SIAM J. NUMER. ANAL.
Vol. 39, No. 4, pp. 1170–1196

A GENERAL FRAMEWORK FOR CONSTRUCTING AND ANALYZING MIXED FINITE VOLUME METHODS ON QUADRILATERAL GRIDS: THE OVERLAPPING COVOLUME CASE*

SO-HSIANG CHOU†, DO Y. KWAK‡, AND KWANG Y. KIM‡

Abstract. We present a general framework for constructing and analyzing finite volume methods applied to the mixed formulation of second-order elliptic problems on quadrilateral grids. The control volumes, or covolumes, in the grids overlap. An overlapping finite volume method of this type was first introduced by Russell in [T. F. Russell, *Tech. report* 3, Reservoir Simulation Research Corp., Tulsa, OK, 1995] and was tested for a variety of problems on rectangular and quadrilateral grids in [Z. Cai et al., *Comput Geosci.*, 1 (1997), pp. 289–315]. Later in [S. H. Chou and D. Y. Kwak, *SIAM J. Numer. Anal.*, 37 (2000), pp. 758–771], Chou and Kwak reformulated it as their mixed covolume method and proved optimal order error estimates using the covolume methodology from [S. H. Chou, *Math. Comp.*, 66 (1997), pp. 85–104] and [S. H. Chou and D. Y. Kwak, *SIAM J. Numer. Anal.*, 35 (1998), pp. 494–507]. However, their treatment was restricted to the case of diagonal coefficient tensor and rectangular grids since a different approach was needed for the quadrilateral (distorted rectangular) case. In this paper we give a new framework, which can handle not only the rectangular anisotropic case but also the anisotropic and irregular grid cases in which the locally supported test functions are images of the natural unit coordinate vectors under the Piola transformation. Our theory sheds light on how to create new test functions using quadratures and now covers Russell's quadrilateral case.

Key words. covolume method, mixed finite element, MAC, control volume finite element method, control volume, finite volume element, error estimate, porous media

AMS subject classifications. 65N30, 65N22, 65F10

PII. S003614290037544X

1. Introduction.

Consider the following second-order elliptic boundary value problem on a bounded polygonal domain Ω in $\mathbb{R}^2$ with the boundary $\partial\Omega$:

$$\begin{cases} -\operatorname{div}(K(\mathbf{x})\nabla p + \mathbf{b}(\mathbf{x})p) + c(\mathbf{x})p = f & \text{in } \Omega, \\ (K\nabla p + \mathbf{b}p)\cdot \mathbf{n} = 0 & \text{on } \partial\Omega. \end{cases} \tag{1.1}$$

Here $\mathbf{n}$ is the outward unit normal vector to $\partial\Omega$, and the coefficient K is a symmetric and uniformly positive-definite matrix; i.e., there exist two positive constants α_1 and α_2 such that

$$\alpha_1 \boldsymbol{\xi}^T\boldsymbol{\xi} \le \boldsymbol{\xi}^T K(\mathbf{x})\boldsymbol{\xi} \le \alpha_2 \boldsymbol{\xi}^T\boldsymbol{\xi} \qquad \forall \boldsymbol{\xi} \in \mathbb{R}^2,\ \mathbf{x} \in \overline{\Omega}. \tag{1.2}$$

For brevity we will often omit dependency of the coefficients $K, \mathbf{b}$, and c on the space variable $\mathbf{x}$.

In some applications, such as mathematical modeling of fluid flow in a porous media or current flow in semiconductor devices, it is often more important to gain

*Received by the editors July 13, 2000; accepted for publication (in revised form) April 16, 2001; published electronically September 5, 2001.

http://www.siam.org/journals/sinum/39-4/37544.html

†Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403-0221 (chou@bgnet.bgsu.edu). The research of this author was supported by NSF grant DMS–0074259.

‡Department of Mathematics and Statistics, Korea Advanced Institute of Science and Technology, Taejon, Korea 305-701 (dykwak@math.kaist.ac.kr, kky@mathx.kaist.ac.kr). The work of these authors was partially supported by KOSEF, Korea.

accurate approximation for the vector variable $\mathbf{u} = -(K\nabla p + \mathbf{b}p)$ rather than the scalar variable p itself. For this reason we rewrite the problem (1.1) as a system of first-order partial differential equations

$$K^{-1}\mathbf{u} + \nabla p + \boldsymbol{\beta} p = 0 \qquad \text{in } \Omega, \tag{1.3a}$$

$$\operatorname{div}\mathbf{u} + cp = f \qquad \text{in } \Omega, \tag{1.3b}$$

$$\mathbf{u}\cdot\mathbf{n} = 0 \qquad \text{on } \partial\Omega, \tag{1.3c}$$

where we set $\boldsymbol{\beta} = K^{-1}\mathbf{b}$. The first equation of (1.3) represents the Darcy law if $\mathbf{b} = \mathbf{0}$, and so in what follows we shall refer to $\mathbf{u}$ and p as the velocity and the pressure variables, resp., as in mathematical modeling of fluid flow in porous media. The homogeneous Neumann condition corresponds to a no-flow condition on the boundary, which is natural for our application here, but all the results in this paper hold for the Dirichlet condition as well.

Now let us introduce the function spaces

$$\mathbf{H}(\operatorname{div};\Omega) = \{\mathbf{v} \in (L^2(\Omega))^2 : \operatorname{div}\mathbf{v} \in L^2(\Omega)\}, \tag{1.4}$$

$$\mathbf{V} = \{\mathbf{v} \in \mathbf{H}(\operatorname{div};\Omega) : \mathbf{u}\cdot\mathbf{n} = 0 \text{ on } \partial\Omega\}, \tag{1.5}$$

$$W = L^2(\Omega). \tag{1.6}$$

Then the associated weak formulation for (1.3) is to find $(\mathbf{u}, p) \in \mathbf{V} \times W$ such that

$$(K^{-1}\mathbf{u}, \mathbf{v}) - (\operatorname{div}\mathbf{v}, p) + (\boldsymbol{\beta} p, \mathbf{v}) = 0 \qquad \forall \mathbf{v} \in \mathbf{V}, \tag{1.7a}$$

$$(\operatorname{div}\mathbf{u}, w) + (cp, w) = (f, w) \qquad \forall w \in W, \tag{1.7b}$$

where $(\cdot,\cdot)$ denotes the standard inner product in $L^2(\Omega)$ or $(L^2(\Omega))^2$. In the case $c \equiv 0$, we take $W = L_0^2(\Omega) = \{w \in L^2(\Omega) : \int_\Omega w = 0\}$ for compatibility.

The standard mixed finite element method chooses a pair of finite element spaces $\mathbf{V}_h \times W_h \subset \mathbf{V} \times W$ satisfying the inf-sup condition which by now is relatively well understood, at least in the case of smooth and isotropic K; for example, see [2, 13, 15, 27] and references therein. On the other hand, in recent years attempts in [4, 7, 10, 21, 25], among others, have been made to develop a parallel theory for mixed finite volume methods. Depending on how they are interpreted, these methods are referred to as mixed covolume methods (a name preferred by us), mixed control volume methods, and mixed balance methods. Regardless of physical interpretations, mathematically this class of methods can be analyzed as Petrov–Galerkin methods with the trial space being related to typical finite element spaces, and the test space being related to finite volumes. Based on early results of Chou and Kwak [6, 7, 8], a unified framework is given in [12] for a number of finite volume or covolume schemes, relating all these schemes to the standard mixed finite element method via an injective *transfer* operator γ_h between the trial space $\mathbf{V}_h$ and the corresponding test spaces $\mathbf{Y}_h$ (e.g., piecewise constant vectors, characteristic functions of finite volumes, etc.). More specifically, all the covolume schemes satisfy the problem of finding $\mathbf{u}_h \in \mathbf{V}_h; p_h \in W_h$ such that

$$(K^{-1}\mathbf{u}_h, \gamma_h\mathbf{v}) + b(\gamma_h\mathbf{v}, p_h) + (\boldsymbol{\beta} p_h, \gamma_h\mathbf{v}) = 0 \qquad \forall \mathbf{v} \in \mathbf{V}_h, \tag{1.8a}$$

$$(\operatorname{div}\mathbf{u}_h, w) + (cp_h, w) = (f, w) \qquad \forall w \in W_h, \tag{1.8b}$$

where the bilinear form b is related to the divergence term. Comparing this with the mixed finite element method associated with the problem (1.7), one sees the similarity.

Thus the most important factor in the above framework is the construction of the operator γ_h and test spaces that will maintain optimal order convergence rates, and superconvergence results in Sobolev spaces, when compared with the corresponding mixed finite element methods. For the nonoverlapping covolume case (see below for the definition) the theory is relatively complete (the main issues are settled). One objective of this paper is to develop a theory for the overlapping covolume case, where the concept of a vector test space must go beyond the simple idea of characteristic functions of finite volumes and where the transfer operator must go beyond the simple idea of lumping.

1.1. The covolume concept. Covolume schemes are popular [5, 18, 23] in practical fluid mechanics computations due to their conservative properties; namely, they represent discrete analogues of the underlying physical conservation laws dictating the behavior of the fluid system. For instance, if the main variables of interest of the underlying fluid system consist of a state variable (concentration, temperature, pressure, etc.) and a flux variable (gradient of the state variable), the covolume method then uses two partitions of the fluid domain to find approximations of these two variables. A conservation law on the primal volumes is used for the state variable and a constitutive law on the dual volumes or covolumes is used for the flux variable. In the case of porous media flow the conservation law for the primal volumes is the mass conservation law, and the constitutive law for the covolumes is the Darcy law. Originally, covolume methods were interpreted as finite difference methods on an irregular network in fluid dynamics that compute only the normal components of the velocity variable. (See the survey paper [23] by Nicolaides, Porsching, and Hall for literature up to 1995.) Later Chou [6] connected covolume methods to the mixed finite element method framework in the case of the Stokes equations.

1.2. Need for analysis of the overlapping covolume case. We now briefly describe how the methodology in [6] was adapted to study elliptic problems. Let $\mathcal{Q}_h$ be a (primal) partition of the fluid domain into a union of triangles or quadrilaterals. In the standard mixed finite element method, one would use only $\mathcal{Q}_h$ to define the discrete weak formulation associated with (1.7). In covolume methods, one uses two partitions: a primal partition $\mathcal{Q}_h$ on which the local mass conservation law (1.3b) holds and a dual partition $\mathcal{Q}_h'$ (a union of covolumes) over which (1.3a) holds in the average sense. The most well-known example is the MAC (Marker and Cell) scheme [19] that uses two staggered rectangular grids. In general, we can classify covolume methods into overlapping and nonoverlapping types according to whether or not covolumes overlap.

For example, we have nonoverlapping covolumes in Figure 1. In the top figure the primal partition is the union of rectangles. A typical interior covolume in the dual partition is the dashed quadrilateral, the closure of the union of the two triangles $T_E^+ \cup T_E^-$ sharing the common side E. The two vertices in the interiors of the two rectangles are their centers. Note that each edge E of the primal element corresponds to a covolume. On the boundary the covolume is either a T_E^+ or a T_E^-. The bottom figure has a similar interpretation. Such covolume methods were analyzed in [6, 8, 9] and finally presented in a unified manner in [12]. On the other hand, in Figure 2, the dashed covolumes are overlapping. This type of staggered grid is also adopted in the MAC method and is particularly of interest in oil recovery simulations. A logically rectangular (quadrilateral) case is further shown in Figure 3. It can be viewed as a distorted figure of the two bottom rectangles in Figure 2.

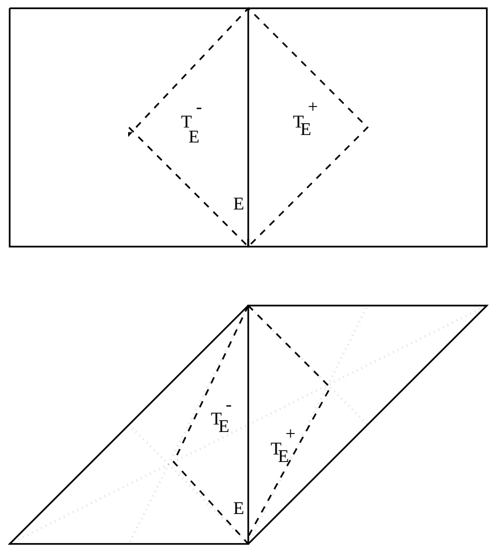

FIG. 1. *Primal and dual domains; the dual element (covolume)* $Q'_E = T_E^+ \cup E \cup T_E^-$.

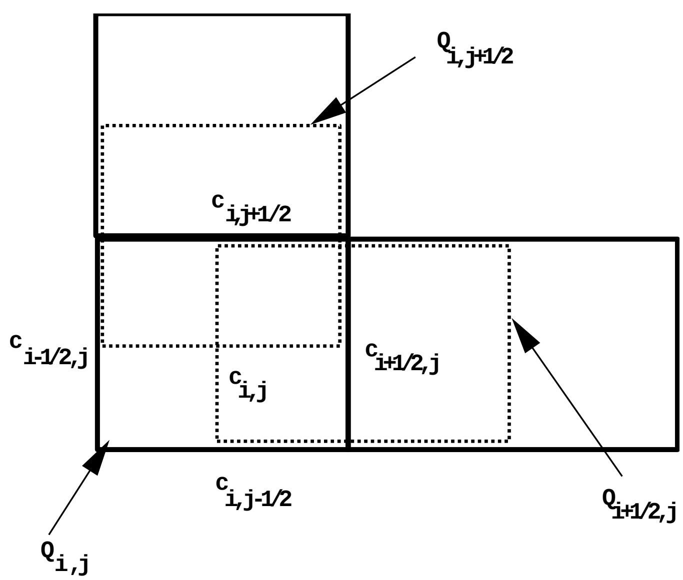

FIG. 2. *Dual domain with overlapping covolumes. The dashed boxes* $Q_{i+1/2}$ *and* $Q_{i,j+1/2}$ *are covolumes;* $c_{i+1/2,j}$ *and* $c_{i,j+1/2}$ *are the midpoints of the edges of the primal volume* $Q_{i,j}$ *whose center is* $c_{i,j}$.

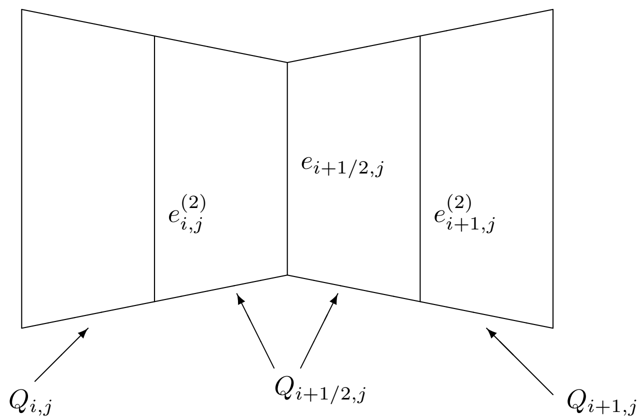

FIG. 3. *Primal and dual domains.*

Interested in providing a new simulation scheme for complex reservoir systems, Russell [25] designed a "control volume finite element" method for the mixed system (1.3) on quadrilateral grids. The method is considered a natural generalization of the traditional rectangular case. He and his coworkers carried out various numerical experiments which show the robustness of the scheme [4, 20]. However, no rigorous error analysis was given. Later in [9], Chou and Kwak, observing the similarity between Russell's method and the covolume method, reformulated his scheme using the covolume methodology from [6, 8, 9] and proved optimal order error estimates for it by comparing the method with the standard mixed finite element method. However, it became clear that there was a major obstacle in extending the techniques of [9] to the quadrilateral case due to the presence of the Piola transformation. In addition, the old idea of using characteristic functions to choose control volumes is no longer correct and the transfer operator mentioned previously was much harder to capture or define. A new approach is called for.

In this paper we extend the covolume approach to quadrilateral grids, or for that matter, to grids defined through a common master domain. We again use the covolume methodology of Chou to maintain the general mixed covolume formulation (1.8), but our main theoretical tool for error analysis will be to develop global estimates similar to that of Douglas and Roberts [13].

Let us give a brief explanation of the role of the quadrilateral grids in the reservoir simulation. Early methods and codes in the field are mostly of the finite difference type [24]; later the finite element type began to emerge [14]. Thus, the natural second step was to go from rectangular grids to the logically rectangular (quadrilateral) case in two or three space dimensions. On the finite element side, Thomas's parallelogram grids [26] appeared in 1977 and Wang–Mathew's theory [30] on quadrilateral Raviart–Thomas spaces in 1994. On the finite volume side, the first true quadrilateral case consistent with the old finite difference approach on rectangular grids seems to be first proposed by Russell [25]. The present paper is a natural next step in setting up a theoretical basis.

It is important to point out that the greatest advantages of mixed finite element or volume methods occur when the tensor K is strongly heterogeneous or discontinuous (e.g., piecewise constant with discontinuities at edges of primal elements). This is when these methods yield substantially better fluxes than standard schemes, such as Galerkin methods. For the important Darcy flow case $\mathbf{b} \equiv \mathbf{0}$, our results cover the discontinuous full tensor case (see Remarks 6.1–6.2 and Theorem 6.7). However, for the general convection-diffusion case ($\mathbf{b}, c$ not identically zero), our theoretical convergence results (e.g., Theorem 6.5 below) in this paper require $K \in W^{1,\infty}$. Of course, we can also obtain superconvergence results under the smooth assumption. For homogeneous or smooth K, mixed methods are generally not worth the overhead of solving for two variables instead of one. A very effective and inexpensive local flux recovery procedure based only on standard Galerkin methods can be found in Chou and Tang [11]. The flux thus obtained has the good feature of conserving mass elementwise.

The rest of the paper is organized as follows. In the next section we introduce some notation about the quadrilateral grids which will be used throughout the paper. In section 3, we show how to define the test function spaces in conjunction with the transfer operator and the Piola transformation from the reference rectangle to the quadrilateral grids. It is shown that the essential properties of the transfer operator are coercivity and strong and weak approximabilities. These are also important for deriving error estimates. In section 4 we describe and reformulate the mixed covolume method in Galerkin form. Then we establish some lemmas in section 5 and prove optimal error estimates for the mixed covolume method using the duality argument of Douglas and Roberts in section 6. We analyze Russell's scheme as a modification of our mixed covolume method in section 7. Finally, in the appendix we show how the anisotropic rectangular case is also solved.

2. Quadrilateral grids. Let $\mathcal{Q}_h$ be a partition of Ω into convex quadrilaterals with diameters less than or equal to h. The intersection, if any, of any two (closed) quadrilaterals in the partition must consist of entirely their own common edges. The partition is thus logically rectangular in the sense that each quadrilateral has unique eastern, western, northern, and southern adjacent neighbors if they exist. Hence one can write $\mathcal{Q}_h = \{Q_{i,j}\}$, indexed by two indices. (This type of simple logic feature is particularly attractive when it comes to coding the three-dimensional case.)

The eastern (resp., northern) and the western (resp., southern) edges of $Q_{i,j}$ are denoted by

$$e_{i\pm 1/2,j} = \partial Q_{i,j} \cap \partial Q_{i\pm 1,j}, \qquad (\text{resp., } e_{i,j\pm 1/2} = \partial Q_{i,j} \cap \partial Q_{i,j\pm 1}),$$

with $e_{i,j}^{(1)}$ (resp., $e_{i,j}^{(2)}$) as the line segment joining their midpoints. The volume formed by the left-hand quadrilateral between the edges $e_{i,j}^{(2)}$ and $e_{i+1/2,j}$ and the right-hand quadrilateral between the edges $e_{i+1,j}^{(2)}$ and $e_{i+1/2,j}$ is denoted by $Q_{i+1/2,j}$ (cf. Figure 3). The volume $Q_{i,j+1/2}$ is defined in a similar way.

Let $\hat{\mathbf{x}} = (\hat{x}, \hat{y})$ and $\mathbf{x} = (x, y)$. We take the unit square $\hat{Q} = [0,1] \times [0,1]$ as the reference element in the $\hat{x}\hat{y}$-plane with vertices denoted by

$$\hat{\mathbf{x}}_1 = (0,0), \quad \hat{\mathbf{x}}_2 = (1,0), \quad \hat{\mathbf{x}}_3 = (1,1), \quad \hat{\mathbf{x}}_4 = (0,1).$$

Let Q be a convex quadrilateral with the $\mathbf{x}_i$'s of the vertices placed counterclockwise. Then there exists a unique invertible bilinear transformation F_Q which maps $\hat{Q}$ onto Q and satisfies

$$\mathbf{x}_i = F_Q(\hat{\mathbf{x}}_i), \qquad i = 1, 2, 3, 4.$$

In fact, it is given by

$$\mathbf{x} = F_Q(\hat{\mathbf{x}}) = \mathbf{x}_1 + \mathbf{x}_{21}\hat{x} + \mathbf{x}_{41}\hat{y} + \mathbf{g}\hat{x}\hat{y}, \tag{2.1}$$

where we set

$$\mathbf{x}_{ij} = \mathbf{x}_i - \mathbf{x}_j, \quad \mathbf{g} = \mathbf{x}_{12} + \mathbf{x}_{34}.$$

By a simple calculation it is easy to see that the Jacobian matrix $\mathcal{J}_Q$ of F_Q is given by

$$\mathcal{J}_Q = \begin{pmatrix} \frac{\partial x}{\partial \hat{x}} & \frac{\partial x}{\partial \hat{y}} \\ \frac{\partial y}{\partial \hat{x}} & \frac{\partial y}{\partial \hat{y}} \end{pmatrix} = (\mathbf{x}_{21} + \mathbf{g}\hat{y}, \mathbf{x}_{41} + \mathbf{g}\hat{x}). \tag{2.2}$$

Denote by S_i the subtriangle of Q with vertices $\mathbf{x}_{i-1}, \mathbf{x}_i$, and $\mathbf{x}_{i+1}$ ($\mathbf{x}_0 = \mathbf{x}_4$). Let h_Q be the diameter of Q and $\rho_Q = 2\min_{1\le i\le 4}$ {diameter of a circle inscribed in S_i}. Throughout the paper we assume a *regular* family of partitions $\mathcal{Q} = \{\mathcal{Q}_h\}$; i.e., there exists a positive constant σ, independent of h, such that

$$\frac{h_Q}{\rho_Q} \le \sigma \quad \forall Q \in \mathcal{Q}_h, \forall \mathcal{Q}_h \in \mathcal{Q}. \tag{2.3}$$

The following upper bounds can be found, e.g., in [17]:

$$\|\mathcal{J}_Q\|_{\infty,\hat{Q}} \le Ch_Q, \qquad \|\mathcal{J}_Q^{-1}\|_{\infty,Q} \le Ch_Q^{-1}, \tag{2.4}$$

where $\|M\|_{\infty,K} := \sup_{\mathbf{x}\in K}\|M(x)\|$, the supremum of the spectral norm of the matrix function M. Hereafter C will denote a generic positive constant which is independent of h. It may have different values in different places, especially when used in proof.

Simple calculation shows that the determinant $J_Q = \det \mathcal{J}_Q$ is a linear function of $\hat{x}$ and $\hat{y}$:

$$J_Q(\hat{x}, \hat{y}) = \alpha + \beta\hat{x} + \gamma\hat{y}, \tag{2.5}$$

where

$$\alpha = \det(\mathbf{x}_{21}, \mathbf{x}_{41}), \quad \beta = \det(\mathbf{x}_{21}, \mathbf{g}), \quad \gamma = \det(\mathbf{g}, \mathbf{x}_{41}).$$

The area of the quadrilateral Q is equal to $J_Q(1/2, 1/2)$, since we have by the midpoint rule that

$$|Q| = \int_Q dxdy = \int_0^1\int_0^1 J_Q(\hat{x}, \hat{y})\, d\hat{x}d\hat{y} = J_Q(1/2, 1/2). \tag{2.6}$$

The following upper bounds for the L_∞-norm of the functions J_Q and J_Q^{-1} also can be found in [17]:

$$|J_Q|_{\infty,\hat{Q}} \le Ch_Q^2, \qquad |J_Q^{-1}|_{\infty,Q} \le Ch_Q^{-2}. \tag{2.7}$$

Furthermore, we shall assume throughout the paper that each quadrilateral in the family of partitions is almost parallelogram :$\|\mathbf{g}\| = O(h_Q^2)$. In other words,

(2.8) *the distance between the midpoints of the two diagonals of Q is $O(h_Q^2)$.*

This condition is easily satisfied if the partitions are obtained by symmetric refinement of quadrilaterals via bisection on edges. A simple consequence of the condition (2.8) that we will use for proving superconvergence results is

$$|\beta| + |\gamma| \le Ch_Q^3. \tag{2.9}$$

REMARK 2.1. *It is well known that regularity is equivalent to the minimum angle condition in the case of a partition with triangular elements. A similar result holds for the regular quadrilateral family satisfying* (2.8). *Specifically, one needs to replace the minimum angle condition with the following two conditions: the uniform boundedness of h_Q/h'_Q, the ratios of diameter to shortest edge h'_Q, and the existence of a positive constant s independent of h such that $|\cos(\theta_Q)| \le s < 1$ for all Q with θ_Q being any interior angle of Q.*

REMARK 2.2. *Condition* (2.8) *is used to obtain superconvergence results. When $\mathbf{b} \equiv \mathbf{0}$ (Darcy flow case), our general optimal order convergence results below will be valid for regular partitions without this condition.*

In three dimensions the reference element becomes a unit cube, while its images under the trilinear mappings are, in general, hexahedra with nonplanar faces [22]. A two-dimensional analogy would be to replace the bilinear transformation with a smooth F whose images are curved quadrilaterals.

REMARK 2.3. *In particular, one may ask if there is a relation between the almost parallelogram grid and a grid that results from a mapping of a uniform grid* [1]. *For example, what should be the smoothness of the mapping? It is easy to see that such a bijective mapping $F : \hat{\Omega} \to \Omega$ should be at least $C^2(\hat{\Omega} \cup \partial\hat{\Omega})$ to generate an almost parallelogram grid. First, if we take F as the bilinear transformation from a square block $\hat{\Omega}$ to a quadrilateral block, sending uniform meshes on $\hat{\Omega}$ to quadrilateral meshes on $\Omega = F(\hat{\Omega})$. Then using the mean value theorem twice shows that* (2.8) *is valid with the constant C dependent on the mixed second derivative of F. Hence an almost parallelogram grid can be generated via uniform meshes. The familiar biquadratic mapping as in the isoparametric family provides another example. Here $\hat{\Omega}$ is a square, $\Omega = F(\hat{\Omega})$, and the points $F(m)$ (m denotes the centers of the square elements in $\hat{\Omega}$) in each curved quadrilateral element K need to be kept within a distance of $O(h_K^2)$ from $\tilde{F}(m)$, the image of m under the corresponding bilinear transformation $\tilde{F}$. See Remark* 2.1 *and Ciarlet* (*page* 247 *of* [3]).

The Piola transformation $\mathcal{P}_Q$ transforms a vector-valued function on $\hat{Q}$ to one on Q by

$$\mathbf{v} = \mathcal{P}_Q \hat{\mathbf{v}} = \frac{1}{J}\mathcal{J}\hat{\mathbf{v}} \circ F^{-1}, \tag{2.10}$$

where we drop the subscript Q for brevity. This transformation preserves the $H(\mathrm{div})$ space on the reference element and has the following well-known properties (cf. [26], [28], [30]): If we let $\hat{p} = p \circ F$, then

$$\int_Q \nabla p \cdot \mathbf{v}\, dxdy = \int_{\hat{Q}} \hat{\nabla}\hat{p} \cdot \hat{\mathbf{v}}\, d\hat{x}d\hat{y}, \tag{2.11}$$

$$\operatorname{div} \mathbf{v} = \frac{1}{J} \operatorname{div} \hat{\mathbf{v}}. \tag{2.12}$$

We also need the following lemma for error analysis.

LEMMA 2.1. *Let* $\mathbf{v}$ *and* $\hat{\mathbf{v}}$ *be related by* (2.10). *For regular partitions, there exist positive constants* C_1 *and* C_2 *such that for every* $\mathbf{v} \in (L^2(Q))^2$*, we have*

$$C_1\|\mathbf{v}\|_{0,Q} \le \|\hat{\mathbf{v}}\|_{0,\hat{Q}} \le C_2\|\mathbf{v}\|_{0,Q}. \tag{2.13}$$

If the regular partition satisfies the almost parallelogram condition (2.8), *then for every* $\mathbf{v} \in (H^1(Q))^2$,

$$|\mathbf{v}|_{1,Q} \le C_1 h^{-1}\|\hat{\mathbf{v}}\|_{1,\hat{Q}}, \qquad |\hat{\mathbf{v}}|_{1,\hat{Q}} \le C_2 h\|\mathbf{v}\|_{1,Q}. \tag{2.14}$$

Proof. Since the first result can be derived in a trivial way by virtue of (2.4) and (2.7), we show only the second one. To derive the left inequality, we begin with the identity

$$\frac{\partial \mathbf{v}}{\partial x} = \frac{\partial}{\partial \hat{x}}\left(\frac{\mathcal{J}}{J}\hat{\mathbf{v}}\right)\frac{\partial \hat{x}}{\partial x} + \frac{\partial}{\partial \hat{y}}\left(\frac{\mathcal{J}}{J}\hat{\mathbf{v}}\right)\frac{\partial \hat{y}}{\partial x} \equiv \mathbf{y}_1 + \mathbf{y}_2,$$

where

$$\mathbf{y}_1 = \left[\frac{1}{J}\frac{\partial \mathcal{J}}{\partial \hat{x}}\hat{\mathbf{v}} - \frac{\mathcal{J}}{J^2}\frac{\partial J}{\partial \hat{x}}\hat{\mathbf{v}} + \frac{\mathcal{J}}{J}\frac{\partial \hat{\mathbf{v}}}{\partial \hat{x}}\right]\frac{\partial \hat{x}}{\partial x},$$
$$\mathbf{y}_2 = \left[\frac{1}{J}\frac{\partial \mathcal{J}}{\partial \hat{y}}\hat{\mathbf{v}} - \frac{\mathcal{J}}{J^2}\frac{\partial J}{\partial \hat{y}}\hat{\mathbf{v}} + \frac{\mathcal{J}}{J}\frac{\partial \hat{\mathbf{v}}}{\partial \hat{y}}\right]\frac{\partial \hat{y}}{\partial x}.$$

From now on we shall abbreviate $|J^{-1}|_{\infty,\hat{Q}}$ as $|J^{-1}|_\infty$ and similarly for the corresponding J and $\mathcal{J}$ terms. Now some trivial computations using (2.2)–(2.7) and the equality $(\frac{\partial \hat{x}}{\partial x}, \frac{\partial \hat{y}}{\partial x})^T = \mathcal{J}^{-1}\binom{1}{0}$ show that

$$\begin{aligned}\|\mathbf{y}_1\| + \|\mathbf{y}_2\| &\le C\Big[|J^{-1}|_\infty\|\mathbf{g}\|\,\|\hat{\mathbf{v}}\| + |J^{-1}|_\infty^2\|\mathcal{J}\|_\infty h\|\mathbf{g}\|\,\|\hat{\mathbf{v}}\| \\ &\qquad\qquad + |J^{-1}|_\infty\|\mathcal{J}\|_\infty\|\hat{\nabla}\hat{\mathbf{v}}\|\Big]\,\|\mathcal{J}^{-1}\|_\infty \\ &\le Ch^{-2}(\|\hat{\mathbf{v}}\| + \|\hat{\nabla}\hat{\mathbf{v}}\|),\end{aligned}$$

which gives

$$\int_Q \left\|\frac{\partial \mathbf{v}}{\partial x}\right\|^2 dxdy \le Ch^{-4}\int_{\hat{Q}}(\|\hat{\mathbf{v}}\|^2 + \|\hat{\nabla}\hat{\mathbf{v}}\|^2)|J|\,d\hat{x}d\hat{y} \le Ch^{-2}\|\hat{\mathbf{v}}\|_{1,\hat{Q}}^2.$$

Likewise we obtain

$$\int_Q \left\|\frac{\partial \mathbf{v}}{\partial y}\right\|^2 dxdy \le Ch^{-2}\|\hat{\mathbf{v}}\|_{1,\hat{Q}}^2.$$

This proves the left inequality of (2.14).

A similar technique can be used for the right inequality. From the identity

$$\frac{\partial \hat{\mathbf{v}}}{\partial \hat{x}} = \frac{\partial}{\partial \hat{x}}(J\mathcal{J}^{-1})(\mathbf{v}\circ F) + J\mathcal{J}^{-1}\left(\frac{\partial \mathbf{v}}{\partial x}\frac{\partial x}{\partial \hat{x}} + \frac{\partial \mathbf{v}}{\partial y}\frac{\partial y}{\partial \hat{x}}\right),$$

it follows from (2.2)–(2.7) and the equality $(\frac{\partial x}{\partial \hat{x}}, \frac{\partial y}{\partial \hat{x}})^T = \mathcal{J}\binom{1}{0}$ that

$$\left\|\frac{\partial \hat{\mathbf{v}}}{\partial \hat{x}}\right\| \le \|\mathbf{g}\|\,\|\mathbf{v}\| + |J|_\infty\|\mathcal{J}^{-1}\|_\infty\|\nabla \mathbf{v}\|\,\|\mathcal{J}\|_\infty \le Ch^2(\|\mathbf{v}\| + \|\nabla \mathbf{v}\|),$$

which gives

$$\int_{\hat{Q}} \left\| \frac{\partial \hat{\mathbf{v}}}{\partial \hat{x}} \right\|^2 d\hat{x}d\hat{y} \le Ch^4 \int_Q \frac{1}{|J|}(\|\mathbf{v}\|^2 + \|\nabla \mathbf{v}\|^2)\, dxdy \le Ch^2 \|\mathbf{v}\|^2_{1,Q}.$$

Similarly, the same result holds for $\frac{\partial \hat{\mathbf{v}}}{\partial \hat{y}}$. This completes the proof. ∎

REMARK 2.4. *Note that, unlike in the affine family of triangular elements, the right sides of* (2.14) *involve full-norm rather than seminorm. This is due to the nonlinearity of Jacobian $\mathcal{J}$ in the quadrilateral case.*

3. Trial and test function spaces. We will choose the lowest-order Raviart–Thomas space on $\mathcal{Q}_h$ as the trial function space $\mathbf{V}_h \times W_h$ in which an approximation for $(\mathbf{u}, p)$ is to be sought. The pressure space W_h is simply the space of piecewise constants

$$W_h = \{w \in W : w \text{ is constant over every } Q \in \mathcal{Q}_h\}, \tag{3.1}$$

and the velocity space $\mathbf{V}_h$ is defined to be

$$\mathbf{V}_h = \{\mathbf{v} \in \mathbf{V} : \mathbf{v}|_Q = \mathcal{P}_Q \hat{\mathbf{v}}\ \forall \hat{\mathbf{v}} \in \mathbf{V}_h(\hat{Q}),\ \text{and } \mathbf{v}\cdot\mathbf{n} = 0 \text{ on } \partial\Omega\}, \tag{3.2}$$

where $\mathbf{V}_h(\hat{Q})$ denotes the local space on $\hat{Q}$,

$$\mathbf{V}_h(\hat{Q}) = \{\hat{\mathbf{v}} : \hat{\mathbf{v}} = (a + b\hat{x}, c + d\hat{y}),\ a, b, c, d \in \mathbb{R}\}.$$

For further properties of these spaces, see [26], [28], [30].

Now if $\mathbf{n}_i$ denotes the unit outward normal to the edge e_i of Q, then for $\hat{\mathbf{v}} \in \mathbf{V}_h(\hat{Q})$,

$$|e_i|\mathbf{v}\cdot\mathbf{n}_i = \hat{\mathbf{v}}\cdot\hat{\mathbf{n}}_i, \qquad i = 1, 2, 3, 4, \tag{3.3}$$

where $\hat{\mathbf{n}}_i$ is the unit exterior normal to $\hat{e}_i$. Due to (3.3) every $\mathbf{v} \in \mathbf{V}_h$ has constant normal components on the edges, which can be used as degrees of freedom. We remind the reader that $\mathbf{v}$ is no longer a polynomial on Q unless Q is a parallelogram, and that its divergence is given by

$$\operatorname{div}\mathbf{v}\Big|_Q = \frac{1}{J}\int_Q \operatorname{div}\mathbf{v}\, dxdy, \tag{3.4}$$

which does not belong to W_h.

Denoting the nodal basis for $\mathbf{V}_h(\hat{Q})$ by

$$\hat{\boldsymbol{\phi}}_{x,0} = \begin{pmatrix} 1-\hat{x} \\ 0 \end{pmatrix}, \quad \hat{\boldsymbol{\phi}}_{x,1} = \begin{pmatrix} \hat{x} \\ 0 \end{pmatrix}, \quad \hat{\boldsymbol{\phi}}_{y,0} = \begin{pmatrix} 0 \\ 1-\hat{y} \end{pmatrix}, \quad \hat{\boldsymbol{\phi}}_{y,1} = \begin{pmatrix} 0 \\ \hat{y} \end{pmatrix}, \tag{3.5}$$

we easily see that the nodal basis for $\mathbf{V}_h$ is given by

$$\boldsymbol{\phi}_{i+1/2,j} = \begin{cases} \mathcal{P}_{Q_{i,j}}\hat{\boldsymbol{\phi}}_{x,1} & \text{on } Q_{i,j}, \\ \mathcal{P}_{Q_{i+1,j}}\hat{\boldsymbol{\phi}}_{x,0} & \text{on } Q_{i+1,j}, \\ 0 & \text{elsewhere} \end{cases} \tag{3.6}$$

and

$$
\boldsymbol{\phi}_{i,j+1/2} = \begin{cases} \mathcal{P}_{Q_{i,j}} \hat{\boldsymbol{\phi}}_{y,1} & \text{on } Q_{i,j}, \\ \mathcal{P}_{Q_{i,j+1}} \hat{\boldsymbol{\phi}}_{y,0} & \text{on } Q_{i,j+1}, \\ 0 & \text{elsewhere.} \end{cases} \tag{3.7}
$$

More precisely, $\boldsymbol{\phi}_{i+1/2,j}$ has *unit* flux through the edge $e_{i+1/2,j}$ and has *zero* flux through all the other edges, and similarly for $\boldsymbol{\phi}_{i,j+1/2}$.

The Raviart–Thomas projection $\Pi_h : H^1(\Omega)^2 \to \mathbf{V}_h$ and the "projection" $P_h : W \to W_h$ are defined as in [30]: Let us define $\hat{\Pi}$ and $\hat{P}$ on $\hat{Q}$ to be

$$
\int_{\hat{e}} \hat{\Pi}\hat{\mathbf{v}} \cdot \hat{\mathbf{n}}\, ds = \int_{\hat{e}} \hat{\mathbf{v}} \cdot \hat{\mathbf{n}}\, ds \qquad \forall \text{ edges } \hat{e} \text{ of } \hat{Q}, \tag{3.8}
$$

$$
\int_{\hat{Q}} \hat{P}\hat{p}\, d\hat{x}d\hat{y} = \int_{\hat{Q}} \hat{p}\, d\hat{x}d\hat{y} \tag{3.9}
$$

and then set

$$
\Pi_Q \mathbf{v} = \mathcal{P}_Q(\hat{\Pi}\hat{\mathbf{v}}) \qquad \forall \mathbf{v} \in (H^1(Q))^2, \tag{3.10}
$$

$$
P_Q \phi = (\hat{P}\hat{\phi}) \circ F_Q^{-1} \qquad \forall \phi \in L^2(Q), \tag{3.11}
$$

where $\mathcal{P}_Q \hat{\mathbf{v}} = \mathbf{v}$ and $\hat{\phi} = \phi \circ F_Q$. Finally, we define

$$
\Pi_h \mathbf{v}|_Q = \Pi_Q \mathbf{v}, \qquad P_h \phi|_Q = P_Q \phi. \tag{3.12}
$$

REMARK 3.1. *Strictly speaking, P_Q is not a projection: $P_Q\phi$ equals $\int_Q J^{-1}\phi$, not the average of ϕ over Q.*

Now we state some properties of Π_h and P_h which are necessary to derive error estimates and which are well known in the rectangular cases.

LEMMA 3.1. *The following orthogonality relations hold:*

$$
(\operatorname{div}(\mathbf{u} - \Pi_h \mathbf{u}), w) = 0 \qquad \forall \mathbf{u} \in \mathbf{V}, \quad \forall w \in W_h, \tag{3.13}
$$

$$
(\operatorname{div} \mathbf{v}, \phi - P_h \phi) = 0 \qquad \forall \mathbf{v} \in \mathbf{V}_h, \quad \forall \phi \in W. \tag{3.14}
$$

Proof. The results follow immediately by transferring the relevant integrals back to $\hat{Q}$ and using the definitions of $\hat{\Pi}$ and $\hat{P}$. ◻

LEMMA 3.2. *The following estimates are valid for regular partitions:*

$$
\|\mathbf{u} - \Pi_h \mathbf{u}\|_0 \le Ch\, \|\mathbf{u}\|_1 \qquad \forall \mathbf{u} \in (H^1(\Omega))^2, \tag{3.15}
$$

$$
\|\operatorname{div}(\mathbf{u} - \Pi_h \mathbf{u})\|_0 \le Ch\, \|\operatorname{div} \mathbf{u}\|_1 \qquad \forall \mathbf{u} \in \mathbf{H}^1(\operatorname{div}; \Omega), \tag{3.16}
$$

$$
\|P_h \phi\|_0 \le C\|\phi\|_0 \qquad \forall \phi \in W, \tag{3.17}
$$

$$
\|\phi - P_h \phi\|_0 \le Ch\|\phi\|_1 \qquad \forall \phi \in H^1(\Omega). \tag{3.18}
$$

If the regular partitions also satisfy the almost parallelogram condition (2.8), *then*

$$
\|\phi - P_h \phi\|_{-1} \le Ch^2\|\phi\|_1 \qquad \forall \phi \in H^1(\Omega). \tag{3.19}
$$

Proof. The only nontrivial part is the estimate $\|\phi - P_h\phi\|_{-1} \le Ch^2\|\phi\|_1$. For a given $\psi \in H^1(\Omega)$, we obtain

$$\begin{aligned}\int_Q (\phi - P_h\phi)\psi\, dxdy &= \int_{\hat{Q}} J(\hat{x},\hat{y})(\hat{\phi} - \hat{P}\hat{\phi})\hat{\psi}\, d\hat{x}d\hat{y} \\ &= \int_{\hat{Q}} [J - J(1/2,1/2)](\hat{\phi} - \hat{P}\hat{\phi})\hat{\psi}\, d\hat{x}d\hat{y} \\ &\qquad + \int_{\hat{Q}} J(1/2,1/2)(\hat{\phi} - \hat{P}\hat{\phi})(\hat{\psi} - \hat{P}\hat{\psi})\, d\hat{x}d\hat{y} \\ &= \int_Q J^{-1}[J - J(1/2,1/2)](\phi - P_h\phi)\psi\, dxdy \\ &\qquad + \int_Q J^{-1}J(1/2,1/2)(\phi - P_h\phi)(\psi - P_h\psi)\, dxdy.\end{aligned}$$

Since we have, by (2.7) and (2.9),

$$|J^{-1}[J - J(1/2,1/2)]| \le C|J^{-1}|_\infty(|\beta| + |\gamma|) \le Ch, \tag{3.20}$$

it follows that

$$\begin{aligned}\int_Q (\phi - P_h\phi)\psi\, dxdy &\le Ch\|\phi - P_h\phi\|_{0,Q}\|\psi\|_{0,Q} + C\|\phi - P_h\phi\|_{0,Q}\|\psi - P_h\psi\|_{0,Q} \\ &\le Ch^2\|\phi\|_{1,Q}\|\psi\|_{1,Q}.\end{aligned}$$

Summing over $Q \in \mathcal{Q}_h$ gives

$$(\phi - P_h\phi, \psi) \le Ch^2\|\phi\|_1\|\psi\|_1. \tag{3.21}$$

The proof is completed by dividing both sides by $\|\psi\|_1$ and then taking the supremum with respect to ψ. $\square$

Now we define the test function space $\mathbf{Y}_h \times W_h$, with W_h defined as before. In light of the rectangular cases (see [7]), it is natural to define $\mathbf{Y}_h$ using the transfer map $\hat{\gamma} : \mathbf{V}_h(\hat{Q}) \to \mathbf{Y}_h(\hat{Q})$. Let

$$\hat{\gamma}\hat{u} = \begin{cases} \hat{u}(0,\hat{y}) & \text{on } [0,1/2]\times[0,1], \\ \hat{u}(1,\hat{y}) & \text{on } [1/2,1]\times[0,1] \end{cases}$$

and

$$\hat{\gamma}\hat{v} = \begin{cases} \hat{v}(\hat{x},0) & \text{on } [0,1]\times[0,1/2], \\ \hat{v}(\hat{x},1) & \text{on } [0,1]\times[1/2,1]. \end{cases}$$

Then we set $\hat{\gamma}\hat{\mathbf{v}} = (\hat{\gamma}\hat{u}, \hat{\gamma}\hat{v})$, $\hat{\mathbf{v}} = (\hat{u},\hat{v})^t$.

We define $\gamma\boldsymbol{\phi}_{i+1/2,j}$ by the formula (3.6) with $\hat{\boldsymbol{\phi}}_{x,1}$ and $\hat{\boldsymbol{\phi}}_{x,0}$ replaced by $\hat{\gamma}\hat{\boldsymbol{\phi}}_{x,1}$ and $\hat{\gamma}\hat{\boldsymbol{\phi}}_{x,0}$, resp., i.e.,

$$\gamma_h\boldsymbol{\phi}_{i+1/2,j} = \begin{cases} \mathcal{P}_{Q_{i,j}}\begin{pmatrix}1\\0\end{pmatrix} & \text{on } Q_{i,j}\cap Q_{i+1/2,j}, \\ \mathcal{P}_{Q_{i+1,j}}\begin{pmatrix}1\\0\end{pmatrix} & \text{on } Q_{i+1,j}\cap Q_{i+1/2,j}, \\ 0 & \text{elsewhere.} \end{cases} \tag{3.22}$$

Similarly, $\gamma_h \boldsymbol{\phi}_{i,j+1/2}$ is defined by the formula (3.7) with $\hat{\boldsymbol{\phi}}_{y,1}$ and $\hat{\boldsymbol{\phi}}_{y,0}$ replaced by $\hat{\gamma}\hat{\boldsymbol{\phi}}_{y,1}$ and $\hat{\gamma}\hat{\boldsymbol{\phi}}_{y,0}$, respectively. The velocity test space $\mathbf{Y}_h$ is defined to be spanned by them. The transfer map $\gamma_h : \mathbf{V}_h \to \mathbf{Y}_h$ is then defined in the obvious way: If $\mathbf{v} \in \mathbf{V}_h$ is expressed in the form

$$\mathbf{v} = \sum_{i,j} (v_{i+1/2,j} \boldsymbol{\phi}_{i+1/2,j} + v_{i,j+1/2} \boldsymbol{\phi}_{i,j+1/2}), \tag{3.23}$$

then we set

$$\gamma_h \mathbf{v} = \sum_{i,j} (v_{i+1/2,j} \gamma_h \boldsymbol{\phi}_{i+1/2,j} + v_{i,j+1/2} \gamma_h \boldsymbol{\phi}_{i,j+1/2}). \tag{3.24}$$

4. Mixed covolume methods. To derive the mixed covolume method for the problem (1.1), we begin by integrating the mixed system (1.3) with respect to the test functions in $\mathbf{Y}_h \times W_h$

$$(K^{-1}\mathbf{u} + \nabla p + \boldsymbol{\beta} p, \gamma_h \mathbf{v}) = 0 \qquad \forall \mathbf{v} \in \mathbf{V}_h, \tag{4.1}$$

$$(\operatorname{div} \mathbf{u} + cp, w) = (f, w) \qquad \forall w \in W_h. \tag{4.2}$$

This is equivalent to integrating over the covolumes $Q_{i+1/2,j}, Q_{i,j+1/2}$, and $Q_{i,j}$, since they are the supports of the basis functions of $\mathbf{Y}_h$ and W_h.

As in the rectangular cases (cf. [7]), it is possible to transform the term $(\nabla p, \gamma_h \mathbf{v})$ into an equivalent form which involves no derivatives of p as follows: Take $\mathbf{v} = \boldsymbol{\phi}_{i+1/2,j}$,

$$(\nabla p, \gamma_h \boldsymbol{\phi}_{i+1/2,j}) = \sum_{k=0}^{1} \int_{Q_{i+k,j} \cap Q_{i+1/2,j}} \nabla p \cdot \mathcal{P}_{Q_{i,j}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} dxdy.$$

Using integration by parts we obtain

$$\begin{aligned} \int_{Q_{i,j} \cap Q_{i+1/2,j}} \nabla p \cdot \mathcal{P}_{Q_{i,j}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} dxdy &= \int_0^1 \int_{1/2}^1 \hat{\nabla} \hat{p} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} d\hat{x} d\hat{y} \\ &= \int_0^1 [\hat{p}(1, \hat{y}) - \hat{p}(1/2, \hat{y})] \, d\hat{y} \\ &= \frac{1}{|e_{i+1/2,j}|} \int_{e_{i+1/2,j}} p \, ds - \frac{1}{|e^{(2)}_{i,j}|} \int_{e^{(2)}_{i,j}} p \, ds, \end{aligned}$$

and similarly,

$$\int_{Q_{i+1,j} \cap Q_{i+1/2,j}} \nabla p \cdot \mathcal{P}_{Q_{i+1,j}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} dxdy = \frac{1}{|e^{(2)}_{i+1,j}|} \int_{e^{(2)}_{i+1,j}} p \, ds - \frac{1}{|e_{i+1/2,j}|} \int_{e_{i+1/2,j}} p \, ds.$$

Thus it follows that

$$(\nabla p, \gamma_h \boldsymbol{\phi}_{i+1/2,j}) = \frac{1}{|e^{(2)}_{i+1,j}|} \int_{e^{(2)}_{i+1,j}} p \, ds - \frac{1}{|e^{(2)}_{i,j}|} \int_{e^{(2)}_{i,j}} p \, ds. \tag{4.3}$$

Likewise we obtain

$$(\nabla p, \gamma_h \boldsymbol{\phi}_{i,j+1/2}) = \frac{1}{|e^{(1)}_{i,j+1}|} \int_{e^{(1)}_{i,j+1}} p \, ds - \frac{1}{|e^{(1)}_{i,j}|} \int_{e^{(1)}_{i,j}} p \, ds. \tag{4.4}$$

So for sufficiently smooth p define

$$b(\gamma_h \mathbf{v}, p) = \sum_{i,j} \left[v_{i+1/2,j} \left(\frac{1}{|e^{(2)}_{i+1,j}|} \int_{e^{(2)}_{i+1,j}} p\, ds - \frac{1}{|e^{(2)}_{i,j}|} \int_{e^{(2)}_{i,j}} p\, ds \right) \right. \\ \left. + v_{i,j+1/2} \left(\frac{1}{|e^{(1)}_{i,j+1}|} \int_{e^{(1)}_{i,j+1}} p\, ds - \frac{1}{|e^{(1)}_{i,j}|} \int_{e^{(1)}_{i,j}} p\, ds \right) \right], \tag{4.5}$$

where $\mathbf{v} \in \mathbf{V}_h$ is of the form (3.23). Then (4.1) is equivalent to

$$(K^{-1}\mathbf{u}, \gamma_h \mathbf{v}) + b(\gamma_h \mathbf{v}, p) + (\boldsymbol{\beta} p, \gamma_h \mathbf{v}) = 0 \qquad \forall \mathbf{v} \in \mathbf{V}_h. \tag{4.6}$$

Now the mixed covolume method for the problem (1.1) is to find $(\mathbf{u}_h, p_h)$ in $\mathbf{V}_h \times W_h$ which satisfies

$$(K^{-1}\mathbf{u}_h, \gamma_h \mathbf{v}) + b(\gamma_h \mathbf{v}, p_h) + (\boldsymbol{\beta} p_h, \gamma_h \mathbf{v}) = 0 \qquad \forall \mathbf{v} \in \mathbf{V}_h, \tag{4.7a}$$

$$(\operatorname{div} \mathbf{u}_h, w) + (c p_h, w) = (f, w) \qquad \forall w \in W_h. \tag{4.7b}$$

LEMMA 4.1. *For all $\mathbf{v} \in \mathbf{V}_h$ and $p_h \in W_h$ we have*

$$b(\gamma_h \mathbf{v}, p_h) = -(\operatorname{div} \mathbf{v}, p_h).$$

Proof. Since p_h is piecewise constant, it follows that

$$b(\gamma_h \mathbf{v}, p_h) = \sum_{i,j} [v_{i+1/2,j}(p_{i+1,j} - p_{i,j}) + v_{i,j+1/2}(p_{i,j+1} - p_{i,j})],$$

which gives by summation by parts

$$\begin{aligned} b(\gamma_h \mathbf{v}, p_h) &= -\sum_{i,j} p_{i,j}(v_{i+1/2,j} - v_{i-1/2,j} + v_{i,j+1/2} - v_{i,j-1/2}) \\ &= -\sum_{i,j} p_{i,j} \int_{Q_{i,j}} \operatorname{div} \mathbf{v}\, dxdy = -(\operatorname{div} \mathbf{v}, p_h). \end{aligned}$$

This proves the desired result. ∎

This lemma implies that the mixed covolume method (4.7) can be rewritten as

$$(K^{-1}\mathbf{u}_h, \gamma_h \mathbf{v}) - (\operatorname{div} \mathbf{v}, p_h) + (\boldsymbol{\beta} p_h, \gamma_h \mathbf{v}) = 0 \qquad \forall \mathbf{v} \in \mathbf{V}_h, \tag{4.8a}$$

$$(\operatorname{div} \mathbf{u}_h, w) + (c p_h, w) = (f, w) \qquad \forall w \in W_h. \tag{4.8b}$$

We observe that the above system differs from the standard mixed finite element method in just one respect: the test function is now $\gamma_h \mathbf{v}$ instead of $\mathbf{v}$. However, showing the existence and uniqueness of its solution will take some effort. The proof will be given in Theorem 6.5.

5. Some properties of $\boldsymbol{\gamma_h}$. Before proceeding to the error analysis of the mixed covolume method, we derive some properties of the operator γ_h that will be of crucial importance in establishing error estimates for (4.8). These properties are much harder to establish than their counterparts in the nonoverlapping case (cf. Lemmas 5.2–5.4 of [12]).

LEMMA 5.1. *There exists a positive constant C_1 such that for every $\mathbf{u}_h \in \mathbf{V}_h$, we have*

$$\|\gamma_h \mathbf{u}_h\|_0 \le C_1 \|\mathbf{u}_h\|_0. \tag{5.1}$$

Let the coefficient $K \in Lip(Q)$; i.e., each entry of K is Lipschitz on $Q \in \mathcal{Q}_h$. Then there exists a positive constant C_2 such that for sufficiently small h

$$(K^{-1}\mathbf{u}_h, \gamma_h \mathbf{u}_h) \ge C_2 \|\mathbf{u}_h\|_0^2. \tag{5.2}$$

Proof. Define $\hat{\mathbf{u}}_h \in \mathbf{V}_h(\hat{Q})$ by $\mathcal{P}_Q \hat{\mathbf{u}}_h = \mathbf{u}_h|_Q$. By definition of γ_h it is easy to see that $\mathcal{P}_Q(\hat{\gamma}\hat{\mathbf{u}}_h) = \gamma_h \mathbf{u}_h$ on Q. Hence using (2.13) we have

$$\|\gamma_h \mathbf{u}_h\|_{0,Q} = \|\mathcal{P}_Q(\hat{\gamma}\hat{\mathbf{u}}_h)\|_{0,Q} \le C' \|\hat{\gamma}\hat{\mathbf{u}}_h\|_{0,\hat{Q}} \le C'' \|\hat{\mathbf{u}}_h\|_{0,\hat{Q}} \le C \|\mathbf{u}_h\|_{0,Q},$$

which gives (5.1) by summing over $Q \in \mathcal{Q}_h$.

To prove (5.2), let $B = J^{-1}\mathcal{J}^T(K^{-1} \circ F)\mathcal{J}$ and $\bar{B}$ be its average over $\hat{Q}$. Let $\mathbf{x} = (x, y)$ and $\hat{\mathbf{x}} = (\hat{x}, \hat{y})$ and observe the following relations with explicit arguments:

$$K^{-1}(\mathbf{x})\mathbf{u}_h(\mathbf{x}) = K^{-1}(F(\hat{\mathbf{x}}))J^{-1}\mathcal{J}\hat{\mathbf{u}}_h(\hat{\mathbf{x}}), \qquad \gamma_h \mathbf{u}_h(\mathbf{x}) = J^{-1}\mathcal{J}\hat{\gamma}_h \hat{\mathbf{u}}_h(\hat{\mathbf{x}}).$$

Now use them to arrive at

$$\begin{aligned}
\int_Q K^{-1}\mathbf{u}_h \cdot \gamma_h \mathbf{u}_h \, dxdy &= \int_{\hat{Q}} B\hat{\mathbf{u}}_h \cdot \hat{\gamma}\hat{\mathbf{u}}_h \, d\hat{x}d\hat{y} \\
&= \int_{\hat{Q}} \bar{B}\hat{\mathbf{u}}_h \cdot \hat{\gamma}\hat{\mathbf{u}}_h \, d\hat{x}d\hat{y} + \int_{\hat{Q}} (B - \bar{B})\hat{\mathbf{u}}_h \cdot \hat{\gamma}\hat{\mathbf{u}}_h \, d\hat{x}d\hat{y}.
\end{aligned}$$

It is easy to verify that $\bar{B}$ is uniformly bounded below (since $\|B^{-1}\| \le C$ uniformly in Q). This fact, together with (2.13) and Lemma A.3 in the appendix, implies that

$$\int_{\hat{Q}} \bar{B}\hat{\mathbf{u}}_h \cdot \hat{\gamma}\hat{\mathbf{u}}_h \, d\hat{x}d\hat{y} \ge C \int_{\hat{Q}} \bar{B}\hat{\mathbf{u}}_h \cdot \hat{\mathbf{u}}_h \, d\hat{x}d\hat{y} \ge C\|\hat{\mathbf{u}}_h\|_{0,\hat{Q}}^2 \ge C\|\mathbf{u}_h\|_{0,Q}^2.$$

Next we show that $\|B - \bar{B}\| \le Ch$. To this end, let $\mathcal{P}$ denote the matrix $J^{-1}\mathcal{J}$. We decompose $B - \bar{B}$ as follows. For $\hat{\mathbf{x}}_0 \in \hat{Q}$, one has, using (2.6) in the second equality below, that

$$\begin{aligned}
(B - \bar{B})(\hat{\mathbf{x}}_0) &= [J\mathcal{P}^T(K^{-1} \circ F)\mathcal{P}](\hat{\mathbf{x}}_0) - \int_{\hat{Q}} J(\hat{x}, \hat{y})\mathcal{P}^T(\hat{x}, \hat{y})(K^{-1} \circ F)(\hat{x}, \hat{y})\mathcal{P}(\hat{x}, \hat{y}) \, d\hat{x}d\hat{y} \\
&= [J(\hat{\mathbf{x}}_0) - J(1/2, 1/2)] \left[\mathcal{P}^T(K^{-1} \circ F)\mathcal{P}\right](\hat{\mathbf{x}}_0) \\
&\quad + \int_Q \left([\mathcal{P}^T(K^{-1} \circ F)\mathcal{P}](\hat{\mathbf{x}}_0) - \mathcal{P}^T(x', y')K^{-1}(x', y')\mathcal{P}(x', y')\right) dx'dy' \\
&= [J(\hat{\mathbf{x}}_0) - J(1/2, 1/2)] \left[\mathcal{P}^T(K^{-1} \circ F)\mathcal{P}\right](\hat{\mathbf{x}}_0) \\
&\qquad + \int_Q [\mathcal{P}(\hat{\mathbf{x}}_0) - \mathcal{P}(x', y')]^T K^{-1}(\mathbf{x}_0)\mathcal{P}(\hat{\mathbf{x}}_0) \, dx'dy' \\
&\qquad + \int_Q \mathcal{P}^T(x', y')[K^{-1}(\mathbf{x}_0) - K^{-1}(x', y')]\mathcal{P}(\hat{\mathbf{x}}_0) \, dx'dy' \\
&\qquad + \int_Q \mathcal{P}^T(x', y')K^{-1}(x', y')[\mathcal{P}(\hat{\mathbf{x}}_0) - \mathcal{P}(x', y')] \, dx'dy' \\
&\equiv I_1 + I_2 + I_3 + I_4,
\end{aligned}$$

where we have used $\mathcal{P}(x', y')$ as an abbreviation for $\mathcal{P} \circ F^{-1}(x', y')$.

We estimate the four terms one by one. For simplicity we will not write out the dependency on $\hat{\mathbf{x}}_0$. The term I_1 can be estimated in a trivial way:

$$\|I_1\| \le C(|\beta| + |\gamma|)|J^{-1}|_\infty^2 \|\mathcal{J}\|_\infty^2 \le Ch. \tag{5.3}$$

To estimate I_2 we note the identity

$$\mathcal{P} - \mathcal{P}(x', y') = \left[\frac{1}{J} - \frac{1}{J(x', y')}\right]\mathcal{J} + \frac{1}{J(x', y')}[\mathcal{J} - \mathcal{J}(x', y')],$$

from which it follows that

$$\|\mathcal{P} - \mathcal{P}(x', y')\| \le C.$$

Thus we obtain

$$\|I_2\| + \|I_4\| \le C|Q|\,|J^{-1}|_\infty \|\mathcal{J}\|_\infty \le Ch. \tag{5.4}$$

Finally, we have

$$\|I_3\| \le Ch|Q|\,|J^{-1}|_\infty^2 \|\mathcal{J}\|_\infty^2 \le Ch. \tag{5.5}$$

Combining (5.3)–(5.5), we obtain $\|B - \bar{B}\| \le Ch$.

Now we complete the proof by observing that

$$(K^{-1}\mathbf{u}_h, \gamma_h \mathbf{u}_h) \ge C'\|\mathbf{u}_h\|_0^2 - C''h\|\mathbf{u}_h\|_0^2 \ge C\|\mathbf{u}_h\|_0^2,$$

provided h is sufficiently small. ∎

LEMMA 5.2. *There exists a positive constant C such that*

$$\|\mathbf{u} - \gamma_h \Pi_h \mathbf{u}\|_0 \le Ch\|\mathbf{u}\|_1 \qquad \forall \mathbf{u} \in (H^1(\Omega))^2. \tag{5.6}$$

Proof. First we note the fundamental relation

$$\gamma_h \Pi_h \mathbf{u}|_Q = \mathcal{P}_Q(\hat{\gamma}\hat{\Pi}\hat{\mathbf{u}}), \qquad \mathcal{P}_Q \hat{\mathbf{u}} = \mathbf{u}|_Q. \tag{5.7}$$

Then it holds that $\mathbf{u} - \gamma_h \Pi_h \mathbf{u} = \mathcal{P}_Q(\hat{\mathbf{u}} - \hat{\gamma}\hat{\Pi}\hat{\mathbf{u}})$, which gives by (2.13), (2.14),

$$\|\mathbf{u} - \gamma_h \Pi_h \mathbf{u}\|_{0,Q} \le C\|\hat{\mathbf{u}} - \hat{\gamma}\hat{\Pi}\hat{\mathbf{u}}\|_{0,\hat{Q}} \le C|\hat{\mathbf{u}}|_{1,\hat{Q}} \le Ch\|\mathbf{u}\|_{1,Q},$$

where we have applied the Bramble–Hilbert lemma for $\hat{\mathbf{u}} - \hat{\gamma}\hat{\Pi}\hat{\mathbf{u}}$. The proof is completed by summing over $Q \in \mathcal{Q}_h$. ∎

LEMMA 5.3. *Suppose that $\mathbf{u} \in (H^1(\Omega))^2$, $p \in H^1(\Omega)$, and $\boldsymbol{\beta} \in (W^{1,\infty}(\Omega))^2$. Then, for every $\mathbf{v} \in \mathbf{V}_h$ we have*

$$|(K^{-1}\mathbf{u} + \boldsymbol{\beta}p, \mathbf{v} - \gamma_h \mathbf{v})| \le Ch(\|\mathbf{u}\|_1 + \|p\|_1)\|\mathbf{v}\|_0. \tag{5.8}$$

Proof. Let $B = J^{-1}\mathcal{J}^T(K^{-1} \circ F)\mathcal{J}$ and $\bar{B}$ be its average over $\hat{Q}$. Then it is easy to check that for any constant vector $\mathbf{c}$

$$\int_{\hat{Q}} \mathbf{c} \cdot (\hat{\mathbf{v}} - \hat{\gamma}\hat{\mathbf{v}})\, d\hat{x} d\hat{y} = 0. \tag{5.9}$$

So applying the above relation for $\bar{B}\mathbf{c}$ we have

$$\begin{aligned}\int_Q K^{-1}\mathbf{u}\cdot(\mathbf{v}-\gamma_h\mathbf{v})\,dxdy &= \int_{\hat{Q}} B\hat{\mathbf{u}}\cdot(\hat{\mathbf{v}}-\hat{\gamma}\hat{\mathbf{v}})\,d\hat{x}d\hat{y}\\ &= \int_{\hat{Q}} B(\hat{\mathbf{u}}-\mathbf{c})\cdot(\hat{\mathbf{v}}-\hat{\gamma}\hat{\mathbf{v}})\,d\hat{x}d\hat{y}\\ &\qquad + \int_{\hat{Q}} (B-\bar{B})\mathbf{c}\cdot(\hat{\mathbf{v}}-\hat{\gamma}\hat{\mathbf{v}})\,d\hat{x}d\hat{y}.\end{aligned}$$

Now take the constant $\mathbf{c}$ to be the constant L^2 projection of $\hat{\mathbf{u}}$, which satisfies $||\mathbf{c}-\hat{\mathbf{u}}\|_0 \le C|\hat{\mathbf{u}}|_{1,\hat{Q}}$. By virtue of $\|B-\bar{B}\| \le Ch$ and (2.14) we obtain

$$\left|\int_Q K^{-1}\mathbf{u}\cdot(\mathbf{v}-\gamma_h\mathbf{v})\,dxdy\right| \le Ch\|\mathbf{u}\|_{1,Q}\|\mathbf{v}\|_{0,Q},$$

which gives $|(K^{-1}\mathbf{u},\mathbf{v}-\gamma_h\mathbf{v})| \le Ch\|\mathbf{u}\|_1\|\mathbf{v}\|_0$.

The other term $(\boldsymbol{\beta}p,\mathbf{v}-\gamma_h\mathbf{v})$ can be treated analogously, yielding

$$|(\boldsymbol{\beta}p,\mathbf{v}-\gamma_h\mathbf{v})| \le Ch\|p\|_1\|\mathbf{v}\|_0.$$

This completes the proof. ∎

6. Error estimates. Let $\boldsymbol{\xi}=\mathbf{u}-\mathbf{u}_h$ and $\tau = P_hp-p_h$. Subtracting (4.8) from (1.7) and using the orthogonality relation (3.14), we obtain the error equations

$$(K^{-1}\boldsymbol{\xi},\gamma_h\mathbf{v}) - (\operatorname{div}\mathbf{v},\tau) + (\boldsymbol{\beta}\tau,\gamma_h\mathbf{v}) = l(\mathbf{v}) \qquad \forall \mathbf{v}\in V_h, \tag{6.1a}$$

$$(\operatorname{div}\boldsymbol{\xi},w) + (c\,\tau,w) = m(w) \qquad \forall w \in W_h, \tag{6.1b}$$

where l and m are the linear functionals on $\mathbf{V}_h$ and W_h, resp., given by

$$l(\mathbf{v}) = (\boldsymbol{\beta}(P_hp-p),\gamma_h\mathbf{v}) + (K^{-1}\mathbf{u}+\boldsymbol{\beta}p,\mathbf{v}-\gamma_h\mathbf{v}), \tag{6.2}$$

$$m(w) = (c(P_hp-p),w). \tag{6.3}$$

We point out that these error equations differ from those of the standard mixed finite element method in only two respects: the test function is now $\gamma_h\mathbf{v}$, and we have an additional term $(K^{-1}\mathbf{u}+\boldsymbol{\beta}p,\mathbf{v}-\gamma_h\mathbf{v})$. This makes it possible to analyze the mixed covolume method the same way as for the standard mixed finite element method.

We say that the domain Ω is 2-regular if the adjoint Dirichlet problem of (1.1) is uniquely solvable for every right-hand side function $f\in L^2(\Omega)$ and the solution satisfies the elliptic regularity condition $\|p\|_2 \le C\|f\|_0$. Now let us apply the duality argument of Douglas and Roberts [13] to the system (6.1) to obtain the following theorem.

THEOREM 6.1. *Let Ω be 2-regular and let K satisfy* (1.2), $K\in W^{1,\infty}, \mathbf{b}\in (W^{1,\infty})^2, c\in W^{1,\infty}, \mathbf{u}\in(H^1)^2, p\in H^1$. *Suppose $(\boldsymbol{\xi},\tau)\in \mathbf{V}\times W_h$ satisfies the system* (6.1). *Then, for sufficiently small h we have*

$$\|\tau\|_0 \le C\big[h\|\boldsymbol{\xi}\|_0 + h\|\operatorname{div}\boldsymbol{\xi}\|_0 + h^2(\|\mathbf{u}\|_1+\|p\|_1)\big]. \tag{6.4}$$

Proof. Given $\psi\in L^2(\Omega)$, let $\phi\in H^1(\Omega)$ be the solution of the adjoint problem

$$\begin{cases} -\operatorname{div}(K\nabla\phi)+\mathbf{b}\cdot\nabla\phi + c\,\phi = \psi & \text{in } \Omega,\\ \phi = 0 & \text{on } \partial\Omega,\end{cases}$$

and let $\boldsymbol{\zeta} = K\nabla\phi$. By assumption, the elliptic regularity $\|\boldsymbol{\zeta}\|_1 + \|\phi\|_1 \le C\|\psi\|_0$ holds for this problem. Then we obtain by (3.13)

$$\begin{aligned}(\tau,\psi) &= (\tau, -\operatorname{div}\boldsymbol{\zeta} + \boldsymbol{\beta}\cdot\boldsymbol{\zeta} + c\,\phi) = (\tau, -\operatorname{div}(\Pi_h\boldsymbol{\zeta}) + \boldsymbol{\beta}\cdot\boldsymbol{\zeta} + c\,\phi)\\ &= l(\Pi_h\boldsymbol{\zeta}) - (K^{-1}\boldsymbol{\xi}, \gamma_h\Pi_h\boldsymbol{\zeta}) + (\boldsymbol{\beta}\tau, \boldsymbol{\zeta} - \gamma_h\Pi_h\boldsymbol{\zeta}) + (c\,\tau,\phi)\\ &= l(\Pi_h\boldsymbol{\zeta}) + (K^{-1}\boldsymbol{\xi} + \boldsymbol{\beta}\tau, \boldsymbol{\zeta} - \gamma_h\Pi_h\boldsymbol{\zeta}) - (K^{-1}\boldsymbol{\xi}, \boldsymbol{\zeta}) + (c\,\tau,\phi)\\ &= l(\Pi_h\boldsymbol{\zeta}) + (K^{-1}\boldsymbol{\xi} + \boldsymbol{\beta}\tau, \boldsymbol{\zeta} - \gamma_h\Pi_h\boldsymbol{\zeta}) + (\operatorname{div}\boldsymbol{\xi}, \phi) + (c\,\tau,\phi)\\ &= l(\Pi_h\boldsymbol{\zeta}) + (K^{-1}\boldsymbol{\xi} + \boldsymbol{\beta}\tau, \boldsymbol{\zeta} - \gamma_h\Pi_h\boldsymbol{\zeta}) + (\operatorname{div}\boldsymbol{\xi} + c\,\tau, \phi - P_h\phi) + m(P_h\phi).\end{aligned}$$

Estimates for the second and third terms can be derived immediately from (3.18), (5.6):

$$(6.5)\qquad \begin{aligned}|(K^{-1}\boldsymbol{\xi} + \boldsymbol{\beta}\tau, \boldsymbol{\zeta} - \gamma_h\Pi_h\boldsymbol{\zeta})| &\le C'(\|\boldsymbol{\xi}\|_0 + \|\tau\|_0)\|\boldsymbol{\zeta} - \gamma_h\Pi_h\boldsymbol{\zeta}\|_0\\ &\le Ch(\|\boldsymbol{\xi}\|_0 + \|\tau\|_0)\|\boldsymbol{\zeta}\|_1,\end{aligned}$$

$$(6.6)\qquad \begin{aligned}|(\operatorname{div}\boldsymbol{\xi} + c\,\tau, \phi - P_h\phi)| &\le C'(\|\operatorname{div}\boldsymbol{\xi}\|_0 + \|\tau\|_0)\|\phi - P_h\phi\|_0\\ &\le Ch(\|\operatorname{div}\boldsymbol{\xi}\|_0 + \|\tau\|_0)\|\phi\|_1.\end{aligned}$$

The other terms $l(\Pi_h\boldsymbol{\zeta})$ and $m(P_h\phi)$ will be estimated in a series of lemmas below. Specifically, we will show that

$$(6.7)\qquad |l(\Pi_h\boldsymbol{\zeta})| \le Ch^2(\|\mathbf{u}\|_1 + \|p\|_1)\|\boldsymbol{\zeta}\|_1, \qquad |m(P_h\phi)| \le Ch^2\|p\|_1\|\phi\|_0.$$

Now combining (6.5)–(6.7) and applying the elliptic regularity, we obtain

$$|(\tau,\psi)| \le C\big[h\|\tau\|_0 + h\|\boldsymbol{\xi}\|_0 + h\|\operatorname{div}\boldsymbol{\xi}\|_0 + h^2(\|\mathbf{u}\|_1 + \|p\|_1)\big]\,\|\psi\|_0.$$

By dividing both sides by $\|\psi\|_0$ and taking the supremum over ψ, it follows that

$$\|\tau\|_0 \le C\big[h\|\tau\|_0 + h\|\boldsymbol{\xi}\|_0 + h\|\operatorname{div}\boldsymbol{\xi}\|_0 + h^2(\|\mathbf{u}\|_1 + \|p\|_1)\big],$$

which gives the desired result if we absorb the term $h\|\tau\|_0$ into $\|\tau\|_0$, provided h is small enough. ∎

To complete the proof of Theorem 6.1 it remains to show the following.

LEMMA 6.2. *Suppose that* $\mathbf{u} \in (H^1(\Omega))^2$, $p \in H^1(\Omega)$, *and* $\boldsymbol{\beta} \in (W^{1,\infty}(\Omega))^2$. *Then, for every* $\boldsymbol{\zeta} \in (H^1(\Omega))^2$ *we have*

$$|(K^{-1}\mathbf{u} + \boldsymbol{\beta}p, \Pi_h\boldsymbol{\zeta} - \gamma_h\Pi_h\boldsymbol{\zeta})| \le Ch^2(\|\mathbf{u}\|_1 + \|p\|_1)\|\boldsymbol{\zeta}\|_1.$$

Proof. Set $\mathbf{v} = \Pi_h\boldsymbol{\zeta}$ in the *proof* of Lemma 5.3. Instead of estimating in terms of $\|\mathbf{v}\|_0$, estimate in terms of $\|\mathbf{v}\|_1$, using the arguments in the proof of Lemma 5.2 to get another power of h. ∎

LEMMA 6.3. *Suppose that* $\boldsymbol{\beta} \in (W^{1,\infty}(\Omega))^2$ *and* $p \in H^1(\Omega)$. *Then, for every* $\boldsymbol{\zeta} \in (H^1(\Omega))^2$ *we have*

$$|(\boldsymbol{\beta}(P_hp - p), \gamma_h\Pi_h\boldsymbol{\zeta})| \le Ch^2\,\|p\|_1\|\boldsymbol{\zeta}\|_1.$$

Proof. Observing that $\boldsymbol{\beta}\cdot\boldsymbol{\zeta} \in H^1(\Omega)$, we obtain by (3.19) that

$$\begin{aligned}|(\boldsymbol{\beta}(P_hp - p), \gamma_h\Pi_h\boldsymbol{\zeta})| &= |(\boldsymbol{\beta}(P_hp - p), \gamma_h\Pi_h\boldsymbol{\zeta} - \boldsymbol{\zeta}) + (\boldsymbol{\beta}(P_hp - p), \boldsymbol{\zeta})|\\ &\le C\|P_hp - p\|_0\|\gamma_h\Pi_h\boldsymbol{\zeta} - \boldsymbol{\zeta}\|_0 + C\|P_hp - p\|_{-1}\|\boldsymbol{\zeta}\|_1\\ &\le Ch^2\|p\|_1\|\boldsymbol{\zeta}\|_1,\end{aligned}$$

which gives the desired result. ∎

LEMMA 6.4. *Suppose that $c \in W^{1,\infty}(\Omega)$ and $p \in H^1(\Omega)$. Then, for every $w \in W_h$ we have*

$$|(c(P_h p - p), w)| \le Ch^2 \|p\|_1 \|w\|_0.$$

Proof. We observe the identity

$$\begin{aligned}
\int_Q c(P_h p - p) w \, dxdy &= \int_{\hat{Q}} \hat{c}(\hat{P}\hat{p} - \hat{p}) w J(\hat{x}, \hat{y}) \, d\hat{x} d\hat{y} \\
&= \int_{\hat{Q}} \hat{c}(\hat{P}\hat{p} - \hat{p}) w [J - J(1/2, 1/2)] \, d\hat{x} d\hat{y} \\
&\qquad + \int_{\hat{Q}} (\hat{c} - \hat{P}\hat{c})(\hat{P}\hat{p} - \hat{p}) w J(1/2, 1/2) \, d\hat{x} d\hat{y} \\
&= \int_Q c(P_h p - p) w J^{-1} [J - J(1/2, 1/2)] \, dxdy \\
&\qquad + \int_Q (c - P_h c)(P_h p - p) w J^{-1} J(1/2, 1/2) \, dxdy.
\end{aligned}$$

Since we have $|J^{-1}[J - J(1/2, 1/2)]| \le Ch$ by (3.20), it follows that

$$\left| \int_Q c(P_h p - p) w \, dxdy \right| \le Ch \|p - P_h p\|_{0,Q} \|w\|_{0,Q} \le Ch^2 \|p\|_{1,Q} \|w\|_{0,Q}.$$

The proof is completed by summing over $Q \in \mathcal{Q}_h$. ∎

Now we estimate $\|\boldsymbol{\xi}\|_0$ and $\|\operatorname{div} \boldsymbol{\xi}\|_0$. First let $\boldsymbol{\sigma} = \Pi_h \mathbf{u} - \mathbf{u}_h$ and rewrite the error equations (6.1) as

$$(K^{-1}\boldsymbol{\sigma}, \gamma_h \mathbf{v}) - (\operatorname{div} \mathbf{v}, \tau) + (\boldsymbol{\beta}\tau, \gamma_h \mathbf{v}) = l(\mathbf{v}) + (K^{-1}(\Pi_h \mathbf{u} - \mathbf{u}), \gamma_h \mathbf{v}) \ \forall \mathbf{v} \in \mathbf{V}_h, \tag{6.8a}$$

$$(\operatorname{div} \boldsymbol{\sigma}, w) + (c\,\tau, w) = m(w) \qquad \forall w \in W_h. \tag{6.8b}$$

Taking $w = P_h(\operatorname{div} \boldsymbol{\sigma})$ and using (3.14), (3.17), we obtain

$$\begin{aligned}
\|\operatorname{div} \boldsymbol{\sigma}\|_0^2 = (\operatorname{div} \boldsymbol{\sigma}, \operatorname{div} \boldsymbol{\sigma}) &= (\operatorname{div} \boldsymbol{\sigma}, P_h(\operatorname{div} \boldsymbol{\sigma})) \\
&\le C(\|\tau\|_0 + \|m\|_0) \|P_h(\operatorname{div} \boldsymbol{\sigma})\|_0 \\
&\le C(\|\tau\|_0 + \|m\|_0) \|\operatorname{div} \boldsymbol{\sigma}\|_0,
\end{aligned}$$

which yields

$$\|\operatorname{div} \boldsymbol{\sigma}\|_0 \le C(\|\tau\|_0 + \|m\|_0) \le C(\|\tau\|_0 + h^2 \|p\|_1). \tag{6.9}$$

Next we take $\mathbf{v} = \boldsymbol{\sigma}$ and use Lemma 5.3 and (3.18) on $\|l\|_0$ to obtain

$$\begin{aligned}
|(K^{-1}\boldsymbol{\sigma}, \gamma_h \boldsymbol{\sigma})| &\le C(\|\operatorname{div} \boldsymbol{\sigma}\|_0 \|\tau\|_0 + \|\tau\|_0 \|\boldsymbol{\sigma}\|_0 + \|l\|_0 \|\boldsymbol{\sigma}\|_0 + \|\Pi_h \mathbf{u} - \mathbf{u}\|_0 \|\boldsymbol{\sigma}\|_0) \\
&\le C(\|\tau\|_0^2 + h\|p\|_1 \|\tau\|_0 + \|\tau\|_0 \|\boldsymbol{\sigma}\|_0 + h\|p\|_1 \|\boldsymbol{\sigma}\|_0 + h\|\mathbf{u}\|_1 \|\boldsymbol{\sigma}\|_0),
\end{aligned}$$

from which it follows by (5.2) that

$$\|\boldsymbol{\sigma}\|_0 \le C(\|\tau\|_0 + h\|\mathbf{u}\|_1 + h\|p\|_1). \tag{6.10}$$

Immediate consequences of (6.9), (6.10) are

$$\|\boldsymbol{\xi}\|_0 \le \|\mathbf{u} - \Pi_h \mathbf{u}\|_0 + \|\boldsymbol{\sigma}\|_0 \le C(\|\tau\|_0 + h\|\mathbf{u}\|_1 + h\|p\|_1), \tag{6.11}$$

and for $s = 0, 1$

$$\begin{aligned} \|\operatorname{div} \boldsymbol{\xi}\|_0 &\le \|\operatorname{div} \mathbf{u} - \operatorname{div}(\Pi_h \mathbf{u})\|_0 + \|\operatorname{div} \boldsymbol{\sigma}\|_0 \\ &\le C(h^s \|\operatorname{div} \mathbf{u}\|_s + \|\tau\|_0 + h\|\mathbf{u}\|_1 + h\|p\|_1), \end{aligned} \tag{6.12}$$

which, when substituted into (6.4), give for sufficiently small h

$$\|\tau\|_0 \le C\big[h^2(\|\mathbf{u}\|_1 + \|p\|_1) + h^{1+s}\|\operatorname{div} \mathbf{u}\|_s\big]. \tag{6.13}$$

Substituting this back into (6.11), (6.12) yields

$$\|\boldsymbol{\xi}\|_0 \le Ch(\|\mathbf{u}\|_1 + \|p\|_1), \tag{6.14}$$

$$\|\operatorname{div} \boldsymbol{\xi}\|_0 \le Ch(\|\mathbf{u}\|_1 + \|p\|_1 + \|\operatorname{div} \mathbf{u}\|_1). \tag{6.15}$$

Now we are ready to present the following theorem.

THEOREM 6.5. *Let there be given a 2-regular domain Ω for* (1.1) *and a regular family of quadrilateral partitions satisfying the almost parallelogram condition* (2.8). *Furthermore, let K satisfy* (1.2), $K \in W^{1,\infty}, \mathbf{b} \in (W^{1,\infty})^2$, *and* $c \in W^{1,\infty}$. *Suppose that $(\boldsymbol{\xi}, \tau) \in \mathbf{V} \times W_h$ satisfies the system* (6.1). *Assume that $(\mathbf{u}, p)$, the solution of the mixed formulation of* (1.1), *exists so that $\mathbf{u} \in H^1(\Omega)^2, \operatorname{div} \mathbf{u} \in H^1(\Omega)$, and $p \in H^1(\Omega)$. Then, for sufficiently small h, there exists a unique solution $(\mathbf{u}_h, p_h)$ in $\mathbf{V}_h \times W_h$ of the system* (4.8) *such that*

$$\|\mathbf{u} - \mathbf{u}_h\|_0 + \|p - p_h\|_0 \le Ch(\|\mathbf{u}\|_1 + \|p\|_1), \tag{6.16}$$

$$\|\operatorname{div}(\mathbf{u} - \mathbf{u}_h)\|_0 \le Ch(\|\mathbf{u}\|_1 + \|p\|_1 + \|\operatorname{div} \mathbf{u}\|_1). \tag{6.17}$$

Moreover, the following superconvergence result holds for the pressure variable:

$$\|P_h p - p_h\|_0 \le Ch^2(\|\mathbf{u}\|_1 + \|p\|_1 + \|\operatorname{div} \mathbf{u}\|_1). \tag{6.18}$$

Proof. It is easy to check that γ_h is one-to-one, since $\gamma_h \mathbf{v} = 0$ implies $\mathbf{v} = 0$. Hence it suffices to show that the homogeneous system has only the trivial solution. Suppose that $(\tilde{\mathbf{u}}_h, \tilde{p}_h)$ satisfies

$$\begin{aligned} (K^{-1}\tilde{\mathbf{u}}_h, \gamma_h \mathbf{v}) - (\operatorname{div} \mathbf{v}, \tilde{p}_h) + (\boldsymbol{\beta}\tilde{p}_h, \gamma_h \mathbf{v}) = 0 \qquad &\forall \mathbf{v} \in \mathbf{V}_h, \\ (\operatorname{div} \tilde{\mathbf{u}}_h, w) + (c\tilde{p}_h, w) = 0 \qquad &\forall w \in W_h. \end{aligned}$$

As before, we take $w = P_h(\operatorname{div} \tilde{\mathbf{u}}_h)$ to obtain

$$\|\operatorname{div} \tilde{\mathbf{u}}_h\|_0 \le C\|\tilde{p}_h\|_0.$$

Theorem 6.1, with $(\tilde{\mathbf{u}}_h, \tilde{p}_h)$ playing the role of (ξ, τ), and with $\mathbf{u} = p = 0$ by uniqueness for (1.7), implies that

$$\|\tilde{p}_h\|_0 \le Ch(\|\tilde{\mathbf{u}}_h\|_0 + \|\tilde{p}_h\|_0),$$

so that for sufficiently small h we have

$$\|\tilde{p}_h\|_0 \le Ch\|\tilde{\mathbf{u}}_h\|_0.$$

Finally, if we take $\mathbf{v} = \tilde{\mathbf{u}}_h$, then it follows by (5.2) that

$$\|\tilde{\mathbf{u}}_h\|_0 \le C\|\tilde{p}_h\|_0 \le Ch\|\tilde{\mathbf{u}}_h\|_0,$$

which yields $\tilde{\mathbf{u}}_h = \tilde{p}_h = 0$, provided h is sufficiently small. The estimates (6.16)–(6.18) now follow immediately by the discussion preceding the theorem. ∎

COROLLARY 6.6. *For $2 < q \le \infty$ the following optimal L^q-error estimates hold for the pressure variable:*

$$\|p - p_h\|_{0,q} \le Ch(\|p\|_{1,q} + \|\mathbf{u}\|_1 + \|p\|_1 + \|\operatorname{div}\mathbf{u}\|_1). \tag{6.19}$$

Proof. The desired results follow immediately by using the inverse inequalities $\|\tau\|_{0,q} \le Ch^{-(q-2)/q}\|\tau\|_0$ for $2 < q < \infty$ and $\|\tau\|_{0,\infty} \le Ch^{-1}\|\tau\|_0$. We refer to [13] for details. ∎

REMARK 6.1. *When the problem* (1.1) *is symmetric, i.e.,* $\mathbf{b} \equiv 0$, *we may obtain the following error estimate through the conventional* inf-sup *condition:*

$$\|p - p_h\|_0 + \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{H}(\mathrm{div};\Omega)} \le Ch(\|\mathbf{u}\|_1 + \|p\|_1 + \|\operatorname{div}\mathbf{u}\|_1) \tag{6.20}$$

without assuming the elliptic regularity $\|\boldsymbol{\zeta}\|_1 + \|\phi\|_1 \le C\|\psi\|_0$. *This approach was taken by Chou and Kwak in* [7].

REMARK 6.2. *Now since the assumption of $K \in W^{1,\infty}$ was used only in the proof of Theorem* 6.1 *to handle* $\|\boldsymbol{\zeta}\|_1$, *this raises the question of whether one can avoid that global smoothness assumption. Any result that avoids that assumption is much more useful in practice. In applications of Darcy flow,* $\mathbf{b} \equiv \mathbf{0}, c \equiv 0$ *generally holds. The answer to our question is yes, thanks to the abstract framework contained in* [7]. *For discontinuous K it is hardly the case that* $\|\operatorname{div}\mathbf{u}\|_1$ *is finite, and hence the estimate* (6.20) *is unrealistic. In the theorem below we assume only* $\mathbf{u} \in H^1(\Omega)^2$, *which is a reasonable assumption. For instance, in the Darcy flow case with rough K, it may occur that the pressure* $p \in H^1(\Omega), p \notin H^2(\Omega)$, *while* $\mathbf{u} \in H^1(\Omega)^2$.

THEOREM 6.7. *Let there be given a regular family of quadrilateral partitions satisfying the almost parallelogram condition* (2.8). *Furthermore, let K satisfy* (1.2), $K \in Lip(Q)\ \forall Q \in \mathcal{Q}_h, \mathbf{b} \equiv \mathbf{0}$, *and* $c \equiv 0$. *Assume that* $(\mathbf{u}, p)$, *the solution of the mixed formulation of* (1.1), *exists so that* $\mathbf{u} \in H^1(\Omega)^2$ *and* $p \in H^1(\Omega)$. *Then, for sufficiently small h, there exists a unique solution* $(\mathbf{u}_h, p_h)$ *in* $\mathbf{V}_h \times W_h$ *of the system* (4.8) *such that*

$$\|p - p_h\|_0 \le Ch(\|p\|_1 + \|\mathbf{u}\|_1), \tag{6.21}$$

$$\|\mathbf{u} - \mathbf{u}_h\|_0 \le Ch\|\mathbf{u}\|_1. \tag{6.22}$$

Proof. We need only to refine the proof of Theorem 3.1 in [7] and we give a sketch here. Let $(\tilde{\mathbf{u}}_h, \tilde{p}_h)$ be the standard mixed finite element solution of (1.1). Then estimates (6.21)–(6.22) hold with the $(\mathbf{u}_h, p_h)$ replaced by $(\tilde{\mathbf{u}}_h, \tilde{p}_h)$. Using the triangle inequality, we see it suffices to estimate $\tilde{\mathbf{e}}_h := \tilde{\mathbf{u}}_h - \mathbf{u}_h$. A simple calculation or copying the error equation (3.9) in [7] gives

$$a(\tilde{\mathbf{e}}_h, \gamma_h\tilde{\mathbf{e}}_h) = -a(\mathbf{u} - \tilde{\mathbf{u}}_h, \gamma_h\tilde{\mathbf{e}}_h) - a(\mathbf{u}, (I - \gamma_h)\tilde{\mathbf{e}}_h) + B(\tilde{\mathbf{e}}_h, p - \tilde{p}_h), \tag{6.23}$$

where $a(\mathbf{w}, \mathbf{v}) = (K^{-1}\mathbf{w}, \mathbf{v})$ and $B(\mathbf{w}, q) = (\mathrm{div}\mathbf{w}, q)$. The last term $B(\tilde{\mathbf{e}}_h, p - \tilde{p}_h) = 0$ is by orthogonality condition (3.14) and the fact that $B(\tilde{\mathbf{e}}_h, q_h) = 0$ for all $q_h \in W_h$. Now use the coercivity and (5.1)–(5.3) to derive the second assertion of the theorem. The first one is by the inf-sup condition. ∎

7. Russell's scheme. In [4, 25] Russell discussed a somewhat different scheme than ours. It has turned out that his scheme corresponds to the following set of test functions:

$$\tilde{\gamma}_h \boldsymbol{\phi}_{i+1/2,j} \equiv \begin{cases} \dfrac{J}{J(3/4,1/2)} \mathcal{P}_{i,j} \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \text{on } Q_{i,j} \cap Q_{i+1/2,j}, \\ \dfrac{J}{J(1/4,1/2)} \mathcal{P}_{i+1,j} \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \text{on } Q_{i+1,j} \cap Q_{i+1/2,j}, \\ 0 & \text{elsewhere,} \end{cases} \tag{7.1}$$

and

$$\tilde{\gamma}_h \boldsymbol{\phi}_{i,j+1/2} \equiv \begin{cases} \dfrac{J}{J(1/2,3/4)} \mathcal{P}_{i,j} \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \text{on } Q_{i,j} \cap Q_{i,j+1/2}, \\ \dfrac{J}{J(1/2,1/4)} \mathcal{P}_{i,j+1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \text{on } Q_{i,j+1} \cap Q_{i,j+1/2}, \\ 0 & \text{elsewhere.} \end{cases} \tag{7.2}$$

For brevity we write $\mathcal{P}_{i,j} = \mathcal{P}_{Q_{i,j}}$ and omit the subscripts for J. It is easy to see that this is equivalent to replacing the nonconstant term J^{-1} in the Piola transformation with its value at the center of one-half of a covolume. In this case the test functions are piecewise linear polynomials, and the resulting scheme is computationally more convenient.

In order to see how these test functions lead to Russell's scheme, we proceed in the same way as in section 4. We begin with the system

$$(K^{-1}\mathbf{u} + \nabla p + \boldsymbol{\beta} p, \tilde{\gamma}_h \mathbf{v}) = 0 \qquad \forall \mathbf{v} \in \mathbf{V}_h, \tag{7.3}$$

$$(\mathrm{div}\, \mathbf{u} + cp, w) = (f, w) \qquad \forall w \in W_h, \tag{7.4}$$

which is obtained by integrating the mixed system (1.3) with respect to the new test functions (7.1), (7.2). We next transform the term $(\nabla p, \tilde{\gamma}_h \mathbf{v})$ into an equivalent form containing no derivatives of p: Take $\mathbf{v} = \boldsymbol{\phi}_{i+1/2,j}$,

$$\begin{aligned}(\nabla p, \tilde{\gamma}_h \boldsymbol{\phi}_{i+1/2,j}) = \int_{Q_{i,j} \cap Q_{i+1/2,j}} \nabla p \cdot \frac{J}{J(3/4,1/2)} \mathcal{P}_{i,j} \begin{pmatrix} 1 \\ 0 \end{pmatrix} dxdy \\ + \int_{Q_{i+1,j} \cap Q_{i+1/2,j}} \nabla p \cdot \frac{J}{J(1/4,1/2)} \mathcal{P}_{i+1,j} \begin{pmatrix} 1 \\ 0 \end{pmatrix} dxdy.\end{aligned}$$

As before, we use integration by parts to obtain

$$\begin{aligned}
\int_{Q_{i,j}\cap Q_{i+1/2,j}} &\nabla p \cdot \frac{J}{J(3/4,1/2)} \mathcal{P}_{i,j} \begin{pmatrix} 1 \\ 0 \end{pmatrix} dxdy \\
&= \int_0^1 \int_{1/2}^1 \frac{J(\hat{x},\hat{y})}{J(3/4,1/2)} \frac{\partial \hat{p}}{\partial \hat{x}}\, d\hat{x} d\hat{y} \\
&= \frac{1}{J(3/4,1/2)} \left[\int_0^1 [J(1,\hat{y})\hat{p}(1,\hat{y}) - J(1/2,\hat{y})\hat{p}(1/2,\hat{y})]\, d\hat{y} \right. \\
&\qquad\qquad \left. - \frac{\partial J}{\partial \hat{x}} \int_0^1 \int_{1/2}^1 \hat{p}(\hat{x},\hat{y})\, d\hat{x} d\hat{y} \right] \\
&= \int_0^1 \hat{p}(1,\hat{y})\, d\hat{y} + \frac{1}{J(3/4,1/2)} \left[\int_0^1 [J(1,\hat{y}) - J(3/4,1/2)]\hat{p}(1,\hat{y})\, d\hat{y} \right. \\
&\qquad \left. - \int_0^1 J(1/2,\hat{y})\hat{p}(1/2,\hat{y})\, d\hat{y} - \frac{\partial J}{\partial \hat{x}} \int_0^1 \int_{1/2}^1 \hat{p}(\hat{x},\hat{y})\, d\hat{x} d\hat{y} \right],
\end{aligned}$$

and similarly,

$$\begin{aligned}
\int_{Q_{i+1,j}\cap Q_{i+1/2,j}} &\nabla p \cdot \frac{J}{J(1/4,1/2)} \mathcal{P}_{i+1,j} \begin{pmatrix} 1 \\ 0 \end{pmatrix} dxdy \\
&= -\int_0^1 \hat{p}(0,\hat{y})\, d\hat{y} + \frac{1}{J(1/4,1/2)} \left[\int_0^1 J(1/2,\hat{y})\hat{p}(1/2,\hat{y})\, d\hat{y} \right. \\
&\quad \left. - \int_0^1 [J(0,\hat{y}) - J(1/4,1/2)]\hat{p}(0,\hat{y})\, d\hat{y} - \frac{\partial J}{\partial \hat{x}} \int_0^1 \int_0^{1/2} \hat{p}(\hat{x},\hat{y})\, d\hat{x} d\hat{y} \right].
\end{aligned}$$

Recalling that the first terms in both integrals canceled each other in the summation over the whole covolume $Q_{i+1/2,j}$, we can write down the expression for the associated bilinear form $\tilde{b}(\tilde{\gamma}_h \mathbf{v}, p)$, which is not given here due to its complexity.

Now let $p = p_h$ be piecewise constant with the value $p_{i,j}$ on $Q_{i,j}$. Then it is easy to see that $\tilde{b}(\tilde{\gamma}_h \boldsymbol{\phi}_{i+1/2,j}, p_h)$ is of the form $Ap_{i,j} + Bp_{i+1,j}$ with the coefficients A, B given by

$$\begin{aligned}
A &= -1 + \frac{1}{J(3/4,1/2)} \left[J(1,1/2) - J(1/2,1/2) - \frac{1}{2}\frac{\partial J}{\partial \hat{x}} \right] = -1, \\
B &= 1 + \frac{1}{J(1/4,1/2)} \left[J(1/2,1/2) - J(0,1/2) - \frac{1}{2}\frac{\partial J}{\partial \hat{x}} \right] = 1.
\end{aligned}$$

This leads to

$$\tilde{b}(\tilde{\gamma}_h \boldsymbol{\phi}_{i+1/2,j}, p_h) = p_{i+1,j} - p_{i,j}, \tag{7.5}$$

and likewise we have

$$\tilde{b}(\tilde{\gamma}_h \boldsymbol{\phi}_{i,j+1/2}, p_h) = p_{i,j+1} - p_{i,j}, \tag{7.6}$$

which implies that Russell's scheme is indeed equivalent to choosing (7.1) and (7.2) as test functions. In addition we have proved the following lemma.

LEMMA 7.1. *We have for all* $\mathbf{v} \in \mathbf{V}_h$ *and* $p_h \in W_h$

$$\tilde{b}(\tilde{\gamma}_h \mathbf{v}, p_h) = -(\operatorname{div} \mathbf{v}, p_h).$$

This lemma shows that Russell's scheme can be written in the Galerkin form

$$(K^{-1}\mathbf{u}_h, \tilde{\gamma}_h \mathbf{v}) - (\operatorname{div} \mathbf{v}, p_h) + (\boldsymbol{\beta} p_h, \tilde{\gamma}_h \mathbf{v}) = 0 \qquad \forall \mathbf{v} \in \mathbf{V}_h, \tag{7.7a}$$

$$(\operatorname{div} \mathbf{u}_h, w) + (c p_h, w) = (f, w) \qquad \forall w \in W_h. \tag{7.7b}$$

Error analysis for Russell's scheme (7.7) can be performed in the same way as we did previously for our original scheme (4.8). This is accomplished by comparing $\gamma_h \mathbf{v}$ and $\tilde{\gamma}_h \mathbf{v}$.

LEMMA 7.2. *For every* $\mathbf{v} \in \mathbf{V}_h$ *we have*

$$\|\gamma_h \mathbf{v} - \tilde{\gamma}_h \mathbf{v}\|_0 \le C h \|\mathbf{v}\|_0. \tag{7.8}$$

Proof. Define $\hat{\mathbf{e}} \in \mathbf{V}_h(\hat{Q})$ by $\mathcal{P}_{i,j} \hat{\mathbf{e}} = \gamma_h \mathbf{v}|_{Q_{i,j}} - \tilde{\gamma}_h \mathbf{v}|_{Q_{i,j}}$ and let

$$\mathbf{v}|_{Q_{i,j}} = v_{i-1/2,j} \boldsymbol{\phi}_{i-1/2,j} + v_{i+1/2,j} \boldsymbol{\phi}_{i+1/2,j} + v_{i,j-1/2} \boldsymbol{\phi}_{i,j-1/2} + v_{i,j+1/2} \boldsymbol{\phi}_{i,j+1/2}.$$

Then we can easily deduce that

$$\begin{aligned}
\hat{\mathbf{e}} = v_{i-1/2,j} &\left[1 - \frac{J}{J(1/4,1/2)}\right] \hat{\gamma} \hat{\boldsymbol{\phi}}_{x,0} + v_{i+1/2,j} \left[1 - \frac{J}{J(3/4,1/2)}\right] \hat{\gamma} \hat{\boldsymbol{\phi}}_{x,1} \\
&+ v_{i,j-1/2} \left[1 - \frac{J}{J(1/2,1/4)}\right] \hat{\gamma} \hat{\boldsymbol{\phi}}_{y,0} + v_{i,j+1/2} \left[1 - \frac{J}{J(1/2,3/4)}\right] \hat{\gamma} \hat{\boldsymbol{\phi}}_{y,1}.
\end{aligned}$$

Note that the four terms make independent contributions to $\|\hat{\mathbf{e}}\|^2_{0,\hat{Q}}$. Namely, we have

$$\|\hat{\mathbf{e}}\|^2_{0,\hat{Q}} = v^2_{i-1/2,j} I_1 + v^2_{i+1/2,j} I_2 + v^2_{i,j-1/2} I_3 + v^2_{i,j+1/2} I_4,$$

where

$$I_1 = \int_0^1 \int_0^{1/2} \left[1 - \frac{J}{J(1/4,1/2)}\right]^2 d\hat{x} d\hat{y},$$

and I_2, I_3, I_4 are similarly defined. A Taylor expansion of J shows that

$$I_1 \le C(\beta^2 + \gamma^2) |J^{-1}|^2_\infty \le C h^2,$$

with the same results for I_2, I_3, and I_4. Consequently, it follows that

$$\|\hat{\mathbf{e}}\|^2_{0,\hat{Q}} \le C h^2 (v^2_{i-1/2,j} + v^2_{i+1/2,j} + v^2_{i,j-1/2} + v^2_{i,j+1/2}) \le C h^2 \|\hat{\mathbf{v}}\|^2_{0,\hat{Q}},$$

which gives

$$\|\gamma_h \mathbf{v} - \tilde{\gamma}_h \mathbf{v}\|^2_{0,Q_{i,j}} \le C \|\hat{\mathbf{e}}\|^2_{0,\hat{Q}} \le C h^2 \|\hat{\mathbf{v}}\|^2_{0,\hat{Q}} \le C h^2 \|\mathbf{v}\|^2_{0,Q_{i,j}}. \tag{7.9}$$

By summing over $Q_{i,j} \in \mathcal{Q}_h$ we obtain the desired result. $\square$

Now we provide the optimal error estimates for Russell's scheme.

THEOREM 7.3. *Under the same hypotheses of Theorem* 6.5, *we have for sufficiently small h that there exists a unique solution $(\mathbf{u}_h, p_h)$ in $\mathbf{V}_h \times W_h$ of the system* (7.7) *such that*

$$\|\mathbf{u}-\mathbf{u}_h\|_0 + \|p-p_h\|_0 \le Ch(\|\mathbf{u}\|_1 + \|p\|_1), \tag{7.10}$$

$$\|\operatorname{div}(\mathbf{u}-\mathbf{u}_h)\|_0 \le Ch(\|\mathbf{u}\|_1 + \|p\|_1 + \|\operatorname{div}\mathbf{u}\|_1). \tag{7.11}$$

Moreover, the following superconvergence result holds for the pressure variable:

$$\|P_h p - p_h\|_0 \le Ch^2(\|\mathbf{u}\|_1 + \|p\|_1 + \|\operatorname{div}\mathbf{u}\|_1). \tag{7.12}$$

Proof. Thanks to Lemma 7.2, it is easy to verify that Lemmas 5.1–5.3 and 6.2–6.4 still hold with γ_h replaced by $\tilde{\gamma}_h$. Since the proof of Theorem 6.5 depends only on these lemmas, we conclude that the results of Theorem 6.5 are valid equally for Russell's scheme. □

Of course, an analogous result to Theorem 6.7 holds here also. A final remark concerning our test functions in (3.22) and Russell's test function is in order here.

REMARK 7.1. *First, it was pointed out by a referee that the test functions in* (3.22) *were also independently suggested in* 1999 *by Garanzha and Konshin* [16]. *Russell's test functions are more computationally convenient (allowing for analytical integration), but those in* (3.22) *perform better in some cases* [22].

Appendix. In this appendix we introduce an important approach in the finite volume method when it comes to proving coercivity in the presence of a nondiagonal coefficient tensor K (anisotropic case). It was first introduced in Chou and Vassilevski [12] for the nonoverlapping covolume case.

Let us first extend the domain of the transfer operator $\hat{\gamma}$ from $V_h(\hat{Q})$ to $P_1(\hat{Q})^2$, the set of linear (vector) polynomials, as follows. First extend it componentwise; i.e., for $\hat{w} \in P_1(\hat{Q})$ define

$$\hat{\gamma}\hat{w} = \begin{cases} \hat{w}(0,0) & \text{on } [0,1/2]\times[0,1/2], \\ \hat{w}(0,1) & \text{on } [0,1/2]\times[1/2,1], \\ \hat{w}(1,0) & \text{on } [1/2,1]\times[0,1/2], \\ \hat{w}(1,1) & \text{on } [1/2,1]\times[1/2,1]. \end{cases}$$

Then define the vector version: $\hat{\gamma}\hat{\mathbf{v}} := (\hat{\gamma}\hat{u}, \hat{\gamma}\hat{v}),\ \hat{\mathbf{v}} = (\hat{u}, \hat{v})^t$. Now by direct calculation one can verify the following good property.

LEMMA A.1. *For the extension operator $\hat{\gamma}$,*

$$\bar{B}\hat{\gamma}\mathbf{v} = \hat{\gamma}\bar{B}\mathbf{v} \quad \forall \mathbf{v} \in P_1(\hat{Q})^2,$$

where $\bar{B}$ is a constant matrix.

Note that the right-hand side is not well defined for the original $\hat{\gamma}$. Let $(\cdot,\cdot)_K$ denote the L^2-inner product on domain K. Next we show another lemma.

LEMMA A.2. *The following coercivity estimate holds:*

$$(\hat{\gamma}\mathbf{v}, \mathbf{v})_{\hat{Q}} \ge C\|\mathbf{v}\|^2_{0,\hat{Q}} \quad \forall \mathbf{v} \in P_1(\hat{Q})^2.$$

Proof. Let $\mathbf{v} = (a+bx+cy, p+qx+ry)^t$. Then after simple but tedious calculations we obtain that

$$\begin{aligned} \textstyle\int_{\hat{Q}} \hat{\gamma}\mathbf{v}\cdot\mathbf{v}\, dxdy &= \big[a^2 + \tfrac{3}{8}b^2 + ab + ac + \tfrac{1}{2}bc + \tfrac{3}{8}c^2 \\ &\qquad + p^2 + \tfrac{3}{8}q^2 + pq + pr + \tfrac{1}{2}qr + \tfrac{3}{8}r^2\big]\cdot \end{aligned}$$

The positive definiteness is then seen by looking at the coefficient matrix of the above quadratic forms,

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{8} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{3}{8} \end{bmatrix}.$$

By direct computation one sees that the eigenvalues are 0.125, 1.544, and 0.809. Hence

$$\int_{\hat{Q}} \hat{\gamma}\mathbf{v}\cdot\mathbf{v}\, dxdy \geq 0.125\left[a^2+b^2+c^2+p^2+q^2+r^2\right] \geq C\|\mathbf{v}\|^2_{0,\hat{Q}} \cdot \qquad \square$$

We finally show the coercivity of $\hat{\gamma}_h$ in the weighted inner product $(\bar{B}\cdot,\cdot)$, where $\bar{B}$ is a symmetric positive-definite matrix.

LEMMA A.3.

$$(\bar{B}\mathbf{v}, \hat{\gamma}\mathbf{v})_{\hat{Q}} \geq C(\bar{B}\mathbf{v},\mathbf{v})_{\hat{Q}} \quad \forall \mathbf{v}\in P_1(\hat{Q})^2.$$

Proof. Since $\bar{B}$ is constant, by Lemma A.1,

$$\bar{B}^{\frac{1}{2}}\hat{\gamma}\mathbf{v} = \hat{\gamma}\bar{B}^{\frac{1}{2}}\mathbf{v}.$$

Therefore, by Lemma A.2 one gets

$$(\bar{B}\hat{\gamma}\mathbf{v},\mathbf{v})_{\hat{Q}} = (\hat{\gamma}\bar{B}^{\frac{1}{2}}\mathbf{v},\ \bar{B}^{\frac{1}{2}}\mathbf{v})_{\hat{Q}} \geq C(\bar{B}^{\frac{1}{2}}\mathbf{v},\ \bar{B}^{\frac{1}{2}}\mathbf{v})_{\hat{Q}} = C(\bar{B}\mathbf{v},\mathbf{v})_{\hat{Q}} \cdot \qquad \square$$

REMARK A.1. *In the rectangular grid case* [7], *Chou and Kwak were only able to cover the diagonal tensor K case due to their way of handling the coercivity analysis in Lemma* 2.2 *there. It is clear now from the above general approach that the nondiagonal case is also settled.*

Acknowledgments. We thank Prof. T. F. Russell for his encouragement and generous comments in connection to our research in this direction. Helpful remarks of the referees are also greatly appreciated.

REFERENCES

[1] T. ARBOGAST, M. F. WHEELER, AND I. YOTOV, *Mixed finite elements for elliptic problems with tensor coefficients as cell-centered finite differences*, SIAM J. Numer. Anal., 34 (1997), pp. 828–852.

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

[3] P. CIARLET, *The Finite Element Method for Elliptic Problems*, North–Holland, Amsterdam, 1978.

[4] Z. CAI, J. E. JONES, S. F. MCCORMICK, AND T. F. RUSSELL, *Control-volume mixed finite element methods*, Comput. Geosci., 1 (1997), pp. 289–315.

[5] J. C. CAVENDISH, C. A. HALL, AND T. A. PORSCHING, *A complementary volume approach for modeling three-dimensional Navier-Stokes equations using dual Delaunay/Voronoi tessellations*, Internat. J. Numer. Methods Heat Fluid Flow, 4 (1994), pp. 329–345.

[6] S. H. CHOU, *Analysis and convergence of a covolume method for the generalized Stokes problem*, Math. Comp., 66 (1997), pp. 85–104.

[7] S. H. CHOU AND D. Y. KWAK, *Mixed covolume methods on rectangular grids for elliptic problems*, SIAM J. Numer. Anal., 37 (2000), pp. 758–771.

[8] S. H. CHOU AND D. Y. KWAK, *Analysis and convergence of a MAC-like scheme for the generalized Stokes problem*, Numer. Methods Partial Differential Equations, 13 (1997), pp. 1–16.

[9] S. H. Chou and D. Y. Kwak, *A covolume method based on rotated bilinears for the generalized Stokes problem*, SIAM J. Numer. Anal., 35 (1998), pp. 494–507.
[10] S. H. Chou, D. Y. Kwak, and P. S. Vassilevski, *Mixed upwinding covolume methods on rectangular grids for convection-diffusion problems*, SIAM J. Sci. Comput., 21 (1999), pp. 145–165.
[11] S. H. Chou and S. Tang, *Conservative P1 conforming and nonconforming Galerkin FEMS: Effective flux evaluation via a nonmixed method approach*, SIAM. J. Numer. Anal., 38 (2000), pp. 660–680.
[12] S. H. Chou and P. S. Vassilevski, *A general mixed covolume framework for constructing conservative schemes for elliptic problems*, Math. Comp., 68 (1999), pp. 991–1011.
[13] J. Douglas, Jr. and J. E. Roberts, *Global estimates for mixed methods for second order elliptic equations*, Math. Comp., 44 (1985), pp. 39–52.
[14] R. E. Ewing, ed., *The Mathematics of Reservoir Simulation*, SIAM, Philadelphia, 1983.
[15] M. Fortin, *An analysis of the convergence of mixed finite element methods*, RAIRO Anal. Numér., 11 (1977), pp. 341–354.
[16] V. Garanzha and V. Konshin, *Approximation Schemes and Discrete Well Models for the Numerical Simulation of the 2-D Non-Darcy Fluid Flows in Porous Media*, Communications on Applied Mathematics, Russian Academy of Sciences, Computer Center, Moscow, 1999.
[17] V. Girault and P. A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer-Verlag, Berlin, 1986.
[18] C. A. Hall, T. A. Porsching, and P. Hu, *Covolume-dual variable method for thermally expandable flow on unstructured triangular grids*, Comp. Fluid Dynam., 2 (1994), pp. 111–139.
[19] F. H. Harlow and F. E. Welch, *Numerical calculations of time dependent viscous incompressible flow of fluid with a free surface*, Phys. Fluids, 8 (1965), pp. 2181–2197.
[20] J. E. Jones, *A Mixed Finite Volume Element Method for Accurate Computation of Fluid Velocities in Porous Media*, Ph.D. thesis, University of Colorado at Denver, 1995.
[21] R. D. Lazarov, I. D. Michev, and P. S. Vassilevski, *Finite volume methods for convection-diffusion problems*, SIAM J. Numer. Anal., 33 (1996), pp. 31–55.
[22] R. L. Naff, T. F. Russell, and J. D. Wilson, *Test functions for three-dimensional control–volume mixed finite-element methods on irregular grids*, in Computational Methods in Water Resources, Vol. 2, L. R. Bentley et al., eds., A. A. Balkma, Rotterdam, 2000, pp. 667–684.
[23] R. A. Nicolaides, T. A. Porsching, and C. A. Hall, *Covolume methods in computational fluid dynamics*, in Computational Fluid Dynamics Review, M. Hafez and K. Oshma, eds., John Wiley and Sons, New York, 1995, pp. 279–299.
[24] D. W. Peaceman, *Fundamentals of Numerical Reservoir Simulation*, Elsevier, Amsterdam, 1977.
[25] T. F. Russell, *Rigorous Block-Centered Discretizations on Irregular Grids: Improved Simulation of Complex Reservoir Systems*, Tech. report 3, Reservoir Simulation Research Corporation, Tulsa, OK, 1995.
[26] P. A. Raviart and J. M. Thomas, *A mixed finite element method for 2nd order elliptic problems*, in Proc. Conf. on Mathematical Aspects of Finite Element Methods, Lecture Notes in Math. 606, Springer-Verlag, Berlin, 1977, pp. 292–315.
[27] J. Roberts and J. Thomas, *Mixed and hybrid methods*, in Handbook of Numerical Analysis, Vol. II, P. G. Ciarlet and J. L. Lions, eds., North-Holland, Amsterdam, 1991, pp. 523–639.
[28] J. Shen, *Mixed Finite Element Methods on Distorted Rectangular Grids*, Tech. report ISC-94-13-MATH, Institute for Scientific Computation, Texas A&M University, College Station, TX, 1994. Available online at http://www.isc.tamu.edu/iscpubs/
[29] E. Suli, *The accuracy of cell vertex finite volume methods on quadrilateral meshes*, Math. Comp., 59 (1992), pp. 359–382.
[30] J. Wang and T. Mathew, *Mixed finite element methods over quadrilaterals*, in Proc. Third International Conf. on Advances in Numerical Methods and Applications, I. T. Dimov, Bl. Sendov, and P. Vassilevski, eds., World Scientific, River Edge, NJ, 1994, pp. 203–214.