IMA Journal of Numerical Analysis (2002) **22**, 231–252

H^1-Galerkin mixed finite element methods for parabolic partial integro-differential equations

AMIYA K. PANI
Department of Mathematics, Industrial Mathematics Group, Indian Institute of Technology, Bombay, Powai, Mumbai 400076, India

AND

GRAEME FAIRWEATHER
Department of Mathematical and Computer Sciences, Colorado School of Mines, Golden, Colorado 80401, USA

[Received on 31 May 2000; revised on 2 July 2001]

H^1-Galerkin mixed finite element methods are analysed for parabolic partial integro-differential equations which arise in mathematical models of reactive flows in porous media and of materials with memory effects. Depending on the physical quantities of interest, two methods are discussed. Optimal error estimates are derived for both semidiscrete and fully discrete schemes for problems in one space dimension. An extension to problems in two and three space variables is also discussed and it is shown that the H^1-Galerkin mixed finite element approximations have the same rate of convergence as in the classical methods without requiring the LBB consistency condition.

Keywords: H^1-Galerkin mixed finite element methods; parabolic partial integro-differential equations; semidiscrete and fully discrete schemes; optional error estimates.

1. Introduction

Mathematical models describing the nonlocal reactive flows in porous media (Cushman & Glinn, 1993; Dagan, 1994) and heat conduction through materials with memory (Renardy *et al.*, 1987; see Ewing *et al.* (2000) for additional references) give rise to parabolic partial integro-differential equations of the form

$$
\begin{aligned}
u_t + \mathcal{A}u + \int_0^t \mathcal{B}(t,s)u(s)\,\mathrm{d}s &= f(x,t), \quad x \in \Omega, \quad t \in (0,T], \\
u(x,t) &= 0, \quad x \in \partial\Omega, \quad t \in [0,T], \\
u(x,0) &= u_0(x), \quad x \in \Omega.
\end{aligned}
\tag{1.1}
$$

Here, Ω is a bounded domain in $R^d (d = 1, 2, 3)$ with boundary $\partial\Omega, T > 0$, $\mathcal{A}$ is a second-order uniformly elliptic and positive definite operator, $\mathcal{B}$ is an arbitrary second-order differential operator, and f and u_0 are known functions. When a classical mixed finite element method is applied to (1.1), we encounter certain difficulties due to the presence of the integral term (Ewing *et al.*, 2001; Jiang, 1999). However, if the operator $\mathcal{B}(t,s)$ is of a special form, namely, $\mathcal{B}(t,s) = \kappa(t,s)\mathcal{A}$, then *a priori* error estimates can be derived for the classical mixed method, provided the finite element spaces satisfy the LBB consistency condition. In the first part of this paper, we apply a recently developed mixed method (Pani, 1998) to a one-dimensional version of (1.1) and examine its convergence without requiring the LBB condition. The specific problem that we consider is

$$u_t - (a(x)u_x)_x - \int_0^t (b(t,s)u_x(s))_x \, \mathrm{d}s = f(x,t), \quad x \in (0,1), \quad t \in (0,T],$$
$$u(0,t) = u(1,t) = 0, \quad t \in [0,T], \tag{1.2}$$
$$u(x,0) = u_0(x), \quad x \in (0,1),$$

where $a(x)$ and $b(t,s) := b(x,t,s)$ are smooth functions with bounded derivatives, $f(x,t)$ and $u_0(x)$ are given functions, and

$$|b(x,t,s)| \leqslant a_1, \quad 0 < a_0 \leqslant a(x) \leqslant a_1, \quad x \in (0,1), \tag{1.3}$$

for positive constants a_0 and a_1. Depending on the physical quantities of interest, we consider two methods. With

$$q = a(x)u_x + \int_0^t b(t,s)u_x(s)\, \mathrm{d}s,$$

we reformulate the parabolic integro-differential equation as the first-order system

$$a(x)u_x + \int_0^t b(t,s)u_x(s)\, \mathrm{d}s = q, \tag{1.4}$$

$$u_t - q_x = f. \tag{1.5}$$

If our concern is to approximate $\sigma = a(x)u_x$ accurately, we rewrite (1.2) as the first-order system

$$a(x)u_x = \sigma, \tag{1.6}$$

$$u_t - \sigma_x - \int_0^t (\beta(t,s)\sigma(s))_x \, \mathrm{d}s = f, \tag{1.7}$$

where $\beta(t,s) = b(t,s)/a$. We apply H^1-Galerkin mixed finite element methods to both (1.4)–(1.7) and then extend these results to problem (1.1) in two and three space variables. The proposed methods can be thought of as nonsymmetric versions of least square methods, and the error analysis yields an optimal rate of convergence for the flux. For related studies on least square mixed finite element methods, see Carey & Shen (1989); Neittaanmäki & Saranen (1981a,b,c); Pehlivanov *et al.* (1993, 1994) and references therein.

Since (1.1) is an integral perturbation of a parabolic problem, it is natural to examine how far the mixed methods for parabolic problems (Johnson & Thomée, 1981) can be extended to the present case. It is to be noted that the classical mixed method must satisfy the LBB consistency condition on the approximating subspaces and this restricts the choice of finite element spaces. For example, the Raviart–Thomas spaces of index $r \geqslant 0$ are commonly used for classical mixed methods. However, when these methods are applied to (1.4)–(1.7), we encounter some difficulties in the error analysis due to the presence of the integral term. We have also observed that when $\mathcal{B}(t,s) = \kappa(t,s)\mathcal{A}$ in (1.1), the classical method yields optimal convergence for approximating q or σ provided the finite element spaces satisfy the LBB consistency condition. In order to circumvent these difficulties, we extend the analysis of Pani (1998) to (1.4), (1.5) and (1.6), (1.7). The other notable advantage of this approach is that the approximating finite element spaces V_h (approximating u) and W_h (approximating q or σ) are allowed to be of different polynomial degrees. Moreover, we note that, for L^2- and H^1-error estimates, we do not require a quasi-uniformity condition on the finite element mesh.

In the literature, there are only a few results concerning the analysis of the classical mixed methods for (1.1); see, for example, Ewing *et al.* (2001); Jiang (1999). In Jiang (1999), the author discussed a special case of (1.1) and derived *a priori* estimates for a multidimensional formulation of (1.4), (1.5). In Ewing *et al.* (2001), a classical mixed formulation is derived for the system

$$u_x = \alpha q + \int_0^t R(t,s)\alpha q(s)\,\mathrm{d}s,$$
$$u_t - q_x = f,$$

(cf., (1.4), (1.5)), where $\alpha = 1/a$ and $R(t,s)$ is the resolvent of $\alpha b(t,s)$ defined by

$$R(t,s) = \alpha b(t,s) + \int_0^t \alpha b(t,\tau) R(\tau,s)\,\mathrm{d}s, \quad t > s > 0.$$

Error estimates are obtained for the multidimensional problem using the LBB consistency condition.

Earlier work on finite element approximations to (1.1) is described in Cannon & Lin (1988, 1990); Lin *et al.* (1991); Pani *et al.* (1992); Sloan & Thomée (1986); Yanik & Fairweather (1988) and references therein. For generalized finite difference schemes, a convergence analysis is given in Pani *et al.* (1991) and finite volume methods applied to (1.2) are studied in Ewing *et al.* (2000).

A brief outline of this paper is as follows. In section 2, semidiscrete H^1-Galerkin mixed finite element approximations are discussed for both (1.4), (1.5) and (1.6), (1.7), and optimal error estimates are derived. Section 3 is devoted to a fully discrete scheme and the related optimal error estimates are obtained. In Section 4, we extend the H^1-Galerkin mixed finite element methods to problems in two and three space variables and derive *a priori* error estimates for the semidiscrete case. Moreover, some comparisons between classical mixed methods and the present ones are discussed.

Throughout this paper, C denotes a generic positive constant which does not depend on the spatial mesh parameter h or the time step Δt.

2. Error estimates for the semidiscrete case

2.1 *An H^1-Galerkin mixed finite element method for the system* (1.4)*,* (1.5)

We denote the natural inner product in $L^2(I)$ by $(\cdot,\cdot)$, and let $H_0^1 = \{v \in H^1(I) : v(0) = v(1) = 0\}$. Further, we use the classical Sobolev spaces $W^{m,p}(I)$, $1 \leqslant p \leqslant \infty$, written as $W^{m,p}$. The norm on $W^{m,p}$ is denoted by $\|\cdot\|_{m,p}$. When $p = 2$, we write $W^{m,2}$ as H^m and denote the norm by $\|\cdot\|_m$.

For use in the formulation of the H^1-Galerkin mixed finite element method for the system (1.4), (1.5), we consider the following weak formulation: find $\{u, q\} : [0, T] \mapsto H_0^1 \times H^1$ such that

$$(au_x, v_x) + \int_0^t (b(t,s)u_x(s), v_x)\,\mathrm{d}s = (q, v_x), \quad v \in H_0^1, \tag{2.1}$$

$$(\alpha q_t, w) + (q_x, w_x) = (\beta(t,t)u_x, w) + \int_0^t (\beta_t(t,s)u_x, w)\,\mathrm{d}s - (f, w_x), \quad w \in H^1. \tag{2.2}$$

Clearly (2.1) is obtained by multiplying (1.4) by v_x and integrating the resulting equation with respect to x. Equation (2.2) is derived by first multiplying (1.5) by w_x and integrating the first term by parts to give

$$(u_{tx}, w) + (q_x, w_x) = -(f, w_x), \tag{2.3}$$

since $u_t(0) = u_t(1) = 0$. Then, on dividing (1.4) by $a(x)$ and differentiating with respect to t, it follows that

$$u_{tx} = -\beta(t,t)u_x - \int_0^t \beta_t(t,s)u_x(s)\,\mathrm{d}s + \alpha q_t.$$

Substituting this expression in (2.3) yields (2.2).

Let V_h and W_h be finite-dimensional subspaces of H_0^1 and H^1, respectively, with the following approximation properties: for $1 \leqslant p \leqslant \infty$ and k, r positive integers,

$$\inf_{v_h \in V_h} \{\|v - v_h\|_{L^p} + h\|v - v_h\|_{W^{1,p}}\} \leqslant Ch^{k+1}\|v\|_{W^{k+1,p}}, \quad v \in H_0^1 \cap W^{k+1,p},$$

and

$$\inf_{w_h \in W_h} \{\|w - w_h\|_{L^p} + h\|w - w_h\|_{W^{1,p}}\} \leqslant Ch^{r+1}\|w\|_{W^{r+1,p}}, \quad w \in W^{r+1,p}.$$

The semidiscrete H^1-Galerkin mixed finite element method for (2.1), (2.2) consists in determining $\{u_h, q_h\} : [0, T] \mapsto V_h \times W_h$ such that

$$\begin{aligned} &(au_{hx}, v_{hx}) + \int_0^t (b(t,s)u_{hx}(s), v_{hx})\,\mathrm{d}s = (q_h, v_{hx}), \quad v_h \in V_h, \\ &(\alpha q_{ht}, w_h) + (q_{hx}, w_{hx}) = (\beta(t,t)u_{hx}, w_h) + \int_0^t (\beta_t(t,s)u_{hx}, w_h)\,\mathrm{d}s \\ &\qquad - (f, w_{hx}), \quad w_h \in W_h, \end{aligned} \tag{2.4}$$

with given $u_h(0)$ and $q_h(0)$.

For use in the error analysis, we define the Ritz–Volterra projection $\tilde{u}_h \in V_h$ by

$$(a(u_x - \tilde{u}_{hx}), v_{hx}) + \int_0^t (b(t,s)(u_x - \tilde{u}_{hx}), v_{hx})\,\mathrm{d}s = 0, \quad v_h \in V_h, \tag{2.5}$$

Cannon & Lin (1988, 1990); Lin *et al.* (1991). Further, we also define an elliptic projection $\tilde{q}_h \in W_h$ of q (Wheeler, 1973) as the solution of

$$A(q - \tilde{q}_h, w_h) = 0, \quad w_h \in W_h, \tag{2.6}$$

where $A(q, w) = (q_x, w_x) + \lambda(q, w)$. Here λ is chosen so that A is H^1-coercive, that is

$$A(w, w) \geqslant \aleph_0 \|w\|_1^2, \quad w \in H^1, \tag{2.7}$$

where $\aleph_0$ is a positive constant. Moreover, it is easy to show that $A(\cdot, \cdot)$ is bounded.

Let $\rho = q - \tilde{q}_h$ and $\eta = u - \tilde{u}_h$. The following estimates for ρ and η are well known (Pani *et al.*, 1992; Wheeler, 1973): for $j = 0, 1$,

$$\|\eta\|_j + \|\eta_t\|_j \leqslant Ch^{k+1-j}\left[\|u\|_{k+1} + \|u_t\|_{k+1} + \int_0^t \|u(s)\|_{k+1}\,\mathrm{d}s\right], \tag{2.8}$$

and

$$\|\rho\|_j \leqslant Ch^{r+1-j}\|q\|_{r+1}, \quad \|\rho_t\|_j \leqslant Ch^{r+1-j}\|q_t\|_{r+1}. \tag{2.9}$$

Further, for $j = 0, 1$, and $1 \leqslant p \leqslant \infty$, we have

$$\|\eta\|_{W^{j,p}} \leqslant Ch^{k+1-j}\left[\|u\|_{W^{k+1,p}} + \int_0^t \|u(s)\|_{W^{k+1,p}}\,\mathrm{d}s\right], \tag{2.10}$$

and

$$\|\rho\|_{W^{j,p}} \leqslant Ch^{r+1-j}\|q\|_{W^{r+1,p}}. \tag{2.11}$$

Note that, for $p = \infty$, we require a quasi-uniformity condition on the finite element mesh.

To determine the desired *a priori* error estimates, we write

$$u - u_h = (u - \tilde{u}_h) + (\tilde{u}_h - u_h) := \eta + \zeta,$$

and

$$q - q_h = (q - \tilde{q}_h) + (\tilde{q}_h - q_h) := \rho + \xi.$$

Using (2.1), (2.2) and the projections (2.5), (2.6), we obtain

$$(a\zeta_x, v_{hx}) + \int_0^t (b(t,s)\zeta_x, v_{hx})\,\mathrm{d}s = (\rho, v_{hx}) + (\xi, v_{hx}), \quad v_h \in V_h, \tag{2.12}$$

and

$$\begin{aligned}(\alpha\xi_t, w_h) + A(\xi, w_h) = &-(\alpha\rho_t, w_h) + \lambda(\rho + \xi, w_h) + (\beta(t,t)(\eta_x + \zeta_x), w_h) \\ &+ \int_0^t (\beta_t(t,s)(\eta_x + \zeta_x), w_h)\,\mathrm{d}s, \quad w_h \in W_h. \end{aligned} \tag{2.13}$$

THEOREM 2.1 If $q(0) = au_{0x}$ and $q_h(0) = \tilde{q}_h(0)$, then

$$\begin{aligned}&\|(u - u_h)(t)\| + \|(q - q_h)(t)\| + h\|(u - u_h)(t)\|_1 \\ &\qquad \leqslant Ch^{\min(k+1,r+1)}[\|u\|_{L^\infty(H^{k+1})} + \|q\|_{L^\infty(H^{r+1})} + \|q_t\|_{L^2(H^{r+1})}].\end{aligned}$$

Proof. Since estimates of η and ρ are given in (2.8), (2.9), it suffices to bound ζ and ξ. To this end, we choose $v_h = \zeta$ in (2.12). Then, using (1.3), the Cauchy–Schwarz inequality and Young's inequality ($ab \leqslant \epsilon a^2 + b^2/4\epsilon,\ a,\ b \in \mathcal{R}, \epsilon > 0$), we obtain

$$(a_0 - \epsilon)\|\zeta_x\|^2 \leqslant C\left[\|\rho\|^2 + \|\xi\|^2 + \int_0^t \|\zeta_x(s)\|^2 \,\mathrm{d}s\right].$$

Using the Gronwall lemma, it follows that, for ϵ sufficiently small, $\epsilon = \alpha_0/2$ say,

$$\|\zeta_x(t)\|^2 \leqslant C\left[\|\rho(t)\|^2 + \|\xi(t)\|^2 + \int_0^t (\|\rho(s)\|^2 + \|\xi(s)\|^2)\,\mathrm{d}s\right]. \tag{2.14}$$

Next, we set $w_h = \xi$ in (2.13). Then, on using the coercivity of A, (2.7),

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\alpha^{1/2}\xi\|^2 + \aleph_0\|\xi\|^2 &\leqslant -(\alpha\rho_t, \xi) + \lambda(\rho + \xi, \xi) + (\beta(t,t)\eta_x, \xi) + (\beta(t,t)\zeta_x, \xi)\\ &\quad + \int_0^t (\beta_t(t,s)\eta_x, \xi)\,\mathrm{d}s + \int_0^t (\beta_t(t,s)\zeta_x, \xi)\,\mathrm{d}s.\end{aligned} \tag{2.15}$$

Using integration by parts with respect to x,

$$(\beta(t,t)\eta_x, \xi) = -(\beta_x(t,t)\eta, \xi) - (\beta(t,t)\eta, \xi_x). \tag{2.16}$$

Similarly,

$$\int_0^t (\beta_t(t,s)\eta_x, \xi)\,\mathrm{d}s = -\int_0^t (\beta_{tx}(t,s)\eta, \xi)\,\mathrm{d}s - \int_0^t (\beta_t(t,s)\eta, \xi_x)\,\mathrm{d}s. \tag{2.17}$$

On substituting (2.16) and (2.17) in (2.15) and using the Cauchy–Schwarz inequality and Young's inequality, we obtain, for sufficiently small ϵ,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\|\alpha^{1/2}\xi\|^2 + (2\aleph_0 - \epsilon)\|\xi\|_1^2 &\leqslant C\bigg[\|\rho_t\|^2 + \|\rho\|^2 + \|\eta\|^2 + \|\xi\|^2 + \|\zeta_x\|^2\\ &\quad + \int_0^t (\|\eta(s)\|^2 + \|\zeta_x(s)\|^2 + \|\xi(s)\|_1^2)\,\mathrm{d}s\bigg]\\ &\leqslant C\bigg[\|\rho_t\|^2 + \|\rho\|^2 + \|\eta\|^2 + \|\xi\|^2\\ &\quad + \int_0^t (\|\eta(s)\|^2 + \|\rho(s)\|^2 + \|\xi(s)\|^2 + \|\xi(s)\|_1^2)\,\mathrm{d}s\bigg],\end{aligned} \tag{2.18}$$

where in the second step we have used (2.14). If we integrate with respect to t, then

$$\begin{aligned}\|\xi\|^2 + \int_0^t \|\xi(s)\|_1^2\,\mathrm{d}s &\leqslant C\bigg[\int_0^t (\|\rho_t(s)\|^2 + \|\rho(s)\|^2 + \|\eta(s)\|^2 + \|\xi(s)\|^2)\,\mathrm{d}s\\ &\quad + \int_0^t\int_0^\tau (\|\eta(s)\|^2 + \|\rho(s)\|^2 + \|\xi(s)\|^2 + \|\xi(s)\|_1^2)\,\mathrm{d}s\,\mathrm{d}\tau\bigg].\end{aligned}$$

On changing the order of integration, it follows that

$$\int_0^t \int_0^\tau \chi^2(s)\,\mathrm{d}s\,\mathrm{d}\tau \leqslant T \int_0^t \chi^2(s)\,\mathrm{d}s, \tag{2.19}$$

where

$$\chi^2(s) = \|\eta(s)\|^2 + \|\rho(s)\|^2 + \|\xi(s)\|^2.$$

Thus, an application of the Gronwall lemma and the use of (2.8) and (2.9) yield

$$\begin{aligned}\|\xi\|^2 + \int_0^t \|\xi(s)\|_1^2\,\mathrm{d}s &\leqslant C\left[\int_0^t (\|\rho_t(s)\|^2 + \|\rho(s)\|^2 + \|\eta(s)\|^2)\,\mathrm{d}s\right]\\ &\leqslant Ch^{2\min(k+1,r+1)} \int_0^T (\|u(s)\|_{k+1}^2 + \|q(s)\|_{r+1}^2 + \|q_t(s)\|_{r+1}^2)\,\mathrm{d}s.\end{aligned}$$

If, in (2.14), we use this estimate for $\|\xi\|$ and (2.9), it follows that

$$\|\zeta_x(t)\| \leqslant Ch^{\min(k+1,r+1)}[\|u\|_{L^2(H^{k+1})} + \|q\|_{L^\infty(H^{r+1})} + \|q_t\|_{L^2(H^{r+1})}]. \tag{2.20}$$

Since $\zeta \in H_0^1$, we have $\|\zeta(t)\| \leqslant \|\zeta_x(t)\|$ from the Poincaré inequality. Applying the triangle inequality with (2.8) completes the proof. □

Note that in Theorem 2.1 it is possible to choose the initial approximation $q_h(0)$ as the L^2 projection of $q(0)$ onto W_h.

THEOREM 2.2 Assume that $q_h(0) = \tilde{q}_h(0)$. Then

$$\|(q - q_h)(t)\|_1 \leqslant Ch^{\min(k+1,r)}[\|u\|_{L^\infty(H^{k+1})} + \|u_t\|_{L^2(H^{k+1})} + \|q\|_{L^\infty(H^r)} + \|q_t\|_{L^2(H^r)}], \tag{2.21}$$

and, for $1 < p \leqslant \infty$,

$$\begin{aligned}\|(u - u_h)(t)\|_{L^p} + \|(q - q_h)(t)\|_{L^p} \leqslant Ch^{\min(k+1,r+1)}[&\|u\|_{L^\infty(W^{k+1,p})} + \|u_t\|_{L^2(H^{k+1})}\\ &+ \|q\|_{L^\infty(W^{r+1,p})} + \|q_t\|_{L^2(H^{r+1})}].\end{aligned} \tag{2.22}$$

Proof. Since we have a superconvergence result for ζ_x from (2.20), it is sufficient to obtain a superconvergence estimate for ξ in the H^1-norm. To this end, we choose $w_h = \xi_t$ in (2.13) to obtain

$$\begin{aligned}\|\alpha^{1/2}\xi_t\|^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}A(\xi,\xi) = &-(\alpha\rho_t, \xi_t) + \lambda(\rho + \xi, \xi_t) + (\beta(t,t)(\eta_x + \zeta_x), \xi_t)\\ &+ \int_0^t (\beta_t(t,s)(\eta_x(s) + \zeta_x(s)), \xi_t)\,\mathrm{d}s.\end{aligned}$$

Using integration by parts with respect to x,

$$(\beta(t,t)\eta_x, \xi_t) = -(\beta_x(t,t)\eta, \xi_t) - (\beta(t,t)\eta, \xi_{tx}),$$

and

$$\int_0^t (\beta_t(t,s)\eta_x(s), \xi_t)\,\mathrm{d}s = -\int_0^t (\beta_{tx}(t,s)\eta(s), \xi_t)\,\mathrm{d}s - \int_0^t (\beta_t(t,s)\eta(s), \xi_{tx})\,\mathrm{d}s.$$

Also, it is easy to show that

$$-(\beta(t,t)\eta,\xi_{tx}) = -\frac{\mathrm{d}}{\mathrm{d}t}(\beta\eta,\xi_x) + (\beta(t,t)\eta_t + \beta_t(t,t)\eta,\xi_x),$$

and

$$\begin{aligned}-\int_0^t (\beta_t(t,s)\eta(s),\xi_{tx})\,\mathrm{d}s = -\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^t (\beta_t(t,s)\eta(s),\xi_x)\,\mathrm{d}s\right] + (\beta_t(t,t)\eta(t),\xi_x)\\ + \int_0^t (\beta_{tt}(t,s)\eta(s),\xi_x)\,\mathrm{d}s.\end{aligned}$$

Thus,

$$\begin{aligned}\|\alpha^{1/2}\xi_t\|^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}A(\xi,\xi) &= -(\alpha\rho_t,\xi_t) + \lambda(\rho+\xi,\xi_t) + (\beta(t,t)\zeta_x - \beta_x(t,t)\eta,\xi_t)\\ &+ \int_0^t (\beta_t(t,s)\zeta_x(s) - \beta_{tx}(t,s)\eta(s),\xi_t)\,\mathrm{d}s - \frac{\mathrm{d}}{\mathrm{d}t}(\beta\eta(t),\xi_x) + (\beta\eta_t + \beta_t\eta,\xi_x)\\ &- \frac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^t (\beta_t(t,s)\eta(s),\xi_x)\,\mathrm{d}s\right] + (\beta_t(t,t)\eta,\xi_x) + \int_0^t (\beta_{tt}(t,s)\eta(s),\xi_x)\,\mathrm{d}s.\end{aligned}$$

On integrating with respect to t and using (2.7), the Cauchy–Schwarz inequality, Young's inequality, and (2.19) with appropriate $\chi(s)$, we obtain

$$\begin{aligned}\int_0^t \|\xi_t(s)\|^2\,\mathrm{d}s + \|\xi(t)\|_1^2 \leqslant C\bigg[\|\eta(t)\|^2 + \int_0^t (\|\rho(s)\|^2 + \|\rho_t(s)\|^2\\ + \|\eta(s)\|^2 + \|\eta_t(s)\|^2 + \|\zeta_x(s)\|^2)\,\mathrm{d}s + \int_0^t (\|\xi(s)\|_1^2 + \|\xi_t(s)\|^2)\,\mathrm{d}s\bigg].\end{aligned}$$

Using the Gronwall lemma followed by (2.8), (2.9) and (2.20) yields

$$\begin{aligned}\|\xi\|_1^2 + \int_0^t \|\xi_t(s)\|^2\,\mathrm{d}s \leqslant Ch^{2\min(k+1,r+1)}\bigg[\|u(t)\|_{k+1}^2 + \int_0^t (\|q(s)\|_{r+1}^2\\ + \|u(s)\|_{k+1}^2 + \|u_t(s)\|_{k+1}^2 + \|q_t(s)\|_{r+1}^2)\,\mathrm{d}s\bigg]. \quad (2.23)\end{aligned}$$

Note that, from (2.23), we obtain a superconvergence result for ξ in the H^1-norm. However, for the derivation of (2.21), it is possible to reduce the regularity of q and q_t by allowing $q \in L^\infty(H^r)$ and $q_t \in L^2(H^r)$. For $1 \leqslant p \leqslant \infty$, we have, from the Sobolev embedding theorem,

$$\|\xi(t)\|_{L^p} \leqslant \|\xi(t)\|_1, \quad \xi \in H^1,$$

and

$$\|\zeta(t)\|_{L^p} \leqslant C\|\zeta_x(t)\|, \quad \zeta \in H_0^1.$$

The use of the superconvergence results (2.20) and (2.23) with (2.10), (2.11) and the triangle inequality completes the proof. □

REMARK 2.1

(i) From Theorem 2.1 with $k = r$ and using the definition of q, (1.4),

$$\|(u - u_h)\|_{L^\infty(L^2)} + \|(q - q_h)\|_{L^\infty(L^2)} \leqslant Ch^{r+1}[\|u\|_{L^\infty(H^{r+2})} + \|u_t\|_{L^2(H^{r+2})}].$$

Further, from the superconvergence result (2.20) and the estimate (2.8), we obtain

$$\|(u - u_h)(t)\|_1 \leqslant Ch^{\min(k,r+1)}[\|u\|_{L^\infty(H^k)} + \|q\|_{L^\infty(H^{r+1})} + \|q_t\|_{L^2(H^{r+1})}].$$

For $r + 1 = k$, we obtain

$$\|u - u_h\|_{L^\infty(H^1)} = \mathrm{O}(h^k),$$

whereas from Theorem 2.2 for $k + 1 = r$, we have the estimate

$$\|q - q_h\|_{L^\infty(H^1)} = \mathrm{O}(h^r).$$

Hence, the order of convergence corresponds to the degree of the polynomials used in the corresponding finite element spaces.

(ii) When $\mathcal{B}(t, s) = \kappa(t, s)\mathcal{A}$ with a smooth kernel κ, it is possible to apply the classical mixed method. The convergence analysis may be obtained using the LBB consistency condition. As was pointed out in Pani (1998) in the case of parabolic problems, the present method provides a better rate of convergence for the error $q - q_h$ than the conventional use of linear elements, without appealing to the LBB consistency condition.

2.2 *An H^1-Galerkin mixed finite element method for the system* (1.6), (1.7)

We now discuss the convergence of an H^1-mixed method applied to the system (1.6), (1.7). This method is based on the following weak formulation: find $\{u, \sigma\} : [0, T] \mapsto H_0^1 \times H^1$ such that

$$(au_x, v_x) = (\sigma, v_x), \quad v \in H_0^1, \tag{2.24}$$

$$(\alpha\sigma_t, w) + (\sigma_x, w_x) + \int_0^t (\beta(t, s)\sigma_x(s) + \beta_x(t, s)\sigma(s), w_x)\,\mathrm{d}s = -(f, w_x), \quad w \in H^1, \tag{2.25}$$

where $\alpha = 1/a$ and $\beta = b/\alpha$ as before. To obtain (2.24), we first multiply (1.7) by $-w_x$ and integrate the first term in the resulting equation by parts, which yields

$$(u_{tx}, w) + (\sigma_x, w_x) + \int_0^t (\beta_x(t, s)\sigma(s) + \beta(t, s)\sigma_x(s), w_x)\,\mathrm{d}s = -(f, w_x),$$

using the fact that $u_t(0) = u_t(1) = 0$. Since $u_t = \alpha\sigma_t$ from (1.6), we then obtain (2.25).

With finite element spaces V_h and W_h, we define the semidiscrete H^1-Galerkin mixed finite element approximation $\{u_h, \sigma_h\} : [0, T] \mapsto V_h \times W_h$ by

$$(au_{hx}, v_{hx}) = (\sigma_h, v_{hx}), \quad v_h \in V_h, \tag{2.26}$$

$$(\alpha\sigma_{ht}, w_h) + (\sigma_{hx}, w_{hx}) + \int_0^t (\beta(t, s)\sigma_{hx}(s) + \beta_x(t, s)\sigma_h(s), w_{hx})\,\mathrm{d}s = -(f, w_{hx}), \quad w_h \in W_h, \tag{2.27}$$

with initial approximations $u_h(0)$ and $\sigma_h(0)$ defined later.

To derive the error estimates, we employ the projections $\hat{u}_h$, $\hat{\sigma}_h$ defined by

$$(a(u_x - \hat{u}_{hx}), v_{hx}) = 0, v_h \in V_h, \tag{2.28}$$

$$A(\sigma - \hat{\sigma}_h, w_h) + \int_0^t (b(t,s)(\sigma_x - \hat{\sigma}_{hx}) + \beta_x(t,s)(\sigma - \hat{\sigma}_h), w_{hx})\,\mathrm{d}s = 0, \quad w_h \in W_h, \tag{2.29}$$

where $A(\phi, \chi) = (\phi_x, \chi_x) + \lambda(\phi, \chi)$. As before, the constant λ is chosen so that the bilinear form $A(\cdot,\cdot)$ is H^1-coercive.

With $\hat{\eta} = u - \hat{u}_h$ and $\hat{\rho} = \sigma - \hat{\sigma}_h$, the following estimates are well known (Pani *et al.*, 1992; Wheeler, 1973): for $j = 0, 1$,

$$\|\hat{\eta}\|_j \leqslant Ch^{k+1-j}\|u\|_{k+1}, \quad \|\hat{\eta}_t\|_j \leqslant Ch^{k+1-j}\|u_t\|_{k+1}, \tag{2.30}$$

and

$$\|\hat{\rho}\|_j + \|\hat{\rho}_t\|_j \leqslant Ch^{r+1-j}\left[\|\sigma\|_{r+1} + \|\sigma_t\|_{r+1} + \int_0^t \|\sigma(s)\|_{r+1}\,\mathrm{d}s\right]. \tag{2.31}$$

Moreover, for $j = 0, 1$, and $1 \leqslant p \leqslant \infty$, we have

$$\|\hat{\rho}\|_{W^{j,p}} \leqslant Ch^{r+1-j}\left[\|\sigma\|_{W^{r+1,p}} + \int_0^t \|\sigma(s)\|_{W^{r+1,p}}\,\mathrm{d}s\right], \tag{2.32}$$

and

$$\|\hat{\eta}\|_{W^{j,p}} \leqslant Ch^{k+1-j}\|u\|_{W^{k+1,p}}. \tag{2.33}$$

Again, for $p = \infty$, the finite element mesh is required to be quasi-uniform.

With $u - u_h = (u - \hat{u}_h) + (\hat{u}_h - u_h) := \hat{\eta} + \hat{\zeta}$, and $\sigma - \sigma_h = (\sigma - \hat{\sigma}_h) + (\hat{\sigma}_h - \sigma_h) := \hat{\rho} + \hat{\xi}$, the equations for $\hat{\zeta}$ and $\hat{\xi}$ may be written as

$$(a\hat{\zeta}_x, v_{hx}) = (\hat{\rho}, v_{hx}) + (\hat{\xi}, v_{hx}), \quad v_h \in V_h, \tag{2.34}$$

and

$$(\alpha\hat{\xi}_t, w_h) + A(\hat{\xi}, w_h) + \int_0^t (\beta(t,s)\hat{\xi}_x(s) + \beta_x(t,s)\hat{\xi}(s), w_{hx})\,\mathrm{d}s$$
$$= -(\alpha\hat{\rho}_t, w_h) + \lambda(\hat{\rho} + \hat{\xi}, w_h), \quad w_h \in W_h. \tag{2.35}$$

THEOREM 2.3 With $\sigma(0) = au_{0x}$, assume that $\sigma_h(0) = \hat{\sigma}_h(0)$. Then

$$\|(\sigma - \sigma_h)(t)\| \leqslant Ch^{r+1}[\|\sigma\|_{L^\infty(H^{r+1})} + \|\sigma_t\|_{L^2(H^{r+1})}],$$
$$\|(u - u_h)(t)\| \leqslant Ch^{\min(k+1,r+1)}[\|u\|_{L^\infty(H^{k+1})} + \|\sigma\|_{L^\infty(H^{r+1})} + \|\sigma_t\|_{L^2(H^{r+1})}],$$

and

$$\|(u - u_h)(t)\|_1 \leqslant Ch^{\min(k,r+1)}[\|u\|_{L^\infty(H^{k+1})} + \|\sigma\|_{L^\infty(H^{r+1})} + \|\sigma_t\|_{L^2(H^{r+1})}].$$

Proof. In light of (2.30), (2.31), it is sufficient to derive estimates for $\hat{\zeta}$ and $\hat{\xi}$. Thus, if we set $v_h = \hat{\zeta}$ in (2.34) and use (1.3), it is easy to show that

$$\|\hat{\zeta}_x\| \leqslant C(\|\hat{\rho}\| + \|\hat{\xi}\|). \tag{2.36}$$

Further, choosing $w_h = \hat{\xi}$ in (2.35) and applying standard arguments, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\alpha^{\frac{1}{2}}\hat{\xi}\|^2 + \|\hat{\xi}\|_1^2 \leqslant C\left[\|\hat{\rho}_t\|^2 + \|\hat{\rho}\|^2 + \|\hat{\xi}\|^2 + \int_0^t \|\hat{\xi}(s)\|_1^2\,\mathrm{d}s\right].$$

If we integrate with respect to t and use the boundedness of α, then an application of the Gronwall lemma yields

$$\|\hat{\xi}(t)\|^2 + \int_0^t \|\hat{\xi}(s)\|_1^2\,\mathrm{d}s \leqslant C\left[\|\hat{\xi}(0)\|^2 + \int_0^t (\|\hat{\rho}_t(s)\|^2 + \|\hat{\rho}(s)\|^2)\,\mathrm{d}s\right].$$

Since $\sigma_h(0) = \hat{\sigma}_h(0)$, we have $\hat{\xi}(0) = 0$. The desired results follow from (2.30), (2.31) and (2.36). □

From the proof of Theorem 2.3, it is clear that we can choose $\sigma_h(0)$ as the L^2 projection of $\sigma(0)$ onto W_h instead of the elliptic projection $\hat{\sigma}_h(0)$.

THEOREM 2.4 Assume that $\sigma_h(0) = \hat{\sigma}_h(0)$. Then, for $1 < p \leqslant \infty$,

$$\|(\sigma - \sigma_h)(t)\|_{L^p} \leqslant Ch^{r+1}[\|\sigma\|_{L^\infty(W^{r+1,p})} + \|\sigma_t\|_{L^2(H^{r+1})}],$$

and

$$\|(u - u_h)(t)\|_{L^p} \leqslant Ch^{\min(k+1,r+1)}[\|u\|_{L^\infty(W^{k+1,p})} + \|\sigma\|_{L^\infty(H^{r+1})} + \|\sigma_t\|_{L^2(H^{r+1})}].$$

Moreover,

$$\|(\sigma - \sigma_h)(t)\|_1 \leqslant Ch^r[\|\sigma\|_{L^\infty(H^{r+1})} + \|\sigma_t\|_{L^2(H^r)}].$$

Proof. Since we have a superconvergence result for $\|\hat{\zeta}_x\|$ from (2.36), it is sufficient to derive a superconvergence estimate for $\hat{\xi}$ in the H^1-norm. Note that using the Sobolev embedding theorem, we obtain

$$\|\hat{\zeta}\|_{L^p} \leqslant C\|\hat{\zeta}_x\|, \quad \hat{\zeta} \in H_0^1,$$

and

$$\|\hat{\xi}\|_{L^p} \leqslant C\|\hat{\xi}\|_1, \quad \hat{\xi} \in H^1.$$

Then the estimates (2.32), (2.33) complete the proof.

To derive an H^1-estimate of $\hat{\xi}$, we set $w_h = \hat{\xi}_t$ in (2.35) and rewrite the resulting equation as

$$\begin{aligned}\|\alpha^{\frac{1}{2}}\hat{\xi}_t\|^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}A(\hat{\xi}, \hat{\xi}) &= -(\alpha\hat{\rho}_t, \hat{\xi}_t) + \lambda(\hat{\rho} + \hat{\xi}, \hat{\xi}_t) \\ &\quad - \frac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^t (\beta(t,s)\hat{\xi}_x(s) + \beta_x(t,s)\hat{\xi}(s), \hat{\xi}_x)\,\mathrm{d}s\right] \\ &\quad + (\beta(t,t)\hat{\xi}_x(t) + \beta_x(t,t)\hat{\xi}(t), \hat{\xi}_x) \\ &\quad + \int_0^t (\beta_t(t,s)\hat{\xi}_x(s) + \beta_{xt}(t,s)\hat{\xi}(s), \hat{\xi}_x)\,\mathrm{d}s.\end{aligned}$$

We now integrate with respect to t and use the boundedness of α. Since $\hat{\xi}_h(0) = 0$, we find that

$$\|\hat{\xi}(t)\|_1^2 \leqslant C\left[\int_0^t (\|\hat{\rho}(s)\|^2 + \|\hat{\rho}_t(s)\|^2 + \|\hat{\xi}(s)\|_1^2)\,\mathrm{d}s\right].$$

An application of the Gronwall lemma with (2.30) yields the superconvergence result

$$\|\hat{\xi}(t)\|_1^2 \leqslant Ch^{2(r+1)} \int_0^t [\|\sigma(s)\|_{r+1}^2 + \|\sigma_t(s)\|_{r+1}^2]\,\mathrm{d}s.$$

□

REMARK 2.2 Note that (2.27) can be solved independently of (2.26). From Theorems 2.3 and 2.4, it follows that, since $\sigma = au_x$,

$$\|(\sigma - \sigma_h)(t)\|_j \leqslant Ch^{r+1-j}[\|u\|_{L^\infty(H^{r+2})} + \|u_t\|_{L^2(H^{r+2-j})}], \quad j = 0, 1,$$

and

$$\|(\sigma - \sigma_h)(t)\|_{L^\infty} \leqslant Ch^{r+1}[\|u\|_{L^\infty(W^{r+2,\infty})} + \|u_t\|_{L^2(H^{r+2})}].$$

Thus, the degree k of V_h does not influence the L^2, H^1 and L^∞ norm estimates of $\sigma - \sigma_h$.

3. Fully discrete scheme

For the temporal discretization, we discuss the backward Euler method, which is first-order accurate in time; second-order schemes such as the Crank–Nicolson and two-step backward difference methods can be considered in a similar fashion (Pani *et al.*, 1992). Moreover, we examine the fully discrete scheme only for (2.1), (2.2); the analysis for (2.24), (2.25) follows in a similar fashion.

For the backward Euler method, let $t_n = n\Delta t$, with $\Delta t = T/M$, for some positive integer M. Further, let $\phi^n = \phi(t_n)$ and $\bar{\partial}_t\phi^n = (\phi^n - \phi^{n-1})/\Delta t$, for some continuous function $\phi \in C^0[0, T]$. For approximating the integrals, we use the composite left rectangle rule,

$$Q_n(\phi) := \Delta t \sum_{j=0}^{n-1} \phi^j \approx \int_0^{t_n} \phi(s)\,\mathrm{d}s.$$

Note that for $\phi \in C^1[0, T]$, the quadrature error satisfies

$$\left| Q_n(\phi) - \int_0^{t_n} \phi(s)\,\mathrm{d}s \right| \leqslant C\Delta t \int_0^{t_n} |\phi_t(s)|\,\mathrm{d}s. \tag{3.1}$$

Then, the fully discrete backward Euler approximation for (2.1), (2.2) is a sequence $\{U^n, Z^n\}_{n=0}^M$ such that

$$(aU_x^n, v_{hx}) + \Delta t \sum_{j=0}^{n-1} (b_{nj}U_x^j, v_{hx}) = (Z^n, v_{hx}), \quad v_h \in V_h, \tag{3.2}$$

$$(\alpha\bar{\partial}_t Z^n, w_h) + (Z_x^n, w_{hx}) = (\beta_{nn}U_x^n, w_h) + \Delta t \sum_{j=0}^{n-1} (\beta_{t,nj}U_x^j, w_h) - (f^n, w_{hx}), \quad w_h \in W_h. \tag{3.3}$$

Here, $b_{nj} = b(t_n, t_j)$, $\beta_{nn} = \beta(t_n, t_n)$ and $\beta_{t,nj} = \beta_t(t_n, t_j)$.

For the error analysis, we write

$$u(t_n) - U^n = (u(t_n) - \tilde{u}_h(t_n)) + (\tilde{u}_h(t_n) - U^n) := \eta^n + \zeta^n,$$

and

$$q(t_n) - Z^n = (q(t_n) - \tilde{q}_h(t_n)) + (\tilde{q}_h(t_n) - Z^n) := \rho^n + \xi^n.$$

Using (2.1), (2.2), (3.2), (3.3), (2.5) and (2.6) at $t = t_n$, we obtain

$$(a\zeta_x^n, v_{hx}) + \Delta t \sum_{j=0}^{n-1} (b_{nj}\zeta_x^j, v_{hx}) = (\rho^n + \xi^n, v_{hx}) + \epsilon_1^n(v_h), \quad v_h \in V_h, \tag{3.4}$$

$$\begin{aligned}(\alpha\bar{\partial}_t\xi^n, w_h) + A(\xi^n, w_h) = &-(\alpha\bar{\partial}_t\rho^n + \alpha\tau^n, w_h) + \lambda(\rho^n + \xi^n, w_h) \\ &+ (\beta_{nn}\zeta_x^n - \beta_{x,nn}\eta^n, w_h) - I^n(w_{hx}) + \epsilon_2^n(w_h) \\ &+ \Delta t \sum_{j=0}^{n-1} (\beta_{t,nj}\zeta_x^j - \beta_{tx,nj}\eta^j, w_h), \quad w_h \in W_h,\end{aligned} \tag{3.5}$$

where

$$\begin{aligned}\tau^n &= u_t(t_n) - \bar{\partial}_t u(t_n), \\ \epsilon_1^n(v_h) &= \Delta t \sum_{j=0}^{n-1} (b_{nj}\tilde{u}_{hx}(t_j), v_{hx}) - \int_0^{t_n} (b(t_n, s)\tilde{u}_{hx}(s), v_{hx})\, \mathrm{d}s, \\ \epsilon_2^n(w_h) &= \Delta t \sum_{j=0}^{n-1} (\beta_{t,nj}u_x(t_j), w_h) - \int_0^{t_n} (\beta_t(t_n, s)u_x(s), w_h)\, \mathrm{d}s,\end{aligned}$$

and

$$I^n(w_{hx}) = (\beta_{nn}\eta^n, w_{hx}) + \Delta t \sum_{j=0}^{n-1} (\beta_{t,nj}\eta^j, w_{hx}).$$

Note that, for the terms containing η_x in the right-hand side of (3.5), we have used integration by parts with respect to x.

We now prove the following theorem.

THEOREM 3.1 Assume that $q_h(0) = \tilde{q}_h(0)$. Then, for $0 \leqslant J \leqslant M$,

$$\begin{aligned}\|q^J - Z^J\| \leqslant C[&h^{\min(r+1,k+1)}(\|u\|_{L^\infty(H^{k+1})} + \|q\|_{L^\infty(H^{r+1})} + \|q_t\|_{L^2(H^{r+1})}) \\ &+ \Delta t(\|u\|_{L^2(H^1)} + \|u_t\|_{L^2(H^1)} + \|q_{tt}\|_{L^2(L^2)})],\end{aligned} \tag{3.6}$$

and

$$\begin{aligned}\|u^J - U^J\|_j \leqslant C[&h^{\min(k+1-j,r+1)}(\|u\|_{L^\infty(H^{k+1})} + \|q\|_{L^\infty(H^{r+1})} + \|q_t\|_{L^2(H^{r+1})}) \\ &+ \Delta t(\|u\|_{L^2(H^1)} + \|u_t\|_{L^2(H^1)} + \|q_{tt}\|_{L^2(L^2)})], \quad j = 0, 1.\end{aligned} \tag{3.7}$$

Proof. With $v_h = \zeta^n$ in (3.4), we obtain

$$\|\zeta_x^n\| \leqslant C\left[\|\rho^n\| + \|\xi^n\| + \|\epsilon_1^n\| + \Delta t\sum_{j=0}^{n-1}\|\zeta_x^j\|\right].$$

Use of the discrete Gronwall lemma and the estimate of the quadrature error for the second last term in the right-hand side yields

$$\|\zeta_x^n\| \leqslant C\left[\|\rho^n\| + \|\xi^n\| + \Delta t\int_0^{t_n}(\|\tilde{u}_{hx}(s)\| + \|\tilde{u}_{h,xt}(s)\|)\,\mathrm{d}s\right]. \tag{3.8}$$

With $w_h = \xi^n$ in (3.5), then using

$$(\alpha\bar{\partial}_t\xi^n, \xi^n) \geqslant \tfrac{1}{2}\bar{\partial}_t\|\alpha^{\frac{1}{2}}\xi^n\|^2,$$

and Young's inequality, we obtain

$$\bar{\partial}_t\|\alpha^{\frac{1}{2}}\xi^n\|^2 + (2\aleph_0 - \epsilon)\|\xi^n\|_1^2 \leqslant C[\|\bar{\partial}_t\rho^n\|^2 + \|\tau^n\|^2 + \|\rho^n\|^2 + \|\zeta_x^n\|^2 + \|\eta^n\|^2 + \|\epsilon_2^n\|^2 + \|I^n\|^2] + \lambda\|\xi^n\|^2 + C\Delta t\sum_{j=0}^{n-1}(\|\zeta_x^j\|^2 + \|\eta^j\|^2),$$

where

$$\|\bar{\partial}_t\rho^n\|^2 \leqslant Ch^{2(r+1)}\frac{1}{\Delta t}\int_{t_{n-1}}^{t_n}\|q_t(s)\|_{r+1}^2\,\mathrm{d}s,$$

$$\|\tau^n\|^2 \leqslant C\Delta t\int_{t_{n-1}}^{t_n}\|q_{tt}(s)\|^2\,\mathrm{d}s,$$

$$\|\epsilon_2^n\|^2 \leqslant C(\Delta t)^2\int_0^{t_n}[\|u_x(s)\|^2 + \|u_{xt}(s)\|^2]\,\mathrm{d}s,$$

and

$$\|I^n\|^2 \leqslant C\left[\|\eta^n\|^2 + \Delta t\sum_{j=0}^{n-1}\|\eta^j\|^2\right].$$

Summing from $n = 1$ to J, $1 \leqslant J \leqslant M$, we obtain using (3.8)

$$(\aleph_0 - \lambda\Delta t)\|\xi^J\|^2 + \Delta t\sum_{n=1}^{J}\|\xi^n\|_1^2 \leqslant C\Bigg[\|\xi^0\|^2 + \Delta t\sum_{n=0}^{J-1}\|\xi^n\|^2 + h^{2\min(r+1,k+1)}(\|u\|_{L^\infty(H^{k+1})}^2 + \|q\|_{L^\infty(H^{r+1})}^2 + \|q_t\|_{L^2(H^{r+1})}^2) + (\Delta t)^2(\|u\|_{L^2(H^1)}^2 + \|u_t\|_{L^2(H^1)}^2 + \|q_{tt}\|_{L^2(L^2)}^2)\Bigg]. \tag{3.9}$$

Choose Δt so that $(\aleph_0 - \lambda\Delta t) > 0$ and apply the discrete Gronwall lemma. Since $q_h(0) = \tilde{q}_h(0)$, we have $\xi^0 = 0$ and use of the triangle inequality completes the proof of (3.6).

To prove (3.7), we substitute (3.9) in (3.8) with $n = J$ to obtain the superconvergence estimate for ζ_x^J,

$$\|\zeta_x^J\| \leqslant C[h^{\min(r+1,k+1)}(\|u\|_{L^\infty(H^{k+1})} + \|q\|_{L^\infty(H^{r+1})} + \|q_t\|_{L^2(H^{r+1})}) + \Delta t(\|u\|_{L^2(H^1)} + \|u_t\|_{L^2(H^1)} + \|q_{tt}\|_{L^2(L^2)})].$$

Using the Poincaré inequality, (2.8)–(2.10) at $t = t_J$, and the triangle inequality, we complete the proof. □

REMARK 3.1

(i) Note that we can choose the L^2 projection or the interpolant of $q(0)$ as the initial approximation $q_h(0)$ instead of the elliptic projection $\tilde{q}_h(0)$. As a result, we may have an additional term $\|q(0)\|_{r+1}$ in the estimates of Theorem 3.1.

(ii) With $w_h = \bar{\partial}_t \xi^n$ in (3.6), it is a routine calculation to show that

$$\|q^J - Z^J\|_1 = \mathrm{O}(h^{\min(k+1,r)} + \Delta t),$$

and, for $1 < p \leqslant \infty$,

$$\|u^J - U^J\|_{L^p} + \|q^J - Z^J\|_{L^p} = \mathrm{O}(h^{\min(k+1,r+1)} + \Delta t).$$

4. Extension to problems in two and three space variables

In this section, we briefly describe the extension of the results for the one space variable case to (1.1) in two and three space variables. The specific problem that we consider is

$$\begin{aligned} u_t - \nabla \cdot (A\nabla u) - \int_0^t \nabla \cdot (B(t,s)\nabla u(s))\,\mathrm{d}s &= f(x,t), \quad x \in \Omega, \quad t \in (0,T], \\ u(x,t) &= 0, \quad x \in \partial\Omega, \quad t \in [0,T], \\ u(x,0) &= u_0(x), \quad x \in \Omega, \end{aligned} \tag{4.1}$$

where Ω is a bounded domain in R^d $(d = 2, 3)$ with boundary $\partial\Omega$. Here, $A = A(x)$ and $B(t,s) = B(x;t,s)$ are $d \times d$ matrices with smooth and bounded entries, and the matrix A is positive definite.

As in the one space variable case, we consider two formulations. Setting

$$A(x)\nabla u + \int_0^t B(t,s)\nabla u(s)\,\mathrm{d}s = \mathbf{q}, \tag{4.2}$$

we obtain

$$u_t - \nabla \cdot \mathbf{q} = f. \tag{4.3}$$

Similarly, with

$$A(x)\nabla u = \boldsymbol{\sigma}, \tag{4.4}$$

we rewrite (4.1) as

$$u_t - \nabla \cdot \boldsymbol{\sigma} - \int_0^t \nabla \cdot (\beta(t,s)\boldsymbol{\sigma}(s))\,\mathrm{d}s = f, \tag{4.5}$$

where $\beta(t,s) = A^{-1}B(t,s)$.

Let $\mathbf{L}^2(\Omega) = (L^2(\Omega))^d$ with inner product and norm

$$(\boldsymbol{\sigma}, \mathbf{w}) = \sum_{i=1}^{d} (\sigma_i, w_i) \quad \text{and} \quad \|\mathbf{w}\| = \left(\sum_{i=1}^{d} \|w_i\|^2 \right)^{1/2}.$$

Further, let $\mathbf{H}^m = (H^m(\Omega))^d$ with the usual inner product and norm. Let

$$\mathbf{H}(\mathrm{div};\Omega) = \{\mathbf{w} \in \mathbf{L}^2(\Omega) : \nabla \cdot \mathbf{w} \in L^2(\Omega)\}$$

with norm

$$\|\mathbf{w}\|_{\mathbf{H}(\mathrm{div};\Omega)} = (\|\mathbf{w}\|^2 + \|\nabla \cdot \mathbf{w}\|^2)^{1/2}.$$

We now discuss the analysis of the H^1-Galerkin mixed finite element method only for problem (4.2), (4.3) and comment on related results for (4.4), (4.5). As in Section 2, the weak form of (4.2), (4.3) is: find $\{u, \mathbf{q}\} : [0,T] \mapsto H_0^1 \times \mathbf{H}(\mathrm{div};\Omega)$ such that

$$(A\nabla u, \nabla v) + \int_0^t (B(t,s)\nabla u(s), \nabla v)\,\mathrm{d}s = (\mathbf{q}, \nabla v), \quad v \in H_0^1, \tag{4.6}$$

$$(\alpha \mathbf{q}_t, \mathbf{w}) + (\nabla \cdot \mathbf{q}, \nabla \cdot \mathbf{w}) = (\beta(t,t)\nabla u, \mathbf{w})$$
$$+ \int_0^t (\beta_t(t,s)\nabla u(s), \mathbf{w})\,\mathrm{d}s - (f, \nabla \cdot \mathbf{w}), \quad \mathbf{w} \in \mathbf{H}(\mathrm{div};\Omega), \tag{4.7}$$

where $\alpha = A^{-1}$ and $\beta(t,s) = A^{-1}B(t,s)$.

Let V_h and $\mathbf{W}_h$ be finite-dimensional subspaces of H_0^1 and $\mathbf{H}(\mathrm{div};\Omega)$, respectively, with the following approximation properties (Pani, 1998; Pehlivanov *et al.*, 1994): for non-negative integers k and r,

$$\inf_{v_h \in V_h} \{\|v - v_h\| + h\|v - v_h\|_1\} \leqslant Ch^{k+1}\|v\|_{k+1}, \quad v \in H_0^1 \cap H^{k+1},$$

and

$$\inf_{\mathbf{w}_h \in \mathbf{W}_h} \{\|\mathbf{w} - \mathbf{w}_h\| + h\|\mathbf{w} - \mathbf{w}_h\|_{\mathbf{H}(\mathrm{div};\Omega)}\} \leqslant Ch^{r+1}\|\mathbf{w}\|_{r+1}, \quad w \in \mathbf{H}^{r+1}.$$

The semidiscrete approximation to (4.6), (4.7) is to determine $\{u_h, \mathbf{q}_h\} : [0,T] \mapsto V_h \times \mathbf{W}_h$ such that

$$(A\nabla u_h, \nabla v_h) + \int_0^t (B(t,s)\nabla u_h(s), \nabla v_h)\,\mathrm{d}s = (\mathbf{q}_h, \nabla v_h), \quad v_h \in V_h, \tag{4.8}$$

$$(\alpha \mathbf{q}_{ht}, \mathbf{w}_h) + (\nabla \cdot \mathbf{q}_h, \nabla \cdot \mathbf{w}_h) = (\beta(t,t)\nabla u_h, \mathbf{w}_h)$$
$$+ \int_0^t (\beta_t(t,s)\nabla u_h(s), \mathbf{w}_h)\,\mathrm{d}s - (f, \nabla \cdot \mathbf{w}_h), \mathbf{w}_h \in \mathbf{W}_h, \tag{4.9}$$

with $u_h(0)$ and $\mathbf{q}_h(0)$ specified later.

Note that the semidiscrete H^1-Galerkin mixed finite element method for the system (4.4), (4.5) consists in determining $\{u_h, \boldsymbol{\sigma}_h\} : [0, T] \mapsto V_h \times \mathbf{W}_h$ such that

$$(A\nabla u_h, \nabla v_h) = (\boldsymbol{\sigma}_h, \nabla v_h), \quad v_h \in V_h, \tag{4.10}$$

$$(\alpha \boldsymbol{\sigma}_{ht}, \mathbf{w}_h) + (\nabla \cdot \boldsymbol{\sigma}_h, \nabla \cdot \mathbf{w}_h) + \int_0^t (\nabla \cdot (\beta(t,s)\boldsymbol{\sigma}_h(s)), \nabla \cdot \mathbf{w}_h)\,\mathrm{d}s = -(f, \nabla \cdot \mathbf{w}_h),\ \mathbf{w}_h \in \mathbf{W}_h, \tag{4.11}$$

with appropriately defined $u_h(0)$ and $\boldsymbol{\sigma}_h(0)$.

To determine the desired error estimates for the semidiscrete Galerkin approximation $\{u_h, \mathbf{q}_h\}$ of (4.6), (4.7), we define Ritz–Volterra projection $\tilde{u}_h \in V_h$ by

$$(A\nabla(u - \tilde{u}_h), \nabla v_h) + \int_0^t (B(t,s)\nabla(u - \tilde{u}_h)(s), \nabla v_h)\,\mathrm{d}s = 0,\ v_h \in V_h, \tag{4.12}$$

(Cannon & Lin, 1988, 1990; Lin *et al.*, 1991; Pani *et al.*, 1992), and the standard finite element interpolant $\tilde{\mathbf{q}}_h \in \mathbf{W}_h$ of $\mathbf{q}$. Let $\eta = u - \tilde{u}_h$ and $\boldsymbol{\rho} = \mathbf{q} - \tilde{\mathbf{q}}_h$. Then, for non-negative integers k and r, the following estimates hold:

$$\|\eta\| + h\|\eta\|_1 \leqslant Ch^{k+1}\left[\|u\|_{k+1} + \int_0^t \|u(s)\|_{k+1}\,\mathrm{d}s\right], \tag{4.13}$$

and

$$\|\boldsymbol{\rho}\| + \|\boldsymbol{\rho}_t\| + h\|\boldsymbol{\rho}\|_{\mathbf{H}(\mathrm{div};\Omega)} \leqslant Ch^{k+1}(\|\mathbf{q}\|_{r+1} + \|\mathbf{q}_t\|_{r+1}). \tag{4.14}$$

We now write

$$u - u_h = (u - \tilde{u}_h) + (\tilde{u}_h - u_h) := \eta + \zeta,$$

and

$$\mathbf{q} - \mathbf{q}_h = (\mathbf{q} - \tilde{\mathbf{q}}_h) + (\tilde{\mathbf{q}}_h - \mathbf{q}_h) := \boldsymbol{\rho} + \boldsymbol{\xi}.$$

From (4.6)–(4.9) and (4.12), we obtain

$$(A\nabla\zeta, \nabla v_h) + \int_0^t (B(t,s)\nabla\zeta(s), \nabla v_h)\,\mathrm{d}s = (\boldsymbol{\rho} + \boldsymbol{\xi}, \nabla v_h), \quad v_h \in V_h, \tag{4.15}$$

and

$$\begin{aligned}(\alpha\boldsymbol{\xi}_t, \mathbf{w}_h) + (\nabla \cdot \boldsymbol{\xi}, \nabla \cdot \mathbf{w}_h) + (\boldsymbol{\xi}, \mathbf{w}_h) &= -(\alpha\boldsymbol{\rho}_t, \mathbf{w}_h) - (\nabla \cdot \boldsymbol{\rho}, \nabla \cdot \mathbf{w}_h) + (\boldsymbol{\xi}, \mathbf{w}_h)\\ &+ (\beta(t,t)(\nabla\zeta + \nabla\eta), \mathbf{w}_h) + \int_0^t (\beta_t(t,s)(\nabla\zeta + \nabla\eta)(s), \mathbf{w}_h)\,\mathrm{d}s,\ \mathbf{w}_h \in \mathbf{W}_h.\end{aligned} \tag{4.16}$$

For the terms involving $\nabla\eta$ in (4.16), we use Green's theorem and rewrite (4.16) as

$$\begin{aligned}(\alpha\boldsymbol{\xi}_t, \mathbf{w}_h) + (\nabla \cdot \boldsymbol{\xi}, \nabla \cdot \mathbf{w}_h) + (\boldsymbol{\xi}, \mathbf{w}_h) &= -(\alpha\boldsymbol{\rho}_t, \mathbf{w}_h) - (\nabla \cdot \boldsymbol{\rho}, \nabla \cdot \mathbf{w}_h)\\ &+ (\boldsymbol{\xi}, \mathbf{w}_h) + (\beta(t,t)\nabla\zeta, \mathbf{w}_h) + \int_0^t (\beta_t(t,s)\nabla\zeta(s), \mathbf{w}_h)\,\mathrm{d}s\\ &- (\eta, \nabla \cdot (\beta(t,t)\mathbf{w}_h)) - \int_0^t (\eta, \nabla \cdot (\beta_t(t,s)\mathbf{w}_h))\,\mathrm{d}s, \quad \mathbf{w}_h \in \mathbf{W}_h.\end{aligned} \tag{4.17}$$

Since A is positive definite, we have

$$\lambda_{\min}\|\mathbf{w}\|^2 \leqslant (A\mathbf{w},\mathbf{w}) \leqslant \lambda_{\max}\|\mathbf{w}\|^2, \text{ and } \lambda_{\max}^{-1}\|\mathbf{w}\|^2 \leqslant (A^{-1}\mathbf{w},\mathbf{w}) \leqslant \lambda_{\min}^{-1}\|\mathbf{w}\|^2, \quad (4.18)$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalues of A, respectively. Since the proof of the following theorem follows essentially the same lines as the proofs of Theorems 2.1 and 2.2, we only present a brief sketch of it.

THEOREM 4.1 With $\mathbf{q}_0 = A\nabla u_0$, assume that $\mathbf{q}_h(0)$ is the standard finite element interpolant of $\mathbf{q}_0$, that is, $\mathbf{q}_h(0) = \tilde{\mathbf{q}}_h(0)$. Then

$$\begin{aligned}\|(u-u_h)(t)\| + \|(\mathbf{q}-\mathbf{q}_h)(t)\| &\leqslant Ch^{\min(k+1,r)}[\|\mathbf{q}_0\|_r + \|u\|_{L^\infty(H^{k+1})} \\ &\quad + \|\mathbf{q}\|_{L^2(H^{r+1})} + \|\mathbf{q}_t\|_{L^2(H^r)}],\\ \|(u-u_h)(t)\|_1 \leqslant Ch^{\min(k,r)}[\|\mathbf{q}_0\|_r &+ \|u\|_{L^\infty(H^{k+1})} + \|\mathbf{q}\|_{L^2(H^{r+1})} + \|\mathbf{q}_t\|_{L^2(H^r)}],\end{aligned}$$

and

$$\|(\mathbf{q}-\mathbf{q}_h)(t)\|_{\mathbf{H}(\mathrm{div};\Omega)} \leqslant Ch^{\min(k+1,r)}[\|\mathbf{q}_0\|_{r+1}+\|u_0\|_{k+1}+\|u_t\|_{L^2(H^{k+1})}+\|\mathbf{q}_t\|_{L^2(H^{r+1})}].$$

Proof. Since estimates of η and $\boldsymbol{\rho}$ are known from (4.13), (4.14), it is sufficient to estimate ζ and $\boldsymbol{\xi}$. Choose $v_h = \zeta$ in (4.15) and use (4.18) with the boundedness of B to obtain

$$\|\nabla\zeta\| \leqslant C\left[\|\boldsymbol{\rho}\| + \|\boldsymbol{\xi}\| + \int_0^t \|\nabla\zeta(s)\|\,\mathrm{d}s\right].$$

An application of the Gronwall lemma yields

$$\|\nabla\zeta\| \leqslant C\left[\|\boldsymbol{\rho}\| + \|\boldsymbol{\xi}\| + \int_0^t (\|\boldsymbol{\rho}(s)\| + \|\boldsymbol{\xi}(s)\|)\,\mathrm{d}s\right]. \quad (4.19)$$

Setting $\mathbf{w}_h = \boldsymbol{\xi}$ in (4.17), we use the Cauchy–Schwarz inequality and Young's inequality with (4.18) to obtain

$$\begin{aligned}\|\boldsymbol{\xi}(t)\|^2 + \int_0^t \|\boldsymbol{\xi}(s)\|^2_{\mathbf{H}(\mathrm{div};\Omega)}\,\mathrm{d}s &\leqslant C\bigg[\|\boldsymbol{\xi}(0)\|^2 + \int_0^t (\|\boldsymbol{\rho}_t\|^2 + \|\nabla\cdot\boldsymbol{\rho}\|^2 \\ &\quad + \|\nabla\zeta\|^2 + \|\eta\|^2 + \|\boldsymbol{\xi}\|^2)\,\mathrm{d}s\bigg] + C\int_0^t\int_0^s (\|\eta\|^2 + \|\boldsymbol{\xi}\|^2_{\mathbf{H}(\mathrm{div};\Omega)})\,\mathrm{d}\tau\,\mathrm{d}s.\end{aligned}$$

On substituting (4.19), an application of the Gronwall lemma then yields

$$\begin{aligned}&\|\boldsymbol{\xi}(t)\|^2 + \int_0^t \|\boldsymbol{\xi}\|^2_{\mathbf{H}(\mathrm{div};\Omega)}\,\mathrm{d}s \\ &\quad\leqslant Ch^{2\min(k+1,r)}\left[\|\mathbf{q}(0)\|_r^2 + \int_0^t (\|u\|^2_{k+1} + \|\mathbf{q}\|^2_{r+1} + \|\mathbf{q}_t\|_r^2)\,\mathrm{d}s\right]. \quad (4.20)\end{aligned}$$

Now substitute (4.20) and (4.14) in (4.19) to derive an estimate of ζ in H^1-norm.

In order to estimate $\|\boldsymbol{\xi}(t)\|_{\mathbf{H}(\mathrm{div};\Omega)}$, we choose $\mathbf{w}_h = \boldsymbol{\zeta}_t$ in (4.17) and repeat the arguments given in the proof of Theorem 2.2, the only exception being that the term $-(\nabla\cdot\boldsymbol{\rho}, \nabla\cdot\boldsymbol{\xi}_t)$ is replaced by

$$-\frac{\mathrm{d}}{\mathrm{d}t}(\nabla\cdot\boldsymbol{\rho}, \nabla\cdot\boldsymbol{\xi}) + (\nabla\cdot\boldsymbol{\rho}_t, \nabla\cdot\boldsymbol{\xi}).$$

Thus we easily obtain the estimate

$$\int_0^t \|\boldsymbol{\xi}_t(s)\|^2\,\mathrm{d}s + \|\boldsymbol{\xi}(t)\|^2_{\mathbf{H}(\mathrm{div};\Omega)} \leqslant C\bigg[\|\boldsymbol{\xi}(0)\|^2 + \|\boldsymbol{\rho}\|^2_{\mathbf{H}(\mathrm{div};\Omega)} + \|\nabla\zeta\|^2 + \|\eta\|^2 + \int_0^t (\|\boldsymbol{\rho}_t\|^2_{\mathbf{H}(\mathrm{div};\Omega)} + \|\eta\|^2 + \|\nabla\zeta\|^2 + \|\boldsymbol{\xi}\|^2_{\mathbf{H}(\mathrm{div};\Omega)})\,\mathrm{d}s\bigg].$$

The use of (4.13), (4.14) and (4.20) along with the triangle inequality completes the proof. □

REMARK 4.1

(i) Now we compare the order of convergence of the mixed finite element method described in Ewing *et al.* (2001); Jiang (1999) and that of the H^1-Galerkin mixed method discussed in Theorem 4.1. Assume that $\{V_h, \mathbf{W}_h\}$ is a pair of Raviart–Thomas spaces of index $r-1$, that is, the components of $\mathbf{W}_h$ contain incomplete polynomials of degree $r = k+1$ on each finite element. From Theorem 4.1 of Jiang (1999) (see also Ewing *et al.*, 2001), we obtain

$$\|(u-u_h)(t)\| + \|(\mathbf{q}-\mathbf{q}_h)(t)\| \leqslant Ch^r[\|u_0\|_r + \|\mathbf{q}_0\|_r + \|\mathbf{q}\|_{L^2(H^r)} + \|u_t\|_{L^2(H^r)} + \|\mathbf{q}_t\|_{L^2(H^r)}].$$

With $k+1 = r$ in Theorem 4.1, we observe that the rate of convergence of both methods is the same, but, for the H^1-Galerkin mixed finite element method, the LBB consistency condition has been avoided.

(ii) For the second formulation (4.4), (4.5) and (4.10), (4.11), we use an elliptic projection instead of the Ritz–Volterra projection as a comparison function, and then appropriately modify Theorem 4.1 to obtain similar *a priori* error estimates for $\boldsymbol{\sigma}-\boldsymbol{\sigma}_h$ and $u-u_h$. Since the arguments are quite routine, we omit the details.

(iii) The error estimates derived in Theorem 4.1 for $\|(u-u_h)(t)\|_1$ and $\|(\mathbf{q}-\mathbf{q}_h)(t)\|_{\mathbf{H}(\mathrm{div};\Omega)}$ with $k+1 = r$ are optimal with respect to the stated norms. However, the error estimate for $\|(\mathbf{q}-\mathbf{q}_h)(t)\|$ is not optimal in the $\mathbf{L}^2(\Omega)$-norm. This is in contrast to the optimal L^2-error estimates obtained in the one space variable case.

In order to derive optimal error estimate for $\mathbf{q}-\mathbf{q}_h$ in the $\mathbf{L}^2(\Omega)$-norm, a modification is proposed for purely parabolic problems in Pani (1998). Because of the presence of a nonlocal term in (1.1), there are certain difficulties in extending the modification to the present problem. However, if we choose the finite-dimensional space $\mathbf{W}_h$ as one of the Raviart–Thomas spaces RT_r with index r (that is, the components of $\mathbf{W}_h$ consist of incomplete polynomials of degree $r+1$ on each finite element) or Brezzi–Douglas–Marini spaces of index r, that is, BDM_r (see Brezzi & Fortin (1991) and Raviart & Thomas (1977)), it is possible to improve the $\mathbf{L}^2(\Omega)$-estimates of $\mathbf{q} - \mathbf{q}_h$. Now, instead of using the finite element interpolant as an auxiliary function, set $\tilde{\mathbf{q}}_h = \Pi_h \mathbf{q}$, where the Raviart–Thomas projection $\Pi_h \mathbf{q} : \mathbf{H}(\text{div}; \Omega) \mapsto \mathbf{W}_h$ is defined by

$$(\nabla \cdot (\mathbf{q} - \Pi_h \mathbf{q}), \nabla \cdot \mathbf{w}_h) = 0, \mathbf{w}_h \in \mathbf{W}_h.$$

With $\boldsymbol{\rho} = \mathbf{q} - \Pi_h \mathbf{q}$, the following error estimates hold (Brezzi & Fortin, 1991; Raviart & Thomas, 1977):

$$\|\boldsymbol{\rho}\| \leqslant C h^{r+1} \|\mathbf{q}\|_{r+1},$$

and

$$\|\nabla \cdot \boldsymbol{\rho}\| \leqslant C h^{r+1} \|\nabla \cdot \mathbf{q}\|_{r+1} \leqslant C h^{r+1} \|\mathbf{q}\|_{r+2}.$$

Since Π_h commutes with the time derivative, we obtain

$$\|\boldsymbol{\rho}_t\| \leqslant C h^{r+1} \|\mathbf{q}_t\|_{r+1}.$$

Now write

$$\mathbf{q} - \mathbf{q}_h = (\mathbf{q} - \Pi_h \mathbf{q}) + (\Pi_h \mathbf{q} - \mathbf{q}_h) := \boldsymbol{\rho} + \boldsymbol{\xi}.$$

The term $-(\nabla \cdot \boldsymbol{\rho}, \nabla \cdot \mathbf{w}_h)$ in (4.17) now vanishes and hence the equation in $\boldsymbol{\xi}$ can be written as

$$\begin{aligned}(\alpha \boldsymbol{\xi}_t, \mathbf{w}_h) + (\nabla \cdot \boldsymbol{\xi}, \nabla \cdot \mathbf{w}_h) + (\boldsymbol{\xi}, \mathbf{w}_h) &= -(\alpha \boldsymbol{\rho}_t, \mathbf{w}_h)(\boldsymbol{\xi}, \mathbf{w}_h) \\ &+ (\beta(t,t)\nabla\zeta, \mathbf{w}_h) + \int_0^t (\beta_t(t,s)\nabla\zeta(s), \mathbf{w}_h)\,\mathrm{d}s - (\eta, \nabla \cdot (\beta(t,t)\mathbf{w}_h)) \\ &- \int_0^t (\eta, \nabla \cdot (\beta_t(t,s)\mathbf{w}_h))\,\mathrm{d}s, \quad \mathbf{w}_h \in \mathbf{W}_h.\end{aligned}$$

Proceeding exactly as in the proof of Theorem 4.1, we finally obtain

$$\begin{aligned}&\|(u - u_h)(t)\| + \|(\mathbf{q} - \mathbf{q}_h)(t)\| \\ &\quad \leqslant C h^{\min(k+1,r+1)} [\|\mathbf{q}_0\|_{r+1} + \|u_0\|_{k+1} + \|u_t\|_{L^2(H^{k+1})} + \|\mathbf{q}_t\|_{L^2(H^{r+1})}].\end{aligned}$$

Moreover,

$$\begin{aligned}&\|(\mathbf{q} - \mathbf{q}_h)(t)\|_{\mathbf{H}(\text{div};\Omega)} \leqslant C h^{\min(k+1,r+1)} \\ &\quad \times [\|\mathbf{q}_0\|_{r+1} + \|u_0\|_{k+1} + \|\mathbf{q}\|_{L^2(H^{r+2})} \|u_t\|_{L^2(H^{k+1})} + \|\mathbf{q}_t\|_{L^2(H^{r+1})}].\end{aligned}$$

Note that the L^2-estimate of $\mathbf{q} - \mathbf{q}_h$ is optimal in the stated norm if $k = r$ and this is achieved provided we use $\mathbf{W}_h$ as the Raviart–Thomas spaces of index r or the BDM spaces of index r. However, it is possible to use other classical mixed finite element spaces (Brezzi & Fortin, 1991) for approximating $\mathbf{q}$ that preserve the L^2-optimality of the error $\mathbf{q} - \mathbf{q}_h$.

Acknowledgements

A.K.P. thanks the Department of Mathematical and Computer Sciences, Colorado School of Mines, for financial support and hospitality during his sabbatical leave from the Indian Institute of Technology, Bombay. The research of G.F. was supported by the National Science Foundation grant DMS9805827. The authors thank the referee for some useful suggestions.

REFERENCES

BREZZI, F. & FORTIN, M. (1991) *Mixed and Hybrid Finite Element Methods*. New York: Springer.

CANNON, J. & LIN, Y. (1988) Non-classical H^1 projection and Galerkin methods for nonlinear parabolic integro-differential equations. *Calcolo*, **25**, 187–201.

CANNON, J. & LIN, Y. (1990) A priori L^2 error estimates for finite element methods for nonlinear differential equations with memory. *SIAM J. Numer. Anal.*, **27**, 595–607.

CAREY, G. F. & SHEN, Y. (1989) Convergence studies of least-square finite elements for first-order systems. *Commun. Appl. Numer. Meth.*, **5**, 427–434.

CUSHMAN, J. & GLINN, T. (1993) Nonlocal dispersion in media with continuously evolving scales of heterogeneity. *Transport in Porous Media*, **13**, 123–138.

DAGAN, G. (1994) The significance of heterogeneity of evolving scales to transport in porous formations. *Water Resour. Res.*, **30**, 3327–3336.

EWING, R., LAZAROV, R. & LIN, Y. (2000) Finite volume element approximations of nonlocal in time one-dimensional flows in porous media. *Computing*, **64**, 157–182.

EWING, R., LIN, Y. & WANG, J. (2001) A numerical approximation of non-Fickian flows with mixing length growth in porous media. *Acta. Math. Univ. Comenian.*, **LXX**, 75–84.

JIANG, Z. (1999) $L^\infty(L^2)$ and $L^\infty(L^\infty)$ error estimates for mixed methods for integro-differential equations of parabolic type. *M^2AN Math. Model. Numer. Anal.*, **33**, 531–546.

JOHNSON, C. & THOMÉE, V. (1981) Error estimates for some mixed finite element methods for parabolic type problems. *RAIRO Numer. Anal.*, **15**, 41–78.

LIN, Y., THOMÉE, V. & WAHLBIN, L. B. (1991) Ritz–Volterra projection onto finite element spaces and application to integro-differential and related equations. *SIAM J. Numer. Anal.*, **28**, 1047–1070.

NEITTAANMÄKI, P. & SARANEN, J. (1981) On finite element approximation of the gradient for solution of Poisson equation. *Numer. Math.*, **37**, 333–337.

NEITTAANMÄKI, P. & SARANEN, J. (1981) On finite element approximation of vector fields by curl and divergence. *Math. Meth. Appl. Sci.*, **3**, 328–335.

NEITTAANMÄKI, P. & SARANEN, J. (1981) A mixed finite element method for the heat flow problem. *BIT*, **21**, 342–346.

PANI, A. K. (1998) An H^1-Galerkin mixed finite element method for parabolic partial differential equations. *SIAM J. Numer. Anal.*, **35**, 721–727.

PANI, A. K., ANDERSSEN, R. S. & CHUNG, S. K. On convergence of finite difference schemes for generalised solutions of parabolic and hyperbolic integro-differential equations. *Research Report CMA-R3-91*. Canberra: Australian National University.

PANI, A. K., THOMÉE, V. & WAHLBIN, L. B. (1992) Numerical methods for hyperbolic and parabolic integro-differential equations. *J. Integral Equ. Appl.*, **4**, 533–584.

PEHLIVANOV, A. I., CAREY, G. F., LAZAROV, R. D. & SHEN, Y. (1993) Convergence analysis of least-square mixed finite elements. *Computing*, **51**, 111–123.

PEHLIVANOV, A. I., CAREY, G. F. & LAZAROV, R. D. (1994) Least-square mixed finite elements for second-order elliptic problems. *SIAM J. Numer. Anal.*, **31**, 1368–1377.

RAVIART, P. A. & THOMAS, J. M. (1977) *A Mixed Finite Element Method for Second-order Elliptic Problems*, Lecture Notes in Mathematics 606. New York: Springer, pp. 293–315.

RENARDY, M., HRUSA, W. & NOHEL, J. (1987) *Mathematical Problems in Viscoelasticity*, Pitman Monographs and Surveys in Pure Appl. Math. No. 35. New York: Wiley.

SLOAN, I. H. & THOMÉE, V. (1986) Time discretization of an integro-differential equation of parabolic type. *SIAM J. Numer. Anal.*, **23**, 1052–1061.

YANIK, E. G. & FAIRWEATHER, G. (1988) Finite element methods for parabolic and hyperbolic integro-differential equations. *Nonlinear Anal. Theor. Meth. Appl.*, **12**, 785–809.

WHEELER, M. F. (1973) *A priori* L^2-error estimates for Galerkin approximations to parabolic differential equations. *SIAM J. Numer. Anal.*, **10**, 723–749.