EXPANDED MIXED FINITE ELEMENT METHODS FOR QUASILINEAR SECOND ORDER ELLIPTIC PROBLEMS II

Zhangxin Chen

Abstract. A new mixed formulation recently proposed for linear problems is extended to quasilinear second order elliptic problems. This new formulation expands the standard mixed formulation in the sense that three variables are explicitly treated, i.e., the scalar unknown, its gradient, and its flux (the coefficients times the gradient). Based on this formulation, mixed finite element approximations of the quasilinear problems are established. Existence and uniqueness of the solution of the mixed formulation and its discretization are demonstrated. Optimal order error estimates in the L^p and H^{-s}-norms are obtained for the mixed approximations. A postprocessing method for improving the scalar variable is analyzed, and superconvergent estimates are derived. Implementation techniques for solving the systems of algebraic equations are discussed. Comparisons between the standard and expanded mixed formulations are given both theoretically and experimentally. The mixed formulation proposed here is suitable for the case where the coefficient of differential equations is a small tensor and does not need to be inverted.

1. Introduction

This is the second paper of a series in which we develop and analyze expanded mixed formulations for numerical solution of second order elliptic problems. This new formulation expands the standard mixed formulation in the sense that three variables are explicitly treated, i.e., the scalar unknown, its gradient, and its flux (the coefficient times the gradient). It is suitable for the case where the coefficient of differential equations is a small tensor and does not need to be inverted. It directly applies to the flow equation with low permeability and to the transport equation with small dispersion in ground water modeling and petroleum reservoir simulation.

In the first paper of the series [7], we analyzed the expanded mixed formulation for linear second order elliptic problems. Optimal order and superconvergent error estimates for mixed approximations were obtained, and various implementation techniques for solving the system of algebraic equations were discussed.

1991 *Mathematics Subject Classification.* Primary 65N30, 65N22, 65F10.

Key words and phrases. expanded mixed method, quasilinear problem, error estimate, implementation, finite difference, superconvergence, postprocessing.

Partly supported by the Department of Energy under contract DE-ACOS-840R21400.

In this paper, we consider the expanded mixed formulation for a general quasilinear second order elliptic problem. The analysis for the nonlinear problem is completely different from that for the linear problem. First, existence and uniqueness of solution to the nonlinear expanded discretization need to be proven explicitly. This is accomplished through the Brouwer fixed point theorem. Second, the nonlinear error analysis heavily depends upon the established existence result, and is much more difficult. Also, the post-processing scheme proposed here for the first time for the nonlinear mixed method is not a straightforward extension of its linear counterpart. Finally, there has been little theory for solving the system of algebraic equations arising from the nonlinear mixed method. In this paper, we discuss implementation techniques for solving the nonlinear mixed method. Special attention is here paid to implementation of the mixed method as a finite difference method.

It was shown [20, 21, 23] that the linear system arising from the usual mixed formulation can be simplified by use of certain quadrature rules for the lowest-order Raviart-Thomas [19] spaces over a rectangular grid. That is, the mixed method system can be written as a cell-centered finite difference method. The same spaces were also considered for the linear expanded mixed method in [3]. We here consider an analogous simplification of the expanded mixed method system as a finite difference method for another widely used space, the lowest order Brezzi-Douglas-Marini space [6] in the planar case or the lowest-order Brezzi-Douglas-Durán-Fortin space [4] in the three-dimensional case.

This paper also gives a comparison between the standard mixed formulation and the expanded one. For certain nonlinear problems, we show that the expanded formulation is superior to the standard one in that the former leads to the derivation of optimal order error estimates, while the latter gives only sub-optimal error estimates for the mixed method solution. This result is also justified through numerical results. In the previous papers [8, 9, 16], only the Raviart-Thomas spaces have been considered for nonlinear problems. Here we are able to consider all the existing mixed finite element spaces [4, 5, 6, 10, 13, 17, 18, 19].

In the next section we develop the expanded mixed formulation for a fairly general nonlinear second order elliptic problem. It is proven that this formulation has a unique solution and is equivalent to the original differential problem. Then, in §3 we show that all the existing mixed finite elements apply to this formulation. In particular, it is demonstrated that the approximation formulation has a unique solution and gives optimal error estimates in the L^p and H^{-s}-norms. In §4, we propose and analyze a postprocessing method for improving the scalar unknown and derive superconvergent estimates. In §5, we extend the analysis to a nonlinear problem and discuss the difference between the usual mixed method and the standard one. Finally, in §6 we discuss implementation techniques for solving the system of algebraic equations arising from the expanded mixed method and present numerical examples to illustrate our theoretical results.

2. Expanded mixed formulation

Let Ω be a bounded domain in $\mathbb{R}^n$, $n = 2$ or 3, with the boundary $\partial\Omega$. We consider the quasilinear problem

$$Lu = -\nabla \cdot (a(u)\nabla u - b(u)) + c(u) = f \quad \text{in } \Omega, \tag{2.1a}$$

$$u = -g \quad \text{on } \partial\Omega, \tag{2.1b}$$

where we assume that the coefficients $a : \bar{\Omega} \times \mathbb{R} \to \mathbb{R}$, $b : \bar{\Omega} \times \mathbb{R} \to \mathbb{R}^n$, and $c : \bar{\Omega} \times \mathbb{R} \to \mathbb{R}$ are twice continuously differentiable with bounded derivatives through second order; moreover, we assume that

$$(a(u)\mu, \mu) \geq a_0 \|\mu\|^2, \qquad u \in \mathbb{R}, \quad \mu \in \left(L^2(\Omega)\right)^n, \ a_0 > 0. \tag{2.1c}$$

($H^k(\Omega) = W^{k,2}(\Omega)$ is the Sobolev space of k differentiable functions in $L^2(\Omega)$ with the norm $\|\cdot\|_k$; we omit k when it is zero). We also assume that for some ε $(0 < \varepsilon < 1)$ and each pair of functions $(f, g) \in H^\varepsilon(\Omega) \times H^{3/2+\varepsilon}(\partial\Omega)$ there exists a unique solution $u \in H^{2+\varepsilon}(\Omega)$ to (2.1).

Let

$$\begin{aligned} V &= H(\mathrm{div}; \Omega) = \left\{ v \in \left(L^2(\Omega)\right)^n : \nabla \cdot v \in L^2(\Omega) \right\}, \\ W &= L^2(\Omega), \\ \Lambda &= \left(L^2(\Omega)\right)^n, \end{aligned}$$

and let $(\cdot,\cdot)_S$ denote the $L^2(S)$ inner product (we omit S if $S = \Omega$). Then (2.1) is formulated in the following expanded mixed form for $(\sigma, \lambda, u) \in V \times \Lambda \times W$:

$$(a(u)\lambda, \mu) - (\sigma, \mu) + (b(u), \mu) = 0, \quad \forall\, \mu \in \Lambda, \tag{2.2a}$$

$$(\lambda, v) - (u, \nabla \cdot v) = (g, v \cdot \nu)_{\partial\Omega}, \quad \forall\, v \in V, \tag{2.2b}$$

$$(\nabla \cdot \sigma, w) + (c(u), w) = (f, w), \quad \forall\, w \in W, \tag{2.2c}$$

where ν is the outer unit normal to the domain Ω.

To analyze (2.2), let $U = W \times \Lambda$ with the usual product norm $\|\tau\|_U^2 = \|w\|^2 + \|\mu\|^2$, $\tau = (w, \mu) \in U$, and introduce the bilinear forms $\mathcal{A}(\cdot,\cdot) : U \times U \to \mathbb{R}$ and $\mathcal{B}(\cdot,\cdot) : U \times V \to \mathbb{R}$ by

$$\begin{aligned} \mathcal{A}(\chi, \tau) &= (a(u)\lambda, \mu), & \chi &= (u, \lambda), \quad \tau = (w, \mu) \in U, \\ \mathcal{B}(\tau, v) &= (w, \nabla \cdot v) - (\mu, v), & \tau &= (w, \mu) \in U, \ v \in V. \end{aligned}$$

Then (2.2) can be written in the form for $(\chi, \sigma) \in U \times V$ such that

$$\mathcal{A}(\chi, \tau) + \mathcal{B}(\tau, \sigma) + \mathcal{C}(\chi, \tau) = \mathcal{F}(\tau), \quad \forall\, \tau \in U, \tag{2.3a}$$

$$\mathcal{B}(\chi, v) = -(g, v \cdot \nu)_{\partial\Omega}, \quad \forall\, v \in V, \tag{2.3b}$$

where

$$\begin{aligned} &\mathcal{C}(\chi,\tau) = (b(u),\mu) + (c(u),w), && \tau = (w,\mu) \in U, \\ &\mathcal{F}(\tau) = (f,w), && \tau = (w,\mu) \in U. \end{aligned}$$

Finally, we define

$$Z = \{\tau \in U : \mathcal{B}(\tau, v) = 0,\ \forall\, v \in V\}.$$

The next result can be found in the first paper [7].

Lemma 2.1. *Let $\tau = (w,\mu) \in U$. Then $\tau \in Z$ if and only if $w \in H_0^1(\Omega)$ and $\mu = -\nabla w$.*

Theorem 2.2. *If $(\chi,\sigma) \in U \times V$ is the solution of (2.3) with $\chi = (u,\lambda)$, then $u \in H^1(\Omega)$ is the solution of (2.1) with $\lambda = -\nabla u$ and $u|_{\partial\Omega} = g$. Conversely, if $u \in H^1(\Omega)$ is the solution of (2.1) with $u|_{\partial\Omega} = g$, then (2.3) has the solution $(\chi,\sigma) \in U \times V$ with $\chi = (u,\lambda)$, $\lambda = -\nabla u$, and $\sigma = -(a(u)\nabla u - b(u))$.*

Proof. First, let $(\chi,\sigma) \in U \times V$ be the solution of (2.3) with $\chi = (u,\lambda)$. Without loss of generality, let $g = 0$ (otherwise, let $u_0 \in H^1(\Omega)$ such that $u_0|_{\partial\Omega} = g$ and consider $u - u_0$ [14]). Then (2.3b) with $g = 0$ implies that $\chi \in Z$ so that, by Lemma 2.1, $u \in H_0^1(\Omega)$ and $\lambda = -\nabla u$. Hence, for all $w \in H_0^1(\Omega)$ and $\mu = -\nabla w$, it follows from Lemma 2.1 that

$$\mathcal{A}(\chi,\tau) + \mathcal{C}(\chi,\tau) = \mathcal{F}(\tau), \quad \forall\, \tau = (w,\mu) \in Z,$$

i.e.,

$$(a(u)\nabla u, \nabla w) + (b(u), \nabla w) + (c(u), w) = (f,w), \quad \forall\, w \in H_0^1(\Omega).$$

Hence u is a weak solution of (2.1), i.e., the solution of (2.1) [11].

Next, we assume that $u \in H_0^1(\Omega)$ is the solution of (2.1). Set $\chi = (u,\lambda)$ with $\lambda = -\nabla u$ and $\sigma = -(a(u)\nabla u - b(u))$. Then it follows from Lemma 2.1 that $\chi \in Z$, so (2.3b) with $g = 0$ holds. Thus it remains to prove (2.3a). For each $\tau \in U$ with $\tau = (w,\mu)$,

$$\begin{aligned} \mathcal{A}(\chi,\tau) + \mathcal{B}(\tau,\sigma) + \mathcal{C}(\chi,\tau) &= (a(u)\lambda,\mu) + (w, \nabla\cdot\sigma) - (\mu,\sigma) + (b(u),\mu) + (c(u),w) \\ &= (w, -\nabla\cdot(a(u)\nabla u - b(u)) + c(u)) \\ &= (f,w), \quad \forall\, w \in W, \end{aligned}$$

which implies (2.3a). $\square$

3. Mixed finite elements

To define a finite element method, we need a partition $\mathcal{E}_h$ of Ω into elements E, say, simplexes, rectangular parallelepipeds, and/or prisms, where only edges or faces on $\partial\Omega$ may be curved. In $\mathcal{E}_h$, it is also necessary that adjacent elements completely share their common edge or face; let $\partial\mathcal{E}_h$ denote the set of all interior edges ($n = 2$) or faces ($n = 3$) e of $\mathcal{E}_h$.

Since mixed finite element spaces are finite dimensional and defined locally on each element, let, for each $E \in \mathcal{E}_h$, $V_h(E) \times W_h(E)$ denote one of the mixed finite element spaces introduced in [4, 5, 6, 10, 13, 17, 18, 19] for second order elliptic problems. Then we define

$$\begin{aligned}\Lambda_h &= \{\mu \in \Lambda : \mu|_E \in V_h(E) \text{ for each } E \in \mathcal{E}_h\},\\ V_h &= \{v \in V : v|_E \in V_h(E) \text{ for each } E \in \mathcal{E}_h\},\\ W_h &= \{w \in W : w|_E \in W_h(E) \text{ for each } E \in \mathcal{E}_h\}.\end{aligned}$$

The expanded mixed finite element method for (2.1) is to find $(\sigma_h, \lambda_h, u_h) \in V_h \times \Lambda_h \times W_h$ such that

$$(a(u_h)\lambda_h, \mu) - (\sigma_h, \mu) + (b(u_h), \mu) = 0, \quad \forall\, \mu \in \Lambda_h, \tag{3.1a}$$

$$(\lambda_h, v) - (u_h, \nabla \cdot v) = (g, v \cdot \nu)_{\partial\Omega}, \quad \forall\, v \in V_h, \tag{3.1b}$$

$$(\nabla \cdot \sigma_h, w) + (c(u_h), w) = (f, w), \quad \forall\, w \in W_h. \tag{3.1c}$$

We shall establish existence, uniqueness, and convergence results for (3.1) in this section. For simplicity, we concentrate on the planar case; an extension to the space case is straightforward. We mention that while an extra unknown is introduced in (3.1), the computational cost for solving (3.1) is the same as that for solving the usual mixed method, as shown in §6.

3.1. Existence. C and C_1 are generic constants below, where C_1 depends on $\|u\|_{2+\varepsilon}$, at most quadratically. Each of our mixed finite element spaces [4, 5, 6, 10, 13, 17, 18, 19] has the property that there are projection operators $\Pi_h : (H^1(\Omega))^n \to V_h$ and $P_h = L^2$-projection : $L^2(\Omega) \to W_h$ such that

$$\|v - \Pi_h v\| \le C\|v\|_r h^r, \quad 1 \le r \le k+1, \tag{3.2a}$$

$$\|\nabla \cdot (v - \Pi_h v)\| \le C\|\nabla \cdot v\|_r h^r, \quad 0 \le r \le k^*, \tag{3.2b}$$

$$\|w - P_h w\|_{-s} \le C\|w\|_r h^{r+s}, \quad 0 \le s,\ r \le k^*, \tag{3.2c}$$

and

$$(\nabla \cdot (v - \Pi_h v), w) = 0, \quad \forall\, w \in W_h, \tag{3.3a}$$

$$(\nabla \cdot v, w - P_h w) = 0, \quad \forall\, v \in V_h, \tag{3.3b}$$

where $k^* = k+1$ for the Raviart-Thomas-Nedelec spaces [19, 17, 18] and the Brezzi-Douglas-Fortin-Marini spaces, $k^* = k$ for the Brezzi-Douglas-Marini spaces and Brezzi-Douglas-Durán-Fortin [6, 4], and the Chen-Douglas spaces include both cases. Also, let R_h be the L^2 projection onto Λ_h. Then we see that

$$(\mu - R_h \mu, \tau) = 0, \quad \forall\, \mu \in \Lambda, \quad \tau \in \Lambda_h, \tag{3.4}$$

and

$$\|\mu - R_h\mu\|_{-s} \le C\|\mu\|_r h^{r+s}, \qquad 0 \le s,\ r \le k+1. \tag{3.5}$$

For the analysis below, we write

$$a(u_h) - a(u) = -\tilde{a}_u(u_h)(u-u_h) = -a_u(u)(u-u_h) + \tilde{a}_{uu}(u_h)(u-u_h)^2, \tag{3.6}$$

where

$$\tilde{a}_u(u_h) = \int_0^1 a_u(u_h + t(u-u_h))dt,$$
$$\tilde{a}_{uu}(u_h) = \int_0^1 (1-t)a_{uu}(u + t(u_h-u))dt,$$

are bounded in $\bar{\Omega}$. Similarly, we write

$$b(u_h) - b(u) = -\tilde{b}_u(u_h)(u-u_h) = -b_u(u)(u-u_h) + \tilde{b}_{uu}(u_h)(u-u_h)^2, \tag{3.7}$$

$$c(u_h) - c(u) = -\tilde{c}_u(u_h)(u-u_h) = -c_u(u)(u-u_h) + \tilde{c}_{uu}(u_h)(u-u_h)^2, \tag{3.8}$$

where $\tilde{b}_u(u_h)$, $\tilde{b}_{uu}(u_h)$, $\tilde{c}_u(u_h)$, and $\tilde{c}_{uu}(u_h)$ are bounded functions in $\bar{\Omega}$. We now subtract (3.1) from (2.2) to obtain the error equations

$$\text{(3.9a)} \quad (a(u)(\lambda-\lambda_h),\mu) - (\sigma-\sigma_h,\mu) + (b(u)-b(u_h),\mu) = ((a(u_h)-a(u))\lambda_h,\mu)\,, \quad \forall\ \mu \in \Lambda_h,$$

$$\text{(3.9b)} \quad (\lambda-\lambda_h,v) - (u-u_h,\nabla\cdot v) = 0, \quad \forall\ v \in V_h,$$

$$\text{(3.9c)} \quad (\nabla\cdot(\sigma-\sigma_h),w) + (c(u)-c(u_h),w) = 0, \quad \forall\ w \in W_h.$$

Substituting (3.6)–(3.8) into (3.9), we see that

$$\begin{aligned} \text{(3.10a)} \quad & (a(u)(\lambda-\lambda_h),\mu) - (\sigma-\sigma_h,\mu) + (\Gamma(u)(u-u_h),\mu) \\ &= \left(\left(\tilde{a}_{uu}(u_h)\lambda + b_{uu}^2(u_h)\right)(u-u_h)^2,\mu\right) + (\tilde{a}_u(u_h)(u-u_h)(\lambda-\lambda_h),\mu)\,, \quad \forall \mu \in \Lambda_h, \end{aligned}$$

$$\text{(3.10b)} \quad (\lambda-\lambda_h,v) - (u-u_h,\nabla\cdot v) = 0, \quad \forall v \in V_h,$$

$$\text{(3.10c)} \quad (\nabla\cdot(\sigma-\sigma_h),w) + (\gamma(u-u_h),w) = \left(\tilde{c}_{uu}(u_h)(u-u_h)^2,w\right), \quad \forall\ w \in W_h,$$

where $\Gamma(u) = a_u(u)\lambda + b_u(u)$ and $\gamma(u) = c_u(u)$. Now let $M : H^2(\Omega) \to L^2(\Omega)$ be the linear operator

$$Mw = -\nabla\cdot(a(u)\nabla w - \Gamma(u)w) + \gamma w,$$

and let

$$\Phi : V_h \times \Lambda_h \times W_h \to V_h \times \Lambda_h \times W_h$$

be given by $\Phi((\tau,\eta,\rho)) = (x,y,z)$, where (x,y,z) is the solution of the system

$$\begin{aligned}
&(3.11a) && (a(u)(\lambda - y), \mu) - (\sigma - x, \mu) + (\Gamma(u - z), \mu) \\
& && \quad = \left(\left(\tilde{a}_{uu}(\rho)\lambda + \tilde{b}_{uu}(\rho) \right) (u-\rho)^2, \mu \right) + (\tilde{a}_u(\rho)(u-\rho)(\lambda - \eta), \mu), \\
& && \qquad\qquad \forall\ \mu \in \Lambda_h, \\
&(3.11b) && (\lambda - y, v) - (u - z, \nabla\cdot v) = 0, \qquad \forall\ v \in V_h, \\
&(3.11c) && (\nabla\cdot(\sigma - x), w) + (\gamma(u-z), w) = \left(\tilde{c}_{uu}(\rho)(u-\rho)^2, w\right), \quad \forall\ w \in W_h.
\end{aligned}$$

We assume that the restrictions of M and M^* (its adjoint) to $H^2(\Omega)\cap H_0^1(\Omega)$ have bounded inverses. This is satisfied if $c_u \geq 0$ [14]. Then existence and uniqueness of solution to (3.11) is known [7] since (3.11) corresponds to the expanded mixed method for the linear operator M. Now we see that existence of a solution to (3.1) is equivalent to the problem that the map Φ has a fixed point. Consequently, the solvability of (3.1) follows from the Brouwer fixed point theorem if we can prove that Φ maps a ball of $V_h \times \Lambda_h \times W_h$ into itself. Toward that end, we need the following definition [12].

We say that Ω is $(s+2,\theta)$-regular with respect to M if the Dirichlet problem

$$\begin{aligned}
&(3.12a) && M^*\varphi = \psi \qquad \text{in } \Omega, \\
&(3.12b) && \varphi = 0 \qquad \text{on } \partial\Omega
\end{aligned}$$

is uniquely solvable for $\psi \in L^2(\Omega)$ and if

$$\|\varphi\|_{s+2,\theta} \leq C\|\psi\|_{s,\theta}. \tag{3.13}$$

Lemma 3.1. *Assume that $2 \leq \theta < \infty$ and Ω is $(s+2,\theta')$-regular with respect to M, where $\theta' = \theta/(\theta-1)$ is the conjugate exponent of θ. Let $\xi \in L^2(\Omega)$, $\phi \in V$, $\zeta \in L^2(\Omega)$, and $r \in L^2(\Omega)$. If $\pi \in W_h$ satisfies the system*

$$\begin{aligned}
&(3.14a) && (a(u)\xi, \mu) - (\phi, \mu) + (\Gamma\pi, \mu) = (\zeta, \mu), \quad \forall \mu \in \Lambda_h, \\
&(3.14b) && (\xi, v) - (\pi, \nabla\cdot v) = 0, \qquad \forall\ v \in V_h, \\
&(3.14c) && (\nabla\cdot\phi, w) + (\gamma\pi, w) = (r, w), \qquad \forall\ w \in W_h,
\end{aligned}$$

then there is a constant $C = C(\theta, a, \Gamma, \gamma, \Omega)$ such that

$$\|\pi\|_{0,\theta} \leq C\left\{ (\|\xi\| + \|\phi\|)\, h^{2/\theta} + h^{\min(1+2/\theta,k^*)}\|\nabla\cdot\phi\| + \|\zeta\| + \|\gamma\| \right\}. \tag{3.15}$$

Moreover, if $\xi \in L^\theta(\Omega)$, $\phi \in W^{0,\theta}(\mathrm{div}\,;\Omega) = \left\{v \in L^\theta(\Omega); \nabla\cdot v \in L^\theta(\Omega)\right\}$, $\zeta \in L^\theta(\Omega)$, and $r \in L^\theta(\Omega)$, then for $0 \leq s \leq 2k^*$

$$\begin{aligned}
\|\pi\|_{-s,\theta} \leq C\big\{ &(\|\xi\|_{0,\theta} + \|\phi\|_{0,\theta})\, h^{\min(s+1,k+1)} + \|\zeta\|_{0,\theta} h^{\min(s+1,k^*)} \\
&+ (\|\nabla\cdot\phi\|_{0,\theta} + \|r\|_{0,\theta})\, h^{\min(s+2,k^*)} \\
&+ \|\zeta\|_{-s-1,\theta} + \|r\|_{-s-2,\theta}\big\}.
\end{aligned} \tag{3.16}$$

Proof. We only prove (3.16); (3.15) can be shown more easily. Let $\psi \in W^{s,\theta'}(\Omega)$ and $\varphi \in W^{s+2,\theta'}(\Omega)$ be the solution of (3.12). Then, by (3.3), (3.14), and integration by parts, we see that

$$
\begin{aligned}
(\pi, \psi) =&(\pi, M^*\varphi) \\
=& (\pi, -\nabla\cdot(a(u)\nabla\varphi) - \Gamma\nabla\varphi + \gamma\varphi) \\
=& -(\xi, \Pi_h(a(u)\nabla\varphi)) - (\Gamma\pi, \nabla\varphi - R_h\nabla\varphi) - (\Gamma\pi, R_h\nabla\varphi) + (\gamma\pi, \varphi) \\
=& (\xi, a(u)\nabla\varphi - \Pi_h(a(u)\nabla\varphi)) - (a(u)\xi, \nabla\varphi - R_h\nabla\varphi) + (\phi, \nabla\varphi - R_h\nabla\varphi) \\
&+ (\nabla\cdot\phi, \varphi - P_h\varphi) + (\zeta, \nabla\varphi - R_h\nabla\varphi) - (\zeta, \nabla\varphi) + (r, R_h\varphi - \varphi) + (r, \varphi) \\
&+ (\Gamma\pi, \nabla\varphi - R_h\nabla\varphi) + (\gamma\pi, \varphi - P_h\varphi).
\end{aligned}
\tag{3.17}
$$

Applying (3.2a), (3.2b), and (3.5), we observe that

$$
\begin{aligned}
&|(\xi, a(u)\nabla\varphi - \Pi_h(a(u)\nabla\varphi))| \le C\|\xi\|_{0,\theta}\ \|\varphi\|_{s+2,\theta'}\ h^{\min(s+1,k+1)}, \\
&|(a(u)\xi, \nabla\varphi - R_h\nabla\varphi)| \le C\|\xi\|_{0,\theta}\ \|\varphi\|_{s+2,\theta'}\ h^{\min(s+1,k+1)}, \\
&|(\phi, \nabla\varphi - R_h\nabla\varphi)| \le C\|\phi\|_{0,\theta}\ \|\varphi\|_{s+2,\theta'}\ h^{\min(s+1,k+1)}, \\
&|(\nabla\cdot\phi, \varphi - P_h\varphi)| \le C\|\nabla\cdot\phi\|_{0,\theta}\ \|\varphi\|_{s+2,\theta'}\ h^{\min(s+2,k^*)}, \\
&|(\zeta, \nabla\varphi - R_h\nabla\varphi)| \le C\|\zeta\|_{0,\theta}\ \|\varphi\|_{s+2,\theta'}\ h^{\min(s+1,k+1)}, \\
&|(\zeta, \nabla\varphi)| \le \|\zeta\|_{-s-1,\theta}\ \|\varphi\|_{s+2,\theta'}, \\
&|(r, P_h\varphi - \varphi)| \le C\|r\|_{0,\theta}\|\ \varphi\|_{s+2,\theta'}\ h^{\min(s+2,k^*)}, \\
&|(r, \varphi)| \le \|r\|_{-s-2,\theta}\ \|\varphi\|_{s+2,\theta'}, \\
&|(\Gamma\pi, \nabla\varphi - R_h\nabla\varphi)| \le C\|\pi\|_{0,\theta}\|\ \varphi\|_{s+2,\theta'}\ h^{\min(s+1,k+1)}, \\
&|(\gamma\pi, \varphi - P_h\varphi)| \le C\|\pi\|_{0,\theta}\ \|\varphi\|_{s+2,\theta'}\ h^{\min(s+2,k^*)}.
\end{aligned}
$$

Substitute these inequalities into (3.17) and use (3.13) to obtain

$$
\begin{aligned}
\|\pi\|_{-s,\theta} \le C\Big(&(\|\xi\|_{0,\theta} + \|\phi\|_{0,\theta} + \|\zeta\|_{0,\theta})\, h^{\min(s+1,k+1)} \\
&+ \|\nabla\cdot\phi\|_{0,\theta}\ h^{\min(s+2,k^*)} + \|\zeta\|_{-s-1,\theta} \\
&+ \|r\|_{0,\theta}\ h^{\min(s+2,k^*)} + \|r\|_{-s-2,\theta} + \|\pi\|_{0,\theta}\ h^{\min(s+1,k^*)}\Big).
\end{aligned}
\tag{3.18}
$$

First, consider $s = 0$; for h sufficiently small, the $h\|\pi\|_{0,\theta}$ term on the right-hand side of (3.18) can be absorbed into the left-hand side, and the result (3.16) has been established for $s = 0$. Then, for $s > 0$, apply (3.18) again, the established result for $s = 0$, and the interpolation result [15]

$$
\|r\|_{-2,\theta} \le C\|r\|_{0,\theta}^{s/(s+2)}\|r\|_{-s-2,\theta}^{2/(s+2)} \le C\left(h\|r\|_{0,\theta} + h^{-s/2}\|r\|_{-s-2,\theta}\right),
$$

to obtain (3.16) since $k^* \le k+1$ and $s \le 2k^*$. $\square$

We now turn to existence of a solution to (3.1). For this we rewrite (3.11) by shifting (u, λ, σ) to $(P_h u, R_h \lambda, \Pi_h \sigma)$ and using (3.3a), (3.3b), and (3.4) as follows:

$$\begin{aligned}
&(3.19a) && (a(u)(R_h\lambda - y), \mu) - (\Pi_h\sigma - x, \mu) + (\Gamma(P_h u - z), \mu) \\
& && \quad = \left(\left(\tilde{a}_{uu}(\rho)\lambda + \tilde{b}_{uu}(\rho)\right)(u-\rho)^2, \mu\right) + (\tilde{a}_u(\rho)(u-\rho)(\lambda-\eta), \mu) \\
& && \qquad + (a(u)(R_h\lambda - \lambda), \mu) + (\sigma - \Pi_h\sigma, \mu) + (\Gamma(P_h u - u), \mu), \\
& && \qquad\qquad\qquad \forall\ \mu \in \Lambda_h, \\
&(3.19b) && (R_h\lambda - y, v) - (P_h u - z, \nabla\cdot v) = 0, \qquad \forall\ v \in V_h, \\
&(3.19c) && (\nabla\cdot(\Pi_h\sigma - x), w) + (\gamma(P_h u - z), w) \\
& && \quad = \left(\tilde{c}_{uu}(\rho)(u-\rho)^2, w\right) + (\gamma(P_h u - u), w), \qquad \forall\ w \in W_h.
\end{aligned}$$

Let $\mathcal{W}_h = W_h$ and $\mathcal{L}_h = \Lambda_h$ with the stronger norms $\|w\|_{\mathcal{W}_h} = \|w\|_{0,\theta}$ and $\|\mu\|_{\mathcal{L}_h} = \|\mu\|_{0,2+\varepsilon}$, respectively, where $\theta = (4+2\varepsilon)/\epsilon > 4$.

Theorem 3.2. *For $\delta > 0$ sufficiently small (dependent of h), Φ maps a ball of radius δ of $V_h \times \mathcal{L}_h \times \mathcal{W}_h$ onto itself.*

Proof. Let

$$(3.20) \qquad \|\Pi_h\sigma - \tau\|_V < \delta, \qquad \|P_h u - \rho\|_{0,\theta} < \delta, \qquad \|R_h\lambda - \eta\|_{0,2+\varepsilon} < \delta.$$

We now apply (3.15) to (3.19) with

$$\begin{aligned}
\zeta &= \left(\tilde{a}_{uu}(\rho)\lambda + \tilde{b}_{uu}(\rho)\right)(u-\rho)^2 + \tilde{a}_u(\rho)(u-\rho)(\lambda-\eta) + a(u)(R_h\lambda - \lambda) \\
&\quad + \sigma - \Pi_h\sigma + \Gamma(P_h u - u), \\
r &= \tilde{c}_{uu}(\rho)(u-\rho)^2 + \gamma(P_h u - u).
\end{aligned}$$

First, note that, by (3.2a), (3.2b), and (3.5),

$$\begin{aligned}
\|\zeta\| + \|r\| &\le C\big\{\|u-\rho\|_{0,4}^2 + \|u-\rho\|_{0,\theta}\ \|\lambda-\eta\|_{0,2+\varepsilon} \\
&\qquad + \|R_h\lambda - \lambda\| + \|\sigma - \Pi_h\sigma\| + \|P_h u - u\|\big\} \\
&\le C\big\{\|u - P_h u\|_{0,\theta}^2 + \|\rho - P_h u\|_{0,\theta}^2 + h\|u\|_2 \\
&\qquad + (\|u - P_h u\|_{0,\theta} + \|P_h u - \rho\|_{0,\theta})\,(\|\lambda - R_h\lambda\|_{0,2+\varepsilon} + \|R_h\lambda - \eta\|_{0,2+\varepsilon})\big\},
\end{aligned}$$

so that, by (3.2), (3.5), (3.20), and the Sobolev embedding inequalities [1]

$$\|u\|_{2,2+\varepsilon} \le C_\varepsilon \|u\|_{2+\varepsilon}, \qquad \|u\|_{1,\theta} \le C_\varepsilon \|u\|_{2+\varepsilon},$$

we see that

$$\|\zeta\| + \|r\| \le C_1(h + \delta^2), \tag{3.21}$$

where $C_1 = C_1(\|u\|_{2+\varepsilon})$. If we take the last term on the left side of equations (3.19a) and (3.19c) over to the right side, the left side in (3.19) beomes exactly the expanded mixed method for the differential operator $-\nabla \cdot (a(u)\nabla)$. It follows from [7] that

$$\|\Pi_h\sigma - x\|_V \le C\left(\|P_h u - z\| + \|\zeta\| + \|r\|\right), \tag{3.22a}$$

$$\|R_h\lambda - y\| \le C\left(\|P_h u - z\| + \|\zeta\| + \|r\|\right). \tag{3.22b}$$

Now, apply (3.15) to (3.19) to obtain

$$\begin{aligned}\|P_h u - z\|_{0,\theta} \le C\big\{&\left(\|R_h\lambda - y\| + \|\Pi_h\sigma - x\|\right) h^{2/\theta} \\ &+ \|\nabla \cdot (\Pi_h\sigma - x)\|\, h^{\min(1+2/\theta, k^*)} + \|\zeta\| + \|r\|\big\}.\end{aligned}$$

Consequently, it follows from (3.21) and (3.22) that

$$\|P_h u - z\|_{0,\theta} \le C_1(h + \delta^2), \tag{3.23}$$

for h sufficiently small. Exploit (3.21)–(3.23) again to see that

$$\|\Pi_h\sigma - x\|_V \le C_1(h + \delta^2), \tag{3.24a}$$

$$\|R_h\lambda - y\| \le C_1(h + \delta^2). \tag{3.24b}$$

Using the quasi-regularity of T_h, we find that

$$\|R_h\lambda - y\|_{0,2+\varepsilon} \le Ch^{-\varepsilon/(2+\varepsilon)}\|R_h\lambda - y\| \le C_1 h^{-\varepsilon/(2+\varepsilon)}(h + \delta^2). \tag{3.25}$$

Finally, let $h < (2C_1)^{-(4+2\varepsilon)/(2-\varepsilon)}$ and choose $\delta = 2C_1 h^{2/(2+\varepsilon)}$. Observe that in order to have $C_1 h^{2/(2+\varepsilon)} \le \delta/2$ and $C_1 h^{-\varepsilon/(2+\varepsilon)}\delta^2 \le \delta/2$, δ must belong to

$$\left[2C_1 h^{2/(2+\varepsilon)}, (2C_1)^{-1} h^{\varepsilon/(2+\varepsilon)}\right] \ne \emptyset,$$

which is satisfied for h and δ as chosen. Now, by (3.23), (3.24a), and (3.25), for such chosen h and δ, we have

$$\|\Pi_h\sigma - x\|_V < \delta, \qquad \|P_h u - z\|_{0,\theta} < \delta, \qquad \|R_h\lambda - y\|_{0,2+\varepsilon} < \delta. \tag{3.26}$$

That is, Φ maps the ball of radius δ, centered at $(\Pi_h\sigma, R_h\lambda, P_h u)$ onto itself. $\square$

3.2. L^2-error estimates. Assume momentarily that (3.1) has a unique solution. That it does will be established later, at least for small h. To obtain error estimates, we rewrite (3.9), by (3.6)–(3.8), as follows:

$$
\begin{aligned}
&(a(u)(\lambda-\lambda_h),\mu)-(\sigma-\sigma_h,\mu)+\Big(\big(\tilde a_u(u_h)\lambda_h+\tilde b_u(u_h)\big)(u-u_h),\mu\Big)=0, && \forall\ \mu\in\Lambda_h,\\
&(\lambda-\lambda_h,v)-(u-u_h,\nabla\cdot v)=0, && \forall\ v\in V_h,\\
&(\nabla\cdot(\sigma-\sigma_h),w)+(\tilde c_u(u_h)(u-u_h),w)=0, && \forall\ w\in W_h.
\end{aligned}
$$

Define

$$
\begin{aligned}
\alpha&=\lambda-\lambda_h, & \beta&=R_h\lambda-\lambda_h,\\
d&=\sigma-\sigma_h, & e&=\Pi_h\sigma-\sigma_h,\\
y&=u-u_h, & z&=P_hu-u_h.
\end{aligned}
$$

We then have with $\tilde\Gamma_h=\tilde a_u(u_h)\lambda_h+\tilde b_u(u_h)$

$$(a(u)\alpha,\mu)-(d,\mu)+(\tilde\Gamma_h z,\mu)=\Big(\tilde\Gamma_h(P_hu-u),\mu\Big),\quad \forall\ \mu\in\Lambda_h, \tag{3.27a}$$

$$(\alpha,v)-(z,\nabla\cdot v)=0,\quad \forall\ v\in V_h, \tag{3.27b}$$

$$(\nabla\cdot d,w)+(\tilde c_u(u_h)z,w)=(\tilde c_u(u_h)(P_hu-u),w),\quad \forall\ w\in W_h. \tag{3.27c}$$

Or, equivalently, as a result of (3.3b) and (3.4),

$$(a(u)\alpha,\mu)-(d,\mu)+(\tilde\Gamma_h\ z,\mu)=\Big(\tilde\Gamma_h(P_hu-u),\mu\Big),\quad \forall\ \mu\in\Lambda_h, \tag{3.28a}$$

$$(\beta,v)-(z,\nabla\cdot v)=0,\quad \forall\ v\in V_h, \tag{3.28b}$$

$$(\nabla\cdot d,w)+(\tilde c_u(u_h)z,w)=(\tilde c_u(u_h)(P_hu-u),w),\quad \forall\ w\in W_h. \tag{3.28c}$$

Observe that (3.27) or (3.28) corresponds to the mixed method for the linear operator $N:H^2(\Omega)\to L^2(\Omega)$ given by $N_w=-\nabla\cdot\Big(a(u)\nabla w-\tilde\Gamma_h w\Big)+\tilde c_u(u_h)w$. As shown in [16], it follows from the results (3.26) in the proof of Theorem 3.2 that there is an h_0 such that the restriction of its adjoint N^* to $H^2(\Omega)\cap H_0^1(\Omega)$ has a bounded inverse for $h<h_0$. Now we prove the next result.

Theorem 3.3. *Assume that Ω is (2,2)-regular with respect to M. Then for h sufficiently small*

$$\|u-u_h\|\le C_1\left(\|u\|_r\ h^r+\|u\|_{r_1+\delta_{1k^*}}\ h^{r_1}\right),\qquad 2\le r\le k+2,\ 1\le r_1\le k^*, \tag{3.29a}$$

$$\|\lambda-\lambda_h\|+\|\sigma-\sigma_h\|\le C_1\Big(\|u\|_{r+1}\ h^r+\|u\|_{r_1}\ h^{r_1}+\|\nabla\cdot\sigma\|_{r_1}\ h^{r_1+\min(2,k^*)}\Big),\qquad 1\le r\le k+1,\ 0\le r_1\le k^*, \tag{3.29b}$$

$$\|\nabla\cdot(\sigma-\sigma_h)\|\le C_1\left(\|u\|_{r+1}\ h^r+\|u\|_{r_1}\ h^{r_1}+\|\nabla\cdot\sigma\|_{r_1}\ h^{r_1}\right),\qquad 1\le r\le k+1,\ 0\le r_1\le k^*. \tag{3.29c}$$

Proof. Using (3.26) with $\delta = 2C_1 h^{2/(2+\varepsilon)}$, the embedding relation $H^{1+\varepsilon}(\Omega) \subset W^{\varepsilon/2,\infty}(\Omega)$, and the quasi-regularity of T_h, we see that

$$\begin{aligned}
\|\lambda_h\|_{0,\infty} \leq & \|\beta\|_{0,\infty} + \|P_h\lambda\|_{0,\infty} \\
\leq & Ch^{-2/(2+\varepsilon)}\|\beta\|_{0,2+\epsilon} + \|\lambda - P_h\lambda\|_{0,\infty} + \|\lambda\|_{0,\infty} \\
\leq & C_1\left(\|u\|_{2+\varepsilon}\right),
\end{aligned} \tag{3.30}$$

so that $\|\tilde{\Gamma}_h\|_{0,\infty}$ is bounded by C_1. Now, apply (3.16) to (3.27) to get

$$\begin{aligned}
\|z\| \leq C\Big\{ & \left(\|\alpha\| + \|d\| + \left\|\tilde{\Gamma}_h(P_hu - u)\right\|\right) h \\
& + \left(\|\nabla\cdot d\| + \|\tilde{c}_u(u_h)(P_hu-u)\|\right) h^{\min(2,k^*)} \\
& + \left\|\tilde{\Gamma}_h(P_hu - u)\right\|_{-1} + \|\tilde{c}_u(u_h)(P_hu-u)\|_{-2}\Big\}.
\end{aligned} \tag{3.31}$$

Furthermore, by (3.2c) and (3.30), we see that

$$\left\|\tilde{\Gamma}_h(P_hu - u)\right\| h + \left\|\tilde{\Gamma}_h(P_hu - u)\right\|_{-1} \leq C_1\|u\|_{r_1}\ h^{r_1+1}, \qquad 0 \leq r_1 \leq k^*, \tag{3.32}$$

$$\|\tilde{c}_u(u_h)(P_hu-u)\|\, h + \|\tilde{c}_u(u_h)(P_hu-u)\|_{-2} \leq C_1\|u\|_{r_1}\ h^{r_1+1}, \quad 0 \leq r_1 \leq k^*. \tag{3.33}$$

It remains to estimate α, d, and $\|\nabla\cdot d\|$. As in the proof of Theorem 3.2, it follows from (3.28) [7] that

$$\|\beta\| + \|e\|_V \leq C_1\left(\|\Pi_h\sigma - \sigma\| + \|\lambda - R_h\lambda\| + \|y\|\right),$$

so that, by (3.2a) and (3.5),

$$\|\beta\| + \|e\|_V \leq C_1\left(\|u\|_{r+1}\ h^r + \|y\|\right), \quad 1 \leq r \leq k+1. \tag{3.34}$$

Now, apply (3.2a), (3.2b), (3.5), and (3.34) to obtain

$$\|\alpha\| \leq C_1\left(\|u\|_{r+1}\ h^r + \|y\|\right), \qquad 1 \leq r \leq k+1, \tag{3.35a}$$

$$\|d\| \leq C_1\left(\|u\|_{r+1}\ h^r + \|y\|\right), \qquad 1 \leq r \leq k+1, \tag{3.35b}$$

$$\|\nabla\cdot d\| \leq C_1\left(\|\nabla\cdot\sigma\|_{r_1}\ h^{r_1} + \|u\|_r\ h^r + \|y\|\right), \quad 0 \leq r_1 \leq k^*,\ 1 \leq r \leq k+1. \tag{3.35c}$$

Substitute (3.35a)–(3.35c) into (3.31) and use (3.2c), (3.32), and (3.33) to obtain

$$\begin{aligned}
\|z\| \leq C_1 & \left(\|u\|_{r+1}\ h^{r+1} + \|u\|_{r_1}\ h^{r_1+1} + \|\nabla\cdot\sigma\|_{r_1}\ h^{r_1+\min(2,k^*)}\right), \\
& 1 \leq r \leq k+1, \quad 0 \leq r_1 \leq k^*,
\end{aligned} \tag{3.36}$$

for h sufficiently small. Now, combine (3.2c), (3.35), and (3.36) to yield the desired result (3.29). □

We remark that the L^2-error estimates in Theorem 3.3 are optimal both in rate (for any h) and in regularity. Also, as a result of (3.36), we have

$$\|P_h u - u_h\| \le C_1 \|u\|_r h^{k^*+1}, \quad r = \max(k^*+1, 3), \tag{3.37}$$

which is a superconvergence result and is needed in the analysis of the later postprocessing method. Note that in the case where $k^* = k+1$ we have the superconvergence order $O(h^{k+2})$, and in the case where $k^* = k$ we have $O(h^{k+1})$. For the linear case where a does not depend on the solution u and $b = c \equiv 0$ in (2.1), we have shown a superconvergence result, which is of order $O(h^{k+2})$ for both cases [7]. The reason that we only have a superconvergence order $O(h^{k+1})$ for the latter case for the present nonlinear problem is that the coefficient a depends on u and b and c are not zero. The same remark applies to the postprocessing method proposed in §4 later.

3.3. Uniqueness. We now demonstrate the uniqueness of the solution to (3.1). Let $(u^i, \lambda^i, \sigma^i) \in W_h \times \Lambda_h \times V_h$ be solutions of (3.1), $i = 1, 2$. Note that it follows from Theorem 3.3 that these two solutions satisfy the error bounds in (3.29) provided that they satisfy (3.26). Then the quasi-regularity of T_h and the error bounds imply that λ^i is bounded by $\|u\|_{2+\varepsilon}$, $i = 1, 2$. Let $\bar{u} = u^1 - u^2$, $\bar{\lambda} = \lambda^1 - \lambda^2$, and $\bar{\sigma} = \sigma^1 - \sigma^2$. Then, by (3.1), we see that

$$\left(a(u^1)\bar{\lambda}, \mu\right) - (\bar{\sigma}, \mu) + \left(\left(\tilde{a}_u(u^2)\lambda^2 + \tilde{b}_u(u_2)\right)\bar{u}, \mu\right) = 0, \quad \forall\, \mu \in \Lambda_h, \tag{3.38a}$$

$$(\bar{\lambda}, v) - (\bar{u}, \nabla \cdot v) = 0, \quad \forall\, v \in V_h, \tag{3.38b}$$

$$(\nabla \cdot \bar{\sigma}, w) + \left(\tilde{c}_u(u^2)\bar{u}, w\right) = 0, \quad \forall w \in W_h. \tag{3.38c}$$

Then, as in the linear case [7], we have

$$\|\bar{\lambda}\| + \|\bar{\sigma}\|_V \le C_1 \|\bar{u}\|. \tag{3.39}$$

Also, we rewrite (3.38) in the form

$$\begin{aligned}
&\left(a(u)\bar{\lambda}, \mu\right) - (\bar{\sigma}, \mu) + \left(\left(\tilde{a}_u(u^2)\lambda^2 + \tilde{b}_u(u_2)\right)\bar{u}, \mu\right) = \left(\tilde{a}_u(u^1)\bar{\lambda}(u - u^1), \mu\right), && \forall\, \mu \in \Lambda_h,\\
&(\bar{\lambda}, v) - (\bar{u}, \nabla \cdot v) = 0, && \forall\, v \in V_h,\\
&(\nabla \cdot \bar{\sigma}, w) + \left(\tilde{c}_u(u^2)\bar{u}, w\right) = 0, && \forall w \in W_h.
\end{aligned}$$

Then, apply (3.15) to this system to see that

$$\|\bar{u}\| \le C_1 \left(\|\bar{\lambda}\| + \|\bar{\sigma}\|_V\right) h,$$

which, together with (3.39), implies that

$$\|\bar{u}\| \le C_1 h \|\bar{u}\|.$$

Namely, $u^1 = u^2$ for h small enough. So, (3.39) yields that $\lambda^1 = \lambda^2$ and $\sigma^1 = \sigma^2$. Hence, the uniqueness is shown.

3.4. $H^{-s}(\Omega)$-error estimates. Apply (3.16) to (3.27) with $\theta = 2$ to see that

$$\begin{aligned} \|z\|_{-s} \le C_1\big\{&(\|\alpha\| + \|d\|)\, h^{\min(s+1,k+1)} + \|P_h u - u\| h^{\min(s+1,k^*)} \\ &+ \|\nabla\cdot d\| h^{\min(s+2,k^*)} + \|P_h u - u\|_{-s-1}\big\}. \end{aligned} \tag{3.40}$$

Then it follows from (3.2c) and (3.29) that

$$\begin{aligned} \|u - u_h\|_{-s} &\le \|u - P_h u\|_{-s} + \|z\|_{-s} \\ &\le C_1 \begin{cases} \|u\|_r\ h^{r+s} + \|u\|_{r_1}\ h^{r_1+s}, \\ \qquad 0 \le s \le k^*-2, \quad 2 \le r \le k+1, \quad 2 \le r_1 \le k^*, \\ \|u\|_{r+1}\ h^{r+k^*} + \|u\|_{r_1+1}\ h^{r_1+k^*-1}, \\ \qquad s = k^*-1, \quad 1 \le r \le k+1, \quad 1 \le r_1 \le k^*, \\ \|u\|_{r+1}\ h^{r+k^*} + \|u\|_{r_1+2}\ h^{r_1+k^*}, \\ \qquad s = k^*, \quad 1 \le r \le k+1, \quad 0 \le r_1 \le k^*. \end{cases} \end{aligned} \tag{3.41}$$

Now, let $\varphi \in H^s(\Omega)$. By (3.3a) and (3.27), we have

$$\begin{aligned} (d,\varphi) =&(d, \varphi - R_h\varphi) + (d, R_h\varphi) \\ =&(d, \varphi - R_h\varphi) + (a(u)\alpha, R_h\varphi) + \left(\tilde\Gamma_h z, R_h\varphi\right) - \left(\tilde\Gamma_h(P_h u - u), R_h\varphi\right) \\ =&(d, \varphi - R_h\varphi) - (a(u)\alpha, \varphi - R_h\varphi) + (\alpha, a\varphi - \Pi_h(a\varphi)) \\ &+ (z, \nabla\cdot(a(u)\varphi)) - \left(\tilde\Gamma_h z, \varphi - R_h\varphi\right) + \left(\tilde\Gamma_h z, \varphi\right) \\ &+ \left(\tilde\Gamma_h(P_h u - u), \varphi - R_h\varphi\right) - \left(\tilde\Gamma_h(P_h u - u), \varphi\right), \end{aligned}$$

so that

$$\begin{aligned} |(d,\varphi)| \le C_1\big\{&(\|d\| + \|\alpha\| + \|z\| + \|P_h u - u\|)\, h^{\min(s,k+1)} \\ &+ \|z\|_{-s+1} + \|P_h u - u\|_{-s}\big\}\|\varphi\|_s. \end{aligned}$$

This inequality, together with (3.29b), (3.2c), and (3.40), implies that

$$\|\sigma - \sigma_h\|_{-s} \le C_1 \begin{cases} \|u\|_{r+1}\ h^{r+s} + \|u\|_{r_1+1}\ h^{r_1+s}, \\ \qquad 0 \le s \le k^*-1, \quad 1 \le r \le k+1, \quad 1 \le r_1 \le k^*, \\ \|u\|_{r+1}\ h^{r+k^*} + \|u\|_{r_1+2}\ h^{r_1+k^*}, \\ \qquad s = k^*, \quad 1 \le r \le k+1, \quad 0 \le r_1 \le k^*. \end{cases} \tag{3.42}$$

The same result holds for $\lambda - \lambda_h$ by means of a similar argument. Finally, using (3.2c) and (3.27c), we see that, for $\varphi \in H^s(\Omega)$,

$$\begin{aligned} (\nabla\cdot d, \varphi) =&(\nabla\cdot d, \varphi - P_h\varphi) + (\nabla\cdot d, P_h\varphi) \\ =&(\nabla\cdot d, \varphi - P_h\varphi) - (\tilde c_u(u_h) z, \varphi) + (\tilde c_u(u_h) z, \varphi - P_h\varphi) \\ &+ (\tilde c_u(u_h)(P_h u - u), \varphi) + (\tilde c_u(u_h)(P_h u - u), P_h\varphi - \varphi). \end{aligned}$$

Consequently, we have

$$\begin{aligned}\|\nabla\cdot(\sigma-\sigma_h)\|_{-s} \leq & C\big\{(\|\nabla\cdot d\|+\|z\|+\|P_h u-u\|)\, h^{\min(s,k^*)} \\ & + \|z\|_{-s}+\|P_h u-u\|_{-s}\big\} \\ \leq & C_1\|u\|_{r+2}h^{r+s}, \quad 0\leq s,\ r\leq k^*. \end{aligned} \tag{3.43}$$

The results in (3.41)–(3.43) can be summarized in the following theorem.

Theorem 3.4. *Let Ω be $(s+2,2)$-regular with respect to M. Then for h sufficiently small the results in* (3.41)–(3.43) *hold.*

3.5. L^p-error estimates. The next theorem can be easily shown from (3.2c), (3.3b), the triangle inequality, and the quasi-regularity of T_h.

Theorem 3.5. *There exists a constant C_1 independent of h such that*

$$\begin{aligned}\|u-u_h\|_{0,p} \leq C_1\left(\|u\|_{r+1}\ h^r+\|u\|_{r_1,p}\ h^{r_1}+\|\nabla\cdot\sigma\|_{r_1}\ h^{r_1+\min(1,k^*-1)}\right),\\ 1\leq r\leq k+1, \quad 0\leq r_1\leq k^*, \quad 2\leq p\leq\infty.\end{aligned}$$

4. Postprocessing and superconvergence

In this section we consider a postprocessing scheme, which leads to a new more accurate approximation to the solution than u_h. The present scheme is an extension to the nonlinear case of the postprocessing procedure considered in [7] for the expanded mixed method for the linear problem. A similar approach for the usual linear mixed method is given in [22]. Let

$$W_h^* = \{w\in W : W|_E\in R(E) \text{ for each } E\in\mathcal{E}_h\},$$

where $R(E)=P_{k^*}(E)$ if $E\in\mathcal{E}_h$ is a triangle and $R(E)=P_{k^*}(E)\otimes P_{k^*}(E)$ if $E\in\mathcal{E}_h$ is a rectangle. Then the postprocessing scheme is given for $u_h^*\in W_h^*$ as the solution of the system

$$(u_h^*,1)_E=(u_h,1)_E, \quad E\in\mathcal{E}_h, \tag{4.1a}$$

$$\begin{aligned}(a(u_h^*)\nabla u_h^*-b(u_h^*),\nabla v)_E+(c(u_h),v)_E=(f,v)_E-\langle\sigma_h\cdot\nu_E,v\rangle_{\partial E},\\ \forall\ v\in R(E), \quad E\in\mathcal{E}_h,\end{aligned} \tag{4.1b}$$

where (u_h,σ_h) is the solution of (3.1) and ν_E is the outer unit normal to E.

To see that there exists at least one solution u_h^* to (4.1), let us consider the map $S: W_h^*\to W_h^*$ defined by

$$(Sy,1)_E=(u_h,1)_E, \quad E\in\mathcal{E}_h, \tag{4.2a}$$

$$\begin{aligned}(a(y)\nabla(Sy)-b(y),\nabla v)_E+(c(u_h),v)_E=(f,v)_E-\langle\sigma_h\cdot\nu_E,v\rangle_{\partial E},\\ \forall\ v\in R(E), \quad E\in\mathcal{E}_h,\end{aligned} \tag{4.2b}$$

for $y \in W_h^*$. Note that, by (3.1c),

$$(c(u_h), v)_E = (f, v)_E - \langle \sigma_h \cdot \nu_E, v \rangle_{\partial E}, \quad \forall\, v \in P_0(E),$$

so that the linear equations (4.2) define S uniquely. Now, choose $v = Sy$ in (4.2b) to see that the range of S is contained in a ball. Since S is clearly continuous, the Brower fixed point theorem implies that (3.4) has a solution, as illustrated in Theorem 3.2. The argument in §3.3 can also be used to show uniqueness of the solution for h sufficiently small.

To carry out an error analysis for (4.1), we also need a family $\{U_h\}_{0<h<1}$ of continuous spaces in $\bar{\Omega}$, which are piecewise polynomials over $\mathcal{E}_h$, such that

$$\inf \left\{ \|v - \xi\| + h \,\|\nabla(v - \xi)\| + h^2 \|v - \xi\|_{1,6} : \xi \in U_h \right\} \leq C \|v\|_s \; h^s, \tag{4.3}$$

if $2 \leq s \leq k^* + 1$. Finally, let P_E denote the L^2-projection onto $P_0(E)$. Because of the finite dimensionality of each U_h, the infimum in (4.3) is achieved. Let $\tilde{u}_h \in V_h$ be such that $\|u - \tilde{u}_h\| + h \,\|\nabla(u - \tilde{u}_h)\| + h^2 \|u - \tilde{u}_h\|_{1,6}$ is minimal. Then it follows from (4.3) that

$$\|\nabla \tilde{u}_h\|_{0,6} \leq C \|u\|_{1,6} \leq C \|u\|_{2+\varepsilon}. \tag{4.4}$$

Theorem 4.1. *Let $u \in H^{2+\varepsilon}(\Omega) \cap H^{k+2}(\Omega)$ be the solution of* (2.1) *and u_h^* be the solution of* (4.1)*. Then*

$$\|u - u_h^*\| \leq C_1 \|u\|_r h^{k^*+1}, \quad r = \max(k^* + 1, 3). \tag{4.5}$$

Proof. By (2.1) and the relation $\sigma = -\left(a(u)\nabla u - b(u)\right)$, we see that

$$(a(u)\nabla u - b(u), \nabla v)_E + (c(u), v)_E = (f, v)_E - \langle \sigma \cdot \nu_E, v \rangle_{\partial E}, \quad \forall\, v \in R(E). \tag{4.6}$$

Consequently, subtract (4.1) from (4.6) to yield the error equation

$$\begin{aligned}
&(a(u)\nabla u - a(u_h^*)\nabla u_h^*, \nabla v)_E - (b(u) - b(u_h^*), v)_E + (c(u) - c(u_h^*), v)_E \\
&\quad = \langle (\sigma - \sigma_h) \cdot \nu_E, v \rangle_{\partial E}, \quad \forall\, v \in R(E).
\end{aligned}$$

This inequality, together with (2.1c), implies that

$$\begin{aligned}
&a_0 \,\|\nabla(\tilde{u}_h - u_h^*)\|_E^2 \\
&\quad = a_0 \,\|\nabla(I - P_E)(\tilde{u}_h - u_h^*)\|_E^2 \\
&\quad \leq (a(u_h^*)\nabla(I - P_E)(\tilde{u}_h - u_h^*), \nabla(I - P_E)(\tilde{u}_h - u_h^*))_E \\
&\quad = (a(u)\nabla(\tilde{u}_h - u), \nabla(\tilde{u}_h - u_h^*))_E + ([a(u_h^*) - a(u)]\, \nabla \tilde{u}_h, \nabla(\tilde{u}_h - u_h^*))_E \\
&\qquad + (b(u) - b(u_h^*), \nabla(\tilde{u}_h - u_h^*))_E - (c(u) - c(u_h), (I - P_E)(\tilde{u}_h - u_h^*))_E \\
&\qquad - \langle (\sigma - \sigma_h) \cdot \nu_E, (I - P_E)(\tilde{u}_h - u_h^*) \rangle_{\partial E} \\
&\quad \leq C \,\|\nabla(\tilde{u}_h - u)\|_E \,\|\nabla(\tilde{u}_h - u_h^*)\|_E \\
&\qquad + \|a(u_h^*) - a(u)\|_{0,3,E} \,\|\nabla \tilde{u}_h\|_{0,6,E} \,\|\nabla(\tilde{u}_h - u_h^*)\| \\
&\qquad + \|b(u) - b(u_h^*)\|_E \,\|\nabla(\tilde{u}_h - u_h^*)\|_E + \|c(u) - c(u_h)\|_E \,\|(I - P_E)(\tilde{u}_h - u_h^*)\|_E \\
&\qquad + \left(h_E \int_{\partial E} |(\sigma_h - \sigma) \cdot \nu_E|^2 \, ds\right)^{1/2} \left(h_E^{-1} \int_{\partial E} |(I - P_E)(\tilde{u}_h - u_h^*)|^2 \, ds\right)^{1/2}.
\end{aligned} \tag{4.7}$$

Note that a scaling argument implies that

$$\|(I-P_E)(\tilde u_h-u_h^*)\|_E \le C h_E \,\|\nabla(I-P_E)(\tilde u_h-u_h^*)\|_E\,. \tag{4.8}$$

Exploit (4.4), (4.7), and (4.8) to obtain

$$\begin{aligned}\|\nabla(\tilde u_h-u_h^*)\|_E \le C_1\Big\{ &\|\nabla(\tilde u_h-u)\|_E+\|a(u_h^*)-a(u)\|_{0,3,E}\\ &+\|b(u)-b(u_h^*)\|_E+h_E\,\|c(u)-c(u_h)\|_E\\ &+\left(h_E\int_{\partial E}|(\sigma_h-\sigma)\cdot\nu_E|^2\,ds\right)^{1/2}\Big\}.\end{aligned} \tag{4.9}$$

Now, using the interpolation result

$$\|\phi\|_{0,3,E}\le C\|\phi\|_E^{1/2}\|\nabla\phi\|_E^{1/2},$$

it follows from (4.8), (4.9), and the assumption on the coefficients a, b, and c that

$$\begin{aligned}\|\tilde u_h-u_h^*\|_E \le C_1 h_E\Big\{ &\|\nabla(\tilde u_h-u)\|_E+\|u-u_h^*\|_E+h_E\|u-u_h\|_E\\ &+\left(h_E\int_{\partial E}|(\sigma_h-\sigma)\cdot\nu_E|^2\,ds\right)^{1/2}\Big\}+\|P_E(\tilde u_h-u_h^*)\|_E\,.\end{aligned} \tag{4.10}$$

Since P_E is bounded, it follows by (4.1a) that

$$\|P_E(\tilde u_h-u_h^*)\|_E\le\|\tilde u_h-u\|_E+\|P_hu-u_h\|_E,$$

which, together with (4.10), yields that

$$\begin{aligned}\|\tilde u_h-u_h^*\|_E \le C_1 h_E\Big\{ &\|\nabla(\tilde u_h-u)\|_E+\|u-u_h^*\|_E+h_E\|u-u_h\|_E\\ &+\left(h_E\int_{\partial E}|(\sigma_h-\sigma)\cdot\nu_E|^2\,ds\right)^{1/2}\Big\}\\ &+\|\tilde u_h-u\|_E+\|P_hu-u_h\|_E.\end{aligned}$$

Sum this expression over all $E\in\mathcal{E}_h$ to obtain

$$\begin{aligned}&\|\tilde u_h-u_h^*\|\\ &\le C_1\Big\{h\big(\|\nabla(\tilde u_h-u)\|+h\|u-u_h\|+\big(\sum_{E\in\mathcal{E}_h}h_E\int_{\partial E}|(\sigma-\Pi_h\sigma)\cdot\nu_E|^2\,ds\big)^{1/2}\\ &\qquad+\big(\sum_{E\in\mathcal{E}_h}h_E\int_{\partial E}|(\Pi_h\sigma-\sigma_h)\cdot\nu_E|^2\,ds\big)^{1/2}\big)+\|\tilde u_h-u\|+\|P_hu-u_h\|\Big\}\\ &\le C_1\Big\{h\big(\|\nabla(\tilde u_h-u)\|+h\|u-u_h\|+\big(\sum_{E\in\mathcal{E}_h}h_E\int_{\partial E}|(\sigma-\Pi_h\sigma)\cdot\nu_E|^2\,ds\big)^{1/2}\\ &\qquad+\|\sigma-\Pi_h\sigma\|+\|\sigma-\sigma_h\|\big)+\|\tilde u_h-u\|+\|P_hu-u_h\|\Big\},\end{aligned}$$

for h sufficiently small. Finally, apply (3.29), (3.37), (4.3), and the approximation property of Π_h to obtain the desired result (4.5). $\square$

5. Extension to a nonlinear problem

In this section we extend the previous analysis to the nonlinear problem

$$-\nabla \cdot A(x, \nabla u) = f(x) \quad \text{in } \Omega, \tag{5.1a}$$

$$u = -g \quad \text{on } \partial\Omega, \tag{5.1b}$$

and point out a difference between the usual mixed method and the expanded mixed method. We assume that $A : \bar{\Omega} \times \mathbb{R}^n \to \mathbb{R}^n$ is twice continuously differentiable with bounded derivatives through second order and that (5.1) is strictly elliptic at λ in the sense that there is a constant $a_0 > 0$ such that

$$\xi^T DA(x, \lambda)\xi \geq a_0 \|\xi\|^2_{\mathbb{R}^n}, \quad \xi \in \mathbb{R}^n, \quad (x, \lambda) \in \bar{\Omega} \times \mathbb{R}^n, \tag{5.2}$$

where $DA(x, \lambda) = (\partial A_i / \partial \lambda_j)$ is the $n \times n$ Jacobian matrix. The variable x is omitted in the notation below.

Using the previous notation, the expanded mixed form for (5.1) is formulated as follows: Find $(\sigma, \lambda, u) \in V \times \Lambda \times W$ such that

$$(A(\lambda), \mu) + (\sigma, \mu) = 0, \quad \forall\, \mu \in \Lambda, \tag{5.3a}$$

$$\begin{aligned} (\lambda, v) + (u, \nabla \cdot v) &= (g, v \cdot \nu)_{\partial\Omega}, && \forall\, v \in V, \\ (\nabla \cdot \sigma, w) &= (f, w), && \forall\, w \in W. \end{aligned} \tag{5.3b}$$

As in Theorem 2.2, it can be shown that (5.3) has a unique solution and is equivalent to (5.1) through the relations

$$\lambda = \nabla u \quad \text{and} \quad \sigma = -A(\nabla u).$$

The expanded mixed solution of (5.1) is $(\sigma_h, \lambda_h, u_h) \in V_h \times \Lambda_h \times W_h$ satisfying

$$(A(\lambda_h), \mu) + (\sigma_h, \mu) = 0, \quad \forall\, \mu \in \Lambda_h, \tag{5.4a}$$

$$(\lambda_h, v) + (u_h, \nabla \cdot v) = 0, \quad \forall\, v \in V_h, \tag{5.4b}$$

$$(\nabla \cdot \sigma_h, w) = (f, w), \quad \forall\, w \in W_h. \tag{5.4c}$$

Also, using the arguments in §3, it can be seen that (5.4) has a unique solution for $h > 0$ sufficiently small and produces optimal error estimates in the L^p and H^{-s}-norms. In particular, we state the L^2-error estimates as follows:

$$\|u-u_h\| \le C_1 \begin{cases} \|u\|_r\ h^r, & 2 \le r \le k^*,\ k \ge 2, \\ \|u\|_2\ h, & k=1, \text{ in the case of } k^*=k, \\ \|u\|_2\ h^{k+1}, & k=0,1, \text{ in the case of } k^*=k+1, \end{cases} \tag{5.5a}$$

$$\|\lambda-\lambda_h\| \le C_1 \|u\|_{r+1}\ h^r, \quad 1 \le r \le k+1, \tag{5.5b}$$

$$\|\sigma-\sigma_h\| \le C_1 \|u\|_{r+1}\ h^r, \quad 1 \le r \le k+1, \tag{5.5c}$$

$$\|\nabla\cdot(\sigma-\sigma_h)\| \le C_1 \|\nabla\cdot\sigma\|_r\ h^r, \quad 0 \le r \le k^*, \tag{5.5d}$$

$$\|u_h-P_hu\| \le C_1 \begin{cases} u\|_{k+2}\ h^{k+2}, & k \ge 2, \\ \|u\|_3\ h^2, & k=1, \text{ in the case of } k^*=k, \\ \|u\|_3\ h^{k+2}, & k=0,1, \text{ in the case of } k^*=k+1. \end{cases} \tag{5.5e}$$

The postprocessing scheme can be easily defined here; using (5.5e), analogous superconvergence results can be obtained. In the present case, we are able to obtain the superconvergence result (5.5e), which is of order $O(h^{k+2})$ in both cases where $k^*=k$ and $k^*=k+1$. The reason for this is that the coefficient A depends on λ instead of u. The vector variable has the error estimate of higher order, as shown in (5.5b).

We point out that attempts at using the usual mixed method based on the Brezzi-Douglas-Marini mixed finite elements ($n=2$) [6] and the Brezzi-Douglas-Durán-Fortin mixed finite elements ($n=3$) [4] (or some of the Chen-Douglas mixed finite elements [10]) for (5.1) are not entirely successful, as shown in [9]. The reason for this is that error equations couple the scalar variable u and the flux variable σ. Consequently, the errors of the scalar influence those of the flux. Hence the error estimates for the flux variable are not optimal since these mixed spaces use higher order polynomials for this variable than for the scalar. However, the expanded mixed method decouples the flux error equations from the scalar equations; as a consequence, optimal error estimates can be obtained for both the flux and scalar variables, as shown in (5.5).

6. Implementation and numerical results

In this section we present numerical results for the model problem

$$-\nabla\cdot(a(u)\nabla u)=f \quad \text{in } \Omega, \tag{6.1a}$$

$$u=-g \qquad \text{on } \partial\Omega. \tag{6.1b}$$

Before this, we need to consider implementation techniques for solving the corresponding mixed method solution $(\sigma_h,\lambda_h,u_h)\in V_h\times\Lambda_h\times W_h$ satisfying

$$(a(u_h)\lambda_h,\mu)-(\sigma_h,\mu)=0, \qquad \forall\ \mu\in\Lambda_h, \tag{6.2a}$$

$$(\lambda_h,v)-(u_h,\nabla\cdot v)=(g,v\cdot\nu)_{\partial\Omega}, \quad \forall\ v\in V_h, \tag{6.2b}$$

$$(\nabla\cdot\sigma_h,w)=(f,w), \qquad \forall\ w\in W_h. \tag{6.2c}$$

A linearized version of (6.2) is constructed as follows. Starting from any $(\sigma_h^0, \lambda_h^0, u_h^0) \in V_h \times \Lambda_h \times W_h$, we construct the sequence $(\sigma_h^m, \lambda_h^m, u_h^m) \in V_h \times \Lambda_h \times W_h$ by solving

$$\big(a(u_h^{m-1})\lambda_h^m, \mu\big) - (\sigma_h^m, \mu) = 0, \qquad \forall\ \mu \in \Lambda_h, \tag{6.3a}$$

$$(\lambda_h^m, v) - (u_h^m, \nabla\cdot v) = (g, v\cdot\nu)_{\partial\Omega}, \quad \forall\ v \in V_h, \tag{6.3b}$$

$$(\nabla\cdot\sigma_h^m, w) = (f, w), \qquad \forall\ w \in W_h. \tag{6.3c}$$

The ideas in [8] can be used to show that the sequence $\{(\sigma_h^m, \lambda_h^m, u_h^m)\}$ converges to $(\sigma_h, \lambda_h, u_h)$. Consequently, since (6.3) is linear for each m, the implementation techniques discussed in [7] for the linear expanded mixed method (e.g., alternating direction iterative methods, hybridization methods, and preconditioned iterative methods) can be applied here.

In this section we concentrate on the implementation of the nonlinear expanded mixed method (6.2) as a finite difference method. As mentioned in the introduction, it was shown [20, 21, 23] that the linear system arising from the usual mixed formulation can be simplified by use of certain quadrature rules for the lowest-order Raviart-Thomas-Nedelec spaces over a rectangular grid. That is, the mixed method system can be written as a cell-centered finite difference method. The same simplification is valid for the expanded Raviart-Thomas-Nedelec method [3]. However, an analogous simplification of the mixed method system as a finite difference method for another widely used space, the lowest order Brezzi-Douglas-Marini space [6] if $n = 2$ or the lowest-order Brezzi-Douglas-Durán-Fortin [4] space if $n = 3$, has not been known. We here derive a finite difference method for this space, without any loss in the rate of convergence. In particular, we show that for a diagonal tensor coefficient, the lowest order Brezzi-Douglas-Marini mixed method can be written as a cell-centered finite difference method with a five point stencil, and the Brezzi-Douglas-Durán-Fortin method can be given with a nine point stencil by using certain quadrature rules. For a full tensor coefficient, these two methods can be written with a nine and nineteen point stencil, respectively. We present the derivation only for the case where $n = 2$ and a is a scalar (or a diagonal tensor); the derivation for other cases is the same.

Throughout this section, we consider a partition of the rectangular domain Ω into rectangles: $\mathcal{E}_h = \{x_{i+1/2}\}_{i=0}^{n_x} \times \{y_{j+1/2}\}_{j=0}^{n_y}$, and define

$$\begin{aligned}
x_i &= \frac{1}{2}(x_{i+1/2} + x_{i-1/2}), \qquad && i = 1, \cdots, n_x,\\
y_j &= \frac{1}{2}(y_{j+1/2} + y_{j-1/2}), \qquad && j = 1, \cdots, n_y,\\
\Delta x_i &= x_{i+1/2} - x_{i-1/2}, &&\\
\Delta y_j &= y_{j+1/2} - y_{j-1/2}, &&\\
\Delta x_{i+1/2} &= \frac{1}{2}(\Delta x_i + \Delta x_{i+1}), &&\\
\Delta y_{j+1/2} &= \frac{1}{2}(\Delta y_j + \Delta y_{j+1}). &&
\end{aligned}$$

For any function $v(x,y)$, let $v_{i,j}$ denote $v(x_i,y_j)$, let $v_{i+1/2,j}$ denote $v(x_{i+1/2},y_j)$, and let $v_{i,j+1/2}$ denote $v(x_i,y_{j+1/2})$. Also, $(\cdot,\cdot)_T$ represents the midpoint rule in each coordinate direction, and $(\cdot,\cdot)_M$ denotes the two-point Gaussian quadrature rule for the x-component on the vertical edges and for the y-component on the horizontal edges of each cell. This choice of quadrature rules is compatible with the nodal basis functions for V_h, as seen in [2]. Also, we take Λ_h to be V_h. Then we introduce the modified method for $(\sigma_h,\lambda_h,u_h)\in V_h\times V_h\times W_h$ satisfying

$$(a(u_h)\lambda_h,\mu)_M-(\sigma_h,\mu)_M=0, \qquad \forall\ \mu\in V_h, \tag{6.4a}$$

$$(\lambda_h,v)_M-(u_h,\nabla\cdot v)=(g,v\cdot\nu)_{\partial\Omega,T}, \quad \forall\ v\in V_h, \tag{6.4b}$$

$$(\nabla\cdot\sigma_h,w)=(f,w)_T, \qquad \forall\ w\in W_h. \tag{6.4c}$$

Lemma 6.1. *The system* (6.4) *has a unique solution.*

Proof. Since (6.4) is a finite dimensional, square, linear system, existence follows from uniqueness. To show uniqueness, let $f=g=0$, $\mu=\lambda_h$, $v=\sigma_h$, and $w=u_h$. Then the three equations in (6.4) imply that

$$(a(u_h)\lambda_h,\lambda_h)_M=0.$$

By (2.1c), this equation means that λ_h vanishes at the two Gaussian points of the edges of each element. Hence, $\lambda_h=0$ everywhere. Then (6.4a) implies that

$$(\sigma_h,\sigma_h)_M=0,$$

so that $\sigma_h=0$. Since $V_h=\nabla\cdot W_h$, (6.4b) with $g=0$ yields that $u_h=0$. □

From (6.4b) we see that the normal component λ_h at any nodal point can be expressed as a difference of the pressure at the midpoints of the two adjacent elements. Namely (h is omitted below and let $\lambda=(\lambda^x,\lambda^y)$),

$$\lambda^x_{i+1/2,j}=-\frac{u_{i+1,j}-u_{i,j}}{\Delta x_{i+1/2}}; \tag{6.5}$$

a similar relation holds for $\lambda^y_{i,j+1/2}$. This corresponds to a finite difference approximation of the equation $\lambda=-\nabla u$.

Next, from (6.4a) it follows that the normal component of σ_h can be expressed at any nodal by the normal components of λ_h at the nodes of the adjacent elements. That is,

$$\sigma^x_{i+1/2,j}=\frac{1}{2}\left(a(u_h)_{i+1/2,j^+}+a(u_h)_{i+1/2,j^-}\right)\lambda^x_{i+1/2,j}, \tag{6.6}$$

where $(i+1/2,j^+)$ and $(i+1/2,j^-)$ denote the nodal points $(x_{i+1/2},y_j+\Delta y_j/(2\sqrt{3}))$ and $(x_{i+1/2},y_j-\Delta y_j/(2\sqrt{3}))$, respectively; an analogous equation is valid for $\sigma^y_{i,j+1/2}$. Thus we have a finite difference approximation of the relation $\sigma=a(u)\lambda$.

Finally, from (6.4c) we have

$$(\sigma^x_{i+1/2,j} - \sigma^x_{i-1/2,j})\Delta y_j + (\sigma^y_{i,j+1/2} - \sigma^y_{i,j-1/2})\Delta x_i = f_{ij}\Delta x_i \Delta y_j. \tag{6.7}$$

Substitute (6.5) and (6.6) into (6.7) to obtain a finite difference stencil for the scalar u, an approximation of the elliptic equation $-\nabla \cdot (a(u)\nabla u) = f$. Efficient iteration methods such as the Newton iteration method can be now used to solve the final nonlinear finite difference system.

Since the cell-centered finite difference methods arising from the Brezzi-Douglas-Marini (or Brezzi-Douglas-Durán-Fortin) and Raviart-Thomas-Nedelec mixed methods have the same form, the convergence results obtained for the latter [3, 21, 23] are also valid for the former. Namely, we have error estimates of order $O(h^2)$ in the L^2-norm for both the scalar and vector variables. For more details, refer to [2].

We now present two two-dimensional problems on the unit square with the Dirichlet boundary condition (5.1b) or (6.1b). In the first example, the coefficient $a(u)$ in (6.1a) is taken to be of the form $a(u) = u$. The true solution is

$$u(x, y) = x^2 + y^2 + \sin(x)\cos(y),$$

with f and g defined accordingly by (6.1). The expanded mixed formulation is discretized by means of the lowest-order Brezzi-Douglas-Marini space [6] on rectangles as in (6.4). Namely, we solve a cell-centered finite difference system for the scalar u over a uniform rectangular decomposition of Ω. In Table 1 we show the errors and convergence rates. Note that the orders of convergence in L^2 and L^∞ are two in all cases. So, in fact, we have a superconvergent result for the scalar u.

$1/h$	L^∞-error ($\times 10^2$)	$L^\infty-$order	L^2-error ($\times 10^2$)	L^2-order
5	1.550	-	1.470	-
10	0.470	1.73	0.380	1.95
20	0.120	1.97	0.091	2.06
40	0.029	2.05	0.022	2.05

Table 1. Convergence rates for the scalar in example one.

In the second example, the coefficient $A(\nabla u)$ in (5.1a) is defined by

$$A(v) = (v_1, 3v_2/2 - \sin(2v_2)/4), \quad v = (v_1, v_2),$$

$g \equiv 0$ in (5.1b), and f in (5.1a) is given by

$$f(x,y) = 2(y-y^2) + (x-x^2)\big(3 - \cos\big(2(x-x^2)(1-2y)\big)\big).$$

Problem (5.1) has a unique solution [14] for such chosen functions. The Brezzi-Douglas-Marini space [6] of lowest order on a uniform triangular decomposition of Ω is exploited this time. Tables 2 and 3 show the errors and convergence rates for the scalar and the flux variable, respectively. The convergence rate for the scalar is $O(h)$, while it is $O(h^2)$ for the flux. The numerical results in Tables 1, 2, and 3 confirm the theoretical results from the previous sections.

$1/h$	L^∞-error ($\times 10^2$)	$L^\infty-$order	L^2-error ($\times 10^2$)	L^2-order
5	3.57	-	2.50	-
10	1.89	0.91	1.20	0.99
20	0.99	0.93	0.63	1.02
40	0.52	0.98	0.30	1.09

Table 2. Convergence rates for the scalar in example two.

$1/h$	L^∞-error ($\times 10^2$)	$L^\infty-$order	L^2-error ($\times 10^2$)	L^2-order
5	1.870	-	1.540	-
10	0.540	1.79	0.430	1.84
20	0.140	1.94	0.110	1.97
40	0.032	2.12	0.027	2.03

Table 3. Convergence rates for the flux in example two.

References

1. D. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

2. T. Arbogast and Z. Chen, *On the implementation of mixed methods as finite difference methods for second order elliptic problems*, Math. Comp. (1995, in press).
3. T. Arbogast, M. Wheeler, and I. Yotov, *Mixed finite elements for elliptic problems with tensor coefficients as finite differences*, TR94-02, Dept. of Comp. and Appl. Math, Rice Univ. (1994).
4. 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), 237–250.
5. F. Brezzi, J. Douglas, Jr., M. Fortin, and L. Marini, *Efficient rectangular mixed finite elements in two and three space variables*, RAIRO Modèl. Math. Anal. Numér **21** (1987), 581–604.
6. F. Brezzi, J. Douglas, Jr., and L. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math. **47** (1985), 217–235.
7. Z. Chen, *Expanded mixed finite element methods for linear second order elliptic problems I*, IMA Preprint Series # 1219, 1994, submitted to RAIRO Modèl. Math. Anal. Numér.
8. Z. Chen, *On the existence, uniqueness and convergence of nonlinear mixed finite element methods*, Mat. Aplic. Comput. **8** (1989), 241–258.
9. Z. Chen, *BDM mixed methods for a nonlinear elliptic problem*, J. Comp. Appl. Math. **53** (1994), 207–223.
10. Z. Chen and J. Douglas, Jr., *Prismatic mixed finite elements for second order elliptic problems*, Calcolo **26** (1989), 135–148.
11. J. Douglas, Jr. and T. Dupont, *A Galerkin method for a nonlinear Dirichlet problem*, Math. Comp. **29** (1975), 689–696.
12. J. Douglas, Jr. and J. Roberts, *Global estimates for mixed methods for second order elliptic problems*, Math. Comp. **45** (1985), 39–52.
13. J. Douglas, Jr. and J. Wang, *A new family of mixed finite element spaces over rectangles*, Mat. Aplic. Comput. **12** (1993), 183–197.
14. D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Grundlehren der Mathematischen Wissenschaften, vol. 224, Springer-Verlag, Berlin, 1977.
15. J. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications, Vol. I*, Springer-Verlag, Berlin, 1970.
16. F. Milner, *Mixed finite element methods for quasilinear second order elliptic problems*, Math. Comp. **44** (1985), 303–320.
17. J. C. Nedelec, *Mixed finite elements in* $\mathbb{R}^3$, Numer. Math. **35** (1980), 315–341.
18. J. C. Nedelec, *A new family of mixed finite elements in* $\mathbb{R}^3$, Numer. Math. **50** (1986), 57–81.
19. P.A. Raviart and J.M. Thomas, *A mixed finite element method for second order elliptic problems*, Lecture Notes in Math. **606**, Springer, Berlin, 1977, pp. 292–315.
20. T. Russell and M. Wheeler, *Finite element and finite difference methods for continuous flows in porous media*, Chapter II, The Mathematics of Reservoir Simulation, R. Ewing, ed., Frontiers in Applied Mathematics 1, Society for Industrial and Applied Mathematics, Philadelphia, 1983, pp. 35–106.
21. J. Shen, *A block finite difference scheme for second order elliptic problems with discontinuous coefficients*, to appear in SIAM J. Numer. Anal.
22. R. Stenberg, *Postprocessing schemes for some mixed finite elements*, RAIRO Modèl. Math. Anal. Numér. **25** (1991), 151–167.
23. A. Weiser and M. Wheeler, *On convergence of block-centered finite-differences for elliptic problems*, SIAM J. Numer. Anal. **25** (1988), 351–375.

Department of Mathematics and the Institute for Scientific Computation, Texas A&M University, College Station, TX 77843.

E-mail address: zchen@isc.tamu.edu