A Nonlinear Mixed Finite Element Method for a Degenerate Parabolic Equation Arising in Flow in Porous Media

Todd Arbogast
Mary F. Wheeler
Nai-Ying Zhang

CRPC-TR94452
May 1994

Center for Research on Parallel Computation
Rice University
P.O. Box 1892
Houston, TX 77251-1892

This work was partially supported by the CRPC through the NSF and by the State of Texas Governor's Energy Office.

A NONLINEAR MIXED FINITE ELEMENT METHOD FOR A DEGENERATE PARABOLIC EQUATION ARISING IN FLOW IN POROUS MEDIA*

TODD ARBOGAST†, MARY F. WHEELER†, AND NAI-YING ZHANG‡

Abstract. We study a model nonlinear, degenerate, advection-diffusion equation having application in petroleum reservoir and groundwater aquifer simulation. The main difficulty is that the true solution is typically lacking in regularity; therefore, we consider the problem from the point of view of optimal approximation. Through time integration, we develop a mixed variational form that respects the known minimal regularity, and then we develop and analyze two versions of a mixed finite element approximation, a simpler semidiscrete (time continuous) version and a fully discrete version. Our error bounds are optimal in the sense that all but one of the bounding terms reduce to standard approximation error. The exceptional term is a nonstandard approximation error term. We also consider our new formulation for the nondegenerate problem, showing the usual optimal L_2-error bounds; moreover, superconvergence is obtained under special circumstances.

Key words. Mixed finite element, degenerate parabolic equation, nonlinear, error estimates, porous media

AMS(MOS) subject classifications. Primary 65M60, 65M12, 65M15; secondary 35K65, 76S05

1. Introduction. Let $\Omega \subset \mathbf{R}^d$, $d = 1$, 2, or 3, be a bounded domain with sufficiently smooth boundary $\partial\Omega$, and let $0 < T < \infty$ and $J = (0, T\,]$. We develop and analyze a mixed finite element approximation to the following nonlinear advection-diffusion problem in $u(x,t)$:

$$\frac{\partial u}{\partial t} - \nabla \cdot [\alpha \nabla P(u) + \beta(P(u))] = \gamma(P(u)), \quad (x,t) \in \Omega \times J, \tag{1.1a}$$

$$u = u_D, \quad (x,t) \in \partial\Omega \times J, \tag{1.1b}$$

$$u = u_0, \quad (x,t) \in \Omega \times \{0\}, \tag{1.1c}$$

where $P(u) = P(x,t;u)$ is strictly monotone increasing in u for each $(x,t) \in \Omega \times J$, $\gamma(u) = \gamma(x,t;u)$, $u_D = u_D(x,t)$, $u_0 = u_0(x)$, $\beta(u) = \beta(x,t;u)$ is a vector, and $\alpha = \alpha(x,t)$ is a $d \times d$ symmetric matrix that is uniformly positive definite with respect to $(x,t) \in \bar{\Omega} \times \bar{J}$. These functions are tacitly assumed to be smooth enough for our purposes.

We concentrate on the case in which $\partial P(u)/\partial u = P_u(u)$ may be zero for some values of u. Since $\nabla P(u) = P_u(x,t;u)\nabla u + \nabla_x P(x,t;u)$, (1.1) is degenerate

* This work was partially supported by the Center for Research on Parallel Computation through NSF Cooperative agreement No. CCR-8809615 and by the State of Texas Governor's Energy Office through contract #1059 for the Geophysical Parallel Computation Project.

† Department of Computational and Applied Mathematics, Rice University, P.O. Box 1892, Houston, Texas 77251-1892

‡ Center for Research on Parallel Computation, P.O. Box 1892, Rice University, Houston, Texas 77251-1892; current address, CLever Sys. Inc., 1334 Stokley Way, Vienna, Virginia 22182

parabolic. Let $(\cdot,\cdot)$ denote the $L_2(\Omega)$-inner product (or sometimes the duality pairing) and $\|\cdot\|$ its norm. Our main assumptions are that there is a constant $C_0 > 0$, independent of time, such that

$$(\mathbf{A1}) \quad \|P(\varphi_1) - P(\varphi_2)\|^2 \le C_0(P(\varphi_1) - P(\varphi_2), \varphi_1 - \varphi_2), \quad \text{for } \varphi_1, \varphi_2 \in L_2(\Omega),$$

and both β and γ are Lipschitz continuous:

$$(\mathbf{A2}) \quad \|\beta(\varphi_1) - \beta(\varphi_2)\| + \|\gamma(\varphi_1) - \gamma(\varphi_2)\| \le C_0\|\varphi_1 - \varphi_2\|, \quad \text{for } \varphi_1, \varphi_2 \in L_2(\Omega).$$

A sufficient condition for $(\mathbf{A1})$ is

$$(\mathbf{A1}') \quad 0 \le P_u(x,t;\varphi) \le C_0 \qquad \text{for } (x,t) \in \Omega \times J \text{ and } \varphi \text{ in the range of the true solution}$$

(when considering numerical schemes, this inequality must hold also on the range of the numerical solution, so extend P in some reasonable way).

To obtain below a mixed formulation, we introduce a new variable

$$\psi = -\alpha\nabla P(u) - \beta(P(u)). \tag{1.2}$$

The main difficulty in approximating (1.1) is that the solution is typically lacking in regularity. According to Alt and Luckhaus [2] (see also [1], [4], and [15]), we have at least that

$$u \in L_\infty(J; L_1(\Omega)), \tag{1.3a}$$

$$u_t \in L_2(J; H^{-1}(\Omega)), \tag{1.3b}$$

$$\psi \in L_2(J; (L_2(\Omega))^d), \tag{1.3c}$$

$$\gamma(P(u)) \in L_2(J; L_2(\Omega)), \tag{1.3d}$$

where H^{-1} is the dual of H_0^1. Furthermore, if we assume that the problem is physically consistent so that a maximum principle holds (e.g., $\beta(P(u))$ is zero for two values of u, our initial and boundary conditions stay between these two values, and the source term $\gamma(P(u))$ respects the range of u), then u remains bounded. Let us simply assume that

$$(\mathbf{A3}) \qquad u \in L_\infty(J; L_\infty(\Omega)).$$

Because of (1.3b)–(1.3c), it is natural to consider conforming finite element discretizations of (1.1). We mention four such works below.

Rose [28], [29] considered a similar problem for flow through porous media. He defined a continuous, piecewise linear finite element Galerkin method and derived rates of convergence based on assumed asymptotic rates of degeneracy. Once such rates are assumed, the solution can be shown to have more regularity (e.g., u_t is a function, not merely a distribution), which he then exploited.

Magenes, Nochetto, and Verdi [20] considered a class of problems including the Stephan problem and the porous medium equation; their results apply also to (1.1). Their scheme is discrete in time only. They relax the strict equality (1.2) by using the asymptotically correct (as the time step tends to zero) Chernoff formulation.

Nochetto and Verdi [25] consider a similar degenerate parabolic equation. They defined a continuous, piecewise linear finite element Galerkin method and proved its convergence; moreover, they extracted error estimates in measure for the free boundaries that appear in the solutions.

Barrett and Knabner [7] considered the problem of solute transport (see Section 2). They also defined a continuous, piecewise linear finite element Galerkin method, and they used a regularization of the problem to obtain their results.

In the petroleum industry, equations similar to (1.1) (see Section 2) are most often discretized by using the cell-centered finite difference method [26]. As shown in [30], [32], [6], this scheme is actually the lowest order Raviart-Thomas mixed finite element method on rectangles [27], combined with special quadrature rules. The mixed method for the nondegenerate problem has been well studied (see, e.g., [27], [14], [21]); however, it appears that no convergence theory has been presented for the fully degenerate problem (1.1).

Let $a = \alpha^{-1}$. For almost every time, a mixed variational form of (1.1) is

$$(u_t, w) + (\nabla \cdot \psi, w) = \big(\gamma(P(u)), w\big), \quad \forall w \in H_0^1(\Omega), \tag{1.4a}$$

$$\begin{aligned} &(a\psi, v) - (P(u), \nabla \cdot v) + \big(a\beta(P(u)), v\big) \\ &\qquad = -\langle P(u_D), v \cdot \nu\rangle, \quad \forall v \in H(\Omega; \mathrm{div}), \end{aligned} \tag{1.4b}$$

$$(u(\cdot, 0), w) = (u_0, w), \quad \forall w \in L_2(\Omega), \tag{1.4c}$$

where $H(\Omega; \mathrm{div}) = \{v \in L_2(\Omega) : \nabla \cdot v \in L_2(\Omega)\}$ and $\langle \cdot, \cdot \rangle$ denotes the usual inner product in $L_2(\partial\Omega)$, or the duality pairing. Since we can only expect in general that $u_t \in L_2(J; H^{-1}(\Omega))$, in the straightforward mixed formulation (1.4), this requires that the trial functions in (1.4a) belong to $H_0^1(\Omega)$. To avoid this, we derive below an alternate variational formulation incorporating an integration in time.

We consider the problem from the point of view of optimal approximation, regardless of the rate at which $P(u)$ tends to zero. We show that our scheme approximates the true solution about as well as can be expected for our approximating spaces. Our error bounds are optimal in the sense that all bounding terms reduce to approximation error, except one. This latter term involves the difference of two discrete projections of the integral time average of the total flux. For the Raviart-Thomas rectangular spaces [27], these two projections are super-close. We can recover actual rates of convergence of the scheme as soon as some regularity is shown for the solution.

The outline of the rest of this paper is as follows. In the next section, we provide two practical examples of (1.1) that serve to motivate our work. In Section 3 we present a different mixed variational formulation and two versions of a mixed finite element method. One is semidiscrete (continuous in time), the other is fully discrete. The semidiscrete version, though not computable, is easier to understand and gives some insight into our treatment. It is analyzed in Section 4, and the fully discrete version is analyzed in Section 5. This appears to be the first proof that the type of discretization schemes used in the petroleum industry–namely mixed methods–converge for the fully degenerate problem. In the last two sections, we consider our new formulation for the nondegenerate problem. The usual optimal L_2 error bounds are derived in Section 6; moreover, in Section 7, we prove that superconvergence is obtained under special circumstances. It appears that we have the first proof of superconvergence for the vector flux variable in the nonlinear problem, though superconvergence had been observed experimentally [31].

2. Two applications. Problem (1.1) appears in many applications; we motivate our work by describing two of them. Petroleum reservoir and groundwater aquifer simulation often requires the solution of a nonlinear, degenerate, advection-diffusion problem describing two-phase flow in porous media [1], [4], [8], [13], [26]. We restrict our discussion here to a model equation possessing the degeneracy emphasized in the introduction:

$$\frac{\partial(\phi s)}{\partial t} + \nabla \cdot [\tilde{\alpha}(s) K \nabla p_c(s) + v\tilde{\beta}(s)] = \tilde{\gamma}(s), \quad (x,t) \in \Omega \times J, \tag{2.1}$$

where $0 \leq s = s(x,t) \leq 1$ stands for the (normalized) wetting fluid phase saturation, ϕ is the porosity (uniformly positive and bounded), p_c is the capillary pressure function, K is the tensor of absolute permeability, v is the total Darcy fluid velocity, $\tilde{\alpha}$ and $\tilde{\beta}$ are related to the phase mobilities, and $\tilde{\gamma}$ models the effect of wells.

For this problem, $\tilde{\alpha}(s) = \tilde{\alpha}(x,t;s) \geq 0$ and vanishes if, and only if, $s = 0$ or $s = 1$; thus, (**A3**) holds. Also $p_c(s) = p_c(x;s)$ is strictly monotone decreasing. Let $A(s) = \tilde{\alpha}(p_c^{-1}(s))$, and denote by $P(s)$ the modified Kirchhoff transformation

$$P(s) = \int_0^{p_c(s)} A(\pi)\, d\pi. \tag{2.2}$$

Since

$$\nabla P(s) = \tilde{\alpha}(s)\nabla p_c(s) + \int_0^{p_c(s)} \nabla_x A(\pi)\, d\pi,$$

setting

$$\beta(P(s)) = v\tilde{\beta}(s) - K\int_0^{p_c(s)} \nabla_x A(\pi)\, d\pi$$

and shifting from s to the unknown ϕs casts (2.1) in the form of (1.1).

If $\tilde{\beta}(P^{-1}(\cdot))$ and $\tilde{\gamma}(P^{-1}(\cdot))$ are Lipschitz continuous, and $|\nabla_x A(\cdot)| \leq CA(\cdot)$, then (**A2**) holds. Although (**A2**) is somewhat artificial for this problem, we can demonstrate that (**A1′**) is physically reasonable.

If $\lambda_o(s)$ and $\lambda_w(s)$ are the oil and water phase mobilities, respectively, then $\tilde{\alpha} = \lambda_o\lambda_w/(\lambda_o + \lambda_w)$. By Leverett's semi-empirical equation, $p_c(x;s) = \kappa(x)J(s)$, and by Burdine's relationship between relative permeability and capillary pressure, we can assume that

$$\lambda_w(s) \sim s^2 \int_0^s \frac{d\zeta}{J^2(\zeta)} \qquad \text{and} \qquad \lambda_o(s) \sim (1-s)^2 \int_s^1 \frac{d\zeta}{J^2(1-\zeta)}$$

(see, e.g., [8], [13]). If, say, $J(s) \sim s^{\delta_1}$, $0 < \delta_1 < 1$, as $s \to 0$ and $J(s) \sim (1-s)^{\delta_2}$, $0 < \delta_2 < 1$, as $s \to 1$, then

$$P_s(s) = \tilde{\alpha}(s)\frac{\partial p_c(s)}{\partial s} \to 0 \text{ as } s \to 0 \text{ or } s \to 1,$$

which establishes (**A1**$'$). (The quantity $P(s)$ can be considered as a "complementary" pressure. See Arbogast [4] for a more detailed discussion.)

Another application is a macroscopic model for the transport of a solute with concentration $c(x,t)$ in a porous medium with an equilibrium adsorption reaction, such as

$$\frac{\partial}{\partial t}\big(\Theta c + \rho\varphi(c)\big) - \nabla\cdot(D\nabla c - vc) = 0 \quad \text{in } \Omega\times J, \tag{2.3}$$

subject to initial and boundary conditions. Here $\Theta = \Theta(x)$ is a function with uniformly positive upper and lower bounds, $\rho = \rho(x) \geq \rho_0 > 0$, D is the diffusion/dispersion tensor, v is the fluid velocity, and $\varphi(\cdot)$ is a the sorption isotherm, a non-decreasing function with $\varphi(0) = 0$ and $\varphi(c) > 0$ for $c > 0$ (see [15], [18]). With $\mu(c) = \Theta c + \rho\varphi(c)$, this equation is

$$\frac{\partial\mu(c)}{\partial t} - \nabla\cdot[D\nabla\mu^{-1}(\mu(c)) - v\mu^{-1}(\mu(c))] = 0 \quad \text{in } \Omega\times J, \tag{2.4}$$

which is (1.1) for the variable μ with $P(\mu) = \mu^{-1}(\mu)$. Easily, (**A1**$'$), (**A2**), and (**A3**) hold, since $0 \leq \partial P/\partial\mu \leq 1/\Theta$, $\beta = vP$, $\gamma = 0$, and $0 \leq c \leq 1$. In fact, the problem is nondegenerate if the Langmuir isotherm is used, or if the exponent for the Freundlich isotherm is greater than or equal to one.

3. The mixed finite element method. In this section, we develop first a semidiscrete (time continuous) mixed finite element method for the degenerate problem, and then a fully discrete version. Our algorithms are well defined even when the true solution is minimally regular, as described in the introduction (recall (1.3)). We begin by deriving an appropriate mixed variational formulation of (1.1).

From (1.3), both u and u_t are in $L_2(J; H^{-1}(\Omega))$; we therefore conclude that $u \in C^0(J; H^{-1}(\Omega))$ (see Chapter 1 of [19]). This gives us $u(\cdot,t)$ pointwise for every $t \in \bar{J}$, first as a distribution in $H^{-1}(\Omega)$, but actually in $L_\infty(\Omega)$ by (**A3**).

We are justified now in integrating (1.1a) in time from 0 to $t \in J$ and using (1.1c) to obtain the equivalent distributional equation

$$u(x,t) + \nabla\cdot\int_0^t \psi\, d\tau = \int_0^t \gamma(P(u))\, d\tau + u_0(x), \quad (x,t)\in\Omega\times J, \tag{3.1a}$$

$$u = u_D, \quad (x,t)\in\partial\Omega\times J. \tag{3.1b}$$

Note that from (1.3), we can conclude that

$$\int_0^t \psi\, d\tau \in H^1(J;(L_2(\Omega))^d) \cap L_2(J; H(\Omega;\mathrm{div}));$$

thus, we have a variational form for almost every time $t \in \bar{J}$ as

$$
\begin{aligned}
&(u(\cdot,t),w) + \left(\nabla \cdot \int_0^t \psi \, d\tau, w\right) \\
&\qquad = \left(\int_0^t \gamma(P(u)) \, d\tau, w\right) + (u_0, w), \quad \forall w \in L_2(\Omega),
\end{aligned} \tag{3.2a}
$$

$$
\begin{aligned}
&(a\psi, v) - (P(u), \nabla \cdot v) + \big(a\beta(P(u)), v\big) \\
&\qquad = -\langle P(u_D), v \cdot \nu \rangle, \quad \forall v \in H(\Omega; \mathrm{div}).
\end{aligned} \tag{3.2b}
$$

In fact, (3.2a) holds for *every* $t \in \bar{J}$, since u is defined for each time; moreover, we can *define* ψ for every time by (3.2b).

Let Ω be partitioned into a conforming finite element mesh with maximal element diameter h. We seek approximate solutions in a mixed finite element space $\mathcal{W}_h \times \mathcal{V}_h \subset L_2(\Omega) \times H(\Omega; \mathrm{div})$ defined over the mesh, e.g. the Raviart-Thomas-Nedelec finite element spaces [27], [23] (or those of [9], [10], [11], [12], or [24]).

A semidiscrete mixed finite element method.

For each $t \in \bar{J}$, let $(U(\cdot,t), \Psi(\cdot,t)) \in \mathcal{W}_h \times \mathcal{V}_h$ be the approximation of $(u(\cdot,t), \psi(\cdot,t))$ such that

$$
\begin{aligned}
&(U(\cdot,t),W) + \left(\nabla \cdot \int_0^t \Psi \, d\tau, W\right) \\
&\qquad = \left(\int_0^t \gamma(P(U)) \, d\tau, W\right) + (u_0, W), \quad \forall W \in \mathcal{W}_h,
\end{aligned} \tag{3.3a}
$$

$$
\begin{aligned}
&(a\Psi, V) - (P(U), \nabla \cdot V) + \big(a\beta(P(U)), V\big) \\
&\qquad = -\langle P(u_D), V \cdot \nu \rangle, \quad \forall V \in \mathcal{V}_h.
\end{aligned} \tag{3.3b}
$$

We turn to our backward Euler, fully discrete scheme. Let $t_0 = 0 < t_1 < \cdots < t_N = T$ partition J, and let $\Delta t^n = t_n - t_{n-1}$ be the nth time step size. For any function φ of time, let φ^n denote $\varphi(t_n)$; we also abuse the notation by writing $P(\varphi^n)$ in place of $P(\cdot, t_n; \varphi^n)$.

A fully discrete mixed finite element method.

For each $n \geq 0$, let $(U^n, \Psi^n) \in \mathcal{W}_h \times \mathcal{V}_h$ be the approximation of (u^n, ψ^n) such that

$$
\begin{aligned}
&(U^n,W) + \left(\nabla \cdot \sum_{j=1}^n \Psi^j \, \Delta t^j, W\right) \\
&\qquad = \left(\sum_{j=1}^n \gamma^j(P(U^j)) \, \Delta t^j, W\right) + (u_0, W), \quad \forall W \in \mathcal{W}_h,
\end{aligned} \tag{3.4a}
$$

$$
\begin{aligned}
&(a^n\Psi^n, V) - (P(U^n), \nabla \cdot V) + \big(a^n\beta^n(P(U^n)), V\big) \\
&\qquad = -\langle P(u_D^n), V \cdot \nu \rangle, \quad \forall V \in \mathcal{V}_h.
\end{aligned} \tag{3.4b}
$$

Note that in practical computation, we would use a more straightforward equivalent form of (3.4a) as follows. Subtract (3.4a) at time level n from that at level $n-1$, and divide by Δt^n to obtain for $n \geq 1$ that

$$\left(\frac{U^n - U^{n-1}}{\Delta t^n}, W\right) + (\nabla \cdot \Psi^n, W) = \big(\gamma^n(P(U^n)), W\big), \quad \forall W \in \mathcal{W}_h, \tag{3.5a}$$

$$(U^0, W) = (u_0, W), \quad \forall W \in \mathcal{W}_h. \tag{3.5b}$$

4. Analysis of the semidiscrete scheme. We begin this section by defining some projection operators. We need a projection operator Π_h mapping into $\mathcal{V}_h$ with the property that for any v in the domain of the operator,

$$\big(\nabla \cdot (\Pi_h v - v), W\big) = 0, \quad \forall W \in \mathcal{W}_h. \tag{4.1}$$

We explicitly assume that

(**A4**) $\quad \nabla \cdot \mathcal{V}_h \subset \mathcal{W}_h$ and there exists $\Pi_h : (H^1(\Omega))^d \to \mathcal{V}_h$ satisfying (4.1).

All the usual mixed spaces (see [9], [10], [11], [12], [14], [23], [24], [27]) satisfy (**A4**); moreover, Π_h approximates I (the identity operator) to the optimal order.

Denote by $\hat{\mathcal{P}}_h : L_2(\Omega) \to \mathcal{W}_h$ the $L_2(\Omega)$-projection operator. For any function $\varphi \in L_2(\Omega)$, $\hat{\varphi}$ denotes $\hat{\mathcal{P}}_h \varphi$ for short. Finally, associated with $\alpha(\cdot, t)$, we introduce the weighted $(L_2(\Omega))^d$-projection operator $\mathcal{P}_h(t) : (L_2(\Omega))^d \to \mathcal{V}_h$ defined, for $v \in (L_2(\Omega))^d$ and all $V \in \mathcal{V}_h$, by

$$\big(a(\cdot, t)(\mathcal{P}_h(t) v - v), V\big) = 0, \quad \forall V \in \mathcal{V}_h. \tag{4.2}$$

We need to apply Π_h to $\int_0^t \psi \, d\tau$, which can be done only if this function is sufficiently smooth. We assume explicitly that

(**A5**) $$\int_0^t \psi \, d\tau \in H^1(J; (L_2(\Omega))^d) \cap L_2(J; (H^1(\Omega))^d) = (H^1(\Omega \times J))^d.$$

Lemma 1. *If $\alpha(x,t) = \alpha_1(t)\,\alpha_2(x)$, where α_1 is a scalar, then* (**A5**) *holds.*

Proof. Recall that (1.2) defines ψ. Since $u \in L_\infty(J; L_\infty(\Omega))$, $\beta(P(u)) \in L_2(J; L_2(\Omega))$; moreover, (1.3c) implies that in fact $P(u) \in L_2(J; H^1(\Omega))$, and then also $\beta(P(u)) \in L_2(J; H^1(\Omega))$. Time integration only helps matters, so we conclude from (3.1a) that

$$\nabla \cdot \int_0^t \alpha \nabla P(u) \, d\tau = \nabla \cdot \alpha_2 \nabla \int_0^t \alpha_1 P(u) \, d\tau \in L_2(J; L_2(\Omega)).$$

We have assumed that $\partial\Omega$ is sufficiently smooth, so elliptic regularity implies that in fact

$$\int_0^t P(u) \, d\tau \in L_2(J; H^2(\Omega)),$$

yielding the lemma. $\square$

Remark. The tensor α in petroleum flow is a function of x only [8], [13], [26]. For solute transport, at least in the nondegenerate case, (**A5**) holds. In general, (**A5**) can be avoided if one changes (**A4**) to state that $\nabla \cdot \mathcal{V}_h = \mathcal{W}_h$. Then Π_h :

$H(\Omega;\mathrm{div}) \to \mathcal{V}_h$ can be defined satisfying (4.1); though, nonuniqueness requires that a choice be made. We no longer know that Π_h approximates I, so define $\Pi_h v \in \mathcal{V}_h$ such that $\|\Pi_h v - v\|$ is minimal subject to the constraint that $\nabla\cdot\Pi_h v = \hat{\mathcal{P}}_h \nabla\cdot v$.

Theorem 1. *Assume* (**A1**)–(**A5**). *Let* (u,ψ) *solve problem* (1.1) *and* (U,Ψ) *solve its semidiscrete mixed finite element approximation* (3.3). *There is some constant* $C>0$ *such that for any* $t\in J$,

$$\int_0^t (U-u, P(U)-P(u))\,d\tau + \left\| \int_0^t \Psi\,d\tau - \mathcal{P}_h(t)\int_0^t \psi\,d\tau \right\|^2$$
$$\le C\biggl\{ \int_0^t \|\hat{\mathcal{P}}_h u - u\|^2\,d\tau + \int_0^t \left\| (\mathcal{P}_h(\tau)-I)\int_0^\tau \psi\,d\sigma \right\|^2 d\tau$$
$$+ \int_0^t \left\| \nabla\cdot(\Pi_h - \mathcal{P}_h(\tau))\int_0^\tau \psi\,d\sigma \right\|^2 d\tau \biggr\}.$$

Remark. The form
$$\left\{ \int_0^t (U-u, P(U)-P(u))\,d\tau \right\}^{1/2}$$
bounds the size of $U-u$; for example, it bounds the norm $\|P(U)-P(u)\|$ by (**A1**). It is *not*, however, a norm itself. It may even fail to be a metric.

Proof. Let $\Phi = \Psi - \psi$, $\mathring{\Phi} = \int_0^t \Phi\,d\tau$, and $\mathring{\psi} = \int_0^t \psi\,d\tau$. By (3.2), (3.3), (4.1), and the fact that $\nabla\cdot\mathcal{V}_h \subset \mathcal{W}_h$, we have that

$$\begin{aligned}&(U-u,W) + (\nabla\cdot\Pi_h\mathring{\Phi}, W)\\ &\quad= \left(\int_0^t [\gamma(P(U)) - \gamma(P(u))]\,d\tau, W\right), \quad \forall W\in\mathcal{W}_h,\end{aligned} \tag{4.3a}$$

$$\begin{aligned}&(a\Phi, V) - (\widehat{P(U)} - \widehat{P(u)}, \nabla\cdot V)\\ &\quad= -(a[\beta(P(U)) - \beta(P(u))], V), \quad \forall V\in\mathcal{V}_h.\end{aligned} \tag{4.3b}$$

Take $W = \widehat{P(U)} - \widehat{P(u)} \in \mathcal{W}_h$ and $V = \mathcal{P}_h\mathring{\Phi} = \Pi_h\mathring{\Phi} + (\Pi_h - \mathcal{P}_h)\mathring{\psi} \in \mathcal{V}_h$ above. Add these two equations together, cancel the two terms $(\widehat{P(U)} - \widehat{P(u)}, \nabla\cdot\Pi_h\mathring{\Phi})$, and use (4.2) to obtain

$$\begin{aligned}&\left(U-u, \widehat{P(U)} - \widehat{P(u)}\right) + (a\mathcal{P}_h\Phi, \mathcal{P}_h\mathring{\Phi})\\ &\quad= (\widehat{P(U)} - \widehat{P(u)}, \nabla\cdot(\Pi_h - \mathcal{P}_h)\mathring{\psi})\\ &\qquad+ \left(\int_0^t [\gamma(P(U)) - \gamma(P(u))]\,d\tau, \widehat{P(U)} - \widehat{P(u)}\right)\\ &\qquad- (a[\beta(P(U)) - \beta(P(u))], \mathcal{P}_h\mathring{\Phi}).\end{aligned} \tag{4.4}$$

We integrate this equation in time from 0 to t. The first term on the left-hand side becomes

$$\int_0^t \Big(U-u, \widehat{P(U)}-\widehat{P(u)}\Big)\,d\tau = \int_0^t (U-u, P(U)-P(u))\,d\tau - T_1, \tag{4.5}$$

where

$$T_1 \equiv \int_0^t (\hat{u}-u, P(U)-P(u))\,d\tau.$$

For the second term on the left-hand side of (4.4), set $v=\mathring{\Phi}$ in (4.2) and differentiate in time to obtain

$$\big(a_t(\cdot,t)(\mathcal{P}_h(t)\mathring{\Phi}-\mathring{\Phi}),V\big)+\big(a(\cdot,t)[(\mathcal{P}_h(t)\mathring{\Phi})_t-\Phi],V\big)=0, \quad \forall V\in\mathcal{V}_h. \tag{4.6}$$

Taking $V=\mathcal{P}_h\mathring{\Phi}$ and using (4.2) gives

$$\begin{aligned}(a\mathcal{P}_h\Phi,\mathcal{P}_h\mathring{\Phi}) &= \big(a(\mathcal{P}_h\mathring{\Phi})_t,\mathcal{P}_h\mathring{\Phi}\big)+\big(a_t(\mathcal{P}_h(t)\mathring{\Phi}-\mathring{\Phi}),\mathcal{P}_h\mathring{\Phi}\big)\\ &= \frac{1}{2}\left[\frac{d}{dt}(a\mathcal{P}_h\mathring{\Phi},\mathcal{P}_h\mathring{\Phi})-\big(a_t\mathcal{P}_h\mathring{\Phi},\mathcal{P}_h\mathring{\Phi}\big)\right]+\big(a_t(\mathcal{P}_h(t)\mathring{\Phi}-\mathring{\Phi}),\mathcal{P}_h\mathring{\Phi}\big).\end{aligned} \tag{4.7}$$

Thus

$$\int_0^t \big(a\mathcal{P}_h\Phi,\mathcal{P}_h\mathring{\Phi}\big)\,d\tau = \tfrac{1}{2}\big\|a^{1/2}(\cdot,t)(\mathcal{P}_h\mathring{\Phi})(t)\big\|^2 - T_2, \tag{4.8}$$

where

$$T_2 \equiv \frac{1}{2}\int_0^t \big(a_t\mathcal{P}_h\mathring{\Phi},\mathcal{P}_h\mathring{\Phi}\big)\,d\tau + \int_0^t \big(a_t(\mathcal{P}_h\mathring{\psi}-\mathring{\psi}),\mathcal{P}_h\mathring{\Phi}\big)\,d\tau.$$

Collecting (4.4)–(4.8) together, we obtain that

$$\int_0^t (U-u,P(U)-P(u))\,d\tau + \tfrac{1}{2}\big\|a^{1/2}(\cdot,t)(\mathcal{P}_h\mathring{\Phi})(t)\big\|^2 = \sum_{k=1}^{5} T_k, \tag{4.9}$$

where

$$\begin{aligned} T_3 &\equiv \int_0^t \big(\widehat{P(U)}-\widehat{P(u)},\nabla\cdot(\Pi_h-\mathcal{P}_h)\mathring{\psi}\big)\,d\tau,\\ T_4 &\equiv \int_0^t \left(\int_0^\tau [\gamma(P(U))-\gamma(P(u))]\,d\sigma, \widehat{P(U)}-\widehat{P(u)}\right)d\tau,\\ T_5 &\equiv -\int_0^t \big(a[\beta(P(U))-\beta(P(u))],\mathcal{P}_h\mathring{\Phi}\big)\,d\tau.\end{aligned}$$

We estimate each T_k, $k=1,\dots,5$. For any $\epsilon>0$, we have

$$|T_1| \le C\int_0^t \|\hat{u}-u\|^2\,d\tau + \epsilon\int_0^t \|P(U)-P(u)\|^2\,d\tau,$$

where C is a generic positive constant independent of any discretization parameters. Easily

$$|T_2| \le C\left\{ \int_0^t \|\mathcal{P}_h \overset{\circ}{\Phi}\|^2 \, d\tau + \int_0^t \|(\mathcal{P}_h - I)\overset{\circ}{\psi}\|^2 \, d\tau \right\},$$

$$|T_3| \le \epsilon \int_0^t \|\widehat{P(U)} - \widehat{P(u)}\|^2 \, d\tau + C \int_0^t \|\nabla \cdot (\Pi_h - \mathcal{P}_h)\overset{\circ}{\psi}\|^2 \, d\tau.$$

By Assumption (**A2**),

$$|T_4| \le \epsilon \int_0^t \|P(U) - P(u)\|^2 \, d\tau + C \int_0^t \int_0^\tau \|P(U) - P(u)\|^2 \, d\sigma \, d\tau,$$

$$|T_5| \le \epsilon \int_0^t \|P(U) - P(u)\|^2 \, d\tau + C \int_0^t \|\mathcal{P}_h \overset{\circ}{\Phi}\|^2 \, d\tau.$$

Combining these estimates and using (**A1**) twice, for ϵ sufficiently small, we obtain

$$\begin{aligned}
&\int_0^t (U - u, P(U) - P(u)) \, d\tau + \|(\mathcal{P}_h \overset{\circ}{\Phi})(t)\|^2 \\
&\quad \le C\left\{ \int_0^t \|\hat{u} - u\|^2 \, d\tau + \int_0^t \|(\mathcal{P}_h - I)\overset{\circ}{\psi}\|^2 \, d\tau + \int_0^t \|\nabla \cdot (\Pi_h - \mathcal{P}_h)\overset{\circ}{\psi}\|^2 \, d\tau \right. \\
&\qquad \left. + \int_0^t \int_0^\tau (U - u, P(U) - P(u)) \, d\sigma \, d\tau + \int_0^t \|\mathcal{P}_h \overset{\circ}{\Phi}\|^2 \, d\tau \right\}.
\end{aligned} \tag{4.10}$$

Use of Gronwall's inequality to remove the last two terms completes the proof. □

Theorem 1 gives a bound for the time integral of $\|P(U) - P(u)\|$, but it does not give a bound for any norm of $U - u$. We give now such a bound for the negative norm, defined by

$$\|\cdot\|_{H^{-1}(\Omega)} = \sup_{\varphi \in H_0^1(\Omega)} \frac{(\,\cdot\,, \varphi)}{\|\varphi\|_{H^1(\Omega)}}.$$

Theorem 2. *Assume* (**A1**)–(**A5**). *Let* (u, ψ) *solve problem* (1.1) *and* (U, Ψ) *solve its semidiscrete mixed finite element approximation* (3.3). *There is some constant* $C > 0$ *such that for any* $t \in J$,

$$\begin{aligned}
&\|U(\cdot, t) - u(\cdot, t)\|_{H^{-1}(\Omega)} \\
&\quad \le C\left\{ h\|\hat{\mathcal{P}}_h u(\cdot, t) - u(\cdot, t)\| + \left\| \int_0^t \Psi \, d\tau - \Pi_h \int_0^t \psi \, d\tau \right\| + \int_0^t \|P(U) - P(u)\| \, d\tau \right\}.
\end{aligned}$$

Proof. Let $\varphi \in H_0^1(\Omega)$ and $\hat{\varphi} = \hat{\mathcal{P}}_h \varphi \in \mathcal{W}_h$. By (4.3a) we have that

$$\begin{aligned}
(U - u, \varphi) &= (U - u, \varphi - \hat{\varphi}) + (U - u, \hat{\varphi}) = (\hat{u} - u, \varphi - \hat{\varphi}) + (U - u, \hat{\varphi}) \\
&= (\hat{u} - u, \varphi - \hat{\varphi}) - \big(\nabla \cdot \Pi_h \overset{\circ}{\Phi}, \hat{\varphi}\big) + \left(\int_0^t [\gamma(P(U)) - \gamma(P(u))] \, d\tau, \hat{\varphi} \right).
\end{aligned} \tag{4.11}$$

Since $\nabla \cdot \Pi_h \overset{\circ}{\Phi} \in \mathcal{W}_h$, integration by parts gives

$$-\big(\nabla \cdot \Pi_h \overset{\circ}{\Phi}, \hat{\varphi}\big) = -\big(\nabla \cdot \Pi_h \overset{\circ}{\Phi}, \varphi\big) = \big(\Pi_h \overset{\circ}{\Phi}, \nabla\varphi\big). \tag{4.12}$$

Thus

$$\begin{aligned}(U-u,\varphi) &= \big(\hat{u}-u, \varphi-\hat{\varphi}\big) + \big(\Pi_h \overset{\circ}{\Phi}, \nabla\varphi\big) + \left(\int_0^t [\gamma(P(U)) - \gamma(P(u))]\,d\tau, \hat{\varphi}\right) \\ &\le C\left\{h\|\hat{u}-u\| + \|\Pi_h \overset{\circ}{\Phi}\| + \int_0^t \|P(U)-P(u)\|\,d\tau\right\}\|\varphi\|_{H^1(\Omega)},\end{aligned} \tag{4.13}$$

and the theorem follows. ∎

Remark. The error $\|U-u\|_{H^{-1}(\Omega)}$ is bounded by Theorems 1 and 2 and (**A1**) in terms of approximation theory.

5. Analysis of the fully discrete scheme. Based on our semidiscrete analysis, we derive analogous results for the fully discrete scheme. We need to assume that there is some $C_1 \ge 1$ such that

$$\Delta t^n \le C_1 \Delta t^{n-1}, \quad \forall n = 2, ..., N. \tag{\textbf{A6}}$$

Theorem 3. *Assume* (**A1**)–(**A6**). *Let* (u,ψ) *solve problem* (1.1) *and* (U^n, Ψ^n) *solve its fully discrete mixed finite element approximation* (3.4). *There is some constant* $C > 0$ *such that if the* Δt^j *are sufficiently small, then for any* n *between 1 and* N,

$$\begin{aligned}&\sum_{j=1}^n (U^j - u^j, P(U^j) - P(u^j))\,\Delta t^j + \left\|\sum_{j=1}^n \Psi^j\,\Delta t^j - \mathcal{P}_h^n \int_0^{t_n} \psi\,d\tau\right\|^2 \\ &\quad\le C\sum_{j=1}^n \Bigg\{\|\hat{\mathcal{P}}_h u^j - u^j\|^2 + \left\|(\mathcal{P}_h^j - I)\int_0^{t_j}\psi\,d\tau\right\|^2 + \left\|\nabla\cdot(\Pi_h - \mathcal{P}_h^j)\int_0^{t_j}\psi\,d\tau\right\|^2 \\ &\qquad + \left\|\sum_{i=1}^j \gamma^i(P(u^i))\,\Delta t^i - \int_0^{t_j}\gamma(P(u))\,d\tau\right\|^2 + \left\|\frac{1}{\Delta t^j}\int_{t_{j-1}}^{t_j}\psi\,d\tau - \psi^j\right\|^2\Bigg\}\Delta t^j.\end{aligned}$$

Moreover,

$$\begin{aligned}\|U^n - u^n\|_{H^{-1}(\Omega)} \le C\Bigg\{& h\|\hat{\mathcal{P}}_h u^n - u^n\| + \left\|\sum_{j=1}^n \Psi^j\,\Delta t^j - \Pi_h\int_0^{t_n}\psi\,d\tau\right\| \\ &+ \sum_{j=1}^n \|P(U^j) - P(u^j)\|\,\Delta t^j + \left\|\sum_{j=1}^n \gamma^j(P(u^j))\,\Delta t^j - \int_0^{t_n}\gamma(P(u))\,d\tau\right\|^2\Bigg\}.\end{aligned}$$

Proof. Let

$$\bar\psi^n = \frac{1}{\Delta t^n}\int_{t_{n-1}}^{t_n} \psi \, d\tau, \qquad \mathring{\bar\psi}^n = \sum_{j=1}^n \bar\psi^j \,\Delta t^j = \int_0^{t_n} \psi\, d\tau,$$
$$\Phi^n = \Psi^n - \bar\psi^n, \qquad \mathring{\Phi}^n = \sum_{j=1}^n \Phi^j\,\Delta t^j = \sum_{j=1}^n \Psi^j\,\Delta t^j - \int_0^{t_n} \psi\, d\tau.$$

Taking together (3.2) at $t = t_n$ and (3.4), and replacing W by $\widehat{P(U^n)} - \widehat{P(u^n)}$ and V by $\mathcal{P}_h^n \mathring{\Phi}^n = \Pi_h \mathring{\Phi}^n + (\Pi_h - \mathcal{P}_h^n)\mathring{\bar\psi}^n$, we obtain the following analogue of (4.4):

$$\begin{aligned}
&(U^n - u^n, \widehat{P(U^n)} - \widehat{P(u^n)}) + (a^n \mathcal{P}_h^n \Phi^n, \mathcal{P}_h^n \mathring{\Phi}^n)\\
&\quad = \big(\widehat{P(U^n)} - \widehat{P(u^n)}, \nabla\cdot(\Pi_h - \mathcal{P}_h^n)\mathring{\bar\psi}^n\big)\\
&\qquad + \Big(\sum_{j=1}^n \gamma^j(P(U^j))\,\Delta t^j - \int_0^{t_n} \gamma(P(u))\, d\tau, \widehat{P(U^n)} - \widehat{P(u^n)}\Big)\\
&\qquad - \big(a^n[\beta^n(P(U^n)) - \beta^n(P(u^n))], \mathcal{P}_h^n \mathring{\Phi}^n\big) - \big(a^n(\bar\psi^n - \psi^n), \mathcal{P}_h^n \mathring{\Phi}^n\big).
\end{aligned} \tag{5.1}$$

Note that $\mathring{\Phi}^{n-1} = \mathring{\Phi}^n - \Phi^n \Delta t^n$; thus, we have by (4.2) the identity

$$\begin{aligned}
&\big((a^n - a^{n-1})(\mathcal{P}_h^{n-1}\mathring{\Phi}^{n-1} - \mathring{\Phi}^{n-1}), V\big)\\
&\qquad + \big(a^n(\mathcal{P}_h^n \mathring{\Phi}^n - \mathcal{P}_h^{n-1}\mathring{\Phi}^{n-1} - \Phi^n \Delta t^n), V\big) = 0, \quad \forall V \in \mathcal{V}_h.
\end{aligned} \tag{5.2}$$

Substitute $V = \mathcal{P}_h^n \mathring{\Phi}^n$ to obtain that

$$\begin{aligned}
&(a^n \Phi^n, \mathcal{P}_h^n \mathring{\Phi}^n)\Delta t^n\\
&\quad = \big((a^n - a^{n-1})(\mathcal{P}_h^{n-1}\mathring{\Phi}^{n-1} - \mathring{\Phi}^{n-1}), \mathcal{P}_h^n \mathring{\Phi}^n\big) + \big(a^n(\mathcal{P}_h^n \mathring{\Phi}^n - \mathcal{P}_h^{n-1}\mathring{\Phi}^{n-1}), \mathcal{P}_h^n \mathring{\Phi}^n\big)\\
&\quad = \tfrac{1}{2}\big[(a^n \mathcal{P}_h^n \mathring{\Phi}^n, \mathcal{P}_h^n \mathring{\Phi}^n) - (a^{n-1}\mathcal{P}_h^{n-1}\mathring{\Phi}^{n-1}, \mathcal{P}_h^{n-1}\mathring{\Phi}^{n-1})\big]\\
&\qquad + \tfrac{1}{2}\big[(a^n \mathcal{P}_h^n \mathring{\Phi}^n, \mathcal{P}_h^n \mathring{\Phi}^n) + (a^{n-1}\mathcal{P}_h^{n-1}\mathring{\Phi}^{n-1}, \mathcal{P}_h^{n-1}\mathring{\Phi}^{n-1})\big]\\
&\qquad - (a^n \mathcal{P}_h^{n-1}\mathring{\Phi}^{n-1}, \mathcal{P}_h^n \mathring{\Phi}^n) + \big((a^n - a^{n-1})(\mathcal{P}_h^{n-1}\mathring{\Phi}^{n-1} - \mathring{\Phi}^{n-1}), \mathcal{P}_h^n \mathring{\Phi}^n\big).
\end{aligned} \tag{5.3}$$

If we replace n by j above, multiply (5.1) through by Δt^j, sum on j from 1 to n, and use (5.3), the first term on the far right-hand side of (5.3) collapses and we obtain that

$$\sum_{j=1}^n (U^j - u^j, P(U^j) - P(u^j))\,\Delta t^j + \tfrac{1}{2}(a^n \mathcal{P}_h^n \mathring{\Phi}^n, \mathcal{P}_h^n \mathring{\Phi}^n) = \sum_{k=1}^6 T_k, \tag{5.4}$$

where, for any $\epsilon > 0$,

$$
\begin{aligned}
T_1 &\equiv \sum_{j=1}^{n} \big(\hat{u}^j - u^j, P(U^j) - P(u^j)\big)\,\Delta t^j \\
&\le C\sum_{j=1}^{n} \|\hat{u}^j - u^j\|^2 \Delta t^j + \epsilon \sum_{j=1}^{n} \|P(U^j) - P(u^j)\|^2 \Delta t^j, \\
T_2 &\equiv \sum_{j=1}^{n} \big\{ \tfrac{1}{2}\big[(a^j \mathcal{P}_h^j \mathring{\Phi}^j, \mathcal{P}_h^j \mathring{\Phi}^j) + (a^{j-1}\mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1}, \mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1})\big] \\
&\quad - (a^j \mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1}, \mathcal{P}_h^j \mathring{\Phi}^j) + \big((a^j - a^{j-1})(\mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1} - \mathring{\Phi}^{j-1}), \mathcal{P}_h^j\mathring{\Phi}^j\big)\big\} \\
&= \sum_{j=1}^{n} \Big\{ \tfrac{1}{2}\big[(a^j \mathcal{P}_h^j \mathring{\Phi}^j, \mathcal{P}_h^j \mathring{\Phi}^j) + (a^j \mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1}, \mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1}) - 2(a^j \mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1}, \mathcal{P}_h^j\mathring{\Phi}^j)\big] \\
&\quad - \tfrac{1}{2}\big((a^j - a^{j-1})\mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1}, \mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1}\big) \Big\} \\
&\quad - \big((a^j - a^{j-1})(\mathcal{P}_h^{j-1}\mathring{\bar{\psi}}^{j-1} - \mathring{\bar{\psi}}^{j-1}), \mathcal{P}_h^j \mathring{\Phi}^j\big) \\
&\le - \sum_{j=1}^{n} \Big\{ \frac{1}{2}\Big(\frac{a^j - a^{j-1}}{\Delta t^j}\mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1}, \mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1}\Big) \\
&\quad + \Big(\frac{a^j - a^{j-1}}{\Delta t^j}(\mathcal{P}_h^{j-1}\mathring{\bar{\psi}}^{j-1} - \mathring{\bar{\psi}}^{j-1}), \mathcal{P}_h^j \mathring{\Phi}^j \Big) \Big\}\,\Delta t^j \\
&\le C \sum_{j=1}^{n} \big[\|\mathcal{P}_h^{j-1}\mathring{\Phi}^{j-1}\|^2 + \|\mathcal{P}_h^j \mathring{\Phi}^j\|^2 + \|\mathcal{P}_h^{j-1}\mathring{\bar{\psi}}^{j-1} - \mathring{\bar{\psi}}^{j-1}\|^2 \big]\,\Delta t^j, \\
T_3 &\equiv \sum_{j=1}^{n} \big(\widehat{P(U^j)} - \widehat{P(u^j)}, \nabla\cdot(\Pi_h - \mathcal{P}_h^j)\mathring{\bar{\psi}}^j \big)\,\Delta t^j \\
&\le \epsilon \sum_{j=1}^{n} \|P(U^j) - P(u^j)\|^2 \Delta t^j + C\sum_{j=1}^{n} \|\nabla\cdot(\Pi_h - \mathcal{P}_h^j)\mathring{\bar{\psi}}^j\|^2 \Delta t^j, \\
T_4 &\equiv \sum_{j=1}^{n} \Big(\sum_{i=1}^{j} \gamma^i(P(U^i))\,\Delta t^i - \int_0^{t_j} \gamma(P(u))\,d\tau, \widehat{P(U^j)} - \widehat{P(u^j)} \Big)\Delta t^j \\
&\le C \sum_{j=1}^{n} \Big\{ \Big\| \sum_{i=1}^{j} \gamma^i(P(u^i))\,\Delta t^i - \int_0^{t_j} \gamma(P(u))\,d\tau \Big\|^2 \\
&\quad + \sum_{i=1}^{j} \|P(U^i) - P(u)\|^2\,\Delta t^i \Big\}\Delta t^j \\
&\quad + \epsilon \sum_{j=1}^{n} \|P(U^j) - P(u^j)\|^2 \Delta t^j,
\end{aligned}
$$

$$T_5 \equiv -\sum_{j=1}^{n} \left(a^j[\beta^j(P(U^j)) - \beta^j(P(u^j))], \mathcal{P}_h^j \mathring{\Phi}^j\right)\Delta t^j$$
$$\le \epsilon \sum_{j=1}^{n} \|P(U^j) - P(u^j)\|^2 \Delta t^j + C \sum_{j=1}^{n} \|\mathcal{P}_h^j \mathring{\Phi}^j\|^2 \Delta t^j,$$
$$T_6 \equiv -\sum_{j=1}^{n} \left(a^j(\bar{\psi}^j - \psi^j), \mathcal{P}_h^j \mathring{\Phi}^j\right) \Delta t^j$$
$$\le C\left\{ \sum_{j=1}^{n} \|\bar{\psi}^j - \psi^j\|^2 \Delta t^j + \sum_{j=1}^{n} \|\mathcal{P}_h^j \mathring{\Phi}^j\|^2 \Delta t^j \right\}.$$

An application of the discrete Gronwall inequality gives the first part of the theorem.

The second part of the theorem can be shown as in the proof of Theorem 2. ∎

6. Analysis of the nondegenerate case. The above results were derived under the assumption that the nonlinear function $P(\cdot)$ may vanish. In the next two sections, we consider the nondegenerate case. We assume that there exist two constants C_2 and C_3 such that

$$\textbf{(A7)} \qquad 0 < C_2 \le P_u(x,t;\varphi) \le C_3 < \infty \quad \text{for } (x,t) \in \Omega \times J \text{ and } \varphi \in \mathbf{R}.$$

Immediately

$$C_2|U-u| \le |P(U) - P(u)| \le C_3|U-u|, \tag{6.1}$$

and so Theorem 1 gives an error bound for $\int_0^t \|U-u\|^2\,d\tau$. With some extra effort, we can present an error estimate for $\sup_{t\in J}\|U(t)-u(t)\|$, and bound $\int_0^t \|\Psi-\psi\|^2\,d\tau$ at the same time. We exploit the fact that **(A7)** easily implies that

$$\tfrac{1}{2}C_2(U-u)^2 \le \int_u^U (P(\mu) - P(u))\,d\mu \le \tfrac{1}{2}C_3(U-u)^2. \tag{6.2}$$

Similar results hold for the fully discrete case as well.

Let us begin with the semidiscrete case. The true solution has the needed regularity, so we write the semidiscrete scheme as

$$(U_t, W) + (\nabla\cdot\Psi, W) = \big(\gamma(P(U)), W\big), \quad \forall W \in \mathcal{W}_h, \tag{6.3a}$$
$$(a\Psi, V) - (P(U), \nabla\cdot V) + \big(a\beta(P(U)), V\big) = -\langle P(u_D), V\cdot\nu\rangle, \quad \forall V \in \mathcal{V}_h, \tag{6.3b}$$
$$U(\cdot,0) = \hat{\mathcal{P}}_h u_0. \tag{6.3c}$$

Let $\Phi = \Psi - \psi$. Together with (1.4), we have the error equation

$$(U_t - u_t, W) + (\nabla\cdot\Pi_h\Phi, W) = \big(\gamma(P(U)) - \gamma(P(u)), W\big), \quad \forall W \in \mathcal{W}_h, \tag{6.4a}$$
$$\begin{aligned}(a\Pi_h\Phi, V) - (\widehat{P(U)} - \widehat{P(u)}, \nabla\cdot V) + \big(a[\beta(P(U)) - \beta(P(u))], V\big)& \\ = \big(a(\psi - \Pi_h\psi), V\big), \quad \forall V \in \mathcal{V}_h.&\end{aligned} \tag{6.4b}$$

Letting $W = \widehat{P(U)} - \widehat{P(u)}$ and $V = \Pi_h \Phi$ and adding the two equations together yields

$$
\begin{aligned}
&(U_t - u_t, P(U) - P(u)) + (a\Pi_h\Phi, \Pi_h\Phi) \\
&\quad = \big(\gamma(P(U)) - \gamma(P(u)), \widehat{P(U)} - \widehat{P(u)}\big) - \big(a[\beta(P(U)) - \beta(P(u))], \Pi_h\Phi\big) \\
&\qquad + (a(\psi - \Pi_h\psi), \Pi_h\Phi) + (\widehat{u}_t - u_t, P(U) - P(u)).
\end{aligned} \tag{6.5}
$$

The difficult term is the first one on the left-hand side. We assume explicitly that there is some positive constant C_4 such that

$$
|P_{uu}(\varphi)| + |P_{tu}(\varphi)| \le C_4 \quad \text{for } (x,t) \in \Omega \times J \text{ and } \varphi \in \mathbf{R}. \tag{\textbf{A8}}
$$

We employ an argument used by Wheeler and Dupont [33] and Arbogast [3] (see also [5; Lemma 2]). By Assumptions (**A7**)–(**A8**) and the Mean Value Theorem, for some $W_1(\mu)$ and W_2 between u and U, we have

$$
\begin{aligned}
&(U_t - u_t)(P(U) - P(u)) \\
&\quad = \frac{\partial}{\partial t}\int_u^U (P(\mu) - P(u))\,d\mu - \int_u^U (P_t(\mu) - P_t(u))\,d\mu \\
&\qquad + \big[(U-u)P_u(u) - (P(U) - P(u))\big]u_t \\
&\quad = \frac{\partial}{\partial t}\int_u^U (P(\mu) - P(u))\,d\mu - \int_u^U P_{tu}(W_1(\mu))(\mu - u)\,d\mu \\
&\qquad + (P_u(u) - P_u(W_2))(U-u)u_t \\
&\quad \ge \frac{\partial}{\partial t}\int_u^U (P(\mu) - P(u))\,d\mu - C\{(U-u)^2 + |(u - W_2)(U-u)|\}.
\end{aligned} \tag{6.6}
$$

Therefore, with (6.2),

$$
\begin{aligned}
&\int_0^t (U_t - u_t, P(U) - P(u))\,d\tau \\
&\quad \ge \int_\Omega \int_u^U (P(\mu) - P(u))\,d\mu\,dx - \int_\Omega \int_{u_0}^{U^0} (P(\mu) - P(u_0))\,d\mu\,dx \\
&\qquad - C\int_0^t \|U - u\|^2\,d\tau \\
&\quad \ge \tfrac{1}{2} C_2 \|U - u\|^2 - C\Big\{\|U^0 - u_0\|^2 + \int_0^t \|U - u\|^2\,d\tau\Big\}.
\end{aligned} \tag{6.7}
$$

Integrating (6.5) from 0 to t and using (6.7), (6.1), and Assumption (**A1**), we obtain

$$
\begin{aligned}
\|U - u\|^2 + \int_0^t \|\Pi_h\Phi\|^2\,d\tau \le C\Big\{&\|U^0 - u_0\|^2 + \int_0^t \|U - u\|^2\,d\tau \\
&+ \int_0^t \|\psi - \Pi_h\psi\|^2\,d\tau + \int_0^t \|\widehat{u}_t - u_t\|^2\,d\tau\Big\}.
\end{aligned} \tag{6.8}
$$

By Gronwall's lemma, we obtain our result.

Theorem 4. *Assume* (**A1**)–(**A5**), (**A7**)–(**A8**). *Let* (u,ψ) *solve problem* (1.1) *and* (U,Ψ) *solve its semidiscrete mixed finite element approximation* (3.3). *There is some constant* $C>0$ *such that for any* $t\in J$,

$$\|U(t)-u(t)\|^2+\int_0^t \|\Psi-\psi\|^2\,d\tau \le C\biggl\{\|\hat{\mathcal{P}}_h u_0-u_0\|^2+\int_0^t \|\psi-\Pi_h\psi\|^2\,d\tau+\int_0^t \|\hat{\mathcal{P}}_h u_t-u_t\|^2\,d\tau\biggr\}.$$

Remark. This result gives optimal order approximation if the solution is smooth enough.

Next we analyze the fully backward Euler discretization. Denoting

$$\bar{\partial}_t\varphi^n=\frac{\varphi^n-\varphi^{n-1}}{\Delta t^n},$$

the difference of (3.4b), (3.5) and (1.4) gives an error equation within which we substitute $W=\widehat{P(U^n)}-\widehat{P(u^n)}$ and $V=\Pi_h\Phi^n$. We then add the two main equations and, as before, obtain

$$\begin{aligned}(\bar{\partial}_t(U-u)^n, &P(U^n)-P(u^n))+(a^n\Pi_h\Phi^n,\Pi_h\Phi^n)\\ &=(\bar{\partial}_t(\hat{u}-u)^n,P(U^n)-P(u^n))-(\bar{\partial}_t u^n-u_t^n,\widehat{P(U^n)}-\widehat{P(u^n)})\\ &\quad+(\gamma^n(P(U^n))-\gamma^n(P(u^n)),\widehat{P(U^n)}-\widehat{P(u^n)})\\ &\quad-(a^n[\beta^n(P(U^n))-\beta^n(P(u^n))],\Pi_h\Phi^n)+(a^n(\psi^n-\Pi_h\psi^n),\Pi_h\Phi^n).\end{aligned}\tag{6.9}$$

To handle the first term on the left-hand side, we need a generalization of (6.6) to the case of discrete time as given in [5; Lemma 2] (see also [4]). The result is

$$\bar{\partial}_t(U-u)^n(P(U^n)-P(u^n))=\bar{\partial}_t\biggl(\int_u^U (P(\mu)-P(u))\,d\mu\biggr)^n-E^n,\tag{6.10}$$

where

$$E^n\le C\{|U^n-u^n|^2+|U^{n-1}-u^{n-1}|^2+(\Delta t^n)^2\}.\tag{6.11}$$

Replace n by j in (6.9), multiply by Δt^j, and sum from 1 to n. Using (6.1)–(6.2), (6.10)–(6.11), and noting that the first term on the right-hand side of (6.10) collapses, we obtain

$$\begin{aligned}\|U^n-u^n\|^2&+\sum_{j=1}^n \|\Pi_h\Phi^j\|^2\,\Delta t^j\\ &\le C\biggl\{\|U^0-u_0\|^2+(\Delta t)^2+\sum_{j=1}^n \|U^j-u^j\|^2\,\Delta t^j\\ &\quad+\sum_{j=1}^n \bigl[\|\bar{\partial}_t(\hat{u}-u)^j\|^2+\|\bar{\partial}_t u^j-u_t^j\|^2+\|\psi^j-\Pi_h\psi^j\|^2\bigr]\,\Delta t^j\biggr\},\end{aligned}\tag{6.12}$$

where Δt is the maximal Δt^j. The second term in the final summation is time truncation error, boundable by $(\Delta t)^2$, while the first such term is bounded as

$$\begin{aligned}\sum_{j=1}^{n} \|\bar{\partial}_t(\hat{u}-u)^j\|^2\,\Delta t^j &= \sum_{j=1}^{n} \frac{1}{\Delta t^j}\left\| \int_{t_{j-1}}^{t_j} (\hat{u}-u)_t\,d\tau \right\|^2 \\ &\le \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} \|(\hat{u}-u)_t\|^2\,d\tau = \int_0^{t_n} \|(\hat{u}-u)_t\|^2\,d\tau.\end{aligned} \tag{6.13}$$

We conclude the following theorem by the discrete Gronwall lemma.

Theorem 5. *Assume* (**A1**)–(**A8**). *Let* (u,ψ) *solve problem* (1.1) *and* (U^n,Ψ^n) *solve its fully discrete mixed finite element approximation* (3.4). *There is some constant* C *such that if the* Δt^j *are sufficiently small, then for* n *between 1 and* N,

$$\begin{aligned}&\|U^n-u^n\|^2 + \sum_{j=1}^{n} \|\Psi^j-\psi^j\|^2\,\Delta t^j \\ &\quad\le C\left\{ \|\hat{\mathcal{P}}_h u_0 - u_0\|^2 + (\Delta t)^2 + \int_0^{t_n} \|\hat{P}_h u_t - u_t\|^2\,d\tau + \sum_{j=1}^{n} \|\psi^j - \Pi_h\psi^j\|^2\,\Delta t^j \right\}.\end{aligned}$$

Remark. Again this result gives optimal order approximation if the solution is smooth enough.

7. Some superconvergence results. In this section, we present some superconvergence results for the nondegenerate case. Such results are known for linear elliptic problems under the hypotheses that α is diagonal and the grid is rectangular [22], [17], [16], or merely that the grid is rectangular [6]. We need to assume that $\beta = 0$, and that there is a constant $C_5 = C_5(u)$ such that

(**A9**) $\qquad |\nabla\hat{\mathcal{P}}_h u| \le C_5(u) \quad$ for $x \in \bar{E}$, and E an element of the mesh.

This holds, for instance, if $\mathcal{W}_h$ consists of piecewise constants defined over the given mesh, e.g., the lowest order Raviart-Thomas spaces, or for other spaces that possess the inverse inequality. We further assume that

(**A10**) $\qquad \chi W \in \mathcal{W}_h \quad$ for $W \in \mathcal{W}_h$ and any piecewise constant χ.

This holds for all the usual spaces. We will use the following general lemma.

LEMMA 2. *Assume* (**A9**)–(**A10**). *If* $f(x;u)$ *and* $g(x,t;u)$ *are functions defined on* $\Omega \times \mathbf{R}$ *and* $\Omega \times J \times \mathbf{R}$, *respectively, for which*

$$\begin{aligned}&|\nabla_x f_u(\varphi)| + |f_{uu}(\varphi)| + |\nabla_x g_u(\varphi)| + |g_{uu}(\varphi)| + |\nabla_x g_{uu}(\varphi)| + |g_{uuu}(\varphi)| \\ &\qquad + |\nabla_x g_{tu}(\varphi)| + |g_{tuu}(\varphi)| \le C' \quad \text{for } (x,t) \in \Omega \times J \text{ and } \varphi \in \mathbf{R}\end{aligned}$$

for some constant C', *then there exists a constant* C *such that for any* $W \in \mathcal{W}_h$,

$$\begin{aligned}&\big(f(\hat{\mathcal{P}}_h u) - f(u), W\big) \le Ch\,\|\hat{\mathcal{P}}_h u - u\|\,\|W\|, \\ &\left(\frac{\partial}{\partial t}[g(\hat{\mathcal{P}}_h u) - g(u)], W\right) \le Ch\left\{ \left\| \frac{\partial}{\partial t}[\hat{\mathcal{P}}_h u - u] \right\| + \|\hat{\mathcal{P}}_h u - u\| \right\} \|W\|,\end{aligned}$$

and, for any n such that $1 \le n \le N$,

$$\begin{aligned}&\big(\bar\partial_t(g(\hat{\mathcal{P}}_h u) - g(u))^n, W\big)\\ &\qquad \le Ch\frac{1}{\Delta t^n}\int_{t_{n-1}}^{t_n}\left(\left\|\frac{\partial}{\partial t}[\hat{\mathcal{P}}_h u - u]\right\| + \|\hat{\mathcal{P}}_h u - u\|\right)dt\,\|W\|.\end{aligned}$$

Proof. For the first result, let

$$\rho = \int_0^1 f_u(\xi\hat u + (1-\xi)u)\,d\xi, \tag{7.1}$$

and introduce ρ_h as the $L_2(\Omega)$-projection of ρ into the finite element space consisting of the piecewise constant functions defined over the given mesh. We have

$$f(\hat u) - f(u) = \rho\,(\hat u - u) = (\rho - \rho_h)(\hat u - u) + \rho_h\,(\hat u - u), \tag{7.2}$$

so with (**A9**)–(**A10**), on an element E,

$$\begin{aligned}\|\rho - \rho_h\|_{L_\infty(E)} \le Ch\|\nabla\rho\|_{L_\infty(E)} \le Ch\bigg\{&\left\|\int_0^1 \nabla_x f_u(\xi\hat u + (1-\xi)u)\,d\xi\right\|_{L_\infty(E)}\\ &+\left\|\int_0^1 f_{uu}(\xi\hat u + (1-\xi)u)(\xi\nabla\hat u + (1-\xi)\nabla u)\,d\xi\right\|_{L_\infty(E)}\bigg\} \le C(u)\,h.\end{aligned} \tag{7.3}$$

From (**A10**),

$$(\rho_h(\hat u - u), W) = (\hat u - u, \rho_h W) = 0. \tag{7.4}$$

The first part of the lemma is a combination of (7.2)–(7.4).

For the second result, note that

$$\begin{aligned}(g(\hat u))_t - (g(u))_t &= g_t(\hat u) - g_t(u) + g_u(\hat u)\hat u_t - g_u(u)u_t\\ &= g_t(\hat u) - g_t(u) + [g_u(\hat u) - g_u(u)]u_t + g_u(\hat u)(\hat u_t - u_t).\end{aligned}$$

The first two differences on the far right-hand side are estimated by the first result of the lemma (with $g_t(\cdot)$ and $u_t g_u(\cdot)$ as our functions). The last term is estimated for a good choice of the piecewise constant function χ as

$$\begin{aligned}(g_u(\hat u)(\hat u_t - u_t), W) &= \big((g_u(\hat u) - \chi)(\hat u_t - u_t), W\big)\\ &\le Ch\,\|\nabla g_u(\hat u)\|_{L_\infty(\Omega)}\|\hat u_t - u_t\|\,\|W\|.\end{aligned} \tag{7.5}$$

Now the second result follows by noting that $\nabla g_u(\hat u) = \nabla_x g_u(\hat u) + g_{uu}(\hat u)\nabla\hat u$.

The third result follows from the second, since

$$\bar\partial_t(g(\hat{\mathcal{P}}_h u) - g(u))^n = \frac{1}{\Delta t^n}\int_{t_{n-1}}^{t_n}\frac{\partial}{\partial t}[g(\hat{\mathcal{P}}_h u) - g(u)]\,dt. \qquad \square$$

Define $\Gamma(\cdot) = \gamma(P(\cdot))$. To use Lemma 2, we assume that for some constant C_6,

$$\begin{aligned}&|\nabla_x P_u(\varphi)| + |P_{uu}(\varphi)| + |\nabla_x P_{tu}(\varphi)| + |P_{tuu}(\varphi)|\\ &\quad + |\nabla_x P_{tuu}(\varphi)| + |P_{tuuu}(\varphi)| + |\nabla_x\Gamma_u(\varphi)| + |\Gamma_{uu}(\varphi)| \le C_6\\ &\qquad\qquad \text{for } (x,t)\in\Omega\times J \text{ and } \varphi\in\mathbf{R}.\end{aligned} \tag{\textbf{A11}}$$

Rewrite (6.5) as

$$
\begin{aligned}
&(U_t - \hat{u}_t, P(U) - P(\hat{u})) + (a\Pi_h\Phi, \Pi_h\Phi) \\
&\quad = (\Gamma(U) - \Gamma(\hat{u}), \widehat{P(U)} - \widehat{P(\hat{u})}) \\
&\qquad + (\widehat{\Gamma(U)} - \widehat{\Gamma(\hat{u})}, P(\hat{u}) - P(u)) + (\Gamma(\hat{u}) - \Gamma(u), \widehat{P(U)} - \widehat{P(u)}) \\
&\qquad + (a(\mathcal{P}_h\psi - \Pi_h\psi), \Pi_h\Phi) + (U_t - \hat{u}_t, P(u) - P(\hat{u})).
\end{aligned}
\tag{7.6}
$$

By Lemma 2,

$$
\begin{aligned}
&(\widehat{\Gamma(U)} - \widehat{\Gamma(\hat{u})}, P(\hat{u}) - P(u)) + (\Gamma(\hat{u}) - \Gamma(u), \widehat{P(U)} - \widehat{P(u)}) \\
&\quad \le Ch\,\|\hat{u} - u\|\,[\|\widehat{\Gamma(U)} - \widehat{\Gamma(\hat{u})}\| + \|\widehat{P(U)} - \widehat{P(u)}\|] \\
&\quad \le Ch\,\|\hat{u} - u\|\big(\|U - \hat{u}\| + \|\hat{u} - u\|\big).
\end{aligned}
\tag{7.7}
$$

For the last term of (7.6), integration by parts and Lemma 2 gives

$$
\begin{aligned}
&\int_0^t (U_t - \hat{u}_t, P(u) - P(\hat{u}))\,d\tau \\
&\quad = (U - \hat{u}, P(u) - P(\hat{u})) - \int_0^t \left(U - \hat{u}, \frac{\partial}{\partial t}[P(u) - P(\hat{u})]\right) d\tau, \\
&\quad \le Ch\left\{\|U - \hat{u}\|\,\|u - \hat{u}\| + \int_0^t \|U - \hat{u}\|(\|u - \hat{u}\| + \|u_t - \hat{u}_t\|)\,d\tau\right\},
\end{aligned}
\tag{7.8}
$$

since $U^0 = \hat{u}_0$, and assuming smoothness of P.

Returning to (7.6), using the above inequalities, and following the argument in (6.6)–(6.7), we are lead to the following analogue of (6.8):

$$
\begin{aligned}
\|U - \hat{u}\|^2 + \int_0^t \|\Pi_h\Phi\|^2\,d\tau \le C\bigg\{ &\int_0^t \|U - \hat{u}\|^2\,d\tau + \int_0^t \|\mathcal{P}_h\psi - \Pi_h\psi\|^2\,d\tau \\
&+ h^2\left[\|\hat{u} - u\|^2 + \int_0^t \|\hat{u}_t - u_t\|^2\,d\tau + \int_0^t \|\hat{u} - u\|^2\,d\tau\right]\bigg\}.
\end{aligned}
\tag{7.9}
$$

An application of Gronwall's lemma gives our first result.

Theorem 6. *Assume* (**A1**)–(**A5**), (**A7**)–(**A11**). *Let* (u, ψ) *solve problem* (1.1) *with* $\beta = 0$, *and let* (U, Ψ) *solve its semidiscrete mixed finite element approximation* (3.3). *There is some constant* $C > 0$ *such that for any* $t \in J$,

$$
\begin{aligned}
&\|U(t) - \hat{\mathcal{P}}_h u(t)\|^2 + \int_0^t \|\Psi - \Pi_h\psi\|^2\,d\tau \\
&\quad \le C\bigg\{\int_0^t \|(\mathcal{P}_h - \Pi_h)\psi\|^2\,d\tau \\
&\qquad + h^2\left[\max_{0\le\tau\le t} \|\hat{\mathcal{P}}_h u(\tau) - u(\tau)\|^2 + \int_0^t \|\hat{\mathcal{P}}_h u_t - u_t\|^2\,d\tau + \int_0^t \|\hat{\mathcal{P}}_h u - u\|^2\,d\tau\right]\bigg\}.
\end{aligned}
$$

For the fully discrete scheme (3.3), using the techniques employed in the last two sections, we can give a similar result.

Theorem 7. *Assume* (**A1**)–(**A11**). *Let* (u,ψ) *solve problem* (1.1) *with* $\beta = 0$, *and let* (U^n, Ψ^n) *solve its fully discrete mixed finite element approximation* (3.4). *There is some constant* C *such that if the* Δt^j *are sufficiently small, then for any* n *between 1 and* N,

$$\|U^n - \hat{\mathcal{P}}_h u^n\|^2 + \sum_{j=1}^{n} \|\Psi^j - \mathcal{P}_h \psi^j\|^2 \, \Delta t^j$$
$$\leq C\Big\{(\Delta t)^2 + \sum_{j=1}^{n} \|\mathcal{P}_h\psi^j - \Pi_h\psi^j\|^2 \, \Delta t^j$$
$$+ h^2\Big[\max_{0\leq j\leq n} \|\hat{\mathcal{P}}_h u^j - u^j\|^2 + \int_0^{t_n} \|\hat{\mathcal{P}}_h u_t - u_t\|^2 \, d\tau + \int_0^{t_n} \|\hat{\mathcal{P}}_h u - u\|^2 \, d\tau\Big]\Big\}.$$

Remark. These last two results gives optimal order approximation in time and superconvergent (one power of h better than optimal) approximation in space if the solution is smooth enough, and if $\mathcal{P}_h\psi$ is superclose to $\Pi_h\psi$. The latter is true on rectangular grids when α is diagonal [22], [16], [17].

REFERENCES

[1] H. W. Alt and E. DiBenedetto, *Nonsteady flow of water and oil through inhomogeneous porous media*, Ann. Scuola Norm. Sup. Pisa Cl. Sci., 12 (1985), pp. 335–392.

[2] H. W. Alt and S. Luckhaus, *Quasilinear elliptic-parabolic differential equations*, Math. Z., 183 (1983), pp. 311–341.

[3] T. Arbogast, *An error analysis for Galerkin approximations to an equation of mixed elliptic-parabolic type*, Technical Report TR 90-33, Department of Computational and Applied Mathematics, Rice University (1990).

[4] T. Arbogast, *The existence of weak solutions to single porosity and simple dual-porosity models of two-phase incompressible flow*, Nonlinear Analysis, Theory, Methods and Applications, 19 (1992), pp. 1009–1031.

[5] T. Arbogast, M. Obeyesekere and M. F. Wheeler, *Numerical methods for the simulation of flow in root-soil systems*, SIAM J. Numer. Anal., 30 (1993), pp. 1677–1702.

[6] T. Arbogast, M. F. Wheeler, and I. Yotov, *Mixed finite elements for elliptic problems with tensor coefficients as finite differences*, to appear.

[7] J. W. Barrett and P. Knabner, *Finite element approximation of transport of reactive solutes in porous media, Part 2, Error estimates for equilibrium adsorption processes*, to appear.

[8] J. Bear, *Dynamics of Fluids in Porous Media*, Dover Publications, 1972.

[9] F. Brezzi, J. Douglas, Jr., R. Durán, and M. Fortin, *Mixed finite elements for second order elliptic problems in three variables*, Numer. Math., 51 (1987), pp. 237–250.

[10] F. Brezzi, J. Douglas, Jr., M. Fortin, and L. D. Marini, *Efficient rectangular mixed finite elements in two and three space variables*, R.A.I.R.O. Modélisation Mathématique et Analyse Numérique, 21 (1987), pp. 581–604.

[11] F. Brezzi, J. Douglas, Jr., and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numerische Mathematik, 47 (1985), pp. 217–235.

[12] Zhangxin Chen and J. Douglas, Jr., *Prismatic mixed finite elements for second order elliptic problems*, Calcolo, 26 (1989), pp. 135–148.

[13] R. E. Collins, *Flow of fluids through porous materials*, Reinhold, New York, 1961.

[14] J. Douglas, Jr. and J. Roberts, *Global estimates for mixed methods for second order elliptic problems*, Math. Comp., 45 (1985), pp. 39–52.

[15] C. J. van Duijn and P. Knabner, *Solute Transport Through Porous Media with Slow Adsorption*, Report No. 54, Institut fur Mathematik, Universitat Augsburg (1988).

[16] R. Durán, *Superconvergence for rectangular mixed finite elements*, Numerische Mathematik, 58 (1990), pp. 287–298.

[17] R. E. Ewing, R. D. Lazarov, and J. Wang, *Superconvergence of the velocity along the Gauss lines in mixed finite element methods*, SIAM J. Numer. Anal., 28 (1991), pp. 1015–1029.

[18] C. W. Fetter, *Contaminant Hydrogeology*, Macmillan, New York, 1993.

[19] J. L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications* 1, Springer-Verlag, New York, 1971.

[20] E. Magenes, R. H. Nochetto and C. Verdi, *Energy error estimates for a linear scheme to approximate nonlinear parabolic problems*, R.A.I.R.O., Mathematical Modelling and Numerical Analysis, 21 (1987), pp. 655–678.

[21] F. A. Milner, *Mixed finite element methods for quasilinear second-order elliptic problems*, Math. Comp., 44 (1985), pp. 303–320.

[22] M. Nakata, A. Weiser, and M.F. Wheeler, *Some superconvergence results for mixed finite element methods for elliptic problems on rectangular domains*, in The Mathematics of Finite Elements and Applications V, J.R. Whiteman, ed., Academic Press, London, 1985.

[23] J. C. Nedelec, *Mixed finite elements in* $\mathbf{R}^3$, Numer. Math., 35 (1980), pp. 315–341.

[24] J. C. Nedelec, *A new family of mixed finite elements in* $\mathbf{R}^3$, Numer. Math., 50 (1986), pp. 57–81.

[25] R. H. Nochetto and C. Verdi, *Approximation of degenerate parabolic problems using numerical integration*, SIAM J. Numer. Anal., 25 (1988), pp. 784–814.

[26] D. W. Peaceman, *Fundamentals of numerical reservoir simulation*, Elsevier, Amsterdam, 1977.

[27] P. A. Raviart and J. M. Thomas, *A mixed finite element method for 2nd order elliptic problems*, in Mathematical Aspects of the Finite Element Method, Lecture Notes in Mathematics 606, Springer-Verlag, New York, 1977, pp. 292–315.

[28] M. E. Rose, *Numerical Methods for Flows Through Porous Media. I*, Mathematics of Computation, 40 (1983), pp. 435–467.

[29] M. E. Rose, *Numerical Methods for Flows Through Porous Media. II*, Computers and Mathematics with Applications, 20 (1980), pp. 99–122.

[30] T. F. Russell and M. F. Wheeler, *Finite element and finite difference methods for continuous flows in porous media*, in The Mathematics of Reservoir Simulation, R. E. Ewing, ed.,, Society for Industrial and Applied Mathematics, Philadelphia, 1983, pp. 35–106.

[31] C. Srinivas, B. Ramaswamy, M. F. Wheeler, *Mixed finite element methods for flow through unsaturated porous media*, in Computational Methods in Water Resources IX, Vol. 1: Numerical Methods in Water Resources, T. F. Russell, R. E. Ewing, C. A. Brebbia, W. G. Gray, and G. F. Pindar, eds., Computational Mechanics Publications, Southampton, U.K., 1992, pp. 239–246.

[32] A. Weiser and M. F. Wheeler, *On convergence of block-centered finite-differences for elliptic problems*, SIAM J. Numer. Anal., 25 (1988), pp. 351–375.

[33] M. F. Wheeler, *A priori* L_2 *error estimates for Galerkin approximations to parabolic partial differential equations*, SIAM J. Numer. Anal., 10 (1973), pp. 723–759.