ICES REPORT 16-18

August 2016

Mixed methods for two-phase Darcy-Stokes mixtures of partially melted materials with regions of zero porosity

by

Todd Arbogast, Marc A. Hesse, and Abraham L. Taicher

The Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, Texas 78712

Reference: Todd Arbogast, Marc A. Hesse, and Abraham L. Taicher, "Mixed methods for two-phase Darcy-Stokes mixtures of partially melted materials with regions of zero porosity," ICES REPORT 16-18, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, August 2016.

MIXED METHODS FOR TWO-PHASE DARCY-STOKES MIXTURES OF PARTIALLY MELTED MATERIALS WITH REGIONS OF ZERO POROSITY*

TODD ARBOGAST†, MARC A. HESSE‡, AND ABRAHAM L. TAICHER§

Abstract. The Earth's mantle (or, e.g., a glacier) involves a deformable solid matrix phase within which a second phase, a fluid, may form due to melting processes. The system is modeled as a dual-continuum mixture, with at each point of space the solid matrix being governed by a Stokes flow and the fluid melt, if it exists, being governed by a Darcy law. This system is mathematically degenerate when the porosity (volume fraction of fluid) vanishes. Assuming the porosity is given, we develop a mixed variational framework for the mechanics of the system by carefully scaling the Darcy variables by powers of the porosity. We prove that the variational problem is well-posed, even when there are regions of one and two phases. We then develop an accurate mixed finite element method for solving this Darcy-Stokes system and prove a convergence result. Numerical results are presented that illustrate and verify the convergence of the method.

Key words. Degenerate Elliptic, Energy Bounds, Mixed Finite Element Method, Mantle Dynamics, Glaciers, Mid-Ocean Ridge

AMS subject classifications. 65N12, 65N30, 35J70, 76M10, 76S05, 76T99

1. Introduction.

The goal of this work is to develop mixed finite element methods for the mechanics part of the equations of mantle dynamics introduced by McKenzie [29]. This multi-phase model is based on a mixture of fluid melt and solid matrix, where both fluid and solid phases are assumed to exist *at each point* of the spatial domain Ω. The fluid melt velocity obeys Darcy's law while the deformable "solid" matrix is governed by a highly viscous Stokes equation. The system is coupled through mass conservation and compaction relations. Together these equations form a dual-continuum of fluid melt and solid matrix mixed together over the domain Ω. A key quantity is the porosity ϕ, which is the volume fraction of fluid melt. It is assumed to be much smaller than one, but it may be zero in parts of the domain where there is no fluid melt.

The equations of mantle dynamics have a wide range of applications in Earth physics [2, 27, 26, 25], such as in modeling mid-ocean ridges, subduction zones, and hot-spot volcanism, as well as to glacier dynamics [22, 8, 39] and other two-phase flows in porous media [12, 28, 13, 18]. The physical problem that guides our work is the modeling of a mid-ocean ridge. Melt is believed to migrate upward until it reaches the lithospheric "tent" where it then moves toward the ridge within a high porosity band. Simulation of this physical phenomenon requires confidence in numerical methods to accurately handle highly heterogeneous porosity and the single-phase to two-phase transition.

*This work was supported by the U.S. National Science Foundation under grants EAR-1025321 and DMS-1418752. The first two authors thank the Isaac Newton Institute for Mathematical Sciences, Cambridge (supported by EPSRC grant no EP/K032208/1), for support during the programme Melt in the Mantle, 15 Feb. to 17 June 2016, when some work on this paper was undertaken.

†Department of Mathematics, University of Texas, 2515 Speedway, C1200, Austin, TX 78712-1202 and Institute for Computational Engineering and Sciences, University of Texas, 201 East 24th St., C0200, Austin, TX 78712-1229 (arbogast@ices.utexas.edu)

‡Jackson School of Geosciences, University of Texas, 2305 Speedway, C1160, Austin, TX 78712-1692 (mhesse@jsg.utexas.edu)

§Institute for Computational Engineering and Sciences, University of Texas, 201 East 24th St., C0200, Austin, TX 78712-1229 (ataicher@ices.utexas.edu)

When coupled with solute transport and thermal evolution, the model transitions dynamically in time from a non-porous single phase Stokes solid to a two-phase porous medium. The model is based on mixture theory, and it has the advantage that the free boundary between the one and two-phase regions need *not* be determined explicitly in the numerical approximation. Unfortunately, the disadvantage is that the Darcy part of the equations is mathematically degenerate in regions where the porosity is zero, since then there is only the one solid phase, even though the model equations continue to describe both phases over the entire domain Ω.

A mixed finite element method (MFEM) is a good candidate for the computational model of this system. MFEMs have an extensive theory for both Darcy and Stokes flow. Moreover, velocity fields computed using MFEM are continuous on each element and have a continuous normal component across element boundaries. This allows coupling with the transport equations of solute and thermal evolution, since the velocities unambiguously determine particle trajectories.

At any instant the porosity field $\phi(\mathbf{x})$ is given and the coupled Darcy-Stokes equations determine the unknown fluid flux or Darcy velocity (relative to the solid) $\mathbf{u}$, the solid matrix velocity $\mathbf{v}_s$, and fluid melt and solid matrix pressure *potentials* q_f and q_s (see (2.8) below), which give rise to the mixture potential $q = \phi q_f + (1-\phi)q_s$. The equations can be written on all of Ω in the form

$$\mathbf{u} + \frac{k_0\phi^{2+2\Theta}}{\mu_f}\nabla q_f = 0, \tag{1.1}$$

$$\mu_s \nabla\cdot\mathbf{u} + \frac{\phi}{1-\phi}(q_f - q) = 0, \tag{1.2}$$

$$\nabla q - \nabla\cdot\hat{\boldsymbol{\sigma}}(\mathbf{v}_s) = -(1-\phi)\,\rho_r\,\mathbf{g}, \tag{1.3}$$

$$\mu_s \nabla\cdot\mathbf{v}_s - \frac{\phi}{1-\phi}(q_f - q) = 0, \tag{1.4}$$

where the deviatoric stress of the mixture $\hat{\boldsymbol{\sigma}}$ is

$$\hat{\boldsymbol{\sigma}} = \hat{\boldsymbol{\sigma}}(\mathbf{v}_s) = 2\mu_s(1-\phi)\big(\mathcal{D}\mathbf{v}_s - \tfrac{1}{3}\nabla\cdot\mathbf{v}_s\mathbf{I}\big), \tag{1.5}$$

wherein $\mathcal{D}\mathbf{v}_s = \frac{1}{2}(\nabla\mathbf{v}_s + \nabla\mathbf{v}_s^T)$ is the symmetric gradient. The model parameters are assumed to be constant and include the (relative) density difference between the fluid and the solid matrix $\rho_r = \rho_f - \rho_s$ and the fluid and solid viscosities μ_f and μ_s, as well as the gravitational acceleration vector $\mathbf{g}$ pointing downwards. For completeness, the system is derived in detail in the next section, where k_0, $\Theta \in [0, 1/2]$, and other quantities are defined. Equation (1.1) represents Darcy's law for an incompressible fluid, (1.3)–(1.5) is a Stokes system for a highly viscous, compressible material (matrix plus fluid), and (1.2) plus (1.4) enforces mass conservation.

As we will show later, the Stokes part of the system is well-behaved, but the Darcy part has difficulties when ϕ vanishes. We will derive stability (also called energy) estimates in (3.15) that show

$$\|\phi^{-1-\Theta}\mathbf{u}\| + \|\phi^{-1/2}\nabla\cdot\mathbf{u}\| + \|\phi^{1/2}q_f\| \le C \tag{1.6}$$

for some constant C, where $\|\cdot\|$ is the $L^2(\Omega)$-norm. These estimates suggest that the fluid pressure may be *unbounded* where porosity vanishes. Indeed, the fluid pressure is no longer a physical variable when there is no fluid, and so it is not possible to have stability control on q_f. Moreover, any numerical method that does not take into

account the degeneracy of ϕ, say by instead imposing a small nonzero porosity ϕ_0 everywhere, is sure to have a condition number that grows as $\phi_0 \to 0$. Our numerical results will show both these issues.

Recently, two of the current authors [5, 4] developed a MFEM and cell-centered finite difference method for a single Darcy system with a similar degeneracy as appears in (1.1)–(1.2). The key is to follow the hint in the stability estimates and scale the fluid pressure and velocity to avoid problems with vanishing porosity. In this paper we apply this idea to the full set (1.1)–(1.5) of mantle dynamics equations.

In the rest of the paper, we first review McKenzie's derivation of the equations for partially molten materials in Section 2. In Section 3 we present several relatively standard mixed variational formulations of the governing equations (1.1)–(1.5), which require that the porosity *not* degenerate to zero. The first is a straightforward mixed variational formulation of the equations assuming positive porosity everywhere, and it allows us to derive stability/energy estimates for the pressure potentials and velocities. We also present two common but inadequate approaches to resolving the problem of a vanishing porosity. In Section 4, we finally present our scaled formulation which directly resolves the issue of degenerate porosity. We prove the existence and uniqueness of a solution to this scaled variational formulation. In Section 5 we define our MFEM for the numerical approximation of the scaled variational formulation and prove its stability and convergence. In section 6, we present a modification of the MFEM that is locally mass conservative. In section 7, we discuss implementation and give a mass lumping modification that results in a relatively simple solution procedure on rectangular meshes. Numerical results illustrating and evaluating the effects of degenerate porosity are given in Sections 8–9. We include tests of a one-dimensional compacting column with various porosity functions, and a two-dimensional test example akin to a mid ocean ridge. We conclude the paper in Section 10.

2. Governing Equations. In this section, for completeness, we recall briefly the derivation of the mechanical system (1.1)–(1.5) according to McKenzie [29]. It is a dual-continuum mixture approach similar to other models of flow in the Earth's mantle, see, e.g., [2, 27, 26, 25], where the mixing parameter is the porosity ϕ.

For a quantity Ψ having both a part in the fluid and in the solid matrix, define the mixture variable

$$\overline{\Psi} = \phi\Psi_f + (1-\phi)\Psi_s = \Psi_s + \phi\Psi_r, \quad \Psi_r = \Psi_f - \Psi_s. \tag{2.1}$$

As the two phases melt or solidify, total mass is conserved according to

$$\frac{\partial \bar{\rho}}{\partial t} + \nabla \cdot \overline{\rho \mathbf{v}} = 0. \tag{2.2}$$

If we apply the *Boussinsq approximation* [36] (constant and equal densities for non-buoyancy terms), this becomes more simply

$$\nabla \cdot \bar{\mathbf{v}} = 0. \tag{2.3}$$

In the case of mantle convection, Table 2.1 lists representative values of parameters for the fluid and solid phases. The fluid and solid matrix Reynolds numbers,

$$\mathrm{Re}_f = \frac{\rho_f |\mathbf{v}_f| a}{\mu_f} \ll 1 \quad \text{and} \quad \mathrm{Re}_s = \frac{\rho_s |\mathbf{v}_s| L}{\mu_s} \ll 1,$$

are both very small, where a or L is a characteristic length, $|\mathbf{v}|$ is a characteristic velocity, ρ is density, and μ is the dynamic viscosity. That is, both the rate of

TABLE 2.1
Representative values of various parameters for the Earth's mantle. Grain size represents the average size of a matrix grain around which fluid flows. The ridge spreading rate is the rate at which ocean floor is spreading at mid-ocean ridges.

Parameter		Value or Range
a	Grain size	10^{-3} m
L	Matrix length scale	10^4 to 10^5 m
$\lvert\mathbf{v}_f\rvert$	Characteristic fluid velocity	10^{-8} to 10^{-6} m/s
$\lvert\mathbf{v}_s\rvert$	Ridge spreading rate	10^{-10} to 10^{-9} m/s
k_0	Permeability constant	10^{-9}–10^{-6} m^2
μ_s	Solid shear viscosity	10^{19} Pa·s
μ_f	Fluid viscosity	1 Pa·s
ρ_s	Solid matrix Density	3300 kg/m^3
ρ_f	Fluid Density	2800 kg/m^3
g	Gravity	9.8 m/s^2

advection of momentum directly through flow and the rate of change of momentum are negligible in comparison to momentum diffusion by viscous effects. This leads to Darcy's law for fluid flow around matrix "grains," which is

$$\mathbf{u} = \phi\mathbf{v}_r = \phi(\mathbf{v}_f - \mathbf{v}_s) = -\frac{k_0\phi^{2+2\Theta}}{\mu_f}(\nabla p_f - \rho_f\mathbf{g}), \tag{2.4}$$

where $\mathbf{v}_f$ and p_f are the fluid velocity and pressure, $\mathbf{v}_r = \mathbf{v}_f - \mathbf{v}_s$ is the relative velocity, and $k(\phi) = k_0\phi^{2+2\Theta}$ is the porosity dependent permeability calculated according to the standard Kozeny-Carmen relationship [12, 40], simplified for small porosity and constant grain size, where Θ is a constant exponent between 0 and 1/2.

The fluid stress is $\boldsymbol{\sigma}_f = -p_f\mathbf{I}$ and the solid matrix stress is

$$\boldsymbol{\sigma}_s = -p_s\mathbf{I} + \mu_s\big(2\mathcal{D}\mathbf{v}_s - \tfrac{2}{3}\nabla\cdot\mathbf{v}_s\mathbf{I}\big),$$

where p_s is the solid matrix pressure. Conservation of momentum for the mixture obeys the Stokes equation

$$\bar{\boldsymbol{\sigma}} = -\bar{p}\,\mathbf{I} + \mu_s(1-\phi)\big(2\mathcal{D}\mathbf{v}_s - \tfrac{2}{3}\nabla\cdot\mathbf{v}_s\mathbf{I}\big), \tag{2.5}$$

$$\nabla\cdot\bar{\boldsymbol{\sigma}} = -\bar{\rho}\,\mathbf{g}, \tag{2.6}$$

where $\bar{\boldsymbol{\sigma}}$ is the mixture stress. The mixture pressure $\bar{p}$ depends on the shear deformation of the solid matrix and the mixture buoyancy of the fluid and matrix.

The mechanical system is closed by relating the solid and fluid pressures through a compaction relation [35]

$$p_s - p_f = -\frac{\mu_s}{\phi}\nabla\cdot\mathbf{v}_s, \tag{2.7}$$

where $\zeta_s = \mu_s/\phi$ is the solid matrix bulk viscosity. That is, the shear viscosity is taken to be constant while the matrix bulk viscosity is proportional to the inverse of porosity. This equation can be interpreted as the compressibility of the matrix. If the solid matrix pressure is larger than the fluid pressure, the matrix volume should contract (i.e., $\nabla\cdot\mathbf{v}_s < 0$).

Introduce the pressure potentials

$$q_f = p_f - \rho_f g z \quad \text{and} \quad q_s = p_s - \rho_f g z, \tag{2.8}$$

where z is depth and indeed q_s is defined using the *fluid* density ρ_f. Also let $q = \bar{q}$ be the mixture potential

$$q = \bar{q} = \phi q_f + (1-\phi) q_s = q_s + \phi(q_f - q_s), \tag{2.9}$$

and note that

$$q_f - q_s = \frac{1}{1-\phi}(q_f - q). \tag{2.10}$$

Then (2.4) becomes (1.1) and (2.7) becomes (1.4). Expanding the conservation equation (2.3) and adding compaction (1.4) leads us to (1.2). Finally, (2.5)–(2.6) can be manipulated into (1.3) and (1.5).

3. Mixed Variational Formulations and Stability for Positive Porosity. In this section, we explore variational formulations of the equations that do not deal explicitly with the possibility that the porosity may vanish. Suppose that the domain Ω is a bounded, simply connected, Lipschitz domain in $\mathbb{R}^d$, $d = 1$, 2, or 3, with outward pointing unit normal vector ν. We impose for simplicity of exposition in this section homogeneous Neumann conditions on the fluid and solid velocity, i.e.,

$$\mathbf{u} \cdot \nu = 0 \quad \text{and} \quad \mathbf{v}_s = \mathbf{0} \quad \text{on } \partial\Omega. \tag{3.1}$$

Before deriving our weak formulations for (1.1)–(1.5), we review our notation for the Hilbert spaces wherein solutions can be found. The space $L^2(\Omega)$ consists of all square integrable, real-valued functions on Ω. It is equipped with the inner product

$$(u, v) = (u, v)_\Omega = \int_\Omega uv \, dx$$

and associated norm $\|u\| = (u,u)^{1/2}$. Denote by $H^1(\Omega)$ all square integrable functions with square integrable weak derivatives. This space has the corresponding norm $\|u\|_1 = \left\{\|u\|^2 + \|\nabla u\|^2\right\}^{1/2}$. Let $H(\mathrm{div};\Omega)$ denote all square integrable vector-valued functions with square integrable weak divergence, and equip it with the norm $\|\mathbf{u}\|_{H(\mathrm{div})} = \left\{\|\mathbf{u}\|^2 + \|\nabla \cdot \mathbf{u}\|^2\right\}^{1/2}$. Note that $(H^1(\Omega))^d \subset H(\mathrm{div};\Omega) \subset (L^2(\Omega))^d$.

For functions defined on a boundary ω, let $\langle \cdot, \cdot \rangle_\omega$ denote the $L^2(\omega)$ inner product or duality pairing. We can restrict functions in $H^1(\Omega)$ to the boundary $\partial\Omega$ using the trace lemma [1, 24]. The space of these restrictions is $H^{1/2}(\partial\Omega) \subset L^2(\partial\Omega)$, and we have the bound

$$\|u\|_{1/2,\partial\Omega} \le C_\Omega \|u\|_1. \tag{3.2}$$

We leave out a definition of the norm $\|u\|_{1/2,\partial\Omega}$ as we are only interested in its existence. A similar lemma holds for functions in $H(\mathrm{div};\Omega)$ [17], and

$$\|\mathbf{u} \cdot \nu\|_{-1/2,\partial\Omega} \le C_\Omega \|\mathbf{u}\|_{H(\mathrm{div})}, \tag{3.3}$$

where $\|\cdot\|_{-1/2,\partial\Omega}$ is the norm of the dual space of $H^{1/2}(\partial\Omega)$.

We also use two Banach spaces. The space $L^\infty(\Omega)$ consists of all essentially bounded functions on Ω equipped with the essential supremum norm $\|\cdot\|_{L^\infty(\Omega)}$. The space $W^{1,\infty}(\Omega)$ consists of the functions in $L^\infty(\Omega)$ that have weak derivatives also in $L^\infty(\Omega)$, and the norm is $\|\cdot\|_{W^{1,\infty}(\Omega)} = \|\cdot\|_{L^\infty(\Omega)} + \|\nabla(\cdot)\|_{(L^\infty(\Omega))^d}$.

3.1. A standard weak formulation for positive porosity. Define the function spaces

$$\begin{aligned}
\mathbb{V}_r &= H(\mathrm{div};\Omega),\\
\mathbb{V}_{r,0} &= H_0(\mathrm{div};\Omega) = \big\{\mathbf{v}\in\mathbb{V}_r : \mathbf{v}\cdot\nu = 0 \text{ on } \Gamma^f_{\mathrm{ess}}\big\},\\
\mathbb{V}_s &= (H^1(\Omega))^d,\\
\mathbb{V}_{s,\mathbf{0}} &= (H^1_0(\Omega))^d = \big\{\mathbf{v}\in\mathbb{V}_s : \mathbf{v} = \mathbf{0} \text{ on } \partial\Omega\big\},\\
\mathbb{W} = \mathbb{W}_f &= L^2(\Omega),\\
\mathbb{W}_0 &= L^2(\Omega)/\mathbb{R} = \Big\{w\in\mathbb{W} : \int_\Omega w\,dx = 0\Big\},
\end{aligned}$$

each with its natural norm. Our first three weak formulations of equations (1.1)–(1.5) will ignore the possibility of ϕ vanishing, so we can freely divide by ϕ. In practice, we might approximate ϕ by adding a small positive constant to it.

Standard formulation. Find $\mathbf{u}\in\mathbb{V}_{r,0}$, $q_f\in\mathbb{W}_f$, $\mathbf{v}_s\in\mathbb{V}_{s,\mathbf{0}}$, and $q\in\mathbb{W}_0$ such that

$$\Big(\frac{\mu_f}{k_0}\phi^{-2-2\Theta}\mathbf{u},\boldsymbol{\psi}_r\Big) - (q_f,\nabla\cdot\boldsymbol{\psi}_r) = 0 \qquad \forall\boldsymbol{\psi}_r\in\mathbb{V}_{r,0}, \tag{3.4}$$

$$(\nabla\cdot\mathbf{u},w_f) + \Big(\frac{\phi}{\mu_s(1-\phi)}(q_f-q),w_f\Big) = 0 \qquad \forall w_f\in\mathbb{W}_f, \tag{3.5}$$

$$-(q,\nabla\cdot\boldsymbol{\psi}_s) + (\hat{\boldsymbol{\sigma}}(\mathbf{v}_s),\nabla\boldsymbol{\psi}_s) = -\big((1-\phi)\rho_r\mathbf{g},\boldsymbol{\psi}_s\big) \qquad \forall\boldsymbol{\psi}_s\in\mathbb{V}_{s,\mathbf{0}}, \tag{3.6}$$

$$(\nabla\cdot\mathbf{v}_s,w) - \Big(\frac{\phi}{\mu_s(1-\phi)}(q_f-q),w\Big) = 0 \qquad \forall w\in\mathbb{W}_0. \tag{3.7}$$

These equations are derived as follows. Multiply through by the coefficient in (1.1), multiply by the vector test function $\boldsymbol{\psi}_r\in\mathbb{V}_{r,0}$, integrate over Ω, and use integration by parts with the boundary condition (3.1) to derive the first equation (3.4). Multiply (1.2) and (1.4) by $w_f\in\mathbb{W}_f$ and $w\in\mathbb{W}_0$, respectively, and integrate to derive (3.5) and (3.7). Finally, (3.6) comes from (1.3) multiplied by $\boldsymbol{\psi}_s\in\mathbb{V}_{s,\mathbf{0}}$, integrated, and integrated by parts (using (3.1)).

3.2. Stability/energy estimates. If porosity is allowed to vanish, the standard theory for proving existence and uniqueness of a solution to (3.4)–(3.7) breaks down. However, we can still obtain stability estimates when $\phi\ge\phi_*>0$ and investigate the behavior of the solution as ϕ_* approaches zero. We will assume that $\phi\le\phi^*<1$ in the analysis, since the model breaks down as ϕ approaches one.

Proceeding formally, assume a solution to (3.4)–(3.7) exists. We begin with an estimate of $(\hat{\boldsymbol{\sigma}}(\mathbf{v}_s),\nabla\mathbf{v}_s)$. Using the definition (1.5), this term is

$$\begin{aligned}
(\hat{\boldsymbol{\sigma}}(\mathbf{v}_s),\nabla\mathbf{v}_s) &= \big(2\mu_s(1-\phi)(\mathcal{D}\mathbf{v}_s - \tfrac{1}{3}\nabla\cdot\mathbf{v}_s\mathbf{I}),\nabla\mathbf{v}_s\big)\\
&= 2\mu_s\big\{\big((1-\phi)\,\mathcal{D}\mathbf{v}_s,\mathcal{D}\mathbf{v}_s\big) - \tfrac{1}{3}((1-\phi)\nabla\cdot\mathbf{v}_s,\nabla\cdot\mathbf{v}_s)\big\}.
\end{aligned}$$

We conclude that

$$(\hat{\boldsymbol{\sigma}}(\mathbf{v}_s),\nabla\mathbf{v}_s)\ge C\|\mathcal{D}\mathbf{v}_s\|^2,$$

for some positive constant C. An application on Korn's inequality [23, 16] results in

$$(\hat{\boldsymbol{\sigma}}(\mathbf{v}_s),\nabla\mathbf{v}_s)\ge C\|\mathcal{D}\mathbf{v}_s\|^2\ge C_1\|\mathbf{v}_s\|_1^2. \tag{3.8}$$

Testing (3.4)–(3.7) with $\boldsymbol{\psi}_r = \mathbf{u}$, $w_f = q_f$, $\boldsymbol{\psi}_s = \mathbf{v}_s$, and $w = q$ and adding the equations leads to four terms canceling, leaving only

$$\Big(\frac{\mu_f}{k_0}\phi^{-2-2\Theta}\mathbf{u},\mathbf{u}\Big) + \Big(\frac{\phi}{\mu_s(1-\phi)}(q_f-q), q_f-q\Big) + (\hat{\boldsymbol{\sigma}}(\mathbf{v}_s), \nabla\mathbf{v}_s) \\ = -\big((1-\phi)\rho_r\mathbf{g}, \mathbf{v}_s\big).$$

The result (3.8) and the elementary inequality $ab \le \epsilon a^2 + b^2/4\epsilon$ give us

$$\|\phi^{-1-\Theta}\mathbf{u}\| + \|\phi^{1/2}(q_f - q)\| + \|\mathbf{v}_s\|_1 \le \epsilon\|\mathbf{v}_s\| + C|\rho_r|, \tag{3.9}$$

for any $\epsilon > 0$ and some constant C depending on ϵ but *not* on porosity.

We will assume that $-2\phi^{1+\Theta}\nabla\phi^{-1/2} = \phi^{\Theta-1/2}\nabla\phi$ is bounded, so that

$$\begin{aligned}\|\phi^{-1/2}\mathbf{u}\|_{H(\mathrm{div})} &\le C\big\{\|\phi^{-1/2}\mathbf{u}\| + \|\nabla\cdot(\phi^{-1/2}\mathbf{u})\|\big\}\\ &\le C\big\{\|\phi^{-1/2}\mathbf{u}\| + \|\nabla\phi^{-1/2}\cdot\mathbf{u}\| + \|\phi^{-1/2}\nabla\cdot\mathbf{u}\|\big\}\\ &\le C\big\{\|\phi^{-1/2}\mathbf{u}\| + \|\phi^{-3/2}\nabla\phi\cdot\mathbf{u}\| + \|\phi^{-1/2}\nabla\cdot\mathbf{u}\|\big\}\\ &\le C\big\{\|\phi^{-1-\Theta}\mathbf{u}\| + \|\phi^{-1/2}\nabla\cdot\mathbf{u}\|\big\}.\end{aligned} \tag{3.10}$$

Substituting the test function $w_f = \phi^{-1}\nabla\cdot\mathbf{u}$ in (3.5) gives

$$(\nabla\cdot\mathbf{u}, \phi^{-1}\nabla\cdot\mathbf{u}) = -\Big(\frac{\phi}{\mu_s(1-\phi)}(q_f-q), \phi^{-1}\nabla\cdot\mathbf{u}\Big),$$

and therefore

$$\|\phi^{-1/2}\nabla\cdot\mathbf{u}\| \le C\|\phi^{1/2}(q_f - q)\|. \tag{3.11}$$

We recall the inf-sup condition [23, 17, 16, 15] for the Stokes problem. There exists $\gamma_{\mathrm{S}} > 0$ such that for any $w \in \mathbb{W}_0 = L^2(\Omega)/\mathbb{R}$,

$$\sup_{\boldsymbol{\psi}_s\in\mathbb{V}_{s,\mathbf{0}}} \frac{(w, \nabla\cdot\boldsymbol{\psi}_s)}{\|\boldsymbol{\psi}_s\|_1} \ge \gamma_{\mathrm{S}}\|w\|. \tag{3.12}$$

Substituting (3.6) into (3.12) shows

$$\|q\| \le C \sup_{\boldsymbol{\psi}_s\in\mathbb{V}_{s,\mathbf{0}}} \frac{1}{\|\boldsymbol{\psi}_s\|_1}\big\{|(\hat{\boldsymbol{\sigma}}, \nabla\boldsymbol{\psi}_s)| + \big|\big((1-\phi)\rho_r\mathbf{g}, \boldsymbol{\psi}_s\big)\big|\big\} \le C(\|\mathbf{v}_s\|_1 + |\rho_r|). \tag{3.13}$$

For the fluid potential, by the triangle inequality,

$$\|\phi^{1/2}q_f\| \le \|\phi^{1/2}q\| + \|\phi^{1/2}(q_f - q)\| \le \|q\| + \|\phi^{1/2}(q_f - q)\|. \tag{3.14}$$

Combining (3.9)–(3.11) and (3.13)–(3.14), we have the stability estimates

$$\|\phi^{-1-\Theta}\mathbf{u}\| + \|\phi^{-1/2}\nabla\cdot\mathbf{u}\| + \|\phi^{1/2}q_f\| + \|\mathbf{v}_s\|_1 + \|q\| \le C|\rho_r|. \tag{3.15}$$

The solid matrix potential is bounded by

$$\|q_s\| \le \|q\| + \|\phi(q_f - q_s)\| \le \|q\| + \|\phi^{1/2}q_f\| + \phi^*\|q_s\|,$$

which implies

$$\|q_s\| \le C\big\{\|q\| + \|\phi^{1/2}q_f\|\big\}. \tag{3.16}$$

We conclude that the two velocities and the solid matrix pressure remain stable, i.e., they are bounded, as the porosity approaches zero. However, in this case the fluid potential *may* become unbounded. This potential loss of stability is a significant issue for numerical modeling, since in many cases (such as the mid-ocean ridge application), we expect a well-defined melted region bounded by regions containing no melt, i.e., regions where $\phi = 0$.

3.3. Two alternative weak formulations. One can envision alternatives to the standard weak formulation (3.4)–(3.7) to avoid dividing by the porosity. A commonly used approach, the expanded mixed formulation [7], introduces an auxiliary velocity variable

$$\tilde{\mathbf{v}}_r^{\mathrm{ex}} = -\frac{k_0}{\mu_f}\nabla q_f \quad \Longleftrightarrow \quad \mathbf{u} = \phi^{2+2\Theta}\tilde{\mathbf{v}}_r^{\mathrm{ex}}. \tag{3.17}$$

In this approach, we use a weak form of this equation and modify (3.4).

Expanded formulation. Find $\mathbf{u} \in \mathbb{V}_{r,0}$, $\tilde{\mathbf{v}}_r^{\mathrm{ex}} \in \mathbb{V}_{r,0}$, $q_f \in \mathbb{W}_f$, $\mathbf{v}_s \in \mathbb{V}_{s,\mathbf{0}}$, and $q \in \mathbb{W}_0$ such that (3.5)–(3.7) hold as well as

$$\left(\phi^{2+2\Theta}\tilde{\mathbf{v}}_r^{\mathrm{ex}}, \tilde{\boldsymbol{\psi}}_r\right) - (\mathbf{u}, \tilde{\boldsymbol{\psi}}_r) \quad = 0 \quad \forall\, \tilde{\boldsymbol{\psi}}_r \in \mathbb{V}_{r,0}, \tag{3.18}$$

$$\left(\frac{\mu_f}{k_0}\tilde{\mathbf{v}}_r^{\mathrm{ex}}, \boldsymbol{\psi}_r\right) - (q_f, \nabla\cdot\boldsymbol{\psi}_r) = 0 \quad \forall\, \boldsymbol{\psi}_r \in \mathbb{V}_{r,0}. \tag{3.19}$$

Alternatively, we could balance the degeneracy by using a square-root scaling of the coefficient $\phi^{2+2\Theta}$ in (1.1) and the corresponding square-root scaling modification of (1.2), as was done in, e.g., [19]. However, we would still need to divide by ϕ in a standard approach, so we use the idea in [5] to modify the test function to avoid division by ϕ. To be precise, we replace the Darcy velocity $\mathbf{u} = \phi\mathbf{v}_r$ by the *scaled relative velocity*

$$\tilde{\mathbf{v}}_r = \phi^{-\Theta}\mathbf{v}_r = -\frac{k_0}{\mu_f}\phi^{1+\Theta}\nabla q_f \quad \Longleftrightarrow \quad \mathbf{u} = \phi^{1+\Theta}\tilde{\mathbf{v}}_r. \tag{3.20}$$

When taking the weak form of this equation we do *not* divide by $\phi^{1+\Theta}$. Rather, $\phi^{1+\Theta}$ multiplies the test function before we integrate by parts, as in [5].

Symmetry preserving formulation. Find $\tilde{\mathbf{v}}_r \in \mathbb{V}_{r,0}$, $q_f \in \mathbb{W}_f$, $\mathbf{v}_s \in \mathbb{V}_{s,\mathbf{0}}$, and $q \in \mathbb{W}_0$ such that (3.6)–(3.7) hold as well as

$$\left(\frac{\mu_f}{k_0}\tilde{\mathbf{v}}_r, \boldsymbol{\psi}_r\right) - \left(q_f, \nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi}_r)\right) = 0 \qquad \forall\, \boldsymbol{\psi}_r \in \mathbb{V}_{r,0}, \tag{3.21}$$

$$\left(\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_r), w_f\right) + \left(\frac{\phi}{\mu_s(1-\phi)}(q_f - q), w_f\right) = 0 \quad \forall w_f \in \mathbb{W}_f. \tag{3.22}$$

We call this the symmetry preserving formulation, since the degeneracy in ϕ appears in a symmetric way between (3.21) and (3.22). While it is perhaps somewhat unorthodox to leave the porosity multiplying the test functions, this approach should be gentler numerically, because it spreads out the porosity over two equations. Of course, we have not really handled the degeneracy, since now (3.22) degenerates to $0 = 0$ as porosity approaches zero. In fact, none of the three methods address the degeneracy in a direct way, nor, therefore, the issue of having no stability control of the fluid potential.

4. A Scaled Mixed Variational Formulation. Following [5], the stability estimates (3.15) suggest a new weak formulation that respects the possibility that the porosity might vanish. We have that the scaled relative velocity $\tilde{\mathbf{v}}_r = \phi^{-1-\Theta}\mathbf{u}$ is stable. If we define the *scaled fluid potential*

$$\tilde{q}_f = \phi^{1/2} q_f, \tag{4.1}$$

then this quantity is also stable. Therefore we reformulate the problem in terms of these scaled quantities, which means that (1.1)–(1.4) become

$$\tilde{\mathbf{v}}_r + \frac{k_0\phi^{1+\Theta}}{\mu_f}\nabla(\phi^{-1/2}\tilde{q}_f) = 0, \tag{4.2}$$

$$\mu_s\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_r) + \frac{1}{1-\phi}\big(\tilde{q}_f - \phi^{1/2}q\big) = 0, \tag{4.3}$$

$$\nabla q - \nabla\cdot\hat{\boldsymbol{\sigma}}(\mathbf{v}_s) = -(1-\phi)\,\rho_r\,\mathbf{g}, \tag{4.4}$$

$$\mu_s\nabla\cdot\mathbf{v}_s - \frac{\phi^{1/2}}{1-\phi}\big(\tilde{q}_f - \phi^{1/2}q\big) = 0, \tag{4.5}$$

wherein we have scaled the entire second equation by $\phi^{-1/2}$ and (1.5) continues to define $\hat{\boldsymbol{\sigma}}$.

The scaled equations make sense provided that the gradient and divergence terms are well-defined when $\phi = 0$. The divergence term in (4.3) expands to

$$\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_r) = \phi^{1/2+\Theta}\nabla\cdot\tilde{\mathbf{v}}_r + \phi^{\Theta-1/2}\nabla\phi\cdot\tilde{\mathbf{v}}_r, \tag{4.6}$$

and it is well defined provided that, for example,

$$\phi^{\Theta-1/2}\nabla\phi \in (L^\infty(\Omega))^d, \tag{4.7}$$

i.e., $|\phi^{\Theta-1/2}\nabla\phi|$ is bounded almost everywhere in the sense of Lebesgue measure. The gradient terms in (4.2) make sense under the same condition. The porosity ϕ in the physical model satisfies the full set of equations, including solute and thermal transport equations. It is *not* clear if we should expect that this porosity satisfies our condition. Nevertheless, we will tacitly assume that this condition holds. Our numerical results suggest that it is not strictly necessary, and perhaps can be weakened (see also [5] for a discussion of the necessity of this condition).

4.1. Boundary conditions and the scaled formulation. We should not expect the scaled velocity to lie in $\mathbb{V}_r$. Rather, $\tilde{\mathbf{v}}_r$ should lie in the space

$$\tilde{\mathbb{V}}_r = H_\phi(\mathrm{div};\Omega) = \big\{\mathbf{v}\in(L^2(\Omega))^d : \phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\mathbf{v})\in L^2(\Omega)\big\}.$$

As discussed in [5], this is a Hilbert space with the inner product

$$(\tilde{\mathbf{u}},\tilde{\mathbf{v}})_{\tilde{\mathbb{V}}_r} = (\tilde{\mathbf{u}},\tilde{\mathbf{v}}) + \big(\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{u}}), \phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}})\big)$$

and induced norm $\|\tilde{\mathbf{v}}\|_{\tilde{\mathbb{V}}_r} = (\tilde{\mathbf{v}},\tilde{\mathbf{v}})^{1/2}_{\tilde{\mathbb{V}}_r}$. Moreover, these vector functions have a well-defined normal trace on $\partial\Omega$, and, similar to (3.3),

$$\|\phi^{1/2+\Theta}\tilde{\mathbf{v}}\cdot\nu\|_{-1/2,\partial\Omega} \le C_\Omega\|\tilde{\mathbf{v}}\|_{\tilde{\mathbb{V}}_r}. \tag{4.8}$$

We also have the space $H_\phi^{-1/2}(\partial\Omega)$, which is the image of this normal trace operator on $\tilde{\mathbb{V}}_r = H_\phi(\mathrm{div})$.

The boundary conditions (3.1) will now be made more general (see [38] for even more generality). In terms of the scaled quantities, we take nonhomogeneous boundary conditions on the Darcy system that are natural (with $\Gamma^f_{\mathrm{nat}} = \partial\Omega$ and $\Gamma^f_{\mathrm{ess}} = \emptyset$) or essential ($\Gamma^f_{\mathrm{ess}} = \partial\Omega$ and $\Gamma^f_{\mathrm{nat}} = \emptyset$) of the form

$$\tilde{q}_f = \phi^{1/2}g_f = \tilde{g}_f \quad \text{on } \Gamma^f_{\mathrm{nat}}, \tag{4.9}$$

$$\phi^{1/2+\Theta}\tilde{\mathbf{v}}_r\cdot\nu = g_r \quad \text{on } \Gamma^f_{\mathrm{ess}}, \tag{4.10}$$

and we impose nonhomogeneous essential conditions on the Stokes system

$$\mathbf{v}_s = \mathbf{g}_s \quad \text{on } \partial\Omega. \tag{4.11}$$

A compatibility condition may be needed:

$$\int_{\partial\Omega} (g_r + \mathbf{g}_s \cdot \nu)\, ds = 0 \quad \text{if } \Gamma^f_{\text{ess}} = \partial\Omega. \tag{4.12}$$

To impose the essential boundary conditions, we define the spaces

$$\tilde{\mathbb{V}}_{r,g_r} = \big\{\mathbf{v} \in \tilde{\mathbb{V}}_r : \phi^{1/2+\Theta}\mathbf{v}\cdot\nu = g_r \text{ on } \Gamma^f_{\text{ess}}\big\},$$
$$\mathbb{V}_{s,\mathbf{g}_s} = \big\{\mathbf{v} \in (H^1(\Omega))^d : \mathbf{v} = \mathbf{g}_s \text{ on } \partial\Omega\big\},$$

and we let $\mathbb{W}_*$ be $\mathbb{W}_0$ if $\Gamma_{\text{ess}} = \partial\Omega$ or ϕ vanishes entirely (i.e., the pressure scale is not set by the boundary conditions) and $\mathbb{W}$ otherwise. As with the other formulations, a similar procedure leads us to the following scaled weak formulation.

Scaled formulation. Find $\tilde{\mathbf{v}}_r \in \tilde{\mathbb{V}}_{r,g_r}$, $\tilde{q}_f \in \mathbb{W}_f$, $\mathbf{v}_s \in \mathbb{V}_{s,\mathbf{g}_s}$, and $q \in \mathbb{W}_*$ such that

$$\begin{aligned}
&\Big(\frac{\mu_f}{k_0}\tilde{\mathbf{v}}_r, \boldsymbol{\psi}_r\Big) - \big(\tilde{q}_f, \phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi}_r)\big) \\
&\quad = -\langle \tilde{g}_f, \phi^{1/2+\Theta}\boldsymbol{\psi}_r\cdot\nu\rangle_{\Gamma^f_{\text{nat}}} && \forall\, \boldsymbol{\psi}_r \in \tilde{\mathbb{V}}_{r,0}, && (4.13)\\
&\big(\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_r), w_f\big) + \Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_f - \phi^{1/2}q), w_f\Big) \\
&\quad = 0 && \forall w_f \in \mathbb{W}_f, && (4.14)\\
&-(q, \nabla\cdot\boldsymbol{\psi}_s) + (\hat{\boldsymbol{\sigma}}(\mathbf{v}_s), \nabla\boldsymbol{\psi}_s) = -\big((1-\phi)\rho_r\mathbf{g}, \boldsymbol{\psi}_s\big) && \forall\, \boldsymbol{\psi}_s \in \mathbb{V}_{s,\mathbf{0}}, && (4.15)\\
&(\nabla\cdot\mathbf{v}_s, w) - \Big(\frac{\phi^{1/2}}{\mu_s(1-\phi)}(\tilde{q}_f - \phi^{1/2}q), w\Big) = 0 && \forall w \in \mathbb{W}_*. && (4.16)
\end{aligned}$$

4.2. Existence and uniqueness of the solution. We can derive stability estimates similar to (3.15) for the scaled formulation. Our proof combines four key results: the Babuška-Lax-Milgram Theorem [9, 10, 31] (Theorem 2 stated below), Korn's inequality in the form (3.8), the Stokes inf-sup condition (3.12), and a special inf-sup condition for the scaled Darcy system (Lemma (3) below). We also use some of the concepts developed in [5], where the degenerate Darcy problem is analyzed.

To impose essential boundary conditions, we assume that $\mathbf{g}_s \in (H^{1/2}(\partial\Omega))^d$ and extend it continuously from the boundary into the domain, so that the extension $\mathbf{g}_s \in \mathbb{V}_s = (H^1(\Omega))^d$ and $\|\mathbf{g}_s\|_1 \le C\|\mathbf{g}_s\|_{1/2,\partial\Omega}$. In a similar way, following [5], we assume that $g_r \in H_\phi^{-1/2}(\Gamma^f_{\text{ess}})$, the image of the scaled normal trace operator on $\tilde{\mathbb{V}}_r = H_\phi(\text{div};\Omega)$ which appears in (4.8). Then g_r has a bounded extension $\mathbf{g}_r \in \tilde{\mathbb{V}}_r$ on Ω such that

$$\phi^{1/2+\Theta}\mathbf{g}_r\cdot\nu = g_r \quad \text{on } \Gamma^f_{\text{ess}}.$$

THEOREM 1. *Assume that* (4.7) *holds on the porosity,* $0 \le \phi \le \phi^* < 1$*, and the extensions* $\mathbf{g}_r \in \tilde{\mathbb{V}}_r$ *and* $\mathbf{g}_s \in \mathbb{V}_s$ *satisfy* (4.12)*. If* $\Gamma^f_{\text{nat}} = \emptyset$ *or* ϕ *is identically zero, let* $\mathbb{W}_* = \mathbb{W}_0$*; otherwise, let* $\mathbb{W}_* = \mathbb{W}$*,* $\tilde{g}_f \in H^{1/2}(\Gamma^f_{\text{nat}})$*,* $g_r = 0$*, and assume the stronger condition* (4.21) *on the porosity. Then there exists a unique solution to the scaled formulation* (4.13)–(4.16)*,* (1.5)*, and it satisfies*

$$\begin{aligned}
&\|\tilde{\mathbf{v}}_r\| + \|\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_r)\| + \|\tilde{q}_f\| + \|\mathbf{v}_s\|_1 + \|q\| \\
&\quad \le C\big\{|\rho_r| + \|\tilde{g}_f\|_{1/2,\Gamma^f_{\text{nat}}} + \|\mathbf{g}_r\|_{\tilde{\mathbb{V}}_r} + \|\mathbf{g}_s\|_1\big\}.
\end{aligned} \tag{4.17}$$

Before proving this theorem, we state two key results. The first is well-known [9, 10, 31], and the second will be proven after we prove Theorem 1.

THEOREM 2 (Babuška-Lax-Milgram). *Let U and V be two real Hilbert spaces. Suppose that $a : U \times V \to \mathbb{R}$ is a continuous bilinear functional such that for some constant $\gamma > 0$ and all $u \in U$ and $v \in V$, $v \neq 0$,*

$$\sup_{\|v\|=1} |a(u,v)| \geq \gamma \|u\| \quad \textit{and} \quad \sup_{\|u\|=1} |a(u,v)| > 0. \tag{4.18}$$

Then, for all $f \in V^*$, there exists a unique solution $u \in U$ to

$$a(u,v) = f(v) \quad \forall v \in V,$$

and

$$\|u\| \leq \frac{1}{\gamma}\|f\|. \tag{4.19}$$

The divergence operator satisfies an inf-sup condition [32, 17, 34]: there is some $\gamma_{\mathrm{D}} > 0$ such that

$$\sup_{\mathbf{v} \in H(\mathrm{div};\Omega)} \frac{(w, \nabla \cdot \mathbf{v})}{\|\mathbf{v}\|_{H(\mathrm{div};\Omega)}} \geq \gamma_{\mathrm{D}} \|w\| \quad \forall w \in L^2(\Omega). \tag{4.20}$$

LEMMA 3. *If*

$$\|(1+\Theta)\phi^{-1}\nabla\phi\|_{(L^\infty(\Omega))^d} < \gamma_{\mathrm{D}}, \tag{4.21}$$

then there is some $\hat{\gamma}_{\mathrm{D}} > 0$ such that

$$\sup_{\mathbf{v} \in H(\mathrm{div};\Omega)} \frac{\big(w, \phi^{-1/2}\nabla \cdot (\phi^{1+\Theta}\mathbf{v})\big)}{\|\mathbf{v}\|_{H(\mathrm{div};\Omega)}} \geq \hat{\gamma}_{\mathrm{D}} \|\phi^{1/2+\Theta} w\| \quad \forall w \in L^2(\Omega). \tag{4.22}$$

Proof of Theorem 1. Let

$$\mathbb{X} = \tilde{\mathbb{V}}_{r,0} \times \mathbb{W}_f \times \mathbb{V}_{s,\mathbf{0}} \times \mathbb{W}_*,$$

and take $U = V = \mathbb{X}$, which is indeed a real Hilbert space. The bilinear form is defined by the equations (4.13)–(4.16) for any $\mathbf{U} = (\tilde{\mathbf{v}}_{r,0}, \tilde{q}_f, \mathbf{v}_{s,0}, q) \in \mathbb{X}$ and $\boldsymbol{\Psi} = (\boldsymbol{\psi}_r, w_f, \boldsymbol{\psi}_s, w) \in \mathbb{X}$ as

$$\begin{aligned} a(\mathbf{U}, \boldsymbol{\Psi}) = &\Big(\frac{\mu_f}{k_0}\tilde{\mathbf{v}}_{r,0}, \boldsymbol{\psi}_r\Big) - \big(\tilde{q}_f, \phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi}_r)\big) + \big(\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_{r,0}), w_f\big) \\ &+ \Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_f - \phi^{1/2}q), w_f\Big) - \Big(\frac{\phi^{1/2}}{\mu_s(1-\phi)}(\tilde{q}_f - \phi^{1/2}q), w\Big) \\ &- (q, \nabla\cdot\boldsymbol{\psi}_s) + (\hat{\boldsymbol{\sigma}}(\mathbf{v}_{s,0}), \nabla\boldsymbol{\psi}_s) + (\nabla\cdot\mathbf{v}_{s,0}, w). \end{aligned}$$

The linear functional is

$$\begin{aligned} f(\boldsymbol{\Psi}) = &-\langle \tilde{g}_f, \phi^{1/2+\Theta}\boldsymbol{\psi}_r\cdot\nu\rangle_{\Gamma^f_{\mathrm{nat}}} - \big((1-\phi)\rho_r\mathbf{g}, \boldsymbol{\psi}_s\big) \\ &- \Big(\frac{\mu_f}{k_0}\mathbf{g}_r, \boldsymbol{\psi}_r\Big) - \big(\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\mathbf{g}_r), w_f\big) - (\hat{\boldsymbol{\sigma}}(\mathbf{g}_s), \nabla\boldsymbol{\psi}_s) - (\nabla\cdot\mathbf{g}_s, w). \end{aligned}$$

Clearly we have continuity (boundedness) of a on $\mathbb{X}\times\mathbb{X}$ and f on $\mathbb{X}$, using the trace result (4.8).

Our scaled formulation is written in the context of the Babuška-Lax-Milgram Theorem as follows. We find $\mathbf{U}\in\mathbb{X}$ such that

$$a(\mathbf{U},\mathbf{\Psi}) = f(\mathbf{\Psi}) \quad \forall \mathbf{\Psi}\in\mathbb{X}, \tag{4.23}$$

and then set $\tilde{\mathbf{v}}_r = \tilde{\mathbf{v}}_{r,0} + \mathbf{g}_r$ and $\mathbf{v}_s = \mathbf{v}_{s,0} + \mathbf{g}_s$.

We turn attention to the inf-sup condition, the first condition in (4.18). For any $q\in\mathbb{W}_*$, let

$$\bar{q} = \frac{1}{|\Omega|}\int_\Omega q\,dx$$

be the average of q (note that $\bar{q}=0$ if $\mathbb{W}_* = \mathbb{W}_0$). Using (3.12) for the Stokes problem, there is $\mathbf{v}_q\in\mathbb{V}_{s,\mathbf{0}}$ normalized so that $\|\mathbf{v}_q\|_1 = \|q-\bar{q}\|$ and satisfying

$$-(q,\nabla\cdot\mathbf{v}_q) \geq \tfrac{1}{2}\gamma_{\mathrm{S}}\|q-\bar{q}\|^2. \tag{4.24}$$

In case $\bar{q}\neq 0$ (i.e, $\Gamma^f_{\mathrm{nat}} = \partial\Omega$), we apply Lemma 3 to any $\tilde{q}_f\in\mathbb{W}_f$ to find $\mathbf{v}_{\tilde{q}_f}\in H(\mathrm{div};\Omega)$ such that $\|\mathbf{v}_{\tilde{q}_f}\|_{H(\mathrm{div};\Omega)} = \|\phi^{1/2+\Theta}\tilde{q}_f\|$ and

$$-\big(\tilde{q}_f,\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\mathbf{v}_q)\big) \geq \tfrac{1}{2}\hat{\gamma}_{\mathrm{D}}\|\phi^{1/2+\Theta}\tilde{q}_f\|^2. \tag{4.25}$$

Note that $\mathbf{v}_{\tilde{q}_f}\in H(\mathrm{div};\Omega)\subset\tilde{\mathbb{V}}_r = \tilde{\mathbb{V}}_{r,0}$ when $\Gamma^f_{\mathrm{ess}}=\emptyset$.

For any $\mathbf{U} = (\tilde{\mathbf{v}}_{r,0},\tilde{q}_f,\mathbf{v}_{s,0},q)\in\mathbb{X}$, we take the test function in (4.23) as $\mathbf{\Psi} = (\boldsymbol{\psi}_r, w_f, \boldsymbol{\psi}_s, w)\in\mathbb{X}$ defined by

$$\begin{gathered}\boldsymbol{\psi}_r = \tilde{\mathbf{v}}_{r,0} + \delta_0\mathbf{v}_{\tilde{q}_f}, \quad w_f = \tilde{q}_f + \delta_1\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_{r,0}),\\ \boldsymbol{\psi}_s = \mathbf{v}_{s,0} + \delta_2\mathbf{v}_q, \quad \text{and} \quad w = q,\end{gathered} \tag{4.26}$$

where $\delta_j>0$, $j=0,1,2$, will be determined below. After combining and canceling some terms,

$$\begin{aligned}a(\mathbf{U},\mathbf{\Psi}) = {}& \frac{\mu_f}{k_0}\|\tilde{\mathbf{v}}_{r,0}\|^2 + \delta_1\big\|\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_{r,0})\big\|^2 + \frac{1}{\mu_s}\Big\|\frac{1}{\sqrt{1-\phi}}\big(\tilde{q}_f-\phi^{1/2}q\big)\Big\|^2\\ &+ (\hat{\boldsymbol{\sigma}}(\mathbf{v}_{s,0}),\nabla\mathbf{v}_{s,0}) - \delta_2(q,\nabla\cdot\mathbf{v}_q) - \delta_0\big(\tilde{q}_f,\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\mathbf{v}_{\tilde{q}_f})\big)\\ &+ \delta_0\Big(\frac{\mu_f}{k_0}\tilde{\mathbf{v}}_{r,0},\mathbf{v}_{\tilde{q}_f}\Big) + \delta_1\Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_f-\phi^{1/2}q),\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_{r,0})\Big)\\ &+ \delta_2(\hat{\boldsymbol{\sigma}}(\mathbf{v}_{s,0}),\nabla\mathbf{v}_q).\end{aligned}$$

There is some $C_2>0$ such that

$$(\hat{\boldsymbol{\sigma}}(\mathbf{v}_{s,0}),\nabla\mathbf{v}_q) \leq C_2\|\mathbf{v}_{s,0}\|_1\|\mathbf{v}_q\|_1 = C_2\|\mathbf{v}_{s,0}\|_1\|q-\bar{q}\|,$$

so using (3.8) with its constant $C_1 > 0$, (4.24), and (4.25), we see that

$$\begin{aligned}
a(\mathbf{U},\mathbf{\Psi}) &\geq \frac{\mu_f}{k_0}\|\tilde{\mathbf{v}}_{r,0}\|^2 + \delta_1\big\|\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_{r,0})\big\|^2 + \frac{1}{\mu_s}\|\tilde{q}_f - \phi^{1/2}q\|^2 \\
&\quad + C_1\|\mathbf{v}_{s,0}\|_1^2 + \tfrac{1}{2}\delta_2\gamma_{\mathrm{S}}\|q-\bar{q}\|^2 + \tfrac{1}{2}\delta_0\hat{\gamma}_{\mathrm{D}}\|\phi^{1/2+\Theta}\tilde{q}_f\|^2 \\
&\quad + \delta_0\Big(\frac{\mu_f}{k_0}\tilde{\mathbf{v}}_{r,0},\mathbf{v}_{\tilde{q}_f}\Big) + \delta_1\Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_f-\phi^{1/2}q),\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_{r,0})\Big) \\
&\quad + \delta_2(\hat{\boldsymbol{\sigma}}(\mathbf{v}_{s,0}),\nabla\mathbf{v}_q) \\
&\geq \frac{\mu_f}{k_0}\Big(1-\delta_0\frac{\mu_f}{\hat{\gamma}_{\mathrm{D}}k_0}\Big)\|\tilde{\mathbf{v}}_{r,0}\|^2 + \tfrac{1}{2}\delta_1\big\|\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_{r,0})\big\|^2 \\
&\quad + \frac{1}{\mu_s}\Big(1-\frac{\delta_1}{2\mu_s(1-\phi^*)^2}\Big)\|\tilde{q}_f-\phi^{1/2}q\|^2 + \Big(C_1-\delta_2\frac{C_2^2}{\gamma_{\mathrm{S}}}\Big)\|\mathbf{v}_{s,0}\|_1^2 \\
&\quad + \tfrac{1}{4}\delta_2\gamma_{\mathrm{S}}\|q-\bar{q}\|^2 + \tfrac{1}{4}\delta_0\hat{\gamma}_{\mathrm{D}}\|\phi^{1/2+\Theta}\tilde{q}_f\|^2.
\end{aligned}$$

Taking δ_0, δ_1, and δ_2 positive but sufficiently small shows that for some $c > 0$,

$$\begin{aligned}
a(\mathbf{U},\mathbf{\Psi}) \geq c\big\{&\|\tilde{\mathbf{v}}_{r,0}\|^2 + \big\|\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_{r,0})\big\|^2 + \|\mathbf{v}_{s,0}\|_1^2 \\
&+ \|\tilde{q}_f-\phi^{1/2}q\|^2 + \|q-\bar{q}\|^2 + \|\phi^{1/2+\Theta}\tilde{q}_f\|^2\big\}.
\end{aligned}$$

In case $\bar{q} \neq 0$ (and then $\phi \not\equiv 0$), we estimate

$$\|q\| \leq \|q-\bar{q}\| + |\Omega|^{1/2}|\bar{q}| \leq \|q-\bar{q}\| + |\Omega|^{1/2}\frac{\|\phi^{1+\Theta}\bar{q}\|}{\|\phi^{1+\Theta}\|}$$

and

$$\begin{aligned}
\|\phi^{1+\Theta}\bar{q}\| &\leq \|\phi^{1+\Theta}(q-\bar{q})\| + \|\phi^{1+\Theta}q\| \\
&\leq \|\phi^{1+\Theta}(q-\bar{q})\| + \|\phi^{1/2+\Theta}\tilde{q}_f\| + \|\phi^{1/2+\Theta}(\tilde{q}_f-\phi^{1/2}q)\| \\
&\leq \|q-\bar{q}\| + \|\phi^{1/2+\Theta}\tilde{q}_f\| + \|\tilde{q}_f-\phi^{1/2}q\|,
\end{aligned}$$

so

$$\|q\| \leq C\big\{\|q-\bar{q}\| + \|\phi^{1/2+\Theta}\tilde{q}_f\| + \|\tilde{q}_f-\phi^{1/2}q\|\big\}.$$

Moreover,

$$\|\tilde{q}_f\| \leq \|\tilde{q}_f-\phi^{1/2}q\| + \|q\|,$$

and we have shown the first condition in (4.18). The second condition in (4.18) follows by symmetry. We have thus met the conditions of the Babuška-Lax-Milgram Theorem, and we conclude that the problem (4.13)–(4.16), (1.5) has a unique solution. Moreover, the bound (4.19) is what is written in Theorem 1. □

Proof of Lemma 3. Given $w \in L^2(\Omega)$, let $(\boldsymbol{\psi},\chi) \in H(\mathrm{div};\Omega)\times L^2(\Omega)$ solve

$$\nabla\cdot\boldsymbol{\psi} + (1+\Theta)\phi^{-1}\nabla\phi\cdot\boldsymbol{\psi} = \phi^{1/2+\Theta}w \quad \text{in } \Omega, \tag{4.27}$$

$$\boldsymbol{\psi} = -\nabla\chi \quad \text{in } \Omega, \tag{4.28}$$

$$\chi = 0 \quad \text{on } \partial\Omega. \tag{4.29}$$

This is possible provided (4.21) holds (by (4.20) and the Lax-Milgram Theorem). Moreover,

$$\|\boldsymbol{\psi}\|_{H(\mathrm{div};\Omega)} \leq \hat{C}\|\phi^{1/2+\Theta}w\|.$$

Finally,

$$\sup_{\mathbf{v}\in\mathbb{V}_{r,0}} \frac{\big(w, \phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\mathbf{v})\big)}{\|\mathbf{v}\|_{H(\mathrm{div};\Omega)}} \geq \frac{\big(w, \phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi})\big)}{\|\boldsymbol{\psi}\|_{H(\mathrm{div};\Omega)}}$$
$$= \frac{\big(\phi^{1/2+\Theta}w, \nabla\cdot\boldsymbol{\psi} + (1+\Theta)\phi^{-1}\nabla\phi\cdot\boldsymbol{\psi}\big)}{\|\boldsymbol{\psi}\|_{H(\mathrm{div};\Omega)}} \geq \frac{1}{\hat{C}}\frac{\|\phi^{1/2+\Theta}w\|^2}{\|\phi^{1/2+\Theta}w\|} = \frac{1}{\hat{C}}\|\phi^{1/2+\Theta}w\|. \quad \square$$

5. Mixed Finite Element Methods. Assume Ω is a polygonal domain in one, two, or three dimensions. Let $\mathcal{T}_h$ be a conforming finite element mesh of simplices or rectangular parallelepipeds covering Ω with maximal spacing h, and let $\mathcal{E}_h$ denote the set of element endpoints, edges, or faces.

To continue the exposition, we will restrict to two dimensions. Extension to one and three dimensions should be clear. Let $\mathbb{P}_n$ denote the space of polynomials of degree n and $\mathbb{P}_{n_1,n_2}$ denote the polynomials of degree n_1 in x and n_2 in z (taking the second coordinate to be the depth z).

5.1. Finite element spaces. To formulate the discrete versions of our weak formulations, we introduce two finite element spaces specially designed for mixed methods. The first space will be used for the Darcy part of the system. We choose the lowest order Raviart-Thomas (RT_0) finite element space $\mathbb{V}_{\mathrm{RT}} \times \mathbb{W}_{\mathrm{RT}}$ [32, 17, 34]. On an element $E \in \mathcal{T}_h$,

$$\mathbb{V}_{\mathrm{RT}}(E) = \begin{cases} (\mathbb{P}_0 \times \mathbb{P}_0) \oplus \binom{x}{z}\mathbb{P}_0, & \text{if } E \text{ is a triangle,} \\ \mathbb{P}_{1,0} \times \mathbb{P}_{0,1}, & \text{if } E \text{ is a rectangle,} \end{cases} \tag{5.1}$$

$$\mathbb{W}_{\mathrm{RT}}(E) = \mathbb{P}_0 \tag{5.2}$$

(so $\mathbb{W}_{\mathrm{RT}}$ is the space of piecewise constant functions). The degrees of freedom for $\mathbb{V}_{\mathrm{RT}}$ are the normal fluxes on the edges, and the degrees of freedom for $\mathbb{W}_{\mathrm{RT}}$ are the average values over the elements, i.e.,

$$\mathbb{V}_{\mathrm{RT}} = \mathrm{span}\Big\{\mathbf{v}_e : \int_f \mathbf{v}_e \cdot \nu_f \, ds = \delta_{e,f} \ \forall e, f \in \mathcal{E}_h\Big\}, \tag{5.3}$$

$$\mathbb{W}_{\mathrm{RT}} = \mathrm{span}\{w_E : w_E|_F = \delta_{E,F} \ \forall E, F \in \mathcal{T}_h\}, \tag{5.4}$$

where $\delta_{i,j}$ is the Kronecker delta function for indices i and j. Raviart-Thomas spaces are commonly used to solve Darcy's equation and RT_0 is first order accurate in $H(\mathrm{div};\Omega)$ for $\mathbb{V}_{\mathrm{RT}}$ and in $L^2(\Omega)$ for $\mathbb{W}_{\mathrm{RT}}$. We remark that we could use quadrilateral elements as well, as long as we substitute the Arbogast-Correa (AC_0) spaces [3] for the Raviart-Thomas spaces.

For the Stokes part of the system, we could choose any reasonable finite element space $\mathbb{V}_{\mathrm{S}} \times \mathbb{W}_{\mathrm{S}}$. A good choice on rectangular meshes is the Bernardi-Raugel (BR) space [14, 6] $\mathbb{V}_{\mathrm{BR}} \times \mathbb{W}_{\mathrm{BR}}$. On a rectangular element $E \in \mathcal{T}_h$,

$$\mathbb{V}_{\mathrm{BR}}(E) = \mathbb{P}_{1,2} \times \mathbb{P}_{2,1}, \tag{5.5}$$

and $\mathbb{W}_{\mathrm{BR}}(E) = \mathbb{W}_{\mathrm{RT}}(E)$. The degrees of freedom for $\mathbb{V}_{\mathrm{BR}}$ are the normal fluxes on the edges and the nodal values on the vertices of each component. The space BR is first order accurate in $(H^1(\Omega))^2$ for $\mathbb{V}_{\mathrm{BR}}$ and in $L^2(\Omega)$ for $\mathbb{W}_{\mathrm{BR}}$. The space was first introduced to solve Stokes equation, and has also been used to solve Darcy's equation

with continuous velocities [6]. It is a natural choice for our coupled Darcy-Stokes system, since the convergence rates of the two spaces match.

We could use more standard Stokes elements such as the lowest order Taylor-Hood elements [23, 17, 21]. On an element $E \in \mathcal{T}_h$, if E is triangular, $\mathbb{V}_{\mathrm{TH}}(E) = \mathbb{P}_2 \times \mathbb{P}_2$ and $\mathbb{W}_{\mathrm{TH}}(E) = \mathbb{P}_1$ and, if E is rectangular, $\mathbb{V}_{\mathrm{TH}}(E) = \mathbb{P}_{2,2} \times \mathbb{P}_{2,2}$ and $\mathbb{W}_{\mathrm{TH}}(E) = \mathbb{P}_{1,1}$. On rectangular meshes, these elements are more accurate than the BR elements, but they also have more degrees of freedom. However, we would not gain any additional overall convergence within the coupled system because of the Darcy part.

5.2. Finite element methods. Finite element methods for the standard, expanded, symmetric, and scaled formulations of Sections 3.1, 3.3, and 4 are given by restricting the solution and test functions to the chosen finite element spaces.

For completeness, we describe the scaled finite element method with the boundary conditions (4.9)–(4.11). In the case of essential boundary conditions, the extensions $\mathbf{g}_r$ and $\mathbf{g}_s$ are projected into the finite element spaces as $\hat{\mathbf{g}}_r \in \mathbb{V}_{\mathrm{RT}}$ and $\hat{\mathbf{g}}_s \in \mathbb{V}_{\mathrm{S}}$ ($\mathbb{V}_{\mathrm{BR}}$ or $\mathbb{V}_{\mathrm{TH}}$) in such a way that the two compatibility conditions hold:

$$\int_{\partial\Omega} (\hat{\mathbf{g}}_s - \mathbf{g}_s) \cdot \nu \, ds = 0 \quad \text{and} \quad \int_{\partial\Omega} (\hat{\mathbf{g}}_r + \hat{\mathbf{g}}_s) \cdot \nu \, ds = 0 \quad \text{if } \Gamma^f_{\mathrm{ess}} = \partial\Omega. \tag{5.6}$$

We also need to define

$$\begin{aligned} \mathbb{V}_{\mathrm{RT},0} &= \{\mathbf{v} \in \mathbb{V}_{\mathrm{RT}} : \mathbf{v} \cdot \nu = 0 \text{ on } \Gamma^f_{\mathrm{ess}}\}, \\ \mathbb{V}_{\mathrm{S},\mathbf{0}} &= \{\mathbf{v} \in \mathbb{V}_{\mathrm{S}} : \mathbf{v} = \mathbf{0} \text{ on } \partial\Omega\}, \\ \mathbb{W}_{\mathrm{S},0} &= \Big\{w \in \mathbb{W}_{\mathrm{S}} : \int_\Omega w \, dx = 0\Big\}, \end{aligned}$$

and $\mathbb{W}_{\mathrm{S},*} = \mathbb{W}_{\mathrm{S},0}$ if $\Gamma^f_{\mathrm{ess}} = \partial\Omega$ or $\phi \equiv 0$ and $\mathbb{W}_{\mathrm{S},*} = \mathbb{W}_{\mathrm{S}}$ otherwise.

Scaled mixed finite element method. Find $\tilde{\mathbf{v}}_{r,h} \in \mathbb{V}_{\mathrm{RT},0} + \hat{\mathbf{g}}_r$, $\tilde{q}_{f,h} \in \mathbb{W}_{\mathrm{RT}}$, $\mathbf{v}_{s,h} \in \mathbb{V}_{\mathrm{S},\mathbf{0}} + \hat{\mathbf{g}}_s$, and $q_h \in \mathbb{W}_{\mathrm{S},*}$ such that

$$\begin{aligned} &\Big(\frac{\mu_f}{k_0}\tilde{\mathbf{v}}_{r,h}, \boldsymbol{\psi}_r\Big) - \big(\tilde{q}_{f,h}, \phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi}_r)\big) \\ &\quad = -\langle \tilde{g}_f, \phi^{1/2+\Theta}\boldsymbol{\psi}_r \cdot \nu\rangle_{\Gamma^f_{\mathrm{nat}}} && \forall \boldsymbol{\psi}_r \in \mathbb{V}_{\mathrm{RT},0}, \quad (5.7) \\ &\big(\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_{r,h}), w_f\big) + \Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_{f,h} - \phi^{1/2}q_h), w_f\Big) \\ &\quad = 0 && \forall w_f \in \mathbb{W}_{\mathrm{RT}}, \quad (5.8) \\ &-(q_h, \nabla\cdot\boldsymbol{\psi}_s) + (\hat{\boldsymbol{\sigma}}(\mathbf{v}_{s,h}), \nabla\boldsymbol{\psi}_s) = -\big((1-\phi)\rho_r\mathbf{g}, \boldsymbol{\psi}_s\big) && \forall \boldsymbol{\psi}_s \in \mathbb{V}_{\mathrm{S},\mathbf{0}}, \quad (5.9) \\ &(\nabla\cdot\mathbf{v}_{s,h}, w) - \Big(\frac{\phi^{1/2}}{\mu_s(1-\phi)}(\tilde{q}_{f,h} - \phi^{1/2}q_h), w\Big) = 0 && \forall w \in \mathbb{W}_{\mathrm{S},*}, \quad (5.10) \end{aligned}$$

where the term $\hat{\boldsymbol{\sigma}}$ is defined by (1.5). While the scaled finite element method is well defined when ϕ vanishes due to the condition (4.7), it is important to avoid division by zero in the implementation. One must evaluate the two divergence terms containing ϕ to a negative power in (5.7)–(5.8) at quadrature points. Because the divergence terms scale with ϕ to the overall power $1/2+\Theta > 0$, these terms should be set to zero when ϕ vanishes. That is, at a quadrature point where $\phi = 0$, take the value of the entire term to be zero at that point.

LEMMA 4. *If* (4.7) *holds and either* $\mathbb{W}_{\mathrm{S},*} = \mathbb{W}_{\mathrm{S},0}$ *or* ϕ *is not identically zero, then there exists a unique solution to the scaled mixed finite element method* (5.7)–(5.10).

Proof. Since the scaled method gives rise to a square linear system when restricted to bases for the finite element spaces, the existence of a solution is equivalent to uniqueness. To show uniqueness, set to zero the quantities $\hat{\mathbf{g}}_r$, $\hat{\mathbf{g}}_s$, $\tilde{g}_f$, and $\mathbf{g}$. The test functions

$$\boldsymbol{\psi}_r = \tilde{\mathbf{v}}_{r,h} \in \mathbb{V}_{\mathrm{RT},0}, \quad w_f = \tilde{q}_{f,h} \in \mathbb{W}_{\mathrm{RT}}, \quad \boldsymbol{\psi}_s = \mathbf{v}_{s,h} \in \mathbb{V}_{\mathrm{S},\mathbf{0}} \quad \text{and} \quad w = q_h \in \mathbb{W}_{\mathrm{S},*},$$

when substituted into (5.7)–(5.10) and after the equations are added, imply that

$$\frac{\mu_f}{k_0}\|\tilde{\mathbf{v}}_{r,h}\|^2 + \Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_{f,h} - \phi^{1/2}q_h), \tilde{q}_{f,h} - \phi^{1/2}q_h\Big) + (\hat{\boldsymbol{\sigma}}(\mathbf{v}_{s,h}), \nabla \mathbf{v}_{s,h}) = 0.$$

Thus $\tilde{\mathbf{v}}_{r,h} = 0$, the estimate (3.8) shows $\mathbf{v}_{s,h} = 0$, and $\tilde{q}_{f,h} = \phi^{1/2}q_h$.

The discrete version of the inf-sup condition (3.12) holds for the BR and TH Stokes elements (and any inf-sup stable pair of Stokes elements) with a possibly smaller constant $0 < \gamma_{\mathrm{S}}^* \le \gamma_{\mathrm{S}}$ independent of the mesh spacing parameter h. Therefore there is some $\mathbf{v}_{q,h} \in \{\mathbf{v} \in \mathbb{V}_{\mathrm{S}} : \mathbf{v} = 0 \text{ on } \partial\Omega\} \subset \mathbb{V}_{\mathrm{S},\mathbf{0}}$ such that $\|\mathbf{v}_{q,h}\|_1 = \|q_h - \bar{q}_h\|$ and

$$-(q_h, \nabla \cdot \mathbf{v}_{q,h}) \ge \tfrac{1}{2}\gamma_{\mathrm{S}}^*\|q_h - \bar{q}_h\|^2, \tag{5.11}$$

where $\bar{q}_h = \int_\Omega q_h\,dx/|\Omega|$. The choice $\boldsymbol{\psi}_s = \mathbf{v}_{q,h}$ in (5.9) shows that $q_h = \bar{q}_h$. If $\mathbb{W}_{\mathrm{S},*} = \mathbb{W}_{\mathrm{S},0}$, $\bar{q}_h = 0$ and we conclude that $q_h = \tilde{q}_{h,f} = 0$. If $\mathbb{W}_{\mathrm{S},*} = \mathbb{W}_{\mathrm{S}}$, then since $\tilde{q}_{f,h} = \phi^{1/2}\bar{q}_h \in W_{\mathrm{RT}}$, we conclude that ϕ, being continuous, must be a positive constant. Similar to the Stokes case, we can then use the discrete version of the inf-sup condition (4.20) for the RT spaces and (5.7) to conclude that $\tilde{q}_h = q_h = 0$. Uniqueness, and therefore also existence, of a solution is established. ∎

5.3. Stability and convergence of the scaled method. In this subsection, we assume that the essential Neumann boundary conditions (4.10)–(4.11) are imposed (i.e., $\Gamma_{\mathrm{ess}}^f = \partial\Omega$). To derive a bound for the error, we first take the difference of (4.13)–(4.16) and (5.7)–(5.10) and add the resulting equations to see that

$$\begin{aligned}
&\Big(\frac{\mu_f}{k_0}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}), \boldsymbol{\psi}_r\Big) - \big(\tilde{q}_f - \tilde{q}_{f,h}, \phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi}_r)\big) \\
&\quad + \big(\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h})), w_f\big) \\
&\quad + \Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_f - \tilde{q}_{f,h} - \phi^{1/2}(q - q_h)), w_f - \phi^{1/2}w\Big) \\
&\quad - (q - q_h, \nabla\cdot\boldsymbol{\psi}_s) + (\hat{\boldsymbol{\sigma}}(\mathbf{v}_s - \mathbf{v}_{s,h}), \nabla\boldsymbol{\psi}_s) + (\nabla\cdot(\mathbf{v}_s - \mathbf{v}_{s,h}), w) = 0
\end{aligned} \tag{5.12}$$

for any $\boldsymbol{\psi}_r \in \mathbb{V}_{\mathrm{RT},0}$, $w_f \in \mathbb{W}_{\mathrm{RT}}$, $\boldsymbol{\psi}_s \in \mathbb{V}_{\mathrm{S},\mathbf{0}}$, and $w \in \mathbb{W}_{\mathrm{S},0}$.

Before defining our choice of test functions, we need the usual projection operators associated with RT_0 (or AC_0). Let $\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}} : L^2(\Omega) \to \mathbb{W}_{\mathrm{RT}}$ denote the $L^2(\Omega)$-projection operator mapping onto the space of piecewise constant functions $\mathbb{W}_{\mathrm{RT}}$. Let $\pi_{\mathrm{RT}} : H(\mathrm{div};\Omega) \cap L^{2+\epsilon}(\Omega) \to \mathbb{V}_{\mathrm{RT}}$ (any $\epsilon > 0$) denote the standard Raviart-Thomas or Fortin operator that preserves element average divergence and average edge normal fluxes [32, 17, 34, 3]. We also need the usual $H^1(\Omega)$-projection $\pi_{\mathrm{S}} : H^1(\Omega) \to \mathbb{V}_{\mathrm{S}}$ and the $L^2(\Omega)$-projection $\mathcal{P}_{\mathbb{W}_{\mathrm{S}}} : L^2(\Omega) \to \mathbb{W}_{\mathrm{S}}$.

Let the function $\mathbf{v}_{q,h} \in \mathbb{V}_{\mathrm{S},\mathbf{0}}$ arise from the discrete version of the inf-sup condition for Stokes (4.24) (as in (5.11)), normalized so that $\|\mathbf{v}_{q,h}\|_1 = \|\mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q - q_h\|$ and satisfying

$$-(\mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q - q_h, \nabla\cdot\mathbf{v}_{q,h}) \ge \tfrac{1}{2}\gamma_{\mathrm{S}}^*\|\mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q - q_h\|^2. \tag{5.13}$$

Similar to the test functions taken in (4.26), we take

$$\begin{aligned}
\boldsymbol{\psi}_r &= (\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}) - (\tilde{\mathbf{v}}_r - \pi_{\mathrm{RT}}\tilde{\mathbf{v}}_r) - (\pi_{\mathrm{RT}}\mathbf{g}_r - \hat{\mathbf{g}}_r) && \in \mathbb{V}_{\mathrm{RT},0},\\
w_f &= \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f - \tilde{q}_{f,h} + \delta_1\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}))\big] && \in \mathbb{W}_{\mathrm{RT}},\\
\boldsymbol{\psi}_s &= (\mathbf{v}_s - \mathbf{v}_{s,h}) - (\mathbf{v}_s - \pi_{\mathrm{S}}\mathbf{v}_s) - (\pi_{\mathrm{S}}\mathbf{g}_s - \hat{\mathbf{g}}_s) + \delta_2\mathbf{v}_{q,h} && \in \mathbb{V}_{\mathrm{S},\mathbf{0}},\\
w &= \mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q - q_h && \in \mathbb{W}_{\mathrm{S},0},
\end{aligned}$$

where $\delta_1 > 0$ and $\delta_2 > 0$ will be determined below. We remark that the term multiplying δ_1 must be projected back into the discrete space, and so our derivation is not completely straightforward.

Introducing $\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}$ thrice into (5.12) yields

$$\begin{aligned}
&\Big(\frac{\mu_f}{k_0}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}), \boldsymbol{\psi}_r\Big) - \big(\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f - \tilde{q}_{f,h}, \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi}_r)\big]\big)\\
&\quad + \big(\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}))\big], w_f\big)\\
&\quad - \big(\tilde{q}_f - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f, \phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi}_r)\big)\\
&\quad + \Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_f - \tilde{q}_{f,h} - \phi^{1/2}(q - q_h)), w_f - \phi^{1/2}w\Big)\\
&\quad - (q - q_h, \nabla\cdot\boldsymbol{\psi}_s) + (\hat{\boldsymbol{\sigma}}(\mathbf{v}_s - \mathbf{v}_{s,h}), \nabla\boldsymbol{\psi}_s) + (\nabla\cdot(\mathbf{v}_s - \mathbf{v}_{s,h}), w)\\
&= T_1 + \cdots + T_8 \text{ (respectively)} = 0.
\end{aligned} \tag{5.14}$$

For the first term in (5.14), we deduce that for some generic constant $C > 0$,

$$\begin{aligned}
T_1 &= \Big(\frac{\mu_f}{k_0}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}), \boldsymbol{\psi}_r\Big)\\
&= \frac{\mu_f}{k_0}\|\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}\|^2 - \Big(\frac{\mu_f}{k_0}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}), \tilde{\mathbf{v}}_r - \pi_{\mathrm{RT}}\tilde{\mathbf{v}}_r + \pi_{\mathrm{RT}}\mathbf{g}_r - \hat{\mathbf{g}}_r\Big)\\
&\geq \frac{\mu_f}{2k_0}\|\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}\|^2 - C\big\{\|\tilde{\mathbf{v}}_r - \pi_{\mathrm{RT}}\tilde{\mathbf{v}}_r\|^2 + \|\pi_{\mathrm{RT}}\mathbf{g}_r - \hat{\mathbf{g}}_r\|^2\big\}.
\end{aligned}$$

For the next two terms, for any $\epsilon > 0$,

$$\begin{aligned}
T_2 + T_3 &= -\big(\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f - \tilde{q}_{f,h}, \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi}_r)\big]\big)\\
&\quad + \big(\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}))\big], w_f\big)\\
&= \big(\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f - \tilde{q}_{f,h}, \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r - \pi_{\mathrm{RT}}\tilde{\mathbf{v}}_r + \pi_{\mathrm{RT}}\mathbf{g}_r - \hat{\mathbf{g}}_r)\big]\big)\\
&\quad + \delta_1\big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}))\big]\big\|^2\\
&\geq \delta_1\big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}))\big]\big\|^2 - \epsilon\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f - \tilde{q}_{f,h}\|^2\\
&\quad - C\big\{\big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r - \pi_{\mathrm{RT}}\tilde{\mathbf{v}}_r))\big]\big\|^2\\
&\qquad + \big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\pi_{\mathrm{RT}}\mathbf{g}_r - \hat{\mathbf{g}}_r))\big]\big\|^2\big\}.
\end{aligned}$$

Skipping T_4 for the moment, the next term is

$$\begin{aligned}
T_5 &= \Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_f - \tilde{q}_{f,h} - \phi^{1/2}(q-q_h)), w_f - \phi^{1/2}w\Big) \\
&= \Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_f - \tilde{q}_{f,h} - \phi^{1/2}(q-q_h)), \tilde{q}_f - \tilde{q}_{f,h} - \phi^{1/2}(q-q_h)\Big) \\
&\quad - \Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_f - \tilde{q}_{f,h} - \phi^{1/2}(q-q_h)), \tilde{q}_f - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f - \phi^{1/2}(q-\mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q)\Big) \\
&\quad + \delta_1\Big(\frac{1}{\mu_s(1-\phi)}(\tilde{q}_f - \tilde{q}_{f,h} - \phi^{1/2}(q-q_h)), \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}))\big]\Big) \\
&\geq \frac{1}{2\mu_s(1-\phi^*)}\|\tilde{q}_f - \tilde{q}_{f,h} - \phi^{1/2}(q-q_h)\|^2 - C\big\{\|\tilde{q}_f - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f\|^2 \\
&\qquad\qquad + \|q - \mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q\|^2 + \delta_1^2\big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r - \tilde{\mathbf{v}}_{r,h}))\big]\big\|^2\big\}.
\end{aligned}$$

Noting that $w = q - q_h - (q - \mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q)$, the sixth and eighth terms satisfy

$$\begin{aligned}
T_6 + T_8 &= -(q-q_h, \nabla\cdot\boldsymbol{\psi}_s) + (\nabla\cdot(\mathbf{v}_s - \mathbf{v}_{s,h}), w) \\
&= (q-q_h, \nabla\cdot(\mathbf{v}_s - \pi_{\mathrm{S}}\mathbf{v}_s + \pi_{\mathrm{S}}\mathbf{g}_s - \hat{\mathbf{g}}_s)) - (\nabla\cdot(\mathbf{v}_s - \mathbf{v}_{s,h}), q - \mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q) \\
&\quad - \delta_2\big[(\mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q - q_h, \nabla\cdot\mathbf{v}_{q,h}) + (q - \mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q, \nabla\cdot\mathbf{v}_{q,h})\big].
\end{aligned}$$

Recalling (5.13) and that $\|\mathbf{v}_{q,h}\|_1 = \|\mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q - q_h\|$, we have

$$\begin{aligned}
T_6 + T_8 &\geq \tfrac{1}{4}\delta_2\gamma_{\mathrm{S}}^*\|\mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q - q_h\|^2 - \tfrac{1}{8}\delta_2\gamma_{\mathrm{S}}^*\|q-q_h\|^2 - \epsilon\|\nabla\cdot(\mathbf{v}_s - \mathbf{v}_{s,h})\|^2 \\
&\quad - C\big\{\|q - \mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q\|^2 + \|\nabla\cdot(\mathbf{v}_s - \pi_{\mathrm{S}}\mathbf{v}_s)\|^2 + \|\nabla\cdot(\pi_{\mathrm{S}}\mathbf{g}_s - \hat{\mathbf{g}}_s)\|^2\big\} \\
&\geq \tfrac{1}{8}\delta_2\gamma_{\mathrm{S}}^*\|q-q_h\|^2 - \epsilon\|\nabla\cdot(\mathbf{v}_s - \mathbf{v}_{s,h})\|^2 \\
&\quad - C\big\{\|q - \mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q\|^2 + \|\nabla\cdot(\mathbf{v}_s - \pi_{\mathrm{S}}\mathbf{v}_s)\|^2 + \|\nabla\cdot(\pi_{\mathrm{S}}\mathbf{g}_s - \hat{\mathbf{g}}_s)\|^2\big\}.
\end{aligned}$$

Finally, for the next to last term, note that $\boldsymbol{\psi}_s = \pi_{\mathrm{S}}\mathbf{v}_s - \mathbf{v}_{s,h} - \pi_{\mathrm{S}}\mathbf{g}_s + \hat{\mathbf{g}}_s + \delta_2\mathbf{v}_{q,h}$, so we have from (3.8) that

$$\begin{aligned}
T_7 &= (\hat{\boldsymbol{\sigma}}(\mathbf{v}_s - \mathbf{v}_{s,h}), \nabla\boldsymbol{\psi}_s) \\
&= (\hat{\boldsymbol{\sigma}}(\pi_{\mathrm{S}}\mathbf{v}_s - \mathbf{v}_{s,h} - \pi_{\mathrm{S}}\mathbf{g}_s + \hat{\mathbf{g}}_s), \nabla\boldsymbol{\psi}_s) + (\hat{\boldsymbol{\sigma}}(\mathbf{v}_s - \pi_{\mathrm{S}}\mathbf{v}_s + \pi_{\mathrm{S}}\mathbf{g}_s - \hat{\mathbf{g}}_s), \nabla\boldsymbol{\psi}_s) \\
&\geq C_1\|\pi_{\mathrm{S}}\mathbf{v}_s - \mathbf{v}_{s,h} - \pi_{\mathrm{S}}\mathbf{g}_s + \hat{\mathbf{g}}_s\|_1^2 - \tfrac{1}{2}C_1\|\pi_{\mathrm{S}}\mathbf{v}_s - \mathbf{v}_{s,h}\|_1^2 \\
&\quad - C\big\{\|\mathbf{v}_s - \pi_{\mathrm{S}}\mathbf{v}_s\|_1^2 + \|\pi_{\mathrm{S}}\mathbf{g}_s - \hat{\mathbf{g}}_s\|_1^2 + \delta_2^2\|\mathbf{v}_{q,h}\|_1^2\big\} \\
&\geq \tfrac{1}{2}C_1\|\mathbf{v}_s - \mathbf{v}_{s,h}\|_1^2 - C\big\{\|\mathbf{v}_s - \pi_{\mathrm{S}}\mathbf{v}_s\|_1^2 + \|\pi_{\mathrm{S}}\mathbf{g}_s - \hat{\mathbf{g}}_s\|_1^2 + \delta_2^2\|\mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q - q_h\|^2\big\}.
\end{aligned}$$

We turn now to the fourth term T_4 in (5.14), which we estimate similarly to a term in [5, Section 7] for the degenerate Darcy system. That is, we introduce the projection $I - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}$ and compute as follows:

$$\begin{aligned}
-T_4 &= \big(\tilde{q}_f - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f, \phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi}_r)\big) \\
&= \big(\tilde{q}_f - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f, (I - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}})\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi}_r)\big) \\
&= \big(\tilde{q}_f - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f, (I - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}})\phi^{1/2+\Theta}\nabla\cdot\boldsymbol{\psi}_r + (I - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}})(1+\Theta)\phi^{\Theta-1/2}\nabla\phi\cdot\boldsymbol{\psi}_r\big) \\
&\leq C\|\tilde{q}_f - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f\|\big\{\|(I - \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}})\phi^{1/2+\Theta}\nabla\cdot\boldsymbol{\psi}_r\| + \|\boldsymbol{\psi}_r\|\big\},
\end{aligned}$$

since we have assumed the bound (4.7) on the term $\phi^{\Theta-1/2}\nabla\phi$. Because $\nabla\cdot\boldsymbol{\psi}_r$ is piecewise constant, we have that

$$\begin{aligned}\|(I-\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}})\phi^{1/2+\Theta}\nabla\cdot\boldsymbol{\psi}_r\| &\le \|(I-\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}})\phi^{1/2+\Theta}\|_{L^\infty(\Omega)}\|\nabla\cdot\boldsymbol{\psi}_r\|\\ &\le Ch\,\|\phi^{1/2+\Theta}\|_{W^{1,\infty}(\Omega)}\|\nabla\cdot\boldsymbol{\psi}_r\|\\ &\le Ch\,\|\nabla\cdot\boldsymbol{\psi}_r\|,\end{aligned}$$

using [20] for the approximation of the L^2-projection in L^∞ and (4.7) again. If we assume that the mesh is quasiuniform, then we can remove the divergence operator in the final expression at the expense of a power of the mesh spacing h. Thus we have

$$\begin{aligned}-T_4 &\le C\|\tilde{q}_f-\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f\|\,\|\boldsymbol{\psi}_r\|\\ &\le \epsilon\|\tilde{\mathbf{v}}_r-\tilde{\mathbf{v}}_{r,h}\|^2+C\{\|\tilde{q}_f-\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f\|^2+\|\tilde{\mathbf{v}}_r-\pi_{\mathrm{RT}}\tilde{\mathbf{v}}_r\|^2+\|\pi_{\mathrm{RT}}\mathbf{g}_r-\hat{\mathbf{g}}_r\|^2\}.\end{aligned}$$

Combining these estimates results in

$$\begin{aligned}&\frac{\mu_f}{2k_0}\|\tilde{\mathbf{v}}_r-\tilde{\mathbf{v}}_{r,h}\|^2+\delta_1\big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r-\tilde{\mathbf{v}}_{r,h}))\big]\big\|^2+\tfrac{1}{2}C_1\|\mathbf{v}_s-\mathbf{v}_{s,h}\|_1^2\\ &\quad+\frac{1}{2\mu_s(1-\phi^*)}\|\tilde{q}_f-\tilde{q}_{f,h}-\phi^{1/2}(q-q_h)\|^2+\tfrac{1}{4}\delta_2\gamma_{\mathrm{S}}^*\|q-q_h\|^2\\ &\le \epsilon\big\{\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f-\tilde{q}_{f,h}\|^2+\|\nabla\cdot(\mathbf{v}_s-\mathbf{v}_{s,h})\|^2+\|\tilde{\mathbf{v}}_r-\tilde{\mathbf{v}}_{r,h}\|^2\big\}\\ &\quad+C\big\{\|\tilde{\mathbf{v}}_r-\pi_{\mathrm{RT}}\tilde{\mathbf{v}}_r\|^2+\big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r-\pi_{\mathrm{RT}}\tilde{\mathbf{v}}_r))\big]\big\|^2\\ &\qquad+\|\pi_{\mathrm{RT}}\mathbf{g}_r-\hat{\mathbf{g}}_r\|^2+\big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\pi_{\mathrm{RT}}\mathbf{g}_r-\hat{\mathbf{g}}_r))\big]\big\|^2\\ &\qquad+\|\tilde{q}_f-\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f\|^2+\|q-\mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q\|^2+\delta_1^2\big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r-\tilde{\mathbf{v}}_{r,h}))\big]\big\|^2\\ &\qquad+\|\mathbf{v}_s-\pi_{\mathrm{S}}\mathbf{v}_s\|_1^2+\|\pi_{\mathrm{S}}\mathbf{g}_s-\hat{\mathbf{g}}_s\|_1^2+\delta_2^2\|\mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q-q_h\|^2\big\}.\end{aligned}\tag{5.15}$$

Note that

$$\|\tilde{q}_f-\tilde{q}_{f,h}\|^2\le\|\tilde{q}_f-\tilde{q}_{f,h}-\phi^{1/2}(q-q_h)\|^2+\|q-q_h\|^2.$$

Therefore, if we take ϵ, δ_1, and δ_2 small enough, we have proven the following theorem and specifically the error bound (5.16).

THEOREM 5. *Assume that* (4.7) *holds on the porosity,* $0\le\phi\le\phi^*<1$, *the mesh is quasiuniform,* $\Gamma_{\mathrm{ess}}^f=\partial\Omega$, *and the extensions* $\mathbf{g}_r\in\tilde{\mathbb{V}}_r$ *and* $\mathbf{g}_s\in\mathbb{V}_s$ *satisfy* (4.12) *and their approximations satisfy* (5.6). *Then the difference of the solution to the scaled formulation* (4.13)–(4.16), (1.5), *and its finite element approximation satisfy*

$$\begin{aligned}&\|\tilde{\mathbf{v}}_r-\tilde{\mathbf{v}}_{r,h}\|+\big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r-\tilde{\mathbf{v}}_{r,h}))\big]\big\|+\|\mathbf{v}_s-\mathbf{v}_{s,h}\|_1\\ &\quad+\|\tilde{q}_f-\tilde{q}_{f,h}\|+\|q-q_h\|\\ &\le C\big\{\|\tilde{\mathbf{v}}_r-\pi_{\mathrm{RT}}\tilde{\mathbf{v}}_r\|+\big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\tilde{\mathbf{v}}_r-\pi_{\mathrm{RT}}\tilde{\mathbf{v}}_r))\big]\big\|\\ &\qquad+\|\pi_{\mathrm{RT}}\mathbf{g}_r-\hat{\mathbf{g}}_r\|+\big\|\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\big[\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}(\pi_{\mathrm{RT}}\mathbf{g}_r-\hat{\mathbf{g}}_r))\big]\big\|\\ &\qquad+\|\tilde{q}_f-\mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\tilde{q}_f\|+\|q-\mathcal{P}_{\mathbb{W}_{\mathrm{S}}}q\|+\|\mathbf{v}_s-\pi_{\mathrm{S}}\mathbf{v}_s\|_1+\|\pi_{\mathrm{S}}\mathbf{g}_s-\hat{\mathbf{g}}_s\|_1\big\}.\end{aligned}\tag{5.16}$$

If the solution is sufficiently smooth, this bound implies first order convergence. It also implies stability of the scheme even when the solution is not very smooth.

6. A Modification for Local Mass Conservation. As in [4], we define a locally mass conservative implementation of the scaled method by using the quantity $\hat{\phi} = \mathcal{P}_{\mathbb{W}_{\mathrm{RT}}}\phi \in \mathbb{W}_{\mathrm{RT}}$ given by taking the average over each element $E \in \mathcal{T}_h$, i.e.,

$$\text{for } \mathbf{x} \in E, \quad \hat{\phi}(\mathbf{x}) = \hat{\phi}_E = \frac{1}{|E|}\int_E \phi \, dx, \tag{6.1}$$

where $|E|$ is the area of E. When $\hat{\phi}|_E = \hat{\phi}_E = 0$ vanishes on an element E, ϕ is identically zero on E. We modify the two divergence terms in the scaled MFEM (5.7)–(5.10) by replacing

$$\phi^{-1/2}\nabla\cdot(\phi^{1+\Theta}\mathbf{v})\big|_E \quad \text{by} \quad \begin{cases} \hat{\phi}_E^{-1/2}\nabla\cdot(\phi^{1+\Theta}\mathbf{v}) & \text{if } \hat{\phi}_E \neq 0, \\ 0 & \text{if } \hat{\phi}_E = 0. \end{cases}$$

We also modify the two terms in (5.8) and (5.10) involving the pressure potentials. These changes make the method locally mass conservative, as we show later.

Locally conservative scaled mixed finite element method. Find $\tilde{\mathbf{v}}_{r,h} \in \mathbb{V}_{\mathrm{RT},0} + \hat{\mathbf{g}}_r$, $\tilde{q}_{f,h} \in \mathbb{W}_{\mathrm{RT}}$, $\mathbf{v}_{s,h} \in \mathbb{V}_{\mathrm{S},\mathbf{0}} + \hat{\mathbf{g}}_s$, and $q_h \in \mathbb{W}_{\mathrm{S},*}$ such that

$$\begin{aligned}
&\Big(\frac{\mu_f}{k_0}\tilde{\mathbf{v}}_{r,h}, \boldsymbol{\psi}_r\Big) - \big(\tilde{q}_{f,h}, \hat{\phi}^{-1/2}\nabla\cdot(\phi^{1+\Theta}\boldsymbol{\psi}_r)\big) \\
&\quad = -\langle \tilde{g}_f, \phi^{1/2+\Theta}\boldsymbol{\psi}_r\cdot\nu\rangle_{\Gamma^f_{\mathrm{nat}}} && \forall\, \boldsymbol{\psi}_r \in \mathbb{V}_{\mathrm{RT},0}, \quad (6.2)\\
&\big(\hat{\phi}^{-1/2}\nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_{r,h}), w_f\big) + \Big(\frac{\phi\hat{\phi}^{-1}}{\mu_s(1-\phi)}(\tilde{q}_{f,h} - \hat{\phi}^{1/2}q_h), w_f\Big) \\
&\quad = 0 && \forall w_f \in \mathbb{W}_{\mathrm{RT}}, \quad (6.3)\\
&-(q_h, \nabla\cdot\boldsymbol{\psi}_s) + (\hat{\boldsymbol{\sigma}}(\mathbf{v}_{s,h}), \nabla\boldsymbol{\psi}_s) = -\big((1-\phi)\rho_r\mathbf{g}, \boldsymbol{\psi}_s\big) && \forall\, \boldsymbol{\psi}_s \in \mathbb{V}_{\mathrm{S},\mathbf{0}}, \quad (6.4)\\
&(\nabla\cdot\mathbf{v}_{s,h}, w) - \Big(\frac{\phi\hat{\phi}^{-1/2}}{\mu_s(1-\phi)}(\tilde{q}_{f,h} - \hat{\phi}^{1/2}q_h), w\Big) = 0 && \forall w \in \mathbb{W}_{\mathrm{S},*}, \quad (6.5)
\end{aligned}$$

where the term $\hat{\boldsymbol{\sigma}}$ is defined by (1.5). On an element $E \in \mathcal{T}_h$ where $\hat{\phi}_E = 0$, the three terms in (6.2)–(6.3), (6.5) involving $\hat{\phi}^{-1/2}$ are set to zero, and in the second term in (6.3), we interpret $\phi\hat{\phi}^{-1}$ as one. Furthermore, we define the discrete Darcy velocity $\mathbf{u}_h \in \mathbb{V}_{\mathrm{RT}}$ and fluid potential $q_{f,h} \in \mathbb{W}_{\mathrm{RT}}$ by their degrees of freedom:

$$\mathbf{u}_h\cdot\nu|_e = \frac{1}{|e|}\int_e \phi^{1+\Theta}\, ds\ \tilde{\mathbf{v}}_{r,h}\cdot\nu|_e \quad \forall e \in \mathcal{E}_h, \tag{6.6}$$

$$q_{f,h}|_E = \hat{\phi}_E^{-1/2}\tilde{q}_{f,h}|_E \quad \forall E \in \mathcal{T}_h, \tag{6.7}$$

wherein we arbitrarily set $q_{f,h}|_E = 0$ if $\hat{\phi}_E = 0$.

The existence of a unique solution can be shown in a way completely analogous to that for the nonconservative scaled MFEM in Section 5. Moreover, one can show that the locally conservative MFEM is stable, provided that $\Gamma^f_{\mathrm{ess}} = \partial\Omega$. We have no proof of convergence of the locally conservative method at this time, but the numerical results show optimal convergence and even superconvergence.

To see local mass conservation of the fluid, let $E \in \mathcal{T}_h$ be any element. With w_E defined in (5.4), the test function $w_f = \hat{\phi}_E^{1/2}\, w_E \in \mathbb{W}_{\mathrm{RT}}$ in (6.3) gives

$$\mu_s\int_E \nabla\cdot(\phi^{1+\Theta}\tilde{\mathbf{v}}_{r,h})\, dx + \int_E \frac{\phi}{1-\phi}\big(\hat{\phi}^{-1/2}\tilde{q}_{f,h} - q_h\big)\, dx = 0.$$

Since $\tilde{\mathbf{v}}_r \cdot \nu$ and $\mathbf{u}_h \cdot \nu$ are constant on each edge $e \subset \partial E$, we see from (6.6) that

$$\int_E \nabla \cdot (\phi^{1+\Theta} \tilde{\mathbf{v}}_{r,h})\, dx = \int_{\partial E} \phi^{1+\Theta} \tilde{\mathbf{v}}_{r,h} \cdot \nu \, ds = \int_{\partial E} \mathbf{u}_h \cdot \nu \, ds = \int_E \nabla \cdot \mathbf{u}_h \, dx.$$

The definition of $q_{f,h}$ (6.7) gives

$$\mu_s \int_E \nabla \cdot \mathbf{u}_h \, dx + \int_E \frac{\phi}{1-\phi}(q_{f,h} - q_h)\, dx = 0, \tag{6.8}$$

which is local mass conservation, i.e., (1.2) holds locally.

We obtain local mass conservation of the solid matrix if we use BR spaces. In that case, we can take the test function $w = w_E \in \mathbb{W}_{\mathrm{BR}}$ in (6.5) to see

$$\mu_s \int_E \nabla \cdot \mathbf{v}_{s,h}\, dx - \int_E \frac{\phi}{1-\phi}\big(q_{f,h} - q_{s,h}\big)\, dx = 0,$$

which is (1.4) holding locally.

7. Implementation of the Methods on Rectangular Meshes. The linear system corresponding to either of the methods (5.7)–(5.10) or (6.2)–(6.5) has the form

$$\begin{pmatrix} A & -B_\phi & 0 & 0 \\ B_\phi^T & C_{f,\phi} & 0 & -C_{f,s,\phi} \\ 0 & 0 & D_\phi & -B \\ 0 & -C_{f,s,\phi}^T & B^T & C_{s,\phi} \end{pmatrix} \begin{pmatrix} \tilde{v}_r \\ \tilde{q}_f \\ v_s \\ q \end{pmatrix} = \begin{pmatrix} a_\phi \\ 0 \\ b_\phi \\ 0 \end{pmatrix}, \tag{7.1}$$

wherein the solution represents the degrees of freedom of $\tilde{\mathbf{v}}_{r,h}$, $\tilde{q}_{f,h}$, $\mathbf{v}_{s,h}$, and q_h with respect to the bases of the finite element spaces. We remark only on the evaluation of B_ϕ and D_ϕ. To avoid approximating derivatives of ϕ, the matrix B_ϕ should be computed using the divergence theorem. For the locally conservative method, for any element $E \in \mathcal{T}_h$ and edge $e \in \mathcal{E}_h$,

$$\begin{aligned} B_{\phi,e,E} &= \big(\hat{\phi}^{-1/2} \nabla \cdot (\phi^{1+\Theta} \mathbf{v}_e), w_E\big) \\ &= \begin{cases} \hat{\phi}_E^{-1/2} \displaystyle\int_e \phi^{1+\Theta}\, ds\, \mathbf{v}_e \cdot \nu_E & \text{if } e \subset \partial E \text{ and } \hat{\phi}_E \neq 0, \\ 0 & \text{if } e \not\subset \partial E \text{ or } \hat{\phi}_E = 0. \end{cases} \end{aligned}$$

A similar expression is used for the nonconservative scaled MFEM of Section 5. The matrix D_ϕ is symmetric, and the (k,ℓ) entry is computed using (1.5) as

$$D_{\phi,k,\ell} = (\hat{\boldsymbol{\sigma}}(\mathbf{v}_{s,k}), \nabla \boldsymbol{\psi}_{s,\ell}) = 2\mu_s \big[\big((1-\phi)\mathcal{D}\mathbf{v}_{s,k}, \mathcal{D}\boldsymbol{\psi}_{s,\ell}\big) - \tfrac{1}{3}\big((1-\phi)\nabla \cdot \mathbf{v}_{s,k}, \nabla \cdot \boldsymbol{\psi}_{s,\ell}\big)\big].$$

We can simplify the implementation when Ω is a union of rectangular subdomains in one, two, or three dimensions, and $\mathcal{T}_h$ is a rectangular finite element mesh. We modify either method by approximating the first integral in (5.7) or (6.2) using what is known as mass lumping. The integral is approximated by a trapezoidal quadrature rule $(\cdot,\cdot)_Q$, so that for any two edges $e, f \in \mathcal{E}_h$,

$$A_{e,f} = \Big(\frac{\mu_f}{k_0}\mathbf{v}_e, \mathbf{v}_f\Big)_Q = \frac{\mu_f}{2k_0}|E_e|\delta_{e,f}, \tag{7.2}$$

where E_e is the element or union of two elements that have e as an edge. This approximation diagonalizes A and enables us to eliminate the scaled relative velocity using the Schur complement from the first row of (7.1),

$$\tilde{v}_r = A^{-1}(B_\phi \tilde{q}_f + a_\phi).$$

What remains is a Stokes-like system with *two* pressure potentials. One can further eliminate $v_s = D_\phi^{-1}(Bq + b_\phi)$ to obtain

$$\begin{pmatrix} B_\phi^T A^{-1} B_\phi + C_{f,\phi} & -C_{f,s,\phi} \\ -C_{f,s,\phi}^T & B^T D_\phi^{-1} B + C_{s,\phi} \end{pmatrix} \begin{pmatrix} \tilde{q}_f \\ q \end{pmatrix} = \begin{pmatrix} -B_\phi^T A^{-1} a_\phi \\ -B^T D_\phi^{-1} b_\phi \end{pmatrix}, \tag{7.3}$$

but the matrix $B^T D_\phi^{-1} B$ is not easily formed. Nevertheless, one can apply this matrix and therefore solve a Schur complement system for the two pressure potentials. The system can be preconditioned by a diagonal preconditioner, using any good preconditioners for the two diagonal blocks, and solved by conjugate gradients. See, e.g., the block preconditioner defined in [33].

8. Numerical Results in One Dimension. In this section we present numerical results for the new locally conservative scaled mixed finite element method (6.2)–(6.5) for the Darcy-Stokes equations by simulating a compacting column in one dimension.

Consider a mantle column [30] for $z \in [-L, L]$ with no flow through the top and bottom boundaries, i.e.,

$$v_s(-L) = v_s(L) = u(-L) = u(L) = 0, \tag{8.1}$$

and also with the fluid potential scale set so that

$$q_f(0) = 0. \tag{8.2}$$

In order to reveal the qualitative nature of the equations for the compacting column, we non-dimensionalize using the compaction length scale [27]

$$L_c = \left(\frac{k_0 \mu_s}{\mu_f}\right)^{1/2} \approx 10^5 \text{ to } 3 \times 10^6 \text{ m}. \tag{8.3}$$

Define the dimensionless variables

$$\begin{aligned} x &= L_c\, \check{x}, \\ q_f &= |\rho_r| g L_c\, \check{q}_f, \quad q_s = |\rho_r| g L_c\, \check{q}_s, \\ u &= \frac{k_0 |\rho_r| g}{\mu_f}\, \check{u}, \quad v_s = \frac{k_0 |\rho_r| g}{\mu_f}\, \check{v}_s. \end{aligned}$$

By reducing to the vertical dimension, using (2.10), and dropping the check accent marks, (1.1)–(1.5) become, after some manipulation,

$$u + \phi^{2+2\Theta} q_f' = 0, \tag{8.4}$$

$$u' + \phi(q_f - q_s) = 0, \tag{8.5}$$

$$[q_s - \tfrac{1}{3}(1 - 4\phi) v_s']' = 1 - \phi, \tag{8.6}$$

$$v_s' - \phi(q_f - q_s) = 0. \tag{8.7}$$

Note that in a one-dimensional problem, there is no distinction between shear deformation and compaction.

Where $\phi > 0$, we can reduce the system to a single equation in terms of u as follows. First, (8.5) and (8.7) imply that $v_s' = -u'$ and $q_s = u'/\phi + q_f$. Equation (8.4) gives $q_f' = -\phi^{-2-2\Theta} u$. Finally, (8.6) reduces to

$$\phi^{2+2\Theta}\Big(\frac{3 + \phi - 4\phi^2}{3\phi}\, u'\Big)' - u = \phi^{2+2\Theta}(1 - \phi). \tag{8.8}$$

On an open interval where $\phi = 0$, the equations reduce to $u = 0$, $q_s' = 1$, and $v_s' = 0$.

8.1. Closed form solutions. We will conduct tests of a compacting column problem using the three porosity functions

$$\phi_0(z) = \phi_0, \tag{8.9}$$

$$\phi_{\mathrm{J}}(z) = \begin{cases} \phi_- & \text{if } z \le 0, \\ \phi_+ & \text{if } z > 0, \end{cases} \tag{8.10}$$

$$\phi_2(z) = \begin{cases} 0 & \text{if } z \le 0, \\ \phi_+ z^2 & \text{if } z > 0, \end{cases} \tag{8.11}$$

where $\phi_0 > 0$ and $\phi_- \neq \phi_+$ gives a discontinuous jump in ϕ_{J}. Note that ϕ_0 and ϕ_2 satisfy the condition (4.7), since in fact $\phi_2^{\Theta-1/2}\nabla\phi_2 = 2\phi_+ z^{2\Theta}$ for $z > 0$ is indeed in $L^\infty(\Omega)$. However, ϕ_{J} does not satisfy this condition.

8.1.1. Constant porosity. Taking the constant porosity $\phi(z) = \phi_0 > 0$, (8.8) reduces to

$$R^{-2}u'' - u = \phi_0^{2+2\Theta}(1-\phi_0),$$

where

$$R = R(\phi_0) = \Big(\frac{3+\phi_0-4\phi_0^2}{3}\phi_0^{1+2\Theta}\Big)^{-1/2}.$$

Solving the differential equation with the potential scale condition (8.2) gives the full solution in terms of the constants a and b as

$$u = -\phi_0^{2+2\Theta}(1-\phi_0)\big[1 + a\cosh(Rz) + b\sinh(Rz)\big], \tag{8.12}$$

$$q_f = (1-\phi_0)\bigg\{z - \frac{b}{R} + \frac{1}{R}\big[a\sinh(Rz) + b\cosh(Rz)\big]\bigg\}, \tag{8.13}$$

$$q_s = (1-\phi_0)\bigg\{z - \frac{b}{R} + \frac{1-4\phi_0}{3+\phi_0-4\phi_0^2}\frac{\phi_0}{R}\big[a\sinh(Rz) + b\cosh(Rz)\big]\bigg\}. \tag{8.14}$$

The boundary conditions (8.1) imply that

$$v_s = -u, \quad a = -\frac{1}{\cosh(RL)}, \quad \text{and} \quad b = 0. \tag{8.15}$$

8.1.2. Discontinuous porosity. For the discontinuous porosity $\phi = \phi_J$ given in (8.10), we can solve (8.8) on each subdomain where ϕ is constant. If both ϕ_+ and ϕ_- are positive, the result is (8.12)–(8.14), i.e.,

$$u_\pm = \phi_\pm^{2+2\Theta}(1-\phi_\pm)\big[1 + a_\pm\cosh(R_\pm z) + b_\pm\sinh(R_\pm z)\big]. \tag{8.16}$$

For the equations (8.4)–(8.7) to make sense at the interface $z = 0$, the functions for which we take derivatives must be continuous. The scale condition (8.2) enforces continuity of q_f. However, we must impose continuity at $z = 0$ on u (and thereby on v_s) and on the quantity

$$\begin{aligned} q_s(0) - \frac{1}{3}(1-4\phi)\,v_s'(0) &= \Big(\frac{1}{\phi} + \frac{1}{3}(1-4\phi)\Big)u'(0) \\ &= R^{-2}\phi^{-2-2\Theta}\,u'(0). \end{aligned} \tag{8.17}$$

With the boundary condition (8.1), i.e., $u_\pm(\pm L) = 0$, we have four conditions that determine $a_\pm$ and $b_\pm$. Letting $\mathcal{F}_\pm = \phi_\pm^{2+2\Theta}(1-\phi_\pm)$, the coefficients are determined by solving the relatively simple linear system

$$\begin{bmatrix} \cosh(R_+L) & 0 & \sinh(R_+L) & 0 \\ 0 & \cosh(R_-L) & 0 & -\sinh(R_-L) \\ \mathcal{F}_+ & -\mathcal{F}_- & 0 & 0 \\ 0 & 0 & R_-(1-\phi_+) & -R_+(1-\phi_-) \end{bmatrix} \begin{bmatrix} a_+ \\ a_- \\ b_+ \\ b_- \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ \mathcal{F}_- - \mathcal{F}_+ \\ 0 \end{bmatrix}. \tag{8.18}$$

In the case that $\phi_- = 0$ but $\phi_+ > 0$, the solution to (8.8) is (8.12)–(8.14) for $z > 0$, but for $z < 0$, $u = v_s = 0$ (i.e., $a_- = b_- = 0$) and $q_s = z + c_-$. The interface conditions imply that

$$a_+ = -1, \quad b_+ = \frac{\cosh(R_+L) - 1}{\sinh(R_+L)}, \quad \text{and} \quad c_- = -b_+(1-\phi_+)/R_+. \tag{8.19}$$

Finally, if $\phi_- > 0$ and $\phi_+ = 0$, then

$$a_- = -1 \quad b_- = \frac{1 - \cosh(R_-L)}{\sinh(R_-L)} \quad \text{and} \quad c_+ = -b_-(1-\phi_-)/R_-, \tag{8.20}$$

where (8.12)–(8.14) gives the solution for $z < 0$ and for $z > 0$, $u = v_s = 0$ (i.e., $a_+ = b_+ = 0$) and $q_s = z + c_+$.

8.1.3. Quadratic porosity approximation. The final closed form solution is an approximation to the system. Set $\Theta = 0$ and take $\phi(z) = \phi_2(z)$ from (8.11). Working on $z > 0$, the differential equation (8.8) reduces to

$$\phi_+^2 z^4 \Big(\frac{3 + \phi_+ z^2 - 4\phi_+^2 z^4}{3\phi_+ z^2} u' \Big)' - u = \phi_+^2 z^4 (1 - \phi_+ z^2).$$

Assuming that $\phi = \phi_+ z^2 \ll 1$, we retain only the lowest order terms, i.e., we approximate this equation as

$$\phi_+ z^4 (z^{-2} u')' - u = \phi_+^2 z^4,$$

which reduces to the Euler equation

$$\phi_+ z^2 \, u'' - 2\phi_+ z \, u' - u = \phi_+^2 z^4.$$

The Euler exponents satisfy $\phi_+ r(r-3) - 1 = 0$, which are

$$r_1 = \frac{3 + \sqrt{9 + 4/\phi_+}}{2} > 3 \quad \text{and} \quad r_2 = \frac{3 - \sqrt{9 + 4/\phi_+}}{2} < 0,$$

and the solution to the homogeneous equation is $U = Az^{r_1} + Bz^{r_2}$. Since the solution is well-behaved, $B = 0$. If $\phi_+ \neq 1/4$, variation of parameters and the boundary condition $u(L) = 0$ give the solution for $z > 0$ as

$$u = -v_s = \frac{\phi_+^2}{1 - 4\phi_+} \big(L^{4-r_1} \, z^{r_1} - z^4 \big), \tag{8.21}$$

$$q_f = \frac{1}{1 - 4\phi_+} \Big(z - \frac{L^{4-r_1} \, z^{r_1 - 3}}{r_1 - 3} \Big), \tag{8.22}$$

$$q_s = z. \tag{8.23}$$

For $z < 0$ where $\phi = 0$, the solution is $u = v_s = 0$. Moreover, $q_s = z$, using the continuity condition (8.17) and noting that $u'(0) = 0$.

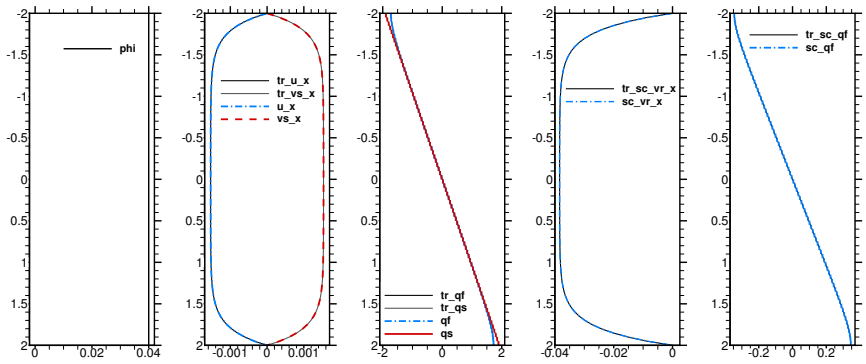


FIG. 8.1. **Constant porosity** (8.9) *with* $\phi_0 = 0.04$. *The computed solution as thick dashed lines, and the closed form solution as thin, solid dark lines. Shown are the porosity* ϕ *(phi),* u*,* $v_s = -u$*,* q_f*, and* q_s*, as well as the scaled* $\tilde{v}_r$ *and* $\tilde{q}_f$*.*

8.2. Verification of the scaled method. In this subsection we present numerical results for the locally conservative scaled mixed method (6.2)–(6.5) and its mass lumped approximation using (7.2). We use the BR spaces for the Stokes part of the system. Note that in one dimension, the velocity part of the RT and BR spaces reduce to piecewise continuous linear functions, and the pressure part is the set of piecewise discontinuous constant functions. The theoretical bound in Theorem 5 would guarantee a convergence rate of $\mathcal{O}(h)$ for the potentials and velocities, provided ϕ satisfies (4.7).

We simulate the dimensionless compacting column test (8.1)–(8.2), (8.4)–(8.7) using the three porosity functions (8.9)–(8.11) defined above. In all tests, we take $L = 2$, so that the domain extends four compaction lengths, and we fix $\Theta = 0$. Each problem is solved on a uniform mesh of n cells. Our computer code is based on the deal.II software library [11].

We chose above to fix the pressure scale of $\tilde{q}_f$ by imposing (8.2) in the interior of the domain. This works well for the closed form solution. However, it does not set the scale properly in our numerical implementation of the problems. Instead, we set the pressure scale of q at the point where it achieves its maximum value. Moreover, in the two cases where the porosity degenerates, we also set the scale of q at the point where it achieves its minimum value.

8.2.1. Constant porosity tests. For the first set of tests, we take the constant porosity $\phi(z) = \phi_0 = 0.04$. This problem tests the overall performance of the code when there is no degeneracy in the porosity. The computed and closed form solutions using $n = 80$ are shown in Fig. 8.1, although the former is so accurate that it obscures the latter.

In Table 8.1 we give the relative errors of the potentials as measured in the L^2-norm for both the scaled mixed method and its mass lumped approximation. The optimal rates of convergence $\mathcal{O}(h)$ are observed for $\tilde{q}_f$, q_f, and q. We also measured the errors in the discrete L^2 norm, which is the usual L^2-norm but evaluated using the midpoint quadrature rule. This is a norm for which one might expect to see superconvergence. Indeed, we see superconvergence for all three potentials. On coarser meshes we see $\mathcal{O}(h^{3/2})$ for the fluid potentials and $\mathcal{O}(h)$ for the mixture, but on fine meshes the rates rise to $\mathcal{O}(h^2)$ for all three variables. Similar superconvergence results hold for the mass lumped approximation.

TABLE 8.1
Constant porosity potential errors. *Relative L^2 errors and convergence rates for the potentials. We show results for the scaled mixed method, the mass lumped approximation, and the scaled mixed method but using the discrete L^2-norm given by using the midpoint rule.*

	$\tilde{q}_f$		q_f		q	
n	L^2 error	rate	L^2 error	rate	L^2 error	rate
Scaled mixed method						
20	1.428e-02	1.00	3.237e-02	1.00	3.435e-02	1.00
40	7.139e-03	1.00	1.618e-02	1.00	1.717e-02	1.00
80	3.569e-03	1.00	8.090e-03	1.00	8.581e-03	1.00
160	1.784e-03	1.00	4.044e-03	1.00	4.290e-03	1.00
Mass lumped method						
20	1.427e-02	1.00	3.236e-02	1.00	3.434e-02	1.00
40	7.139e-03	1.00	1.618e-02	1.00	1.717e-02	1.00
80	3.569e-03	1.00	8.090e-03	1.00	8.581e-03	1.00
160	1.784e-03	1.00	4.044e-03	1.00	4.290e-03	1.00
Scaled mixed method, discrete norm (midpoint rule)						
20	8.597e-04	1.46	1.949e-03	1.46	1.411e-03	0.91
40	2.794e-04	1.62	6.334e-04	1.62	5.422e-04	1.38
80	8.263e-05	1.76	1.873e-04	1.76	1.708e-04	1.67
160	2.271e-05	1.86	5.149e-05	1.86	4.813e-05	1.83
320	5.972e-06	1.93	1.354e-05	1.93	1.279e-05	1.91
640	1.532e-06	1.96	3.472e-06	1.96	3.297e-06	1.96

TABLE 8.2
Constant porosity velocity errors. *Relative L^2 errors and convergence rates for the velocities using the scaled mixed method and the mass lumped approximation.*

	$\tilde{v}_f$		u		v_s	
n	L^2 error	rate	L^2 error	rate	L^2 error	rate
Scaled mixed method						
20	1.147e-03	1.82	4.897e-05	1.82	4.897e-05	1.82
40	2.972e-04	1.95	1.269e-05	1.95	1.269e-05	1.95
80	7.500e-05	1.99	3.203e-06	1.99	3.203e-06	1.99
160	1.879e-05	2.00	8.027e-07	2.00	8.027e-07	2.00
Mass lumped method						
20	1.650e-03	1.72	7.047e-05	1.72	7.047e-05	1.72
40	4.381e-04	1.91	1.871e-05	1.91	1.871e-05	1.91
80	1.113e-04	1.98	4.753e-06	1.98	4.753e-06	1.98
160	2.794e-05	1.99	1.193e-06	1.99	1.193e-06	1.99

In Table 8.2 we give the relative errors of the velocities in the L^2-norm for both the scaled mixed method and its mass lumped approximation. The velocities are approximated by piecewise linears, so the optimal rates of convergence would be $\mathcal{O}(h^2)$. This is precisely what is observed for each of the velocities $\tilde{v}_f$, u, and v_s.

8.2.2. Discontinuous porosity tests. For the next set of tests, we use the discontinuous porosity function (8.10) with $\phi_- = 0$ and $\phi_+ = 0.04$. Not only is there a jump in porosity, but it is also degenerate for $z < 0$. Note that the discontinuity will land on a mesh point if we use an even number of mesh cells n, and it will land

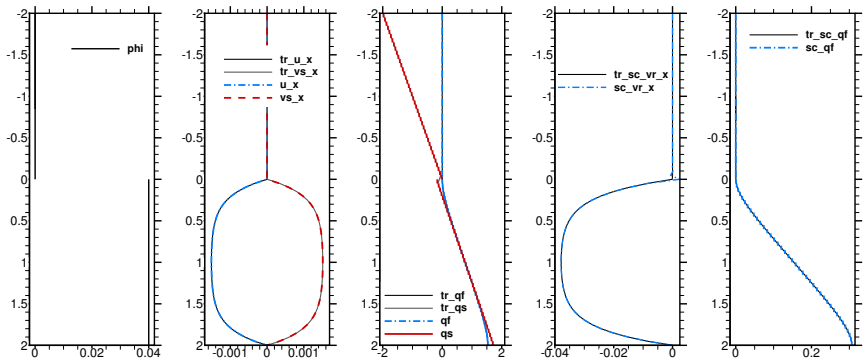


FIG. 8.2. **Discontinuous porosity** (8.10) *with* $\phi_- = 0$ *and* $\phi_+ = 0.04$. *The computed solution as thick dashed lines, and the closed form solution as thin, solid dark lines. Shown are the porosity* ϕ *(phi),* u, $v_s = -u$, q_f, *and* q_s, *as well as the scaled* $\tilde{v}_r$ *and* $\tilde{q}_f$.

in the center of a cell when n is odd.

The computed and closed form solutions using $n = 80$ are shown in Figure 8.2. Note that the discontinuity in q_s is clearly evident (and approximated well). Note also that the computed solution $\tilde{v}_r$ has some difficulty near $z = 0$ where the porosity is discontinuous. This difficulty is not seen in u and v_s, since compared to $\tilde{v}_r$, these velocities are multiplied by ϕ. Overall, the computed solution is an excellent match to the closed form one.

In Table 8.3 we give convergence results for the potentials using the scaled mixed method. The mass lumped approximation has nearly identical results. Even though ϕ does not satisfy the condition (4.7) and is in fact discontinuous, we see good convergence results. When n is even and the grid resolves the discontinuity in ϕ, we see optimal convergence rates $\mathcal{O}(h)$ for all three potentials and superconvergence $\mathcal{O}(h^2)$ when using the discrete norm.

When n is odd and the discontinuity in ϕ is not resolved, we see some degradation in the convergence rate for q and no superconvergence in the discrete norm. To test whether the error near the discontinuity pollutes the solution, we computed the *interior* errors. These are given by computing the error in all cells of the mesh except the five near the discontinuity. That is, we restrict the domain of integration of the L^2-norm to be interior to where ϕ is smooth by removing the center cell and its two neighbors on each side. This mesh dependent norm shows good $\mathcal{O}(h)$ convergence, and so indeed the error is localized to the region of the discontinuity. We do not, however, observe superconvergence in the discrete interior norm when n is odd.

The errors in the velocities are given in Table 8.4. We see good rates of convergence when n is even, being $\mathcal{O}(h^2)$ for all cases except the scaled method's $\tilde{\mathbf{v}}_f$, which is still $\mathcal{O}(h^{3/2})$. When n is odd, we observe $\mathcal{O}(h)$ convergence (we show only the scaled method, but the mass lumped approximation is similar).

8.2.3. Quadratic porosity tests. For the final set of tests, we use the quadratic porosity function (8.11) with $\phi_+ = 0.001$, i.e., $\phi_2(z) = 0.001\, z^2$ for $z > 0$ and $\phi_2(z) = 0$ for $z \le 0$. The maximal value of ϕ is $\phi(2) = 0.004$, so the analytic solution (8.21)–(8.23) should approximate the true solution reasonably well, at least if n is not too large.

The computed and closed form solutions using $n = 80$ are shown in Figure 8.3. Note that there is a boundary layer near $z = 2$ in the velocities that is difficult to

TABLE 8.3

Discontinuous porosity potential errors. *Relative L^2 errors and convergence rates for the potentials. We show results for the scaled mixed method, including two cases using the discrete L^2-norm given by using the midpoint rule, and one case with the L^2-norm restricted to the interior (i.e., away from the discontinuity). When n is even, the discontinuity is at a computational mesh point, but not when n is odd.*

	$\tilde{q}_f$		q_f		q	
n	L^2 error	rate	L^2 error	rate	L^2 error	rate
Scaled mixed method, n even						
20	1.040e-02	1.00	2.852e-02	1.00	3.622e-02	1.00
40	5.202e-03	1.00	1.426e-02	1.00	1.811e-02	1.00
80	2.601e-03	1.00	7.133e-03	1.00	9.055e-03	1.00
160	1.301e-03	1.00	3.567e-03	1.00	4.527e-03	1.00
Scaled mixed method, n odd						
21	9.961e-03	0.98	2.744e-02	0.98	3.955e-02	0.87
41	5.140e-03	0.99	1.416e-02	0.99	2.321e-02	0.80
81	2.611e-03	0.99	7.184e-03	1.00	1.428e-02	0.71
161	1.316e-03	1.00	3.615e-03	1.00	9.218e-03	0.64
Scaled mixed method, n even, discrete norm (midpoint rule)						
20	9.536e-04	1.64	2.615e-03	1.64	1.231e-04	1.46
40	2.529e-04	1.91	6.935e-04	1.91	3.860e-05	1.67
80	6.225e-05	2.02	1.707e-04	2.02	1.102e-05	1.81
160	1.505e-05	2.05	4.128e-05	2.05	2.964e-06	1.89
Scaled mixed method, n odd, discrete norm (midpoint rule)						
21	1.553e-03	0.50	5.136e-03	0.39	2.691e-02	0.52
41	9.191e-04	0.78	2.912e-03	0.85	1.917e-02	0.51
81	4.900e-04	0.92	1.473e-03	1.00	1.361e-02	0.50
161	2.512e-04	0.97	7.263e-04	1.03	9.647e-03	0.50
Scaled mixed method, n odd, interior						
21	9.126e-03	0.68	2.502e-02	0.68	3.016e-02	0.74
41	4.983e-03	0.90	1.367e-02	0.90	1.658e-02	0.89
81	2.572e-03	0.97	7.052e-03	0.97	8.670e-03	0.95
161	1.304e-03	0.99	3.576e-03	0.99	4.431e-03	0.98

resolve, but the computed and closed form solutions agree quite well.

The convergence results for the potentials and velocities are given in Tables 8.5 and 8.6. We give only results for the scaled method, since the mass lumped approximation gives nearly identical results. We expect convergence only while the approximate true solution is adequate. Indeed, we see some degradation of the results as n becomes too fine. The potentials converge to $\mathcal{O}(h)$; however, the discrete norm does not display superconvergence for this test problem (however, the errors are much smaller, as can be seen in Table 8.5). The rates of convergence for the velocities are at least $\mathcal{O}(h)$, and may approach $\mathcal{O}(h^2)$ before the grid becomes too fine.

8.3. Condition number as positive ϕ tends to zero. We now turn our attention to the nondegenerate problem, so that we can solve the system of equations using all four of our mixed finite element formulations: the standard, expanded, symmetric, and scaled formulations. We consider the same three nonconstant porosities (8.9)–(8.11), but add a small positive constant $\phi_\epsilon > 0$ to each. In this test we take

TABLE 8.4
Discontinuous porosity velocity errors. *Relative L^2 errors and convergence rates for the velocities. We show results for the scaled mixed method and the mass lumped approximation using (7.2). When n is even, the discontinuity is at a computational mesh point, but not when n is odd.*

	$\tilde{v}_f$		u		v_s	
n	L^2 error	rate	L^2 error	rate	L^2 error	rate
Scaled mixed method, n even						
10	7.364e-03	—	1.700e-04	—	1.700e-04	—
20	2.929e-03	1.33	4.714e-05	1.85	4.714e-05	1.85
40	1.116e-03	1.39	1.213e-05	1.96	1.213e-05	1.96
80	4.120e-04	1.44	3.090e-06	1.97	3.090e-06	1.97
160	1.491e-04	1.47	7.850e-07	1.98	7.850e-07	1.98
Mass lumped method, n even						
10	5.608e-03	—	2.341e-04	—	2.341e-04	—
20	1.695e-03	1.73	7.076e-05	1.73	7.076e-05	1.73
40	4.499e-04	1.91	1.878e-05	1.91	1.878e-05	1.91
80	1.143e-04	1.98	4.770e-06	1.98	4.770e-06	1.98
160	2.869e-05	1.99	1.197e-06	1.99	1.197e-06	1.99
Scaled mixed method, n odd						
11	4.417e-03	—	2.027e-04	—	2.027e-04	—
21	2.027e-03	1.20	9.004e-05	1.25	9.004e-05	1.25
41	1.055e-03	0.98	4.524e-05	1.03	4.524e-05	1.03
81	5.614e-04	0.93	2.368e-05	0.95	2.368e-05	0.95
161	2.925e-04	0.95	1.227e-05	0.96	1.227e-05	0.96

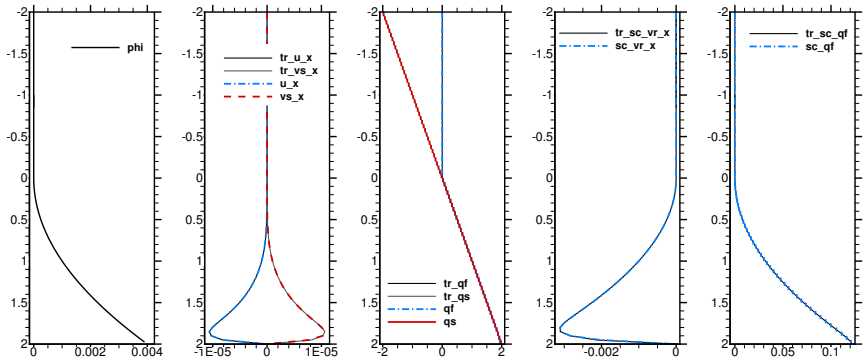


FIG. 8.3. **Quadratic porosity** (8.11) *with $\phi_+ = 0.001$. The computed solution as thick dashed lines, and the closed form solution as thin, solid dark lines. Shown are the porosity ϕ (phi), u, $v_s = -u$, q_f, and q_s, as well as the scaled $\tilde{v}_r$ and $\tilde{q}_f$.*

$\phi_\epsilon \to 0^+$ and observe the condition number of the linear system that is solved by each formulation.

In Fig. 8.4 we show a graph of the results of one of the porosity functions (all three show nearly identical behavior). The condition number of each method increases rapidly as $\phi_\epsilon \to 0$ with the exception of the scaled method, which remains stable. Indeed, as we saw in the previous section, the scaled method works well even when the porosity is identically zero in parts of the domain.

TABLE 8.5

Quadratic porosity potential errors. *Relative L^2 errors and convergence rates for the potentials for the scaled mixed method, including two cases using the discrete L^2-norm given by using the midpoint rule. When n is even, the transition to positive porosity is at a computational mesh point, but not when n is odd.*

	$\tilde{q}_f$		q_f		q	
n	L^2 error	rate	L^2 error	rate	L^2 error	rate
Scaled mixed method, n even						
20	5.326e-03	1.01	3.037e-02	1.01	3.490e-02	1.00
40	2.656e-03	1.00	1.520e-02	1.00	1.747e-02	1.00
80	1.329e-03	1.00	7.667e-03	0.99	8.768e-03	0.99
160	6.675e-04	0.99	3.963e-03	0.95	4.457e-03	0.98
Scaled mixed method, n odd						
21	5.070e-03	1.01	2.893e-02	1.01	3.324e-02	1.00
41	2.592e-03	1.00	1.484e-02	1.00	1.704e-02	1.00
81	1.313e-03	1.00	7.575e-03	0.99	8.660e-03	0.99
161	6.634e-04	0.99	3.940e-03	0.95	4.430e-03	0.98
Scaled mixed method, n even, discrete norm (midpoint rule)						
20	2.649e-04	0.65	1.979e-03	0.57	7.769e-04	-0.27
40	1.236e-04	1.10	1.243e-03	0.67	8.537e-04	-0.14
80	7.927e-05	0.64	1.135e-03	0.13	8.961e-04	-0.07
160	7.481e-05	0.08	1.152e-03	-0.02	9.191e-04	-0.04
Scaled mixed method, n odd, discrete norm (midpoint rule)						
21	2.524e-04	0.68	1.886e-03	0.64	7.529e-04	-0.32
41	1.204e-04	1.11	1.215e-03	0.66	8.390e-04	-0.16
81	7.903e-05	0.62	1.125e-03	0.11	8.880e-04	-0.08
161	7.481e-05	0.08	1.147e-03	-0.03	9.148e-04	-0.04

TABLE 8.6

Quadratic porosity velocity errors. *Relative L^2 errors and convergence rates for the velocities for the scaled mixed method. When n is even, the transition to positive porosity is at a computational mesh point, but not when n is odd.*

	$\tilde{v}_f$		u		v_s	
n	L^2 error	rate	L^2 error	rate	L^2 error	rate
Scaled mixed method, n even						
20	3.663e-04	1.23	1.546e-06	1.14	1.546e-06	1.14
40	1.166e-04	1.65	5.104e-07	1.60	5.104e-07	1.60
80	3.252e-05	1.84	1.429e-07	1.84	1.429e-07	1.84
160	1.079e-05	1.59	4.342e-08	1.72	4.342e-08	1.72
Scaled mixed method, n odd						
21	3.408e-04	1.27	1.444e-06	1.19	1.444e-06	1.19
41	1.115e-04	1.67	4.886e-07	1.62	4.886e-07	1.62
81	3.179e-05	1.84	1.397e-07	1.84	1.397e-07	1.84
161	1.071e-05	1.58	4.304e-08	1.71	4.304e-08	1.71

9. Numerical Results in Two Dimensions. In this section we present numerical results for the mass lumped approximation of the scaled mixed method (6.2)–(6.5), (7.2). We use the TH spaces for the Stokes part of the system and the deal.II software library [11]. We choose a two-dimensional problem related to simulation of

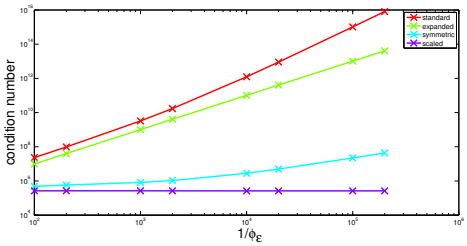


FIG. 8.4. *Compacting column. Condition numbers for the standard, expanded, symmetric, and scaled formulations as* $\phi_\epsilon \to 0^+$*, for the porosities defined in* (8.9)–(8.11) *plus* ϕ_ϵ.

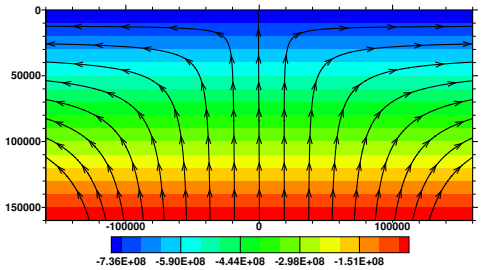


FIG. 9.1. *MOR-like example with zero porosity. We show the solid matrix potential* q_s *as a contour in* $\mathrm{Kg}/(\mathrm{m}\cdot\mathrm{s}^2)$ *and the velocity* v_s *as streamlines.*

the mantle near a mid-ocean ridge (MOR).

If porosity is constant, $\phi = \phi_0$, and one sets $\nabla \cdot \mathbf{u} = \nabla \cdot \mathbf{v}_s = 0$, then (1.1)–(1.5) can be solved in an infinite quarter-plane $\{x > 0, z > 0\}$ [37]. This problem describes viscous corner flow if, at the top of the mantle $\{z = 0\}$, one sets the MOR spreading rate as a boundary condition $\mathbf{v}_s \cdot \tau = \mathbf{v}_s \cdot \hat{\mathbf{x}} = U_0$, and on the ridge axis $\{x = 0\}$, one sets the symmetry condition $\mathbf{v}_f \cdot \nu = \mathbf{v}_f \cdot \hat{\mathbf{z}} = 0$ and $\partial \mathbf{v}_s / \partial x = 0$. The solution is

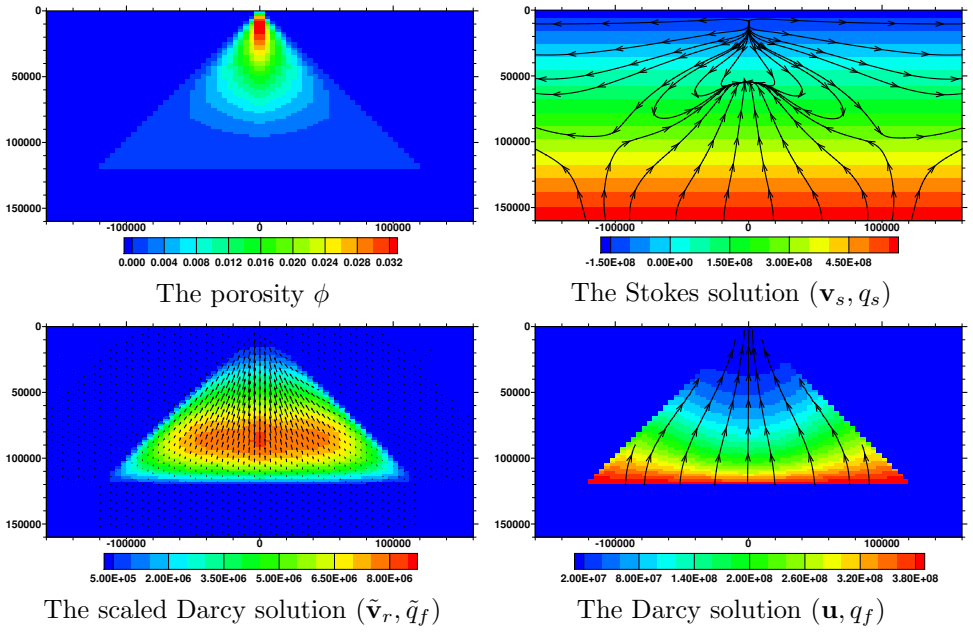
$$q = q_s = q_f = (1 - \phi_0)\Big(\frac{4\mu_s U_0}{\pi(x^2 + z^2)} + |\rho_r|\, g\Big) z, \tag{9.1}$$

$$\mathbf{v}_s = \frac{2U_0}{\pi(x^2 + z^2)} \begin{pmatrix} \tan^{-1}(x/z)\,(x^2 + z^2) - xz \\ -z^2 \end{pmatrix}, \tag{9.2}$$

$$\mathbf{u} = \frac{k_0(1 - \phi_0)\phi_0^{2+2\Theta}}{\mu_f} \Big\{ \frac{4\mu_s U_0}{\pi(x^2 + z^2)^2} \begin{pmatrix} 2xz \\ z^2 - x^2 \end{pmatrix} + \rho_r \mathbf{g} \Big\}. \tag{9.3}$$

We solve the full system of equations (i.e., the unmodified (1.1)–(1.5)) on a rectangular domain $\{(x, z) : -160\,\mathrm{km} < x < 160\,\mathrm{km}, 0 < z < 160\,\mathrm{km}\}$. The MOR is at $(0, 0)$. We take the physical quantities defined in Table 2.1, using a permeability of $10^{-8}\,\mathrm{m}^2$, $\Theta = 0$, and $U_0 = 10^{-9}\,\mathrm{m/s} = 3.1536\,\mathrm{cm/yr}$. We use boundary conditions defined by the corner flow problem. We impose the nonhomogeneous essential boundary condition defined by (9.2) on the Stokes velocity $\mathbf{v}_s$ and the nonhomogeneous natural boundary condition defined by (9.1) on the potential $q = q_f$. However, to avoid the singularity at the corner, we translate x to $x - \ell$ when $x < 0$ and $x + \ell$ when $x > 0$ before evaluating (9.1)–(9.2), where $\ell = 20\,\mathrm{km}$. We use a mesh of 160×80 elements.

In Fig. 9.1 we show the Stokes solution using $\phi = 0$. Note that the mantle flows up to the MOR and outward from there. There is no melt in this computation, although our code solves for the Darcy system as well as the Stokes system. Rather than normalizing the average of q to zero, we set a single point to zero.



The porosity ϕ | The Stokes solution $(\mathbf{v}_s, q_s)$

The scaled Darcy solution $(\tilde{\mathbf{v}}_r, \tilde{q}_f)$ | The Darcy solution $(\mathbf{u}, q_f)$

FIG. 9.2. *MOR-like example. We show the porosity and the solutions* $(\mathbf{v}_s, q_s)$ *to the Stokes system,* $(\tilde{\mathbf{v}}_r, \tilde{q}_f)$ *to the scaled Darcy system, and* $(\mathbf{u}, q_f)$ *to the Darcy system. The porosity and potentials in* $\mathrm{Kg}/(\mathrm{m}\cdot\mathrm{s}^2)$ *are shown as contours, the relative velocity* $\tilde{\mathbf{v}}_r$ *as arrows, and the other velocities as streamlines.*

We then set a nonvanishing porosity by the formula

$$\phi(x,z) = \begin{cases} 0.05\Big(\dfrac{120\,\mathrm{km} - z}{120\,\mathrm{km}}\Big)^2\Big(1 - \dfrac{|x|}{z+\ell}\Big) & \text{if } z \le 120\,\mathrm{km} \text{ and } |x| \le z + \ell, \\ 0 & \text{otherwise,} \end{cases} \tag{9.4}$$

where we took $\ell = 20\,\mathrm{km}$.

In Fig. 9.2, we show the porosity and the solutions $(\mathbf{v}_s, q_s)$ to the Stokes system, $(\tilde{\mathbf{v}}_r, \tilde{q}_f)$ to the scaled Darcy system, and $(\mathbf{u}, q_f)$ to the Darcy system. The form of the solution is dictated by our (arbitrary) choice of ϕ. The scaled and unscaled Darcy solution shows fluid melt rising and focusing into the MOR, and some melt leaving the domain to form new crust. The Stokes solution varies significantly from the case in Fig. 9.1 where there is no melt. The solid matrix rises at the bottom, but it falls at the top near the MOR to compensate for the rise of fluid melt to the surface. We note that the scaled potential $\tilde{q}_f$ is much smoother than q_f. The porosity vanishes in a significant portion of the domain; nevertheless, there is no difficulty solving the system accurately.

10. Conclusions. We developed a mathematically well-posed, mixed variational framework for McKenzie's equations governing the mechanics of a mixture of molten and solid materials [29], assuming that the porosity ϕ is given and satisfies the hypothesis (4.7). Our formulation handles the regions where there are two phases (i.e., the mixture variable $\phi > 0$) as well as the mathematically degenerate regions where there is only the single solid matrix phase (i.e., $\phi = 0$). The formulation is based on a careful scaling of the Darcy variables by powers of the porosity [5]. We also

discussed three other standard formulations, but ϕ must be positive simply for these formulations to be well-defined.

We defined a mixed finite element method (MFEM) based on our scaled variational formulation and proved its stability and optimal order convergence. We also presented a modification that is locally mass conservative, and a modification involving mass lumping (7.2) to simplify and increase solver efficiency of the implementation on rectangular meshes.

Numerical results of a one-dimensional compacting column with various porosity functions showed an excellent match to the closed form solutions for ϕ_0 and ϕ_{J}, as well as a good match to the approximate solution for ϕ_2. Degeneracies in the porosity posed no difficulties for the simulations; in fact, the condition number of the linear system is nearly insensitive to degeneracies in ϕ. The results showed that the method indeed achieves optimal convergence and that the mass lumping approximation does not degrade the results in any way.

The nondegenerate constant porosity example showed $\mathcal{O}(h)$ convergence for the potentials and superconvergence of order $\mathcal{O}(h^2)$ when measured in the discrete midpoint rule norm. The velocity achieved the optimal $\mathcal{O}(h^2)$ convergence for this one-dimensional problem. The degenerate, quadratic porosity example also showed optimal $\mathcal{O}(h)$ convergence of the potentials and perhaps $\mathcal{O}(h^2)$ convergence of the velocities, regardless of whether the computational mesh resolved the point where ϕ transitions from zero to positive.

The degenerate and discontinuous porosity example had an interesting set of results. Even though the porosity does not satisfy (4.7), the MFEM achieved good, but not necessarily optimal, convergence in all cases. When the computational mesh resolved the transition point of ϕ, we saw $\mathcal{O}(h)$ convergence for the potentials and superconvergence of order $\mathcal{O}(h^2)$ when measured in the discrete midpoint rule norm. We also saw $\mathcal{O}(h^{3/2})$ convergence for $\tilde{v}_r$ and $\mathcal{O}(h^2)$ convergence for u and v_s. The mass lumped approximation actually improved the convergence to $\mathcal{O}(h^2)$ for all three velocities. However, when the computational mesh did not resolve the transition point in ϕ, we saw $\mathcal{O}(h)$ convergence for the potentials $\tilde{q}_f$ and q_f, but only $\mathcal{O}(h^{1/2})$ for q. The discrete norm did not help, but we did verify that the main errors were localized to a region near ϕ_2 the transition point, since removing the error there led to $\mathcal{O}(h)$ convergence for all three potentials. The velocities converged to order $\mathcal{O}(h)$. This example suggests that the condition (4.7) may not be strictly necessary.

In the full model of mantle dynamics, the porosity evolves and so must be approximated. In a finite element or discontinuous Galerkin method, one would naturally approximate ϕ by continuous or discontinuous polynomials on each element of the computational mesh. Any jumps in the porosity will then naturally lie on the boundaries of the elements, and so we would expect our method to perform well.

A two-dimensional test example akin to a mid ocean ridge showed the strong effect that melt can have on the velocity field. Even though the porosity vanished in much of the domain, our locally conservative scaled finite element method showed good results. Using the mass lumped approximation, the method easily reduces to a single Stokes system with two potentials, and the efficiency of the linear solver is fairly insensitive to the absence of melt. We believe that our method is highly suited to realistic problems of the mechanics of mantle dynamics, and that it can be used effectively as a component of the full mantle dynamics problem.

REFERENCES

[1] R. A. Adams, *Sobolev Spaces*, Academic Press, 1975.
[2] E. Aharonov, J. A. Whitehead, P. B. Kelemen, and M. Spiegelman, *Channeling instability of upwelling melt in the mantle*, J. Geophysical Research, 100 (1995), pp. 20,433–20,450.
[3] T. Arbogast and M. R. Correa, *Two families of H(div) mixed finite elements on quadrilaterals of minimal dimension*, SIAM J. Numer. Anal., (2016, to appear).
[4] T. Arbogast and A. L. Taicher, *A cell-centered finite difference method for a degenerate elliptic equation arising from two-phase mixtures*, Submitted, (2016).
[5] ———, *A linear degenerate elliptic equation arising from two-phase mixtures*, SIAM J. Numer. Anal., (2016, to appear).
[6] T. Arbogast and M. F. Wheeler, *A family of rectangular mixed elements with a continuous flux for second order elliptic problems*, SIAM J. Numer. Anal., 42 (2005), pp. 1914–1931.
[7] 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.
[8] A. Aschwanden, E. Bueler, C. Khroulev, and H. Blatter, *An enthalpy formulation for glaciers and ice sheets*, J. Glaciology, 58 (2012), pp. 441–457.
[9] I. Babuška, *Error-bounds for finite element method*, Numer. Math., 16 (1971), pp. 322–333.
[10] ———, *The finite element method with Lagrangian multipliers*, Numer. Math., 20 (1973), pp. 179–192.
[11] W. Bangerth, T. Heister, L. Heltai, G. Kanschat, M. Kronbichler, M. Maier, and B. Turcksin, *The `deal.II` library, version 8.3*, Archive of Numer. Software, 4 (2016), pp. 1–11.
[12] J. Bear, *Dynamics of Fluids in Porous Media*, Dover, New York, 1972.
[13] J. Bear and A. H.-D. Cheng, *Modeling Groundwater Flow and Contaminant Transport*, Springer, New York, 2010.
[14] C. Bernardi and G. Raugel, *Analysis of some finite elements for the Stokes problem*, Math. Comp., 44 (1985), pp. 71–79.
[15] J. H. Bramble, *A proof of the inf-sup condition for the Stokes equations on Lipschitz domains*, Math. Models Methods Appl. Sci., 13 (2003), pp. 361–371.
[16] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer, New York, 1994.
[17] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer, New York, 1991.
[18] Z. Chen, G. Huan, and Y. Ma, *Computational Methods for Multiphase Flows in Porous Media*, vol. 2 of Computational Science and Engineering Series, SIAM, Philadelphia, 2006.
[19] B. Cockburn and C.-W. Shu, *Runge–Kutta discontinuous Galerkin methods for convection-dominated problems*, J. Sci. Comput., 16 (2001), pp. 173–261.
[20] J. Douglas, Jr., T. Dupont, and L. Wajlbin, *The stability in L^q of the L^2-projection into finite element function spaces*, Numer. Math., 23 (1975), pp. 193–197.
[21] A. Ern and J.-L. Guermond, *Theory and practice of finite elements*, Applied mathematical sciences, Springer, New York, 2004.
[22] A. C. Fowler, *On the transport of moisture in polythermal glaciers*, Geophys. Astrophys. Fluid Dynamics, 28 (1984), pp. 99–140.
[23] V. Girault and P. A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer-Verlag, Berlin, 1986.
[24] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.
[25] M. A. Hesse, A. R. Schiemenz, Y. Liang, and E. M. Parmentier, *Compaction-dissolution waves in an upwelling mantle column*, Geophysical J. Int., 187 (2011), pp. 1057–1075.
[26] I. J. Hewitt and A. C. Fowler, *Partial melting in an upwelling mantle column*, Proc. R. Soc. A, 464 (2008), pp. 2467–2491.
[27] R. F. Katz, *Magma dynamics with the enthalpy method: Benchmark solutions and magmatic focusing at mid-ocean ridges*, J. Petrology, 49 (2008), pp. 2099–2121.
[28] L. W. Lake, *Enhanced Oil Recovery*, Prentice Hall, Englewood Cliffs, New Jersey, 1989.
[29] D. McKenzie, *The generation and compaction of partially molten rock*, J. Petrology, 25 (1984), pp. 713–765.
[30] ———, *Th-U disequilibrium and the melting processes beneath ridge axes*, Earth and Planetary Sci. Lett., 72 (1985), pp. 149–157.
[31] J. T. Oden and L. F. Demkowicz, *Applied Functional Analysis*, CRC Press, 1996.
[32] R. A. Raviart and J. M. Thomas, *A mixed finite element method for 2nd order elliptic problems*, in Mathematical Aspects of Finite Element Methods, I. Galligani and E. Magenes, eds., no. 606 in Lecture Notes in Math., Springer-Verlag, New York, 1977, pp. 292–315.
[33] S. Rhebergen, G. N. Wells, R. F. KATZ, and A. J. Wathen, *Analysis of block precondi-*

tioners for models of coupled magma/mantle dynamics, SIAM J. Sci. Comput., 36 (2014), pp. A1960–A1977.

[34] J. E. Roberts and J.-M. Thomas, *Mixed and hybrid methods*, in Handbook of Numerical Analysis, P. G. Ciarlet and J. L. Lions, eds., vol. 2, North-Holland, Amsterdam, 1991, pp. 523–639.

[35] N. Sleep, *Tapping of melt by veins and dikes*, J. Geophys. Res., 93 (1988), pp. 255–272.

[36] E. A. Spiegel and G. Veronis, *On the boussinesq approximation for a compressible fluid*, Astophysical J., 131 (1960), pp. 442–447.

[37] M. Spiegelman and D. McKenzie, *Simple 2-D models for melt extraction at mid-ocean ridges and island arcs*, Earth and planetary sci. lett., 83 (1987), pp. 137–152.

[38] A. L. Taicher, *Mixed framework for Darcy-Stokes mixtures*, PhD thesis, Univ. of Texas at Austin, 2014.

[39] C. J. van der Veen, *Fundamentals of Glacier Dynamics*, CRC Press, second ed., 2013.

[40] W. Zhu, G. Gaetani, F. Fusseis, L. Montesi, and F. D. Carlo, *Microtomography of partiallty molten rocks: three dimensional melt distribution in mantle peridotite*, Science, 332 (2011), pp. 88–91.