Comput Geosci (2007) 11:145–158
DOI 10.1007/s10596-007-9044-z

ORIGINAL PAPER

A coupling of mixed and continuous Galerkin finite element methods for poroelasticity II: the discrete-in-time case

Phillip Joseph Phillips · Mary F. Wheeler

Received: 24 May 2005 / Accepted: 1 December 2006 / Published online: 23 March 2007

Abstract In this paper, we formulate a finite element procedure for approximating the coupled fluid and mechanics in Biot's consolidation model of poroelasticity. Here, we approximate the pressure by a mixed finite element method and the displacements by a Galerkin method. Theoretical convergence error estimates are derived in a discrete-in-time setting. Of particular interest is the case when the lowest-order Raviart–Thomas approximating space or cell-centered finite differences are used in the mixed formulation and continuous piecewise linear approximations are used for displacements. This approach appears to be the one most frequently applied to existing reservoir engineering simulators.

Keywords Continuous Galerkin · Discrete-in-time a priori error estimates · Mixed finite elements · Poroelasticity

1 Introduction

In part 1 of this two-part paper [11], we provide an overview of the growing importance of poroelasticity modeling in a diverse range of fields. Indeed, ideas from poroelasticity may have emerged naturally from concerns in the soil mechanics community, but today, those same ideas are contributing to important achievements in reservoir engineering, earthquake engineering, and biomedicine.

The contribution of the prequel was the novel algorithm based on a coupling of continuous Galerkin and mixed methods. Therein, we present optimal a priori continuous-in-time error estimates, and discuss the mass conservation feature of our algorithm.

Here, we turn to the discrete time setting and present optimal a priori discrete-in-time error estimates. We also compare our results against the well-known 1-D problem of Terzaghi, and we also take a look at the more recent 2-D problem of Barry and Mercer.

But first, for convenience, we recapitulate the essential aspects of the Biot's consolidation model and our finite element formulation found in part 1.

P. J. Phillips
Center for Subsurface Modeling (CSM),
Institute for Computational Engineering and Sciences
(ICES), University of Texas at Austin, Austin, Texas, USA

M. F. Wheeler (✉)
CSM, ICES, Department of Aerospace Engineering
and Engineering Mechanics, Department of Petroleum
Engineering and Geosystems Engineering,
University of Texas at Austin, Austin, Texas, USA
e-mail: mfw@ices.utexas.edu

2 Model formulation

2.1 Derivation of the model

The poroelasticity equations consist of a momentum and a mass conservation equation and are derived at the macroscopic scale in the work of Terzaghi [17] and M. A. Biot [3, 4]. In our discussion of the equations, we follow the presentation of Showalter [15].

The momentum conservation is similar to that found in linear elasticity, the exception being the addition of a fluid pressure term. Because the deformation of the material is usually much slower than the flow rate, a quasistatic assumption is made, so that the second time derivative for displacements is ignored. In order

to derive the momentum equation, let $\Omega \subset R^d$, and choose V to be a fixed, arbitrary open subset of Ω. Then, for the total stress tensor, $\tilde{\boldsymbol{\sigma}}$, and a body force, $\mathbf{f}$, we have

$$-\int_{\partial V} \tilde{\boldsymbol{\sigma}}\, \boldsymbol{\nu}\, ds = \int_V \mathbf{f}\, dV,$$

where $\boldsymbol{\nu}$ is the outward normal. Using the divergence theorem on the left side allows us to conclude:

$$-\int_V \nabla \cdot \tilde{\boldsymbol{\sigma}}\, dV = \int_V \mathbf{f}\, dV.$$

Because V was chosen arbitrarily, it follows that $-\nabla \cdot \tilde{\boldsymbol{\sigma}} = \mathbf{f}$ over Ω.

Turning to the mass conservation equation, we again let $V \subset \Omega$ be a fixed, arbitrary open subset of Ω. We refer to η as the fluid content of the medium, $\mathbf{v}_f$ as its flux, and h as the volumetric fluid source term. We then conclude from elementary conservation principles

$$\frac{\partial}{\partial t}\int_V \eta\, dV = -\int_{\partial V} \mathbf{v}_f \cdot \boldsymbol{\nu}\, ds + \int_V h\, dV.$$

Then, the divergence on the first term on the right side of the above equation and the fact that V was chosen arbitrarily imply that

$$\frac{\partial \eta}{\partial t} = -\nabla \cdot \mathbf{v}_f + h.$$

In order to close the model, constitutive relations must be formulated to relate the total stress, $\tilde{\boldsymbol{\sigma}}$; flux, $\mathbf{v}_f$; and fluid content, η to the primary variables pressure, p, and deformation, $\mathbf{u}$. The total stress must account for the usual material stress as in elasticity and for the fluid pressure; consequently, we set $\tilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - \alpha \nabla p$. Here, $\boldsymbol{\sigma}$ is the standard stress tensor from elasticity (which we assume to be a linear function of $\mathbf{u}$), and is referred to as the "effective stress" in the field of poroelasticity. The pressure term measures the effect of fluid of the material medium; an increase in pressure generally causes an expansion. The standard assumption of *Darcy's* law from porous media holds for the flux, and we set $\mathbf{v}_f = -\frac{1}{\mu_f}\boldsymbol{\kappa}(\nabla p - \rho_f \mathbf{g})$. The permeability tensor, $\boldsymbol{\kappa}$, is assumed to be uniformly bounded and uniformly elliptic; that is, there exist positive constants λ_{min} and λ_{max} such that for all $x \in \Omega$, the following relation holds:

$$\lambda_{min}||\boldsymbol{\xi}||^2 \;\leq\; \boldsymbol{\xi}^t \boldsymbol{\kappa}(\mathbf{x})\boldsymbol{\xi} \;\leq\; \lambda_{max}||\boldsymbol{\xi}||^2. \tag{2.1}$$

It is natural to assume that the amount of fluid content, η, depends on the fluid pressure, p, and the material volume, which is measured locally by $\nabla \cdot \mathbf{u}$. More specifically, we set $\eta = c_o p + \alpha \nabla \cdot \mathbf{u}$, and observe that $c_o p$ measures the amount of fluid that can be injected into a fixed material volume, and $\alpha \nabla \cdot \mathbf{u}$ measures the amount of fluid that can be squeezed out. The constrained specific storage coefficient, c_o, may, in general, be zero, but in this paper, it is assumed to be strictly positive:

$$c_0 \geq \gamma_o > 0 \tag{2.2}$$

for some positive constant γ_o.

The momentum and mass conservation equations are coupled through the *Biot–Willis* constant, α, which is usually close to unity. The material parameters are usually found experimentally or through some homogenization technique (Tables 1 and 2).

Table 1 Summary of constitutive relations

$\tilde{\sigma}_{ij}(\mathbf{u}, p) = \sigma_{ij}(\mathbf{u}) - \alpha\delta_{ij}p$	Total stress
$\sigma_{ij}(\mathbf{u}) = \lambda\delta_{ij}\epsilon_{kk}(\mathbf{u}) + 2\mu\epsilon_{ij}(\mathbf{u})$	Effective stress
$\eta = c_o p + \alpha \nabla \cdot \mathbf{u}$	Fluid content
$\mathbf{v}_f = -\frac{1}{\mu_f}\boldsymbol{\kappa}(\nabla p - \rho_f \mathbf{g})$	Volumetric fluid flux

Table 2 Summary of physical parameters

Parameter	Description
λ, μ	Positive Lamé constants
c_o	Constrained Specific Storage Coefficient
α	Biot–Willis
$\boldsymbol{\kappa}$	Symmetric permeability tensor
μ_f	Fluid viscosity

2.2 Summary of the poroelasticity equations

We summarize the governing equations below, and complete the model by including the necessary boundary and initial conditions. There are two distinct sets of boundary conditions, one corresponding to the flow and one corresponding to the deformation.

$$-(\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) - \mu\nabla^2\mathbf{u} + \alpha\nabla p = \mathbf{f}, \tag{2.3a}$$

$$\frac{\partial}{\partial t}(c_o p + \alpha \nabla \cdot \mathbf{u}) - \frac{1}{\mu_f}\nabla \cdot \boldsymbol{\kappa}(\nabla p - \rho_f \mathbf{g}) = h, \tag{2.3b}$$

$$p(t) = p_o \quad \text{on} \quad \Gamma_p, \tag{2.4a}$$

$$-\frac{1}{\mu_f}\boldsymbol{\kappa}(\nabla p - \rho_f \mathbf{g}) \cdot \boldsymbol{\nu} = q \quad \text{on} \quad \Gamma_f, \tag{2.4b}$$

$$\mathbf{u}(t) = \mathbf{u}_D \quad \text{on} \quad \Gamma_o, \tag{2.4c}$$

$$\tilde{\boldsymbol{\sigma}}\, \boldsymbol{\nu} = \mathbf{t}_N \quad \text{on} \quad \Gamma_t, \tag{2.4d}$$

$$p(0) = p^o, \tag{2.4e}$$

$$\mathbf{u}(0) = \mathbf{u}^o, \tag{2.4f}$$

where $\partial\Omega = \Gamma_p \bigcup \Gamma_f$ and $\partial\Omega = \Gamma_t \bigcup \Gamma_o$. Also, $\boldsymbol{\nu}$ represents the outward normal to $\partial\Omega$.

Remark 2.1 Often in practice, initial conditions are found by first setting $p(0)$ equal to the hydrostatic pressure (which amounts to solving $\nabla p(0) = \rho_f \mathbf{g}$), and then using $p(0)$ in Eq. 2.3a to solve for $\mathbf{u}(0)$.

3 The coupled formulation

3.1 Bilinear form and function spaces

For the mixed formulation developed herein, the appropriate function space for the pressure is $L^2(\Omega)$. The space used for the flux variable is $H(\text{div}) \equiv \{\mathbf{s} \in (L^2(\Omega))^d : \nabla \cdot \mathbf{s} \in L^2(\Omega)\}$. With $H(\text{div})$, we then define the following subset:

$$\mathbf{S}_0 \equiv \{\mathbf{s} \in H(\text{div}) : \mathbf{s} \cdot \boldsymbol{\nu}|_{\Gamma_t} = 0\}.$$

The function space relevant to the deformation is

$$\mathbf{V}_0 \equiv \{\mathbf{v} \in (H^1(\Omega))^d : \mathbf{v}|_{\Gamma_o} = 0\}.$$

Associated with this space is the bilinear form defined as

$$a_{\mathbf{u}}(\mathbf{u}, \mathbf{v}) \equiv \int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \epsilon(\mathbf{v}) dx.$$

The linearized *strain* tensor is given by $\epsilon_{kl} \equiv \frac{1}{2}(\partial_k u_l + \partial_l u_k)$, and thus,

$$a_{\mathbf{u}}(\mathbf{u}, \mathbf{v}) = \int_\Omega (2\mu(\epsilon(\mathbf{u}) : \epsilon(\mathbf{v})) + \lambda(\nabla \cdot \mathbf{u})(\nabla \cdot \mathbf{v})) dx. \tag{3.1}$$

Here, $(\sigma : \tau) = \sum_i \sum_j \sigma_{ij}\tau_{ij}$. Clearly $a_{\mathbf{u}}(,)$ is symmetric and continuous. Furthermore, given that $|\Gamma_o| > 0$, *Korn's inequality* shows that $a_{\mathbf{u}}(,)$ is also coercive on $\mathbf{V}_0$ (see Brenner and Scott [5]). With the energy norm defined as $||\mathbf{u}||^2_{a_{\mathbf{u}}} \equiv a_{\mathbf{u}}(\mathbf{u}, \mathbf{u})$, the following inequalities hold for some positive real numbers C_{cont} and C_{coer}:

$$a_{\mathbf{u}}(\mathbf{u}, \mathbf{v}) \le C_{\text{cont}} ||\mathbf{u}||_{H^1} ||\mathbf{v}||_{H^1}, \quad \forall \mathbf{u}, \mathbf{v} \in H^1, \tag{3.2}$$

$$C_{\text{coer}} ||\mathbf{u}||^2_{H^1} \le ||\mathbf{u}||^2_{a_{\mathbf{u}}}, \quad \forall \mathbf{u} \in \mathbf{V}_0. \tag{3.3}$$

We now provide some definitions important for the development of our finite element scheme. Our notation follows Rivière and Wheeler [14]. Let $\mathcal{E}_h = \{E_1, E_2, ..., E_{M_h}\}$ be a nondegenerate subdivision of Ω, where E_j is a triangle or a convex quadrilateral for $d = 2$, or a tetrahedron if $d = 3$. Let $h_j = \text{diam}(E_j)$; then, nondegeneracy requires the existence of $\rho > 0$ such that E_j contains a ball of radius ρh_j. We also set $h = \max\{h_j : j = 1, ..., M_h\}$.

The finite dimensional approximating spaces are defined as follows. Let $(W_h, \mathbf{S}_h) \subset (L^2 \times H(\text{div}))$ denote a standard mixed finite element space defined on $\mathcal{E}_h$; in particular, let $\mathbf{S}_{h,0} \equiv \{\mathbf{s} \in \mathbf{S}_h : \mathbf{s} \cdot \boldsymbol{\nu}|_{\Gamma_f} = 0\}$. Let k refer to the order of this space. Additionally, these spaces are required to be endowed with two linear operators, $\Pi_h, H(\text{div}) \to \mathbf{S}_h$, and the L^2 projection, $L^2 \to W_h$, which satisfy the following properties:

$$(\nabla \cdot (\mathbf{s} - \Pi_h \mathbf{s}), w) = 0, \quad \forall w \in W_h, \tag{3.4a}$$

$$||\mathbf{s} - \Pi_h \mathbf{s}||_{L^2} \le C h^r ||\mathbf{s}||_{H^r}, \quad 1 \le r \le k+1, \tag{3.4b}$$

$$\nabla \cdot \Pi_h = P_h \nabla \cdot, \tag{3.4c}$$

$$(\nabla \cdot \mathbf{s}_h, p - P_h p) = 0, \quad \forall \mathbf{s}_h \in \mathbf{S}_h, \tag{3.4d}$$

$$||p - P_h p||_{L^2} \le C h^r ||p||_{H^r}, \quad 0 \le r \le k+1. \tag{3.4e}$$

The Raviart–Thomas–Nedelec space of order $k = 0$ (see Raviart and Thomas [12] and Nedelec [10]) is one example and is used for the numerical examples that follow. Note that not all mixed-spaces operators satisfy each of the above properties, in particular, Eq. 3.4b, where the upper bound for r is sometimes only k.

Regarding the deformation space, we let $\mathbf{V}_h$ be the space of continuous piecewise polynomials of degree r; set $\mathbf{V}_{h,0} \equiv \{\mathbf{v} \in \mathbf{V}_h : \mathbf{v}|_{\Gamma_o} = 0\}$. The elliptic projection operator $\tilde{P}$, $(H^1)^d \to \mathbf{V}_h$ (see Wheeler [18]), will be used in the following theorem; $\tilde{P}$ satisfies

$$a_{\mathbf{u}}(\mathbf{u} - \tilde{P}\mathbf{u}, \mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in \mathbf{V}_h.$$

The following inequality describes its approximation properties:

$$||\mathbf{u} - \tilde{P}\mathbf{u}||_{a_{\mathbf{u}}} \le C h^{s-1}, \quad 1 \le s \le r+1. \tag{3.5}$$

3.2 Coupled algorithm

Corresponding to Eqs. 2.3a and 2.3b, Phillips and Wheeler [11] (submitted for publication) produce a continuous-in-time, fully coupled algorithm. They begin by making the following assumptions on the problem data:

$$\mathbf{f} \in C^1([0, T]; (H^{-1}(\Omega))^d), \tag{3.6}$$

$$h \in C^0([0, T]; L^2(\Omega)), \tag{3.7}$$

$$p_o \in C^0([0,\,T];\,L^2(\Gamma_p)), \tag{3.8}$$

$$q \in C^0([0,\,T];\,TrS), \tag{3.9}$$

$$\mathbf{u}_D \in C^1([0,\,T];\,(H^{1/2}(\Gamma_o))^d), \tag{3.10}$$

$$\mathbf{t}_N \in C^1([0,\,T];\,(H^{-1/2}(\Gamma_t))^d), \tag{3.11}$$

where $TrS \equiv \{\mathbf{s}\cdot\boldsymbol{\nu}|_{\Gamma_f} : \mathbf{s} \in H(\mathrm{div})\}$. As for the initial time data, they also assume that $\mathbf{u}^o \in (H^1)^d$ and $p^o \in L^2$. They work in affine spaces because the essential boundary conditions for the displacement and flux are allowed to be inhomogeneous. So, for each $t \geq 0$, they select some $\mathbf{u}_d(t,\mathbf{x}) \in (H^1)^d$ such that $\mathbf{u}_d(t,\mathbf{x})|_{\Gamma_o} = \mathbf{u}_D(t,x)$, and choose some $\mathbf{z}_d(t,\mathbf{x}) \in H(\mathrm{div})$ such that $\mathbf{z}_d(t,\mathbf{x})\cdot\boldsymbol{\nu}|_{\Gamma_f} = q(t,\mathbf{x})$. Then, they define $\bar{\mathbf{u}}_d(t,\mathbf{x}) = \tilde{P}\mathbf{u}_d(t,\mathbf{x})$ and $\bar{\mathbf{z}}_d(t,\mathbf{x}) = \Pi_h\mathbf{z}_d(t,\mathbf{x})$. Their scheme then becomes: Find $\bar{\mathbf{u}} \in \bar{\mathbf{u}}_d + H^1([0,\,T];\,\mathbf{V}_{h,0})$, $\bar{p} \in H^1([0,\,T];\,W_h)$ and $\bar{\mathbf{z}} \in \bar{\mathbf{z}}_d + L^2([0,\,T];\,\mathbf{S}_{h,0})$ such that

$$a_{\mathbf{u}}(\bar{\mathbf{u}},\mathbf{v}) - \alpha(\nabla\cdot\mathbf{v},\,\bar{p}) = l_1(\mathbf{v}), \tag{3.12}$$

$$(c_o\bar{p}_t,\,w) + \alpha(\nabla\cdot\bar{\mathbf{u}}_t,\,w) + (\nabla\cdot\bar{\mathbf{z}},\,w) = l_2(w), \tag{3.13}$$

$$(\tilde{\boldsymbol{\kappa}}^{-1}\bar{\mathbf{z}},\,\mathbf{s}) - (\bar{p},\,\nabla\cdot\mathbf{s}) = l_3(\mathbf{s}), \tag{3.14}$$

holds for every $t \in [0,\,T]$ and for all $(\mathbf{v},\,w,\,\mathbf{s}) \in (\mathbf{V}_{h,0},\,W_h,\,\mathbf{S}_{h,0})$. Additionally, they choose the following initial conditions for $\bar{\mathbf{u}}$ and $\bar{p}$:

$$a_{\mathbf{u}}(\bar{\mathbf{u}},\mathbf{v})|_{t=0} = a_{\mathbf{u}}(\mathbf{u}^o,\mathbf{v}), \quad \forall\mathbf{v} \in \mathbf{V}_h, \tag{3.15}$$

$$(\bar{p},\,w)|_{t=0} = (p^o,\,w), \quad \forall w \in W_h. \tag{3.16}$$

Here, l_1, l_2, and l_3 are bounded linear functionals defined as lf1

$$l_1(\mathbf{v}) = \int_\Omega \mathbf{f}\cdot\mathbf{v} + \int_{\Gamma_t}\mathbf{t}_N\cdot\mathbf{v}, \tag{3.17}$$

$$l_2(w) = \int_\Omega h\,w, \tag{3.18}$$

$$l_3(\mathbf{s}) = -\int_{\Gamma_p} p_o\mathbf{s}\cdot\boldsymbol{\nu} + \int_\Omega \rho_f\mathbf{g}\cdot\mathbf{s}. \tag{3.19}$$

We now complete the numerical formulation by discretizing our algorithm (Eqs. 3.12–3.14) in time. To do so, we use the theta method from numerical differential equations. We start with some definitions. Let $\Delta t = T/N$ where N is a positive integer and let $t_j = j\Delta t$. We then use the following notation for $\theta \in [0,\,1]$:

$$g_j = g(x,t_j), \quad 0 \leq j \leq N, \tag{3.20}$$

$$g_{j,\theta} = \frac{1}{2}(1+\theta)g_{j+1} + \frac{1}{2}(1-\theta)g_j, \quad 0 \leq j \leq N-1. \tag{3.21}$$

With this, Eqs. 3.12–3.14 become: Find $\bar{\mathbf{u}}_j \in \bar{\mathbf{u}}_{d,j} + \mathbf{V}_{h,0}$, $\bar{p}_j \in W_h$, and $\bar{\mathbf{z}}_j \in \bar{\mathbf{z}}_{d,j} + \mathbf{S}_{h,0}$, such that

$$a_{\mathbf{u}}(\bar{\mathbf{u}}_{j,\theta},\mathbf{v}_h) - \alpha(\nabla\cdot\mathbf{v}_h,\,\bar{p}_{j,\theta}) = l_{1\,j,\theta}(\mathbf{v}_h), \tag{3.22}$$

$$c_o\left(\frac{\bar{p}_{j+1}-\bar{p}_j}{\Delta t},\,w_h\right) + \alpha\left(\nabla\cdot\left(\frac{\bar{\mathbf{u}}_{j+1}-\bar{\mathbf{u}}_j}{\Delta t}\right),\,w_h\right) + (\nabla\cdot\bar{\mathbf{z}}_{j,\theta},\,w_h) = l_{2\,j,\theta}(w_h), \tag{3.23}$$

$$(\tilde{\boldsymbol{\kappa}}^{-1}\bar{\mathbf{z}}_{j,\theta},\,\mathbf{s}_h) - (\bar{p}_{j,\theta},\,\nabla\cdot\mathbf{s}_h) = l_{3\,j,\theta}(\mathbf{s}_h), \tag{3.24}$$

for all $(\mathbf{v}_h,\,w_h,\,\mathbf{s}_h) \in (\mathbf{V}_{h,0},\,W_h,\,\mathbf{S}_{h,0})$. The right-hand sides are determined by Eqs. 3.17–3.19 in accordance with the notation set forth in Eq. 3.21. We mention that $\theta = 1$ corresponds to the backward Euler method and $\theta = 0$ to the Crank–Nicolson (C-N) scheme.

We now write, for each time $t = j\Delta t$, the functions $\bar{\mathbf{u}}_j(\mathbf{x})$, $\bar{p}_j(\mathbf{x})$, and $\bar{\mathbf{z}}_j(\mathbf{x})$ as components in their respective finite element basis functions, $\mathbf{N}_{\mathbf{u}} = [N_{\mathbf{u},1},\ldots,N_{\mathbf{u},n_{\mathbf{u}}}]$, $\mathbf{N}_p = [N_{p,1},\ldots,N_{p,n_p}]$, and $\mathbf{N}_{\mathbf{z}} = [N_{\mathbf{z},1},\ldots,N_{\mathbf{z},n_{\mathbf{z}}}]$:

$$\begin{aligned}\bar{\mathbf{u}}_j(\mathbf{x}) &= \sum_k u_{j,k}N_{\mathbf{u},k}(\mathbf{x}) + \sum_k u_{d,j,k}(t)N_{\mathbf{u},k}(\mathbf{x})\\ &= \mathbf{u}_{h,j}(\mathbf{x})\cdot\mathbf{N}_{\mathbf{u}}(\mathbf{x}) + \mathbf{u}_{dh,j}(t)\cdot\mathbf{N}_{\mathbf{u}}(\mathbf{x}),\end{aligned}$$

$$\begin{aligned}\bar{p}_j(\mathbf{x}) &= \sum_k p_{j,k}N_{p,k}(\mathbf{x})\\ &= \mathbf{p}_{h,j}\cdot\mathbf{N}_p(\mathbf{x}),\end{aligned}$$

$$\begin{aligned}\bar{\mathbf{z}}_j(\mathbf{x}) &= \sum_k z_{j,k}N_{\mathbf{z},k}(\mathbf{x}) + \sum_k z_{d,j,k}N_{\mathbf{z},k}(\mathbf{x})\\ &= \mathbf{z}_{h,j}\cdot\mathbf{N}_{\mathbf{z}}(\mathbf{x}) + \mathbf{z}_{dh,j}\cdot\mathbf{N}_{\mathbf{z}}(\mathbf{x}).\end{aligned}$$

Here, the vectors $\mathbf{u}_{h,j} = [u_{j,1},\ldots,u_{j,n_{\mathbf{u}}}]^T$, $\mathbf{p}_{h,j} = [p_{j,1},\ldots,p_{j,n_p}]^T$, and $\mathbf{z}_{h,j} = [z_{j,1},\ldots,z_{j,n_{\mathbf{z}}}]^T$. The vectors $\mathbf{u}_{dh,j} = [u_{d,j,1},\ldots,u_{d,j,n_{\mathbf{u}}}]^T$ and $\mathbf{z}_{dh,j} = [z_{d,j,1},\ldots,z_{d,j,n_{\mathbf{u}}}]^T$ are the time $t = j\Delta t$ components of known functions $\bar{\mathbf{u}}_d$ and $\bar{\mathbf{z}}_d$, respectively, that come from the inhomogeneous essential conditions.

Rearranging Eqs. 3.22–3.24 and reverting to matrix form yields

$$\begin{pmatrix} \frac{1}{2}(1+\theta)A_{\mathbf{uu}} & -\alpha\frac{1}{2}(1+\theta)A_{\mathbf{u}p}^T & 0 \\ \alpha A_{\mathbf{u}p} & c_o A_{pp} & \frac{\Delta t}{2}(1+\theta)A_{p\mathbf{z}}^T \\ 0 & -\frac{1}{2}(1+\theta)A_{p\mathbf{z}} & \frac{1}{2}(1+\theta)A_{\mathbf{zz}} \end{pmatrix} \begin{pmatrix} \mathbf{u}_{h,j+1} \\ \mathbf{p}_{h,j+1} \\ \mathbf{z}_{h,j+1} \end{pmatrix} 0 = \begin{pmatrix} b_{1,j} \\ b_{2,j} \\ b_{3,j} \end{pmatrix},$$

where the right-hand side $\left(b_{1,j}\, b_{2,j}\, b_{3,j}\right)^T$ satisfies:

$$\begin{pmatrix} b_{1,j} \\ b_{2,j} \\ b_{3,j} \end{pmatrix} = \begin{pmatrix} l_{1j,\theta}(\mathbf{v}_h) \\ (\Delta t) l_{2j,\theta}(w_h) \\ l_{3j,\theta}(\mathbf{s}_h) \end{pmatrix} - \begin{pmatrix} \frac{1}{2}(1+\theta)A_{\mathbf{uu}} & -\alpha\frac{1}{2}(1+\theta)A_{\mathbf{u}p}^T & 0 \\ \alpha A_{\mathbf{u}p} & c_o A_{pp} & \frac{\Delta t}{2}(1+\theta)A_{p\mathbf{z}}^T \\ 0 & -\frac{1}{2}(1+\theta)A_{p\mathbf{z}} & \frac{1}{2}(1+\theta)A_{\mathbf{zz}} \end{pmatrix} \begin{pmatrix} \mathbf{u}_{dh,j+1} \\ \mathbf{0} \\ \mathbf{z}_{dh,j+1} \end{pmatrix}$$
$$- \begin{pmatrix} \frac{1}{2}(1-\theta)A_{\mathbf{uu}} & -\alpha\frac{1}{2}(1-\theta)A_{\mathbf{u}p}^T & 0 \\ -\alpha A_{\mathbf{u}p} & -c_o A_{pp} & \frac{\Delta t}{2}(1-\theta)A_{p\mathbf{z}}^T \\ 0 & -\frac{1}{2}(1-\theta)A_{p\mathbf{z}} & \frac{1}{2}(1-\theta)A_{\mathbf{zz}} \end{pmatrix} \begin{pmatrix} \mathbf{u}_{h,j} + \mathbf{u}_{dh,j} \\ \mathbf{p}_{h,j} \\ \mathbf{z}_{h,j} + \mathbf{z}_{dh,j} \end{pmatrix}.$$

We remark that the matrix $A_{\mathbf{uu}}$ can be found by using existing linear elasticity software. Likewise, the matrices $A_{\mathbf{zz}}$ and $A_{p\mathbf{z}}$ can be obtained using existing software for mixed formulations of elliptic equations. The only matrices that must be independently assembled are A_{pp} and the coupling matrix, $A_{\mathbf{u}p}$.

We also comment on existence and uniqueness. Part 1 of our paper proved the existence and uniqueness of the continuous-in-time scheme (Eqs. 3.12–3.14). For existence and uniqueness of Eqs. 3.22–3.24, one examines at the existence and uniqueness at each time step and proceeds by induction. In particular, assume that all of the data are zero and solve for the first time step. We see from the first and third rows of the matrix formulation that

$$\mathbf{u}_{h,1} = c_1 A_{\mathbf{uu}}^{-1} A_{\mathbf{u}p}^T \mathbf{p}_{h,1}, \tag{3.25}$$

$$\mathbf{z}_{h,1} = c_2 A_{\mathbf{zz}}^{-1} A_{p\mathbf{z}} \mathbf{p}_{h,1}. \tag{3.26}$$

Here, c_1 and c_2 are constants related to the physical parameters and θ. Plugging in these values for $\mathbf{u}_{h,1}$ and $\mathbf{z}_{h,1}$ into the second row of the matrix and solving for $\mathbf{p}_{h,1}$ leads to

$$\left(c_o A_{pp} + c_3 A_{\mathbf{u}p} A_{\mathbf{uu}}^{-1} A_{\mathbf{u}p}^T + c_4 A_{p\mathbf{z}}^T A_{\mathbf{zz}}^{-1} A_{p\mathbf{z}}\right)\mathbf{p}_{h,1} = \mathbf{0}, \tag{3.27}$$

where we make the same comments regarding the constants c_3 and c_4. Now, the matrix multiplying $\mathbf{p}_{h,1}$ in Eq. 3.27 is the sum of three symmetric matrices. In addition, the first matrix is positive definite (recall also c_o is assumed to be positive), and the second and third matrices in the sum are nonnegative definite. Thus, the matrix multiplying $\mathbf{p}_{h,1}$ is symmetric and positive definite, and thus invertible. Hence, we conclude that $\mathbf{p}_{h,1} = \mathbf{0}$. This establishes uniqueness for $\mathbf{p}_{h,1}$, and because we are dealing with a finite-dimensional square system, the existence of $\mathbf{p}_{h,1}$ also follows. Likewise, from Eqs. 3.25 and 3.26, the uniqueness and existence of $\mathbf{u}_{h,1}$ and $\mathbf{z}_{h,1}$, respectively, also follow.

Now, for the inductive step, assume that, for any time $t = j\Delta t$, all data are zero and $\mathbf{u}_{h,j} = \mathbf{p}_{h,j} = \mathbf{z}_{h,j} = \mathbf{0}$. Then, the exact same argument as above proves the uniqueness and existence of $\mathbf{u}_{h,j+1}$, $\mathbf{p}_{h,j+1}$, and $\mathbf{z}_{h,j+1}$.

Alternatively, Lipnikov [8] proves the unique solvability of Eqs. 3.22–3.24 at each time step by showing that certain bilinear operators satisfy Brezzi's first stability condition and the Ladyshenskaya–Babuska–Brezzi inf – sup condition. He also demonstrates that our choice of finite-dimensional spaces satisfies the discrete forms of the above conditions, and thus, our

problem has a unique solution. His results are of special interest because they incorporate the case where $c_o \to 0$.

3.3 Error estimates

We first recall here some notation and the continuous-in-time error estimates found in the prequel to this paper:

$$E_p^A = P_h p - \bar{p}, \qquad E_p^I = p - P_h p,$$

$$E_{\mathbf{z}}^A = \Pi \mathbf{z} - \bar{\mathbf{z}}, \qquad E_{\mathbf{z}}^I = \mathbf{z} - \Pi \mathbf{z},$$

$$E_{\mathbf{u}}^A = \tilde{P}\mathbf{u} - \bar{\mathbf{u}}, \qquad E_{\mathbf{u}}^I = \mathbf{u} - \tilde{P}\mathbf{u}.$$

Theorem 3.1 (Continuous-in-time auxiliary error estimate) *Let r_1 be associated with the degree of the polynomials used in the mixed space $(W_h, \mathbf{S}_h)$ satisfying Eqs. 3.4a–3.4e, and let r_2 be the degree of the polynomials used in the displacement space $\mathbf{V}_h$ satisfying Eq. 3.5. Then, assuming Eqs. 2.1 and 2.2 and sufficient regularity in the true solution,*

$$\begin{aligned} &||E_{\mathbf{u}}^A||^2_{L^\infty(H^1)} + ||E_p^A||^2_{L^\infty(L^2)} + ||E_{\mathbf{z}}^A||^2_{L^2(L^2)} \\ &\quad \le C(h^{2r_1+2} + h^{2r_2})b \end{aligned} \tag{3.28}$$

where $C = C(T, \boldsymbol{\kappa}, c_o, C_{\text{coer}}, p, p_{,t}, \mathbf{z}, \mathbf{u}_{,t})$.

The importance of auxiliary error estimates is found when coupled with the triangle inequality, which is summarized in the following corollary:

Corollary 3.2 (Continuous-in-time finite element error estimate) *With the same conditions as in the preceding theorem, the following finite element error estimate holds:*

$$\begin{aligned} &||\mathbf{u} - \bar{\mathbf{u}}||^2_{L^\infty(H^1)} + ||p - \bar{p}||^2_{L^\infty(L^2)} + ||\mathbf{z} - \bar{\mathbf{z}}||^2_{L^2(L^2)} \\ &\quad \le C(T, \boldsymbol{\kappa}, c_o, C_{\text{coer}}, p, p_{,t}, \mathbf{z}, \mathbf{u}_{,t})(h^{2r_1+2} + h^{2r_2}) \end{aligned} \tag{3.29}$$

Discrete-in-time error estimates

For convenience, we now make the following definitions:

$$||g||_{l^\infty(L^2)} = \max_{j=0,\ldots,N} ||g_j||_{L^2}, \quad ||g||_{l^2(H^s)} = \left(\sum_{j=0}^{N-1} ||g_{j,\theta}||^2_{H^s} \Delta t\right)^{\frac{1}{2}}.$$

For any sufficiently smooth function $g(t)$, by using the Taylor series expansion about $t = t_{j,\theta}$, we find:

$$\begin{aligned} g_{j+1} &= g|_{t=t_{j,\theta}} + \frac{1}{2}(1-\theta)\Delta t \frac{\partial g}{\partial t}|_{t=t_{j,\theta}} \\ &\quad + \frac{1}{8}(1-\theta)^2(\Delta t)^2 \frac{\partial^2 g}{\partial t^2}|_{t=t_{j,\theta}} + O(\Delta t^3), \\ g_j &= g|_{t=t_{j,\theta}} - \frac{1}{2}(1+\theta)\Delta t \frac{\partial g}{\partial t}|_{t=t_{j,\theta}} \\ &\quad + \frac{1}{8}(1+\theta)^2(\Delta t)^2 \frac{\partial^2 g}{\partial t^2}|_{t=t_{j,\theta}} + O(\Delta t^3). \end{aligned}$$

Therefore, if we multiply the first equation by $\frac{1}{2}(1+\theta)$ and the second equation by $\frac{1}{2}(1-\theta)$ and then sum, we arrive at

$$g_{j,\theta} = g|_{t=t_{j,\theta}} + \frac{1}{8}(\Delta t)^2(1-\theta)(1+\theta)\frac{\partial^2 g}{\partial t^2}|_{t=t_{j,\theta}} + O(\Delta t^3),$$

with the obvious exception that, for $\theta = 1$, $g_{j,\theta} = g|_{t=t_{j,\theta}}$. The above expression allows us to approximate to order $O(\Delta t^2)$, $\mathbf{u}(\mathbf{x}, t_{j,\theta}) \approx \mathbf{u}_{j,\theta}$, $p(\mathbf{x}, t_{j,\theta}) \approx p_{j,\theta}$, and $\mathbf{z}(\mathbf{x}, t_{j,\theta}) \approx \mathbf{z}(\mathbf{x}, t_{j,\theta})$.

We also remark that the following relationships hold as a result of Taylor expansion (see Rivière and Wheeler [13]):

$$\frac{p_{j+1} - p_j}{\Delta t} = p_t(x, t_{j,\theta}) + \Delta t\, \rho_{p,j,\theta}, \quad \forall x \in \Omega, \tag{3.30}$$

$$\frac{\mathbf{u}_{j+1} - \mathbf{u}_j}{\Delta t} = \mathbf{u}_t(x, t_{j,\theta}) + \Delta t\, \rho_{\mathbf{u},j,\theta}, \quad \forall x \in \Omega, \tag{3.31}$$

where $\rho_{p,j,\theta}$ and $\rho_{\mathbf{u},j,\theta}$ depend on time-derivatives of p and $\mathbf{u}$, respectively.

Theorem 3.3 (Auxiliary discrete time error estimate) *Let r_1 be associated with the degree of the polynomials used in the mixed space $(W_h, \mathbf{S}_h)$ satisfying Eqs. 3.4a–3.4e, and let r_2 be the degree of the polynomials used in the displacement space $\mathbf{V}_h$ satisfying Eq. 3.5. Let $\Delta t > 0$ be a sufficiently small, given time-step. Then, assuming Eqs. 2.1 and 2.2 and sufficient regularity of the true solution,*

$$\begin{aligned} &||E_{\mathbf{u}}^A||^2_{l^\infty(H^1)} + ||E_p^A||^2_{l^\infty(L^2)} + ||E_{\mathbf{z}}^A||^2_{l^2(L^2)} \\ &\quad \le C(h^{2r_1+2} + h^{2r_2}) + O(\Delta t^2), \end{aligned} \tag{3.32}$$

where $C = C(T, \alpha, c_o, \boldsymbol{\kappa}, C_{\text{coer}}, p, p_{,t}, \mathbf{z}, \mathbf{u}_{,t})$.

Proof Here, for simplicity, we assume homogeneous essential conditions for the displacement and flux. Note that this assumption affects only the value of C in Eq. 3.32, but not the rate of convergence.

We plug the true solution $(\mathbf{u}, \mathbf{z}, p)$ at time $t = t_{j,\theta}$ into the continuous-in-time, spatially discrete scheme (Eqs. 3.12–3.14). If we utilize Eqs. 3.30 and 3.31 and the other remarks concerning the Taylor series expansion, the following holds up to order $O(\Delta t)$:

$$a_{\mathbf{u}}(\mathbf{u}_{j,\theta}, \mathbf{v}_h) - \alpha(\nabla \cdot \mathbf{v}_h, p_{j,\theta}) = l_{1\,j,\theta}(\mathbf{v}_h), c_o\left(\frac{p_{j+1} - p_j}{\Delta t}, w_h\right) + \alpha\left(\nabla \cdot \left(\frac{\mathbf{u}_{j+1} - \mathbf{u}_j}{\Delta t}\right), w_h\right) + (\nabla \cdot \mathbf{z}_{j,\theta}, w_h) \tag{3.33}$$

$$= l_{2j\theta}(w_h) + c_o \Delta t(\rho_{p,\,j,\theta}, w_h) + \alpha \Delta t(\nabla \cdot \rho_{\mathbf{u},\,j,\theta}, w_h), \tag{3.34}$$

$$(\tilde{\boldsymbol{\kappa}}^{-1}\mathbf{z}_{j,\theta}, \mathbf{s}_h) - (p_{j,\theta}, \nabla \cdot \mathbf{s}_h) = l_{3\,j,\theta}(\mathbf{s}_h). \tag{3.35}$$

Next, subtracting Eqs. 3.22–3.24 from Eqs. 3.33–3.35, we have a set of equations reminiscent of Galerkin orthogonality (except for the time-derivative errors):

$$a_{\mathbf{u}}(\mathbf{u}_{j,\theta} - \bar{\mathbf{u}}_{j,\theta}, \mathbf{v}_h) - \alpha(\nabla \cdot \mathbf{v}_h, p_{j,\theta} - \bar{p}_{j,\theta}) = 0,$$
$$c_o\left(\tfrac{(p_{j+1} - \bar{p}_{j+1}) - (p_j - \bar{p}_j)}{\Delta t}, w_h\right) + \alpha\left(\nabla \cdot \left(\tfrac{(\mathbf{u}_{j+1} - \bar{\mathbf{u}}_{j+1}) - (\mathbf{u}_j - \bar{\mathbf{u}}_j)}{\Delta t}\right), w_h\right) + (\nabla \cdot (\mathbf{z}_{j,\theta} - \bar{\mathbf{z}}_{j,\theta}), w_h) \tag{3.36}$$

$$= c_o \Delta t(\rho_{p,\,j,\theta}, w_h) + \alpha \Delta t(\nabla \cdot \rho_{\mathbf{u},\,j,\theta}, w_h), \tag{3.37}$$

$$(\tilde{\boldsymbol{\kappa}}^{-1}(\mathbf{z}_{j,\theta} - \bar{\mathbf{z}}_{j,\theta}), \mathbf{s}_h) - (p_{j,\theta} - \bar{p}_{j,\theta}, \nabla \cdot \mathbf{s}_h) = 0. \tag{3.38}$$

To continue, we set $w_h = E^A_{p\,j,\theta}$ and $\mathbf{s}_h = E^A_{\mathbf{z}\,j,\theta}$ in Eqs. 3.37 and 3.38. Then, after adding and subtracting the appropriate projection operators, Eqs. 3.37 and 3.38 become

$$c_o\left(\frac{\left(E^I_{p\,j+1} + E^A_{p\,j+1}\right) - \left(E^I_{p\,j} + E^A_{p\,j}\right)}{\Delta t}, E^A_{p_{j,\theta}}\right) + \alpha\left(\nabla \cdot \left(\frac{\left(E^I_{\mathbf{u}\,j+1} + E^A_{\mathbf{u}\,j+1}\right) - \left(E^I_{\mathbf{u}\,j} + E^A_{\mathbf{u}\,j}\right)}{\Delta t}\right), E^A_{p\,j,\theta}\right) + \left(\nabla \cdot \left(E^I_{\mathbf{z}\,j,\theta} + E^A_{\mathbf{z}\,j,\theta}\right), E^A_{p_{j,\theta}}\right) = c_o \Delta t\left(\rho_{p,\,j,\theta}, E^A_{p\,j,\theta}\right) \tag{3.39}$$

$$\left(\tilde{\boldsymbol{\kappa}}^{-1}\left(E^I_{\mathbf{z}\,j,\theta} + E^A_{\mathbf{z}\,j,\theta}\right), E^A_{\mathbf{z}\,j,\theta}\right) - \left(E^I_{p\,j,\theta} + E^A_{p_{j,\theta}}, \nabla \cdot E^A_{\mathbf{z}\,j,\theta}\right) = 0. \tag{3.40}$$

By noting orthogonality relationships of the projection operators Eqs. 3.4a and 3.4d, we may then sum Eqs. 3.39 and 3.40 to get

$$c_o\left(\frac{\left(E^A_{p\,j+1} - E^A_{p\,j}\right)}{\Delta t}, E^A_{p\,j,\theta}\right) + \alpha\left(\nabla \cdot \left(\frac{\left(E^I_{\mathbf{u}\,j+1} + E^A_{\mathbf{u}\,j+1}\right) - \left(E^I_{\mathbf{u}\,j} + E^A_{\mathbf{u}\,j}\right)}{\Delta t}\right), E^A_{p\,j,\theta}\right) + \left(\tilde{\boldsymbol{\kappa}}^{-1}\left(E^I_{\mathbf{z}j,\theta} + E^A_{\mathbf{z}\,j,\theta}\right), E^A_{\mathbf{z}\,j,\theta}\right) = c_o \Delta t\left(\rho_{p,\,j,\theta}, E^A_{p\,j,\theta}\right) + \alpha \Delta t\left(\nabla \cdot \rho_{\mathbf{u},\,j,\theta}, E^A_{p\,j,\theta}\right). \tag{3.41}$$

Similarly, if we let $\mathbf{v}_h = \frac{E^A_{\mathbf{u}\,j+1} - E^A_{\mathbf{u}\,j}}{\Delta t}$, and use the properties of the elliptic projection, Eq. 3.36 becomes

$$a_{\mathbf{u}}\left(E^A_{\mathbf{u}\,j,\theta}, \frac{E^A_{\mathbf{u}\,j+1} - E^A_{\mathbf{u}\,j}}{\Delta t}\right) - \alpha\left(\nabla \cdot \left(\frac{E^A_{\mathbf{u}\,j+1} - E^A_{\mathbf{u}\,j}}{\Delta t}\right), E^I_{p\,j,\theta}\right) - \alpha\left(\nabla \cdot \left(\frac{E^A_{\mathbf{u}\,j+1} - E^A_{\mathbf{u}\,j}}{\Delta t}\right), E^A_{p\,j,\theta}\right) = 0. \tag{3.42}$$

Summing Eqs. 3.41 and 3.42 and rearranging, we find an equation for the auxiliary error:

$$\left[a_{\mathbf{u}}\left(E^A_{\mathbf{u}\,j,\theta}, \frac{E^A_{\mathbf{u}\,j+1} - E^A_{\mathbf{u}\,j}}{\Delta t}\right) + c_o\left(E^A_{p\,j,\theta}, \frac{E^A_{p\,j+1} - E^A_{p\,j}}{\Delta t}\right) \times \left(\tilde{\boldsymbol{\kappa}}^{-\frac{1}{2}} E^A_{\mathbf{z}\,j,\theta}, \tilde{\boldsymbol{\kappa}}^{-\frac{1}{2}} E^A_{\mathbf{z}\,j,\theta}\right)\right]$$
$$= \alpha\left(\nabla \cdot \left(\frac{E^A_{\mathbf{u}\,j+1} - E^A_{\mathbf{u}\,j}}{\Delta t}\right), E^I_{p\,j,\theta}\right) - \alpha\left(\nabla \cdot \left(\frac{E^I_{\mathbf{u}\,j+1} - E^I_{\mathbf{u}\,j}}{\Delta t}\right), E^A_{p\,j,\theta}\right) - \left(\tilde{\boldsymbol{\kappa}}^{-1} E^I_{\mathbf{z}\,j,\theta}, E^A_{\mathbf{z}\,j,\theta}\right) + c_o \Delta t\left(\rho_{p,\,j,\theta}, E^A_{p\,j,\theta}\right) + \alpha \Delta t\left(\nabla \cdot \rho_{\mathbf{u},\,j,\theta}, E^A_{p\,j,\theta}\right).$$

To provide bounds for the above, we start with the following easily proved inequalities:

$$a_{\mathbf{u}}\left(E^A_{\mathbf{u}\,j,\theta}, \frac{E^A_{\mathbf{u}\,j+1} - E^A_{\mathbf{u}\,j}}{\Delta t}\right) = a_{\mathbf{u}}\left(\frac{1+\theta}{2} E^A_{\mathbf{u}\,j+1} + \frac{1-\theta}{2} E^A_{\mathbf{u}\,j}, \frac{E^A_{\mathbf{u}\,j+1} - E^A_{\mathbf{u}\,j}}{\Delta t}\right)$$
$$= \frac{1+\theta}{2\Delta t} a_{\mathbf{u}}\left(E^A_{\mathbf{u}\,j+1}, E^A_{\mathbf{u}\,j+1}\right) - \frac{1+\theta}{2\Delta t} a_{\mathbf{u}}\left(E^A_{\mathbf{u}\,j+1}, E^A_{\mathbf{u}\,j}\right) + \frac{1-\theta}{2\Delta t} a_{\mathbf{u}}\left(E^A_{\mathbf{u}\,j}, E^A_{\mathbf{u}\,j+1}\right) - \frac{1-\theta}{2\Delta t} a\left(E^A_{\mathbf{u}\,j}, E^A_{\mathbf{u}\,j}\right)$$

$$
\begin{aligned}
&= \frac{1}{2\Delta t}||E^A_{\mathbf{u}\,j+1}||^2_{a_\mathbf{u}} - \frac{1}{2\Delta t}||E^A_{\mathbf{u}\,j}||^2_{a_\mathbf{u}} \\
&\quad + \frac{1}{2\Delta t}\theta\Big(||E^A_{\mathbf{u}\,j+1}||^2_{a_\mathbf{u}} - 2a\Big(E^A_{\mathbf{u}\,j+1}, E^A_{\mathbf{u}\,j}\Big) \\
&\quad + ||E^A_{\mathbf{u}\,j}||^2_{a_\mathbf{u}}\Big) \\
&= \frac{1}{2\Delta t}||E^A_{\mathbf{u}\,j+1}||^2_{a_\mathbf{u}} - \frac{1}{2\Delta t}||E^A_{\mathbf{u}\,j}||^2_{a_\mathbf{u}} \\
&\quad + \frac{1}{2\Delta t}\theta||E^A_{\mathbf{u}\,j+1} - E^A_{\mathbf{u}\,j}||^2_{a_\mathbf{u}} \\
&\geq \frac{1}{2\Delta t}\Big(||E^A_{\mathbf{u}\,j+1}||^2_{a_\mathbf{u}} - ||E^A_{\mathbf{u}\,j}||^2_{a_\mathbf{u}}\Big).
\end{aligned}
$$

Likewise,

$$
\Big(E^A_{p\,j,\theta}, \tfrac{E^A_{p\,j+1} - E^A_{pj}}{\Delta t}\Big) \geq \tfrac{1}{2\Delta t}\Big(||E^A_{pj+1}||^2_0 - ||E^A_{p\,j}||^2_0\Big).
$$

With the above identities, we may multiply the auxiliary error by $2\Delta t$ and then sum from 0 to $N-1$ to get:

$$
\begin{aligned}
&\Big(\,||E^A_{\mathbf{u}\,N}||^2_{a_\mathbf{u}} - \overbrace{||E^A_{\mathbf{u}\,0}||^2_{a_\mathbf{u}}}^{=0\text{by Eq.(3.15)}} + c_o(||E^A_{p\,N}||^2_0 - \overbrace{||E^A_{p\,0}||^2_0}^{=0\text{by Eq.(3.16)}}\,\Big) \\
&\quad + 2\sum_{j=0}^{N-1}||\tilde{\boldsymbol{\kappa}}^{-\frac{1}{2}}E^A_{\mathbf{z}\,j,\theta}||^2_0\Delta t \leq \Psi_1 + \Psi_2 + \Psi_3 \\
&\quad + \Psi_4 + \Psi_5, \qquad (3.43)
\end{aligned}
$$

where

$$
\Psi_1 = 2\alpha\sum_{j=0}^{N-1}\Big(\nabla\cdot\Big(\frac{E^A_{\mathbf{u}\,j+1} - E^A_{\mathbf{u}\,j}}{\Delta t}\Big), E^I_{pj,\theta}\Big)\Delta t, \qquad (3.44)
$$

$$
\Psi_2 = -2\alpha\sum_{j=0}^{N-1}\Big(\nabla\cdot\Big(\frac{E^I_{\mathbf{u}\,j+1} - E^I_{\mathbf{u}\,j}}{\Delta t}\Big), E^A_p j,\theta\Big)\Delta t, \qquad (3.45)
$$

$$
\Psi_3 = -2\sum_{j=0}^{N-1}\Big(\tilde{\boldsymbol{\kappa}}^{-1}E^I_{\mathbf{z}\,j,\theta}, E^A_{\mathbf{z}\,j,\theta}\Big)\Delta t, \qquad (3.46)
$$

$$
\Psi_4 = 2c_o\sum_{j=0}^{N-1}\Big(\Delta t\rho_{p,\,j,\theta}, E^A_{p\,j,\theta}\Big)\Delta t, \qquad (3.47)
$$

$$
\Psi_5 = 2\alpha\sum_{j=0}^{N-1}\Big(\Delta t\nabla\cdot\rho_{\mathbf{u},\,j,\theta}, E^A_{p\,j,\theta}\Big)\Delta t, \qquad (3.48)
$$

In order to bound Ψ_1, we utilize discrete integration by parts formula for grid functions f_j and g_j, $\sum_{j=0}^{N-1} g_j\frac{f_{j+1}-f_j}{\Delta t}\Delta t = (f_N g_N - f_0 g_0) - \sum_{j=0}^{N-1} f_{j+1}\frac{g_{j+1}-g_j}{\Delta t}\Delta t$.

We also use the approximation $\frac{E^I_{p\,j+1,\theta} - E^I_{p\,j,\theta}}{\Delta t} \approx (E^I_{p\,j,\theta})_t + O(\Delta t)$.

$$
\begin{aligned}
\Psi_1 &= 2\alpha\sum_{j=0}^{N-1}\Big(\nabla\cdot\Big(\frac{E^A_{\mathbf{u}\,j+1} - E^A_{\mathbf{u}\,j}}{\Delta t}\Big), E^I_{pj,\theta}\Big)\Delta t \\
&= 2\alpha\Big[\Big(\nabla\cdot E^A_{\mathbf{u}\,N}, E^I_{p\,N,\theta}\Big) - \Big(\overbrace{\nabla\cdot E^A_{\mathbf{u}\,0}}^{=0}, E^I_{p\,0,\theta}\Big)\Big] \\
&\quad - 2\alpha\sum_{j=0}^{N-1}\Big(\nabla\cdot E^A_{\mathbf{u}\,j+1}, \frac{E^I_{pj+1,\theta} - E^I_{pj,\theta}}{\Delta t}\Big)\Delta t \\
&\leq 2\alpha||\nabla\cdot E^A_{\mathbf{u}\,N}||_0||E^I_{p\,N,\theta}||_0 \\
&\quad + 2\alpha\sum_{j=0}^{N-1}||\nabla\cdot E^A_{\mathbf{u}\,j+1}||_0||(E^I_{pj,\theta})_t||_0\Delta t + O(\Delta t^2) \\
&\leq \epsilon||E^A_{\mathbf{u}\,N}||^2_1 + C||E^I_{p\,N,\theta}||^2_0 \\
&\quad + \alpha\sum_{j=0}^{N}||E^A_{\mathbf{u}\,j}||^2_1\Delta t + \alpha\sum_{j=0}^{N-1}||(E^I_{p_{j,\theta}})_t||^2_0\Delta t + O(\Delta t^2) \\
&\leq \epsilon||E^A_{\mathbf{u}\,N}||^2_1 + C||E^I_{p\,N,\theta}||^2_0 \\
&\quad + \alpha\sum_{j=0}^{N}||E^A_{\mathbf{u}\,j}||^2_1\Delta t + \alpha\sum_{j=0}^{N}||(E^I_{pj})_t||^2_0\Delta t + O(\Delta t^2),
\end{aligned}
$$

For Eq. 3.45, we likewise approximate the difference with a derivative to discover the bound

$$
\begin{aligned}
\Psi_2 &= -2\alpha\sum_{j=0}^{N-1}\Big(\nabla\cdot\Big(\frac{E^I_{\mathbf{u}\,j+1} - E^I_{\mathbf{u}\,j}}{\Delta t}\Big), E^A_{p\,j,\theta}\Big)\Delta t \\
&\quad \times 2\alpha\sum_{j=0}^{N-1}||(\nabla\cdot E^I_{\mathbf{u}\,j})_t||_0||E^A_{p\,j,\theta}||_0\Delta t + O(\Delta t^2) \\
&\leq \alpha\sum_{j=0}^{N}||(E^I_{\mathbf{u}\,j})_t||^2_1\Delta t + \alpha\sum_{j=0}^{N-1}||E^A_{p\,j,\theta}||^2_0\Delta t + O(\Delta t^2) \\
&\leq \alpha\sum_{j=0}^{N}||(E^I_{\mathbf{u}\,j})_t||^2_1\Delta t + \alpha\sum_{j=0}^{N}||E^A_{p\,j}||^2_0\Delta t + O(\Delta t^2)
\end{aligned}
$$

To bound Eq. 3.46, we use the assumption that $\boldsymbol{\kappa}$ is bounded below.

$$
\begin{aligned}
\Psi_3 &= -2\sum_{j=0}^{N-1}\Big(\tilde{\boldsymbol{\kappa}}^{-1}E^I_{\mathbf{z}\,j,\theta}, E^A_{\mathbf{z}\,j,\theta}\Big)\Delta t \\
&\leq 2\sum_{j=0}^{N-1}||\tilde{\boldsymbol{\kappa}}^{-1}E^I_{\mathbf{z}\,j,\theta}||_0||E^A_{\mathbf{z}\,j,\theta}||_0\Delta t \\
&\leq \frac{C}{\lambda^2_{\min}}\sum_{j=0}^{N}||E^I_{\mathbf{z}j}||^2_0\Delta t + \epsilon\sum_{j=0}^{N-1}||E^A_{\mathbf{z}\,j,\theta}||^2_0\Delta t,
\end{aligned}
$$

The bounds for Eqs. 3.47 and 3.48 are found by simple use of the Cauchy–Schwarz and Young inequalities.

$$\begin{aligned}\Psi_4 &= 2c_o \sum_{j=0}^{N-1} (\Delta t \rho_{p,j,\theta},\, E^A_{p\,j,\theta})\Delta t \\ &\le 2c_o \sum_{j=0}^{N-1} ||\Delta t \rho_{p,j,\theta}||_0 ||E^A_{p\,j,\theta}||_0 \Delta t \\ &\le c_o \sum_{j=0}^{N-1} ||\rho_{p,j,\theta}||_0^2 \Delta t^2 + c_o \sum_{j=0}^{N} ||E^A_{p\,j}||_0^2 \Delta t, \\ \Psi_5 &= 2\alpha \sum_{j=0}^{N-1} (\Delta t \nabla \cdot \rho_{\mathbf{u},j,\theta},\, E^A_{p\,j,\theta})\Delta t \\ &\le 2\alpha \sum_{j=0}^{N-1} ||\Delta t \nabla \cdot \rho_{\mathbf{u},j,\theta}||_0 ||E^A_{p\,j,\theta}||_0 \Delta t \\ &\le \alpha \sum_{j=0}^{N-1} ||\rho_{\mathbf{u},j,\theta}||_1^2 \Delta t^2 + \alpha \sum_{j=0}^{N} ||E^A_{p\,j}||_0^2 \Delta t.\end{aligned}$$

Utilizing this inequality in conjunction with the bounds for $\Psi_1 - \Psi_5$, the coercivity of $a_{\mathbf{u}}$ and the boundedness of κ allows us to rewrite Eq. 3.43 as

$$\begin{aligned}&(C_{\text{coer}} - \epsilon)||E^A_{\mathbf{u}\,N}||_1^2 + c_o||E^A_{p\,N}||_0^2 \\ &\quad + \Big(\frac{1}{\lambda_{\max}} - \epsilon\Big) \sum_{j=0}^{N} ||E^A_{\mathbf{z}\,j,\theta}||_0^2 \Delta t \\ &\quad \le C\Delta t \Big(\sum_{j=0}^{N} ||E^A_{\mathbf{u}\,j}||_1^2 + \sum_{j=0}^{N} ||E^A_{p\,j}||_0^2 \Big) \\ &\qquad + C\Big(||E^I_{p\,N,\theta}||_0^2 + \sum_{j=0}^{N} ||(E^I_{p\,j})_t||_0^2 \Delta t \\ &\qquad\quad + \sum_{j=0}^{N} ||(E^I_{\mathbf{u}\,j})_t||_1^2 \Delta t \\ &\qquad\quad + \sum_{j=0}^{N} ||E^I_{\mathbf{z}\,j,\theta}||_0^2 \Delta t \Big) + O(\Delta t^2).\end{aligned}$$

where $C = C(T, \alpha, c_o, \kappa)$. Then, if ϵ is small enough, all terms on the left-hand side of the above equation are positive. We let $c_{\min}$ to be the minimum value of the left-side coefficients and set the value of each left-side coefficient to $c_{\min}$; the inequality is then preserved. If we then divide both sides by $c_{\min}$, we find

$$\begin{aligned}&||E^A_{\mathbf{u}\,N}||_1^2 + ||E^A_{p\,N}||_0^2 + \sum_{j=0}^{N} ||E^A_{\mathbf{z}\,j,\theta}||_0^2 \Delta t \\ &\quad \le C\Delta t \Big(\sum_{j=0}^{N} ||E^A_{\mathbf{u}\,j}||_1^2 + \sum_{j=0}^{N} ||E^A_{p\,j}||_0^2 \Big) \\ &\qquad + C\Big(||E^I_{p\,N,\theta}||_0^2 + \sum_{j=0}^{N} ||(E^I_{p\,j})_t||_0^2 \Delta t \\ &\qquad\quad + \sum_{j=0}^{N} ||(E^I_{\mathbf{u}\,j})_t||_1^2 \Delta t + \sum_{j=0}^{N} ||E^I_{\mathbf{z}\,j,\theta}||_0^2 \Delta t \Big) + O(\Delta t^2).\end{aligned}$$

Thus, for Δt that is sufficiently small, we may apply the discrete version of Gronwall's lemma (see Gautschi [7]) to find

$$\begin{aligned}&||E^A_{\mathbf{u}\,N}||_1^2 + ||E^A_{p\,N}||_0^2 + \sum_{j=0}^{N} ||E^A_{\mathbf{z}\,j}||_0^2 \Delta t \\ &\quad \le C(T, \alpha, c_o, \kappa, C_{\text{coer}}) \Big(||E^I_{p\,N,\theta}||_0^2 + \sum_{j=0}^{N} ||(E^I_{p\,j})_t||_0^2 \Delta t \\ &\qquad\quad + \sum_{j=0}^{N} ||(E^I_{\mathbf{u}\,j})_t||_1^2 \Delta t \\ &\qquad\quad + \sum_{j=0}^{N} ||E^I_{\mathbf{z}\,j,\theta}||_0^2 \Delta t \Big) \\ &\qquad + O(\Delta t^2).\end{aligned}$$

Because the above holds for all $j = 0, 1, ..., N$, an application of approximation results yields the theorem. □

The auxiliary discrete-in-time error estimate and the triangle inequality allow us to conclude:

Corollary 3.4 (Discrete time finite element error estimate) *With the same conditions as in the preceeding theorem, to leading order in time the following estimate holds:*

$$\begin{aligned}&||\mathbf{u} - \bar{\mathbf{u}}||^2_{l^\infty(H^1)} + ||p - \bar{p}||^2_{l^\infty(L^2)} + ||\mathbf{z} - \bar{\mathbf{z}}||^2_{l^2(L^2)} \\ &\quad \le C(h^{2r_1+2} + h^{2r_2}) + O(\Delta t^2),\end{aligned}$$

where $C = C(T, \kappa, c_o, C_{\text{coer}}, p, p_{,t}, \mathbf{z}, \mathbf{u}_{,t})$.

Proof By the triangle inequality, the following holds:

$$||\mathbf{u}-\bar{\mathbf{u}}||^2_{l^\infty(H^1)} + ||p-\bar{p}||^2_{l^\infty(L^2)} + ||\mathbf{z}-\bar{\mathbf{z}}||^2_{l^2(L^2)}$$
$$\leq \overbrace{||E^A_{\mathbf{u}}||^2_{l^\infty(H^1)} + ||E^A_p||^2_{l^\infty(L^2)} + ||E^A_{\mathbf{z}}||^2_{l^2(L^2)}}^{\text{Auxiliary error}}$$
$$\overbrace{||E^I_{\mathbf{u}}||^2_{l^\infty(H^1)} + ||E^I_p||^2_{l^\infty(L^2)} + ||E^I_{\mathbf{z}}||^2_{l^2(L^2)}}^{\text{Interpolation error}}. \quad (3.49)$$

The preceding theorem and interpolation error estimates complete the corollary. □

4 Numerical results

In this section, we examine the accuracy of our proposed numerical schemes for solving the equations of poroelasticity. The numerical algorithm will be tested against two poroelastic problems with known analytical solutions that were not covered in the first paper. We begin with the standard 1-D consolidation problem known as Terzaghi's problem. We then turn to the more recent problem of Barry and Mercer's point-source problem. This problem gives us a chance to test our algorithm in a 2-D setting and it also provides an example reminiscent of an oil injection/extraction problem.

4.1 Terzaghi's problem

Terzaghi [16] provided one of the first solutions to the classic consolidation problem[1] in poroelasticity. The problem refers a surface load being applied to a fully saturated poroelastic domain. As a result of the initial loading F, an instantaneous positive overpressure occurs throughout the domain; more specifically, theory predicts that $p^+ \equiv \lim_{t\to 0} p(\mathbf{x}, t)$. Therefore, the pressure is discontinuous at $t = 0$ because $0 = p(\mathbf{x}, 0) < p^+$. However, because of the drainage at the surface, there is also a diffusion of pressure back toward its initial state. Regarding the deformation, of primary interest is the surface settlement, $s(t) = u_y(y = H, t)$, where u_y is the vertical component of the displacement and H is the height of the domain. Like the pressure, the settlement responds instantaneously to the applied load and is predicted to satisfy $s_0 \equiv \lim_{t\to 0} s(t)$. Moreover, theory also predicts asymptotic settlement to be $s_\infty \equiv \lim_{t\to\infty} s(t)$. This deformation is the very problem that produces so much of the costly damage that plagues some overland structures. So, despite the simplicity of its 1-D formulation, Terzaghi's problem does illustrate some interesting aspects of solid–fluid interactions and it is, in fact, applicable as a good approximation in a number of real-world circumstances.

For a precise recent formulation of the problem and presentation of the analytical solutions for the pressure and the settlement, we refer the reader to Coussy [6].

4.1.1 Results

Figure 1 shows the instantaneous increase in pressure typical for Terzaghi's problem. This leads to an increasingly large pressure gradient and higher derivatives as time approaches zero. To help determine the regularity of the pressure, $|\frac{\partial p(t)}{\partial y}|^2$ and $|\frac{\partial^2 p(t)}{\partial y^2}|^2$ are integrated over the spatial dimension to give values of $||p_y(t)||^2_0$ and $||p_{yy}(t)||^2_0$. Figure 2 shows their respective values as a function of time. It is discovered that

$$||p_y(t)||^2_0 \approx O(t^{-0.496}) \quad (4.1)$$
$$||p_{yy}(t)||^2_0 \approx O(t^{-1.475}) \quad (4.2)$$

The order approximation Eq. 4.1 implies that $p \in L^2(H^1)$, but Eq. 4.2 shows that $p \notin L^2(H^2)$. One may then assume that $p \in L^2(H^{1+s})$ for some $s \in [0, 1)$. It is this value of s, therefore, which equals the best rate of convergence one can guarantee using the continuous-in-time CG/Mixed algorithm discussed by Phillips and Wheeler [11] (submitted for publication). For the discrete-in-time GG/Mixed scheme presented herein, it may be possible to do better, and some results are presented below.

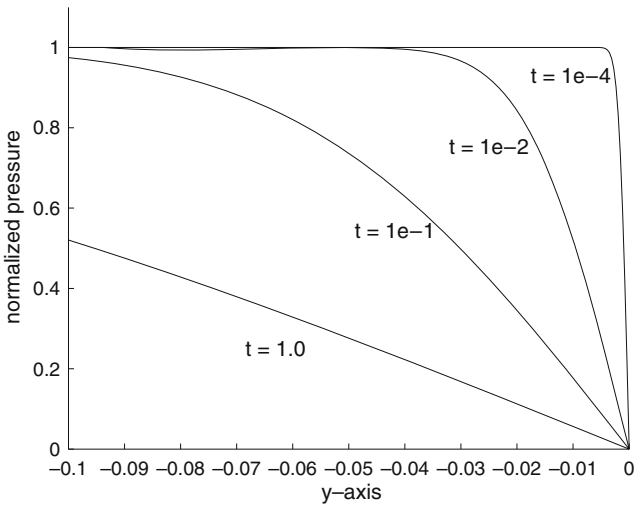

Fig. 1 Terzaghi's pressure at various times. As $t \to 0$ the pressure solution begins to experience very large derivatives near the boundary at $y = 0$

[1] Our discussion of the model and review of analytical results follow that found in the work by Coussy [6].

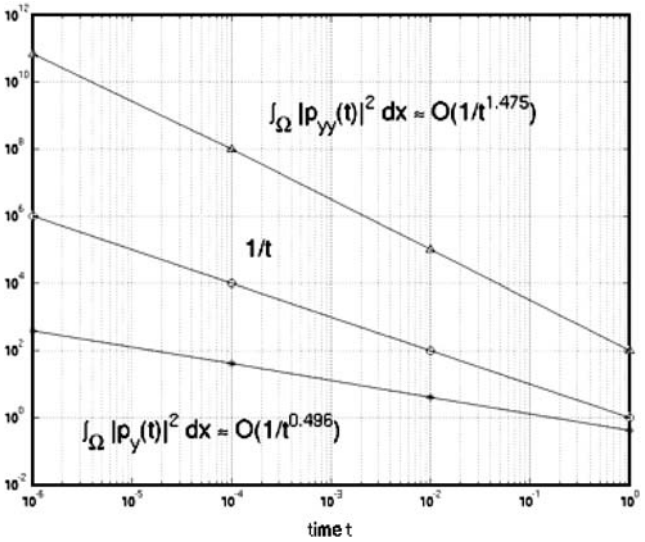

Fig. 2 Terzaghi's problem. The pressure integrals $||p_y(t)||_0^2$ and $||p_{yy}(t)||_0^2$

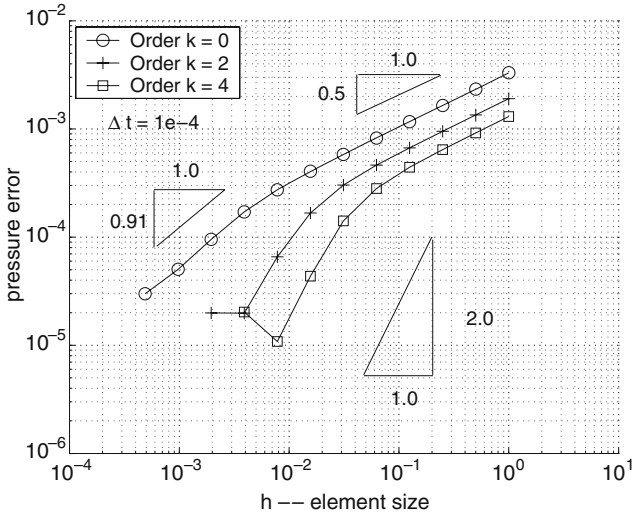

Fig. 3 Pressure error measured in the $l^\infty(L^2)$-norm for Terzaghi's problem. Here, schemes of *order* $k = 0, 2$ and 4 are used, and the time step $\Delta t = 1e-4$ is used

Figure 3 shows how the computed pressure error, $||p - \bar{p}||_{l^\infty(L^2)}$, varies with element size, h, and for the time step $\Delta t = 1e-4$. Figure 4 shows the results for an equivalent numerical experiment, except with the time step $\Delta t = 1e-2$.

For numerical experimentation, we use schemes[2] of *order* k, which we define to be continuous element-wise polynomials of degree $k+1$ to approximate the displacement, and the Raviart–Thomas space of order k for the flow variables. We also utilize the backward Euler time-stepping method ($\theta = 1$). We select the the unit height, $H = 1$, and force, $F = 1{,}000$, and choose the following set of material parameters:

$$\alpha = 1, \quad c_o = 1e-2, \quad \kappa = 1e-4.$$

4.1.2 A breakdown of the error

The convergence properties depicted in Figs. 3 and 4 can be grouped into roughly three categories. The first is characterized by an element size much larger than length over which large pressure gradient occurs. In this case, both the finite element solution and the L^2 projection of the solution approximate the instantaneous overpressure, p^+, well, but do not adequately capture the information about the solution near the large gradient. This means that the auxiliary pressure error, $E_p^A = P_h p - \bar{p}$, is small. Therefore, Eq. 3.49 roughly implies that $||p - \bar{p}||_{l^\infty(L^2)} \approx ||p - P_h p||_{l^\infty(L^2)}$. But because this error is reflective mainly of the one element containing the large gradient, $||p - P_h p||_{l^\infty(L^2)} \approx C(h \times (p^+ - 0)^2)^{1/2} \approx Ch^{1/2}$. This is what is seen for large element sizes.

There is a second, intermediate element size that is close to the size over which the large pressure gradient occurs. For these elements, both the finite element solution and projection begin to approximate the gradient. Therefore, the auxiliary and interpolation errors start to decrease, with the errors using the higher order schemes decreasing more rapidly.

The third range of element sizes are those which are small enough to refine the area with the large

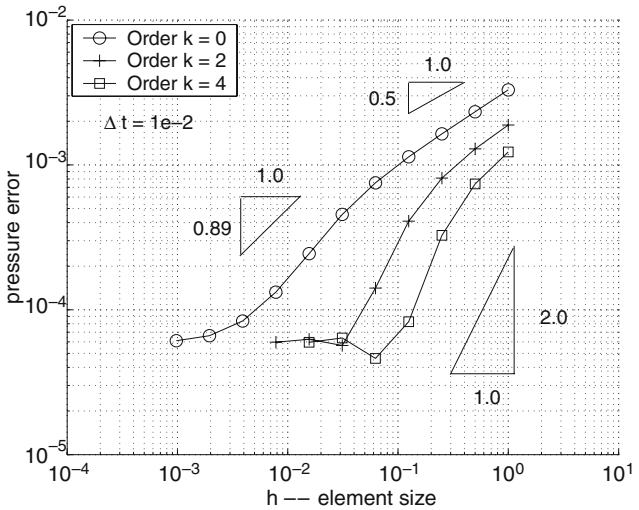

Fig. 4 Pressure error measured in the $l^\infty(L^2)$-norm for Terzaghi's problem. Here, schemes of *order* $k = 0, 2$ and 4 are used, and the time step $\Delta t = 1e-2$ is used

[2]The code for this experiment was implemented in C++ using the class libraries of the free and open-source finite element library, `deal.II` [1].

pressure gradient. For this case, the interpolation error becomes small so that $p \approx P_h p$. But the finite element solution approximates a time-error biased solution. To understand this bias, recall that, for the initial time step, the true solution satisfies[3] the variational formulation:

$$c_o(p_1, w) + \alpha(\nabla \cdot \mathbf{u}_1, w) + \Delta t(\nabla \cdot \mathbf{z}_1, w) = -\frac{1}{2}c_o(\Delta t)^2(p_{1,tt}, w) - \frac{1}{2}\alpha(\Delta t)^2(\nabla \cdot \mathbf{u}_{1,tt}, w) + O(\Delta t^3), \quad (4.3)$$

However, the finite element solution satisfies:

$$c_o(\bar{p}_1, w) + \alpha(\nabla \cdot \bar{\mathbf{u}}_1, w) + \Delta t(\nabla \cdot \bar{\mathbf{z}}_1, w) = 0.$$

So, in effect, the finite element scheme actually approximates a problem in which an *artificial mass source*, $\frac{1}{2}c_o(\Delta t)^2(p_{1,tt}, w) + \frac{1}{2}\alpha(\Delta t)^2(\nabla \cdot \mathbf{u}_{1,tt}, w)$, has been added to the right-hand side of Eq. 4.3 and acts as a mass source term. Additionally, given the rapid initial changes of the pressure, $p_{1,tt}$ can be significant, and thus, the artificial mass source can produce a noticeably biased solution. Figure 5 shows how this time-error-induced artificial mass source affects the finite element solution. The element size is smaller than the zone with the large pressure gradient, but the *order k* schemes progressively refine perfectly around the biased solution. At this point, further refinement or further increasing of the order of the scheme would only more accurately approximate the biased solution, but would do little to improve the accuracy with respect to the true solution.

Somewhat counterintuitively, a higher order time stepping scheme may not improve the situation. For example, the C-N scheme is one order higher in time than the backward Euler scheme used in the examples here. The C-N scheme was shown in experiments to exhibit a similar convergence profile, but with a slightly larger error. The reason is probably that the C-N scheme approximates the time derivative at time $t = \Delta t$ using information at the initial time, when the pressure solution is discontinuous.

Thus, other means may need to be employed in order to restore optimal discrete-in-time convergence. One attempt (for a scheme without the flux as a finite element variable, and only for the case $c_o = 0$) was shown in Murad and Loula [9]. Those authors showed improved rates of convergence for the pressure through postprocessing. Another interesting possibility is to try to approximate the artificial mass source and subtract it from the finite element source, thus reducing the bias. One potential way to accomplish this is to use initial finite element solutions (without an approximate artificial mass source) to approximate $p_{1,tt}$ and $\nabla \cdot \mathbf{u}_{1,tt}$, and then find a new finite element solution using these approximates as an artificial mass source. One could, of course, iterate this process.

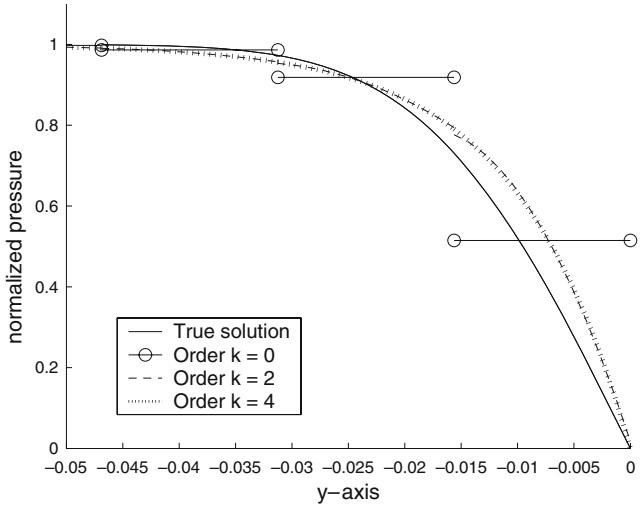

Fig. 5 True pressure solution and computed pressure solutions for Terzaghi's problem using schemes of *order* $k = 0, 2$, and 4. Solutions are shown at the first time step, $t = 1e - 2$. Beyond the large gradient region shown, the computed solutions of all orders match the true solution very well

4.2 The problem of Barry and Mercer

A more recent problem in poroelasticity with an analytical solution on a finite 2-D domain is the point-source problem and analytical series solution of Barry and Mercer [2]. The problem corresponds to a point-source sine wave on a rectangular domain. The boundary conditions are contrived in order to admit a solution, and thus, the problem does not correspond to a realistic relevant situation. However, the problem and solution have interesting mathematical properties that make it particularly suitable for numerical validation. For one thing, this problem offers a solution only for the case that fluid compressibility is zero, $c_o = 0$. Another interesting feature is that the nondimensionalized deformation and pressure solution are each independent of the dimensionless parameter $m = 1 + \frac{\lambda}{\mu}$.

The boundary of the domain $(0, a) \times (0, b)$ is drained on all sides, $p|_{\partial\Omega} = 0$. Along the boundaries $y = 0$, $y = b$, the deformation constraint is $\mathbf{u}_x = 0$, $\frac{\partial \mathbf{u}_y}{\partial y} = 0$, and along the boundaries $x = 0$, $x = a$, the deformation constraint is $\mathbf{u}_y = 0$, $\frac{\partial \mathbf{u}_x}{\partial x} = 0$. In addition, there is a source term,

$$h(t) = 2\beta F(\beta t),$$

[3] This analysis is for the backward Euler method.

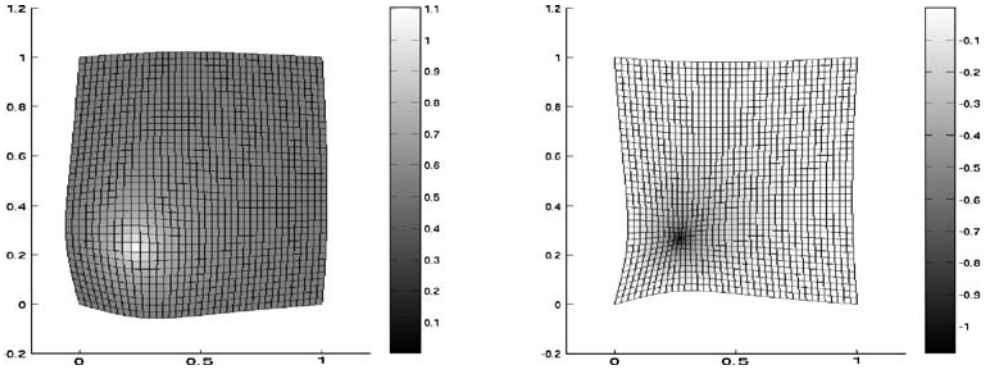

Fig. 6 The problem of Barry and Mercer: CG/Mixed solution at times $\hat{t} = \frac{\pi}{2}$ (*left*), and $\hat{t} = \frac{3\pi}{2}$ (*right*)

where $\beta = (\lambda + 2\mu)\kappa^4$ is the material-dependent source strength. We will also refer to $\hat{t} = \beta t$ as the *normalized time*.

4.2.1 Results

For numerical verification of our algorithms, we pose our problem on the unit domain and choose the following set of material parameters:

$$E = 1e+5, \quad \nu = 0.1, \quad \alpha = 1, \quad c_o = 0, \quad \kappa = 1e-2.$$

Here, E stands for *Young's* modulus and ν refers to *Poisson's* ratio. These constants can be found in terms of the model parameters λ and μ introduced earlier.

Using the prescribed set of physical parameters, Fig. 6 shows the numerical solution of the pressure (shaded field) and deformation produced by the CG/Mixed algorithm. At $\hat{t} = \frac{\pi}{2}$ (left), the source is positive; this fluid injection causes an expansion of the poroelastic medium. On the other hand, at $\hat{t} = \frac{3\pi}{2}$ (right), the source is negative; the resultant fluid withdrawal causes a contraction of the medium.

Figure 7 illustrates the effect of element size on the error of the pressure and displacement. The backward Euler time-stepping scheme is used and a normalized time step of $\Delta\hat{t} = \frac{0.1\pi}{2}$ is used for this test case. The graphic shows the error of the normalized pressure (top), $||\hat{p} - \hat{p}_h||_0$, and normalized displacement (bottom), $||\hat{\mathbf{u}} - \hat{\mathbf{u}}||_0$, using various element sizes. As one can see, a smaller step size leads to a reduction in error at most times. However, for a given time step, the error seems to be independent of element size (and minimal) at times $\hat{t} = n\pi$. This is precisely when the source term vanishes.

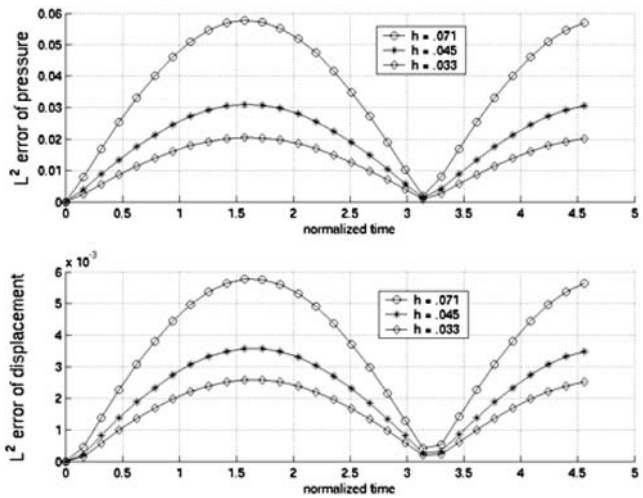

Fig. 7 CG/Mixed error in Barry and Mercer's problem

4.2.2 Independence of $1 + \frac{\lambda}{\mu}$

We mentioned in the opening discussion of Barry and Mercer's problem that the normalized analytical pressure and deformation solutions were independent on

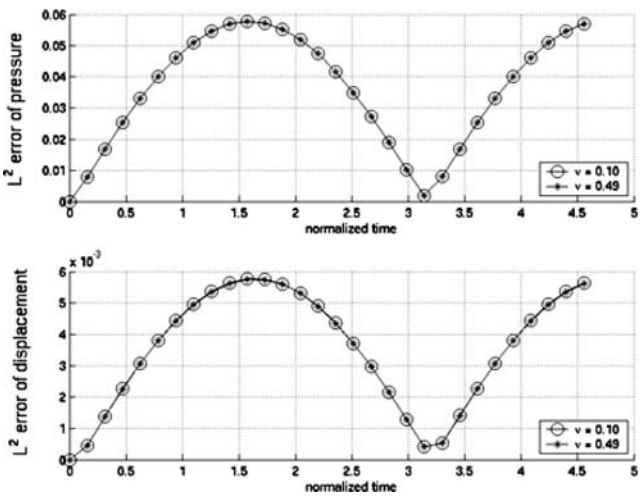

Fig. 8 Independence of $1 + \frac{\lambda}{\mu} = \frac{1}{1-2\nu}$

[4]Here, κ is a constant, though we retain the tensorial notation for uniformity.

$m = 1 + \frac{\lambda}{\mu}$. Furthermore, because $1 + \frac{\lambda}{\mu} = \frac{1}{1-2\nu}$, this implies an independence of Poisson's ratio, ν, over its acceptable range.

For concreteness, we include Fig. 8, which demonstrates that this independence carries over to the numerical algorithm. For the experiment, we use two cases: $\nu = 0.1$ and $\nu = 0.49$. The top graph shows the error of the normalized pressure, $||\hat{p_h}^{\nu} - \hat{p}||_0$, and shows identical error estimates for the normalized pressure errors. The bottom graph shows the normalized displacement, $||\hat{\tilde{\mathbf{u}}}^{\nu} - \hat{\mathbf{u}}||_0$, and shows very good agreement between the normalized deformation errors (the maximum difference is $1.1e-3$). Thus, because the analytical solution is independent of ν, we have by the triangle inequality,

$$||\hat{p_h}^{\nu=0.1} - \hat{p_h}^{\nu=0.49}||_0 \leq ||\hat{p} - \hat{p_h}^{\nu=0.1}||_0 + ||\hat{p} - \hat{p_h}^{\nu=0.49}||_0,$$

and a similar result for the deformation. Because each term is quite small on the right-hand side of the above equation, we conclude that $\hat{p_h}^{\nu=0.1} \approx \hat{p_h}^{\nu=0.49}$. Thus, we conclude that the numerical solution for the normalized pressure (and deformation) retains the material parameter independence.

References

1. Bangerth, W., Hartmann, R., Kanschat, G.: deal.II—A general purpose object oriented finite element library. Report ISC-06-02-MATH, Institute for Scientific Computation, Texas A&M University, College Station (2006)
2. Barry, S., Mercer, G.: Exact solutions for two-dimensional time dependent flow and deformation within a poroelastic medium. J. Appl. Mech. **66**, 536–540 (1999)
3. Biot, M.: General theory of three-dimensional consolidation. J. Appl. Phys. **2**, 155–164 (1941)
4. Biot, M.: Theory of elasticity and consolidation for a porous anisotropic media. J. Appl. Phys. **26**(2), 182–185 (1955)
5. Brenner, S., Scott, L.: The Mathematical Theory of Finite Element Methods. Springer, Berlin Heidelberg New York (1994)
6. Coussy, O.: Poromechanics. Wiley, New York (2004)
7. Gautschi, W.: Numerical Analysis: An Introduction. Birkhäuser, Boston (1997)
8. Lipnikov, K.: Numerical methods for the Biot model in poroelasticity. Ph.D. thesis, University of Houston (2002)
9. Murad, M., Loula, A.: Improved accuracy in finite element analysis of Biot's consolidation problem. Comput. Methods Appl. Mech. Eng. **95**, 359–382 (1992)
10. Nedelec, J.: Mixed finite elements in $\mathbb{R}3$. Numer. Math. **35**, 315–341 (1980)
11. Phillips, P.J., Wheeler, M.F.: A coupling of mixed and continuous Galerkin finite elements for poroelasticity I: the continuous in time case. Comput. Geosci. doi:10.1007/s10596-007-9045-y (2007)
12. Raviart, R., Thomas, J.: Mixed finite element method for second order elliptic problems. In: Galligani, I., Magenes, E. (eds.) Mathematical Aspects of the Finite Element Method. Lecture Notes in Mathematics, vol. 606, pp. 292–315. Springer, Berlin Heidelberg New York (1977)
13. Rivière, B., Wheeler, M.: Discontinuous Galerkin method applied to non- linear parabolic problems. In: Cockburn, B., Karniadakis, G., Shu, C.-W. (eds.) Discontinuous Galerkin Methods: Theory, Computation, and Applications. Lecture Notes in Computational Science and Engineering Edition, vol. 11. Springer, Berlin Heidelberg New York (1999)
14. Rivière, B., Wheeler, M.: Optimal error estimates applied to linear elasticity. Technical Report, ICES Report. ICES, Austin (2000)
15. Showalter, R.E.: Diffusion in poro-elastic media'. J. Math. Anal. Appl. **251**, 310–340 (2000)
16. Terzaghi, K.: Die Berechnung der Durchlassigkeitsziffer des Tones aus dem Verlauf der hydrodynamischen Spannungserscheinungen. Sitzung Berichte. Akadamie der Wissenschaften, Wien Mathematisch-Naturwissenschaftliche Klasse, Abteilung IIa. **132**, 105–124 (1923)
17. Terzaghi, K.: Theoretical Soil Mechanics. Wiley, New York (1943)
18. Wheeler, M.F.: A priori L^2 error estimates for Galerkin approximations to parabolic partial differential equations. SIAM J. Numer. Anal. **10**, 723–759 (1973)