for the solution needed further in our analysis. The estimates related to mixed Ritz-Volterra projection are carried out in Section 3. Section 4 is devoted to the error estimates for smooth initial data. Finally, Section 5 deals with the error estimates with nonsmooth initial data.

2 Mixed finite element formulation and some a priori estimates

In this section, we introduce the mixed form of the problem (1.1) and prove some useful a priori estimates. In addition, we recall some known basic estimates for the solution.

To describe the weak mixed formulation, let $W = L^2(\Omega)$ be the L^2 space on Ω with standard inner product $(\cdot,\cdot)$ and norm $\|\cdot\|$. Let

$$V = H(div, \Omega) = \{\sigma \in (L^2(\Omega))^d : \ \nabla\cdot\sigma \in L^2(\Omega)\}$$

be the Hilbert space equipped with the norm $\|\sigma\| = (\|\sigma\|^2 + \|\nabla\cdot\sigma\|^2)^{\frac{1}{2}}$.

Following [9], we now recall the weak mixed formulation of (1.1) as follows: Find $(u,\sigma) \in W \times V$ such that

$$(u_t, w) - (\nabla\cdot\sigma, w) = (f, w) \quad \forall w \in W, \tag{2.1}$$

$$(\alpha\sigma, v) + \int_0^t (M(t,s)\sigma(s), v)ds + (\nabla\cdot v, u) = 0 \quad \forall v \in V. \tag{2.2}$$

with $u(x,0) = u_0(x)$. Here, $\alpha = A^{-1}$, $M(t,s) = R(t,s)A^{-1}$, and $R(t,s)$ is the resolvent of the matrix $A^{-1}B(t,s)$ and is given by

$$R(s,t) = A^{-1}B(t,s) + \int_s^t A^{-1}B(t,\tau)R(\tau,s)d\tau, \ \ t > s \geq 0.$$

Since the matrix A is positive definite then obviously there exist positive constants C_1 and C_2 such that

$$C_1\|\sigma\| \leq \|\sigma\|^2_{A^{-1}} \leq C_2\|\sigma\|, \quad \text{where} \quad \|\sigma\|^2_{A^{-1}} := (A^{-1}\sigma, \sigma). \tag{2.3}$$

Below, we shall prove some a priori estimates for u and σ satisfying (2.1) and (2.2). These estimates will be useful in our subsequent analysis.

Lemma 2.1 *Let (u, σ) satisfy (2.1)-(2.2) with $f = 0$ and let $0 \le i, j \le 2$.*
If $0 \le 2j - i \le 2$, then

$$t^i \left\| \frac{\partial^j u}{\partial t^j}(t) \right\|^2 \le C \|u_0\|^2_{2j-i} \quad and \quad \int_0^t s^i \left\| \frac{\partial^j \sigma}{\partial s^j}(s) \right\|^2 ds \le C \|u_0\|^2_{2j-i}. \tag{2.4}$$

Further, if $0 \le 2j - i - 1 \le 2$, then

$$t^i \left\| \frac{\partial^j \sigma}{\partial t^j}(t) \right\|^2 \le C \|u_0\|^2_{2j-i-1} \quad and \quad \int_0^t s^i \left\| \frac{\partial^j u}{\partial s^j}(s) \right\|^2 ds \le C \|u_0\|^2_{2j-i-1}. \tag{2.5}$$

Proof. For brevity, we shall refer to the first and second inequalities in (2.4) as $F_1(u; i, j)$ and $F_2(\sigma; i, j)$, respectively. Similarly, the first and second inequalities of (2.5) be denoted by $S_1(u; i, j)$ and $S_2(\sigma; i, j)$, respectively. Choose $w = u$ and $v = \sigma$ in (2.1) and (2.2), respectively. Then we obtain from their sum

$$\frac{1}{2} \frac{d}{dt} \|u\|^2 + \|\sigma\|^2_{A^{-1}} \le C \left(\int_0^t \|\sigma\| ds \right) \|\sigma\|.$$

Integrate from 0 to t. Then use (2.3) to get

$$\|u(t)\|^2 + \int_0^t \|\sigma(s)\|^2 ds \le C \left(\|u_0\|^2 + \int_0^t \int_0^s \|\sigma(\tau)\|^2 d\tau ds \right).$$

An application of Gronwall's lemma leads to the estimates $F_1(u; 0, 0)$ and $F_2(\sigma; 0, 0)$. Differentiate (2.2) with respect to time to have

$$(\alpha \sigma_t, v) + (M(t,t)\sigma(t) + \int_0^t M_t(t,s)\sigma(s), v) ds + (\nabla \cdot v, u_t) = 0, \quad \forall v \in V. \tag{2.6}$$

Here, $M_t(t, s)$ is obtained by differentiating $M(t, s)$ with respect to t. Taking $w = u_t$ and $v = \sigma$ in (2.1) and (2.6), respectively and noting the fact that $\|\sigma(0)\| \le C \|\nabla u_0\| \le C \|u_0\|_1$ we obtain $S_1(\sigma; 0, 0)$ and $S_2(u; 0, 1)$. Similarly, the choice $w = t u_t$ and $v = t\sigma$ in (2.1) and (2.6), respectively will lead to the estimates $S_1(\sigma; 1, 0)$ and $S_2(u; 1, 1)$. Next, differentiating (2.1) with respect to t we obtain for $f = 0$

$$(u_{tt}, w) - (\nabla \cdot \sigma_t, w) = 0. \tag{2.7}$$

Taking $w = t^2 u_t$ and $v = t^2 \sigma_t$ in (2.7) and (2.6), respectively we obtain from their sum

$$\frac{1}{2} \frac{d}{dt} \{t^2 \|u_t\|^2\} + t^2 \|\sigma_t\|^2_{A^{-1}} \le C t^2 \left(\|\sigma(t)\| + \int_0^t \|\sigma(s)\| ds \right) \|\sigma_t\| + t \|u_t\|^2.$$

Integration from 0 to t and a standard kickback argument leads to

$$t^2 \|u_t\|^2 + \int_0^t s^2 \|\sigma_s\|^2 \le C \left(t^2 \|\sigma(t)\|^2 + \int_0^t \{\|\sigma\|^2 + s \|u_s\|^2\} ds \right).$$

Use previously proved estimates $S_1(\sigma; 1, 0)$, $F_2(\sigma; 0, 0)$ and $S_2(u; 1, 1)$ to obtain $F_1(u; 2, 1)$ and $F_2(\sigma; 2, 1)$. The remaining cases will not be discussed in details, but the following table summerizes the necessary techniques: That is, the equations and the choice of w and v that would lead to the desired estimate.

Equations	w	v	Estimates
(2.1), (2.2)	u	σ	$F_1(u;0,0),\ F_2(\sigma;0,0)$
(2.1), $(2.2)^1$	u_t	σ	$S_1(\sigma;0,0),\ S_2(u;0,1)$
(2.1), $(2.2)^1$	tu_t	$t\sigma_t$	$S_1(\sigma;1,0),\ S_2(u;1,1)$
$(2.1)^1$, $(2.2)^1$	t^2u_t	$t^2\sigma_t$	$F_1(u;2,1),\ F_2(\sigma;2,1)$
$(2.1)^1$, $(2.2)^1$	tu_t	$t\sigma_t$	$F_1(u;1,1),\ F_2(\sigma;1,1)$
$(2.1)^1$, $(2.2)^2$	tu_{tt}	$t\sigma_t$	$S_1(\sigma;1,1),\ S_2(u;1,2)$
$(2.1)^1$, $(2.2)^2$	t^2u_{tt}	$t^2\sigma_t$	$S_1(\sigma;2,1),\ S_2(u;2,2)$
$(2.1)^2$, $(2.2)^2$	t^2u_{tt}	$t^2\sigma_{tt}$	$F_1(u;2,2),\ F_2(\sigma;2,2)$

Note that $(\,\cdot\,)^k$ is obtained by k times differentiating equation $(\,\cdot\,)$ with respect to t.

Lemma 2.2 *Let u satisfy (1.1) with $f=0$. If $u_0 \in H^2(\Omega)\cap H^1_0(\Omega)$, then*

$$\|u(t)\|_2^2 + t^2\|u_t\|_2^2 \le C\|u_0\|_2^2.$$

Further, if $u_0 \in L^2(\Omega)$, we have

$$t^2\|u(t)\|_2^2 + t^4\|u_t\|_2^2 \le C\|u_0\|^2, \quad t \in J.$$

Proof. For a proof, see [17] and [21].

Let T_h be a quasiuniform triangulation of Ω. Let $V_h \times W_h$ denote a pair of finite element spaces satisfying the follwing conditions:
(i) $\nabla\cdot V_h \subset W_h$, and
(ii) there exists a linear operator $\Pi_h : V \to V_h$ such that $\nabla\cdot\Pi_h = P_h\nabla$,
where $P_h : W \to W_h$ is the L^2-projection defined by

$$(\phi - P_h\phi, w_h) = 0, \quad \forall w_h \in W_h,\ \phi \in W.$$

Further, we shall assume that the finite element spaces satisfy the following approximation properties:

$$\|\sigma - \Pi_h\sigma\| \le Ch\|\sigma\|_1, \tag{2.8}$$

$$\|u - P_hu\| \le Ch^r\|u\|_r, \quad r = 1,2. \tag{2.9}$$

For examples of such finite element spaces, we refer to Raviart-Thomas [18], Brezzi, Douglas and Marini [1] and Brezzi and Fortin [2]. Note that Π_h and P_h satisfy

$$(\nabla\cdot(\sigma - \Pi_h\sigma), w_h) = 0, \quad w_h \in W_h; \quad (u - P_hu, \nabla\cdot v_h) = 0, \ \ v_h \in V_h. \tag{2.10}$$

Then the corresponding semidiscrete mixed finite element approximation is defined as follows: Find a pair $(u_h, \sigma_h) \in W_h \times V_h$ such that

$$(u_{h,t}, w_h) - (\nabla\cdot\sigma_h, w_h) = (f, w_h) \quad \forall w_h \in W_h, \tag{2.11}$$

$$(\alpha\sigma_h, v_h) + \int_0^t (M(s,t)\sigma_h(s), v_h)ds + (\nabla\cdot v_h, u_h) = 0 \quad \forall v_h \in V_h, \tag{2.12}$$

with $u_h(x,0) = u_{0,h}(x)$, where $u_{0,h}$ is a suitable approximation of the initial function $u_0(x)$ to be defined later. The pair (u_h, σ_h) is a semidiscrete approximation of the true solution of (1.1) in the finite element space $W_h \times V_h$, where $\sigma_h(x,0)$ is chosen to satisfy (2.12) with $t = 0$ and it is related to $u_{0,h}$ as follows:

$$(\alpha\sigma_h(0), v_h) + (u_{0,h}, \nabla\cdot v_h) = 0.$$

Throughout this paper C denotes a generic positive constant which does not depend on the mesh parameter h but may depend on T.

3 Mixed Ritz-Volterra projection and its properties

Following [10], we now define mixed Ritz-Volterra projection as the pair $(\tilde{u}_h, \tilde{\sigma}_h) : [0,T] \to W_h \times V_h$ such that

$$(\alpha(\sigma - \tilde{\sigma}_h, v_h)) + \int_0^t (M(t,s)(\sigma - \tilde{\sigma}_h)(s), v_h) ds + (\nabla \cdot v_h, u - \tilde{u}_h) = 0, \;\; v_h \in V_h, \tag{3.1}$$

$$(\nabla \cdot (\sigma - \tilde{\sigma}_h), w_h) = 0, \;\; w_h \in W_h. \tag{3.2}$$

Set $\rho = (u - \tilde{u}_h)$ and $\eta = (\sigma - \tilde{\sigma}_h)$. We now rewrite the equations (3.1)-(3.2) as

$$(\alpha\eta, v_h) + \int_0^t (M(t,s)\eta(s), v_h)\, ds + (\nabla \cdot v_h, \rho) = 0, \quad v_h \in V_h, \tag{3.3}$$

$$(\nabla \cdot \eta, w_h) = 0, \quad w_h \in W_h. \tag{3.4}$$

From [10] (see, Theorems 2.5-2.6), we now recall the following estimates of ρ and η

$$\|\rho(t)\| + h\|\eta(t)\| \leq Ch^2 \left(\|u(t)\|_2 + \int_0^t \|u(s)\|_2 ds \right) \tag{3.5}$$

and

$$\|\rho_t\| \leq Ch^2 \left(\|u\|_2 + \|u_t\|_2 + \int_0^t (\|u(s)\|_2 + \|u_t(s)\|_2) ds \right). \tag{3.6}$$

Note that the estimate of ρ_t contains a term $\|u_t\|_2$ under the integral sign. However, an inspection of the proof (cf. [10]) shows that it is not necessary. The elimination of this term is very crucial for the nonsmooth data error estimates. In order to analyze this we need the following result which is a particular case of [8] (cf. Lemma 3.1).

Lemma 3.1 *Let the index of $V_h \times W_h$ be at least one. Assume that Ω is 2-regular [8]. Let $\eta \in V$, $g \in L^2(\Omega)$ and $f = \{f_0, f_1\}$ with $f_0 \in (L^2(\Omega))^2$, $f_1 \in L^2(\Omega)$ and*

$$f(v) = (f_0, v) + (f_1, \nabla \cdot v), \quad v \in V.$$

If $z \in W_h$ satisfies the relations

$$\begin{aligned} (\alpha\eta, v_h) + (\nabla \cdot v_h, z) &= f(v_h), \quad v_h \in V_h, \\ (\nabla \cdot \eta, w_h) + (cz, w_h) &= g(w_h), \quad w_h \in W_h, \end{aligned}$$

then there exists $h_0 > 0$ sufficiently small such that, for all $0 < h \leq h_0$,

$$\|z\| \leq C \left(h\|\eta\| + h^2\|\nabla \cdot \eta\| + \|f_0\|_{-1} + h\|f_0\| + \|f_1\| + \|g\|_{-2} + h^2\|g\| \right).$$

Instead of (3.6) we prove the following result.

Lemma 3.2 *Let $(\tilde{u}_h, \tilde{\sigma}_h)$ be the mixed Ritz-Volterra projection of $(u, \sigma) \in W \times V$ defined by (3.1)-(3.2). Then, for small h, there is a positive constant C independent of h such that*

$$\|(u - \tilde{u}_h)_t\| \equiv \|\rho_t(t)\| \leq Ch^2 \left(\|u(t)\|_2 + \|u_t(t)\|_2 + \int_0^t \|u(s)\|_2 ds \right).$$

Proof. We borrow the proof technique from [10]. We first split ρ_t as

$$\rho_t = (u_t - P_h u_t) + \tau_{h,t}, \tag{3.7}$$

where $\tau_h = (P_h u - \tilde{u}_h)$. We now estimate $\|\tau_{h,t}\|$. Differentiating (3.3)-(3.4) with respect to time t to have

$$(\alpha\eta_t, v_h) + (\nabla \cdot v_h, \rho_t) = -\left(M(t,t)\eta(t) + \int_0^t M_t(t,s)\eta(s)ds, v_h \right), \; v_h \in V_h, \tag{3.8}$$

$$(\nabla \cdot \eta_t, w_h) = 0, \;\; w_h \in W_h. \tag{3.9}$$

We now apply Lemma 3.1 to (3.8)-(3.9) with $c \equiv 0$, $f_1 \equiv 0$,

$$f(v_h) = -\left(M(t,t)\eta(t) + \int_0^t M_t(s,t)\eta(s)ds, v_h \right), \;\text{ and }\; g \equiv 0.$$

Since

$$\|f\| \leq C\left(\|\eta\| + \int_0^t \|\eta\| ds \right) \;\text{ and }\; \|f\|_{-1} \leq C\left(\|\eta\|_{-1} + \int_0^t \|\eta\|_{-1} ds \right),$$

we obtain

$$\begin{aligned} \|\tau_{h,t}\| &\leq C\left\{ h\|\eta_t\| + h^2\|\nabla \cdot \eta_t\| + \|f\|_{-1} + h\|f\| \right\} \\ &\leq C\left\{ h\|\eta_t\| + h^2\|\nabla \cdot \eta_t\| + (\|\eta\|_{-1} + h\|\eta\|) + \int_0^t (\|\eta\|_{-1} + h\|\eta\|)ds \right\}. \end{aligned} \tag{3.10}$$

It follows from ([10, p.1544]) that

$$\|\eta\|_{-1} \leq C\left\{ \|\rho\| + Ch(\|\eta\| + \int_0^t \|\eta(s)\|ds) \right\}. \tag{3.11}$$

A substitution of (3.11) into (3.10) yields

$$\|\tau_{h,t}\| \leq C\left\{ h\|\eta_t\| + h(\|\eta\| + \int_0^t \|\eta(s)\|ds) + h^2\|\nabla \cdot \eta_t\| + \|\rho(t)\| + \int_0^t \|\rho(s)\|ds \right\},$$

which together with (3.7), the triangle inequality, and the estimate of $\|\rho\|$ leads to

$$\|\rho_t(t)\| \;\; \leq \;\; Ch\left\{ \|\eta_t\| + \|\eta\| + \int_0^t \|\eta(s)\|ds + h\|\nabla \cdot \eta_t\| + h(\|u\|_2 + \int_0^t \|u\|_2 ds) \right\}. \tag{3.12}$$

Since the estimate of $\|\eta\|$ is already known, it remains to estimate the terms $\|\nabla \cdot \eta_t\|$ and $\|\eta_t\|$. In view of (2.10), (3.2) and (3.9), it is easy to see that

$$\begin{aligned} \|\nabla \cdot \eta_t\|^2 &= (\nabla \cdot (\sigma_t - \tilde{\sigma}_{h,t}), \nabla \cdot (\sigma_t - \tilde{\sigma}_{h,t})) \\ &= (\nabla \cdot (\sigma_t - \tilde{\sigma}_{h,t}), \nabla \cdot (\sigma_t - \Pi_h \sigma_t)) \leq C\|\nabla \cdot \sigma_t\|\|\nabla \cdot \eta_t\|, \end{aligned}$$

so we get

$$\|\nabla \cdot \eta_t\| \leq C\|\sigma_t\|_1. \tag{3.13}$$

Next, to estimate $\|\eta_t\|$, we note that

$$\|\eta_t\| \leq \|\Pi_h \sigma_t - \tilde{\sigma}_{h,t}\| + \|\Pi_h \sigma_t - \sigma_t\| \leq C(\|\psi_{h,t}\| + h\|\sigma_t\|_1). \tag{3.14}$$

where $\psi_h = \Pi_h\sigma - \tilde{\sigma}_h$. For the estimation of $\|\psi_{h,t}\|$, we first differentiate (3.1) with respect to t to get

$$
\begin{aligned}
&\left(\alpha\psi_{h,t} + M(t,t)\psi_h + \int_0^t M_t(t,s)\psi_h(s)ds, \psi_{h,t}\right) \\
&= \left(\alpha\eta_t + M(t,t)\eta_h + \int_0^t M_t(t,s)\eta_h(s)ds, \psi_{h,t}\right) \\
&\quad + \left(\alpha(\Pi_h\sigma_t - \sigma_t) + M(t,t)(\Pi_h\sigma - \sigma) + \int_0^t M_t(t,s)(\Pi_h\sigma - \sigma)(s)ds, \psi_{h,t}\right) \\
&= -(\nabla\cdot\psi_{h,t}, \rho_t) + \left(\alpha(\Pi_h\sigma_t - \sigma_t) + M(t,t)(\Pi_h\sigma - \sigma) + \int_0^t M_t(t,s)(\Pi_h\sigma - \sigma)(s)ds, \psi_{h,t}\right).
\end{aligned}
$$

Then we apply Cauchy-Schwarz inequality to have

$$
\begin{aligned}
\|\psi_{h,t}\|^2 \quad \leq \quad & C\left(\|\psi_h\| + \int_0^t \|\psi_h\|ds\right)\|\psi_{h,t}\| + \|\nabla\cdot\psi_{h,t}\|\|\rho_t\| \\
& + C\left(\|\Pi_h\sigma_t - \sigma_t\| + \|(\Pi_h\sigma - \sigma)\| + \int_0^t \|(\Pi_h\sigma - \sigma)(s)\|ds\right)\|\psi_{h,t}\|.
\end{aligned}
$$

Kickback $\|\psi_{h,t}\|$ to obtain

$$
\begin{aligned}
\|\psi_{h,t}\| \quad \leq \quad & C\left(\|\psi_h\| + \int_0^t \|\psi_h\|ds + \|\nabla\cdot\psi_{h,t}\| + \|\rho_t\|\right) \\
& + C\left(\|\Pi_h\sigma_t - \sigma_t\| + \|(\Pi_h\sigma - \sigma)\| + \int_0^t \|(\Pi_h\sigma - \sigma)(s)\|ds\right). \qquad (3.15)
\end{aligned}
$$

Note that $\|\nabla\cdot\psi_{h,t}\| = 0$ and it follows from [10] (see, page 1545) that

$$
\|\psi_h\| \leq C(\|\rho\| + h(\|\sigma\|_1 + \int_0^t \|\sigma\|_1 ds)). \qquad (3.16)
$$

Putting (3.16) into (3.15) we have

$$
\|\psi_{h,t}\| \quad \leq \quad C\left\{\|\rho\| + \|\rho_t\| + h\left(\|\sigma_t\|_1 + \|\sigma\|_1 + \int_0^t \|\sigma\|_1 ds\right)\right\},
$$

and this combined with (3.14) yields

$$
\|\eta_t\| \leq C\left\{\|\rho\| + \|\rho_t\| + h(\|\sigma_t\|_1 + \|\sigma\|_1 + \int_0^t \|\sigma\|_1 ds)\right\}. \qquad (3.17)
$$

Finally, using (3.17), (3.13) and the estimate of $\|\eta\|$ in (3.12), for small h we obtain

$$
\|\rho_t\| \leq Ch^2\left\{\|u\|_2 + \|u_t\|_2 + \int_0^t \|u\|_2 + \left(\|\sigma_t\|_1 + \|\sigma\|_1 + \int_0^t \|\sigma\|_1 ds\right)\right\},
$$

which completes the proof. □

4 L^2-error estimates with smooth initial data

In this section, we shall derive optimal L^2-error estimates for the solutions u and σ assuming the initial function $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$. Using the mixed Ritz-Volterra projection $(\tilde{u}_h, \tilde{\sigma}_h)$ we first write the errors as

$$e_1(t) := u - u_h = (u - \tilde{u}_h) + (\tilde{u}_h - u_h) := \rho + \rho_h, \tag{4.1}$$

$$e_2(t) := \sigma - \sigma_h = (\sigma - \tilde{\sigma}_h) + (\tilde{\sigma}_h - \sigma_h) := \eta + \theta_h. \tag{4.2}$$

Since the estimates of ρ and η are already known it is enough to have estimates for ρ_h and θ_h. Using (2.1), (2.2), (2.11), (2.12), (3.1) and (3.2), we note that (ρ_h, θ_h) satisfies the following error equations

$$(\alpha\theta_h, v_h) + \int_0^t (M(t,s)\theta_h(s), v_h)ds) + (\nabla \cdot v_h, \rho_h) \;=\; 0, \quad v_h \in V_h, \tag{4.3}$$

$$(\rho_{h,t}, w_h) - (\nabla \cdot \theta_h, w_h) \;=\; -(\rho_t, w_h), \;\; w_h \in W_h. \tag{4.4}$$

For a function ϕ defined on $[0,T]$, we define $\hat{\phi}(t)$ as

$$\hat{\phi}(t) = \int_0^t \phi(\tau)d\tau.$$

Clearly, $\hat{\phi}(0) = 0$ and $\hat{\phi}_t(t) = \phi(t)$. Integrate (3.3) and (3.4) from 0 to t to get

$$(\alpha\hat{\eta}, v_h) + \int_0^t (M(s,s)\hat{\eta}(s), v_h) - \int_0^t\int_0^s (M_\tau(s,\tau)\hat{\eta}(\tau), v_h)d\tau ds + (\nabla \cdot v_h, \hat{\rho}) \;=\; 0, \tag{4.5}$$

$$(\nabla \cdot \hat{\eta}, w_h) \;=\; 0, \tag{4.6}$$

satisfied for $v_h \in V_h$ and $w_h \in W_h$, respectively Similarly, integrate equations (2.1), (2.2), (2.11) and (2.12) from 0 to t. Then using the resulting equations, (4.5), (4.6) and $u_h(0) = P_h u_0$, it is easy to verify that $(\hat{\rho}_h, \hat{\theta}_h)$ satisfies the following equations

$$(\alpha\hat{\theta}_h, v_h) + \int_0^t (M(s,s)\hat{\theta}_h(s), v_h)ds - \int_0^t\int_0^s (M_\tau(s,\tau)\hat{\theta}_h(\tau), v_h)d\tau + (\nabla \cdot v_h, \hat{\rho}_h) = 0, \tag{4.7}$$

$$(\hat{\rho}_{h,t}, w_h) - (\nabla \cdot \hat{\theta}_h, w_h) = -(\rho, w_h) \tag{4.8}$$

with $v_h \in V_h$ and $w_h \in W_h$.

Now we state the main results of this section.

Theorem 4.1 *Let (u, σ) and (u_h, σ_h) be the solutions of (2.1)-(2.2) and (2.11)-(2.12), respectively with $f = 0$. Further, let $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$ and $u_h(0) = P_h u_0$. Then there is a positive generic constant C independent of h such that*

$$\|u(t) - u_h(t)\| \le Ch^2 \|u_0\|_2, \tag{4.9}$$

and

$$\|\sigma(t) - \sigma_h(t)\| \le Ct^{-1/2} h \|u_0\|_2, \quad t \in J \tag{4.10}$$

hold true.

The proof requires some preparatory results that are established below in a sequence of lemmas.

Lemma 4.1 *Let $(\hat{\rho}_h, \hat{\theta}_h)$ satisfy (4.7)-(4.8) and $u_h(0) = P_h u_0$. Then there is a positive constant C independent of h such that*

$$\|\hat{\rho}_h\|^2 + \int_0^t \|\hat{\theta}_h\|^2 ds \le C \int_0^t \|\rho\|^2 ds.$$

Proof. Set $w_h = \hat{\rho}_h$ and $v_h = \hat{\theta}_h$ in (4.7) and (4.8), respectively. Then sum the resulting equations to have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\hat{\rho}_h\|^2 + \|\hat{\theta}_h\|^2_{A^{-1}} &= -\left(\int_0^t M(s,s)\hat{\theta}_h(s)ds - \int_0^t\int_0^s M_\tau(s,\tau)\hat{\theta}_h(\tau)d\tau, \hat{\theta}_h\right) - (\rho,\hat{\rho}_h)\\ &\leq C\left(\int_0^t \{\|\hat{\theta}_h(s)\| + \int_0^s \|\hat{\theta}_h(\tau)\|d\tau\}ds\right)\|\hat{\theta}\| + \|\rho\|\|\hat{\rho}_h\|.\end{aligned}$$

In view of (2.3) we obtain

$$\frac{1}{2}\frac{d}{dt}\|\hat{\rho}_h\|^2 + \frac{C_1}{2}\|\hat{\theta}_h\|^2 \leq C\left(\|\hat{\rho}_h\|^2 + \|\rho\|^2 + \int_0^t \|\hat{\theta}_h(s)\|^2 ds\right)$$

Integrating from 0 to t, we have

$$\|\hat{\rho}_h\|^2 + \int_0^t \|\hat{\theta}_h\|^2 ds \leq C\left(\int_0^t \{\|\hat{\rho}_h\|^2 + \int_0^s \|\hat{\theta}_h(\tau)\|^2 d\tau\}ds + C\int_0^t \|\rho\|^2 ds\right)$$

An application of Gronwall's lemma completes the rest of the proof. □

Lemma 4.2 *Let the hypotheses in Lemma 4.1 hold true. Then there is a positive constant C independent of h such that*

$$\|\hat{\theta}_h(t)\|^2 + \int_0^t \|\rho_h(s)\|^2 ds \leq C\int_0^t \|\rho(s)\|^2 ds.$$

Proof. Setting $w_h = \rho_h$ and $v_h = \hat{\theta}_h$ in (4.8) and (4.3), respectively. Then we obtain from their sum

$$\begin{aligned}\|\rho_h\|^2 + \frac{1}{2}\frac{d}{dt}\|\hat{\theta}_h\|^2_{A^{-1}} &= -\left(M(t,t)\hat{\theta}_h(t) - \int_0^t M_s(t,s)\hat{\theta}_h(s)ds, \hat{\theta}_h\right) - (\rho,\rho_h)\\ &\leq C\left(\|\hat{\theta}_h(t)\| + \int_0^t \|\hat{\theta}_h(s)\|ds\right)\|\hat{\theta}\| + \|\rho\|\|\rho_h\|.\end{aligned}$$

Kickback the term $\|\rho_h\|$ to have

$$\|\rho_h\|^2 + \frac{1}{2}\frac{d}{ds}\|\hat{\theta}_h\|^2_{A^{-1}} \leq C\left(\|\hat{\theta}_h(t)\|^2 + \int_0^t \|\hat{\theta}_h(s)\|^2 ds\right) + C\|\rho\|^2.$$

Integrating from 0 to t and further using (2.3) and Lemma 4.1 the desired estimate is easily obtained. This completes the rest of the proof. □

Lemma 4.3 *Let (ρ_h, θ_h) satisfy (4.3), (4.4) and $u_h(0) = P_h u_0$. Then there is a positive constant C independent of h such that*

$$t\|\rho_h\|^2 + \int_0^t s\|\theta_h\|^2 ds \leq C\int_0^t (\|\rho(s)\|^2 + s^2\|\rho_s(s)\|^2)ds.$$

Proof. Choose $w_h = t\rho_h$ and $v_h = t\theta_h$ in (4.4) and (4.3), respectively. Then sum the resulting equations to have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\{t\|\rho_h\|^2\} + t\|\theta_h\|^2_{A^{-1}} &= -\left(M(t,t)\hat{\theta}_h(t) - \int_0^t M_s(t,s)\hat{\theta}_h(s)ds, t\theta_h\right) - t(\rho_t,\rho_h)\\ &\leq C\left(\|\hat{\theta}_h(s)\| + \int_0^t \|\hat{\theta}_h(s)\|ds\right)t\|\theta_h\| + t\|\rho_t\|\|\rho_h\|.\end{aligned}$$

By (2.3) and kicking back $t\|\theta_h\|$ it now follows that

$$\frac{d}{dt}\{t\|\rho_h\|^2\} + t\|\theta_h\|^2 \leq C\left(\|\rho_h\|^2 + \|\hat{\theta}_h(t)\|^2 + \int_0^t \|\hat{\theta}_h(s)\|^2 ds + t^2\|\rho_t\|^2\right).$$

Integration from 0 to t leads to

$$t\|\hat{\rho}_h\|^2 + \int_0^t s\|\theta_h\|^2 ds \leq C\left(\int_0^t \{\|\rho_h\|^2 + s^2\|\rho_s\|^2 + \|\hat{\theta}_h\|^2 + \int_0^s \|\hat{\theta}_h(\tau)\|^2 d\tau\} ds\right).$$

An application of Lemma 4.1 and Lemma 4.2 completes the rest of the proof. □

In order to obtain an estimate for θ_h we differentiate (4.3) with respect to t to have

$$(\alpha\theta_{h,t}, v_h) + (M(t,t)\theta_h, v_h) + \int_0^t (M_t(t,s)\theta_h(s), v_h) ds, +(\nabla\cdot v_h, \rho_{h,t}) = 0, \;\; v_h \in V_h. \tag{4.11}$$

Lemma 4.4 *Let the hypotheses in Lemma 4.3 hold true. Then there is a positive constant C independent of h such that*

$$\int_0^t s^2\|\rho_{h,s}\|^2 ds + t^2\|\theta_h(t)\|^2 \leq C\int_0^t (\|\rho\|^2 + s^2\|\rho_s\|^2) ds.$$

Proof. Setting $v_h = t^2\theta_h$ in (4.11) and $w_h = t^2\rho_{h,t}$ in (4.4). Then we obtain from their sum

$$\begin{aligned} t^2\|\rho_{h,t}\|^2 \quad &+ \frac{1}{2}\frac{d}{dt}\{t^2\|\theta_h\|^2_{A^{-1}}\} = -t^2(M(t,t)\theta_h, \theta_h) + (M_t(t,t)\hat{\theta}(t), \theta_h) \\ &- \int_0^t (M_{ts}(t,s)\hat{\theta}_h(s), \theta_h) ds + t\|\theta_h\|^2_{A^{-1}} - t^2(\rho_t, \rho_{h,t}). \end{aligned}$$

Integrating from 0 to t and use standard kickback argument to have

$$\begin{aligned} \int_0^t s\|\rho_{h,s}\|^2 ds + t^2\|\theta_h(t)\|^2 \quad &\leq \quad C\left(\int_0^t s\|\theta_h\|^2 ds + \int_0^t \int_0^s \|\hat{\theta}_h(s)\|^2 d\tau ds\right) \\ &+ C\int_0^t s^2\|\rho_s\|^2 ds + C\int_0^t s^2\|\theta_h(s)\|^2 ds. \end{aligned}$$

Finally, use Lemma 4.1, Lemma 4.3 and Gronwall's lemma to complete the rest of the proof. □

Proof of Theorem 4.1. By triangle inequality, we have

$$\|u(t) - u_h(t)\| := \|e_1(t)\| \leq \|\rho(t)\| + \|\rho_h\|.$$

For the first term on the right of the above inequality, we use (3.5) and Lemma 2.2 to have

$$\|\rho(t)\| \leq Ch^2\left(\|u\|_2 + \int_0^t \|u\|_2 ds\right) \leq Ch^2\|u_0\|_2.$$

Using Lemma 4.3, Lemma 3.2, (3.5) and Lemma 2.2, it now follows that

$$t\|\rho_h\|^2 \leq C\left(\int_0^t \{\|\rho\|^2 + s^2\|\rho_s\|^2\}\right) \leq Ch^4\left(\int_0^t \{\|u\|_2^2 + s^2\|u_s\|_2^2\}\right) \leq Ch^4 t\|u_0\|_2^2.$$

Altogether these estimates proves the first statement of the theorem. Combining (3.5), Lemma 3.2, Lemma 4.4 and Lemma 2.2, the second statement is easily obtained and this completes the proof. □

Remark 4.1. (i) Note that Theorem 4.1 yields optimal order of convergence assuming $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$. Compared to [10](see, Theorem 3.1) the results presented in Theorem 4.1 require less regularity assumption on the initial function u_0. In [10], one requires u_0 to be atleast in $H^3(\Omega)$.

(ii) In contrast to [10], we do not require the assumptions

$$\|P_h u_0 - u_h(0)\| \leq Ch^2 \|u_0\|_2 \text{ and } \|\Pi_h \sigma(0) - \sigma_h(0)\| \leq Ch\|u_0\|_2$$

but $u_h(0) = P_h u_0$ suffices for the present analysis.

5 L^2-error estimates with nonsmooth initial data

This section is devoted to the error estimates for the semidiscrete Galerkin method for the homogeneous equation (1.1) with nonsmooth initial data. In particular, for homogeneous equations, optimal order error estimates for the solutions are shown to hold assuming $u_0 \in L^2(\Omega)$. First objective of this section is to prove the following theorem.

Theorem 5.1 *Let (u, σ) and (u_h, σ_h) be the solutions of (2.1), (2.2) and (2.11), (2.12), respectively with $f = 0$. Assume that $u_0 \in L^2(\Omega)$. Then*

$$\|\sigma(t) - \sigma_h(t)\| \leq Ct^{-1} h \|u_0\|, \quad t \in J. \tag{5.1}$$

The proof of the above theorem require some preparations. For this purpose we shall first establish a sequence of lemmas which will lead to the desired result. Using (2.1), (2.2), (2.11) and (2.12), we obtain the following error equations

$$(e_{1,t}, w_h) - (\nabla \cdot e_2, w_h) = 0, \;\; \forall w_h \in W_h, \tag{5.2}$$

$$(\alpha e_2, v_h) + \int_0^t (M(t,s) e_2(s), v_h) ds + (\nabla \cdot v_h, e_1) = 0, \;\; \forall v_h \in V_h. \tag{5.3}$$

Lemma 5.1 *Assume that $u_0 \in L^2(\Omega)$ and $u_h(0) = P_h u_0$. Then we have*

$$\|P_h \hat{u} - \hat{u}_h\|^2 + \int_0^t \|\Pi_h \hat{\sigma} - \hat{\sigma}_h\|^2 \leq Cth^2 \|u_0\|^2.$$

Proof. Integrating (5.2), (5.3) with respect to t and using the fact that $u_h(0) = P_h u_0$, we get

$$(e_1, w_h) \quad - \quad (\nabla \cdot \hat{e}_2, w_h) = 0, \tag{5.4}$$

$$(\alpha \hat{e}_2, v_h) + \int_0^t (M(s,s) \hat{e}_2(s), v_h) ds \quad - \quad \int_0^t \int_0^s (M_\tau(s,\tau) \hat{e}_2(\tau), v_h) d\tau ds, +(\nabla \cdot v_h, \hat{e}_1) = 0. \tag{5.5}$$

Choose $w_h = P_h \hat{u} - \hat{u}_h$ and $v_h = \Pi_h \hat{\sigma} - \hat{\sigma}_h$ in (5.4) and (5.5), respectively and add these equalities. Then using (2.10) we obtain

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\{\|P_h \hat{u} - \hat{u}_h\|^2\} \quad + \quad & \|\Pi_h \hat{\sigma} - \hat{\sigma}_h\|^2_{A^{-1}} \leq C\|\hat{\sigma} - \Pi_h \hat{\sigma}\| \|\Pi_h \hat{\sigma} - \hat{\sigma}_h\| \\ & + C\left(\int_0^t (\hat{\sigma} - \|\Pi_h \hat{\sigma}\| + \int_0^s \|\hat{\sigma} - \Pi_h \hat{\sigma}\| d\tau) ds\right) \|\Pi_h \hat{\sigma} - \hat{\sigma}_h\| \\ & + C\left(\int_0^t (\|\Pi_h \hat{\sigma} - \hat{\sigma}_h\| + \int_0^s \|\Pi_h \hat{\sigma} - \hat{\sigma}_h\| d\tau) ds\right) \|\Pi_h \hat{\sigma} - \hat{\sigma}_h\|.\end{aligned}$$

Apply (2.3), kickback $\|\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|$ and then integrate from 0 to t to get

$$\|P_h\hat{u}-\hat{u}_h\|^2+\int_0^t\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|^2ds\le C\int_0^t\|\hat{\sigma}-\Pi_h\hat{\sigma}\|^2ds+C\int_0^t\int_0^s\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|^2d\tau ds.$$

With an aid of (2.8) and Gronwall's lemma, it follows that

$$\|P_h\hat{u}-\hat{u}_h\|^2+\int_0^t\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|^2ds\le Ch^2\int_0^t\|\hat{\sigma}\|_1^2ds. \tag{5.6}$$

Now it remains to estimate $\|\hat{\sigma}\|_1$. Integrating (1.2) by parts we have

$$\sigma(t)=A\nabla u-\int_0^t B(t,s)\nabla\hat{u}_s(s)ds=A\nabla u-B(t,t)\hat{u}+\int_0^t B_s(t,s)\nabla\hat{u}(s)ds,$$

and hence

$$\|\hat{\sigma}\|_1\le C\left(\|\hat{u}(t)\|_2+\int_0^t\|\hat{u}(s)\|_2ds\right). \tag{5.7}$$

From (1.1) with $f=0$, we have

$$\begin{aligned}-\nabla\cdot(A\nabla u) &= -u_t-\int_0^t\nabla\cdot(B(t,s)\nabla\hat{u}_s(s))ds\\ &= -u_t-\nabla\cdot(B(t,t)\nabla\hat{u}(t))+\int_0^t\nabla\cdot(B_s(t,s)\nabla\hat{u}(s))ds.\end{aligned}$$

Integrating from 0 to t and then using elliptic regularity and Lemma 2.1 we obtain

$$\|\hat{u}\|_2\le\|u_0\|+\|u(t)\|+C\int_0^t\|\hat{u}\|_2ds\le C\|u_0\|+C\int_0^t\|\hat{u}\|_2ds.$$

Now application of Gronwall's lemma yields

$$\|\hat{u}\|_2\le C\|u_0\|. \tag{5.8}$$

Now combine (5.6), (5.7) and (5.8) to complete the proof. □

Lemma 5.2 *Assume that $u_0\in L^2(\Omega)$ and $u_h(0)=P_hu_0$. Then we have*

$$t\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|^2+\int_0^t s\|P_hu-u_h\|^2ds\le Cth^2\|u_0\|^2.$$

Proof. Taking $w_h=t(P_hu-u_h)$ and $v_h=t(\Pi_h\hat{\sigma}-\hat{\sigma}_h)$ in (5.4) and (5.3), respectively. Then using (2.10) we obtain from their sum

$$\begin{aligned}t\|P_hu-u_h\|^2 &+ \frac{1}{2}\frac{d}{dt}\{t\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|_{A^{-1}}^2\}\le Ct\|\sigma-\Pi_h\sigma\|\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|+\frac{1}{2}\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|_{A^{-1}}^2\\ &+C\left(\|\hat{\sigma}-\Pi_h\hat{\sigma}\|+\int_0^t\|\hat{\sigma}-\Pi_h\hat{\sigma}\|ds\right)(t\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|)\\ &+C\left(\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|+\int_0^t\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|ds\right)(t\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|).\end{aligned}$$

Again use of (2.3) and integration from 0 to t now leads to

$$\begin{aligned} t\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|^2+\int_0^t s\|P_hu-u_h\|^2ds &\leq C\left(\int_0^t s^2\|\sigma-\Pi_h\sigma\|^2ds+\int_0^t \|\Pi_h\hat{\sigma}-\hat{\sigma}\|^2ds\right) \\ &\quad +C\int_0^t \|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|^2ds. \end{aligned}$$

Here it is important to emphasize that the multiplication by s^2 is very crucial for the estimation of the first term of the above inequality. We apply Lemma 5.1, (2.8), a priori estimate in Lemma 2.2 and (5.8) to complete the rest of the proof. □

Lemma 5.3 *Assume that $u_0 \in L^2(\Omega)$ and $u_h(0) = P_hu_0$. Then there is a positive constant C independent of h such that*

$$t^2\|P_hu-u_h\|^2+\int_0^t s^2\|\Pi_h\sigma-\sigma_h\|^2ds \leq Cth^2\|u_0\|^2.$$

Proof. Setting $w_h = t^2(P_hu-u_h)$ and $v_h = t^2(\Pi_h\sigma-\sigma_h)$ in (5.2) and (5.3), respectively. Then using (2.10) we obtain from their sum

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\{t^2\|P_hu-u_h\|^2\} &+ t^2\|\Pi_h\sigma-\sigma_h\|^2_{A^{-1}} \leq Ct^2\|\sigma-\Pi_h\sigma\|\|\Pi_h\sigma-\sigma_h\|+t\|P_hu-u_h\|^2 \\ &+C\left(\|\hat{\sigma}-\Pi_h\hat{\sigma}\|+\int_0^t\|\hat{\sigma}-\Pi_h\hat{\sigma}\|ds\right)(t^2\|\Pi_h\sigma-\sigma_h\|) \\ &+C\left(\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|+\int_0^t\|\Pi_h\hat{\sigma}-\hat{\sigma}_h\|ds\right)(t^2\|\Pi_h\sigma-\sigma_h\|). \end{aligned}$$

Integrate from 0 to t and then use standard kickback argument to have

$$\begin{aligned} t^2\|P_hu-u_h\|^2 &+ \int_0^t s^2\|\Pi_h\sigma-\sigma_h\|^2ds \leq C\left(\int_0^t s^2\|\sigma-\Pi_h\sigma\|^2ds\right. \\ &\left.+\int_0^t s\|P_hu-u_h\|^2ds+\int_0^t\|\Pi_h\hat{\sigma}-\hat{\sigma}\|^2ds\right). \end{aligned}$$

An application of (2.8), Lemmas 5.1-5.2, a priori estimates in Lemma 2.2 and (5.8) yield the desired estimate and this completes the proof. □

Lemma 5.4 *With $u_0 \in L^2(\Omega)$ and $u_h(0) = P_hu_0$, we have*

$$t^3\|\Pi_h\sigma-\sigma_h\|^2+\int_0^t s^3\|P_hu_t-u_{h,t}\|^2ds \leq Cth^2\|u_0\|^2.$$

Proof. Differentiate (5.3) with respect to t to have

$$(\alpha e_{2,t}(t),v_h)+(M(t,t)e_2(t),v_h)+\int_0^t(M_t(t,s)e_2(s),v_h)ds+(\nabla\cdot v_h,e_{1,t}(t))=0. \tag{5.9}$$

Setting $w_h = t^3(P_h u_t - u_{h,t})$ and $v_h = t^3(\Pi_h\sigma - \sigma_h)$ in (5.2) and (5.9), respectively and using (2.10) we obtain from their sum

$$\begin{aligned} t^3\|P_h u_t - u_{h,t}\|^2 \quad + \quad & \frac{1}{2}\frac{d}{dt}\{t^3\|\Pi_h\sigma - \sigma_h\|^2_{A^{-1}}\} \leq Ct^3\|\sigma_t - \Pi_h\sigma_t\|\|\Pi_h\sigma - \sigma_h\| + \frac{3}{2}t^2\|\Pi_h\sigma - \sigma_h\|^2_{A^{-1}} \\ & + Ct^3\|\sigma - \Pi_h\sigma\|\|\Pi_h\sigma - \sigma_h\| + Ct^3\|\Pi_h\sigma - \sigma_h\|^2 \\ & + C\left(\|\hat\sigma - \Pi_h\hat\sigma\| + \int_0^t \|\hat\sigma - \Pi_h\hat\sigma\|ds\right)(t^3\|\Pi_h\sigma - \sigma_h\|) \\ & + C\left(\|\Pi_h\hat\sigma - \hat\sigma_h\| + \int_0^t \|\Pi_h\hat\sigma - \hat\sigma_h\|ds\right)(t^3\|\Pi_h\sigma - \sigma_h\|). \end{aligned}$$

Here, we have used the identity

$$\int_0^t (M_t(t,s)e_2(s), v_h)ds = (M_t(t,t)\hat e_2(t), v_h) - \int_0^t M_{ts}(t,s)\hat e_2(s), v_h)ds.$$

and use the same argument as in Lemma 5.3 to get

$$\begin{aligned} t^3\|\Pi_h\sigma - \sigma_h\|^2 \quad + \quad & \int_0^t s^3\|P_h u_t - u_{h,t}\|^2 ds \leq C\left(\int_0^t s^4\|\sigma_t - \Pi_h\sigma_t\|^2 ds\right. \\ & + \int_0^t s^2\|\Pi_h\sigma - \sigma_h\|^2 ds + \int_0^t s^3\|\sigma - \Pi_h\sigma\|^2 ds \\ & \left. + \int_0^t \|\Pi_h\hat\sigma - \hat\sigma\|^2 ds\right) + C\int_0^t s^3\|\Pi_h\sigma - \sigma_h\|^2 ds. \end{aligned}$$

Apply (2.8), Lemma 5.1, Lemma 5.3, a priori estimates in Lemma 2.2 and (5.8) to obtain

$$t^3\|\Pi_h\sigma - \sigma_h\|^2 + \int_0^t s^3\|P_h u - u_{h,t}\|^2 ds \leq Cth^2\|u_0\|^2 + C\int_0^t s^3\|\Pi_h\sigma - \sigma_h\|^2 ds.$$

Finally, an application of Gronwall's lemma completes the rest of the proof. □

Lemma 5.5 *Assume that $u_0 \in L^2(\Omega)$ and $u_h(0) = P_h u_0$. Then there is a positive constant C independent of h such that*

$$t^4\|P_h u_t - u_{h,t}\|^2 + \int_0^t s^4\|\Pi_h\sigma - \sigma_h\|^2 ds \leq Cth^2\|u_0\|^2.$$

Proof. Differentiate (5.2) with respect to t and set $w_h = t^4(P_h u_t - u_{h,t})$ in the resulting equation and $v_h = t^4(\Pi_h\sigma_t - \sigma_{h,t})$ in (5.9). A similar argument as before now leads to

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\{t^4\|P_h u_t - u_{h,t}\|^2\} \quad + \quad & t^4\|\Pi_h\sigma_t - \sigma_{h,t}\|^2_{A^{-1}} \leq Ct^4\|\sigma_t - \Pi_h\sigma_t\|\|\Pi_h\sigma_t - \sigma_{h,t}\| \\ & + 2t^3\|P_h u_t - u_{h,t}\|^2 + C\{\|\sigma - \Pi_h\sigma\| + \|\Pi_h\sigma - \sigma_h\|\}(t^4\|\Pi_h\sigma_t - \sigma_{h,t}\|) \\ & + C\left(\|\hat\sigma - \Pi_h\hat\sigma\| + \int_0^t \|\hat\sigma - \Pi_h\hat\sigma\|ds\right)(t^4\|\Pi_h\sigma_t - \sigma_{h,t}\|) \\ & + C\left(\|\Pi_h\hat\sigma - \hat\sigma_h\| + \int_0^t \|\Pi_h\hat\sigma - \hat\sigma_h\|ds\right)(t^4\|\Pi_h\sigma_t - \sigma_{h,t}\|). \end{aligned}$$

Integrate from 0 to t and then use standard kickback argument to have

$$\begin{aligned} t^4\|P_h u_t - u_{h,t}\|^2 \quad + \quad & \int_0^t s^4 \|\Pi_h \sigma_t - \sigma_{h,t}\|^2 ds \le C \left(\int_0^t s^4 \|\sigma_t - \Pi_h \sigma_t\|^2 ds \right. \\ & + \int_0^t s^3 \|P_h u_t - u_{h,t}\|^2 ds + \int_0^t s^4 \|\sigma - \Pi_h \sigma\|^2 ds \\ & \left. \int_0^t s^4 \|\Pi_h \sigma - \sigma_h\|^2 ds + \int_0^t \|\Pi_h \hat{\sigma} - \hat{\sigma}\|^2 ds \right). \end{aligned}$$

Application of (2.8), Lemma 5.1, Lemma 5.3, Lemma 5.4, a priori estimates in Lemma 2.2 and (5.8) to obtain desired result and this completes the rest of the proof. □

Remark 5.1. Note that Lemma 5.3 and Lemma 5.5 yield the following estimates

$$\|P_h u - u_h\| \le C h t^{-1/2} \|u_0\| \tag{5.10}$$

and

$$\|P_h u_t - u_{h,t}\| \le C h t^{-3/2} \|u_0\|. \tag{5.11}$$

In case of purely parabolic problem (i.e., $B(t,s) = 0$), similar estimates are derived in [cf. Lemmas 7-8, 5] via semigroup theoretic approach. In contrast to [5], the present analysis uses only elementary energy technique.

Define $\hat{e}_2(t) = \int_0^t e_2(s) ds$. In order to derive optimal L^2-error estimate for e_2, we first prove the following result.

Lemma 5.6 *Assume that $u_0 \in L^2(\Omega)$. Then there is a positive constant C such that*

$$\|\hat{e}_2(t)\| \le C h \|u_0\|.$$

Proof. By triangle inequality, we have

$$\|\hat{e}_2(t)\| \le \|\hat{\sigma} - \Pi_h \hat{\sigma}\| + \|\Pi_h \hat{\sigma} - \hat{\sigma}_h\|.$$

Now use Lemma 5.2, (2.8) and (5.8) to obtain the desired estimate which completes the proof.
□

Proof of Theorem 5.1. With $v_h = \Pi_h e_2 \equiv \Pi_h \sigma - \sigma_h$, we obtain using (5.3)

$$\begin{aligned} (\alpha e_2, e_2) &= (\alpha e_2, e_2 - \Pi_h e_2) + (\alpha e_2, \Pi_h e_2) \\ &= (\alpha e_2, e_2 - \Pi_h e_2) - \int_0^t (M(t,s) e_2(s), \Pi_h e_2 - e_2) ds - (e_1, \nabla \cdot (\Pi_h e_2 - e_2)) \\ & \quad - \int_0^t (M(t,s) e_2(s), e_2) ds - (e_1, \nabla \cdot e_2) \\ &= (\alpha e_2, \sigma - \Pi_h \sigma) - \int_0^t (M(t,s) e_2(s), \Pi_h \sigma - \sigma) ds - (e_1, \nabla \cdot (\Pi_h \sigma - \sigma)) \\ & \quad - \int_0^t (M(t,s) e_2(s), e_2) ds - (e_1, \nabla \cdot e_2). \end{aligned} \tag{5.12}$$

Using the definition of P_h, we note that

$$\begin{aligned} -\{(e_1, \nabla \cdot (\Pi_h \sigma - \sigma)) \quad + \quad & (e_1, \nabla \cdot e_2)\} = -(u - u_h, \nabla \cdot (\Pi_h \sigma - \sigma_h)) \\ = \quad & -(P_h u - u_h, \nabla \cdot (\Pi_h \sigma - \sigma_h)). \end{aligned}$$

and hence, using (2.1) and (2.11), we obtain

$$\begin{aligned}(P_hu - u_h, \nabla\cdot(\Pi_h\sigma - \sigma_h)) &= (P_hu - u_h, \nabla\cdot(\Pi_h\sigma - \sigma)) + (P_hu - u_h, \nabla\cdot e_2)\\ &= (u_t - u_{h,t}, P_hu - u_h).\end{aligned}$$

The remaining terms in (5.12) are discussed below. Further, integrating by parts we have

$$\begin{aligned}-\int_0^t (M(t,s)e_2(s), \Pi_h\sigma - \sigma)ds &= -\int_0^t (M(t,s)\hat{e}_{2,s}(s), \Pi_h\sigma - \sigma)ds\\ &= -(M(t,t)\hat{e}_2(t), \Pi_h\sigma - \sigma) + \int_0^t (M_s(t,s)\hat{e}_2(s), \Pi_h\sigma - \sigma)ds.\end{aligned}$$

Similarly,

$$-\int_0^t (M(t,s)e_2(s), e_2)ds = -(M(t,t)\hat{e}_2(t), e_2) + \int_0^t (M_s(t,s)\hat{e}_2(s), e_2)ds.$$

Putting these together and using (2.3), we get from (5.12)

$$\begin{aligned}\|e_2\|^2 \leq\ & C\left(\|e_2\| + \int_0^t \|\hat{e}_2(s)\|ds\right)\|\sigma - \Pi_h\sigma\| + C\left(\|\hat{e}_2\| + \int_0^t \|\hat{e}_2(s)\|ds\right)\|e_2\|\\ &+\|u_t - u_{h,t}\|\|P_hu - u_h\|.\end{aligned}$$

Kicking back $\|e_2\|$ and using Lemma 5.6 we get

$$\|e_2(t)\|^2 \leq C\left(\|\sigma - \Pi_h\sigma\|^2 + \|P_hu - u_h\|\|u_t - u_{h,t}\| + h^2\|u_0\|^2\right).$$

Finally, an use of (2.8), (2.9), a priori estimates in Lemmas 2.1 -2.2, (5.10) and (5.11) completes the rest of the proof. ☐

Remark 5.2. As a consequence of (5.10), (2.9) and Lemma 2.1 it is easy to obtain

$$\|u(t) - u_h(t)\| \leq Cht^{-1/2}\|u_0\|, \quad t \in J. \tag{5.13}$$

Note that the estimate (5.13) is not optimal with respect to the approximation property. One should expect $O(h^2t^{-1})$ order of convergence. However, under certain assumptions on the coefficient matrices it is possible to obtain optimal order of convergence with $u_0 \in L^2(\Omega)$. For this purpose we now consider the following backward problem: For fixed $t > 0$, find $(p(s), \zeta(s)) \in W \times V$ such that

$$(p_s, w) + (\nabla\cdot\zeta, w) = 0 \quad \forall w \in W,\ s < t, \tag{5.14}$$

$$(\alpha\zeta, v) + \int_s^t (M^*(\tau,s)\zeta(\tau), v)d\tau + (\nabla\cdot v, p) = 0 \quad \forall v \in V,\ s < t, \tag{5.15}$$

$$p(t) = g.$$

Here,

$$\zeta(s) = A\nabla p(s) - \int_s^t B^*(\tau,s)\nabla p(\tau)d\tau,$$

$\alpha = A^{-1}$. The matrices $M^*(\tau,s)$ and $B^*(\tau,s)$ denote transposed of $M(\tau,s)$ and $B(\tau,s)$, respectively. The corresponding semidiscrete version seeks a pair $(p_h(s), \zeta_h(s)) \in W_h \times V_h$ such that

$$(p_{h,s}, w_h) + (\nabla\cdot\zeta_h, w_h) = 0 \quad \forall w_h \in W_h, \quad s < t, \tag{5.16}$$

$$(\alpha\zeta_h, v_h) + \int_s^t (M^*(\tau,s)\zeta_h(\tau), v_h)d\tau + (\nabla\cdot v_h, p_h) = 0 \quad \forall v_h \in V_h,\ s < t, \tag{5.17}$$

$$p_h(t) = P_hg.$$

Using (2.1), (2.2), (2.11), (2.12) with $f=0$ and (5.14)-(5.17) we obtain

$$\begin{aligned}
\frac{d}{ds}\{(u,p) &- (u_h,p_h)\} = \{(\nabla\cdot\sigma,p)-(\nabla\cdot\zeta,u)\} - \{(\nabla\cdot\sigma_h,p_h)-(\nabla\cdot\zeta_h,u_h)\} \\
&= -\int_s^t (M^*(\tau,s)\zeta(\tau),\sigma(s))d\tau + \int_0^s (M(s,\tau)\sigma(\tau),\zeta(s))d\tau \\
&\quad + \int_s^t (M^*(\tau,s)\zeta_h(\tau),\sigma_h(s)d\tau - \int_0^s (M(s,\tau)\sigma_h(\tau),\zeta_h(s))d\tau \qquad (5.18)
\end{aligned}$$

The following two lemmas are proved to be convenient for error estimates with nonsmooth initial data. In the lemma below, we first establish the negative norm estimate for $e_1 = u - u_h$.

Lemma 5.7 *Let (u,σ) and (u_h,σ_h) be the solutions of (2.1)-(2.2) and (2.11)-(2.12), respectively with $f=0$. Assume that $u_0 \in L^2(\Omega)$. Then*

$$\|u(t)-u_h(t)\|_{-2} \le Ch^2\|u_0\|.$$

Proof. Integrate the identity (5.18) from 0 to t. Using the fact that

$$\int_0^t\int_0^s (M(s,\tau)\phi(\tau),\psi(s))d\tau ds = \int_0^t\int_s^t (M^*(\tau,s)\psi(\tau),\phi(s))d\tau ds,$$

we get

$$(u(t),p(t)) - (u_h(t),p_h(t)) = (u_0,p(0)) - (u_h(0),p_h(0)).$$

With $u_h(0) = P_h u_0$ and $g_h = P_h g$, we have

$$(e_1(t),g) = (u_0,(p-p_h)(0)).$$

Applying estimate (4.9) of Theorem 4.1 to the backward problem with $g \in H^2(\Omega)\cap H^1_0(\Omega)$ we obtain

$$(e_1(t),g) \le \|u_0\|\|(p-p_h)(0)\| \le Ch^2\|u_0\|\|g\|_2,$$

and this completes the proof. □

We now state our second main result of this section in the following theorem.

Theorem 5.2 *Let (u,σ) and (u_h,σ_h) be the solutions of (2.1)-(2.2) and (2.11)-(2.12), respectively with $f=0$. Further, assume that $u_0 \in L^2(\Omega)$ and*

$$A = aI \quad \text{and} \quad B = b(t,s)I,$$

where a is a positive constant and $b(t,s)$ is a scalar function of s and t. Then the following estimate

$$\|u(t)-u_h(t)\| \le Ct^{-1}h^2\|u_0\|, \quad t\in J,$$

holds true.

Proof. For a function $\psi(s)$ defined on $[s,t]$, we set

$$\tilde\psi(s) = -\int_s^t \psi(\tau)d\tau, \quad s\le t.$$

Note that $\tilde{\psi}(t) = 0$ and $\tilde{\psi}_s(s) = \psi(s)$. Analogous to estimate (5.8) it is easy to show that the solution $p(s)$ of the backward problem (5.14)-(5.15) satisfies

$$\|\tilde{p}(s)\|_2 \leq C\|g\|. \tag{5.19}$$

With $\bar{e}_1(s) = p(s) - p_h(s)$ and $\bar{e}_2(s) = \zeta(s) - \zeta_h(s)$, integrate (5.18) from $\frac{t}{2}$ to t to obtain

$$\begin{aligned}
(e_1(t), g) &= (e_1(t/2), p(t/2)) - (e_1(t/2), \bar{e}_1(t/2)) + (u(t/2), \bar{e}_1(t/2)) \\
&\quad + \int_0^{\frac{t}{2}} \int_{\frac{t}{2}}^t (M(\tau, s)e_2(s), \zeta(\tau))d\tau ds - \int_0^{\frac{t}{2}} \int_{\frac{t}{2}}^t (M(\tau, s)e_2(s), \bar{e}_2(\tau))d\tau ds \\
&\quad + \int_0^{\frac{t}{2}} \int_{\frac{t}{2}}^t (M(\tau, s)\sigma(s), \bar{e}_2(\tau))d\tau ds \\
&=: I_1 + I_2 + I_3 + I_4 + I_5 + I_6.
\end{aligned}$$

We now proceed to estimate each term separately. For the term I_1, apply Lemma 5.7 and *a priori* estimates for the backward problem to have

$$|I_1| = |(e_1(t/2), p(t/2))| \leq \|e_1(t/2)\|_{-2}\|p(t/2)\|_2 \leq Ch^2t^{-1}\|u_0\|\|g\|.$$

Apply (5.13) to the backward error $\bar{e}_1$ to have

$$\|\bar{e}_1(s)\| \leq Ch(t-s)^{-1/2}\|g\|. \tag{5.20}$$

Using (5.13) and (5.20), I_2 can be estimated as

$$|I_2| = |(e_1(t/2), \bar{e}_1(t/2))| \leq \|e(t/2)\|\|\bar{e}_1(t/2)\| \leq Ch^2t^{-1}\|u_0\|\|g\|.$$

For I_3, we apply Lemma 5.7 to the backward problem to obtain

$$\|\bar{e}_1(s)\|_{-2} \leq Ch^2\|g\|. \tag{5.21}$$

Thus, using (5.21) and Lemma 2.2 it now follows that

$$|I_3| = |(u(t/2), \bar{e}_1(t/2))| \leq \|u(t/2)\|_2\|\bar{e}_1(t/2)\|_{-2} \leq Ch^2t^{-1}\|u_0\|\|g\|.$$

To estimate the remaining terms we first note the following: Since the matrices $A = aI$ and $B(t,s) = b(t,s)I$ are independent of x, we set $\zeta(s) = \nabla w(s)$, so that

$$\begin{aligned}
w(s) &= Ap(s) - \int_s^t B(\tau, s)p(\tau)d\tau \\
&= Ap(s) + B(s,s)\tilde{p}(s) + \int_s^t B_\tau(\tau, s)\tilde{p}(\tau)d\tau,
\end{aligned}$$

where $\tilde{p}(s) = -\int_s^t p(\tau)d\tau$. Using (5.19) it is easy to verify that

$$\|\tilde{w}(s)\|_2 \leq C\left(\|\tilde{p}(s)\|_2 + \int_s^t \|\tilde{p}(\tau)\|_2 d\tau\right) \leq C\|g\|. \tag{5.22}$$

Now to estimate I_4, we first rewrite it as (recall that $\zeta(s) = \nabla w(s)$ and $M(t,s)$ is a scalar function)

$$\begin{aligned}
I_4 &= -\int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (M(\tau,s)\nabla\cdot(\sigma-\sigma_h)(s), w(\tau))d\tau ds \\
&= -\int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (\nabla\cdot(\sigma-\sigma_h)(s), M(\tau,s)(w-P_h w)(\tau))d\tau ds \\
&\quad -\int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (\nabla\cdot(\sigma-\sigma_h)(s), M(\tau,s)P_h w(\tau))d\tau ds \\
&= -\int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (\nabla\cdot(\sigma-\sigma_h)(s), M(\tau,s)(w-P_h w)(\tau))d\tau ds \\
&\quad -\int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (e_{1,s}(s), M(\tau,s)(P_h w - w)(\tau))d\tau ds - \int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (e_{1,s}(s), M(\tau,s)w(\tau))d\tau ds \\
&=: I_4^1 + I_4^2 + I_4^3.
\end{aligned}$$

Here in the last step we have used (5.2). We now proceed to estimate each term separately. For I_4^1, we use the definition of P_h operator and integration by parts formula to have

$$\begin{aligned}
I_4^1 &= -\int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (\nabla\cdot(\hat\sigma-\Pi_h\hat\sigma)_s(s), M(\tau,s)(\tilde w - P_h\tilde w)_\tau(\tau))d\tau ds \\
&= \int_0^{\frac{t}{2}} (\nabla\cdot(\hat\sigma-\Pi_h\hat\sigma)_s(s), M(t/2,s)(\tilde w - P_h\tilde w)(t/2))ds \\
&\quad +\int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (\nabla\cdot(\hat\sigma-\Pi_h\hat\sigma)_s(s), M_\tau(\tau,s)(\tilde w - P_h\tilde w)(\tau))d\tau ds \\
&= (\nabla\cdot(\hat\sigma-\Pi_h\hat\sigma)(\frac{t}{2}), M(t/2,t/2)(\tilde w - P_h\tilde w)(t/2)) \\
&\quad -\int_0^{\frac{t}{2}} (\nabla\cdot(\hat\sigma-\Pi_h\hat\sigma)(s), M_s(t/2,s)(\tilde w - P_h\tilde w)(t/2))ds \\
&\quad +\int_{\frac{t}{2}}^t (\nabla\cdot(\hat\sigma-\Pi_h\hat\sigma)(t/2), M_\tau(\tau,t/2)(\tilde w - P_h\tilde w)(\tau))d\tau \\
&\quad -\int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (\nabla\cdot(\hat\sigma-\Pi_h\hat\sigma)(s), M_{\tau,s}(\tau,s)(\tilde w - P_h\tilde w)(\tau))d\tau ds.
\end{aligned}$$

Here, we have used to fact that $\hat\sigma(0) = 0$ and $\tilde w(t) = 0$. Using (2.9), a priori estimates (5.7), (5.8) and (5.22) it now follows that

$$\begin{aligned}
|I_4^1| &\le Ch^2\Bigg(\|\hat\sigma(t/2)\|_1\|\tilde w(t/2)\|_2 + \int_0^{\frac{t}{2}} \|\hat\sigma(s)\|_1\|\tilde w(t/2)\|_2 ds \\
&\qquad \int_{\frac{t}{2}}^t \|\hat\sigma(t/2)\|_1\|\tilde w(\tau)\|_2 d\tau + \int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t \|\hat\sigma(s)\|_1\|\tilde w(\tau)\|_2 d\tau ds\Bigg) \\
&\le Ch^2\|u_0\|\|g\|.
\end{aligned}$$

Similarly, for I_4^2, integration by parts formula leads to

$$\begin{aligned}
I_4^2 &= -\int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (e_{1,s}(s), M(\tau,s)(\tilde{w}-P_h\tilde{w})_\tau(\tau))d\tau ds \\
&= \int_0^{\frac{t}{2}} (e_{1,s}(s), M(t/2,s)(\tilde{w}-P_h\tilde{w})(t/2))ds + \int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (e_{1,s}(s), M_\tau(\tau,s)(\tilde{w}-P_h\tilde{w})(\tau))d\tau ds \\
&= (e_1(t/2), M(t/2,t/2)(\tilde{w}-P_h\tilde{w})(t/2)) - (e_1(0), M(t/2,0)(\tilde{w}-P_h\tilde{w})(t/2)) \\
&\quad - \int_0^{\frac{t}{2}} (e_1(s), M_s(t/2,s)(\tilde{w}-P_h\tilde{w})(t/2))ds + \int_{\frac{t}{2}}^t (e_1(t/2), M_\tau(\tau,t/2)(\tilde{w}-P_h\tilde{w})(\tau))d\tau \\
&\quad - \int_{\frac{t}{2}}^t (e_1(0), M_\tau(\tau,0)(\tilde{w}-P_h\tilde{w})(\tau))d\tau - \int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (e_1(s), M_{\tau,s}(\tau,s)(\tilde{w}-P_h\tilde{w})(\tau))d\tau ds.
\end{aligned}$$

Using the fact that $u_h(0) = P_h u_0$ and applying (2.9), Lemma 2.1 and (5.22) we obtain

$$\begin{aligned}
|I_4^2| &\leq Ch^2 \Bigg(\{\|e_1(t/2)\| + \|e_1(0)\|\}\|\tilde{w}(t/2)\|_2 + \int_0^{\frac{t}{2}} \|e_1(s)\|\|\|\tilde{w}(t/2)\|_2 ds \\
&\qquad + \int_{\frac{t}{2}}^t \{\|e_1(t/2)\| + \|e_1(0)\|\}\|\tilde{w}(\tau)\|_2 d\tau + \int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t \|e_1(s)\|\|\tilde{w}(\tau)\|_2 d\tau ds \Bigg). \\
&\leq Ch^2\|u_0\|\|g\|.
\end{aligned}$$

As before integrating by parts we rewrite the term I_4^3 as

$$\begin{aligned}
I_4^3 &= -\int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (e_{1,s}(s), M(\tau,s)\tilde{w}_\tau(\tau))d\tau ds \\
&= \int_0^{\frac{t}{2}} (e_{1,s}(s), M(t/2,s)\tilde{w}(t/2))ds + \int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (e_{1,s}(s), M_\tau(\tau,s)\tilde{w}(\tau))d\tau ds \\
&= (e_1(t/2), M(t/2,t/2)\tilde{w}(t/2)) - (e_1(0), M(t/2,0)(\tilde{w}-P_h\tilde{w})(t/2)) \\
&\quad - \int_0^{\frac{t}{2}} (e_1(s), M_s(t/2,s)\tilde{w}(t/2))ds + \int_{\frac{t}{2}}^t (e_1(t/2), M_\tau(\tau,t/2)\tilde{w}(\tau))d\tau \\
&\quad - \int_{\frac{t}{2}}^t (e_1(0), M_\tau(\tau,0)(\tilde{w}-P_h\tilde{w})(\tau))d\tau - \int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t (e_1(s), M_{\tau,s}(\tau,s)\tilde{w}(\tau))d\tau ds.
\end{aligned}$$

Here, in the last step we have used the definition of P_h operator. Now using Lemma 5.7 and a priori estimate in (5.22) we obtain

$$\begin{aligned}
|I_4^3| &\leq C \Bigg(\{\|e_1(t/2)\|_{-2} + h^2\|u_0\|\}\|\tilde{w}(t/2)\|_2 + \int_0^{\frac{t}{2}} \|e_1(s)\|_{-2}\|\tilde{w}(t/2)\|_2 ds \\
&\qquad + \int_{\frac{t}{2}}^t \{\|e_1(t/2)\|_{-2} + h^2\|u_0\|\}\|\tilde{w}(\tau)\|_2 d\tau + \int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^t \|e_1(s)\|_{-2}\|\tilde{w}(\tau)\|_2 d\tau ds \Bigg). \\
&\leq Ch^2\|u_0\|\|g\|.
\end{aligned}$$

Hence,

$$|I_4| \leq Ch^2\|u_0\|\|g\|.$$

The term I_6 is estimated in a manner similar to I_4 and hence, we get

$$|I_6| \le Ch^2\|u_0\|\|g\|.$$

Finally, it remains to estimate the term I_5. Again integtating by parts we obtain

$$\begin{aligned}
I_5 &= -\int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^{t}(M(\tau,s)\hat{e}_{2,s}(s),\tilde{\bar{e}}_{2,\tau}(\tau))d\tau ds \\
&= \int_0^{\frac{t}{2}}(M(t/2,s)\hat{e}_{2,s}(s),\tilde{\bar{e}}_2(t/2))ds + \int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^{t}(M_\tau(\tau,s)\hat{e}_{2,s}(s),\tilde{\bar{e}}_2(\tau))d\tau ds \\
&= (M(t/2,t/2)\hat{e}_2(t/2),\tilde{\bar{e}}_2(t/2)) - \int_0^{\frac{t}{2}}(M_s(t/2,s)\hat{e}_2(s),\tilde{\bar{e}}_2(t/2))ds \\
&\quad + \int_{\frac{t}{2}}^{t}(M_\tau(\tau,t/2)\hat{e}_2(t/2),\tilde{\bar{e}}_2(\tau))d\tau - \int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^{t}(M_{\tau,s}(\tau,s)\hat{e}_2(s),\tilde{\bar{e}}_2(\tau))d\tau ds.
\end{aligned}$$

Before estimating the term I_5 we note the following. Analogous to Lemma 5.6, we obtain (with time reverse)

$$\|\tilde{\bar{e}}_2(s)\| \le Ch\|g\|. \tag{5.23}$$

Applying Lemma 5.6 and (5.23) it now follows that

$$\begin{aligned}
|I_5| &\le C\Bigg(\|\hat{e}_2(t/2)\|\|\tilde{\bar{e}}_2(t/2)\| + \int_0^{\frac{t}{2}}\|\hat{e}_2(s)\|\|\tilde{\bar{e}}_2(t/2)\|ds \\
&\qquad + \int_{\frac{t}{2}}^{t}\|\hat{e}_2(t/2)\|\|\tilde{\bar{e}}_2(\tau)\|d\tau + \int_0^{\frac{t}{2}}\int_{\frac{t}{2}}^{t}\|\hat{e}_2(s)\|\tilde{\bar{e}}_2(\tau)\|d\tau ds.\Bigg) \\
&\le Ch^2\|u_0\|\|g\|.
\end{aligned}$$

Altogether these estimates yield the desired result and this completes the proof. □

Acknowledgments. The research of the first author has been supported by the Department of Science and Technology, Govt. of India under BOYSCAST fellowship. The first author also thanks the Institute for Scientific Computations and the Department of Mathematics, Texas A&M University for providing necessary facilities for carrying out the research work.

References

[1] F. Brezzi, J. Douglas, and L. Marini, *Two families of mixed elements for second order elliptic problems*, Numer. Math., 88 (1985), pp. 217-235.

[2] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[3] J. R. Cannon and Y. Lin, *Nonclassical H^1 projections and Galerkin methods for nonlinear parabolic integro-differential equations*, Calcolo 25 (1988), pp. 187-201.

[4] J. R. Cannon and Y. 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] H. Chen, R. E. Ewing, R. D. Lazarov, *Superconvergence of mixed finite element methods for parabolic problems with nonsmooth initial data*, Numer. Math., 78 (1998), pp. 495-521.

[6] J. H. Cushman and T. R. Glinn, *Nonlocal dispersion in media with continuously evolving scales of heterogeneity*, Transport in Porous Media, 13(1993), pp. 123-138.

[7] G. Dagan, *The significance of heterogeneity of evolving scales to transport in porous formations*, Water Resour. Res., 30 (1994), pp. 3327-3336.

[8] J. Douglas, JR., J. E. Roberts, *Global estimates for mixed methods for second order elliptic equations*, Math. Comp., 44 (1985), pp. 39-52.

[9] R.E. Ewing, Y. Lin and J. Wang, *A numerical approximation of non-Fickian flows with mixing length growth in porous media*, Acta. Math. Univ. Comenian. 70 (2001), pp. 75-84.

[10] R.E. Ewing, Y. Lin, T. Sun, J. Wang and S. Zhang, *Sharp L^2-error estimates and superconvergence of mixed finite element methods for non-Fickian flows in porous media*, SIAM J. Numer. Anal., 40 (2002), pp. 1538-1560.

[11] M. Huang and V. Thomée, *Some convergence estimates for semidiscrete type schemes for time-dependent nonselfadjoint parabolic equations*, Math. Comp. 37 (1981), pp. 327-346.

[12] Z. Jiang, *$L^\infty(L^2)$ and $L^\infty(L^\infty)$ error estimates for mixed methods for integro-differential equations of parabolic type*, M2AN Math. Model. Numer. Anal., 33 (1999), pp. 531-546.

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

[14] M. Luskin and R. Rannacher, *On the smoothing property of the Galerkin method for parabolic equations*, SIAM J. Numer. Anal. 19 (1981), pp. 93-113.

[15] A. K. Pani, V. Thomée and L. B. Wahlbin, *Numerical methods for hyperbolic and parabolic integro-differential equations*, J. Integral Equations and Appl. 4 (1992), pp. 533-584.

[16] A. K. Pani and T. E. Peterson, *The effect of numerical quadrature on semidiscrete finite element methods for parabolic integro-differential equations*, SIAM J. Numer. Anal. 33 (1996), pp. 1084-1105.

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

[18] P. A. Raviart and J. M. Thomas, *A mixed finite element method for second order elliptic problems*, Lecture Notes in Mathematics 606, I. Galligani and E. Magenes, eds., Springer-Verlag, Berlin, New York, 1997, pp. 292-315.

[19] M. Renardy, W. Hrusa and J. Nohel, *Mathematical Problems in Viscoelasticity*, Pitman Monographs and Surveys in Pure Appl. Math. No. 35 (1987), New York, Wiley.

[20] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Lecture Notes in Math. 1054 (1984), Springer-Verlag.

[21] 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.